# Gene Expression Patterns Associated With Tumor-Infiltrating CD4+ and CD8+ T Cells In Invasive Breast Carcinomas

**Zhanwei Wang**
  Huzhou Central Hospital

**Jiamin Xu**
  Huzhou University

**Xi Yang**
  Huzhou Central Hospital

**Yuefen Pan**
  Huzhou Central Hospital

**Junjun Shen**
  Huzhou Central Hospital

**Jin Liu**
  Huzhou Central Hospital

**Shuwen Han** ( ✉ shuwenhan985@163.com )
  https://orcid.org/0000-0001-6180-9565

---

---

# Abstract

**Objective** This study investigated the gene expression patterns associated with tumor-infiltrating CD4 + and CD8 + T cells in invasive breast carcinomas.

**Methods** The gene expression data and corresponding clinical phenotype data from the Cancer Genome Atlas Breast Invasive Carcinoma (TCGA-BRCA) were downloaded. The stromal and immune score were calculated using ESTIMATE. The differentially expressed genes (DEGs) with a high vs. low stromal score and a high vs. low immune score were screened and then functionally enriched. The tumor-infiltrating immune cells were investigated using the Cibersort algorithm, and the CD4 + and CD8 + T cell-related genes were identified using a Spearman correlation test of infiltrating abundance with the DEGs. Moreover, the miRNA-mRNA pairs and lncRNA-miRNA pairs were predicted to construct the competing endogenous RNAs (ceRNA) network. Kaplan-Meier (K-M) survival curves were also plotted.

**Results** In total, 478 DEGs with a high vs. low stromal score and 796 DEGs with a high vs. low immune score were identified. In addition, 39 CD4 + T cell-related genes and 78 CD8 + T cell-related genes were identified; of these, 14 genes were significantly associated with the prognosis of BRCA patients. Moreover, for CD4 + T cell-related genes, the chr22-38_28785274-29006793.1−miR-34a/c-5p−CAPN6 axis was identified from the ceRNA network, whereas the chr22-38_28785274-29006793.1−miR-494-3p−SLC9A7 axis was identified for CD8 + T cell-related genes.

**Conclusions** The chr22-38_28785274-29006793.1−miR-34a/c-5p−CAPN6 axis and the chr22-38_28785274-29006793.1−miR-494-3p−SLC9A7 axis might regulate cellular activities associated with CD4 + and CD8 + T cell infiltration, respectively, in BRCA.

# Background

Breast carcinoma is one of the most common tumors in women, accounting for approximately 25% of all cancers; it is the second leading cause of cancer-related death in women worldwide [1]. Breast carcinomas can be classified into two main categories, including *in situ* carcinomas and invasive carcinomas [2]. Invasive breast carcinoma (BRCA) refers to those tumors that show tumor cell invasion to adjacent tissues of the mammary ducts and display a trend of regional lymph node metastasis [2]. The clinical course and outcome of breast carcinoma differs with diverse immunohistochemical characteristics and histopathological subtypes [3]. Despite the great progress in diagnosis and therapy of breast carcinomas, clinical outcomes remain unsatisfactory, and the 5-year survival rate of female patients with metastatic breast cancer is 27% [4].

Recently, the therapeutic prospects of malignancies, including breast carcinomas, have been remarkably altered with a deeper understanding of tumor-immune interactions and the development of immunological checkpoint inhibitors [5, 6]. Immune checkpoints are defined as various inhibitory pathways mediating self-tolerance and immune responses to limit collateral tissue damage [7]. These pathways can be utilized by tumor cells to escape detection and elimination by the immune system [8].

Programmed death 1 (PD-1) is an immune checkpoint; it suppresses the biological functions of effector T cells. Tumor cells can express PD-L1, a ligand of PD-1, which can inhibit the antitumor immune response by binding to PD-1 [9, 10]. Tumor-infiltrating lymphocytes (TILs) are immune cells that have migrated to the tumor tissue microenvironment, indicating an antitumor immune response [6, 11]. Shi et al. revealed that the distribution of TILs differs in diverse subtypes; patients with triple-negative breast carcinomas show more PD-1 + exhausted TILs, suggesting an immunosuppressive microenvironment [12]. CD4 + and CD8 + T cells are two major lymphocyte cell types. Matsumoto et al. suggested that increased CD4 + and CD8 + T cell infiltration indicates good prognosis in triple-negative breast carcinomas [13]. Additionally, Su et al. indicated that immunosuppression in breast carcinomas can be reversed by blocking the recruitment of naive CD4 + T cells, which might be a promising strategy for anticancer immunotherapy [14].

The growth and metastasis of breast carcinomas involves dynamic processes that are affected by the tumor-immune microenvironment, in which the gene expression of resident cells exhibit obvious alterations [8, 15]. Genetic alterations in tumor progression facilitate the ectopic expression of normally silent genes in tumors, with potential carcinogenesis [16]. Masjedi et al. demonstrated that olfactory receptor genes are upregulated in BRCA and are implicated in the proliferation and progression of BRCA [17]. The role of tumor-immune interactions in breast carcinomas has received more attention. Nevertheless, the major molecular changes in the immune response accompanying breast carcinoma progression, especially in CD4 + and CD8 + T cells, is still largely unknown. Specific TIL subsets were indicated to be clinically significant and could be used to predict treatment responses. Cytotoxic CD8 + T lymphocytes could alternatively identify and eliminate tumor cells. While it could be deactivated by T reg cells, and the role of co-stimulatory proteins on antigen-presenting cells (APC) would subsequently decrease. That is to say, CD8 + T lymphocytes can be inhibited by T reg cells [18]. Breast carcinomas can generate substances that affect APCs and alter T cell types [19]. Disis et al. showed that the cause of robust TILs in triple-negative breast carcinomas might be the elevated mutations producing a neoantigen signature [20]. Therefore, we explored the gene expression pattern associated with tumor-infiltrating CD4 + and CD8 + T cells of invasive breast carcinomas. The results are expected to provide therapeutic targets for clinical treatment of BRCA.

# Materials And Methods

### Data acquisition

The RNA-sequencing (RNA-seq) Fragments per Kilobase of transcript per Million mapped reads (FPKM) and corresponding clinical phenotype data from The Cancer Genome Atlas Breast Invasive Carcinoma (TCGA-BRCA) data collection were downloaded from the University of California Santa Cruz (UCSC, https://xenabrowser.net/) Genome Browser database. There were 1194 samples, including 1082 BRCA samples and 112 normal samples (data acquisition on 07-18-2019). Those data were analyzed according to the presupposed workflow (Figure 1).

## Differentially expressed gene (DEG) screening and functional enrichment

The probes were first annotated according to the annotation files, and were filtered based on whether they matched the gene symbol. The mean value was selected when multiple probes matched to the same gene. The stromal score and immune score of tumor tissue were calculated using the ESTIMATE (Estimation of Stromal and Immune cells in Malignant Tumor tissues using Expression data) package [21]. Then, samples were split into high stromal/immune score and low stromal/immune score groups according to the median value of the stromal/immune score, respectively. Then, the DEGs with a high stromal/immune score vs. a low stromal/immune score were screened using the Limma package [22] (Version 3.10.3) with |log fold change (FC)| > 0.263 and $P$ value < 0.05. Finally, the overlapping DEGs between the two groups were selected and used in the following analysis.

## Functional enrichment analysis

The biological processes terms of the Gene Ontology annotation and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were analyzed to investigate the functions of the upregulated and downregulated overlapping DEGs using clusterProfiler [23] (Version 3.2.11) in R package. The number of enriched genes was set as: count ≥ 2. The terms with $P < 0.05$ were considered to be significant results.

## Identification of CD4+ and CD8+ T cell-related genes

Based on the Cibersort algorithm [24], the tumor-infiltrating immune cells in BRCA were investigated to estimate the infiltrating abundance of six immune cell types, including B cells, CD4+ T cells, CD8+ T cells, neutrophils, macrophages, and dendritic cells. Then, the correlation coefficient (r) between overlapping DEGs and infiltrating abundance of immune cells (CD4+ and CD8+ T cells) was calculated by using the Spearman correlation test, and CD4+ and CD8+ T cell-related genes were identified with |r| > 0.15.

