

CD52 is a prognostic biomarker and correlated with immune features in breast cancer

Jianxin Wang

First Affiliated Hospital of Harbin Medical University

Yang Liu

Tumor Hospital of Harbin Medical University

Zhuowen Yang

First Affiliated Hospital of Harbin Medical University

Yang Sui

First Affiliated Hospital of Harbin Medical University

Jie Tian

First Affiliated Hospital of Harbin Medical University

Lin Tao

First Affiliated Hospital of Harbin Medical University

Jianting Yao

First Affiliated Hospital of Harbin Medical University

Changjun Wu (✉ cjw1964@yeah.net)

First Affiliated Hospital of Harbin Medical University

Research article

Keywords: CD52, breast cancer, immune features, prognosis, biomarker

Posted Date: June 2nd, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-31586/v1>

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

[Read Full License](#)

Abstract

Background: Breast cancer (BRCA) is the most commonly diagnosed cancer of women, which is aggressive cancer and has a mortality rate. CD52 and its monoclonal antibody (Alemtuzumab) play a critical role in inflammatory diseases, but the relationship between CD52 and BRCA is not clear.

Methods: We first used the random forest algorithm to find the most critical genes related to the prognosis of BRCA patients. Then, according to the analysis of RNA sequence and clinical data of the TCGA dataset, we explored the relationship between CD52 with immune response-related pathways and immune metagenes. The pan-cancer analysis shows the importance of CD52 in a variety of tumors

Results: CD52 was related to the prognosis of BRCA patients ($p < 0.001$). Subsequent analysis based on RNA-seq and clinical data from the TCGA dataset revealed that CD52 is positively correlated with immune response-related pathways and immune metagenes. TIMER analysis showed that CD52 expression was positively correlated with immune infiltrating levels of B, CD4+ T, and CD8+ T cells, macrophages, neutrophils, and dendritic cells (DCs) in BRCA ($r = 0.466$, $r = 0.645$, $r = 0.483$, $r = 0.149$, $r = 0.542$, $r = 0.665$, respectively; $p < 0.001$). CpG sites (cg16068833, cg19743891, cg19743891, cg16664472, cg19677267, cg22517705, and cg27430637) were negatively correlated with CD52 expression ($r = -0.662$, $r = -0.629$, $r = -0.598$, $r = -0.519$, $r = -0.492$, $r = -0.445$, respectively; $p < 0.001$). Furthermore, the expression of CD52 was significantly correlated with the following pathological stages (T stage, N stage, and survival state; $p = 0.024$, $p = 0.047$, and $p = 0.007$, respectively). The results of the pan-cancers study suggest that CD52 may play an important role in the occurrence, development, and prognosis of multiple tumors.

Conclusions: These findings suggested that CD52 is a promising immunotherapy target and prognostic prediction value for BRCA.

Background

Breast cancer (BRCA) is the most commonly diagnosed cancer of women, which is second-deadliest cancer after lung cancer. The overall death rate of breast cancer increased by 0.4% per year in 2 decades since 1975. Even though up to 2017, the total fatalities have declined rapidly by 40%. About 13% of women are likely to be diagnosed with aggressive breast cancer during their lifetime, according to the estimate of the American Cancer Society in 2019[1]. Invasive breast cancer accounts for a significant part [2, 3]. Early-stage (stage I and II) perform favorable prognosis with a 5-year survival rate of 98% and 92% benefit by the popularization of mammography and the progression of targeted-therapy.

Nevertheless, the prognosis of breast cancer and the five-year survival rate reveal significant disparity on account of the variety of scales, districts, ages, clinical stages, molecular phenotypes, and local immune infiltration. As a result, poor prognosis is not rare [1, 4]. Thus, exploring more acute and useful biomarkers as a predictor is still around the corner.

Tumor-associated macrophages (TAMs) are the most critical components of tumor-infiltrating immune cells in the tumor microenvironment [5]. There are two principal functional states in TAMs: pro-inflammatory M1 macrophages which are indicated as protective factors for obliterating tumor cells, and alternatively activated M2 macrophages (which are considered as unfavorable factors for prompting tumor proliferation)[3, 6, 7]. Previous research has established that macrophages can decrease the expression of estrogen and progesterone receptors, whereas increasing the expression of urokinase-type plasminogen activator receptor and Ki67 in breast cancer. In addition their results demonstrated a significant positive association of TAMs and poor prognosis in breast cancer patients [8].

In this study, we first performed univariate Cox proportional analysis for selecting prognostic macrophage-related gene signatures. Then, the random forest was recognized to build a 13-gene signature for BRCA, and the variable importance suggested that CD52 is the most critical for further analysis. CD52 (CAMPATH-1 antigen) is a glycosylphosphatidylinositol (GPI) -anchored protein of 12 amino acids present on the cell surface of immune cells, including monocytes/macrophages[9, 10]. Piccaluga PP *et al*/ found that CD52 was up-regulated in peripheral T-cell lymphoma, and the estimation of CD52 expression might provide a theoretical basis for the efficacy of treatment response [11]. Moreover, Alemtuzumab, an anti-CD52 monoclonal antibody, has been investigated as a molecular target for immunotherapy to treat acute myeloid leukemia [12]. Nevertheless, the relationship between CD52 expression, prognostic value, and immune infiltration in BRCA is not clear. Therefore, we analyzed the clinical and molecular data of CD52 in BRCA samples from the TCGA dataset to explore the expression of CD52 and its relationship with immune-related molecules. It may provide a possible basis for the use of Alemtuzumab in the treatment of BRCA patients.

Methods

Data source and downloaded

We downloaded available RNA-sequence and clinical data of invasive breast cancer patients from the TCGA database (<https://portal.gdc.cancer.gov>). The RNA-seq results were combined into the gene expression matrix. We obtained all methylation information from patients with BRCA and normal tissue controls from the UCSC Xena browser (<https://xenabrowser.net/>).

Extraction of macrophage-related gene matrix and selection process of the target gene

We extracted macrophage-related gene expression patterns according to the gene signatures of M1 and M2 macrophages in the literature [13]. We performed a univariate Cox proportional hazard regression to identify the differentially expressed hypoxia-related genes associated with overall survival time ($P < 0.05$ was considered statistically significant). Then we use a random forest to establish a prognosis model. The most crucial gene was selected as the target gene for further analysis according to the importance of variables. The Kaplan-Meier (KM) method was used to evaluate survival differences. The receiver operating characteristic (ROC) curve identifies the accuracy of the model prediction. We used the STRING

database (<https://string-db.org/>) version 11.0 to assess the protein-protein interaction network information of the target gene.

Relationship between CD52 expression and clinical symptoms

The "ggstatsplot" package validated the relationship between and expression of CD52 in the TCGA database and six clinical symptoms (age, survival state, stage, T stage, M stage, N stage).

GSEA-based enriched KEGG analysis

To detect significant differences differentially activated in BRCA, we performed GSEA (Gene Set Enrichment Analysis)-based enriched KEGG (Kyoto Encyclopedia of Genes and Genomes) analysis between low and high CD52 expression phenotype using GSEA software. The enrichment score (ES) >0.4 as a filter and false discovery rate (FDR) value <0.05 was considered to be statistically significant.

ssGSEA analysis revealed the immune features of CD52 in BRCA

We downloaded 16 signatures from the literature [14], including immune-relevant signature, stromal-relevant signature, and mismatch-relevant signature. The list of these genes is shown in **Supplementary Table 1**. We performed ssGSEA (Single-Sample GSEA) analysis to determine the enrichment scores of immune features using the GSVA package of R language [15]. We calculated Pearson correlation values between CD52 expression and immune features based on correlation analysis.

