Identification of prognostic genes from the breast cancer microenvironment

CURRENT STATUS: Posted

Xiao-lu Zhang, Jia Wang, Cun Liu, Chun-di Gao, Jing Zhuang, Li-juan Liu, Xiao-ran Ma, Ji-biao Wu, Peng Sun, Changgang sun

Xiao-lu Zhang
College of First Clinical Medicine, Shandong University of Traditional Chinese Medicine

Jia Wang
College of Traditional Chinese Medicine, Shandong University of Traditional Chinese Medicine

Cun Liu
College of First Clinical Medicine, Shandong University of Traditional Chinese Medicine

Chun-di Gao
College of First Clinical Medicine, Shandong University of Traditional Chinese Medicine

Jing Zhuang
Department of Oncology, Weifang Traditional Chinese Hospital

Li-juan Liu
Department of Oncology, Weifang Traditional Chinese Hospital

Xiao-ran Ma
College of First Clinical Medicine, Shandong University of Traditional Chinese Medicine

Ji-biao Wu
College of Traditional Chinese Medicine, Shandong University of Traditional Chinese Medicine

Peng Sun
College of Pharmacy, Shandong University of Traditional Chinese Medicine

Changgang sun
Department of Oncology, Affiliated Hospital of Shandong University of Traditional Chinese Medicine

Email: scgdoctor@126.com
Corresponding Author
ORCiD: https://orcid.org/0000-0002-6648-3602
Subject Areas

**Oncology**

Keywords

*Tumor microenvironment, Breast cancer, Immune scores, Stromal sores, Protein-protein interaction, Prognosis,*
Abstract

Background Prognostic evaluation of breast cancer is crucial in deciding the course of clinical treatment. The levels of immune cells and stromal cells in tumor microenvironment provide a new approach to evaluate the prognosis of breast cancer.

Method: We obtained mRNA gene expression profiles of breast cancer patients from The Cancer Genome Atlas (TCGA), evaluated immune and stromal scores in tumor microenvironment (TME) using the ESTIMATE algorithm. We also constructed a protein-protein interaction (PPI) network, and performed function enrichment analysis and prognostic analysis.

Results A total of 898 samples were divided into two groups corresponding to high and low score for both stromal as well as immune scores. 247 differentially expressed genes were identified, most of which were associated with T cell activation, lymphocyte differentiation, and cytokine receptor activity. We finally found 10 hub genes. Among them, we were able to confirm the following genes significantly association with prognosis: CCR5, CD2, CD3E, and CD5. Two important modules were isolated from the PPI network and for further analysis.

Conclusion This study identified a group of genes significantly associated with the prognosis of breast cancer, and their efficacy still requires further clinical data validation.

Background

In recent years, therapies of breast cancer have evolved to include both clinical pathology as well as the molecular features of tumor tissue to formulate a treatment regimen. These treatment regimens combine chemotherapy, anti-hormones, and targeted drugs to induce tumor cell death [1]. These strategies rely on intrinsic behavior of cancer cells. However, in the later stages of the treatment, many breast cancer patients face a poor prognosis due to drug resistance, recurrence, and metastasis and heterogeneity of the disease, all of which are considered as serious challenges to the treatment of breast cancer [2]. Along with cancer cells, the TME also plays a role in driving tumor progression. This association has given rise to the term, “the hallmarks of cancer” [3, 4]. Tumor develops in a complex tissue environment, and tumor cells adapt to adverse conditions more rapidly than normal cells, leading to rapid progression of the disease. Similarly, other cells and molecular "participants" in the environment also affect tumor cells [5].

