



Identification and validation of tumor microenvironment-related prognostic biomarkers in breast cancer

Shanliang Zhong^{1#}, Zhangjun Jia^{2#}, Heda Zhang³, Zhen Gong⁴, Jifeng Feng⁵, Hanzi Xu⁶

¹Center of Clinical Laboratory Science, The Affiliated Cancer Hospital of Nanjing Medical University & Jiangsu Cancer Hospital & Jiangsu Institute of Cancer Research, Nanjing, China; ²Department of Clinical Laboratory, The Affiliated Cancer Hospital of Nanjing Medical University & Jiangsu Cancer Hospital & Jiangsu Institute of Cancer Research, Nanjing, China; ³Department of General Surgery, The First Affiliated Hospital of Nanjing Medical University, Nanjing, China; ⁴Department of Gynecology, The Affiliated Obstetrics and Gynecology Hospital of Nanjing Medical University & Nanjing Maternity and Child Health Care Hospital, Nanjing, China; ⁵Department of Medical Oncology, The Affiliated Cancer Hospital of Nanjing Medical University & Jiangsu Cancer Hospital & Jiangsu Institute of Cancer Research, Nanjing, China; ⁶Department of Radiation Oncology, The Affiliated Cancer Hospital of Nanjing Medical University & Jiangsu Cancer Hospital & Jiangsu Institute of Cancer Research, Nanjing, China

Contributions: (I) Conception and design: J Feng, H Xu; (II) Administrative support: S Zhong; (III) Provision of study materials or patients: H Xu; (IV) Collection and assembly of data: H Zhang, Z Gong; (V) Data analysis and interpretation: S Zhong, Z Jia; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Jifeng Feng. Department of Medical Oncology, The Affiliated Cancer Hospital of Nanjing Medical University & Jiangsu Cancer Hospital & Jiangsu Institute of Cancer Research, Baiziting 42, Nanjing 210009, China. Email: feng_jifeng@sina.com; Hanzi Xu. Department of Radiation Oncology, The Affiliated Cancer Hospital of Nanjing Medical University & Jiangsu Cancer Hospital & Jiangsu Institute of Cancer Research, Nanjing 210009, China. Email: xuhanzi@njmu.edu.cn.

Background: Stromal cells and immune cells in tumor microenvironment (TME) have been reported to have significant value in the diagnosis and prognosis of cancers. We aimed to identify key biomarkers predicting survival in the TME of breast cancer.

Methods: Cell type enrichment analysis was performed to estimate cell scores using the xCell method with gene expression data from public database. Least absolute shrinkage and selection operator (LASSO) regression was used to identify key signature from the cell scores.

Results: Totally, 50 cells in TME had different scores between 1,078 breast cancer tissues and 112 adjacent normal tissues. We identified a 4-cell signature predicting breast cancer survival, including myocytes, natural killer T cell (NKT), conventional dendritic cell (cDC) and sebocytes, which was validated in the test set. Further analysis showed that cDC score was a key signature predicting prognosis of breast cancer. cDC score was significantly associated with molecular classification and stage of breast cancer, as well as expression level of Ki67. Spearman's correlation analysis found that cDC score was inversely correlated with the expression level of HER2. High cDC score may predicate better pathological complete response rate. Mechanism analysis indicated high cDC score was associated with elevated immune activity; IL-2 was a key gene associated with high cDC score; and Breast cancer patients with high IL-2 expression had a longer survival time.

Conclusions: In conclusion, cDC score was a key signature predicting prognosis for breast cancer. cDCs may exert antitumor effects by upregulating IL-2.

Keywords: Microenvironment; stromal cells; immune cells; survival; breast cancer

Submitted Jul 11, 2021. Accepted for publication Aug 26, 2021.

doi: 10.21037/tcr-21-1248

View this article at: <https://dx.doi.org/10.21037/tcr-21-1248>

Introduction

Breast cancer is commonest malignancy in women worldwide. Only in the United States, more than 250 thousand new cases are diagnosed with breast cancer each year and 66 thousand cases die from this malignancy (1). In China, it is estimated 208 thousand new cases and 55.5 thousand deaths in 2010 (2). Although significant progress has been made over the past decade in understanding of cancer biology, identification of risk factors, and new treatments of breast cancer, it is still insufficient in terms of early diagnosis, therapy, prevention and prediction of prognosis (3). Therefore, it is needed to find new markers for breast cancer.

The “seed and soil” hypothesis postulates that the tumor microenvironment provides fertile soil for the growth of tumor cells (4,5). Emerging evidence indicates that the cross-talk between tumor cells and tumor microenvironment exerts important effects on initiation, progression and metastasis of tumor (6). For example, tumors, as key drivers, control the differentiation of precursors of cancer-associated fibroblasts by secreting factors. Once present in the developing tumor, cancer-associated fibroblasts shape the tumor microenvironment to support tumor cell survival, dissemination, immune suppression, angiogenesis, and therapy resistance (7). Tumor microenvironment is a complex ecosystem of stromal cells and immune cells (8). Stromal cells and immune cells have been reported to have significant value in the diagnosis and prognosis of various cancers including breast cancer (9).

In the present study, we used xCell to perform cell type enrichment analysis from gene expression data (10). xCell is a gene signatures-based method learned from thousands of pure cell types from various sources and infers 64 immune and stromal cell types. xCell signatures have been validated extensively, and were shown to outperform other methods. We predicted cell scores using xCell with the data from The Cancer Genome Atlas (TCGA), Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) and Gene Expression Omnibus (GEO). Then, we selected and validated prognostic biomarkers from cell scores for breast cancer using the least absolute shrinkage and selection operator (LASSO) (11). We present the following article in accordance with the TRIPOD reporting checklist (available

at <https://dx.doi.org/10.21037/tcr-21-1248>).

Methods

Data acquisition

We downloaded RNA-Seq data and clinical data for 1,097 female breast cancer patients from the data portal for TCGA (accessed October 2020) (12). The data of 1,904 breast cancer patients were obtained from METABRIC database. Datasets of GSE96058, GSE20194, GSE22358, GSE25066 and GSE32646 were downloaded from GEO. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Cell type estimation

Cell type enrichment analysis was performed using the xCell (R package version 1.1) (10) with gene expression data. We compared the cell scores of tumors to those of adjacent normal tissues. Violin plot was drawn using vioplot package (version 0.3.5) in R software. The difference between two groups in cell scores was assessed using Wilcoxon test, and a P value less than 0.05 was considered significant.

Analysis of the cell score according to breast cancer pathological features

We investigated the relationships between cell scores and pathological features of breast cancer, including molecular classification [luminal (HER2⁻/ER⁺), HER2⁺(HR⁺), HER2⁺(HR⁻), and HER2⁻/HR⁻], stage (stage 1, stage 2, stage 3 and stage 4) and expression level of Ki67. R package of “beeswarm” (version 0.2.3) was used to illustrate the distribution of cell score in different groups.

Integration of protein-protein interaction (PPI) network

The PPI network was constructed as previously described (13). Briefly, we mapped a set of genes to Search Tool for the Retrieval of Interacting Genes (STRING) (14). Then, cytoscape software (version 3.8) was used to construct the network (15). The hub gene was the gene with highest connectivity.

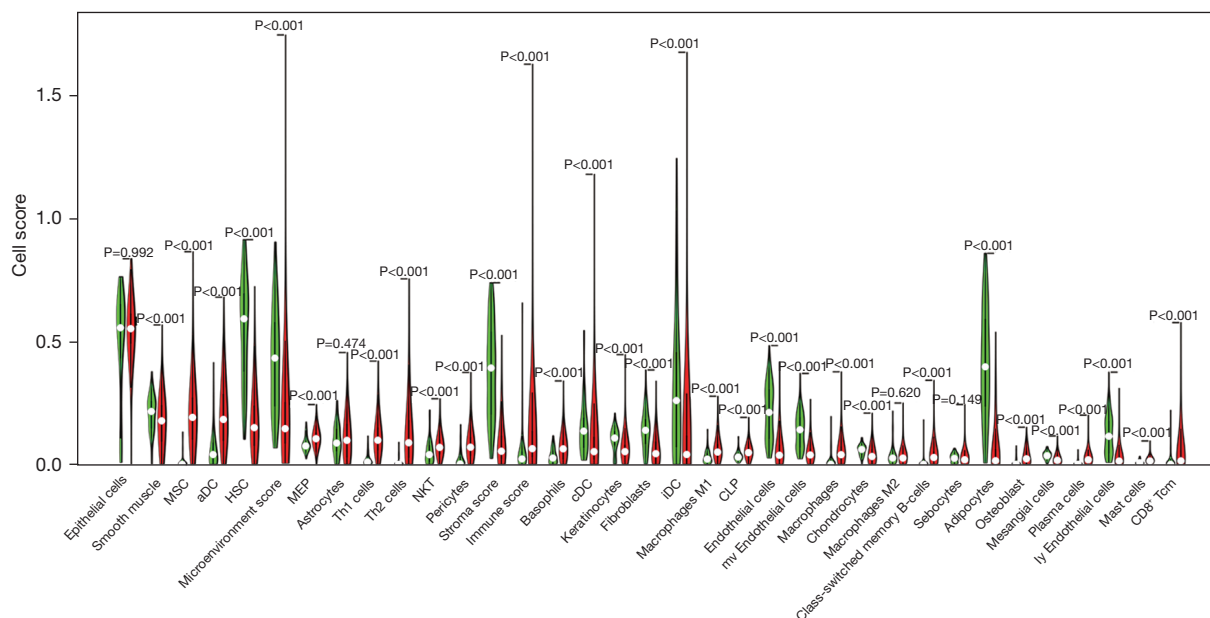


Figure 1 Cell score difference between breast cancer tissues and adjacent normal tissues in the TCGA cohort. Only the cells with a median score larger than the median of all the cell score are presented. Red, breast cancer tissues; and Green, adjacent normal tissues. aDC, activated dendritic cell; cDC, conventional dendritic cell; CLP, common lymphoid progenitor; mv, microvascular; ly, lymphatic; HSC, hematopoietic stem cell; iDC, immature dendritic cell; MEP, megakaryocyte-erythroid progenitor; NKT, natural killer T cell; MSC, mesenchymal stem cell; TCGA, The Cancer Genome Atlas.

