

RESEARCH

Open Access



Comprehensive analysis of the cancer driver genes in breast cancer demonstrates their roles in cancer prognosis and tumor microenvironment

Xiao-wei Du^{1†}, Gao Li^{2†}, Juan Liu³, Chun-yan Zhang⁴, Qiong Liu⁵, Hao Wang^{2*} and Ting-song Chen^{2*}

Abstract

Background: Breast cancer is the most common malignancy in women. Cancer driver gene-mediated alterations in the tumor microenvironment are critical factors affecting the biological behavior of breast cancer. The purpose of this study was to identify the expression characteristics and prognostic value of cancer driver genes in breast cancer.

Methods: The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) datasets are used as the training and test sets. Classified according to cancer and paracancerous tissues, we identified differentially expressed cancer driver genes. We further screened prognosis-associated genes, and candidate genes were submitted for the construction of a risk signature. Functional enrichment analysis and transcriptional regulatory networks were performed to search for possible mechanisms by which cancer driver genes affect breast cancer prognosis.

Results: We identified more than 200 differentially expressed driver genes and 27 prognosis-related genes. High-risk group patients had a lower survival rate compared to the low-risk group ($P < 0.05$), and risk signature showed high specificity and sensitivity in predicting the patient prognosis (AUC 0.790). Multivariate regression analysis suggested that risk scores can independently predict patient prognosis. Further, we found differences in PD-1 expression, immune score, and stromal score among different risk groups.

Conclusion: Our study confirms the critical prognosis role of cancer driver genes in breast cancer. The cancer driver gene risk signature may provide a novel biomarker for clinical treatment strategy and survival prediction of breast cancer.

Keywords: Cancer driver genes, Prognostic signature, Breast cancer, Tumor microenvironment, TCGA

* Correspondence: Howard_123@163.com; cts552052597@163.com

[†]Xiao-wei Du and Gao Li contributed equally to this work.

²Department of Oncology, Seventh People's Hospital of Shanghai University of Traditional Chinese Medicine, Shanghai 200000, China

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

In the past few years, the incidence of breast cancer has been on the rise. Among all malignant tumors in women, breast cancer ranks first with an incidence rate of 30% [1]. Women with breast cancer have the second-highest mortality rate, and the prognosis of patients with different molecular types varies significantly [2–4]. For example, the choice of drugs varies among patients with different hormone receptor status, and patients with triple-negative breast cancer usually have a poorer prognosis [5–7]. Besides, traditional tumor staging sometimes cannot objectively represent the tumor status of patients. Patients with the same tumor stage sometimes have significant survival differences [8]. The heterogeneity of breast cancer poses a challenge for clinical treatment strategy choice and prognosis prediction, while studies of different molecular typing are expected to elucidate the differences in heterogeneity and better guide clinical treatment of breast cancer.

The journal of *Nature Reviews Cancer* recently reported on the important role played by cancer driver genes in the progression of tumor malignancy [9]. 568 cancer driver genes were identified through a large-scale transcriptome analysis, which may mediate complex molecular regulatory networks and changes in the tumor microenvironment. Aberrantly expressed cancer driver genes may result in multiple processes such as uncontrolled tumor cell proliferation, invasion, recurrence, and drug resistance [10, 11]. Zhang's study reported the frequency of mutations in the whole genes of breast cancer [12]. The most frequently mutated genes were TP53 (45%), followed by PIK3CA (44%), GATA3 (18%), and MAP3K1 (10%), all of which also were identified as cancer driver genes. Kruse's study [13] identified metastasis driver genes in breast cancer by massive parallel sequencing, and the cancer driver genes DCC and CREBBP were also identified as key metastatic genes.

Tumorigenesis is often associated with alterations in the stromal environment and immune status, mainly manifested as changes in the tumor microenvironment (TME) [14, 15]. TME plays a key role in several steps of tumor development, including local drug resistance, immune escape, recurrence, and distant metastasis [16, 17]. Tumor cell exposure to TME also contributes to shaping the tissue-specificity of driver genes [18]. Antibodies or inhibitors that target driver gene-mediated signaling pathways may effectively inhibit tumor growth and prolong patient survival [19]. KRAS mutations may affect the TME and patient response to immunotherapy [20]. Cancer driver genes and TME determine the type of liver cancer and can be considered as predictors of patient survival outcomes [21, 22]. All these studies provide a reference for clinical decisions. However, there are fewer studies on cancer driver genes affecting the TME and prognosis of breast cancer.