Soluble factors, stromal cells, extracellular matrix, and the physical state of the TME can affect the behavior of a solid tumor in a complex manner. Therefore, TME is now considered as a hallmark of cancer [2]. Studies suggest that the cytokine TGFβ which is a soluble factor possesses tumor suppressive properties in the normal physiological setting, but promotes growth, progression, and invasion of the tumor during malignant progression [6]. Stromal cells include fibroblasts, (pre-) adipocytes, cells of the vascular system (endothelial cells), and immune cells [7]. Among the stromal cells, cancer associated fibroblasts excrete TGFβ via autocrine and paracrine secretion, thus giving rise to a tumor-promoting microenvironment [8]. In addition to that, a variety of tumor-associated leukocytes and stromal cells are also associated with tumor angiogenesis. Numerous tumor-associated immune cells can help tumors to escape immunosurveillance and facilitate maintenance of an immunosuppressive environment [5]. Taken together, the interactions between TME and tumor cells can represent the evasion of key immune factors, physiological tolerance, and local and systemic infiltration of malignant cells.

Studies have shown that immune or stromal "signatures" of breast cancer could be associated with clinical features of the disease [9, 10], and high levels of immune cell infiltration are associated with better prognosis [11]. Therefore, we hypothesize that the immune and stromal components in the TME may influence specific subtypes of breast cancer and patients’ outcomes, but we are not certain about the extent of their influence. At the same time, under the influence of environment and self-injury factors, the stable expression of cell genome plays an important role in controlling the occurrence and development of a tumor. A study has shown that the
mRNA expression profiles in breast cancer patients contain both ESR1 and EGFR2 using prognostic and predictive tests on hormones and chemotherapy agents [12]. These findings prompted us to study the mRNA expression levels for predicting the prognosis in breast cancer patients.

The “ESTIMATE” algorithm has been designed and developed based on the global gene expression profiles in the TCGA database [13]. According to the gene expression characteristics of immune cells and stromal cells in TME, the algorithm calculates the corresponding score to estimate the level of immune cells and stromal cells in tumor tissue, and predicts the immune infiltration of non-tumor cells. This information can then be used to understand disease progression and to identify the patients which would benefit from current treatment. Subsequent studies have also shown the effectiveness of this algorithm [14, 15]. Therefore, we designed a procedure for transcriptome gene analysis to identify the prognostic genes of breast cancer, which can help in selecting a clinical treatment that is best suited for a patient.

**Methods**

2.1 Data collection

We obtained the mRNA-seq data of breast cancer patients from TCGA database through the Genomic Data Commons database (GDC, https://portal.gdc.cancer.gov/). Clinical information such as TNM stages, pathological stage, and survival status was also collected. Samples containing two or more tumor tissues and incomplete clinical data were excluded.

2.2 Correlation and survival analysis

The ESTIMATE algorithm was used to calculate the immune and stromal scores of 898 breast cancer patients. Based on those scores, we separately analyzed T, N, M, stages as well as different pathological stages of the samples. Additionally, we used the “survival” R package to assess the correlation between immune scores, stromal scores, and overall survival.

2.3 Differential expression analysis

Patients with breast cancer were grouped into high immune/stromal and low immune/stromal groups according to the median values obtained for the above scores. In the next step, differentially expressed genes (DEGs) from high and low score groups were identified using the “Limma” R package. In accordance with the previously published studies, adjusted p values were used in the false discovery rate (FDR) to obtain accurate results [16]. The screening criteria were set as FDR < 0.05 and |logFC (fold change)| > 1. The “PHEATMAP” R package was used to draw a heat map to indicate the significant differences observed in the two groups.

2.4 Functional enrichment analysis

Functional enrichment analysis of DEGs was performed using the R package “Cluster Profiler” to identify the Gene Ontology (GO) categories; including biological processes (BP), molecular functions (MF), and cellular components (CC). This package can also be utilized to carry out pathway enrichment analysis using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database, with p<0.05 considered as statistically significant.

2.5 Construction of protein-protein interaction network

In order to understand the potential relationship between DEGs, we obtained the protein-protein interactions (reliability > 0.7) from the STRING (https://string-db.org/) database, and also extracted the relevant PPI network data. The concrete default parameters to identify densely connected modules in the PPI network were set as “Degree Cutoff = 2”, “Node Score Cutoff = 0.2”, “K-Core = 2”, and “Max. Depth = 100” in the Molecular Complex Detection (MCODE) plugin of Cytoscape. The cytoHubba plugin in Cytoscape software was used to obtain the top 10 hub genes using degree ranking.

