

# CLEC10A is a Potential Prognostic Biomarker and Immune Regulator of Tumor Microenvironment in Breast Cancer

**Jincheng He**

First Hospital of Lanzhou University

**Lei Jiang**

First Hospital of Lanzhou University

**Jun Wang**

First Hospital of Lanzhou University

**Guangtao Min**

First Hospital of Lanzhou University

**Xiangwen Wang**

First Hospital of Lanzhou University

**Wei Chen**

First Hospital of Lanzhou University

**Hongpeng Wang**

First Hospital of Lanzhou University

**Nan Yao** (✉ [yaonanlanzhou@163.com](mailto:yaonanlanzhou@163.com))

First Hospital of Lanzhou University

---

## Research Article

**Keywords:** CLEC10A, tumor microenvironment, tumor-infiltrating immune cells, breast cancer, CIBERSORT, ESTIMATE

**Posted Date:** March 17th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-295872/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

---

## Abstract

The communication between tumor cells and immune cells influences the ecology of the tumor microenvironment in breast cancer, as well as the disease progression and clinical outcome. The aim of this study was to investigate the prognostic value of the immunomodulatory factor CLEC10A in breast cancer. We applied the CIBERSORT and ESTIMATE calculation methods to calculate the proportion of tumor-infiltrating immune cells (TICs) and the amount of immune and stromal components in 1053 BRCA cases from The Cancer Genome Atlas (TCGA) database. The differentially expressed genes (DEGs) were analyzed by COX regression analysis and protein-protein interaction (PPI) network construction. Then, CLEC10A was identified as a prognostic factor by the intersection analysis of univariate COX and PPI. Further analysis revealed that CLEC10A expression was negatively correlated with the clinical pathologic characteristics (age, clinical stage) and positively correlated with survival of BRCA patients. Gene set enrichment analysis (GSEA) showed that genes in the high CLEC10A expression group were mainly enriched in immune-related activities. Genes in the low CLEC10A expression group were enriched in biochemical functions. CIBERSORT analysis of the proportion of TICs revealed that Macrophages M1, B cells memory, B cells naive, T cells CD4<sup>+</sup> memory activated, T cells CD8<sup>+</sup>, and T cells gamma delta were positively correlated with CLEC10A expression, and Macrophages M0, Macrophages M2, Neutrophils, and NK cells resting were positively correlated with CLEC10A expression was negatively correlated, suggesting that CLEC10A may be an important factor in the immune regulation of the tumor microenvironment, especially in mediating the anti-tumor immune response of tumor-infiltrating immune cells at the tumor initiation stage. Therefore, CLEC10A expression may contribute to the prognosis of BRCA patients and provide a new idea for the immunotherapy of BRCA.

## Introduction

Breast cancer is the most common type of cancer in women and the leading cause of cancer death <sup>[1]</sup>. Although advances in diagnosis and treatment have improved the survival rate of patients with early-stage breast cancer, the survival rate of patients with advanced stages remains poor <sup>[2]</sup>. There is increasing amount of evidence indicate that the tumor microenvironment (TME), especially tumor-infiltrating immune cells (TICs), plays an important role in tumor development and is closely related to patient prognosis <sup>[3, 4]</sup>. The co-evolution of the interaction between cancer cell and the tumor microenvironment leads to the malignant phenotype of cancer, such as malignant proliferation, anti-apoptosis and immune surveillance evasion. Immune cells play an important anti-cancer role, although some cancer cell evade immune surveillance <sup>[5, 6]</sup>. Therefore, a comprehensive assessment of the tumor microenvironment and its level of immune infiltration, and further screening for prognostic markers associated with tumor immune infiltration can help predict the prognosis of breast cancer patients and guide individualized immunotherapy.

Breast cancer is generally considered to be non-immunogenic, so its response to immunotherapy is considered limited. However, recent reports have found that breast cancer also has a rich immune microenvironment, which is significantly related to the prognosis of patients. Tumor infiltrating lymphocytes (TILs) are the most common mononuclear immune cells, which have been reported in most breast cancer tissues. Study have shown that high CD8<sup>+</sup> T cell ratio and high CD8/FOXP3 ratio in residual cancer after neoadjuvant chemotherapy can accurately predict the prognosis of patients with triple negative breast cancer (TNBC) <sup>[8]</sup>. Another study showed that macrophages can affect tumor progression and may change the proliferation activity of cancer cells in the metastatic microenvironment <sup>[9]</sup>. These studies have revealed the potential prognostic value of immune cells in breast cancer.

In the present study, we analyzed differentially expressed genes (DEGs) generated by comparing immune components and stromal components in BRCA samples from The Cancer Genome Atlas (TCGA) database, and identify the prognostic biomarker C-type lectin domain containing 10A (CLEC10A), also known as macrophage galactose-type lectin (MGL). Furthermore, we calculated TICs proportions and proportions of immune and stromal components of BRCA samples from The Cancer Genome Atlas (TCGA) database by applying ESTIMATE and CIBERSORT computational methods, and revealed that CLEC10A is a potential biomarker of TME immunomodulation in BRCA.

# 1. Materials And Methods

## 1.1 Raw data

Transcriptomic RNA-seq data of 1164 breast invasive carcinoma (BRCA) cases, all female, containing 111 normal and 1053 tumor samples and their corresponding clinical data were downloaded from the TCGA database (<https://portal.gdc.cancer.gov/>).

## 1.2 Generation of Immunescore, StromalScore and ESTIMATEScore

The scores of the immune, stromal components of the TME were estimated for each sample by the ESTIMATE algorithm by using the R software loaded with the ESTIMATE package (version 4.0.2), presented in the form of three scores: ImmuneScore, StromalScore and ESTIMATEScore, respectively, with immune, stromal and the sum of both positively correlated, which means that the higher the corresponding score, the greater the proportion of the corresponding component in the TME.

## 1.3 Survival Analysis

Survival curves were plotted by using the `sursvive` and `survminer` packages in R software. Survival analysis was performed by Kaplan-Meier method on 1034 samples with detailed survival times out of 1053 tumor samples downloaded from TCGA database. Survival curves were compared using log-rank test. The test criterion was  $\alpha = 0.05$ .

## 1.4 Generation of DEGs in ImmuneScore and StromalScore

The 1034 tumor samples were labeled as high score or low score by comparison with the median score of ImmuneScore and StromalScore, respectively. Differential analysis of gene expression was performed with package `limma`, and then DEGs was generated by comparing high score and low score samples.  $\log_2FC > 1$  and  $FDR < 0.05$  were the inclusion criteria (FC = high score group/low score group).

## 1.5 GO and KEGG enrichment analysis of shared DEGs and Heatmaps

GO and KEGG enrichment analysis was performed on 224 shared DEGs by using the `clusterProfiler`, `richplot` and `ggplot2` packages in R software. Only enrichment terms with  $p < 0.05$  &  $FDRq\text{-value} < 0.05$  were statistically significant. Heatmaps of DEGs were produced using R software and the package `pheatmap`.

## 1.6 Correlation analysis of clinicopathological parameters

Clinical data of BRCA cases were downloaded from the TCGA database and correlations between clinicopathological parameters and TME scores or CLEC10A expression levels were analyzed using R software with the Wilcoxon test or Kruskal-Wallis test. The Bonferroni method was used to correct for significance levels.

## 1.7 PPI network construction

The PPI protein interaction network was constructed from the STRING database, followed by the use of nodes with protein interaction relationship confidence greater than 0.70 for reconstructing the network with Cytoscape version 3.6.1.

## 1.8 COX regression analysis

Univariate COX regression analysis was performed using R software and showed the top 16 significant genes.

## 1.9 Gene set enrichment analysis

The C2 KEGG and C7 gene sets v7.2 sets were downloaded from the Molecular Signatures database as target sets for GSEA, which were performed using the `gsea_4.1.0` software downloaded from the Broad Institute.

## 1.10 About TICs

The proportion of TICs in all tumor samples was estimated using the CIBERSORT calculation method, and only 854 tumor samples with  $p < 0.01$  were selected for the following analysis.