In this study, we systematically analyzed the expression characteristics of 568 cancer driver genes in breast cancer. We screened for risk driver molecules that affect breast cancer prognosis, and the relationship between risk groups with the tumor microenvironment. Our study is expected to find novel molecular subtypes of breast cancer for better guiding clinical treatment strategies.

Materials and methods

Data sources

Breast cancer transcriptome data and corresponding clinical data were downloaded from The Cancer Genome Atlas (TCGA) (<https://portal.gdc.cancer.gov/>) database. As a training set, TCGA transcriptome data contained a total of 112 cases of paracancerous tissues and 1096 cases of cancerous tissues, and the clinical information of patients is shown in Table 1.

The Gene Expression Omnibus (GEO) (<https://www.ncbi.nlm.nih.gov/geo/>) database serves as an external validation of the TCGA dataset and contains two subsets: GSE7390 and GSE42568. GSE42568 dataset contained 17 cases of paracancerous tissues and 104 cases of cancerous tissues. GSE7390 dataset only contained 198 cases of paracancerous tissues. Clinical information of the GEO test set is shown in Table 2. The RNA

Table 1 Clinical features of BRCA patients ($n = 1096$) from TCGA database

Variables	N	%
Total	1096	100.0%
Age, years		
<60	582	53.1%
≥60	514	46.9%
Stage		
Stage i	183	16.7%
Stage ii	621	56.7%
Stage iii	248	22.6%
Stage iv	20	1.8%
T classification		
T1	281	25.6%
T2	635	57.9%
T3	138	12.6%
T4	39	3.6%
ER status		
Negative	796	72.6%
Positive	233	21.3%
PR status		
Negative	688	62.8%
Positive	338	30.8%

Table 2 Clinical features of BRCA patients ($n = 302$) from GEO database

Variables	GSE7390		GSE42568	
	N	%	N	%
Total	198	100.0%	104	100.0%
Age, years				
<60	195	98.5%	59	56.7%
≥60	3	1.5%	45	43.3%
Grade				
G1	30	15.2%	11	10.5%
G2	83	41.9%	40	38.5%
G3	83	41.9%	53	51.0%
ER status				
Negative	64	32.3%	34	32.7%
Positive	134	67.7%	67	64.4%

transcriptome data from different platforms were normalized using the “limma” and “sva” R packages.

Expression of cancer driver genes in breast cancer tissues

The list of 568 cancer driver genes was obtained from the Integrative OncoGenomics platform (<https://www.intogen.org/search>). The differentially expressed genes (DEGs) were analyzed according to the classification of cancerous and paracancerous tissues by using the “limma” R package ($P < 0.05$) in GSE42568 and TCG datasets. Venn diagrams were drawn to summarize differentially expressed genes both in TCGA and GEO datasets. These DEGs are considered to be associated with tumor progression and are used for further prognostic molecular screening.

Construction and validation of the risk signature

We used the univariate Cox and Kaplan-Meier method to screen for cancer driver genes associated with overall survival (OS) ($P < 0.05$) in breast cancer patients. Candidate genes that were both differentially expressed and associated with OS were substituted into least absolute shrinkage and selection operator (Lasso) Cox regression and stepwise multivariate Cox proportional regression to construct the risk signature. According to the risk driver genes expression of the prognostic signature, R package Stats was used for Principal Component Analysis (PCA). PCA could confirm the clustering ability of the cancer driver genes risk signature.

We used several methods to analyze the clinical value of the risk signature. Survival curves and receiver operating characteristic (ROC) curves were used to verify the prognostic value of different risk groups. Univariate and multifactorial regression analyses were performed to explore the independent prognostic role of risk scores.

Besides, we further analyzed whether age, tumor stage, and other clinicopathology were associated with risk scores.

Construction of transcription factor regulatory networks and functional enrichment analysis

To clarify the possible mechanisms by which cancer driver genes affect the progression and prognosis of breast cancer, we constructed a transcription factor regulatory network and gene set enrichment analysis (GSEA). The cancer-associated transcription factors were obtained from the Cistrome platform (<http://cistrome.org/>). We screened for cancer driver genes associated with transcription factors ($|R^2| > 0.3$ and $P < 0.05$) and constructed a regulatory network using Cytoscape software (<https://cytoscape.org/>). Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were performed with the “enrichplot” R package, which suggests possible functional and pathway sets of cancer driver genes affecting breast cancer prognosis.

Correlation analysis of tumor microenvironment

The immune and TME scores were calculated for each sample using single-sample GSEA (ssGSEA) algorithm by the “gsva” R package. We first analyzed the expression levels of PD-1, stromal score, immune score, and estimate score in different risk groups. We likewise analyzed the correlation of risk groups with 16 immune cell scores and 13 immune-related pathway scores. These results contribute to confirm the interaction of tumor microenvironment with cancer driver genes.