Statistical analysis

We performed all the analyses using R software (version 4.0). For LASSO penalized regression analysis, we firstly calculated the cell scores using HTSeq-FPKM data from TCGA with the xCell. Only 1,063 breast cancer patients with a survival time longer than 0 day were included in the present analysis. Then the patients were randomly separated into two sets, training set and test set. We performed LASSO Cox regression with cell scores of the training-set patients. Depending on the regulation weight λ , all regression coefficients are shrunken to towards zero in LASSO, and the irrelevant features are set exactly to zero. Risk scores were calculated by as our previously study (3,16). We used “glmnet” package (R package version 4.0-2) to conduct the LASSO analysis and a P value <0.05 was considered statistically significance. Receiver operating characteristic (ROC) curve was drawn and the corresponding area under the ROC curve (AUC) was calculated to evaluate the prognostic value of the risk score by using ROCR package (R package version 1.0-11).

Regarding survival analysis, the patients were separated into two groups according to the expression level of a gene

or the risk scores, and the median was used as cut-off. Then, the log-rank test was used to assess the overall survival (OS) with survival package (R package version 3.1-7). We calculated hazard ratios (HRs) and their 95% confidence intervals (CIs) using Cox proportional hazards.

We performed enrichment analyses using Gene Set Enrichment Analyses (GSEA) software (version 4.1). False discovery rate (FDR) <0.05, nominal P value <0.05 and normalized enrichment score >1 were used as criteria of statistical significance.

Results

Assessment of cell scores in tumor microenvironment

To identify signature predicting breast cancer survival in tumor microenvironment, we investigated the gene expression profile of 1,097 female breast cancer patients from TCGA. We firstly estimated the cell scores for each sample, and then compared the scores between tumors and adjacent normal tissues. Figure S1 presents the scores of all the cells. Figure 1 presents the cells with a median score larger than the median of all the cell scores.

Conventional dendritic cell (cDC) predicts survival of breast cancer patients

The patients were randomly allocated to two sets, training set and test set. The characteristics of the two set patients were similar (Table S1). Key biomarkers associated breast cancer survival were screened from cell scores in training cohort using LASSO regression. We obtained four biomarkers including myocytes, natural killer T cell (NKT), cDC and sebocytes, and their coefficients were 0.098, -0.131, -0.021 and 0.012. The constructed risk score model divided patients into low- and high-risk groups according to their risk scores. Low-risk group had longer survival and less deaths (Figure 2A). Survival analysis indicated that high risk patients had a HR of 1.75 (95% CI: 1.08–2.83; $P=0.022$) (Figure 2B). ROC curve analysis indicated that the AUC was 0.90, 0.63 and 0.58 for 3-, 5- and 10-year survival (Figure 2C).

The analysis of the test cohort corroborated the findings in training cohort (Figure 2D). The HR of high-risk patients was 1.68 (95% CI: 1.07–2.66, $P=0.024$; Figure 2E) compared to low-risk score patients; The AUC was 0.71, 0.66 and 0.65 for 3-, 5- and 10-year survival (Figure 2F).

LASSO bootstrapping showed that cDC was a promising biomarker surpassing more than 4,000 counterparts covering the memory B-cells, keratinocytes and NKT (Figure 2G). Kaplan-Meier survival analysis indicated that higher cDC score was associated with longer OS in overall TCGA cohort (HR =0.61, 95% CI: 0.44–0.85; $P=0.0033$) (Figure 2H). We further validated it in the other two larger datasets, METABRIC (n=1,904) and GSE96058 (n=3,409). Breast cancer patients with high cDC score in METABRIC had a HR of 0.65 (95% CI: 0.58–0.73; $P<0.001$), and a similar result was found in GSE96058 (HR =0.70, 95% CI: 0.57–0.87; $P<0.001$) (Figure S2).

cDC score is associated with pathological features in breast cancer

We investigated the relationships between cDC score and pathological features of breast cancer with the data from TCGA and METABRIC (Figure 3). Regarding TCGA, cDC score had a significant association with the molecular classification ($P<0.001$; Figure 3A) and stage ($P=0.004$; Figure 3B), but not the expression level of Ki67 ($P=0.254$; Figure 3C). With respect to METABRIC, cDC score was significantly associated with stage ($P<0.001$; Figure 3E) and the expression level of Ki67 ($P<0.001$; Figure 3F) but not

the molecular classification ($P=0.348$; Figure 3D). From Figure 3A,3D, we noted that HER2-positive breast cancer had lower cDC score than HER2-negative breast cancer. Therefore, we explored whether a difference in cDC score exists between HER2-positive and HER2-negative breast cancer. TCGA ($P<0.001$) but not METABRIC ($P=0.712$) showed a significant association (Figure S3A,S3B). Spearman's correlation analysis found that cDC score was inversely correlated with the expression level of HER2 (also known as ERBB2) in both TCGA ($P<0.001$) and METABRIC ($P<0.001$) (Figure S3C,S3D). After adjusted for ERBB2 in the survival analysis, cDC score still had a significant association with breast cancer survival (TCGA: HR =0.61, 95% CI: 0.44–0.86, $P=0.004$; METABRIC: HR =0.67, 95% CI: 0.59–0.75, $P<0.001$).

High cDC score may predicate better pathological complete response (pCR) rate

We also explored the association between cDC score and pCR rate using datasets from GEO including GSE20194, GSE22358, GSE25066 and GSE32646 (Figure 4). GSE25066 indicated that high cDC score was significantly associated with pCR ($P=0.008$) (Figure 4C). GSE22358 ($P=0.173$) and GSE32646 ($P=0.170$) showed a borderline association (Figure 4B,4D). However, we failed to find a significant association between cDC score and pCR using GSE20194 (Figure 4A).

Tumors with high cDC score have elevated immune activity

In order to uncover the mechanisms of cDC score on outcome of breast cancer patients, GSEA was performed to compare the transcriptomes of high- and low-score tumors at the level of Hallmarks gene sets. We noted that all the 19 significant Hallmark gene sets enriched in high-score tumors (Table S2). The top six gene sets were IL-2 STAT5 signaling, complement, inflammatory response, KRAS signaling up, allograft rejection, and IL-6 JAK STAT3 signaling (Figure S4).

IL-2 was a key gene associated with high cDC score

To identify the key genes associated with high cDC score from Table S2, we firstly analyzed the associations between the genes and survival of breast cancer patients, and obtained 135 genes with a significant association (Table S3).

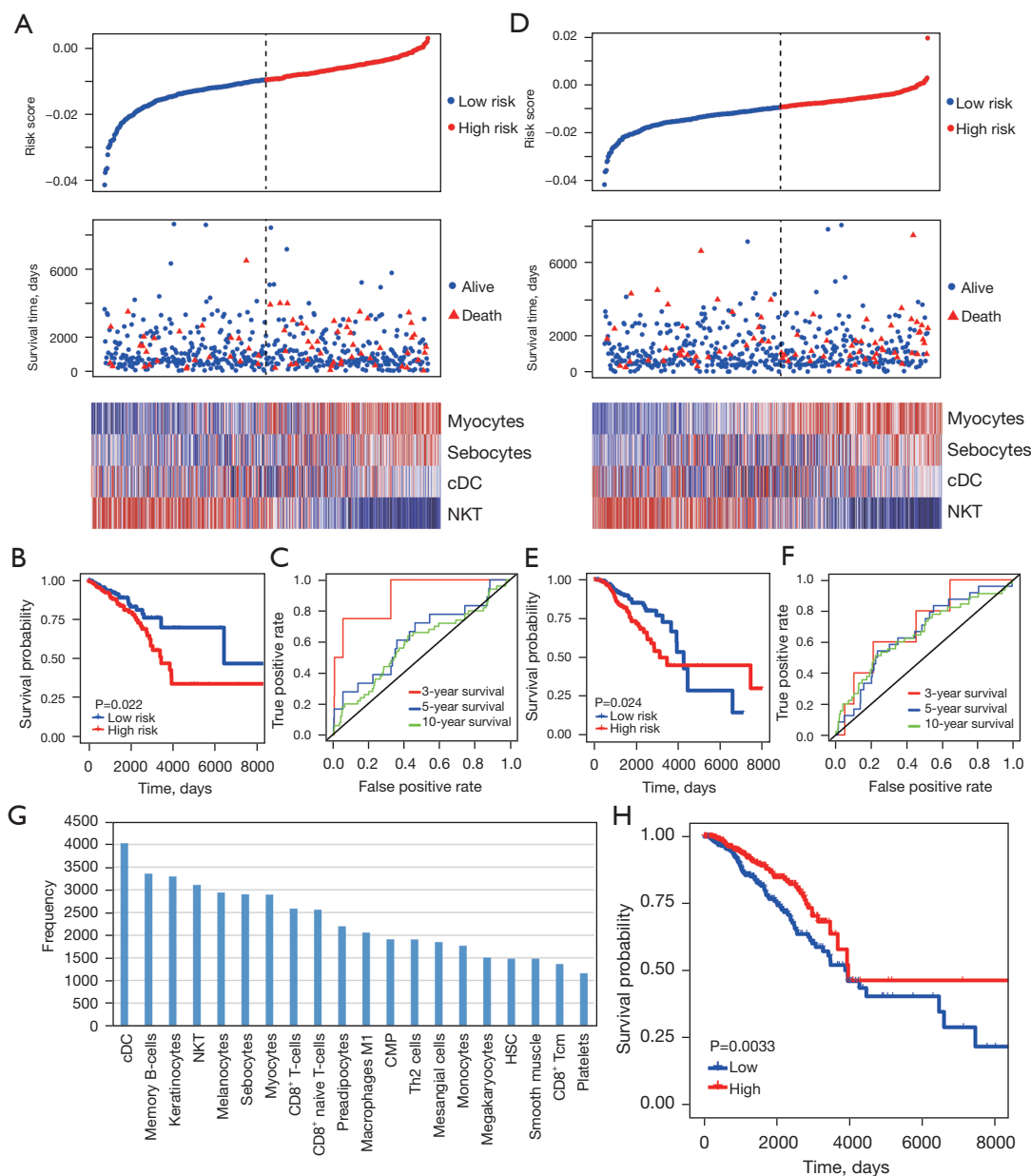


Figure 2 LASSO Cox regression of the cell scores with TCGA cohort. (A) The characteristics of the patients in training cohort order by the risk score calculated from the 4-cell signature. Dotted line: the median of the risk score. Top panel: the distribution of the risk score for the patients. Middle panel: survival time and status of the patients in the high- and low-risk groups, as defined by the risk score. Bottom panel: heatmap of the 4 cell scores. (B) The high-risk group has significantly shorter survival times than low-risk group in training cohort (HR =1.75, 95% CI: 1.08–2.83, P=0.022). (C) Receiver operating characteristic (ROC) curve of the 4-cell signature in test cohort. The area under the ROC curve (AUC) was 0.90, 0.63 and 0.58 for 3-, 5- and 10-year survival. (D) The characteristics of the patients in test cohort order by the risk score. (E) The high-risk group has significantly shorter survival times than low-risk group in test cohort (HR =1.68, 95% CI: 1.07–2.66, P=0.024). (F) ROC curve of the 4-cell signature in test cohort. The AUC was 0.71, 0.66 and 0.65 for 3-, 5- and 10-year survival. (G) Frequency of gene signatures selected by Lasso-bootstrapping highlighted the prominence of cDC contributing to predictive model. (H) Higher cDC score associated with longer overall survival in overall cohort (HR =0.61, 95% CI: 0.44–0.85; P=0.003). cDC, conventional dendritic cell; CI, confidence interval; CMP, common myeloid progenitor; HR, hazard ratio; HSC, hematopoietic stem cell; LASSO, least absolute shrinkage and selection operator; TCGA, The Cancer Genome Atlas.