Association between CD52 expression and immune metagenes

We downloaded seven immune metagenes from the literature [16], including IgG, hemopoietic cell kinase [HCK], MHC-I (major histocompatibility complex-I), MHC-II (major histocompatibility complex-II), LCK (lymphocyte-specific kinase), STAT1 (signal transducer and activator of transcription 1), and Interferon. The list of these genes is shown in **Supplementary Table 2**. We calculated Pearson correlation values between CD52 expression and immune features based on correlation analysis.

Correlation of CD52 expression and immune infiltration level in BRCA

Tumor Immune Estimation Resource (TIMER, <https://cistrome.shinyapps.io/timer/>) [17, 18] is a comprehensive resource to estimate the abundance of immune infiltrates of various cancer types based on the TCGA database. We analyzed the relationship between CD52 expression and immune infiltration in BRCA via gene modules, explicitly including the following immune cell: B cells, CD4+ T cells, CD8+ T cells, neutrophils, macrophages, and dendritic cells. Besides, we also explored the comparison of tumor infiltration levels among BRCA with different somatic copy number alterations for CD52 by using the SCNA module, including deep deletion, arm-level deletion, diploid/normal, arm-level gain, and high amplification.

Correlation of CD52 expression and methylation

We obtained DNA methylation data from the UCSC Xena browser. We calculated the correlation between the CD52 expression and the methylation of the CpG sites by using Pearson's correlation analysis. 0.4 as a filter value of the correlation coefficient.

Assessment of the expression and prognostic importance of CD52 in pan-cancers

We used the TIMER database to evaluate the expression of CD52 in pan-cancers. Moreover, the present study assessed the prognostic importance of CD52 in pan-cancers, using the TCGA analysis database, the Kaplan-Meier Plotter database (<http://kmplot.com/analysis/>).

Results

Selection of CD52 as a prognostic marker in BRCA

A univariate Cox proportional hazard regression analysis showed that 237 genes associated with overall survival time (**Supplementary figure 3**). The random forest prognosis model identified 13 genes (HEXA, TIE1, TMED5, SNX5, CD52, DLAT, IFNAR1, XPNPEP2, STX4, LONRF3, SLAMF8, EXOC5, CLCN7) (**Fig.1A, B**). We calculated the risk score of each sample and grouped the samples into high-risk and low-risk according to the median risk score (cutoff = 43.255). The prognoses of the high-risk and low-risk groups significantly differed (**Supplementary figure 1, Supplementary figure 2**). The average 12-, 36-, and 60-month Area Under Curve (AUC) for the CD52 was 0.98 (**Fig.1C**). Based on the variable importance in **Fig.1B**, CD52 is the most crucial gene, and the survival analysis results indicated that the prognosis value of CD52 was significant (**Fig.1D**). The PPI of the CD52 protein showed its value (**Fig.1E**). High expression of CD52 was associated with low risk and was a protective factor. These results indicated that CD52 is a prognostic marker for further analysis.

Relationship between expression of CD52 and clinical symptoms

By exploring the association between clinical symptoms and expression of CD52 in the TCGA database, we found that there was a very significant correlation between CD52 expression and the T stage, N stage, and survival state ($p = 0.024$, $p = 0.047$ and $p = 0.007$) (**Fig.2A-C**). Age, N stage, and stage were not significantly correlated with the CD52 expression (**Supplementary figure 3**).

GSEA analysis of CD52-related pathway

GSEA analysis results showed that B cell receptor, chemokine, NOD-like receptor, Toll-like receptor, and T cell receptor signaling pathways were significantly enriched in the CD52 high expression group, all of which are strictly related to tumor immunity. In contrast, Glycosylphosphatidylinositol (GPI)-anchor biosynthesis and metabolism-related pathways were significantly enriched in CD52 low expression samples (**Fig.2D**).

CD52 was closely related to different signatures

The correlation analysis result demonstrated that CD52 expression was positively correlated with the activation of immune-relevant signature (**Fig.3A**), especially CD8 T effector (cor=0.758), Immune checkpoint(cor=0.758), and Antigen processing machinery signature(cor=0.483) (**Fig.3B-D**).

CD52 was closely related to immune metagenes

The relevant analysis result had indicated that CD52 expression was positively correlated with the activation of immune metagenes (**Fig.4A**), especially LCK (cor=0.838), MHC-I (cor=0.688), and HCK (cor=0.677) (**Fig.4B-D**).

CD52 was correlated with immune infiltration level in BRCA

TIMER database was used to evaluate the relationship between CD52 expression and immune infiltration level. CD52 expression was positively correlated with immune infiltrating levels of B cells($r = 0.466, p = 6.29e-78$), CD8+ T cells($r = 0.483, p = 4.59e-58$), CD4+ T cells ($r = 0.645, p = 3.79e-114$), macrophages ($r = 0.149, p = 2.77e-06$), neutrophils ($r = 0.542, p = 1.45e-73$) and dendritic cells ($r = 0.665, p = 1.11e-122$) in BRCA (**Fig.5A**). We also provided the comparison of tumor infiltration levels among BRCA with different somatic copy number alterations for CD52 by using the SCNA module, including deep deletion, arm-level deletion, diploid/normal, arm-level gain, and high amplification. **Fig.5B** showed CD52 was related to infiltration level for each SCNA category (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$).

Correlation of CD52 expression and methylation

Seven methylation sites (cg00813993, cg16068833, cg19743891, cg19743891, cg16664472, cg19677267, cg22517705, and cg27430637) were identified. There are six methylation sites with a correlation coefficient higher than 0.4 and a P value less than 0.05 (**Fig.5A-F**).

Assessment of the expression and prognostic importance of CD52 in pan-cancers

We used TIMER and Kaplan Meier Plotter databases to evaluate the expression and prognostic value of CD52 in pan-cancers. CD52 expression was significantly higher in BRCA (Breast invasive carcinoma), CHOL (Cholangiocarcinoma), ESCA (Esophageal Squamous Cell Carcinoma), HNSC (Head-neck squamous cell carcinoma), KIRC (Kidney renal clear cell carcinoma), KIRP (Kidney renal papillary cell carcinoma). However, CD52 expression was significantly lower in BLCA (Bladder Urothelial Carcinoma), COAD (Colon adenocarcinoma), KICH (Kidney chromophobe), LUAD (Lung adenocarcinoma), LUSC (Lung squamous cell carcinoma), PRAD (Prostate adenocarcinoma), READ (Rectum adenocarcinoma) (Figure 7A). Furthermore, the results show that CD52 has prognostic value in eight kinds of cancers, including BRCA, CESC (Cervical squamous cell carcinoma), ESCA, HNSC, LUAD, SARC (Sarcoma), THYM(Thymoma), and UCEC (Uterine corpus endometrial carcinoma) (Figure 7B-G).