2.6 Survival analysis
We used the survival package in R to analyze the relationship between expression of the hub genes and overall survival rates of breast cancer patients with the help of the logrank test. p<0.05 was considered to be statistically significant.

### Results

3.1 Relationship between T, N, M stages (or pathological stages) and immune, stromal, and ESTIMATE scores

The primary pathologic and clinical data of 898 cases with breast cancer were obtained from the TCGA database after exclusion of samples with non-conforming tumor tissue and incomplete clinical data. We used the ESTIMATE algorithm to estimate the range of immune score (-1284.33 to 3659.11), the range of stromal score -(2068.96 to 2098.24), and the range of estimation score -(2926.51 to 5331.17) (Table 1). Additionally, the immune score remained unchanged in T and M stage or pathological stage, while the stromal score showed a decrease with the increase in T stage and pathological stage, and no correlation was seen with the M stage. Besides a decrease in the ESTIMATE score as T stage progressed, there was no other correlation with changes in M stage and pathological stage (Figure 1).

Table 1 Immune scores, stromal scores, and clinical data of patients with breast cancer.
| Characteristic | n   | Stromal score (range) | Immune score (range) |
|---------------|-----|-----------------------|----------------------|
| **Age**       |     |                       |                      |
| ≤60           | 510 | -2068.96 to 2056.97   | -1284.33 to 3113.22 |
| 60            | 388 | -1784.53 to 2098.24   | -1070.49 to 3659.11 |
| **Stage**     |     |                       |                      |
| I             | 158 | -1784.53 to 2098.24   | -1070.49 to 3070.83 |
| II            | 524 | -2068.96 to 2094.93   | -1284.33 to 3659.11 |
| III           | 199 | -1522.73 to 1931.95   | -762.70 to 3113.22  |
| IV            | 17  | -783.05 to 1627.01    | -734.57 to 2490.13  |
| **T stage**   |     |                       |                      |
| T1            | 233 | -1784.53 to 2098.24   | -1070.49 to 3070.83 |
| T2            | 531 | -2068.96 to 2094.93   | -1284.33 to 3659.11 |
| T3            | 102 | -1185.72 to 1702.80   | -734.57 to 3113.22  |
| T4            | 32  | -1522.73 to 1299.00   | -553.68 to 2059.34  |
| **N stage**   |     |                       |                      |
| N0            | 445 | -2068.96 to 2098.24   | -1284.33 to 3659.11 |
| N1            | 296 | -1267.08 to 2056.97   | -1128.40 to 3113.22 |
| N2            | 103 | -1142.45 to 1931.95   | -762.70 to 2798.21  |
| N3            | 54  | -1055.12 to 1563.09   | -734.57 to 2407.86  |
| **M stage**   |     |                       |                      |
| M0            | 881 | -2068.96 to 2098.24   | -1284.33 to 3659.11 |
| M1            | 17  | -783.05 to 1627.01    | -734.57 to 2490.13  |

3.2 Higher immune score is associated with better prognosis

In order to explore the potential correlation between the differences in immune, stromal, and ESTIMATE scores
and the prognosis of breast cancer patients, and previous studies have suggested that overall survival and immune score have statistical significance [11]. The samples were then divided into two groups according to their immune scores as high immune score group and low immune score group. According to the Kaplan-Meier survival curve, higher immune score predicts better prognosis for breast cancer patients (Figure 2a, p = 0.018). In contrast, there was no correlation between stromal (Figure 2b, p = 0.379) and ESTIMATE score (Figure 2c, p = 0.281) and the overall survival of breast cancer patients.