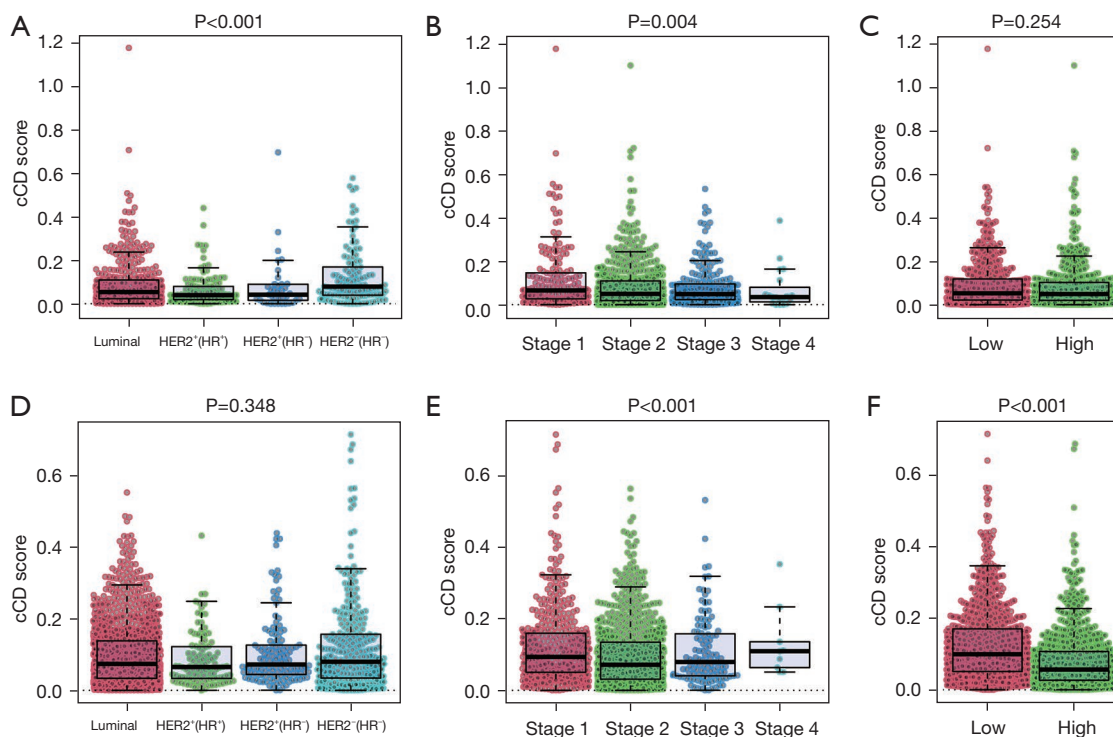


Figure 3 cDC score is associated with pathological features in breast cancer. (A) cDC score in different molecular subtypes of breast cancer in TCGA. (B) cDC score in different stages of breast cancer in TCGA. (C) Distribution of cDC score according to the expression of Ki67 in TCGA. The median was used as cut-off. (D) cDC score in different molecular subtypes of breast cancer in METABRIC. (E) cDC score in different stages of breast cancer in METABRIC. (F) Distribution of cDC score according to the expression of Ki67 in METABRIC. The median was used as cut-off. cDC, conventional dendritic cell; TCGA, The Cancer Genome Atlas.

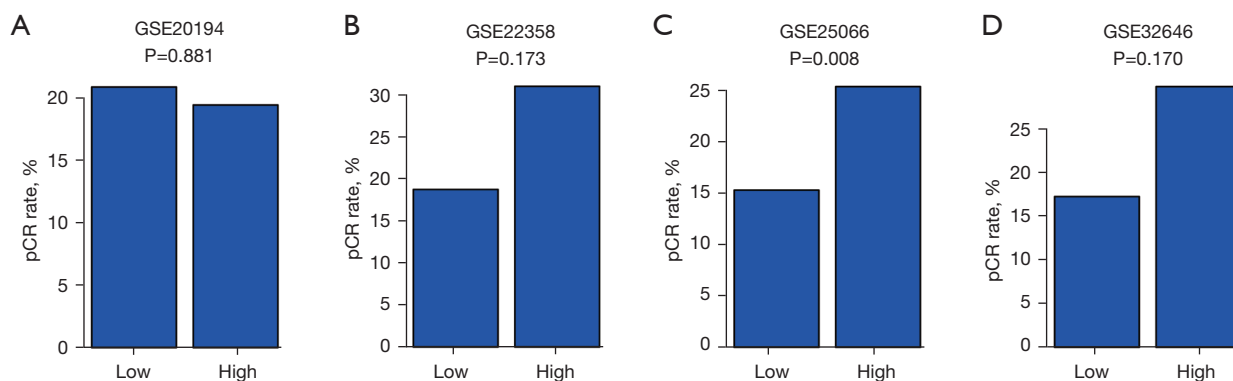


Figure 4 High cDC score may predicate better pathological complete response (pCR) rate. (A) The pCR rate in GSE20194. (B) The pCR rate in GSE22358. (C) The pCR rate in GSE25066. (D) The pCR rate in GSE32646.

Then, we constructed a network with the 135 genes using STRING. *Figure 5A* shows that IL-2 was the hub gene of the network. Breast cancer patients with high IL-2 expression had a HR of 0.56 (95% CI: 0.40–0.78; P < 0.001)

(*Table S2* and *Figure 5B*). Spearman's correlation analysis found that the expression level of IL-2 was positively correlated with cDC score (P < 0.001) (*Figure 5C*). We further analyzed the relationship between IL-2 and other

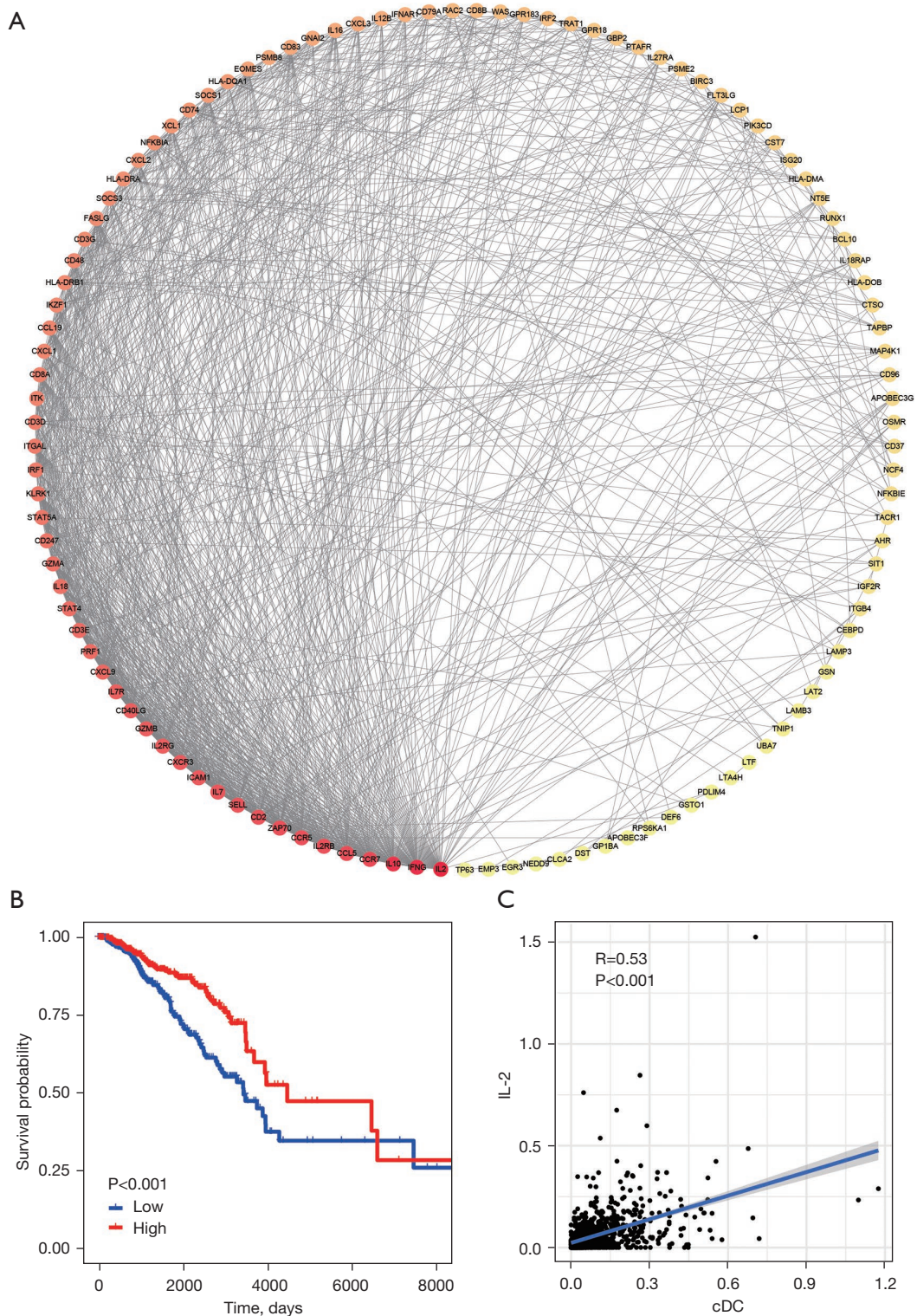


Figure 5 IL-2 is a key gene associated with high cDC score. (A) IL-2 is the hub gene of the protein-protein interaction (PPI) network. (B) High IL-2 expression is associated with a better overall survival (HR =0.56; 95% CI: 0.40–0.78; P<0.001). (C) The expression level of IL-2 was positively correlated with cDC score. cDC, conventional dendritic cell; CI, confidence interval; HR, hazard ratio.

cells. All the cells with a P value less than 0.05 are listed in Table S4.