## 2. Results

### 2.1 Analysis process of this study

The transcriptome RNA sequence data of 1153 cases of BRCA downloaded from the TCGA database was calculated by CIBERSORT and ESTIMATE algorithms to estimate the proportion of TICs in the sample and the amount of immune and stromal components. Protein-protein interaction (PPI) networks and univariate COX regression analyses were constructed by using the DEG shared by ImmuneScore and StromalScore. Then, Intersection analysis was performed by using the core nodes in the PPI network and the significant genes obtained from the univariate COX regression analysis, resulting in 20 genes containing CLEC10A and others. We focused on the subsequent analysis of CLEC10A, including survival analysis and clinicopathological characteristics correlation analysis, gene set enrichment analysis (GSEA), and correlation with TICs.

### 2.2 Immunoscore correlates with prognosis of breast cancer patients

To determine the correlation between the proportion of immune and stromal components and survival, Kaplan-Meier survival analysis was performed on ImmuneScore, StromalScore and ESTIMATEScore, respectively. Higher scores on the ImmuneScore or StromalScore indicated a greater amount of the immune or stromal component in TME. ESTIMATEScore was the sum of ImmuneScore and StromalScore, which represents the combined proportion of these two components in TME. The results showed that although StromalScore and ESTIMATEScore were not significantly correlated with overall survival, the proportion of immune components was positively correlated with overall survival (Fig. 1). These results suggested that the immune component of TME may be a valuable prognostic indicator for BRCA patients.

### 2.3 Tumor microenvironment score correlates with clinicopathological characteristics of breast cancer patients

We downloaded the corresponding clinical information of breast cancer cases from the TCGA database and analyzed the correlation between the proportion of immune and stromal components in TME and clinicopathological characteristics. As shown in Fig. 2 and Supplementary Table 2, ImmuneScore was negatively correlated with age ( $p = 0.045$ ); StromalScore was negatively correlated with age ( $p = 0.033$ ), T classification ( $p = 0.002$ ), and TMN stage ( $p = 0.015$ ), and ESTIMATEScore decreased significantly with the depth of tumor infiltration ( $p = 0.022$ ). These results suggested that the proportion of stromal components is associated with BRCA progression, including invasion and metastasis, while age may be an influential factor for the immune component to affect tumor prognosis.

### 2.4 Enrichment analysis of differentially expressed genes shared by immune score and stromal score

DEGs were obtained by comparing high score samples to low score samples in the immune and stromal components, which were grouped into by comparison with the median. The results showed 818 DEGs were obtained from ImmuneScore, of which 728 genes were up-regulated, while 90 genes were down-regulated (Fig. 3A, C, D). Similarly, 782 DEGs were obtained from StromalScore, including 658 up-regulated and 124 down-regulated genes (Fig. 3B-D). ImmuneScore and StromalScore were crossed by Venn diagram to obtain 203 up-regulated genes, and 21 down-regulated genes. These DEGs (224 genes in total) shared by both in ImmuneScore and StromalScore may be influential factors for TME status. The results of gene ontology (GO) enrichment analysis showed that DEGs are mainly involved in immune-related biological functions, such as T cell activation (Fig. 3E). The Kyoto Encyclopedia Genes and Genomes (KEGG) enrichment analysis revealed that the molecular mechanisms involved in DEGs include chemokine signaling pathways, cytokine-cytokine receptor interactions, and hematopoietic cell lineage (Fig. 3F). Therefore, the function of DEG seems to be related to immune-related activities, which means that the participation of immune factors may have an important impact on the formation of TME in BRCA.

## 2.5 Intersection of PPI network and univariate COX regression analysis (key genes)

We constructed a PPI network based on the STRING database using Cytoscape software to further explore the underlying mechanisms. Figure 4A shows the interactions between the 224 differential genes, and the bar graph demonstrates part of the genes ranked by the number of interacting nodes (Fig. 4B). We performed a univariate COX regression analysis of survival in BRCA patients and identified the top 30 prognostically significant genes among the 224 DEGs (Fig. 4C). Then, the intersection analysis between key genes in the PPI network and univariate COX regression analysis of significant genes was performed, and the analysis identified 20 genes including CLEC10A (Fig. 4D).

## 2.6 The Correlation of CLEC10A Expression with Prognosis and Clinicopathological Characteristics of Breast Cancer Patients

CLEC10A plays a key role in tumor-infiltrating lymphocyte-mediated antitumor immune responses. CLEC10A is an antigen-presenting cell (APC)-specific antigen recognition receptor, which mediates immune-related tissue remodeling [10]. In this study, all BRCA samples were grouped into CLEC10A high expression and CLEC10A low expression groups by comparison with the median CLEC10A expression values. Survival analysis showed that BRCA patients in the CLEC10A high expression group survived longer than those with low CLEC10A expression (Fig. 5C). Subsequently, Wilcoxon rank sum test showed that the expression of CLEC10A was significantly lower in tumor samples than in normal samples (Fig. 5A). Similar results were observed in paired analysis between normal and tumor tissues from the same patient (Fig. 5B). The above results clearly indicated that CLEC10A expression is positively correlated with the prognosis of BRCA patients.

To further investigate the significance of CLEC10A in the development of breast cancer, we performed a correlation analysis with clinicopathological characteristics (Supplementary Table 1), which showed that CLEC10A expression was negatively correlated with age, TNM stage, and T classification of BRCA patients (Fig. 5D-H), especially CLEC10A expression decreased with the progression of TNM stage.

## 2.7 CLEC10A is a potential immunomodulatory factor in breast cancer

Given that the expression level of CLEC10A was negatively correlated with survival and TNM stage in BRCA patients, we applied the high and low expression groups compared with the median value of CLEC10A expression to GSEA enrichment analysis to study the gene function of CLEC10A, respectively. As shown in Fig. 6A, genes in the high CLEC10A expression group were mainly enriched in immune-related activities such as antigen processing and presentation, B-cell signaling pathway, Fc gamma R-mediated phagocytosis, and hematopoietic cell lineage. As for the genes in the low CLEC10A expression group were enriched in biologic recognition pathways such as glycosylphosphatidylinositol (Fig. 6B). For the set of C7 immunologic signatures genes defined by MSigDB, multiple immune function gene sets were enriched in the high CLEC10A expression group (Fig. 6C and Supplementary Table 4). In the low CLEC10A expression group, there was almost no gene set enrichment. These results suggest that CLEC10A may be a potential immune regulator of TME in breast cancer.

## 2.8 Correlation of CLEC10A with the proportion of TICs

We analyzed the proportion of tumor-infiltrating immune cell subpopulations by using the CIBERSORT algorithm to further ascertain the correlation between CLEC10A expression and the immune microenvironment, and constructed 22 kinds of immune cell profiles in BRCA samples (Fig. 7). The results of the difference and correlation analysis showed that a total of 12 TICs were correlated with CLEC10A expression (Fig. 8). Among them, eight kinds of TICs were positively correlated with CLEC10A expression, including memory B cells, naive B cells, resting dendritic cells, M1 macrophages, activated memory CD4<sup>+</sup> T cells, resting memory CD4<sup>+</sup> T cells, CD8<sup>+</sup> T cells, and  $\gamma\delta$  T cells; four kinds of TICs were negatively correlated with CLEC10A expression, including M0 macrophages, M2 macrophages, neutrophils, and resting NK cells. These results further supported that the level of CLEC10A influences the immune activity of TME and suggested that CLEC10A is an important immune regulator of TME in breast cancer.

### 3. Discussion

In the present study, we identify TME-related genes from the TCGA database that contribute to the evaluation of survival in BRCA patients. CLEC10A is considered to be involved in the immune activity of the breast cancer tumor microenvironment. Importantly, a series of analyses suggested that CLEC10A may be a potential indicator of TME immune activity evaluation in BRCA patients.