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

The interactions between CD4+ T cell-related genes were retrieved from the STRING database [25] with a PPI score setting as 0.15 (low confidence), and the species was set as human. Based on the retrieved PPIs, CD4+ T cell-related PPI network visualization was performed using Cytoscape [26] (version: 3.2.0). The CD8+ T cell-related PPI network was also constructed using the same method.

## Construction of the competing endogenous RNAs (ceRNA) network

The prediction of miRNA-mRNA interactions was conducted for CD4+ T cell-related genes using miRWalk 3.0 [27], and the species was set as human. The miRNA-mRNA interactions with a score > 0.95, and those that existed in both the TargetScan and miRDB databases were selected. Next, based on DIANA-LncBase v.2 [28], the prediction of lncRNA-miRNA interactions was conducted and the interactions with score = 1 were selected. The construction of a CD4+ T cell-related ceRNA network was completed by integrating the obtained lncRNA-miRNA interactions and miRNA-mRNA interactions. Similarly, a CD8+ T cell-related ceRNA network was also constructed using the same method.

## Construction of a chemical-target network

The genes and chemicals associated with breast neoplasms were explored from the comparative toxicogenomics database (CTD) [29] using "breast neoplasms" as the search keywords. Then, the overlapping genes between breast neoplasm-associated genes and genes in the CD4+ T cell-related ceRNA network were selected and used to screen chemical-target pairs. Then, the CD4+ T cell-related chemical-target network was visualized using Cytoscape. Similarly, the CD8+ T cell-related chemical-target network was also constructed using the same method.

## Survival analysis

The overall survival (OS) and OS status in the clinical phenotype data were used to perform survival analysis. Briefly, the immune cell-related genes (CD4+ and CD8+ T cell-related genes) were split into high-expression and low-expression groups based on the median expression value, together with a log-rank statistical test. The cut-off was set as a $P$ value < 0.05 to select the genes significantly associated with prognosis, and then Kaplan-Meier (K-M) survival curves were plotted.

# Results

### Differences in gene expression between high and low stromal scores

The stromal score was calculated to predict the non-tumor cell infiltration [21]. Based on the stromal score, the samples were split into high and low stromal score groups, and a total of 478 genes were significantly differentially expressed in the two groups, including 104 upregulated and 374 downregulated genes (Figure 2A). These genes were considered to be associated with the stroma in tumor tissue.

### Differences in gene expression between high and low immune scores

The immune score refers to the infiltration of immune cells in tumor tissue, which is considered an available indicator of prognosis in tumors [30]. Based on the immune score, the samples were split into high and low immune score groups; 796 genes were found to be significantly differentially expressed in the 2 groups (Figure 2B). Of these, 503 genes were upregulated whereas 293 genes were downregulated, suggesting that these genes might be implicated in the immune status in the tumor microenvironment.

### Overlapping genes between the stroma score-related and immune score-related genes

The overlapping genes between the stroma score-related and immune score-related genes were identified by VENN analysis; 167 genes were obtained, including 58 upregulated and 109 downregulated genes (Figure 2C). The functions of these overlapping genes were further investigated, and the downregulated genes were found to be significantly implicated in one KEGG pathway (hsa05146, amoebiasis) and five biological process terms, for example: GO:0006885~regulation of pH (Table 1).

### Immune cell infiltration in BRCA

The infiltration abundance of six immune cell types in BRCA was analyzed using the Cibersort algorithm. From the bar charts of immune cell subset proportions (Figure 3), CD8+ T cells and activated memory CD4+ T cells accounted for a large proportion of immune cell infiltration in BRCA. Therefore, CD8+ and CD4 T+ cells were selected in the following analysis.

## Identification of CD4+ and CD8+ T cell-related genes

The infiltrating abundance of CD4+ and CD8+ T cells is related to prognosis in BRCA [13]. The infiltrating abundance of CD4+ and CD8+ T cells was calculated using the Cibersort algorithm. Then, the immune cell-related genes were identified using the Spearman correlation test between overlapping genes and the infiltrating abundance of immune cells. A total of 39 CD4+ T cell-related genes and 78 CD8+ T cell-related genes were identified. These genes were regarded as crucial genes involved in CD4+ and CD8+ T cell infiltration.

## PPI network of CD4+ and CD8+ T cell-related genes

Proteins and their functional interactions form the backbone of cellular machinery, and connectivity networks are conducive to fully understand biological phenomena [31]. For CD4+ T cell-related genes, the PPI network contained 13 genes and 11 interactions (Figure 4A). Of the 13 genes, 5 genes were upregulated whereas 8 genes were downregulated. For CD8+ T cell-related genes, the PPI network consisted of 56 genes (16 upregulated and 40 downregulated) and 76 interactions (Figure 4B).

## ceRNA network of CD4+ and CD8+ T cell-related genes

For CD4+ T cell-related genes, 17 miRNA-mRNA interactions (Supplemental Figure 1A), and 13 lncRNA-mRNA interactions (Supplemental Table 1) were finally predicted. The ceRNA network is shown in Figure 5A, and contains 2 lncRNAs, 12 miRNAs, and 5 mRNAs, consisting of 25 ceRNA regulatory axes. For example: the chr22-38_28785274-29006793.1−miR-34a/c-5p−CAPN6 axis was identified.

For CD8+ T cell-related genes, a total of 51 miRNA-mRNA pairs (Supplemental Figure 1B), and 43 lncRNA-mRNA pairs were predicted (Supplemental Table 2). The ceRNA network contained 10 lncRNAs, 32 miRNAs, and 14 mRNAs, consisting of 81 ceRNA regulatory axes (Figure 5B). For example: the chr22-38_28785274-29006793.1−miR-494-3p−SLC9A7 axis was identified.

## Breast neoplasm-related chemicals target mRNAs in the ceRNA network

The etiology of many diseases involves the interactions between environmental chemicals and genes that regulate physiological processes [32]. The CTD provides information about chemical−gene/protein-disease relationships [32]. In this study, the chemical−gene interactions involving breast neoplasms were identified from the CTD, and were filtered using the mRNAs in the ceRNA network. A total of 31 chemicals were found to interact with the five genes in the CD4+ T cell-related ceRNA network (Figure 6A, Supplemental Table 3). For example, atrazine (CAS, 1912-24-9) was predicted to result in the increased expression of CAPN6 mRNA.

Similarly, 57 chemicals were found to interact with the 12 genes in the CD8+ T cell-related ceRNA network (Figure 6B, Supplemental Table 4). For example, arsenic (CAS, 7440-38-2) was predicted to affect the methylation of the SLC9A7 gene.

### CD4+ and CD8+ T cell-related genes associated with prognosis of BRCA

To explore the prognostic value of CD4+ and CD8+ T cell-related genes, survival analysis was performed. In total, 14 genes were found to be significantly related to prognosis of BRCA patients; of these, eight genes were commonly related to both CD4+ and CD8+ T cells (Figure 7, Table 2), for example, MUC2. Among the 14 genes, 6 genes were found to be specific to CD4+ or CD8+ T cells; METTL5 and MSTN were CD4+ T cell-related genes, whereas NMNAT3, GNAL, ZBED4, and RDH12 were CD8+ T cell-related genes.

# Discussion

In this study, the RNA-seq data and clinical phenotype data from TCGA-BRCA were used to investigate the gene expression pattern in the immune response in BRCA progression. A total of 478 DEGs with a high vs. low stromal score, and 796 DEGs with a high vs. low immune score were identified, and the overlapping DEGs were found to be implicated in ion transmembrane transport, regulation of pH, and extrinsic apoptotic signaling pathways. In addition, a total of 39 CD4 + T cell-related genes and 78 CD8 + T cell-related genes were identified, of which 14 genes were related to prognosis of BRCA patients, for example, MUC2. Moreover, for CD4 + T cell-related genes, the chr22-38_28785274-29006793.1−miR-34a/c-5p−CAPN6 axis was identified in the ceRNA network and for CD8 + T cell-related genes, the chr22-38_28785274-29006793.1−miR-494-3p−SLC9A7 axis was identified.