3.3 Identification of differentially expressed genes

Figure 3a shows the heat map of 888 DEGs in the high and low immune score groups. A total of 790 genes were up-regulated and 98 genes were down-regulated in the high score group when compared with those seen in the low score group. Figure 3b shows the heat map of 797 DEGs in the high and low stromal score groups, including 675 up-regulated and 122 down-regulated genes. Venn diagrams showed that 247 genes were common in both the groups, including 221 up-regulated (Figure 3c) and 26 down-regulated (Figure 3d) genes in the immune and stromal group.

3.4 Enrichment analysis of common genes

GO analyses and KEGG pathways were used to analyze the biological functions of 247 common genes. Most of these genes were associated with several biological processes related to T cell activation. The molecular function process analysis revealed their significant correlation with the cytokine receptor activity. Data from the cellular components process analysis revealed that most of the target genes were localized to the external region of plasma membrane (Figure 4a). Moreover, the KEGG pathways revealed that these genes were enriched for cytokine-cytokine receptor interaction (Figure 4b).

3.5 Protein-protein interaction (PPI) analysis of common genes

We constructed a protein-protein interaction (PPI) network using the STRING database. Screening was carried out with the minimum required interaction score of 0.700 presents interconnections systemly. Cytoscape software was used for visualizing the data (Figure 5a). PTPRC, IL6, CCR5, CD2, CD28, CCR2, CD5, CD3E, CXCR3, and CSF1R were found as the top 10 hub genes in the network by CytoHubba plugin. MCODE was used to identify 135 genes and 9 function modules from the PPI network. We chose the top two modules for further analysis. The major module was called module a (Figure 5b), which included 12 nodes and 66 edges and all the nodes had the same node degree. PTPRC, CD2, and CD28 were connected more effectively in another module, called module b (Figure 5c). Moreover, we also analyzed the biological processes, cellular components, and molecular functional processes associated with the genes in module a. The results showed that these genes were enriched in the cell chemotaxis pathway, external region of plasma membrane, and C-C chemokine receptor activity (Supplementary Figure 1a). The genes in module b are mainly enriched in the biological process of T cell and lymphocyte differentiation and T cell activation (Supplementary Figure 1b).

3.6 Survival analysis of hub genes

Further survival analysis of the hub genes was essential to predict the prognosis of breast cancer patients clinically. The Kaplan-Meier survival curves were constructed for the hub genes. CCR5, CD2, CD3E and CD5 displayed a significant correlation with the prognosis of patients (p<0.05 was considered as statistically significant). The survival curves are shown in figure 6.

Discussion

Past years have witnessed numerous advances in the "precision" medicine for breast cancer treatment. Advances in sequencing technology have made it possible to analyze genetic maps to provide effective predictions for clinical practice. In our current study, we analyzed gene expression profiles of mRNA transcriptome in breast cancer patients to screen for prognostic genes. Previous studies have shown that immune or stromal cell levels in breast cancer are associated with certain clinical features [9, 10]. We used
ESTIMATE algorithm to calculate the effect of immune and stromal scores on clinical features of the disease. The results revealed that there was no significant correlation between the immune scores and tumor stages and grades, whereas stromal scores were associated with a specific tumor stage and pathological stage. In addition to that, other studies have suggested that high levels of immune cell infiltration are associated with a better survival rate [11]. Our results from the survival analysis on the immune score are in agreement with these observations, and also indicate that stromal score is not associated with survival rate of the patients. However, we were unable to determine the association between immune-specific and stromal-specific genes and prognosis of breast cancer.

We also explored the predictive value of immune and stromal cells. A total of 247 common genes were identified by comparing the DEGs between the high and low immune score groups and the high and low stromal score groups, revealing 221 up-regulated and 26 down-regulated genes. We found that these common genes were involved in T cell activation and lymphocyte differentiation occurring within the TME. The molecular functional processes of these genes were primarily associated with cytokine receptor activity. T cell activation fractions are positively correlated with overall survival of breast cancer patients [17], and studies have also demonstrated the use of chimeric cytokines that can enhance the proliferation and activation of tumor-specific T cells exclusively in TME, resulting in excellent antitumor activity [18]. The metabolic pathway analysis suggested that target genes were enriched in the cytokine-cytokine receptor interaction pathway. Recent studies suggest that unique metabolic changes affect T cell activation and differentiation, and the effect of TME on T cell metabolic reprogramming may provide excellent anticancer strategies to enhance T cell immunity [19]. Further research is required on how to support the clinically. These processes may provide theoretical basis for understanding the role of immune and stromal cells in TME.