TME plays a crucial role in the initiation and progression of tumorigenesis. Exploring potential therapeutic targets can help improve TME remodeling and the prognosis of BRCA patients. A large number of studies have revealed the importance of the immune microenvironment in tumorigenesis. The results of our transcriptomic analysis of BRCA data from the TCGA database suggested that the immune component of TME contributes to patient prognosis, although the stromal component may dominate BRCA progression (e.g., invasion). These results highlighted that the exchange between cancer cells, infiltrating immune cells and stromal cells affects the ecology of the breast cancer microenvironment, influencing disease progression and clinical outcome<sup>[11]</sup>. Immune infiltration plays an important role in breast cancer patients, especially in triple-negative breast cancer (TNBC) or her2-positive breast cancer patients<sup>[12]</sup>. New studies have confirmed that an increase in the percentage of tumor infiltrating lymphocytes (TIL) is associated with survival in patients with smaller volumes of primary breast cancer, with survival rates approaching 100% in patients with a TIL score of 30%, even in the absence of chemotherapy<sup>[13]</sup>. The efficacy of pembrolizumab monotherapy in patients with metastatic TNBC confirms that enrichment of CD8<sup>+</sup> T cells predicts response to antitumor therapy<sup>[14]</sup>. Based on the transcriptomic analysis of BRCA in the TCGA database, we further found that the decreased expression of CLEC10A is closely related to the clinicopathological characteristics and poor prognosis of breast cancer. Therefore, it suggests that CLEC10A may be a potential prognostic marker and therapeutic target of TME in BRCA.

Protein glycosylation is known to have important effects in cell biological processes (such as proliferation, adhesion, differentiation, cell-cell interaction and immune response), and abnormal glycosylation is considered to be a hallmark of cancer<sup>[15]</sup>. Certain glycans with increased expression in tumor tissues compared to normal tissues are known as tumor-associated carbohydrate antigens (TACA)<sup>[16]</sup>. During tumor development, immune cells can detect the presence of TACA, which can enhance or weaken the immune response depending on the nature of the interaction<sup>[17]</sup>. The recognition of TACA is mediated by glycan binding protein, also known as lectin, whose binding specificity depends on their carbohydrate recognition domain (CRD). Macrophage galactose-type lectin (MGL; also known as CLEC10A or CD301), a member of the C-type lectin receptor (CLR) family<sup>[18]</sup>, recognizes glycan structures in a Ca<sup>2+</sup>-dependent manner through the carbohydrate recognition domain (CRD) and plays an important role in both innate and adaptive immune responses<sup>[6]</sup>. CLEC10A is mainly expressed by antigen-presenting cells (APCs), such as tumor-associated dendritic cells (DCs) and macrophages<sup>[19]</sup>, and which can specifically recognize terminal GalNAc residues ( $\alpha$ - or  $\beta$ -linked) found in the (S)Tn antigens and LacdiNAc epitopes.<sup>[20]</sup> It was found that the tumor-associated mucin MUC1 (MUC1-Tn) containing Tn antigen is frequently overexpressed in malignantly transformed epithelial adenocarcinomas (such as breast, pancreatic, colon), and binding to lectin receptors on mature dendritic cells activates T-cell activity to initiate anti-tumor immune responses, however, binding to immature dendritic cells may lead to immunosuppressive effects<sup>[21]</sup>. Our results suggest that CLEC10A expression is downregulated and associated with poor disease prognosis in the advanced stage of BRCA patients. The same results have been reported in hepatocellular carcinoma and colon cancer<sup>[22][23]</sup>. Interestingly, in the MGL1 (human CLEC10A homolog) deficient mouse strain, mouse mite induced dermatitis is exacerbated<sup>[24]</sup> and tissue remodeling associated with chronic immune responses is almost completely absent<sup>[10]</sup>, indicating that CLEC10A mediated immune infiltration may be involved in regulating stromal remodeling. Therefore, we further analyzed the relationship between CLEC10A expression and TME, and the GSEA results showed that the CLEC10A high expression group was significantly enriched in immune-related signaling pathways, such as antigen processing and presentation, B-cell signaling pathway, complement signaling pathway, and hematopoietic cell lineage. In the CLEC10A low expression group was enriched in glycosylphosphatidylinositol. All these results implied that CLEC10A may be involved in the regulation of immune infiltration in TME, especially the antitumor immune response.

Importantly, the receptor ligand svL4 of CLEC10A (a peptide ligand mimic of MGL2) as a monotherapy for ovarian cancer has been proven to cause multiple-fold expansion of mature immune cell populations in the abdominal cavity and effectively inhibit the development of ascites in a mouse ovarian cancer model [25]. This may be an effective anti-cancer therapy.

In this article, CIBERSORT analysis of the proportion of TICs showed that activated memory CD4<sup>+</sup> T cells, resting memory CD4<sup>+</sup> T cells, and CD8<sup>+</sup> T cells were positively correlated with CLEC10A expression in BRCA patients. In general, CD8<sup>+</sup> cytotoxic T cells have been recognized as a key component of effective antitumor immunity, and breast cancer patients with higher CD8<sup>+</sup> T cell infiltration have better survival [26]. However, lacking the help of sufficient CD4<sup>+</sup> T cells, CD8<sup>+</sup> T cells often fail to fully exploit their antitumor immunity *in vivo* [27]. As the main role of the immune system, CD4<sup>+</sup> T cells play a vital role in many aspects of their recruitment, activation and regulation of adaptive immune responses. Their auxiliary functions in B cell and CD8<sup>+</sup> cytotoxic T cell-mediated immune responses has been fully proven. However, tumor regulatory T cells (Treg) eliminate the early anti-tumor immune response induced by tumor-specific CD4<sup>+</sup> T cells and CD8<sup>+</sup> T cells and suppress multiple immune cells, such as CD4<sup>+</sup> and CD8<sup>+</sup> T cells, B cells, NK cells, NKT cells, dendritic cells (DCs) and macrophages [28]. Therefore, we predict that the downregulation of CLEC10A expression in advanced breast cancer tumors may be caused by the reduction of APC induced by regulatory T cells, which often leads to tumor cell immune tolerance. related studies are in progress.

In conclusion, our research indicates that CLEC10A is a valuable prognostic biomarker and immunotherapeutic target for breast cancer, and reveals that it may play an important role in the initiation of tumor immune response. With the development of sequencing technology, we believe that our model has great potential in clinical practice, thus further validating our results with clinical samples.

## Abbreviations

BRCA  
Breast invasive carcinoma  
TME  
Tumor microenvironment  
TIC  
Tumor-infiltrating immune cell  
TCGA  
The Cancer Genome Atlas  
GO  
Gene ontology  
KEGG  
Kyoto encyclopedia of genes and genomes  
GSEA  
Gene Set Enrichment Analysis  
TACA  
Tumor-associated carbohydrate antigens  
TNBC  
Triple negative breast cancer  
CRD  
Carbohydrate recognition domain

## Declarations

## Acknowledgements

The author sincerely appreciates the support of all members of the laboratory for this research.

## Funding

This work was supported by National Natural Science Foundation of China (Grant No. 82060527).

## Availability of data and materials

The datasets generated and/or analyzed during the current study are available in the TCGA database (<https://portal.gdc.cancer.gov/>).

## Ethics approval and consent to participate

Not applicable.

## Competing interests

The authors declare that they have no competing interests.

## Consent for publication

Written informed consent for publication was obtained from all participants.

## Authors Contributions

JCH, LJ, and NY came up with the design and conception. Data analysis and visualization was undertaken by JCH. The original manuscript was written and revised by JCH, LJ, and GTM. JW, XWW, WC, and HPW provided guiding suggestions and assistance. The manuscript was reviewed by all authors.