Discussion

Based on the expression of macrophage related genes in TCGA of BRCA patients, univariate Cox proportional analysis and the random forest algorithm were performed to build a prognostic model. The variable importance suggests that CD52 is the most critical gene. Moreover, patients with high CD52 expression have a better prognosis. CpG methylation typically results in abnormal gene expression [19]. In our study, six CpG sites (cg16068833, cg19743891, cg19743891, cg16664472, cg19677267, cg22517705, and cg27430637) were negatively correlated with CD52 expression ($r = -0.662$, $r = -0.629$, $r = -0.598$, $r = -0.519$, $r = -0.492$, $r = -0.445$, respectively; $p < 0.001$). DNA methylation is most common in CpG dinucleotide and is related to the clinicopathological features of BRCA patients, including stage and histological grade [20, 21]. Furthermore, the CD52 expression was significantly correlated with the following pathological stages (T stage, N stage, and survival state). These results suggest that CD52 is important for the prognosis of BRCA patients.

CD52 epitopes were expressed on the surface membrane of peripheral lymphocytes, monocytes and macrophages, and on the epithelial membrane of the male reproductive system [22]. Rashidi M *et al* have demonstrated that CD52 can inhibit the activation of NF - κ B by inhibiting the signal transduction of Toll-like receptor and tumor necrosis factor receptor and thus inhibit the production of inflammatory cytokines by macrophages, monocytes, and dendritic cells [23]. As the GSEA results show, CD52 was significantly enriched in a variety of immune-related pathways, such as B cell receptor, chemokine, NOD-like receptor, Toll-like receptor, and T cell receptor signaling pathways. These enrichment pathways are closely related to the immune infiltration in cancer [24, 25].

Previous studies demonstrated has demonstrated that immune cell infiltration was associated with activation of the immune response, and it will contribute to anti-tumor effects and get a better prognosis in BRCA [26–28]. To clarify the correlation between CD52 expression and immune features based on ssGSEA analysis, we found that CD52 expression is related to T cell and macrophage related pathways and functions. Previous studies have confirmed that immune infiltration is widespread in breast cancer tissues and affects patient prognosis [29, 30]. We further investigated the correlation of CD52 expression with multiple levels of immune infiltration in BRCA. Our results indicated a significant positive correlation between the infiltration levels of CD8 + T cells, CD4 + T cells, B cells, DC cells, and neutrophils in BRCA and the expression of CD52. Thus, a large number of data confirmed that CD52 plays a role in tumor immunology in regulating BRCA.

Drug Alemtuzumab, an anti-CD52 monoclonal antibody, has been used in the treatment of various immune-related diseases, including multiple sclerosis and inflammatory myopathy [31, 32]. We found that CD52 expression is abnormal in breast cancer patients and its role in regulating tumor immunity. Meanwhile, the results of the pan-cancer study suggest that CD52 differentially expressed in multiple tumors, which may play an essential role in the occurrence, development, and prognosis of multiple tumors. These results may indicate the possibility of expanding the application of Alemtuzumab in tumor immunotherapy.

However, the current study was limited by the absence of experimental evidence. Our PPI results suggested the protein-protein interaction of CD52 related-proteins, which may provide the basis for further mechanism study.

Conclusion

These findings will deepen our understanding of CD52 expression, prognosis, and immune-related features in BRCA. CD52 is a promising immunotherapy target for most cancer patients, and the drug (Alemtuzumab) may also bring new hope for immunotherapy of cancer.

Supplementary Information

supplementary Table 1 List of the genes of immune-relevant signature, stromal-relevant signature, and mismatch-relevant signature

supplementary Table 2 List of the genes of seven immune metagenes

supplementary Table 3 List of 237 genes associated with overall survival time

Supplementary figure 1 The distribution of 13-gene prognosis model, survival status of patients, and gene expression signature.

Supplementary figure 2 The Kaplan-Meier survival analysis with log-rank test. AUC, area under the curve.

Supplementary figure 3 The relationship between CD52 and clinical symptoms. Relationship between CD52 and age, M stage and stage.

Abbreviations

BRCA: breast cancer; DCs: dendritic cells; TAMs: Tumor-associated macrophages; GPI: glycosylphosphatidylinositol; KM: Kaplan-Meier; ROC: receiver operating characteristic; GSEA: Gene Set Enrichment Analysis; ES: enrichment score; ssGSEA; Single-Sample GSEA; FDR: false discovery rate; HCK :hemopoietic cell kinase; MHC-I :major histocompatibility complex- I; MHC-II :major histocompatibility complex-II; LCK: lymphocyte-specific kinase; STAT1: signal transducer and activator of transcription 1; TIMER :Tumor Immune Estimation Resource

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

All authors agree to publish.

Availability of data and materials

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

Competing interests

The authors declare that they have no conflict of interest.

Funding

Not applicable.

Authors' contributions

JXW, YL, ZWY and YS wrote the manuscript. YL, JXW, ZWY, JT and YTY analyzed data. YS and JT were responsible for preparing figures. CJW designed the research and revised the manuscript.

Acknowledgements

Not applicable.

References

1. E DC, Jiemin M, M GM, A NL, D MK, Ann GS, Ahmedin J and L SR. Breast cancer statistics, 2019. CA: a cancer journal for clinicians. 2019; 69(6).
2. Helen T, Meagan R and Kara B. The Immune Microenvironment of Breast Cancer Progression. Cancers. 2019; 11(9).
3. Hendry S, Pang JMB, Byrne D, Lakhani SR, Cummings MC, Campbell IG, Mann GB, Goringe KL and Fox SB. Relationship of the breast ductal carcinoma immune microenvironment with clinico-pathological and genetic features. Clinical Cancer Research—An Official Journal of the American Association for Cancer Research. 2017; 23(17):5210-5217.
4. Emens and A L. Breast cancer immunobiology driving immunotherapy: vaccines and immune checkpoint blockade. Expert Rev Anticancer Ther. 2012; 12(12):1597-1611.
5. Dehne N, Mora J, Namgaladze D, Weigert A and Brüne B. Cancer cell and macrophage cross-talk in the tumor microenvironment. Current Opinion in Pharmacology. 2017; 35:12-19.
6. Si-Qi Q, H WSJ, C ZM, E dVEG, Bert vdV and P SC. Tumor-associated macrophages in breast cancer: Innocent bystander or important player? Cancer treatment reviews. 2018; 70:178-179.
7. Mantovani A and Locati M. Tumor-Associated Macrophages as a Paradigm of Macrophage Plasticity, Diversity, and Polarization: Lessons and Open Questions. Arterioscler Thromb Vasc Biol. 2013; 33(7):1478-1483.