MUC2, also termed mucin-2, belongs to the mucin protein family, which are high molecular weight glycoproteins secreted to form an insoluble mucous barrier that protects the gut lumen [33]. MUC2 was found to be expressed in breast mucinous carcinomas, and Matsukita et al. indicated that overexpression of MUC2 in mucinous carcinoma might serve as a barrier to attenuate tumor aggressiveness [34]. Astashchanka et al. showed that MUC2 participates in the regulation of proliferation, apoptosis, and metastasis of breast carcinoma cells, indicating roles of MUC2 in therapy and in clinical outcome prediction in breast carcinoma [35]. In our study, the expression of MUC2 was correlated with both CD4 + and CD8 + T cell infiltration and was associated with prognosis of BRCA patients. Therefore, we speculated that MUC2 might regulate the aggressiveness of BRCA; this process is probably accompanied by T cell infiltration.

For CD4 + T cell-related genes, the chr22-38_28785274-29006793.1−miR-34a/c-5p−CAPN6 axis was identified from the ceRNA network. The miR-34 family (miR-34a/b/c) is composed of tumor suppressors, functioning in inhibiting proliferation, migration and inducing apoptosis, and are also important mediators of the p53 signaling pathway [36, 37]. It has been reported that miR-34a serves as a tumor suppressor in triple-negative breast carcinomas by targeting the proto-oncogene c-SRC [38]. Similarly, Xiao et al. found that miR-34a can inhibit glycolysis and proliferation of breast carcinoma cells in breast

carcinomas by directly targeting lactate dehydrogenase A, whose overexpression is associated with tumor growth and metastasis [39]. CAPN6, also termed calpain-6, is a member of an intracellular cysteine protease family, and this family is found to be abnormally expressed in malignant tumors [40]. Calpains are reported to be involved in various cellular activities in breast carcinomas, including cellular survival, apoptosis and migration [41]. MacLeod et al. suggested that calpain 1 and 2 play a pro-tumorigenic role in HER2 + breast cancer, whereas tumorigenesis can be delayed by disrupting calpain 1 and 2 expression [42]. Calpain 6 has been reported to relate to tumorigenesis and unfavorable prognosis in head and neck squamous cell carcinoma [40], and is considered to be a possible target in the treatment of sarcomas [43]. The role of calpain 6 in breast carcinomas was rarely reported. A previous study showed that calpain 1 and miR-34a/c were associated with kanamycin-induced inner ear cell apoptosis [44]. In our study, CAPN6/calpain 6 was predicted to be a target of miR-34a/c-5p, which was regulated by the lncRNA chr22-38_28785274-29006793.1. The role of the lncRNA chr22-38_28785274-29006793.1 has not yet been reported. We speculated that the chr22-38_28785274-29006793.1−miR-34a/c-5p−CAPN6 axis might regulate cellular activities associated with CD4 + T cell infiltration in BRCA.

The abnormal expression of miR-494 has been reported in various cancers. However, the role of miR-494 in carcinogenesis is contradictory, including a tumor suppressor role [45, 46] and an oncogenic role [47, 48]. Zhan et al. revealed that miR-494 can inhibit the progression and metastasis of breast carcinomas by targeting P21 (RAC1) activated kinase 1 [49]. The proliferation and migration of MDA-MB-231 and MDA-MB-468 breast carcinoma cells can be promoted by highly expressing miR-183 or miR-494 [50]. SLC9A7, also termed NHE7, is a (Na+, K+)/H + exchanger, functioning in regulating cellular pH and ion homeostasis [51]. pH has a crucial role in regulating cell motility and metastasis. The metastatic potential of breast carcinoma cells can be enhanced by exposure to alkaline pH [52]. Onishi et al. found that SLC9A7/NHE7 can promote adhesion, invasion, and oncogenesis of MDA-MB-231 breast carcinoma cells [53]. In our study, the chr22-38_28785274-29006793.1−miR-494-3p−SLC9A7 axis was identified from the CD8 + T cell-related ceRNA network. Hence, we suggested that the chr22-38_28785274-29006793.1−miR-494-3p−SLC9A7 axis might regulate cellular activities associated with CD8 + T cell infiltration in BRCA.

Although several novel findings were found in this study, there were some limitations. (1) Our study preliminarily analyzed the gene expression pattern of tumor-infiltrating CD4 + and CD8 + T cells of BRCA. However, further experiments are needed to confirm the expression of the DEGs and the predicted ceRNA axes. (2) The correlations between prognosis and the 14 identified genes should be further investigated using clinical trials. (3) The predicted chemical−gene interactions should be confirmed to provide research topics for the treatment of BRCA.

## Conclusions

In conclusion, the gene expression patterns of tumor-infiltrating CD4 + and CD8 + T cells in BRCA were identified. MUC2 might be a biomarker to predict the prognosis of BRCA. The chr22-38_28785274-29006793.1−miR-34a/c-5p−CAPN6 axis and the chr22-38_28785274-29006793.1−miR-494-3p−SLC9A7

axis might regulate the cellular activities associated with CD4 + and CD8 + T cell infiltration, respectively, in BRCA.

# Declarations

## Ethics approval and consent to participate

Not applicable.

## Consent for publication

Not applicable.

## Availability of data and materials

The Cancer Genome Atlas Breast Invasive Carcinoma (TCGA-BRCA) data collection were downloaded from the University of California Santa Cruz (UCSC, https://xenabrowser.net/) and Comparative Toxicogenomics Database (http://ctdbase.org/) are available by contacting the author.

## Competing Interests

The authors declare that no conflicts of interest exist.

## Author's Contributions

All authors participated in the conception and design of the study;
Conceived and drafted the manuscript: Wang Zhanwei and Xu Jiamin;

Analyzed data: Yang Xi, Zhou Qing and Liu Jin;
Collated and proofread the literature: Yang Xi and Zhou Qing;

Wrote the paper: Liu Jin and Han Shuwen;
All authors read and approved the paper.

# References

1. DeSantis C, Ma J, Bryan L, Jemal A. Breast cancer statistics, 2013, CA: A Cancer Journal for Clinicians, 64 (2014) 52–62.

2. Bacchi CE, Viana CR, Histopathological and Immunohistochemical Classification of Invasive Breast Carcinomas, in: Breast Diseases, Springer, 2019, pp. 237–247.

3. Makki J, Diversity of breast carcinoma: histological subtypes and clinical relevance, Clinical Medicine Insights: Pathology, 8 (2015) CPath. S31563.

4. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019, CA: a cancer journal for clinicians, 69 (2019) 7–34.

5. Pusztai L, Karn T, Safonov A, Abu-Khalaf MM, Bianchini G. New strategies in breast cancer: immunotherapy. Clin Cancer Res. 2016;22:2105–10.

6. Vonderheide RH, Domchek SM, Clark AS, Immunotherapy for breast cancer: what are we missing?, in, AACR, 2017.

7. Binnewies M, Roberts EW, Kersten K, Chan V, Fearon DF, Merad M, Coussens LM, Gabrilovich DI, Ostrand-Rosenberg S, Hedrick CC. Understanding the tumor immune microenvironment (TIME) for effective therapy. Nature medicine. 2018;24:541–50.

8. Tower H, Ruppert M, Britt K, The Immune Microenvironment of Breast Cancer Progression, Cancers, 11 (2019) 1375.

9. Blank C, Gajewski TF, Mackensen A. Interaction of PD-L1 on tumor cells with PD-1 on tumor-specific T cells as a mechanism of immune evasion: implications for tumor immunotherapy. Cancer Immunol Immunother. 2005;54:307–14.

10. Blank C, Mackensen A, Contribution of the PD-L1/PD-1 pathway to T-cell exhaustion: an update on implications for chronic infections and tumor evasion, Cancer immunology, immunotherapy, 56 (2007) 739–745.

11. Miyan M, Schmidt-Mende J, Kiessling R, Poschke I, de Boniface J. Differential tumor infiltration by T-cells characterizes intrinsic molecular subtypes in breast cancer. Journal of translational medicine. 2016;14:227.