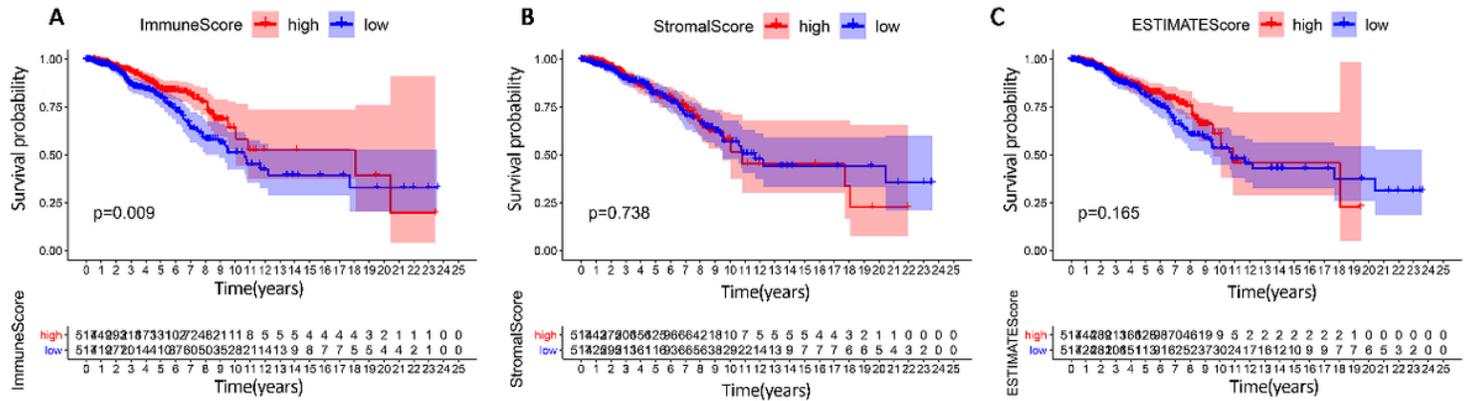
## References

1. Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, *68*(6), 394–424. doi:10.3322/caac.21492
2. Foster, T. S., Miller, J. D., Boye, M. E., Blieden, M. B., Gidwani, R., & Russell, M. W. (2011). The economic burden of metastatic breast cancer: a systematic review of literature from developed countries. *Cancer treatment reviews*, *37*(6), 405–415. doi: 10.1016/j.ctrv.2010.12.008
3. Chen, X., & Song, E. (2019). Turning foes to friends: targeting cancer-associated fibroblasts. *Nature reviews. Drug discovery*, *18*(2), 99–115. doi:10.1038/s41573-018-0004-1
4. Ascierto, M. L., De Giorgi, V., Liu, Q., Bedognetti, D., Spivey, T. L., Murtas, D., et al. (2011). An immunologic portrait of cancer. *Journal of translational medicine*, *9*, 146. doi:10.1186/1479-5876-9-146
5. Spitzer, M. H., Carmi, Y., Reticker-Flynn, N. E., Kwek, S. S., Madhiredy, D., Martins, M. M., et al. (2017). Systemic Immunity Is Required for Effective Cancer Immunotherapy. *Cell*, *168*(3), 487–502.e15. doi: 10.1016/j.cell.2016.12.022
6. Gajewski, T. F., Schreiber, H., & Fu, Y. X. (2013). Innate and adaptive immune cells in the tumor microenvironment. *Nature immunology*, *14*(10), 1014–1022. doi: 10.1038/ni.2703
7. Mittal, S., Brown, N. J., & Holen, I. (2018). The breast tumor microenvironment: role in cancer development, progression and response to therapy. *Expert review of molecular diagnostics*, *18*(3), 227–243. doi: 10.1080/14737159.2018.1439382
8. Miyashita, M., Sasano, H., Tamaki, K., Hirakawa, H., Takahashi, Y., Nakagawa, S., et al. (2015). Prognostic significance of tumor-infiltrating CD8 + and FOXP3 + lymphocytes in residual tumors and alterations in these parameters after neoadjuvant chemotherapy in triple-negative breast cancer: a retrospective multicenter study. *Breast cancer research: BCR*, *17*(1), 124. doi: 10.1186/s13058-015-0632-x

9. Yang, M., Ma, B., Shao, H., Clark, A. M., & Wells, A. (2016). Macrophage phenotypic subtypes diametrically regulate epithelial-mesenchymal plasticity in breast cancer cells. *BMC cancer*, *16*, 419. doi: 10.1186/s12885-016-2411-1
10. Sato, K., Imai, Y., Higashi, N., Kumamoto, Y., Onami, T. M., Hedrick, S. M., et al. (2005). Lack of antigen-specific tissue remodeling in mice deficient in the macrophage galactose-type calcium-type lectin 1/CD301a. *Blood*, *106*(1), 207–215. doi: 10.1182/blood-2004-12-4943
11. Wagner, J., Rapsomaniki, M. A., Chevrier, S., Anzeneder, T., Langwieder, C., Dykgers, A., et al. (2019). A Single-Cell Atlas of the Tumor and Immune Ecosystem of Human Breast Cancer. *Cell*, *177*(5), 1330–1345.e18. doi: 10.1016/j.cell.2019.03.005
12. Savas, P., Salgado, R., Denkert, C., Sotiriou, C., Darcy, P. K., Smyth, M. J., et al. (2016). Clinical relevance of host immunity in breast cancer: from TILs to the clinic. *Nature reviews. Clinical oncology*, *13*(4), 228–241. doi: 10.1038/nrclinonc.2015.215
13. Park, J. H., Jonas, S. F., Bataillon, G., Criscitiello, C., Salgado, R., Loi, S., et al. (2019). Prognostic value of tumor-infiltrating lymphocytes in patients with early-stage triple-negative breast cancers (TNBC) who did not receive adjuvant chemotherapy. *Annals of oncology: official journal of the European Society for Medical Oncology*, *30*(12), 1941–1949. doi: 10.1093/annonc/mdz395
14. Byrne, A., Savas, P., Sant, S., Li, R., Virassamy, B., Luen, S. J., et al. (2020). Tissue-resident memory T cells in breast cancer control and immunotherapy responses. *Nature reviews. Clinical oncology*, *17*(6), 341–348. doi: 10.1038/s41571-020-0333-y
15. Pirro, M., Mohammed, Y., van Vliet, S. J., Rombouts, Y., Sciacca, A., de Ru, A. H., et al. (2020). N-Glycoproteins Have a Major Role in MGL Binding to Colorectal Cancer Cell Lines: Associations with Overall Proteome Diversity. *International journal of molecular sciences*, *21*(15), 5522. doi: 10.3390/ijms21155522
16. Beckwith, D. M., & Cudic, M. (2020). Tumor-associated O-glycans of MUC1: Carriers of the glyco-code and targets for cancer vaccine design. *Seminars in immunology*, *47*, 101389. doi: 10.1016/j.smim.2020.101389
17. Pinho, S. S., & Reis, C. A. (2015). Glycosylation in cancer: mechanisms and clinical implications. *Nature reviews. Cancer*, *15*(9), 540–555. doi: 10.1038/nrc3982
18. van Vliet, S. J., García-Vallejo, J. J., & van Kooyk, Y. (2008). Dendritic cells and C-type lectin receptors: coupling innate to adaptive immune responses. *Immunology and cell biology*, *86*(7), 580–587. doi: 10.1038/icb.2008.55
19. Heger, L., Balk, S., Lühr, J. J., Heidkamp, G. F., Lehmann, C., Hatscher, L., et al. (2018). CLEC10A Is a Specific Marker for Human CD1c<sup>+</sup> Dendritic Cells and Enhances Their Toll-Like Receptor 7/8-Induced Cytokine Secretion. *Frontiers in immunology*, *9*, 744. doi:10.3389/fimmu.2018.00744
20. van Vliet, S. J., van Liempt, E., Saeland, E., Aarnoudse, C. A., Appelmelk, B., Irimura, T., et al. (2005). Carbohydrate profiling reveals a distinctive role for the C-type lectin MGL in the recognition of helminth parasites and tumor antigens by dendritic cells. *International immunology*, *17*(5), 661–669. doi: 10.1093/intimm/dxh246
21. Saeland, E., van Vliet, S. J., Bäckström, M., van den Berg, V. C., Geijtenbeek, T. B., Meijer, G. A., & van Kooyk, Y. (2007). The C-type lectin MGL expressed by dendritic cells detects glycan changes on MUC1 in colon carcinoma. *Cancer immunology, immunotherapy: CII*, *56*(8), 1225–1236. doi: 10.1007/s00262-006-0274-z
22. Zhu, G. Q., Yang, Y., Chen, E. B., Wang, B., Xiao, K., Shi, S. M., et al. (2019). Development and validation of a new tumor-based gene signature predicting prognosis of HBV/HCV-included resected hepatocellular carcinoma patients. *Journal of translational medicine*, *17*(1), 203. doi: 10.1186/s12967-019-1946-8
23. Zhang, X., Zhao, H., Shi, X., Jia, X., & Yang, Y. (2020). Identification and validation of an immune-related gene signature predictive of overall survival in colon cancer. *Aging*, *12*(24), 26095–26120. doi: 10.18632/aging.202317
24. Kanemaru, K., Noguchi, E., Tahara-Hanaoka, S., Mizuno, S., Tateno, H., Denda-Nagai, K., et al. (2019). Clec10a regulates mite-induced dermatitis. *Science immunology*, *4*(42), eaax6908. doi: 10.1126/sciimmunol.aax6908
25. Eggink, L. L., Roby, K. F., Cote, R., & Kenneth Hooper, J. (2018). An innovative immunotherapeutic strategy for ovarian cancer: CLEC10A and glycomimetic peptides. *Journal for immunotherapy of cancer*, *6*(1), 28. doi: 10.1186/s40425-018-