Discussion

In present study, we investigated the association between cell enrichment in microenvironment and survival of breast cancer. We estimated cell scores of each sample, and found that 50 cells had different scores between tumors and adjacent normal tissues, suggesting cells in tumor microenvironment may have important roles in breast cancer. Then, we used LASSO regression analysis to identify signatures predicting breast cancer survival from cell scores, and obtained a 4-cell signature. In addition, cDC score was found to be a key signature associated with longer survival time. Further analyses showed that cDC score was significantly associated with stage of breast cancer and expression level of Ki67. Since Ki67 is a nuclear marker expressed in all phases of the cell cycle other than the G0 phase, Ki67 is usually used to assess cell proliferation (17,18). Our results indicated that low cDC score may be associated with high cell proliferation. We also found that cDC score was inversely correlated with the expression level of HER2. HER2 is a member of the epidermal growth factor family of receptor tyrosine kinases (19). Presence of HER2 protein overexpression or ERBB2 gene amplification has a prognostic and predictive role in breast cancer (20). Breast cancer with HER2 overexpression is associated with malignant tumor growth and poor clinical outcome (21). Finally, we explored the association between cDC score and pCR, and high cDC score may predicate better pCR rate. Taking together, high cDC score was associated with better prognosis of breast cancer.

Dendritic cells (DCs) are major antigen-presenting cells in tumor immune microenvironment and exert anti-cancer activity (22,23). DCs are divided into two principal cell populations, including cDCs and plasmacytoid DCs (24,25). cDCs are subdivided into two subtypes, type 1 cDC (cDC1) and type 2 cDC (cDC2) (23,24). Both cDC1 and cDC2 had antitumor activities. cDC1s are involved in cross-presenting tumor-associated antigens to CD8⁺ T cells which are important effector cells involving tumor cell elimination (24). cDC2s at least exert antitumor effects by inducing the proliferation of antitumor CD4⁺ T cells (26). Ferris *et al.* reported that both CD4⁺ and CD8⁺ T cells are needed for tumor rejection, selective deletion of major histocompatibility complex (MHC) class II in cDC1s, and inhibition of early CD4⁺ T-cell priming, indicating

that cDC1s and cDC2s are necessary for CD4⁺ T-cell priming (27). Therefore, cDCs play immunoregulatory roles in the tumor microenvironment. However, Oshi *et al.* did not find an association between cDC and disease-specific survival (DSS) with data from TCGA, but METABRIC showed a similar result as ours (22). Using different outcome may contribute to this difference. We used OS but Oshi *et al.* used DSS for TCGA. Combining the findings in our study, cDC score may be a key signature predicting OS of breast cancer.

We explored the potential mechanisms underlining the cDCs. GSEA analysis showed that tumors with high cDC score have elevated immune activity. We further found that IL-2 was a key gene associated with high cDC score. IL-2, the first of a series of lymphocytotropic hormones to be recognized, is released from activated T lymphocytes and rapidly cleared from human circulation with a half-life of 3–22 min (28). IL-2 is pivotal for the generation and regulation of the immune response (28,29). IL-2 is predominately secreted by CD4⁺ and CD8⁺ T cells (30,31), and DCs also secrete a lesser amount of IL-2 (32). As the first cytokine approved by the FDA to be used in cancer treatment, IL-2 exerts antitumor effects through immunomodulating T and NK cells in the tumor microenvironment.

When interpreting the results, the limitations of the present study need to be considered. We did not validate our results using experimental method. The mechanisms underlying cDCs are needed to further confirmed. We raised cDCs as a key signature, however, myocytes, NKT and sebocytes do indeed have their effects on breast cancer patients. Although our results suggested that cDC score was a key signature predicting prognosis for breast cancer, it is difficult to specifically profile cDCs from transcriptomic data (33).

Conclusions

In conclusion, cDC score was a key signature predicting prognosis for breast cancer and high cDC score was associated with elevated immune activity and better prognosis of breast cancer. cDCs may exert antitumor effects by upregulating IL-2.

Acknowledgments

Funding: This work was supported by the National Natural Science Foundation of China (grant number 81602551, 81872485), the High-level Innovative and

Entrepreneurial Talent Introduction Plan of Jiangsu Province (303073540ER21) and the young talents program of Jiangsu Cancer Hospital (QL201810).

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at <https://dx.doi.org/10.21037/tcr-21-1248>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://dx.doi.org/10.21037/tcr-21-1248>). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). Institutional ethical approval and informed consent were waived.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA Cancer J Clin* 2019;69:7-34.
2. Zeng H, Zheng R, Zhang S, et al. Female breast cancer statistics of 2010 in China: estimates based on data from 145 population-based cancer registries. *J Thorac Dis* 2014;6:466-70.
3. Zhong S, Chen H, Yang S, et al. Identification and validation of prognostic signature for breast cancer based on genes potentially involved in autophagy. *PeerJ* 2020;8:e9621.
4. Mao Y, Keller ET, Garfield DH, et al. Stromal cells in tumor microenvironment and breast cancer. *Cancer Metastasis Rev* 2013;32:303-15.
5. Paget S. The distribution of secondary growths in cancer of the breast. 1889. *Cancer Metastasis Rev* 1989;8:98-101.
6. Wang J, Zhang Q, Wang D, et al. Microenvironment-induced TIMP2 loss by cancer-secreted exosomal miR-4443 promotes liver metastasis of breast cancer. *J Cell Physiol* 2020;235:5722-35.
7. Houthuijzen JM, Jonkers J. Cancer-associated fibroblasts as key regulators of the breast cancer tumor microenvironment. *Cancer Metastasis Rev* 2018;37:577-97.
8. Jia D, Li S, Li D, et al. Mining TCGA database for genes of prognostic value in glioblastoma microenvironment. *Aging (Albany NY)* 2018;10:592-605.
9. Chu DT, Phuong TNT, Tien NLB, et al. The Effects of Adipocytes on the Regulation of Breast Cancer in the Tumor Microenvironment: An Update. *Cells* 2019;8:857.
10. Aran D, Hu Z, Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol* 2017;18:220.
11. Tibshirani R. Regression Shrinkage and Selection via the Lasso. *J R Stat Soc Series B Stat Methodol* 1996;58:267-88.
12. Tomczak K, Czerwińska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemp Oncol (Pozn)* 2015;19:A68-77.
13. Zhong S, Ma T, Zhang X, et al. MicroRNA expression profiling and bioinformatics analysis of dysregulated microRNAs in vinorelbine-resistant breast cancer cells. *Gene* 2015;556:113-8.
14. Szklarczyk D, Gable AL, Lyon D, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* 2019;47:D607-13.
15. Otasek D, Morris JH, Bouças J, et al. Cytoscape Automation: empowering workflow-based network analysis. *Genome Biol* 2019;20:185.
16. Liu M, Zhou S, Wang J, et al. Identification of genes associated with survival of breast cancer patients. *Breast Cancer* 2019;26:317-25.
17. Penault-Llorca F, Radosevic-Robin N. Ki67 assessment in breast cancer: an update. *Pathology* 2017;49:166-71.
18. Nielsen TO, Leung SCY, Rimm DL, et al. Assessment of Ki67 in Breast Cancer: Updated Recommendations From the International Ki67 in Breast Cancer Working Group. *J Natl Cancer Inst* 2021;113:808-19.
19. Redmond AM, Omarjee S, Chernukhin I, et al. Analysis of HER2 genomic binding in breast cancer cells identifies a global role in direct gene regulation. *PLoS One* 2019;14:e0225180.

20. Trapani D, Rajasekar AKA, Mathew A. More options for adjuvant treatment of HER2-positive breast cancer: How to choose wisely? *Int J Cancer* 2019;145:2901-6.
21. Wei T, Xing H, Wang H, et al. Bovine serum albumin encapsulation of near infrared fluorescent nano-probe with low nonspecificity and cytotoxicity for imaging of HER2-positive breast cancer cells. *Talanta* 2020;210:120625.
22. Oshi M, Newman S, Tokumaru Y, et al. Plasmacytoid Dendritic Cell (pDC) Infiltration Correlate with Tumor Infiltrating Lymphocytes, Cancer Immunity, and Better Survival in Triple Negative Breast Cancer (TNBC) More Strongly than Conventional Dendritic Cell (cDC). *Cancers (Basel)* 2020;12:3342.
23. Michea P, Noël F, Zakine E, et al. Adjustment of dendritic cells to the breast-cancer microenvironment is subset specific. *Nat Immunol* 2018;19:885-97.
24. Kim CW, Kim KD, Lee HK. The role of dendritic cells in tumor microenvironments and their uses as therapeutic targets. *BMB Rep* 2021;54:31-43.
25. Fu C, Peng P, Loschko J, et al. Plasmacytoid dendritic cells cross-prime naive CD8 T cells by transferring antigen to conventional dendritic cells through exosomes. *Proc Natl Acad Sci U S A* 2020;117:23730-41.
26. Binnewies M, Mujal AM, Pollack JL, et al. Unleashing Type-2 Dendritic Cells to Drive Protective Antitumor CD4+ T Cell Immunity. *Cell* 2019;177:556-571.e16.
27. Ferris ST, Durai V, Wu R, et al. cDC1 prime and are licensed by CD4+ T cells to induce anti-tumour immunity. *Nature* 2020;584:624-9.
28. Waldmann TA. The biology of interleukin-2 and interleukin-15: implications for cancer therapy and vaccine design. *Nat Rev Immunol* 2006;6:595-601.
29. Yang Y, Lundqvist A. Immunomodulatory Effects of IL-2 and IL-15; Implications for Cancer Immunotherapy. *Cancers (Basel)* 2020;12:3586.
30. Liao W, Lin JX, Leonard WJ. Interleukin-2 at the crossroads of effector responses, tolerance, and immunotherapy. *Immunity* 2013;38:13-25.
31. Paliard X, de Waal Malefijt R, Yssel H, et al. Simultaneous production of IL-2, IL-4, and IFN-gamma by activated human CD4+ and CD8+ T cell clones. *J Immunol* 1988;141:849-55.
32. Granucci F, Vizzardelli C, Pavelka N, et al. Inducible IL-2 production by dendritic cells revealed by global gene expression analysis. *Nat Immunol* 2001;2:882-8.
33. Sturm G, Finotello F, Petitprez F, et al. Comprehensive evaluation of transcriptome-based cell-type quantification methods for immuno-oncology. *Bioinformatics* 2019;35:i436-45.