12. Shi F, Chang H, Zhou Q, Zhao Y-J, Wu G-J, Song Q-K. Distribution of CD4(+) and CD8(+) exhausted tumor-infiltrating lymphocytes in molecular subtypes of Chinese breast cancer patients. Onco Targets Ther. 2018;11:6139–45.

13. Matsumoto H, Thike AA, Li H, Yeong J, Koo S-l, Dent RA, Tan PH, Iqbal J. Increased CD4 and CD8-positive T cell infiltrate signifies good prognosis in a subset of triple-negative breast cancer. Breast cancer research treatment. 2016;156:237–47.

14. Su S, Liao J, Liu J, Huang D, He C, Chen F, Yang L, Wu W, Chen J, Lin L, Zeng Y, Ouyang N, Cui X, Yao H, Su F, Huang J-d, Lieberman J, Liu Q, Song E. Blocking the recruitment of naive CD4 + T cells reverses immunosuppression in breast cancer. Cell Res. 2017;27:461–82.

15. Ma X-J, Dahiya S, Richardson E, Erlander M, Sgroi DC. Gene expression profiling of the tumor microenvironment during breast cancer progression. Breast Cancer Res. 2009;11:R7.

16. Wang J, Rousseaux S, Khochbin S. Sustaining cancer through addictive ectopic gene activation. Current opinion in oncology. 2014;26:73–7.

17. Masjedi S, Zwiebel LJ, Giorgio TD. Olfactory receptor gene abundance in invasive breast carcinoma. Scientific reports. 2019;9:1–12.

18. Asano Y, Kashiwagi S, Goto W, Kurata K, Noda S, Takashima T, Onoda N, Tanaka S, Ohsawa M, Hirakawa K. Tumour-infiltrating CD8 to FOXP3 lymphocyte ratio in predicting treatment responses to neoadjuvant chemotherapy of aggressive breast cancer. BJS (British Journal of Surgery). 2016;103:845–54.

19. Pedroza-Gonzalez A, Xu K, Wu T-C, Aspord C, Tindle S, Marches F, Gallegos M, Burton EC, Savino D, Hori T. Thymic stromal lymphopoietin fosters human breast tumor growth by promoting type 2 inflammation. J Exp Med. 2011;208:479–90.

20. Disis ML, Stanton SE. Triple-negative breast cancer: immune modulation as the new treatment paradigm. American Society of Clinical Oncology Educational Book. 2015;35:e25–30.

21. Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-Garcia W, Treviño V, Shen H, Laird PW, Levine DA. Inferring tumour purity and stromal and immune cell admixture from expression data. Nature communications. 2013;4:2612.

22. Smyth GK, Ritchie M, Thorne N, Wettenhall J, LIMMA: linear models for microarray data. In Bioinformatics and Computational Biology Solutions Using R and Bioconductor. Statistics for Biology and Health, (2005).

23. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. Omics: a journal of integrative biology. 2012;16:284–7.

24. 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. Nature methods. 2015;12:453.

25. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP. STRING v10: protein–protein interaction networks, integrated over the tree of life. Nucleic acids research. 2014;43:D447–52.

26. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome research. 2003;13:2498–504.

27. Dweep H, Gretz N. miRWalk2. 0: a comprehensive atlas of microRNA-target interactions. Nature methods. 2015;12:697.

28. Paraskevopoulou MD, Vlachos IS, Karagkouni D, Georgakilas G, Kanellos I, Vergoulis T, Zagganas K, Tsanakas P, Floros E, Dalamagas T. DIANA-LncBase v2: indexing microRNA targets on non-coding transcripts. Nucleic acids research. 2015;44:D231–8.

29. Davis AP, Grondin CJ, Johnson RJ, Sciaky D, McMorran R, Wiegers J, Wiegers TC, Mattingly CJ. The comparative toxicogenomics database: update 2019. Nucleic acids research. 2018;47:D948–54.

30. Galon J, Pagès F, Marincola FM, Thurin M, Trinchieri G, Fox BA, Gajewski TF, Ascierto PA. The immune score as a new possible approach for the classification of cancer, in, BioMed Central, 2012.

31. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic acids research. 2018;47:D607–13.

32. Davis AP, Murphy CG, Rosenstein MC, Wiegers TC, Mattingly CJ. The Comparative Toxicogenomics Database facilitates identification and understanding of chemical-gene-disease associations: arsenic as a case study. BMC Med Genom. 2008;1:48.

33. Patel DS, Khandeparkar SGS, Joshi AR, Kulkarni MM, Dhande B, Lengare P, Phegade LA, Narkhede K. Immunohistochemical study of MUC1, MUC2 and MUC5AC expression in primary breast carcinoma. Journal of Clinical Diagnostic Research: JCDR. 2017;11:EC30.

34. Matsukita S, Nomoto M, Kitajima S, Tanaka S, Goto M, Irimura T, Kim YS, Sato E, Yonezawa S. Expression of mucins (MUC1, MUC2, MUC5AC and MUC6) in mucinous carcinoma of the breast: comparison with invasive ductal carcinoma. Histopathology. 2003;42:26–36.

35. Astashchanka A, Shroka TM, Jacobsen BM. Mucin 2 (MUC2) modulates the aggressiveness of breast cancer. Breast cancer research treatment. 2019;173:289–99.

36. Engkvist ME, Stratford EW, Lorenz S, Meza-Zepeda LA, Myklebost O, Munthe E. Analysis of the miR-34 family functions in breast cancer reveals annotation error of miR-34b. Sci Rep. 2017;7:9655.

37. Chang T-C, Wentzel EA, Kent OA, Ramachandran K, Mullendore M, Lee KH, Feldmann G, Yamakuchi M, Ferlito M, Lowenstein CJ. Transactivation of miR-34a by p53 broadly influences gene expression and promotes apoptosis. Molecular cell. 2007;26:745–52.

38. Adams BD, Wali VB, Cheng CJ, Inukai S, Booth CJ, Agarwal S, Rimm DL, Győrffy B, Santarpia L, Pusztai L, Saltzman WM, Slack FJ. miR-34a Silences c-SRC to Attenuate Tumor Growth in Triple-Negative Breast Cancer. Can Res. 2016;76:927.

39. Xiao X, Huang X, Ye F, Chen B, Song C, Wen J, Zhang Z, Zheng G, Tang H, Xie X. The miR-34a-LDHA axis regulates glucose metabolism and tumor growth in breast cancer. Sci Rep. 2016;6:21735.

40. Xiang Y, Li F, Wang L, Zheng A, Zuo J, Li M, Wang Y, Xu Y, Chen C, Chen S. Decreased calpain 6 expression is associated with tumorigenesis and poor prognosis in HNSCC. Oncology letters. 2017;13:2237–43.

41. Storr SJ, Thompson N, Pu X, Zhang Y, Martin SG, Calpain in Breast Cancer: Role in Disease Progression and Treatment Response, Pathobiology, 82 (2015) 133–141.

42. MacLeod JA, Gao Y, Hall C, Muller WJ, Gujral TS, Greer PA, Genetic disruption of calpain-1 and calpain-2 attenuates tumorigenesis in mouse models of HER2 + breast cancer and sensitizes cancer cells to doxorubicin and lapatinib, Oncotarget, 9 (2018) 33382–33395.

43. Andrique C, Morardet L, Linares LK, Cissé MY, Merle C, Chibon F, Provot S, Haÿ E, Ea H-K, Cohen-Solal M, Modrowski D. Calpain-6 controls the fate of sarcoma stem cells by promoting autophagy and preventing senescence. JCI Insight. 2018;3:e121225.

44. Yu L, Tang H, Jiang XH, Tsang LL, Chung YW, Chan HC. Involvement of calpain-I and microRNA34 in kanamycin-induced apoptosis of inner ear cells. Cell Biol Int. 2010;34:1219–25.

45. Chen S-M, Wang B-Y, Lee C-H, Lee H-T, Li J-J, Hong G-C, Hung Y-C, Chien P-J, Chang C-Y, Hsu L-S, Chang W-W. Hinokitiol up-regulates miR-494-3p to suppress BMI1 expression and inhibits self-renewal of breast cancer stem/progenitor cells. Oncotarget. 2017;8:76057–68.