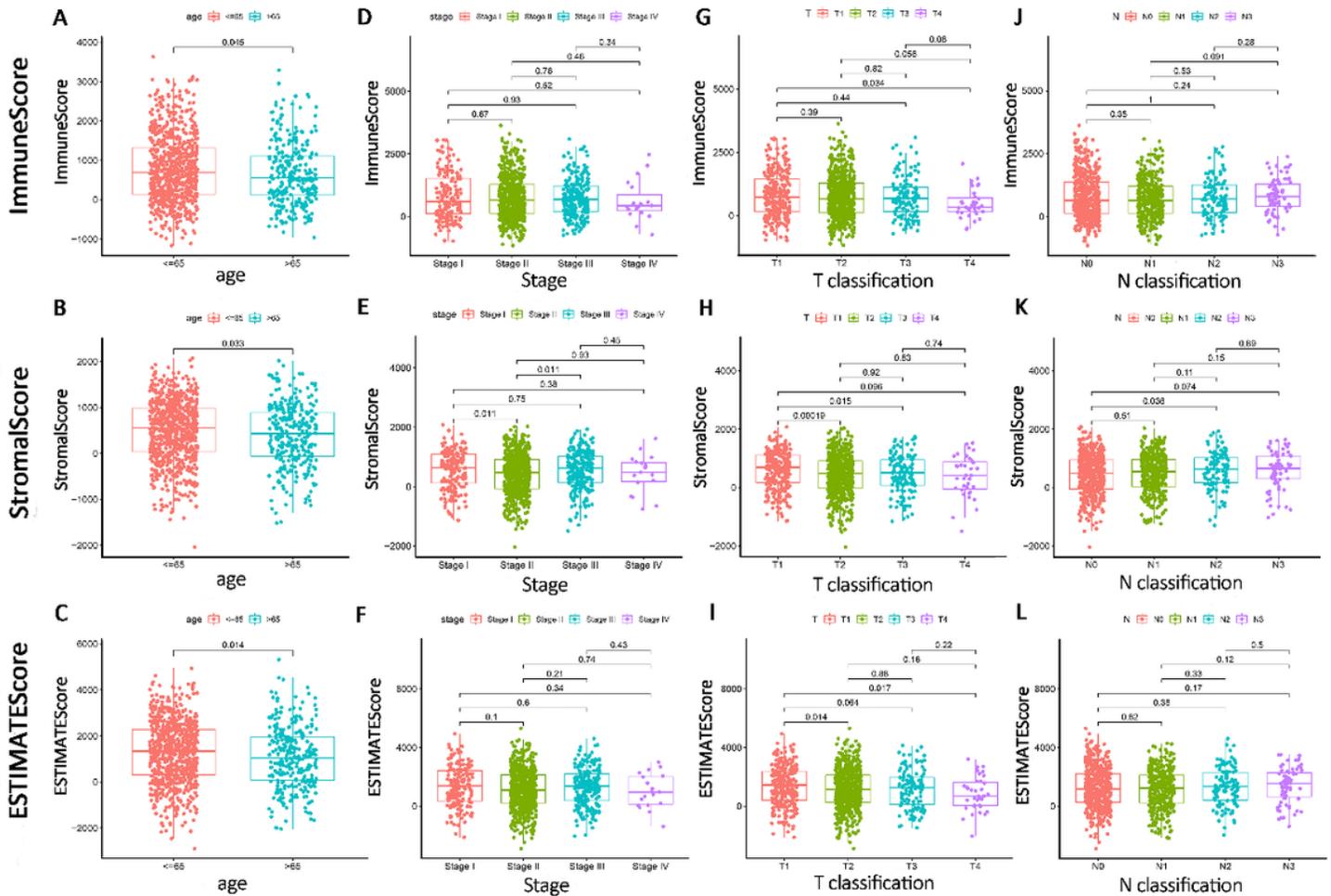
26. Mahmoud, S., Lee, A., Ellis, I., & Green, A. (2012). CD8(+) T lymphocytes infiltrating breast cancer: A promising new prognostic marker?. *Oncoimmunology*, 1(3), 364–365. doi: 10.4161/onci.18614
27. Bos, R., Marquardt, K. L., Cheung, J., & Sherman, L. A. (2012). Functional differences between low- and high-affinity CD8(+) T cells in the tumor environment. *Oncoimmunology*, 1(8), 1239–1247. doi: 10.4161/onci.21285
28. Liu, Z., Kim, J. H., Falo, L. D., Jr, & You, Z. (2009). Tumor regulatory T cells potentially abrogate antitumor immunity. *Journal of immunology (Baltimore, Md.: 1950)*, 182(10), 6160–6167. doi: 10.4049/jimmunol.0802664