Cite this article as: Zhong S, Jia Z, Zhang H, Gong Z, Feng J, Xu H. Identification and validation of tumor microenvironment-related prognostic biomarkers in breast cancer. *Transl Cancer Res* 2021;10(10):4355-4364. doi: 10.21037/tcr-21-1248

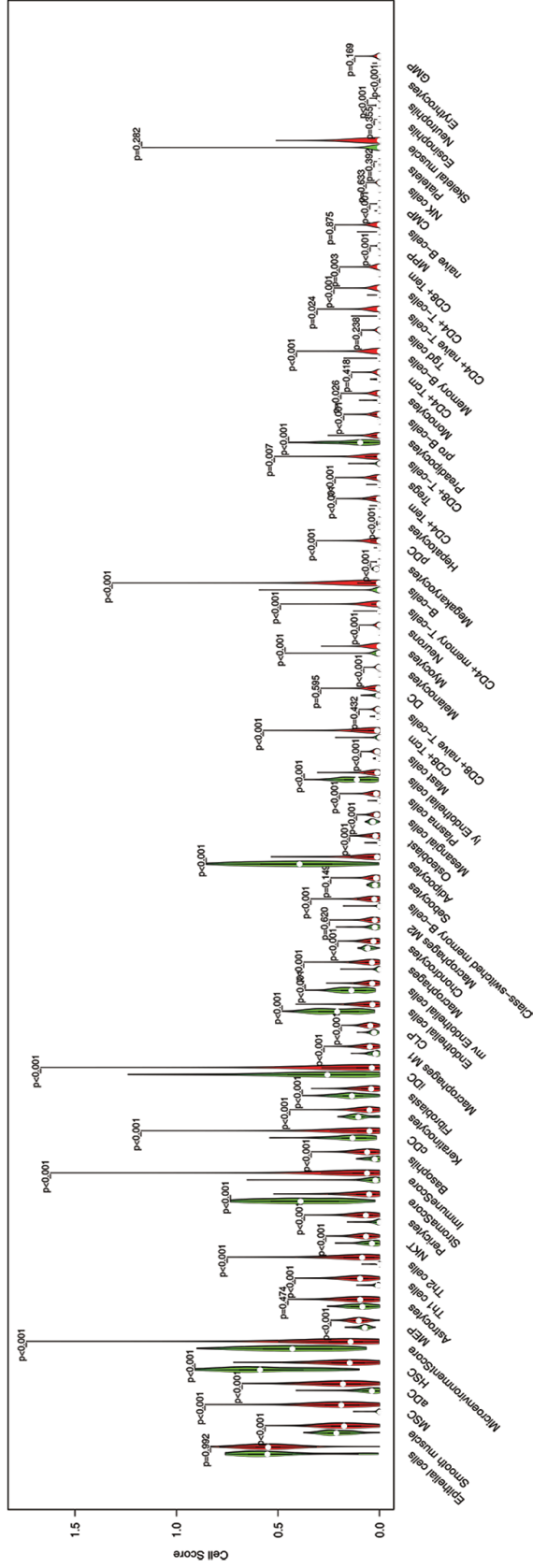


Figure S1 Cell score difference between breast cancer tissues and adjacent normal tissues in TCGA cohort. Red: breast cancer tissues; green: adjacent normal tissues. aDC, activated dendritic cell; cDC, conventional dendritic cell; CLP, common lymphoid progenitor; CMP, common myeloid progenitor; DC, dendritic cell; GMP, granulocyte-macrophage progenitor; HSC, hematopoietic stem cell; iDC, immature dendritic cell; MEP, megakaryocyte-erythroid progenitor; MPP, multipotent progenitor; MSC, mesenchymal stem cell; NK1, natural killer T cell; pDC, plasmacytoid dendritic cell; TCGA, The Cancer Genome Atlas.

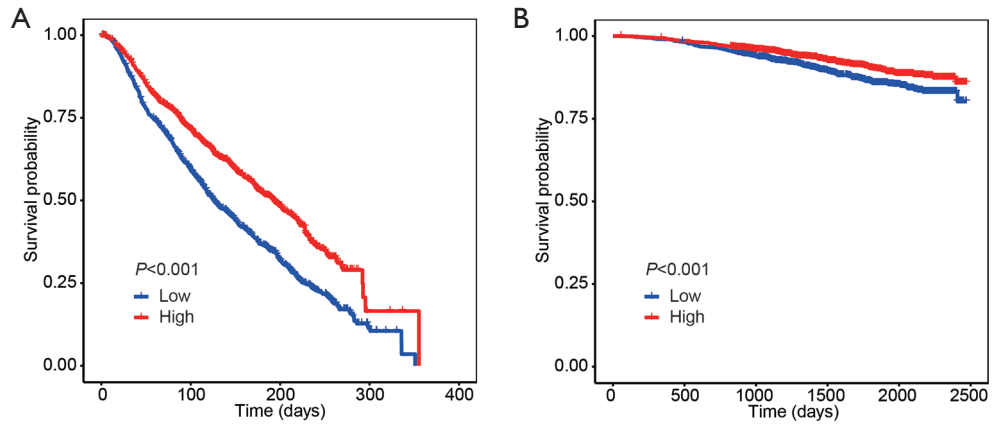


Figure S2 High cDC score is associated with a better overall survival. (A) METABRIC (HR =0.65; 95% CI: 0.58–0.73; $P < 0.001$); (B) GSE96058 (HR =0.70, 95% CI: 0.57–0.87; $P < 0.001$). cDC, conventional dendritic cell; CI, confidence interval; HR, hazard ratio.

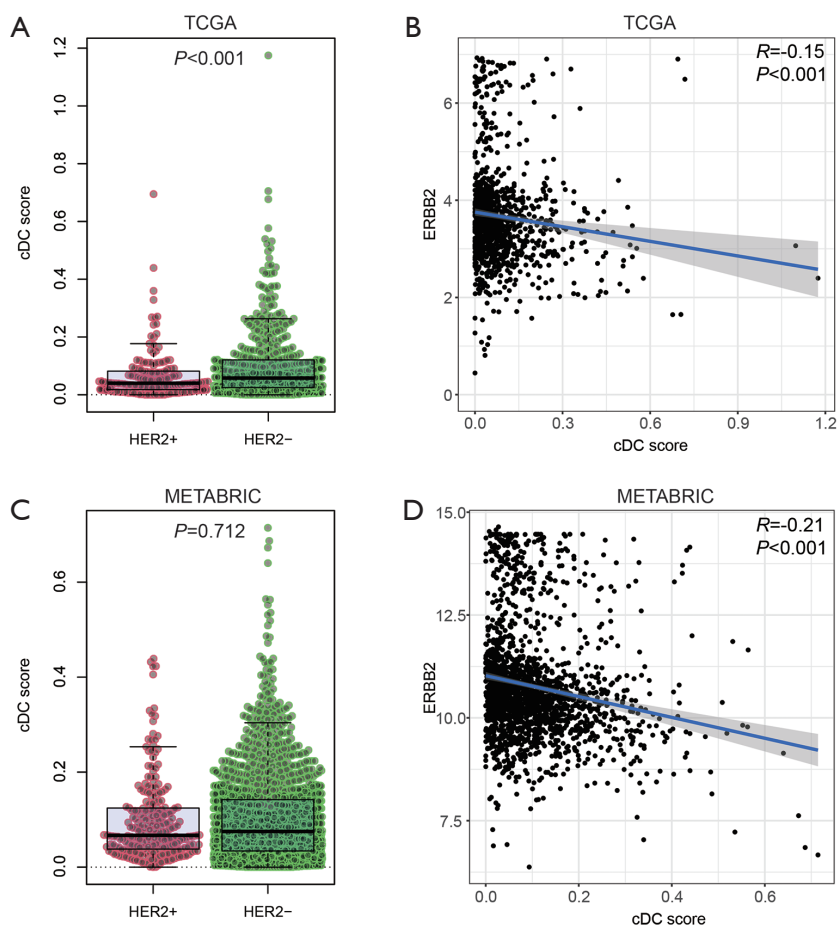


Figure S3 cDC score is inversely correlated with the expression level of HER2. (A) Distribution of cDC score according to HER2 status in TCGA; (B) distribution of cDC score according to HER2 status in METABRIC; (C) Spearman's correlation analysis between cDC score and expression level of HER2 in TCGA; (D) Spearman's correlation analysis between cDC score and expression level of HER2 in METABRIC. cDC, conventional dendritic cell; TCGA, The Cancer Genome Atlas.