46. Song L, Liu D, Wang B, He J, Zhang S, Dai Z, Ma X, Wang X. miR-494 suppresses the progression of breast cancer in vitro by targeting CXCR4 through the Wnt/β-catenin signaling pathway. Oncol Rep. 2015;34:525–31.

47. Romano G, Acunzo M, Garofalo M, Di Leva G, Cascione L, Zanca C, Bolon B, Condorelli G, Croce CM, MiR-494 is regulated by ERK1/2 and modulates TRAIL-induced apoptosis in non–small-cell lung cancer through BIM down-regulation, Proceedings of the National Academy of Sciences, 109 (2012) 16570–16575.

48. Sun HB, Chen X, Ji H, Wu T, Lu HW, Zhang Y, Li H, Li YM. miR–494 is an independent prognostic factor and promotes cell migration and invasion in colorectal cancer by directly targeting PTEN. Int J Oncol. 2014;45:2486–94.

49. Zhan M-N, Yu X-T, Tang J, Zhou C-X, Wang C-L, Yin Q-Q, Gong X-F, He M, He J-R, Chen G-Q, Zhao Q. MicroRNA-494 inhibits breast cancer progression by directly targeting PAK1. Cell Death Dis. 2018;8:e2529–9.

50. Macedo T, Silva–Oliveira RJ, Silva VA, Vidal DO, Evangelista AF, Marques M. Overexpression of mir-183 and mir-494 promotes proliferation and migration in human breast cancer cell lines. Oncology letters. 2017;14:1054–60.

51. Milosavljevic N, Monet M, Léna I, Brau F, Lacas-Gervais S, Feliciangeli S, Counillon L, Poët M. The intracellular Na+/H + exchanger NHE7 effects a Na+-coupled, but not K+-coupled proton-loading mechanism in endocytosis. Cell reports. 2014;7:689–96.

52. Khajah MA, Almohri I, Mathew PM, Luqmani YA. Extracellular alkaline pH leads to increased metastatic potential of estrogen receptor silenced endocrine resistant breast cancer cells. PLoS One. 2013;8:e76327.

53. Onishi I, Lin PJ, Numata Y, Austin P, Cipollone J, Roberge M, Roskelley CD, Numata M. Organellar (Na+, K+)/H + exchanger NHE7 regulates cell adhesion, invasion and anchorage-independent growth of breast cancer MDA-MB-231 cells, Oncology reports, 27 (2012) 311–317.

# Tables

Table 1. The significantly enriched GO terms and KEGG pathways

| ID | Term | Count | P value | Genes |
|---|---|---|---|---|
| GO:0034220 | ion transmembrane transport | 7 | 0.0061074 | GRIK1, AQP8, ATP1A3, ATP6V0A4, ATP12A, FXYD6, FXYD7 |
| GO:0006885 | regulation of pH | 3 | 0.0067471 | SLC9A7, ATP6V0A4, ATP12A |
| GO:0015991 | ATP hydrolysis coupled proton transport | 3 | 0.0257158 | ATP1A3, ATP6V0A4, ATP12A |
| GO:0008625 | extrinsic apoptotic signaling pathway via death domain receptors | 3 | 0.0353638 | TNFRSF10C, DEDD2, CD27 |
| GO:0006355 | regulation of transcription, DNA-templated | 19 | 0.044434 | ZNF44, ZBTB21, ZBTB8B, ZNF132, ZNF284, ZKSCAN1, VENTX, ZNF34, CITED1, YBX2, ZNF439, MEOX2, POU5F1, PERM1, ZNF850, DMRTC1, MTERF3, ZNF319, ATOH7 |
| hsa05146 | Amoebiasis | 4 | 0.033452 | C8A, GNAL, MUC2, GNA15 |

ID: The ID of the enriched GO annotation or KEGG pathway terms; Term: the name of the enriched GO annotation or KEGG pathway; Count: the number of genes enriched in GO or KEGG term; P value: the significance of the enriched terms; Genes: the genes enriched in GO or KEGG term.

Table 2. The genes significantly associated with survival of BRCA patients

| Genes | P value | High.median | Low.median |
|-------|---------|-------------|------------|
| (a) CD4+ T cells related genes | | | |
| MRPS21 | 0.003368158 | 148.5333333 | 113.6333333 |
| GNG8 | 0.011001026 | 131.3666667 | 131.9666667 |
| METTL5 | 0.011379507 | 215.2 | 115.7333333 |
| FAM129C | 0.013358227 | 122.3 | 148.5333333 |
| RNASE12 | 0.013759917 | 122.3 | 142.2333333 |
| NCOA4 | 0.021167879 | 142.2333333 | 131.5 |
| MSTN | 0.032129699 | 131.9666667 | 115.4 |
| MUC2 | 0.03310499 | 131.5 | 131.3666667 |
| DNAI2 | 0.036270816 | 131.5 | 122.3 |
| KRTAP7-1 | 0.038981646 | 142.2333333 | 131.5 |
| (b) CD8+ T cells related genes | | | |
| NMNAT3 | 0.001104224 | 148.5333333 | 115.4 |
| MRPS21 | 0.003368158 | 148.5333333 | 113.6333333 |
| GNAL | 0.00887105 | 148.5333333 | 130.8666667 |
| GNG8 | 0.011001026 | 131.3666667 | 131.9666667 |
| FAM129C | 0.013358227 | 122.3 | 148.5333333 |
| RNASE12 | 0.013759917 | 122.3 | 142.2333333 |
| NCOA4 | 0.021167879 | 142.2333333 | 131.5 |
| ZBED4 | 0.028738163 | 131.3666667 | 131.5 |
| MUC2 | 0.03310499 | 131.5 | 131.3666667 |
| DNAI2 | 0.036270816 | 131.5 | 122.3 |
| KRTAP7-1 | 0.038981646 | 142.2333333 | 131.5 |
| RDH12 | 0.042798068 | 215.2 | 129.1 |

(a) Prognosis associated genes from the CD4+ T cells related genes; (b) Prognosis associated genes from the CD8+ T cells related genes; The genes marked in red represent the overlapped genes from both CD4+ T cells related genes and CD8+ T cells related genes.