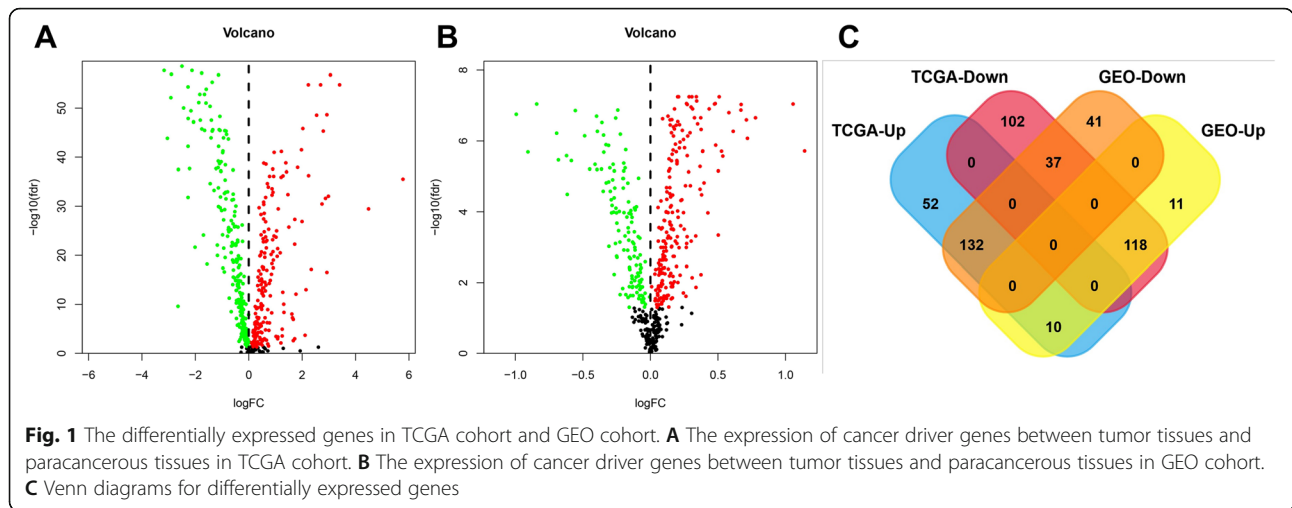
Statistical analysis

All data processing was performed on R v3.4.1 (<https://www.r-project.org/>). Differences in gene expression between cancerous and paracancerous tissues were analyzed by the Wilcoxon method. The correlation between transcription factors and cancer driver gene expression was performed using the Pearson’s correlation coefficient method. Kaplan-Meier method verifies the impact of the candidate driver genes on patient OS. Mann-Whitney U test was used to compare immune or TME scores between the high-risk and low-risk groups. Univariate and multifactorial regression were used to analyze the prognostic value of risk scores.

Results

Differentially expressed cancer driver genes in TCGA and GEO datasets

In the TCGA cohort, the expression of 194 cancer driver genes was higher in cancer tissues than in paracancerous tissues ($P < 0.05$), and 257 cancer driver genes were lower in cancer tissues than in paracancerous tissues ($P < 0.05$).



(Fig. 1A). The results of the GEO cohort were consistent with the TCGA cohort, with 210 genes highly expressed and 139 genes lowly expressed in cancer tissues ($P < 0.05$) (Fig. 1B, C). Most cancer driver genes are specifically expressed in cancerous tissues, which suggest that driver genes may be involved in multiple biological processes in breast cancer.

Construction of a thirteen-mRNA signature for predicting patient prognosis

Through univariate Cox and Kaplan-Meier analysis ($P < 0.05$), we screened 27 cancer driver genes associated with OS in breast cancer patients (Fig. 2A). Substituting these

candidate molecules in lasso Cox regression and stepwise multivariate Cox proportional regression, we finally constructed a thirteen-mRNA risk signature (Fig. 2B, C). The coefficients for each risk genes are shown in Table 3. The risk score for each patient can be calculated by the risk formula, risk score = $e^{(\text{expression of BRD4}^* (-0.072) + \text{expression of BRD7}^* 0.049 + \text{expression of BTG1}^* (-0.014) + \text{expression of CCR7}^* 0.023 + \text{expression of DAXX}^* (-0.026) + \text{expression of DDX3X}^* 0.024 + \text{expression of EGR2}^* (-0.024) + \text{expression of FLT3}^* (-0.096) + \text{expression of IKZF3}^* (-0.089) + \text{expression of JAK1}^* (-0.020) + \text{expression of MAX}^* (-0.034) + \text{expression of NFKBIA}^* (-0.009) + \text{expression of UBE2A}^* 0.051)}$. According to the median values of the risk scores in the TCGA cohort, all patients were divided into high-risk and low-risk groups. PCA