8. Lindsten TS, Hedbrant A, Ramberg A, Wijkander J, Solterbeck A, Eriksson M, Delbro D and Erlandsson A. Effect of macrophages on breast cancer cell proliferation, and on expression of hormone receptors, uPAR and HER-2. *International Journal of Oncology*. 2017; 51(1):104-114.
9. Lively and R. M. Primary Structure of CD52. *Journal of Biological Chemistry*. 1995; 270(11):6088-6099.
10. Junfeng D, Xiaodong Z, Huiting S, Xiaofei S and Yong Z. The immunological function of CD52 and its targeting in organ transplantation. *Inflammation Research Official Journal of the European Histamine Research Society*. 2017; 66(7):571-578.
11. Piccaluga PP, Agostinelli C, Righi S, Zinzani PL and Pileri SA. Expression of CD52 in peripheral T-cell lymphoma. *Haematologica*. 2007; 92(4):566-567.
12. Saito Y, Nakahata S, Yamakawa N, Kaneda K and Morishita K. CD52 as a molecular target for immunotherapy to treat acute myeloid leukemia with high EVI1 expression. *Leukemia Official Journal of the Leukemia Society of America Leukemia Research Fund U K*. 2011; 25(6):921-931.
13. Aran D, Hu Z and Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biology*. 2017; 18(1):220.
14. Zeng D, MeiyiLi, Zhou R, Zhang J and Sun H. Tumor Microenvironment Characterization in Gastric Cancer Identifies Prognostic and Immunotherapeutically Relevant Gene Signatures. 2019. 2019; 7(5):737-750.
15. Hänzelmann S, Castelo R and Guinney J. GSVA: Gene set variation analysis for microarray and RNA-Seq data. *Bmc Bioinformatics*. 2013; 14(1):7.
16. Rody A, Holtrich U, Pusztai L and Liedtke... C. T-cell metagene predicts a favorable prognosis in estrogen receptor-negative and HER2-positive breast cancers. *Breast Cancer Research*. 2009; 11(2).
17. Li B, Severson E, Pignon JC, Zhao H, Li T, Novak J, Jiang P, Shen H, Aster JC and Rodig S. Comprehensive analyses of tumor immunity: implications for cancer immunotherapy. *Genome Biology*. 2016; 17(1):174.
18. Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, Li B and Liu XS. TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. *Cancer Research*. 2017; 77(21): e108-e110.
19. Keller S, Sarchiapone M, Zarrilli F, et al. Increased BDNF promoter methylation in the Wernicke area of suicide subjects. *Arch Gen Psychiatry*. 2010;67(3):258–267.
20. Fleischer T, Edvardsen H, Solvang HK, Daviaud C and Tost J. Integrated analysis of high-resolution DNA methylation profiles, gene expression, germline genotypes and clinical end points in breast cancer patients. *International Journal of Cancer*. 2014; 134(11):2615-2625.
21. Fleischer T, Frigessi A, Johnson KC, Edvardsen H, Touleimat N, Klajic J, Riis ML, Haakensen VD, Warnberg F, Naume B, Helland A, Borresen-Dale AL, Tost J, et al. Genome-wide DNA methylation profiles in progression to invasive carcinoma of the breast with impact on gene transcription and prognosis *Genome Biology*. 2014; 15(8):435.
22. Domagała A and Kurpisz M. CD52 antigen—a review. *Med Sci Monit*. 2001; 7(2):325-331.

23. Maryam R, Esther B-S, E LK, Yuxia Z, M NA, L VS, Robert OD, M WJ, E AT, E VJ and C HL. CD52 inhibits Toll-like receptor activation of NF-kappa B and triggers apoptosis to suppress inflammation. *Cell death and differentiation*. 2018; 25(2):392-405.
24. Liu XS and Mardis ER. Applications of Immunogenomics to Cancer. *Cell*. 2017; 168(4):600-612.
25. Yin T, Zhiqian G, Youwei F and Junjie W. CXCR3 from chemokine receptor family correlates with immune infiltration and predicts poor survival in osteosarcoma. . *Bioscience reports*. 2019; 39(11).
26. Xavier T, Tonje L, Hagen RA, Daniel N, Elin B, Oma OH, Amund KJ, Johan V-C, Marie F, Undlien DE, Gregusson SL, Tu SMA, Øystein G, et al. An independent poor-prognosis subtype of breast cancer defined by a distinct tumor immune microenvironment. *Nature communications*. 2019; 10(1).
27. C DA, L HT, Muhammad A, H FM, J WS, A BR, Alvaro P, M CJ, L S-MM, M ML, E ME, A DL, M VC, et al. Alterations in the Immune Cell Composition in Premalignant Breast Tissue that Precede Breast Cancer Development. *Clinical cancer research: an official journal of the American Association for Cancer Research*. 2017; 23(14).
28. Sawe RT, Kerper M, Badve S, Li J, Sandoval-Cooper M, Xie J, Shi Z, Patel K, Chumba D and Ofulla A. Aggressive breast cancer in western Kenya has early onset, high proliferation, and immune cell infiltration. *Bmc Cancer*. 2016; 16(1):204.
29. Desmedt C, Salgado R, Fornili M, Pruneri G and Sotiriou C. Immune Infiltration in Invasive Lobular Breast Cancer. *Journal of the National Cancer Institute*. 2017; 110(4).
30. Karn T, Jiang T, Hatzis C, Sängner N and Pusztai L. Association Between Genomic Metrics and Immune Infiltration in Triple-Negative Breast Cancer. *Jama Oncol*. 2017; 3(12).
31. Li Z, Richards S, Surks HK, Jacobs A, Panzara MA. Clinical pharmacology of alem-tuzumab, an anti-CD52 immunomodulator, in multiple sclerosis. *Clin Exp Immunol*. 2018; 194:295–314.
32. Dalakas MC, Rakocevic G, Schmidt J, et al. Effect of Alemtuzumab (CAMPATH 1-H) in patients with inclusion-body myositis. *Brain* 2009; 132: 1536–1544