# Figures
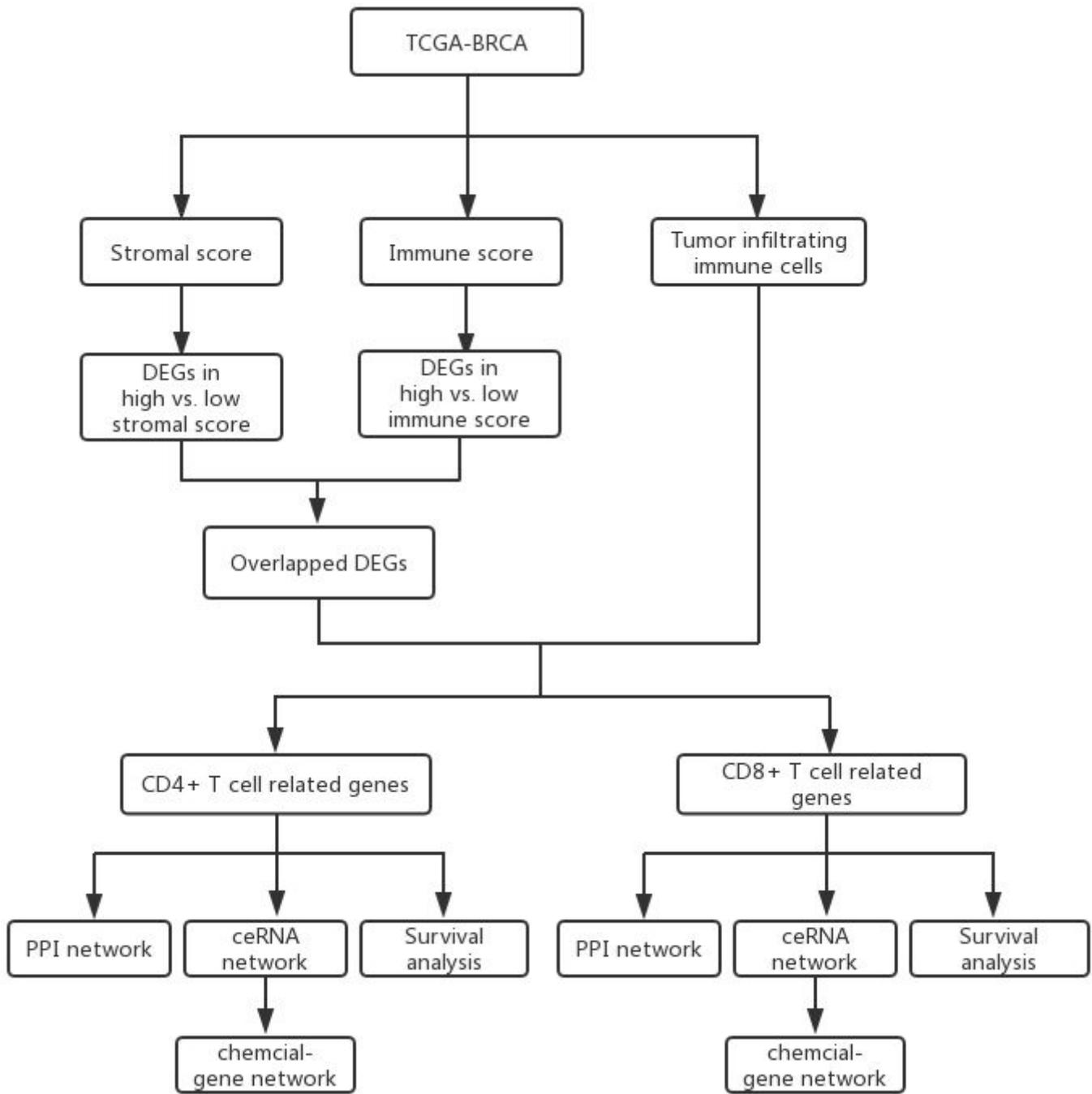


## Figure 1

The workflow of the study.
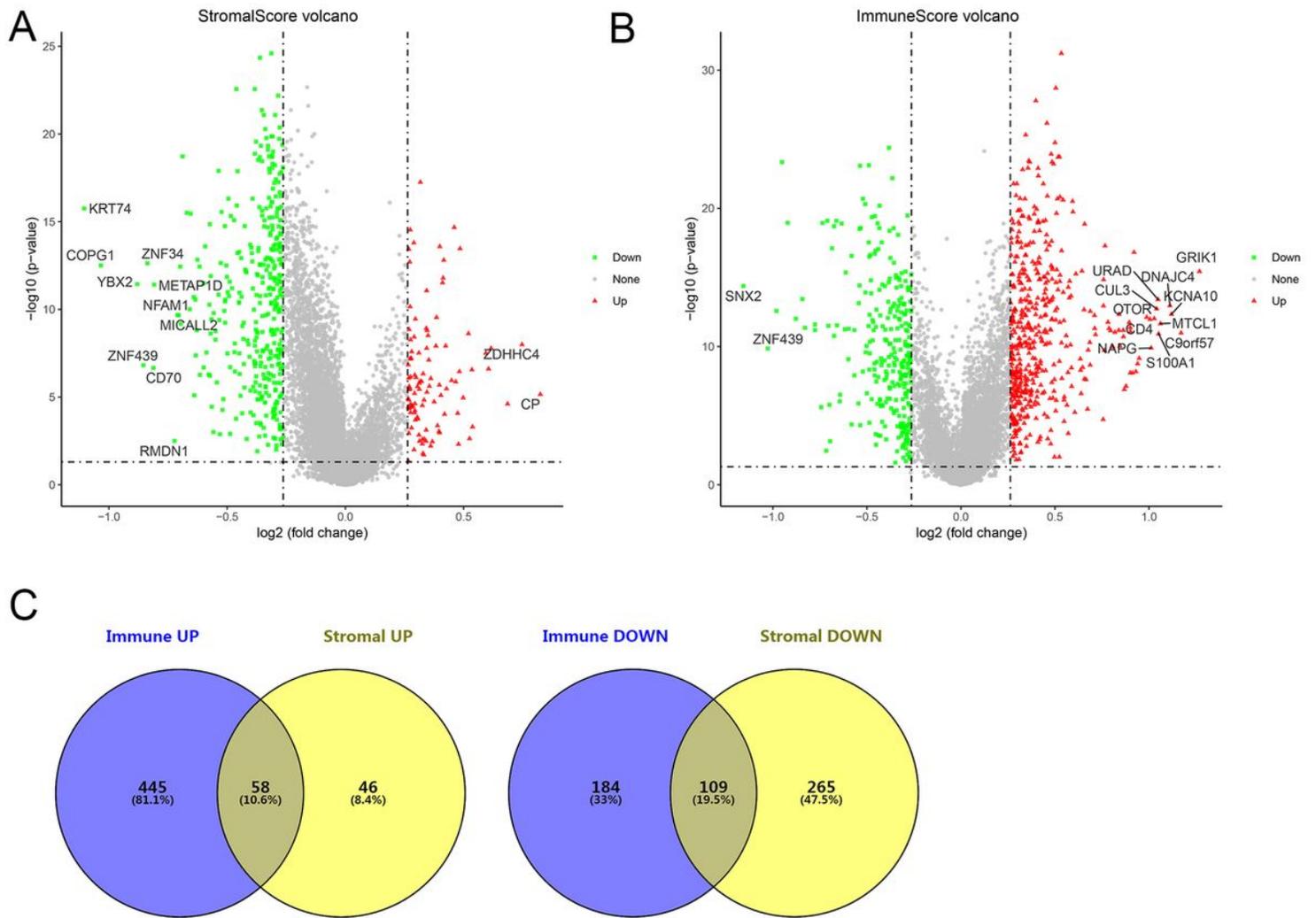
**Figure 2**

Expression profiles of genes in BRCA. Volcano plot showing the expression profile of genes with a high vs. low stromal score (A) and with a high vs. low immune score (B). Red and green dots in the volcano plot represent upregulated and downregulated genes, respectively. VENN diagram (C) showing the overlapping genes in the stromal and immune score-related genes.
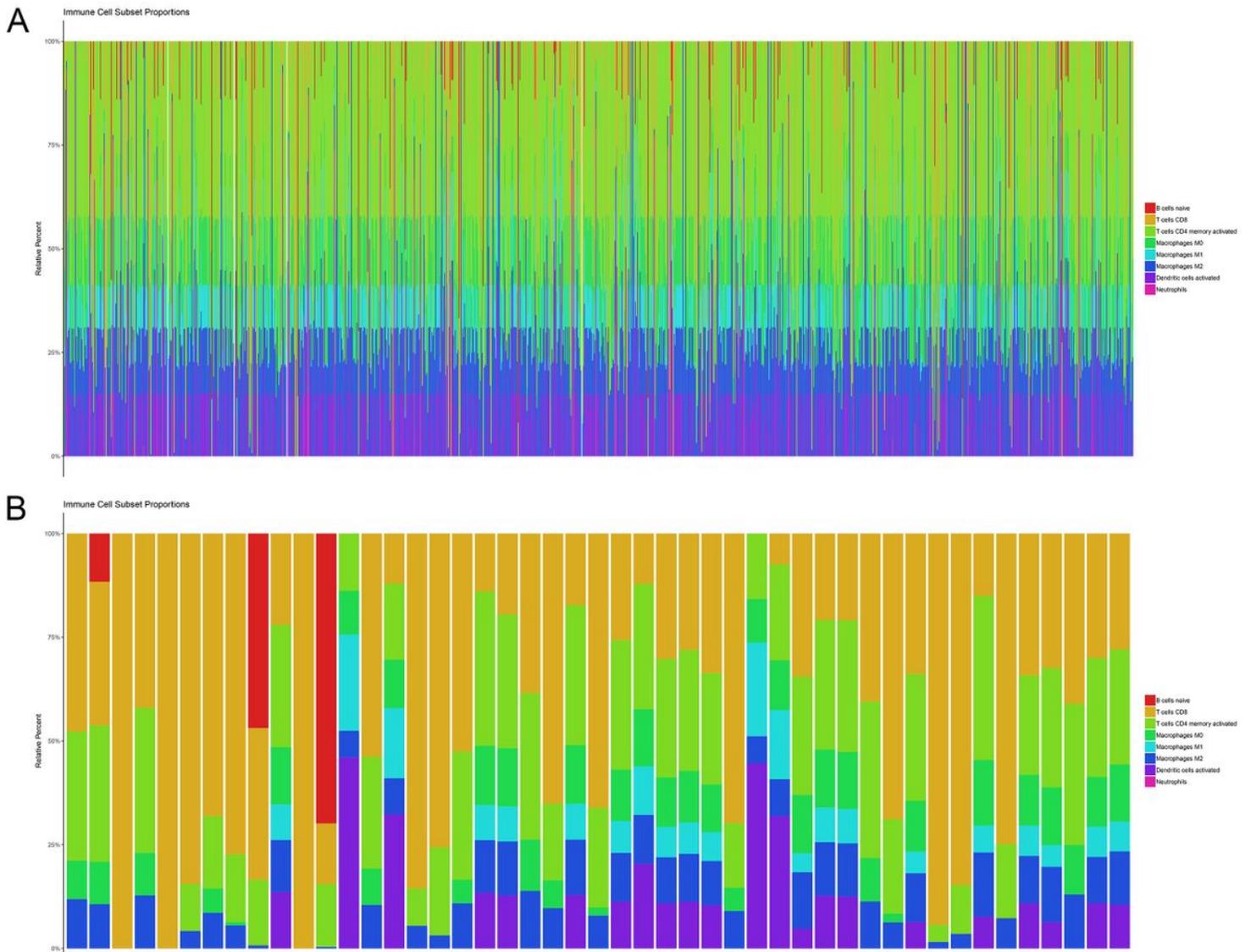
**Figure 3**