## Figures



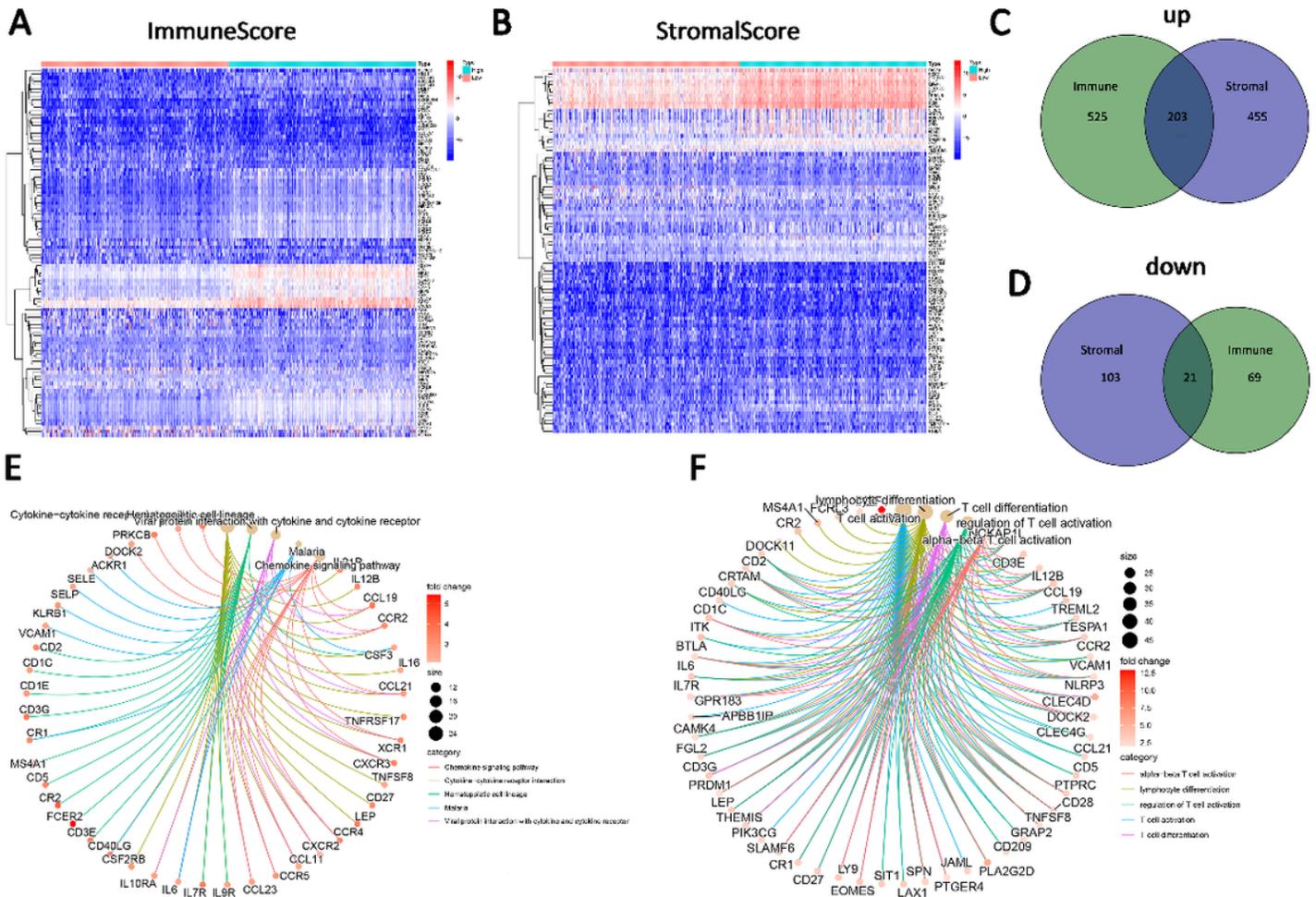
**Figure 1**

Correlation between score and survival of BRCA patients. (A) Kaplan-Meier survival analysis of BRCA patients grouped into high or low score groups in ImmuneScore by comparing with the median.  $p = 0.009$  by log-rank test. (B) Kaplan-Meier survival curves for the StromalScore with  $p = 0.738$  by log-rank test. (C) Kaplan-Meier survival curve of ESTIMATEScore in TME with  $p = 0.165$  by log-rank test.



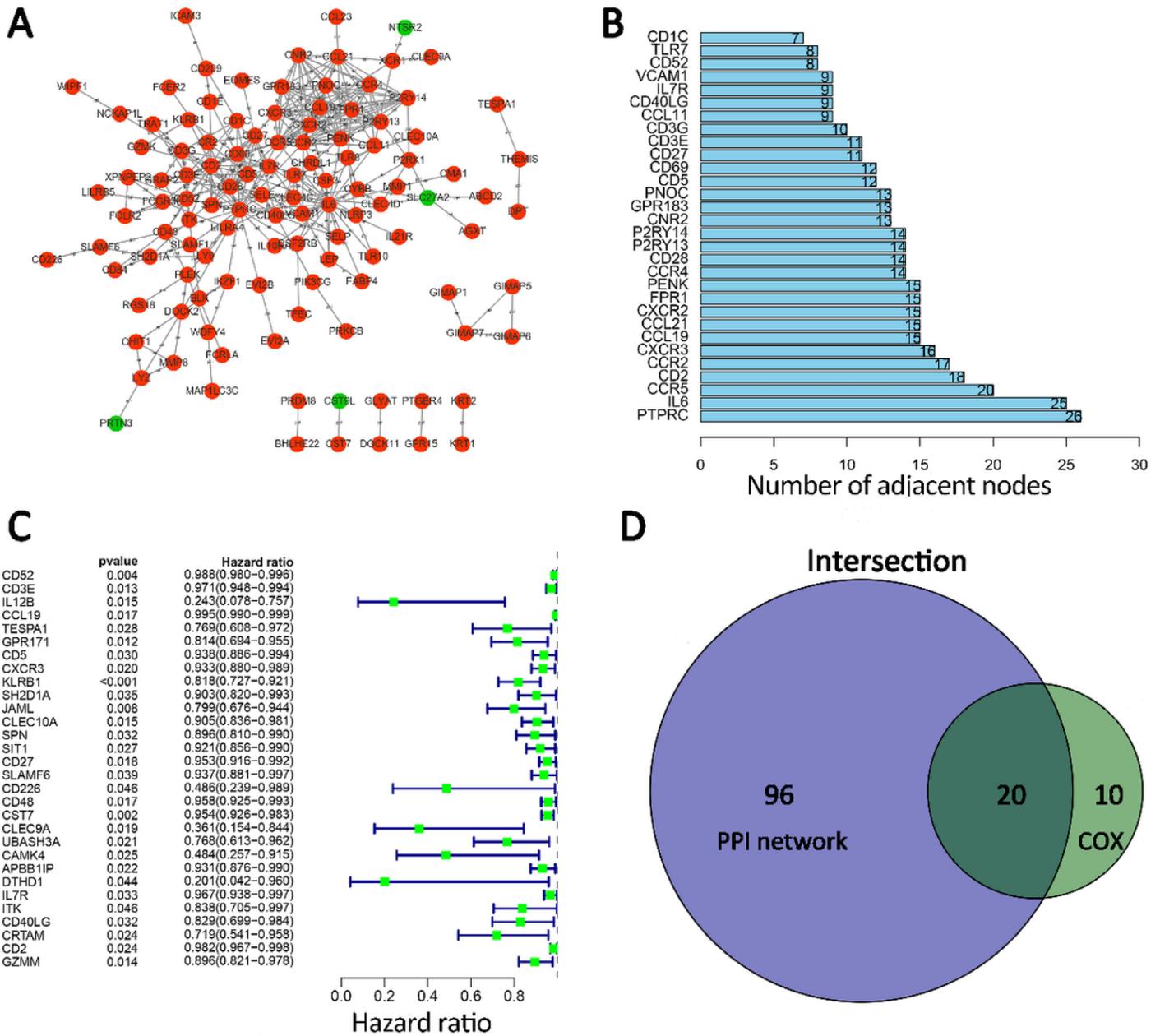
**Figure 2**

Correlation of scores with clinicopathological characteristics. (A-C) Immunescore, Stromalscore and ESTIMATEScore with age, the  $p = 0.045$ ,  $0.033$  and  $0.014$  respectively by Wilcoxon rank sum test. (D-F) Immunescore, Stromalscore and ESTIMATEScore with TNM stage,  $p = 0.84$ ,  $0.015$  and  $0.284$  respectively by Kruskal-Wallis rank sum test. (G-I) Immunescore, Stromalscore and ESTIMATEScore with T classification,  $p = 0.176$ ,  $0.002$  and  $0.022$  respectively by Kruskal-Wallis rank sum test. (J-L) Immunescore, Stromalscore and ESTIMATEScore with N classification,  $p = 0.387$ ,  $0.082$  and  $0.389$  respectively by Kruskal-Wallis rank sum test. Detailed data in Supplementary Table 2.