Figures

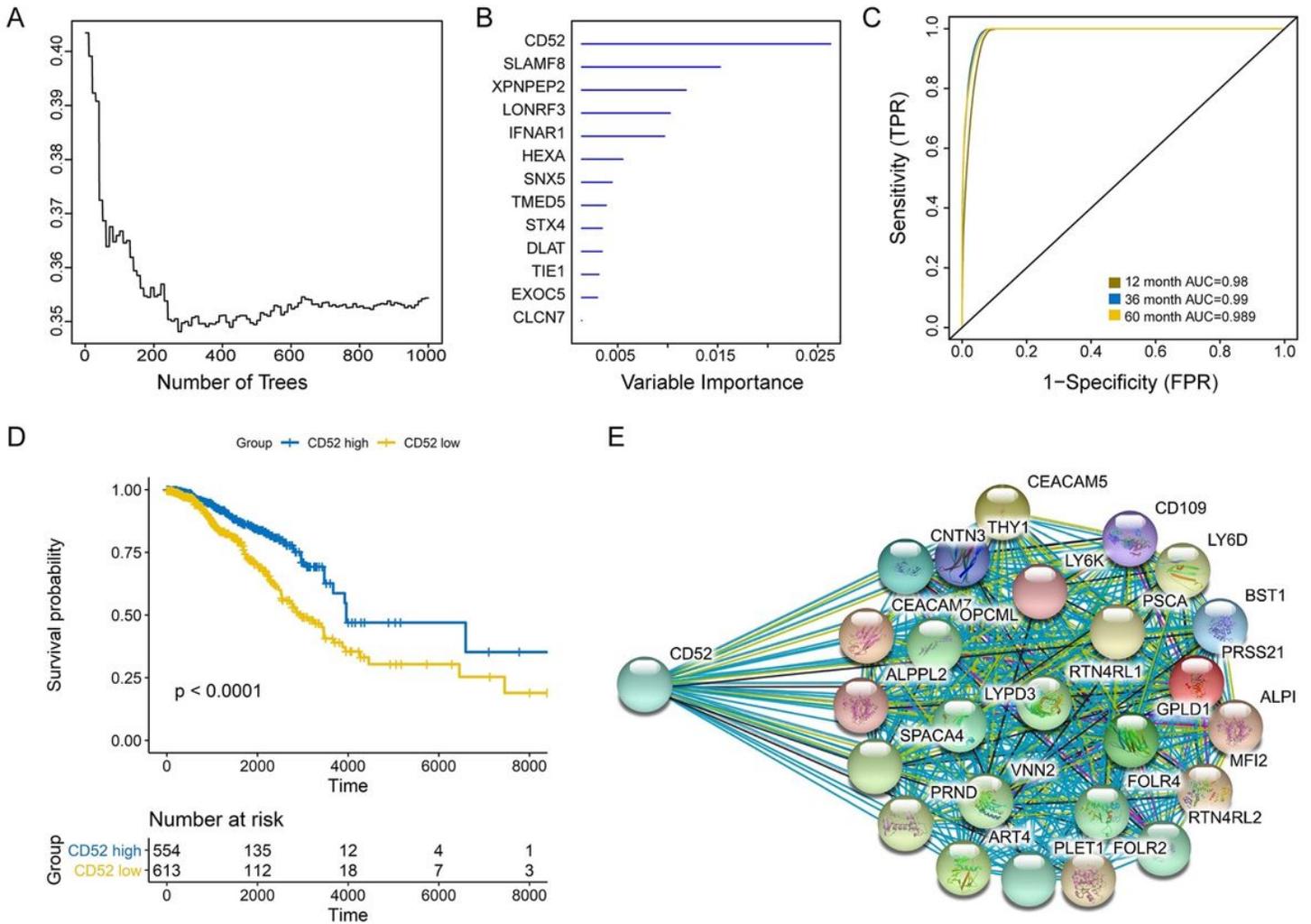


Figure 1

Selection process of CD52. (A) Function tree's error rate. There were 1000 trees (Ntree). (B) The variable importance values of 13 gene signatures. (C) The ROC curve of the prognosis model based on the 13 gene signatures. (D) The Kaplan-Meier survival analysis of CD52. (E) PPI network of CD52. AUC, area under the curve; ROC, receiver operating characteristic; PPI, protein-protein interaction.

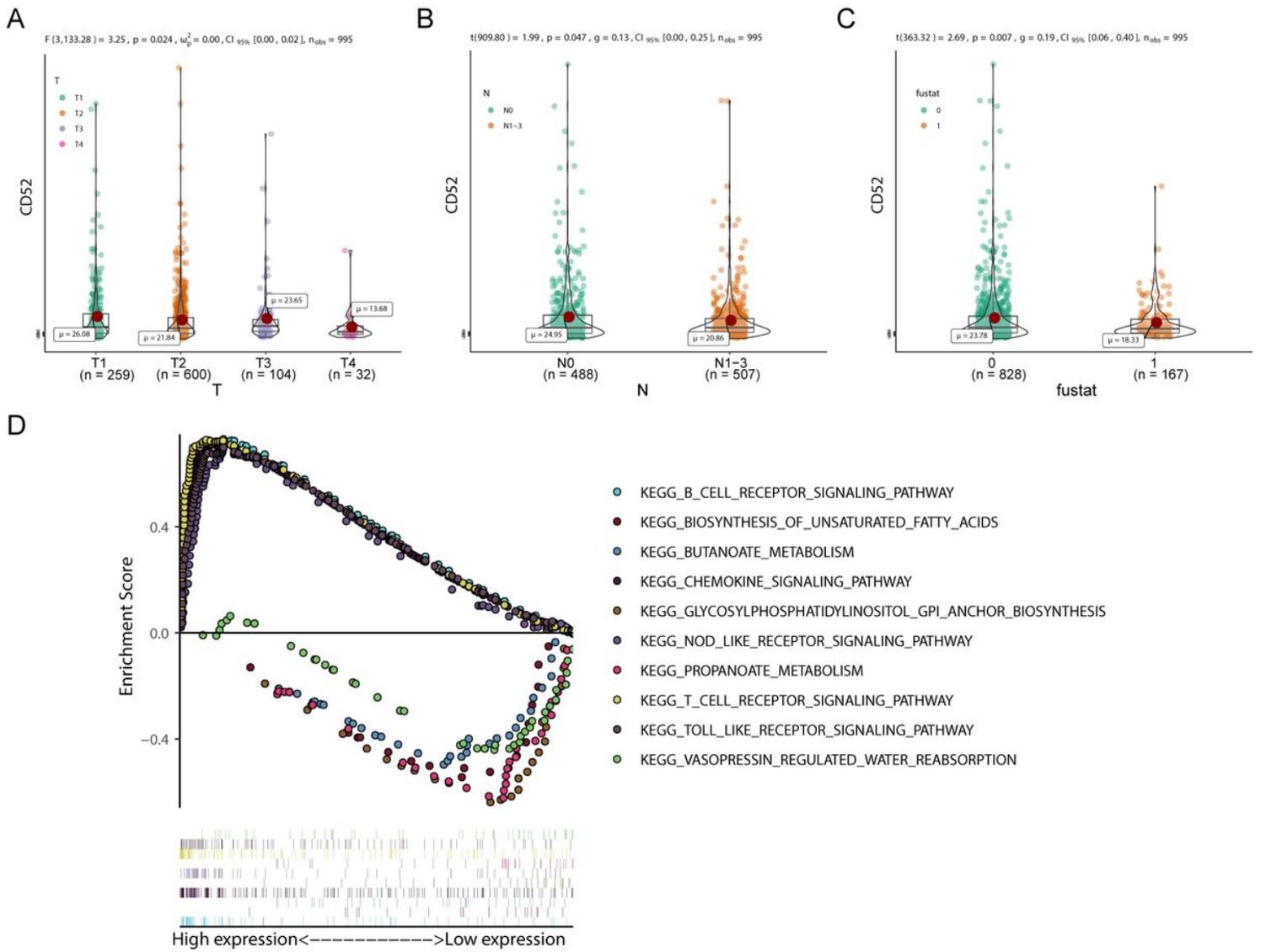


Figure 2

The relationship between CD52 and clinical symptoms and GSEA enrichment analysis. Relationship between CD52 and T stage (A), N stage (B) and fustat (C). KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway enrichment analysis of CD52 by GSEA (D). GSEA, Gene Set Enrichment Analysis.