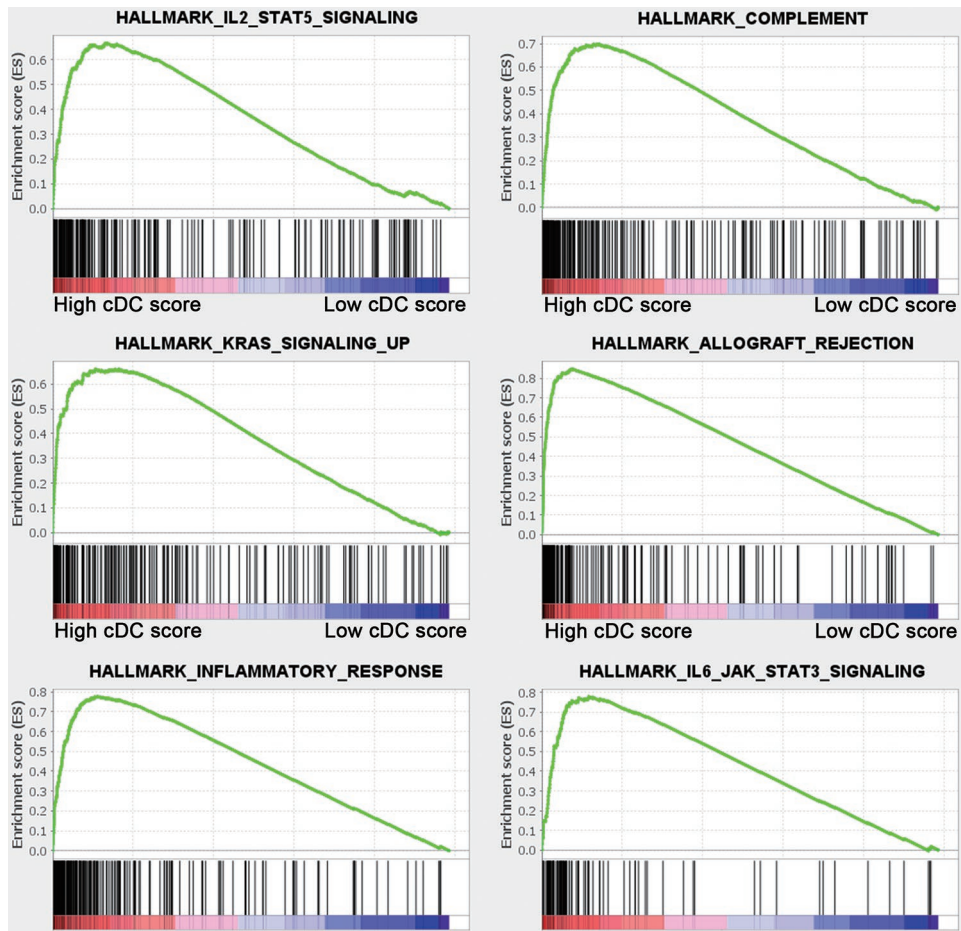


Figure S4 Gene set enrichment analysis (GSEA) with Hallmarks gene sets for which highly enriched in the high cDC tumors compared to low tumors in the TCGA. cDC, conventional dendritic cell; TCGA, The Cancer Genome Atlas.

Table S1 Characteristics of breast cancer patients

Characteristic	All (n=1,056)	Training set (n=528)	Test set (n=528)	P value
Age, years, mean \pm SD	58.05 \pm 12.93	58.15 \pm 13.13	57.94 \pm 12.74	0.781
Age, n (%)				
\leq 65 years	759 (71.88)	375 (71.02)	384 (72.73)	0.538
>65 years	297 (28.12)	153 (28.98)	144 (27.27)	
Race, n (%)				
White	736 (69.70)	373 (70.64)	363 (68.75)	0.503
Others	320 (30.30)	155 (29.36)	165 (31.25)	
Stage, n (%)				
I	181 (17.50)	94 (18.18)	87 (16.83)	0.513
II	597 (57.74)	289 (55.90)	308 (59.57)	
III	237 (22.92)	126 (24.37)	111 (21.47)	
IV	19 (1.84)	8 (1.55)	11 (2.13)	
Tumor size, n (%)				
T1	278 (26.40)	141 (26.76)	137 (26.05)	0.137
T2	606 (57.55)	290 (55.03)	316 (60.08)	
T3	132 (12.54)	72 (13.66)	60 (11.41)	
T4	37 (3.51)	44 (8.55)	13 (2.47)	
Node status, n (%)				
N0	498 (47.93)	262 (50.38)	236 (45.47)	0.070
N1	350 (33.69)	165 (31.73)	185 (35.65)	
N2	119 (11.45)	65 (12.50)	54 (10.40)	
N3	72 (6.93)	28 (5.38)	44 (8.48)	
Metastases, n (%)				
M0	873 (97.65)	436 (98.20)	437 (97.11)	0.283
M1	21 (2.34)	8 (1.80)	13 (2.89)	
Status, n (%)				
Alive	908 (85.98)	457 (86.55)	451 (85.41)	0.595
Death	148 (14.02)	71 (13.45)	77 (15.58)	
Survival time (days), mean \pm SD	1,259.53 \pm 1,193.34	1,237.46 \pm 1,239.282	1,281.60 \pm 1,146.32	0.548

SD, standard deviation.

Table S2 Gene set enrichment analysis (GSEA) with Hallmarks gene sets for which highly enriched in the high cDC tumors compared to low tumors in the TCGA cohort

Name	Size	ES	NES	NOM p-val	FDR q-val	FWER p-val	Genes of core enrichment
HALLMARK_IL2_STAT5_SIGNALING	199	0.667	2.779	0.000	0.000	0.000	IL10RA, TNFRSF1B, FGL2, CCR4, CD48, CD86, IL18R1, TNFRSF9, IRF8, GPR65, CST7, TRAF1, TNFRSF8, ICOS, CTLA4, EOMES, SELP, SYT11, CDC42SE2, TLR7, IRF4, ABCB1, LRRRC8, CD83, PRKCH, IL3RA, CD79B, IL10, IFNGR1, IL4R, BATF3, LTB, ADAM19, CSF1, TNFRSF4, SLC2A3, F2RL2, FLT3LG, PIM1, CTSZ, NRP1, MYO1E, SOCS1, GBP4, SWAP70, TGM2, SHE, PRNP, AHR, ITGAE, KLF6, IL2RA, PLAGL1, RORA, NFIL3, CXCL10, IL13, NT5E, GALM, IL2RB, PLPP1, PTCH1, IL1RL1, DENND5A, GUCY1B1, CSF2, LIF, GLIPR2, RGS16, GABARAPL1, BMPR2, PLSCR1, GADD45B, COL6A1, ANXA4, MAP3K8, EMP1, IKZF4, PLIN2, IGF2R, ITGAV, TNFRSF21, BMP2, GATA1, HOPX, GSTO1, ARL4A, CD81, CAPG, IKZF2, WLS, SELL, CYFIP1
HALLMARK_INFLAMMATORY_RESPONSE	200	0.776	2.743	0.000	0.000	0.000	GPR183, NLRP3, PTAFFR, IL7R, SLAMF1, EB13, CYBB, IL12B, CXCR6, CD69, LTA, C3AR1, IL18RAP, PTGER4, KCNA3, CD40, FPR1, CXCL9, CMKLR1, TLR1, SEMA4D, RGS1, IL15, IL15RA, SCARF1, P2RX7, EMP3, STAB1, PTPRE, CCL22, AXL, MEFV, C5AR1, CCL17, CCL2, IL1B, IL18, GPR132, APLNR, MSR1, NFKBIA, NMI, TACR1, ADGRE1, TLR2, ICAM1, HBEGF, NFKB1, PTGIR, PDPN, HRR1, CD82, OSMR, CXCL11, CCR2, IFNGR2, AQP9, SELE, HAS2, IRAK2, CALCRL, IL1R1, CHST2, SLC31A2, OSM, CD14, ACVR2A, BEST1, ABCA1, TLR3, CLEC5A, CCL24, LAMP3, NAMPT, CCR7, TNFAIP6, TAPBR, LPAR1, EDN1, ITGA5, CX3CL1, TNFSF15, HIF1A, MARCO, GCH1, PDE4B, MEP1A, IFNAR1, ITGB8, ITGB3, GNA15
HALLMARK_COMPLEMENT	200	0.697	2.739	0.000	0.000	0.000	LCP2, PLEK, PIK3R5, WAS, GNGT2, FCN1, LCK, PIK3CG, TNFAIP3, CTSS, GZMK, CD40LG, GZMA, SH2B3, FCER1G, APOBEC3G, CCL5, C1QA, LYN, C3, C1QC, SPOCK2, CASP4, SERPING1, IRF1, CRT, CASP10, CASP1, DOCK10, GZMB, C1S, C1R, CTSC, PLA2G7, LGMN, JAK2, FYN, CFH, GP1BA, GNG2, GNAI2, RHOA, CTSS, LIPA, PSMB9, CASP5, LAP3, DOCK4, OLR1, APOC1, DPP4, ITGAM, LRP1, ZEB1, F10, CTSL, ERAP2, CTSH, IRF2, TIMP2, APOBEC3F, CTSS, MMP8, CBLB, C2, DGKG, CR2, MMP12, EHD1, ITIH1, XPNPEP1, IL6, FDX1, NOTCH4, ANXA5, SERPINE1, PLAUR, PDP1, F3, LTF, USP15, L3MBTL4, GCA, GNB4, DGKH, HPCAL4, COL4A2, CPM, ZFPM2, LGALS3, MMP14, CEBPB, CXCL1, PRCP, LTA4H, KIF2A, CASP7, CD59, DOCK9
HALLMARK_KRAS_SIGNALING_UP	200	0.660	2.655	0.000	0.000	0.000	CLEC4A, DOCK2, IKZF1, LAPTM5, ITGB2, ETS1, IL2RG, LAT2, TLR8, MAP4K1, PDCC1LG2, LY96, PRDM1, CD37, MAFB, LCP1, CSF2RA, BIRC3, ENG, ADAMDEC1, PECAM1, GYPC, ADGRA2, IL33, GFPT2, ADGRL4, SPARCL1, CXCR4, TSPAN7, GPNMB, CBL, PLVAP, TMEM176B, EPHB2, PRRX1, NIN, ADAM8, HSD11B1, FLT4, PPP1R15A, GNG11, SPRY2, F13A1, ST6GAL1, PSMB8, GUCY1A1, MALL, KLF4, AKAP12, TMEM176A, ANKH, SPON1, GPRC5B, ADAM17
HALLMARK_ALLOGRAFT_REJECTION	200	0.847	2.621	0.000	0.000	0.000	CD4, HCLS1, HLA-DRA, HLA-DMB, SRGN, PTPRC, CCR5, CD3G, CD2, HLA-DQA1, HLA-DOA, NCF4, PRKCB, CRTAM, CD1D, ITK, CD247, CD96, IL16, FGR, HLA-DMA, CD3E, FYB1, SPI1, CCR2, CD3D, STAT4, CD28, CXCR3, IGSF6, CD74, ST8SIA4, CD8A, SIT1, LY86, ITGAL, HLA-E, PRF1, CFP, KLRD1, ZAP70, FASLG, NCR1, IL2, CD7, CCL19, FCGR2B, IFNAR2, CD8B, CD79A, B2M, HLA-DOB, LY75, TRAT1, TLR6, F2R, GBP2, IFNG, CCL13, TAP2, PSMB10, NCK1, PTPN6, WARSI, CSK, SOCS5, FLNA, HLA-A, ELF4, CD47, IL27RA
HALLMARK_IL6_JAK_STAT3_SIGNALING	87	0.779	2.596	0.000	0.000	0.000	IL12RB1, CSF2RB, CD38, FAS, ACVRL1, IL9R, CCR1, IL17RA, ITGA4, TGFB1, IL7, IL10RB, CRLF2, MYD88, STAT1, STAT2, A2M, TNFRSF1A, LEPR, PTPN2, HMOX1, BAK1, TYK2
HALLMARK_TNFA_SIGNALING_VIA_NFKB	200	0.734	2.543	0.000	0.000	0.000	TNFAIP8, CCL4, CFLAR, GEM, RNFB19B, CD80, SGK1, ETS2, RELB, SOD2, FUT4, NFKBIE, SERPINB8, EGR2, DUSP2, PNRC1, PLPP3, CCN1, NFKB2, TAP1, MCL1, STAT5A, KLF2, SOCS3, SLC2A6, TANK, B4GALT5, CXCL3, KLF9, BIRC2, DUSP1, NR4A3, MSC, ZFP36, TNF, DNAJB4, TNIP1, CCN1, ATF3, BTG1, EGR3, DRAM1, NFE2L2, CXCL2, JUN, TNFAIP2, TRIP10, MARCKS, JUNB, EGR1, SPSB1, NR4A1, PER1, JAG1, FOS, NFAT5, FOSL2, RCAN1, TIPARP, LAMB3, SAT1, ZC3H12A, TNC, FOSB, CEBPD, IFIH1, LITAF, BCL6, RIPK2, FOSL1, BCL2A1, MAP2K3, KDM6B, IL23A, TSC22D1, BTG2, BTG3