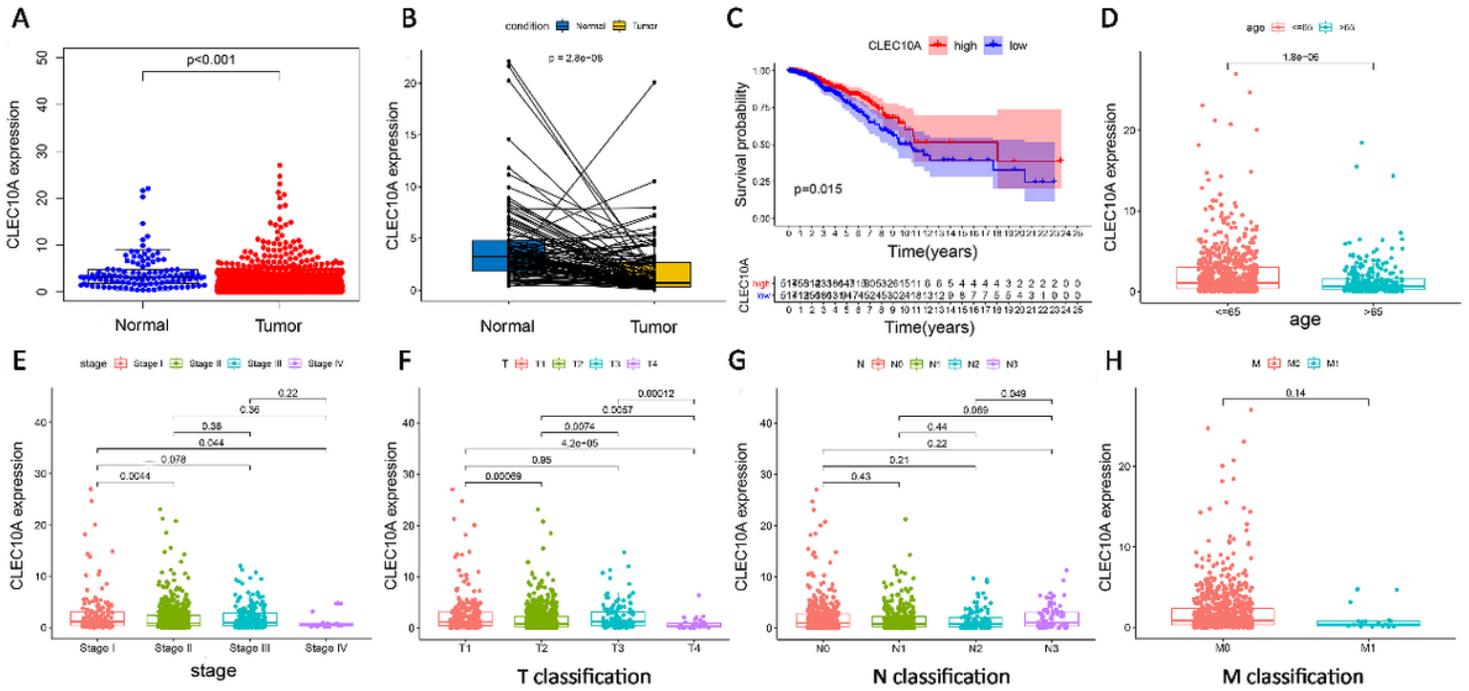
**Figure 3**

Heatmap, Venn plots and GO, KEGG enrichment analysis of DEGs. (A-B) Heatmap of differentially expressed genes for Immunescore or Stromalscore by Wilcoxon rank sum test with  $|\log_2FC| > 1$  &  $FDR < 0.05$  as significance criteria. Row names of heat map are gene names and column names are sample IDs that not shown in plot. (C-D) Venn plots showing common up-regulated and down-regulated DEGs shared by Immunescore or Stromalscore,  $|\log_2FC| > 1$  &  $FDR < 0.05$  as DEGS entry criteria. (E-F) GO and KEGG enrichment analysis of 224 DEGs. Terms with p and FDRq-value  $< 0.05$  were considered as significantly enriched.



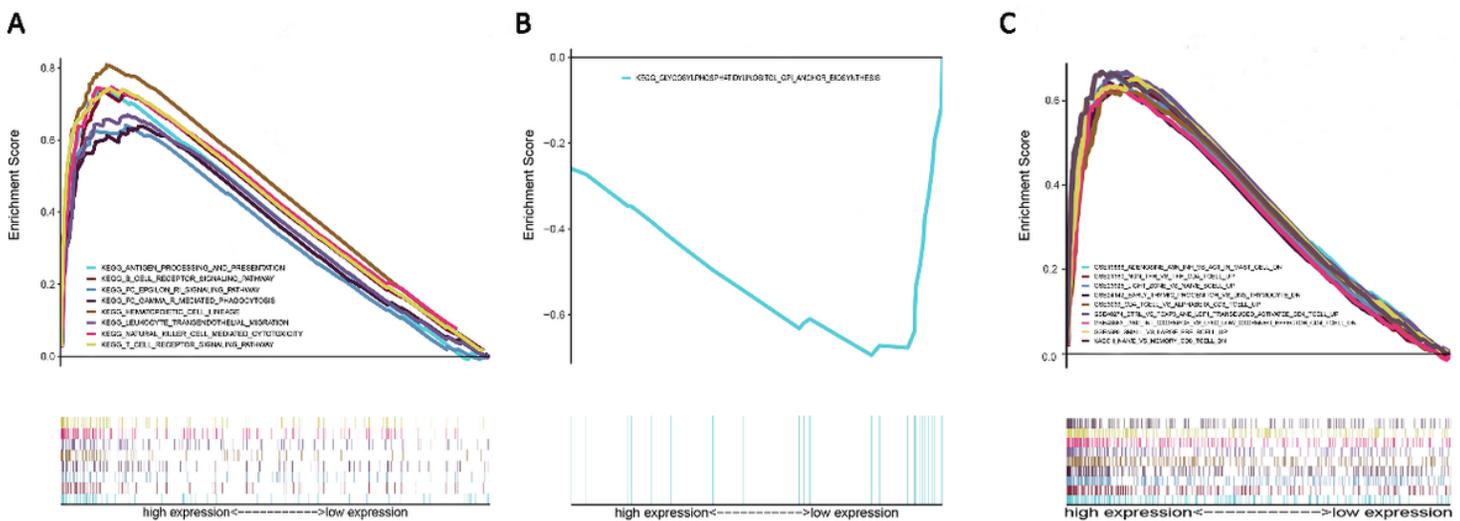
**Figure 4**

Protein interaction network and univariate COX analysis. (A) Protein interaction networks were constructed with nodes with confidence levels >0.7. (B) The top 30 genes ordered according to the number of protein interaction network nodes. (C) Univariate COX regression analysis of DEGs listing the top 30 significant genes with  $p < 0.05$ . (D) Venn plot showing the common factors shared by nodes in PPI and top significant factors in univariate COX.



**Figure 5**

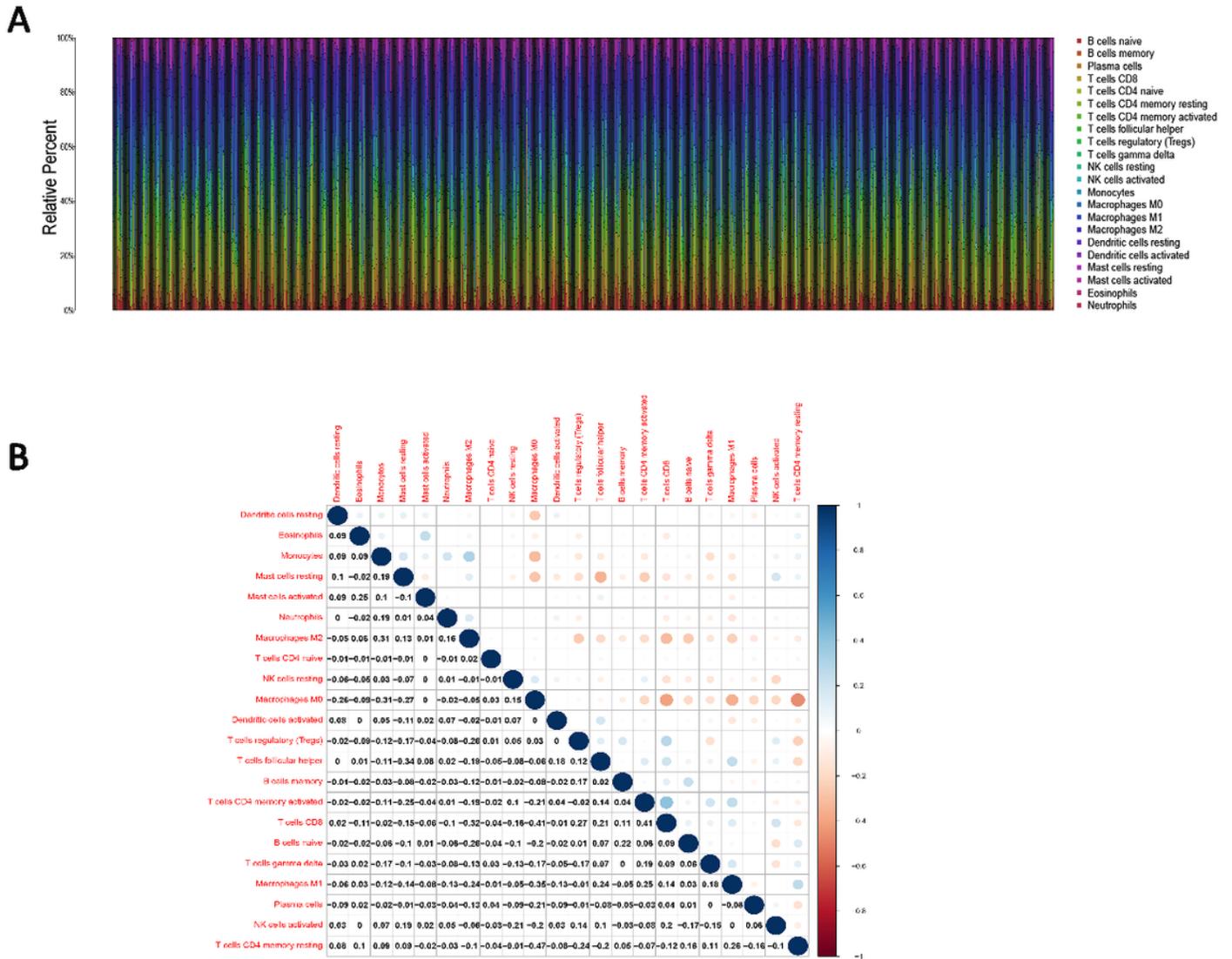
Differential expression of CLEC10A in the samples and correlation with survival analysis and clinicopathological parameters of breast cancer patients. (A) Differential expression of CLEC10A in breast cancer tissues and normal tissues by Wilcoxon rank sum test. (B) Differential expression of CLEC10A in breast cancer tissues and their paired normal tissues by Wilcoxon rank sum test. (C) Kaplan-Meier survival analysis of BRCA patients grouped into high and low expression groups by comparing with the median of CLEC10A.  $p = 0.015$  by log-rank test. (D-H) Correlation of CLEC10A expression with clinicopathological characteristics of BRCA patients. Wilcoxon rank sum or Kruskal–Wallis rank sum test served as the statistical significance test. Detailed data in Supplementary Table 3.