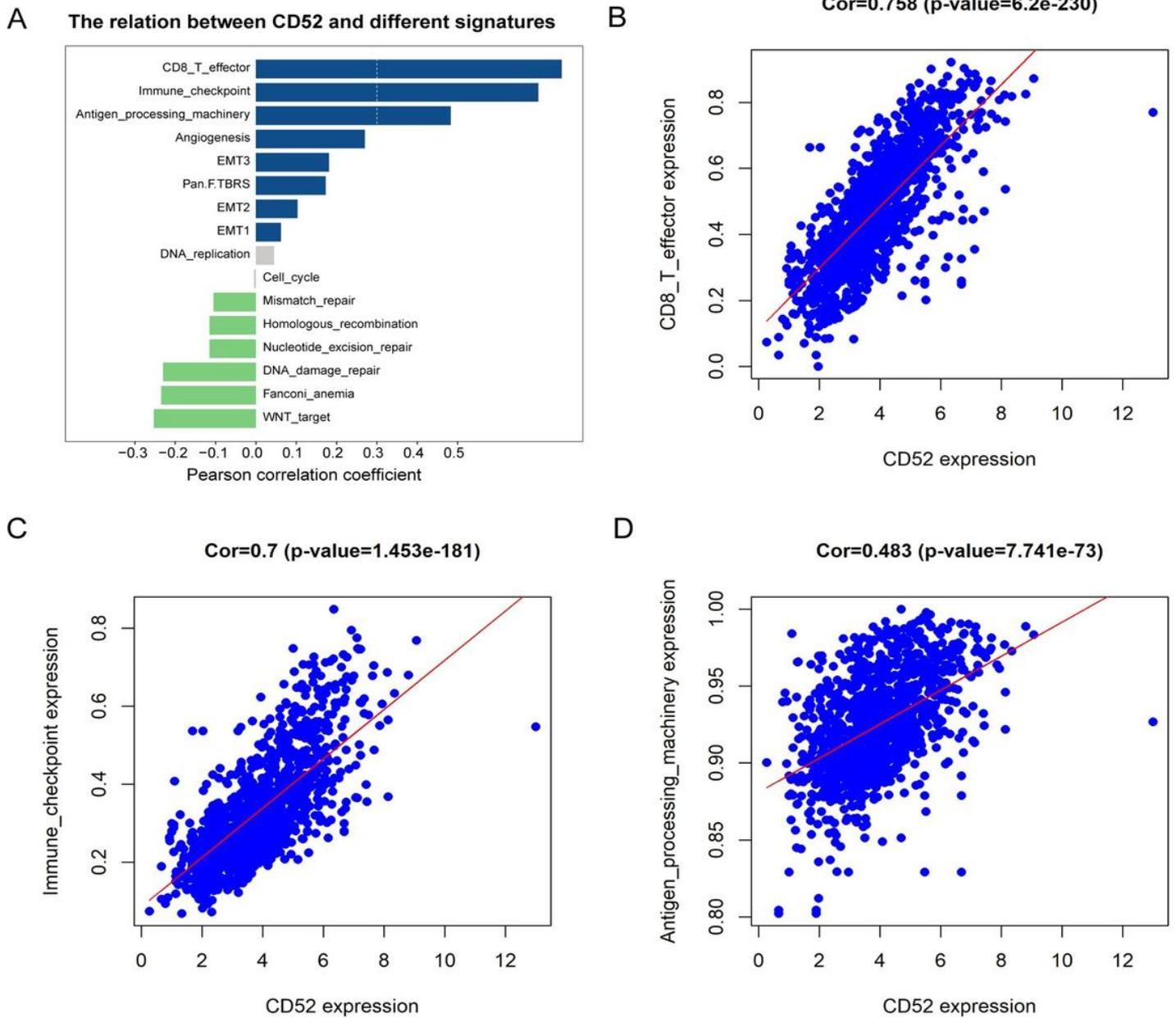


Figure 3

The relation between CD52 and different signatures. (A) The bar chart shows the relationship between CD52 and different signatures. (B–D) Top three CD52 related signatures.

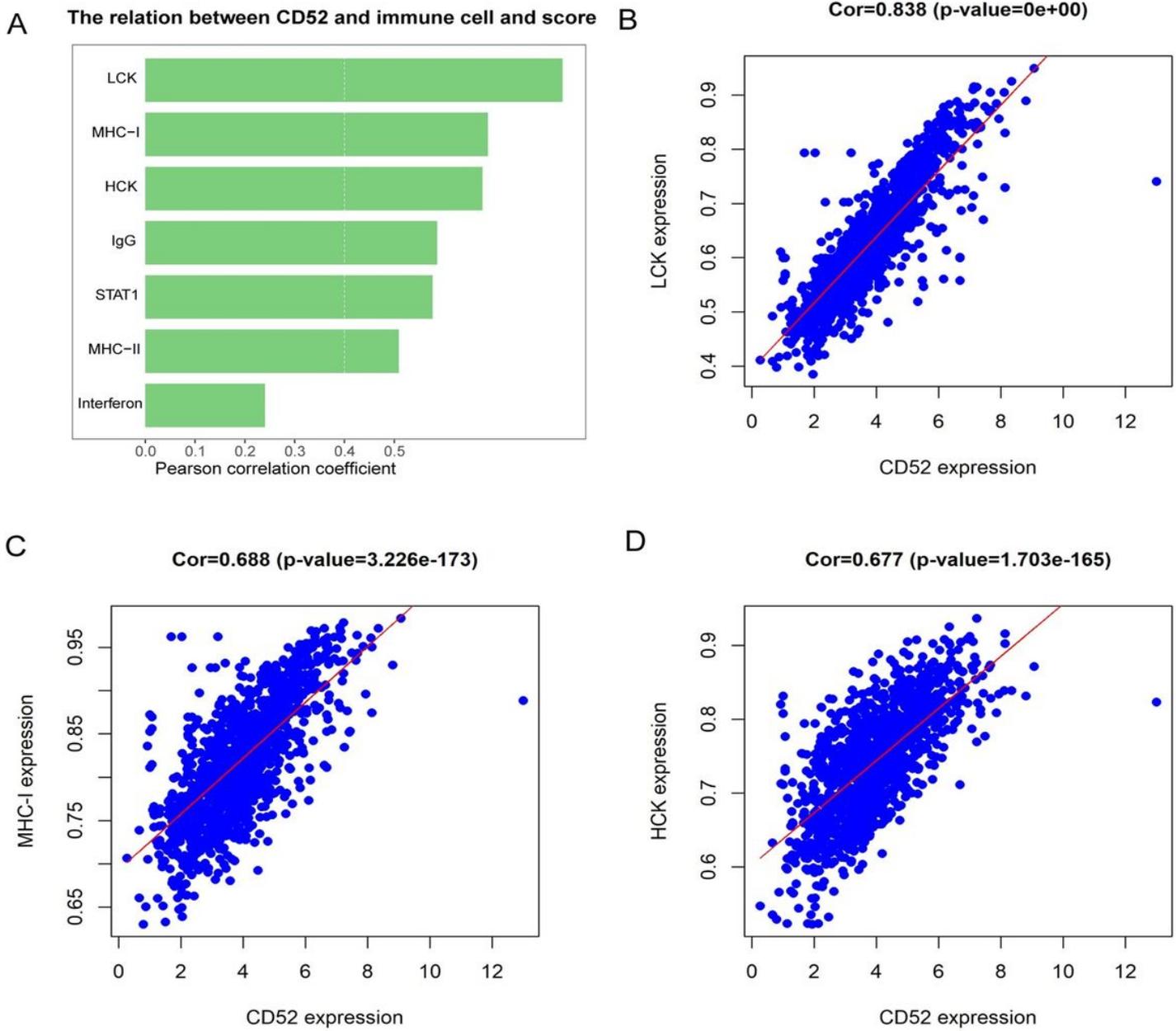


Figure 4

The relation between CD52 and immune cell and score (immune metagenes). (A) The bar chart shows the relationship between CD52 and immune metagenes. (B–D) Top three CD52 immune metagenes.