Table S2 (continued)

Table S2 (continued)

Name	Size	ES	NES	NOM p-val	FDR	q-val	FWER	p-val	Genes of core enrichment
HALLMARK_INTERFERON_GAMMA_RESPONSE	200	0.791	2.325	0.000	0.000	0.000	0.001	0.001	HLA-DRB1, CITA, ITGB7, SAMHD1, IL18BP, SLAMF7, P2RY14, KLRK1, XCL1, MARCHF1, FCGR1A, EPSTH1, NLRC5, CD274, GPR18, IFI30, VAMP5, APOLE6, BANK1, LATS2, SAMD9L, NOD1, ARID5B, IRF5, HLA-B, PML, PELI1, CASP8, VCAM1, TXNIP, TRIM21, ZBP1, SP110, LYSDMD2, TRIM14, SSPN, UPP1, PARRP14, IDO1, XAF1, MX2, ISG20, RIPK1, GBP6, BMTF2, ST3GAL5, ZNFY1, EIF4E3, PARRP12
HALLMARK_APOPTOSIS	161	0.521	2.296	0.000	0.000	0.000	0.001	0.001	PLCB2, DPYD, HGF, CYLD, MMP2, NEDD9, PDGFBR, DCN, ANXA1, ROCK1, GSN, LUM, BGN, GPX1, FEZ1, AVPR1A, BCL10, GPX3, SATB1, CAV1, CTNNB1, CCND2, APP, IL1A
HALLMARK_COAGULATION	137	0.578	2.250	0.000	0.000	0.000	0.001	0.001	PROS1, CFI, VWF, CTSK, KLF7, THBD, ISCU, P2RY1, GNG12, RAPGEF3
HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION	200	0.677	2.209	0.000	0.000	0.000	0.002	0.002	WIPF1, GLIPR1, ABI3BP, DAB2, SLIT2, IL32, FSTL1, LAMA2, CXCL12, FBLN2, SFRP4, VIM, DPYSL3, NNMT, CCN2, ADAM12, GAS1, CALD1, HTRA1, COL16A1, COL6A3, MXRA5, LAMC1, SNAI2, FBLN1, COL3A1, MMP3, FBN1, FBLN5, LOX, SERPINE2, NID2, EFEMP2, COL6A2, BMP1, FAP, MYLK, ELN, DST, CTHRC1, SLIT3, FERMT2, VCAN, GPX7, ITGB1, P3HT, SPARC, COL1A1, COLGALT1, SGCD, FOXC2, CDH11, THY1, LOXL1, COL5A1, FUCA1, PCOLCE, COL1A2, TNFRSF11B, POSTN, TGFB1, VEGFC, SFRP1, COL5A2, NOTCH2, ECM2, CALU, WNT5A, IGFBP3, CDH6, TPM1, PDLIM4, MATN2, LRRCL15, CXCL8, PMP22, TAGLN
HALLMARK_APICAL_JUNCTION	200	0.521	2.162	0.000	0.000	0.000	0.004	0.004	RAC2, MSN, SYK, ICAM2, ARPC2, SIRPA, CD209, EPB41L2, ZYX, CNN2, RHOF, CADM3, CLDN11, CD34, SKAP2, SDC3, CLDN15, ACTB, AKT3, CLDN5, CAP1, ITGA9, ADAM23, JAM3, ACTN1, CDH3, LAYN, NECTIN3, NRXN2, ADAMTS5, MYH9, CD99, VCL, MAPK11, TSPAN4, DSC1, MAPK14, DSC3, PARVA, ITGB4, NECTIN1, MPZL1, SGCE, YWHAH, GNAI1, PARD6G, RSU1, COL17A1, IKBKG, VASR, MMP9, CLDN19, NLGN3, PFN1, FSCN1
HALLMARK_UV_RESPONSE_DN	144	0.522	1.955	0.006	0.007	0.007	0.052	0.052	PIK3CD, CELF2, TGFB2, NR3C1, GRK5, ADD3, RND3, PRDM2, RASA2, SYNE1, DLC1, PTPRM, PTGFR, PTPN21, PRKCA, ATXN1, ID1, KIT, SCHIP1, EFEMP1, RUNX1, KALRN, TFPI, ATP2B4, NFIB, WDR37, MAPIB
HALLMARK_INTERFERON_ALPHA_RESPONSE	97	0.725	1.947	0.010	0.007	0.007	0.055	0.055	TMEM140, LPAR6, NCOA7, HLA-C, UBA7, NUB1, PROCR, TRIM5, UBE2L6, PSME2
HALLMARK_ANGIOGENESIS	36	0.567	1.873	0.002	0.013	0.013	0.102	0.102	SLCO2A1, PRG2, S100A4, SPP1, TIMP1, PDGFA
HALLMARK_PIBK_AKT_MTOR_SIGNALING	105	0.433	1.839	0.002	0.016	0.016	0.126	0.126	SLA, CAMK4, DAPP1, RPS6KA3, ACTR3, MKNK1, CAB39, PIKFYVE, ACTR2, RALB, RPS6KA1
HALLMARK_P53_PATHWAY	199	0.410	1.816	0.004	0.019	0.019	0.151	0.151	TCN2, RAP2B, DGKA, GM2A, PLK3, DEF6, ADA, ZFP36L1, NOTCH1, ZMAT3, EPHA2, PTPN14, TP63, RALGDS, RB1, VDR, FBXW7, TAX1BP3, RHBDF2, STEAP3, POLH, POM121, HDAC3, ZBTB16, TGFA, VWA5A, IRAK1, TNFSF9, CEBPA, CCND3, CLCA2, PRMT2, SPHK1
HALLMARK_HYPOXIA	200	0.434	1.805	0.004	0.020	0.020	0.158	0.158	PLAC8, AMPD3, SLC2A5, HS3ST1, SRPX, CHST3, TES, NDST2, CAVIN1, TGFB3, CAVIN3, NAGK, SLC6A6, HAS1, TPST2, CCN5, PGF, EXT1, GLRX, RBPJ, NDST1, KLHL24, HEXA, MAFF, PGM2, ERF1
HALLMARK_TGF_BETA_SIGNALING	54	0.477	1.689	0.021	0.041	0.041	0.282	0.282	VWTR1, LTBP2, ACVR1, SKI, ID3, TGFB1, SMURF1, FKBP1A, SPTBN1, KLF10, NOG, TJP1, MAP3K7, SMAD1, SMURF2, RHOA

cDC, conventional dendritic cell; ES, enrichment score; FDR, false discovery rate; FWER, family-wise error rate; NES, normalized enrichment score; NOM, normal; TCGA, The Cancer Genome Atlas.

Table S3 Survival analysis with the key genes associated with high cDC score in gene set enrichment analysis

Genes	P	HR (95% CI)
<i>PDP1</i>	0.000	1.969 (1.403–2.763)
<i>RAC2</i>	0.000	0.531 (0.380–0.741)
<i>CEBPD</i>	0.000	0.533 (0.382–0.744)
<i>IL2</i>	0.000	0.557 (0.401–0.774)
<i>SPOCK2</i>	0.001	0.579 (0.416–0.805)
<i>CD3E</i>	0.001	0.577 (0.414–0.804)
<i>DEF6</i>	0.001	0.576 (0.412–0.804)
<i>CD3D</i>	0.001	0.577 (0.413–0.805)
<i>APOBEC3G</i>	0.001	0.585 (0.421–0.812)
<i>CD247</i>	0.002	0.587 (0.421–0.819)
<i>APOBEC3F</i>	0.002	0.593 (0.426–0.825)
<i>IL27RA</i>	0.002	0.595 (0.428–0.826)
<i>HLA-DRA</i>	0.002	0.596 (0.428–0.831)
<i>EGR3</i>	0.002	0.601 (0.434–0.834)
<i>BTG1</i>	0.002	0.602 (0.435–0.834)
<i>UBA7</i>	0.002	0.595 (0.426–0.832)
<i>TRAT1</i>	0.002	0.599 (0.430–0.835)
<i>IFNG</i>	0.003	0.601 (0.432–0.836)
<i>ITIH1</i>	0.003	0.604 (0.435–0.838)
<i>TNIP1</i>	0.003	0.602 (0.433–0.838)
<i>FLT3LG</i>	0.003	0.600 (0.430–0.838)
<i>TAPBP</i>	0.003	0.608 (0.438–0.844)
<i>PSME2</i>	0.003	0.607 (0.437–0.845)
<i>STAT4</i>	0.003	0.606 (0.433–0.846)
<i>CXCL1</i>	0.003	0.613 (0.443–0.850)
<i>GZMA</i>	0.003	0.609 (0.438–0.849)
<i>IL12B</i>	0.004	0.613 (0.442–0.852)
<i>IRF2</i>	0.004	0.618 (0.446–0.854)
<i>CD40LG</i>	0.004	0.612 (0.439–0.854)
<i>CD79A</i>	0.004	0.618 (0.445–0.858)
<i>PSMB8</i>	0.004	0.621 (0.447–0.862)
<i>CD96</i>	0.005	0.619 (0.444–0.863)
<i>FASLG</i>	0.005	0.620 (0.444–0.865)
<i>NFKBIA</i>	0.005	0.628 (0.453–0.872)
<i>MAP4K1</i>	0.005	0.625 (0.448–0.871)
<i>IGF2R</i>	0.006	1.594 (1.145–2.218)
<i>CD2</i>	0.006	0.628 (0.451–0.874)
<i>CD48</i>	0.006	0.628 (0.451–0.875)
<i>GSTO1</i>	0.006	0.630 (0.453–0.876)
<i>TGFB1</i>	0.006	1.580 (1.138–2.194)
<i>HLA-DQA1</i>	0.007	0.638 (0.459–0.887)
<i>LAMB3</i>	0.007	0.638 (0.458–0.887)
<i>RPS6KA1</i>	0.008	0.639 (0.460–0.887)
<i>RUNX1</i>	0.008	0.639 (0.460–0.888)
<i>CD8B</i>	0.008	0.638 (0.458–0.888)