**Figure 6**

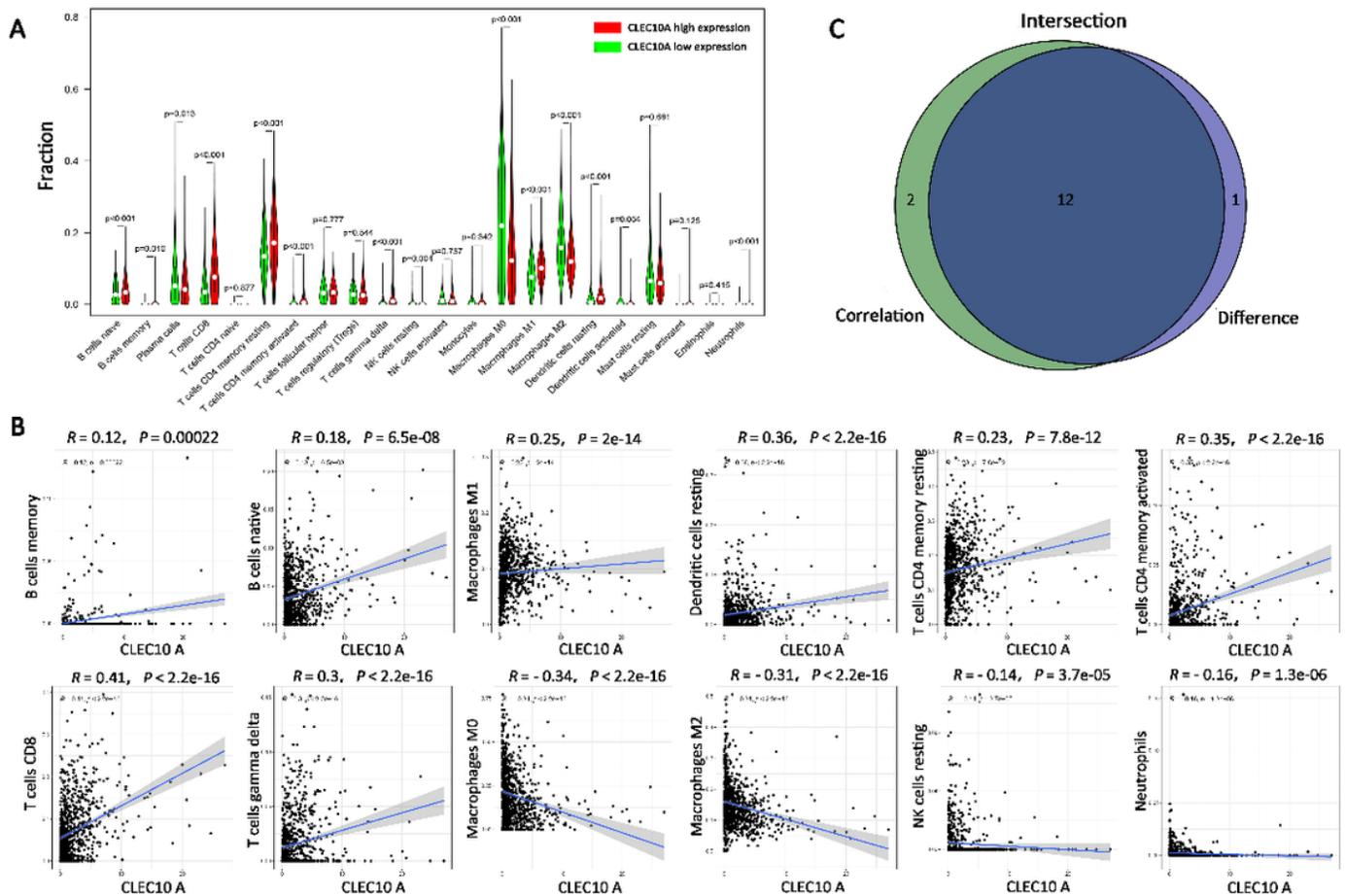
GSEA for samples with high CLEC10A expression and low expression. (A) The enriched gene sets in C2 KEGG collection by the high CLEC10A expression sample. Only several leading gene sets are shown in plot. (B) Gene sets enriched in C2 KEGG collection for samples with low CLEC10A expression. (C) Enriched gene set in C7 collection and immunological gene set by samples with high CLEC10A expression. Only several leading gene sets are shown in plot. Each line represents a specific gene

set with a unique color, while the up-regulated genes are located on the left side approaching the origin of the coordinates, by contrast, the down-regulated genes are located on the right side of the x-axis. Only gene sets with NOM  $p < 0.05$  and FDR  $q < 0.05$  were considered significant.



**Figure 7**

TICs profile in tumor samples and correlation analysis. (A) Bar graph showing the proportion of 22 kinds of TICs in BRCA tumor samples. Column names are sample IDs. (B) Heatmap showing the correlation between 22 kinds of TICs and numeric in each box indicated the p value of correlation between two kinds of cells. The shade of each tiny color box represented corresponding correlation value between two cells, and Pearson coefficient was used for significance test.



**Figure 8**

Correlation of the proportion of TICs with CLEC10A expression. (A) Violin plot showed the ratio differentiation of 22 kinds of immune cells between LUAD tumor samples with low or high BTK expression relative to the median of BTK expression level, and Wilcoxon rank sum was used for the significance test. (B) Scatter plot demonstrating the correlation between 12 kinds of immune cells with CLEC10A expression. The blue line in each graph fits a linear model that indicates the linear relationship between the proportion of immune cells and BTK expression, and the Pearson coefficient was used for correlation testing. (C) Venn plot demonstrated the immune cells shared in the violin plot and scatter plot.

## Supplementary Files

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

- [SupplementTable1.ClinicopathologicalcharacteristicsstatisticsofBRCApatientsfromTCGA..docx](#)
- [SupplementTable2.Correlationoftumormicroenvironmentscorewithclinicopathologicalcharacteristicsofbreastcancer..docx](#)
- [SupplementTable3.CorrelationofCLEC10Aexpressionwithclinicopathologicalcharacteristicsofbreastcancer..docx](#)
- [SupplementTable4.Enrichedgenesets..docx](#)