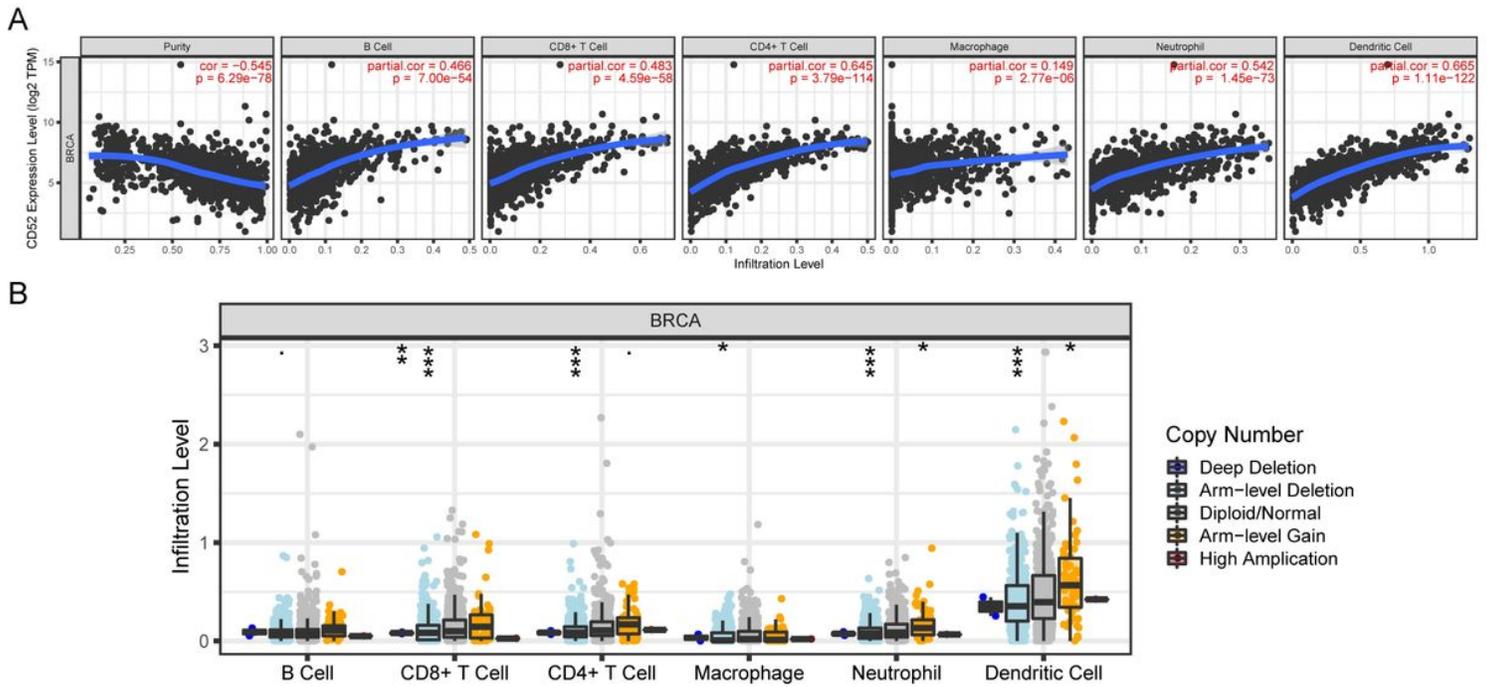


Figure 5

Correlation of CD52 expression with immune infiltration level in BC. (A) CD52 expression was positively correlated with immune infiltrating levels of B cells, CD8+ T cells, CD4+ T cells, macrophages, neutrophils and dendritic cells in BC. (B) CD52 was related to infiltration level for each SCNA category.

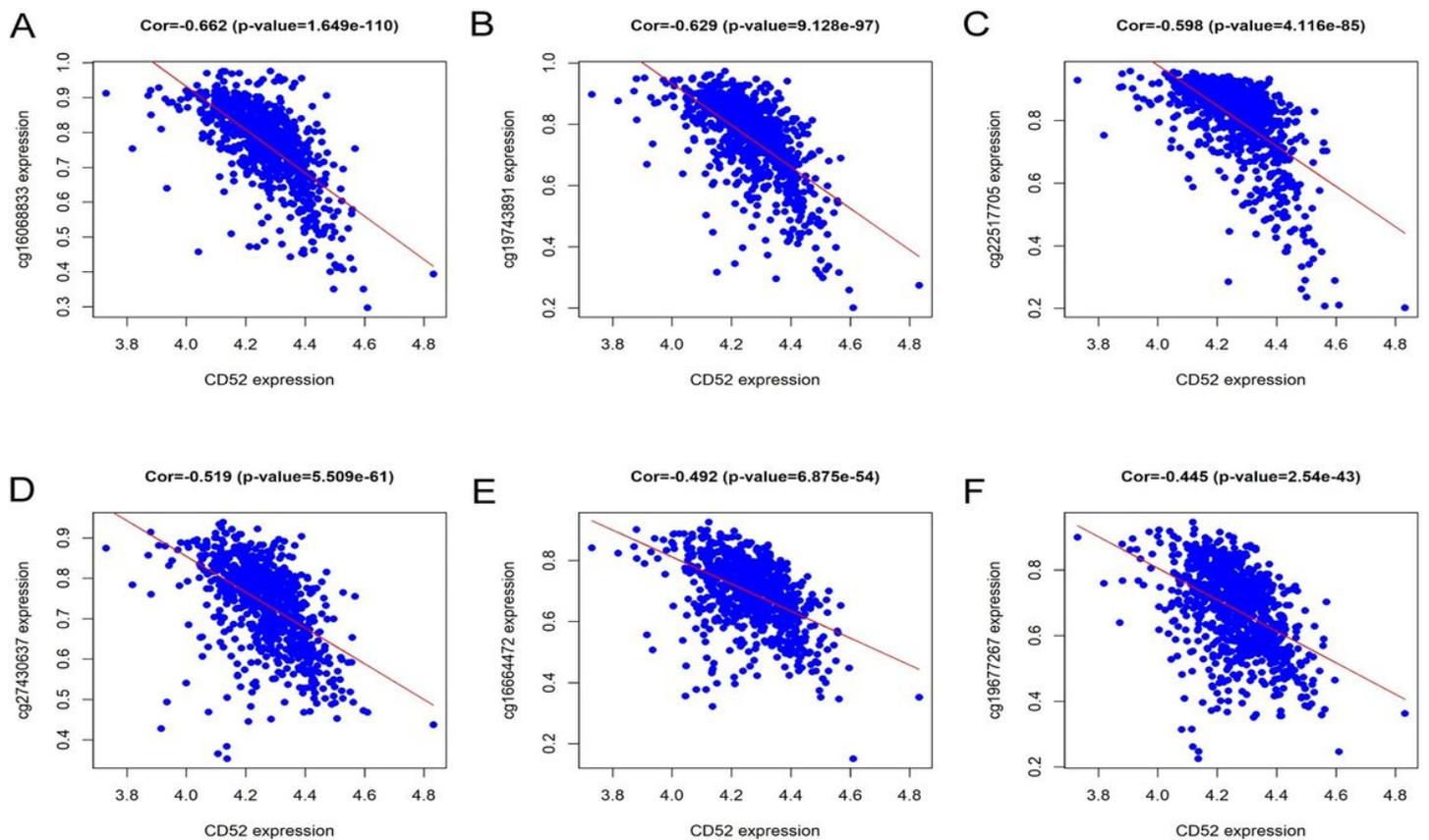


Figure 6

The correlation of CD52 expression and methylation. The correlation of CD52 expression and cg16068833(a), cg19743891(b), cg22517705(c), cg27430637 (d), cg16664472(e), and cg19677267(f).