We used the high conservation of hub genes for prognostic analysis. CCR5, CD2, CD5, and CD3E were identified as the genes associated with overall survival. Studies have shown that CCR5 targeting could be an effective strategy against metastasis of breast cancer into bones. Blocking CCR5 inhibits the proliferation, colony formation, and migration of metastatic breast cancer cells, and also induces G1 cell cycle arrest and apoptosis [20]. However, in our survival curve, the survival rate of patients with high CCR5 expression increased significantly during the 10-year survival period, and decreased thereafter. Studies have shown that intraosseous injection of the 4T1.3 clone causes smaller tumors in mice lacking CCR5 or receiving the CCR5 antagonist, when compared with those seen in the wild-type mice [21]. Reintroducing CCR5 expression into CCR5-negative breast cancer cells can promote tumor metastasis and induces the expression and activity of DNA repair genes that are selectively expressed in cancerous cells [22]. These studies may provide explanation for the decline in survival rates. CD2 is also significantly associated with distant metastasis-free survival in HR/HER2+ breast cancer [23], suggesting that CD2 could be a valuable target gene for developing breast cancer treatments. Interestingly, experiments have shown that both the extracellular domain and the cytoplasmic domain of CD2 interact with CD5, suggesting a specific and close association between the two molecules, which might be related to the fine tuning of the lymphocyte signaling [24]. CD5(+) B cells are widely believed to be related to the poor survival rate seen in diffuse large B-cell tumors and play a key role in promoting tumor growth [25, 26]. But, a study about enhanced antitumor immunity in CD5 knockout mice states that CD5 enhances TCR signaling (ERK activation) and markers of T cell activation [27]. There is a need to study the potential role of the functional blockade of CD5 in increasing the antitumor activity of tumor-specific T cells. CD3E, which is associated with tumor degradation, is also involved in immune response or lymphocyte activation in locally advanced breast cancer [28]. All of these observations confirm the predictive value of the four prognostic genes.

We constructed a PPI network to understand the interactions and functions of the common genes. We also identified two crucial gene modules, which showed enrichment for the “immune/inflammatory response”, “cell chemotaxis”, and “chemokine receptor activity”. Chemotaxis is the directional movement of cells toward a chemical stimulation. It is important in immune response, tumor metastasis, and other physiological processes [29]. CCR2, CCR5, CXCR3, PTPRC, CD2, and CD28 were the key nodes in modules a and b. PTPRC, CD28, CCR2, and CXCR3 affect the prognosis of multiple cancers such as papillary thyroid carcinoma [30], pancreatic cancer [31], non-Hodgkin lymphoma [31], ovarian cancer [32], and breast cancer [33, 34] through different immune mechanisms. It is worth noting that down-regulation of CD28 is a hallmark of aging T cells [35], which indicates that the immune functions in the TME could be weakened. A study has also shown that aging is a high risk factor for the occurrence and poor prognosis of breast cancer [36], and changes in the TME due to aging may reduce
the effectiveness of immunotherapy [35]. We are not certain as to whether aging can affect the prognostic factors in our study.

The breast cancer microenvironment needs more comprehensive scientific assessment and research. Fortunately, the high-throughput sequencing data available in the databases were an excellent set of resources for our study. However, our study has some drawbacks. In order to reduce the error rate of the algorithm, we plan to test its validity in clinical experiments. In addition to that, we believe that combining transcriptome genes from this study with epigenetic mechanisms and proteomics can improve the accuracy of the screening evaluation.