The landscape of immune infiltration in BRCA. Bar charts of immune cell subset proportions showing the infiltration abundance of different immune cells. (A) Bar charts of immune cell subset proportions of all BRCA samples, and (B) bar charts of immune cell subset proportions of BRCA samples with P < 0.3. The Y-axis represents the relative percent of immune cell subset infiltration; X-axis represents BRCA samples.
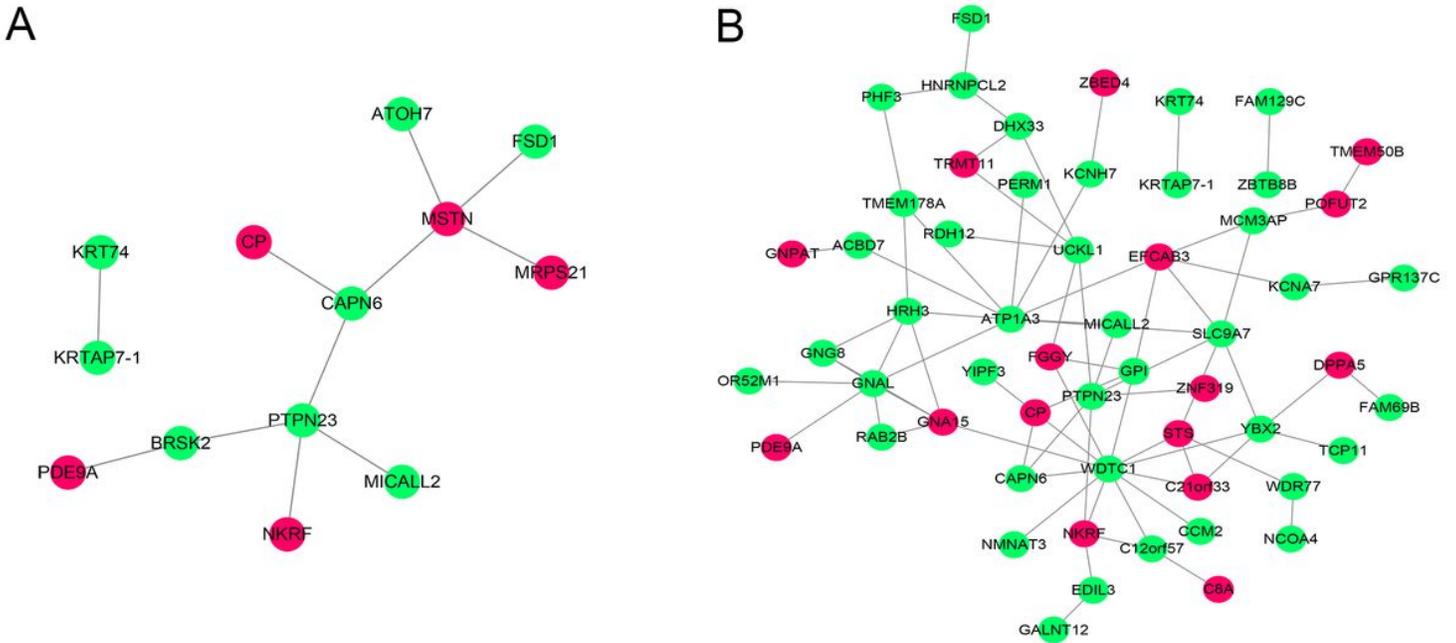
**Figure 4**

Protein-protein interaction networks Protein-protein interaction network of CD4+ T cell-related genes (A) and CD8+ T cell-related genes (B). Red and green nodes represent upregulated and downregulated genes, respectively.
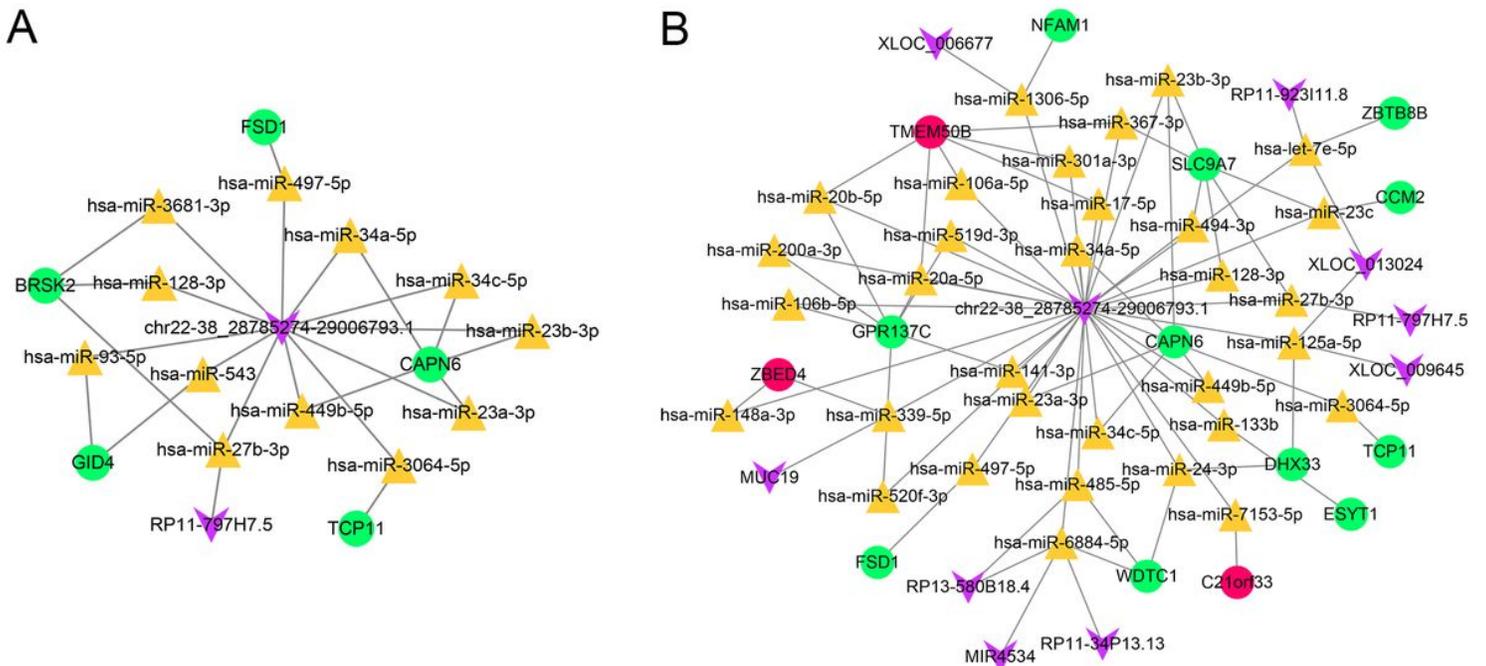