A

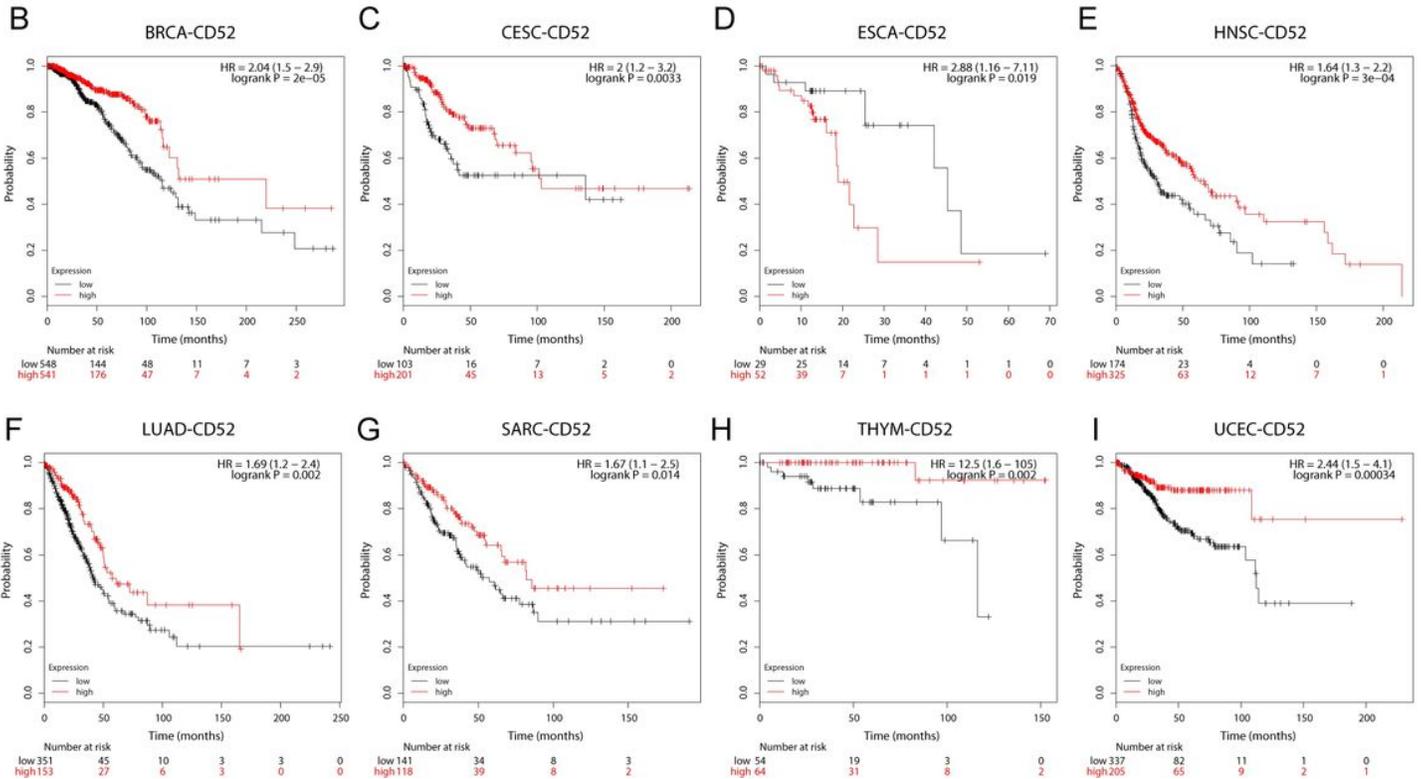
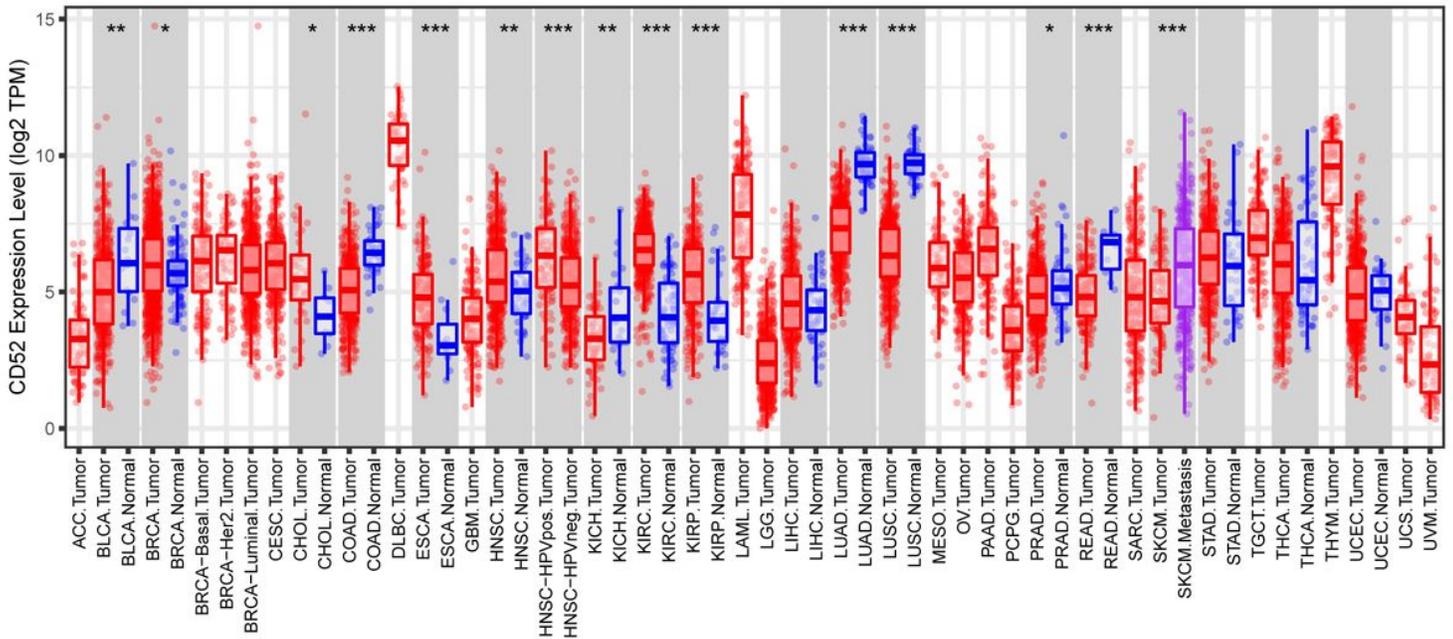


Figure 7

The expression and prognostic importance of CD52 in pan-cancers. (A) Human CD52 expression levels in different tumor types from TCGA database were determined by TIMER (*P < 0.05, **P < 0.01, ***P <

0.001).Kaplan-Meier survival curves comparing the high and low expression of CD52 in different types of cancer in the Kaplan-Meier plotter databases (B–G).

Supplementary Files

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

- [Supplementaryfigure1.tif](#)
- [supplementaryTable1.xlsx](#)
- [Supplementaryfigure2.tif](#)
- [supplementaryTable2.xlsx](#)
- [supplementaryTable3.xlsx](#)
- [Supplementaryfigure3.tif](#)