Table S3 (continued)

Table S3 (continued)

Genes	P	HR (95% CI)
<i>HLA-DOB</i>	0.008	0.642 (0.463–0.889)
<i>SOCS3</i>	0.008	0.643 (0.464–0.890)
<i>CXCL3</i>	0.008	0.644 (0.465–0.893)
<i>IL18RAP</i>	0.008	0.638 (0.458–0.891)
<i>CD74</i>	0.009	0.645 (0.465–0.895)
<i>WLS</i>	0.009	0.646 (0.466–0.895)
<i>CCR7</i>	0.009	0.644 (0.463–0.895)
<i>CST7</i>	0.009	0.645 (0.464–0.897)
<i>CAB39</i>	0.009	1.561 (1.117–2.181)
<i>PLAC8</i>	0.009	0.647 (0.466–0.898)
<i>SELL</i>	0.009	0.648 (0.468–0.899)
<i>CCL19</i>	0.010	0.652 (0.470–0.903)
<i>BIRC3</i>	0.011	0.653 (0.471–0.906)
<i>NFKBIE</i>	0.011	0.653 (0.470–0.906)
<i>CAMK4</i>	0.011	0.651 (0.468–0.906)
<i>CXCL2</i>	0.011	0.655 (0.473–0.908)
<i>HLA-DMA</i>	0.011	0.655 (0.471–0.909)
<i>CXCL9</i>	0.012	0.654 (0.470–0.909)
<i>SOCS1</i>	0.013	0.659 (0.475–0.914)
<i>PDLIM4</i>	0.013	0.661 (0.477–0.916)
<i>IL7</i>	0.013	0.660 (0.475–0.916)
<i>CD3G</i>	0.013	0.659 (0.473–0.917)
<i>EMP3</i>	0.013	0.661 (0.476–0.918)
<i>IRF1</i>	0.014	0.665 (0.480–0.922)
<i>ERRFI1</i>	0.015	0.667 (0.482–0.924)
<i>GBP2</i>	0.016	0.668 (0.481–0.927)
<i>IL2RB</i>	0.016	0.667 (0.480–0.928)
<i>WAS</i>	0.017	0.668 (0.481–0.929)
<i>TP63</i>	0.017	0.672 (0.486–0.931)
<i>SFRP1</i>	0.017	0.673 (0.487–0.931)
<i>IL2RG</i>	0.017	0.670 (0.482–0.930)
<i>ITK</i>	0.019	0.674 (0.485–0.937)
<i>LTF</i>	0.019	0.678 (0.489–0.938)
<i>CD8A</i>	0.020	0.676 (0.486–0.941)
<i>C1S</i>	0.022	0.680 (0.489–0.945)
<i>DST</i>	0.022	0.682 (0.492–0.947)
<i>IKZF1</i>	0.022	0.680 (0.488–0.947)
<i>DENND5A</i>	0.022	0.684 (0.494–0.948)
<i>PRF1</i>	0.023	0.680 (0.488–0.948)
<i>GPR18</i>	0.023	0.685 (0.494–0.949)
<i>CCNL1</i>	0.023	0.685 (0.495–0.950)
<i>CLCA2</i>	0.024	0.685 (0.494–0.951)
<i>NT5E</i>	0.024	1.457 (1.051–2.019)
<i>TACR1</i>	0.024	0.689 (0.498–0.953)
<i>OSMR</i>	0.026	0.690 (0.499–0.956)
<i>DGKA</i>	0.026	0.690 (0.497–0.957)

Table S3 (continued)

Table S3 (continued)

Genes	P	HR (95% CI)
<i>IL18</i>	0.027	0.691 (0.498–0.958)
<i>CD83</i>	0.027	0.691 (0.498–0.959)
<i>L3MBTL4</i>	0.028	0.693 (0.500–0.961)
<i>LTA4H</i>	0.028	0.695 (0.502–0.962)
<i>PIK3CD</i>	0.030	0.696 (0.502–0.965)
<i>ISG20</i>	0.030	0.695 (0.500–0.965)
<i>EOMES</i>	0.030	0.693 (0.498–0.966)
<i>IL7R</i>	0.031	0.695 (0.499–0.967)
<i>SIT1</i>	0.032	0.698 (0.503–0.970)
<i>GSN</i>	0.032	0.696 (0.499–0.970)
<i>ITGB4</i>	0.033	0.700 (0.504–0.971)
<i>ZAP70</i>	0.033	0.700 (0.504–0.972)
<i>FBLN5</i>	0.034	0.702 (0.507–0.973)
<i>HLA-DRB1</i>	0.034	0.699 (0.503–0.973)
<i>PTAFR</i>	0.034	0.700 (0.503–0.974)
<i>CD37</i>	0.035	0.700 (0.503–0.975)
<i>IL16</i>	0.035	0.701 (0.503–0.976)
<i>NCF4</i>	0.036	0.704 (0.507–0.977)
<i>CXCR3</i>	0.036	0.704 (0.507–0.978)
<i>EIF4E3</i>	0.037	0.709 (0.512–0.980)
<i>GP1BA</i>	0.037	0.703 (0.504–0.979)
<i>CLDN11</i>	0.039	0.710 (0.513–0.982)
<i>STAT5A</i>	0.039	0.709 (0.512–0.982)
<i>XCL1</i>	0.039	0.709 (0.512–0.982)
<i>MAFF</i>	0.039	0.711 (0.514–0.983)
<i>IFNAR1</i>	0.039	1.412 (1.017–1.959)
<i>MKKN1</i>	0.040	0.708 (0.509–0.984)
<i>CCL5</i>	0.040	0.710 (0.512–0.985)
<i>NEDD9</i>	0.041	0.712 (0.514–0.986)
<i>LAMP3</i>	0.042	0.714 (0.516–0.988)
<i>CCR5</i>	0.042	0.710 (0.511–0.988)
<i>AHR</i>	0.042	0.714 (0.516–0.989)
<i>GNAI2</i>	0.043	0.714 (0.515–0.989)
<i>ITGAL</i>	0.043	0.713 (0.514–0.990)
<i>BCL10</i>	0.044	0.716 (0.518–0.990)
<i>PLPP3</i>	0.044	0.714 (0.515–0.991)
<i>ICAM1</i>	0.044	0.715 (0.516–0.991)
<i>CTSO</i>	0.046	0.715 (0.514–0.994)
<i>GPR183</i>	0.046	0.714 (0.513–0.995)
<i>LCP1</i>	0.047	0.720 (0.520–0.996)
<i>KLRK1</i>	0.048	0.718 (0.518–0.996)
<i>IL10</i>	0.048	1.394 (1.002–1.938)
<i>LAT2</i>	0.049	0.718 (0.516–0.999)
<i>GZMB</i>	0.049	0.719 (0.517–0.999)

cDC, conventional dendritic cell; HR, hazard ratio; CI, confidence interval.

Table S4 Spearman's correlation analysis between IL-2 and cells

Cell type	P value*	R
CD8 ⁺ Tcm	9.65E-106	0.599
aDC	9.75E-105	0.596
DC	4.04E-78	0.527
cDC	8.65E-78	0.526
CD8 ⁺ T-cells	3.29E-66	0.490
CD4 ⁺ memory T-cells	1.49E-44	0.408
CD4 ⁺ Tem	3.09E-42	0.398
B-cells	5.24E-37	0.373
Megakaryocytes	1.99E-35	0.366
CD4 ⁺ naive T-cells	5.46E-35	0.363
CD8 ⁺ naive T-cells	3.13E-29	0.332
Class-switched memory B-cells	5.30E-29	0.331
Stroma score	3.20E-27	0.321
Tregs	4.75E-25	0.308
Macrophages M1	7.49E-25	0.306
Fibroblasts	3.28E-22	0.289
iDC	2.70E-20	0.276
Adipocytes	2.95E-20	0.276
CD4 ⁺ T-cells	3.05E-20	0.276
Mesangial cells	5.71E-20	0.274
Endothelial cells	1.65E-18	0.263
Memory B-cells	3.04E-18	0.261
pDC	1.94E-17	0.255
HSC	2.62E-17	0.254
ly Endothelial cells	3.28E-17	0.253
mv Endothelial cells	5.23E-16	0.243
Macrophages	6.96E-15	0.234
CD8 ⁺ Tem	7.64E-15	0.234
Monocytes	9.99E-12	0.205
Tgd cells	1.26E-10	0.194
Macrophages M2	1.75E-10	0.193
naive B-cells	3.55E-10	0.190
Chondrocytes	8.00E-10	0.186
Plasma cells	7.14E-09	0.175
Astrocytes	1.07E-07	0.161
Basophils	6.28E-05	0.122
Th2 cells	7.39E-05	0.120
Keratinocytes	3.79E-04	0.108
Sebocytes	4.53E-04	0.107
Melanocytes	7.04E-04	0.103
GMP	1.14E-02	0.077
MEP	1.31E-03	-0.098
MSC	1.68E-04	-0.114
Th1 cells	1.11E-05	-0.133

Table S4 (continued)**Table S4** (continued)

Cell type	P value*	R
Smooth muscle	1.38E-11	-0.204
CLP	1.22E-16	-0.249
Myocytes	5.38E-23	-0.294
Neurons	1.96E-25	-0.310
Osteoblast	5.08E-41	-0.392

*, only cells with a P value <0.05 are listed. aDC, activated dendritic cell; cDC, conventional dendritic cell; CLP, common lymphoid progenitor; DC, dendritic cell; GMP, granulocyte-macrophage progenitor; HSC, hematopoietic stem cell; iDC, immature dendritic cell; MEP, megakaryocyte-erythroid progenitor; MSC, mesenchymal stem cell; pDC, plasmacytoid dendritic cell.