**Conclusion**

Our study identified a group of prognostic genes from the breast cancer microenvironment using ESTIMATE algorithm. We also constructed a relevant PPI network and performed survival analysis. Further studies are required to understand the physiological mechanism of these prognostic genes. We hope that our results can provide a novel perspective for the intervention of breast cancer.

**Declarations**

**Ethics approval and consent to participate**

Not applicable.

**Consent for publication**

The authors consent for publication.

**Availability of data and materials**

All data analyzed in this study are from open data (freely available to anyone) at TCGA database: “https://xenabrowser.net/datapages/”

**Competing interests**

The authors declare that they have no competing interests.

**Funding**

This work is supported by the grants from National Natural Science Foundation of China (81673799,81703915,81973677).

**Authors' contributions**

CGS, XLZ and JW design research topic. XLZ, XRM, JZ and CL obtained data on Breast cancer. JW, CDG, PS and JBW build the related network and analyze the data. XLZ and JW write the manuscript. CGS, CL, LjL, and CDG revise the manuscript. All authors read and approved the final manuscript.

**Acknowledgements**

Not applicable.

**Author information**

Xiaolu Zhang and Jia Wang contributed equally to this work
Affiliations

College of First Clinical Medicine, Shandong University of Traditional Chinese Medicine, Jinan250014, Shandong, China
Xiao-lu Zhang, Cun Liu, Chun-di Gao and Xiao-ran Ma

College of Traditional Chinese Medicine, Shandong University of Traditional Chinese Medicine, Jinan250014, Shandong, China
Jia Wang and Ji-biao Wu

Department of Oncology, Weifang Traditional Chinese Hospital, Weifang261000, Shandong, China
Jing Zhuang and Li-juan Liu

College of Pharmacy, Shandong University of Traditional Chinese Medicine, Jinan250014, Shandong, China
Peng Sun

Innovative Institute of Chinese Medicine and Pharmacy, Shandong University of Traditional Chinese Medicine, Jinan250014, Shandong, China
Chang-gang Sun

Abbreviations

TME tumor microenvironment
TCGA The Cancer Genome Atlas
PPI protein-protein interaction
DEGs differentially expressed genes
FDR false discovery rate
GO Gene Ontology
KEGG Kyoto Encyclopedia of Genes and Genomes
BP biological processes
CC cellular components
MF molecular functions

References

1. BM T, DG H: Pathologic diagnosis of breast cancer patients: evolution of the traditional clinical-pathologic paradigm toward "precision" cancer therapy. Biotechnic & histochemistry : official publication of the Biological Stain Commission 2017, 92(3):175-200.
2. K V, N S, B B, J SR: Tumor microenvironment-mediated chemoresistance in breast cancer. Breast (Edinburgh, Scotland) 2016, 30:92-100.
3. D H, LM C: Accessories to the crime: functions of cells recruited to the tumor
26. **Zhang C, Xin H, Zhang W, Yazaki P, Zhang Z, Ke L, Li W, Lee H, Kwak L, Forman S et al:** CD5 Binds to Interleukin-6 and Induces a Feed-Forward Loop with the Transcription Factor STAT3 in B Cells to Promote Cancer. *Immunity* 2016, 44(4):913-923.

27. **P Z, L L, S Z, L Q, Z Q, X L, B M, H Z:** CD5 expression correlates with inferior survival and enhances the negative effect of p53 overexpression in diffuse large B-cell lymphoma.
27. F A, M R, R F, R Z, M Z, P J F, S M V, Y N, M E-H, X Z et al: **CD5 blockade enhances ex vivo CD8 T cell activation and tumour cell cytotoxicity.** European journal of immunology 2020, 50(5):695-704.

28. MLH K, RADC V, VP A, RA R, LGCA L, LM K, AP C, CAB P, PAMP S, G E et al: **Stromal Cell Signature Associated with Response to Neoadjuvant Chemotherapy in Locally Advanced Breast Cancer.** Cells 2019, 8(12).