**Figure 5**

Competing endogenous RNAs (ceRNA) networks The ceRNA network of CD4+ T cell-related genes (A) and CD8+ T cell-related genes (B). Red and green nodes represent upregulated genes and downregulated genes, respectively. Yellow triangles represent microRNAs; purple nodes represent long non-coding RNAs.
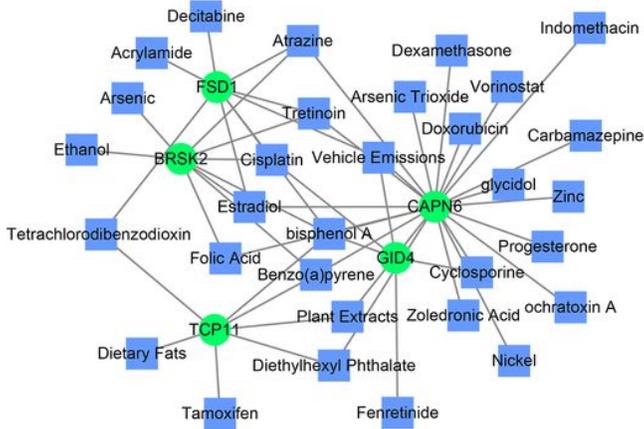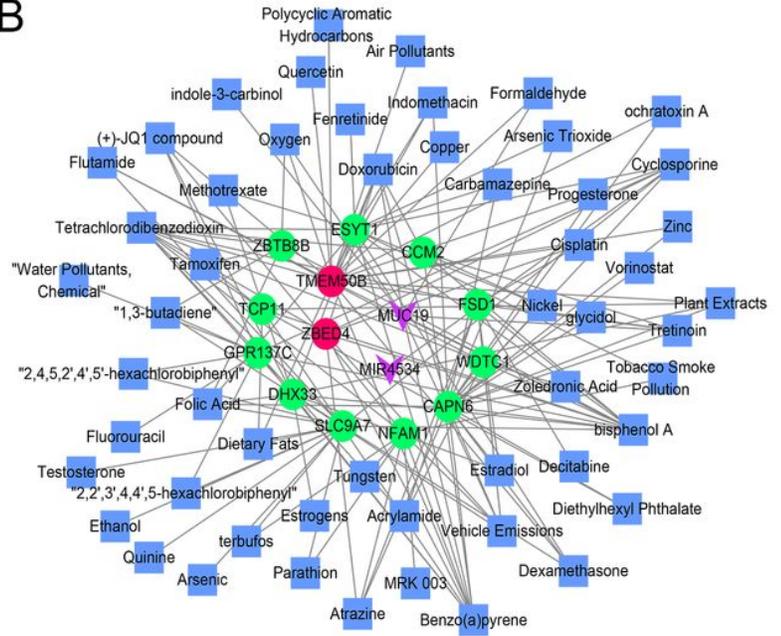
**Figure 6**

Chemical−gene interaction networks Chemical−gene interactions are predicted for the genes in the competing endogenous RNAs network based on the CTD. Chemical−gene interaction network of genes in the CD4+ T cell-related ceRNA network (A) and the CD8+ T cell-related ceRNA network (B). Red and green nodes represent upregulated and downregulated genes, respectively. Blue squares represent chemicals.
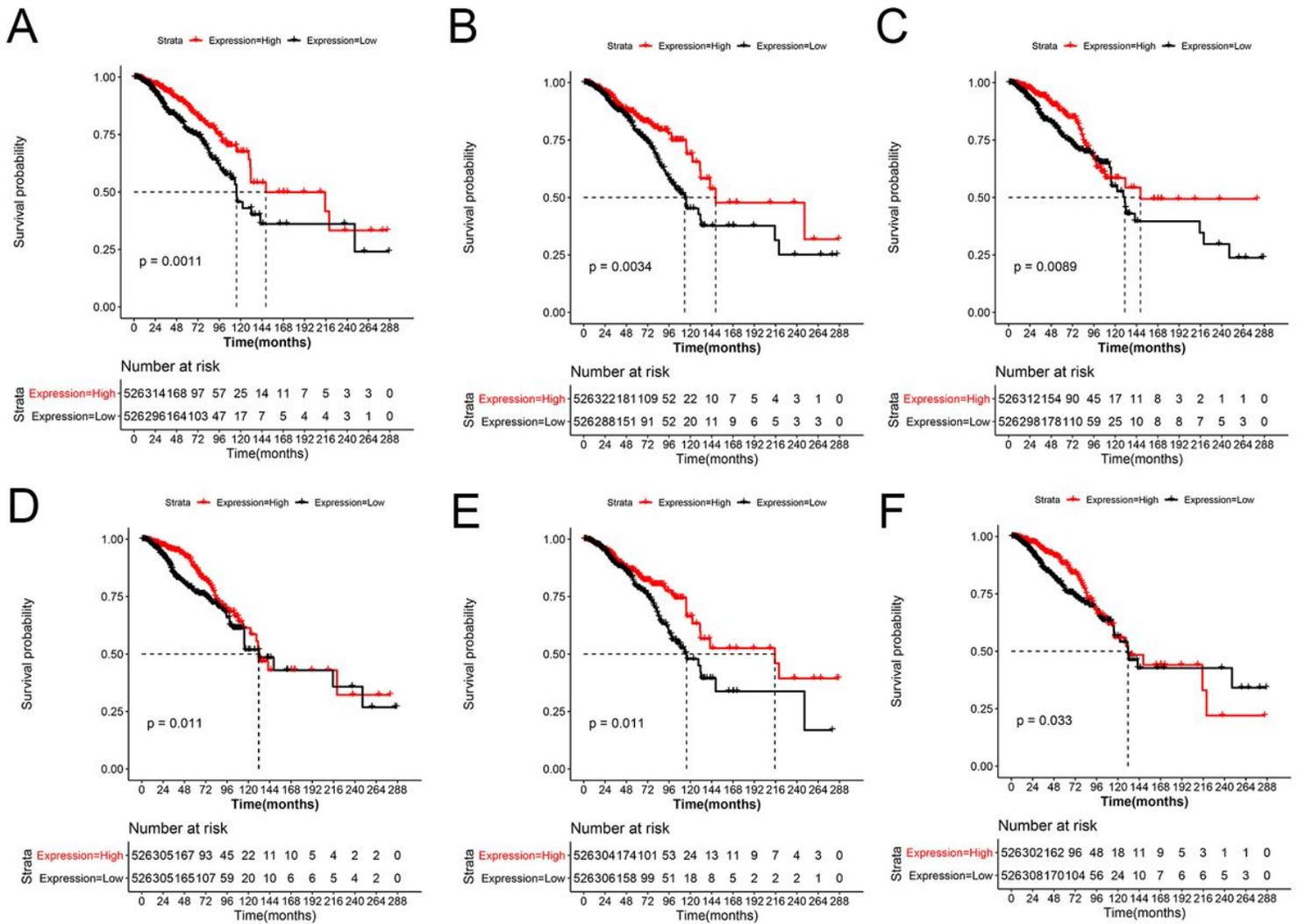
**Figure 7**

Kaplan-Meier curves of overall survival in patients with BRCA Kaplan-Meier curves of overall survival showing the prognosis value of NMNAT3 (A), MRPS21 (B), GNAL (C), GNG8 (D), METTL5 (E), and MUC2 (F).

# Supplementary Files

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

- SupplementalTable4.xls
- SupplementalTable3.xls
- SupplementalTable2.xls
- SupplementalTable1.xls
- Supplementalfiles.doc
- SupplementalFigure1.tif

- OriginalityReport.pdf