29. J C, B R, H L, L M, P C, N S, J L, L Z, ME C: **Leveraging the IncuCyte Technology for Higher-Throughput and Automated Chemotaxis Assays for Target Validation and Compound Characterization.** SLAS discovery : advancing life sciences R & D 2018, 23(2):122-131.

30. Y W, J H, KE V, S Z, W L, Y Z, E J, J S, Q F, J M et al: **Upregulation Of Protein Tyrosine Phosphatase Receptor Type C Associates To The Combination Of Hashimoto’s Thyroiditis And Papillary Thyroid Carcinoma And Is Predictive Of A Poor Prognosis.** OncoTargets and therapy 2019, 12:8479-8489.

31. F R, FBT K, M C, S G, C H, BL C, P D, S E, S K: **PD1-CD28 Fusion Protein Enables CD4+ T Cell Help for Adoptive T Cell Therapy in Models of Pancreatic Cancer and Non-hodgkin Lymphoma.** Frontiers in immunology 2018, 9:1955.

32. S D, X D, J H, J J, W W, J Y, W Z, C C: **MicroRNA-135a-3p is downregulated and serves as a tumour suppressor in ovarian cancer by targeting CCR2.** Biomedicine & pharmacotheraphy = Biomedecine & pharmacotherapie 2018, 107:712-720.

33. J H, X L, X G, Q G, C X, Z Z, Y X, T X, L Z: **CCR2The 3’UTR functions as a competing endogenous RNA to inhibit breast cancer metastasis.** Journal of cell science 2017, 130(19):3399-3413.

34. K G, G F, Q Y, L S, K Z, F S, M S, S R: **CXCR4 and CXCR3 are two distinct prognostic biomarkers in breast cancer: Database mining for CXCR family members.** Molecular medicine reports 2019, 20(6):4791-4802.

35. WX H, JH K, M H, K F, M D: **The Evolving Role of CD8CD28 Immunosenescent T Cells in Cancer Immunology.** International journal of molecular sciences 2019, 20(11).

36. X M, C L, X X, L L, C G, J Z, H L, F F, C Z, Z L et al: **Biomarker expression analysis in different age groups revealed age was a risk factor for breast cancer.** Journal of cellular physiology 2020, 235(5):4268-4278.

**Figures**
The relationship between immune, stromal, and ESTIMATE scores and T stage, M stage, and tumor stage. Distribution of immune scores plotted against T stage (a), M stage (d), and pathological stage (g). Distribution of stromal scores plotted against T stage (b), M stage (e), and pathological stage (h). Distribution of ESTIMATE scores plotted against T stage (c), M stage (f), and pathological stage (i).
Association of immune, stromal, and ESTIMATE scores with overall survival. (a) Elevated immune scores were associated with a better prognosis. Stromal scores (b) and ESTIMATE scores (c) were not associated with overall survival.
Comparison of the gene expression profile with immune and stromal scores. (a) Based on immune score comparisons, 790 genes upregulated and 98 genes downregulated were in the high score group when compared with those seen in the low score group. (b) Based on stromal score comparisons, 675 genes upregulated and 122 genes downregulated were in the high score group when compared with those seen in the low score group. (c) A total of 221 genes were upregulated in both immune and stromal score groups. (d) A total of 26 genes were downregulated in both immune and stromal score groups.
Figure 4
GO(a) and KEGG(b) analysis of the 247 common genes.
Figure 5

PPI network and two clustering module. (a) The PPI network of 135 DEGs. (b) Module of the highest MCODE scores (MCODE score = 10.154). (c) Module of the second MCODE scores (MCODE score = 5.4). The node color deepens as the number of interacting proteins increases.
Figure 6
Four hub genes significantly associated with prognosis of BRCA patients: (a) CCR5, (b) CD2, (c) CD3E and, (d) CD5.

Supplementary Files

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

- [SupplementaryFigure1.pdf](#)