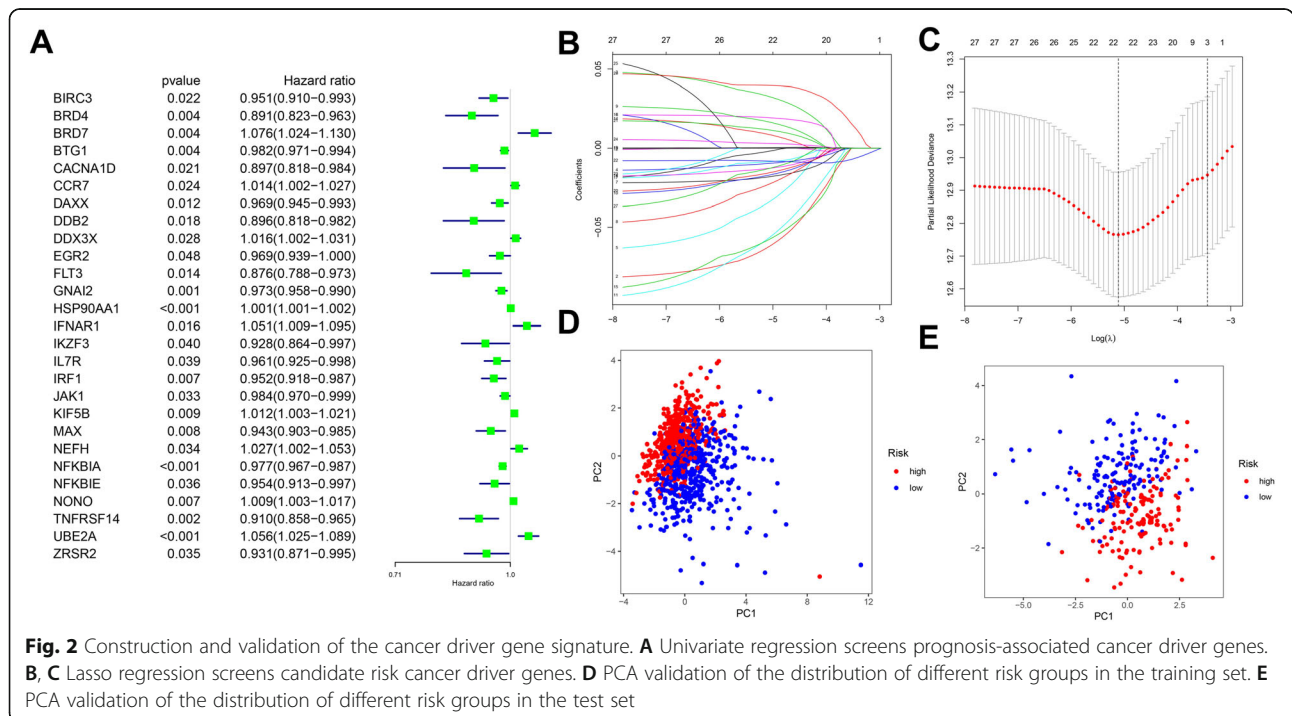


Table 3 Each gene of the thirteen-mRNA signature

mRNA	Coefficient	HR	95%CI	P
BRD4	-0.072	0.931	0.870–0.996	0.037
BRD7	0.049	1.050	0.989–1.116	0.109
BTG1	-0.014	0.986	0.974–0.998	0.019
CCR7	0.023	1.023	1.015–1.031	<0.001
DAXX	-0.026	0.975	0.950–0.999	0.049
DDX3X	0.024	1.024	1.007–1.041	0.007
EGR2	-0.024	0.976	0.948–1.006	0.115
FLT3	-0.096	0.909	0.824–1.002	0.054
IKZF3	-0.089	0.915	0.851–0.984	0.017
JAK1	-0.020	0.981	0.962–0.999	0.042
MAX	-0.034	0.967	0.925–1.010	0.132
NFKBIA	-0.009	0.991	0.980–1.002	0.092
UBE2A	0.051	1.052	1.012–1.094	0.010

validates that the different risk groups showed a two-way distribution, indicating the high specificity of the risk signature (Fig. 2D, E).

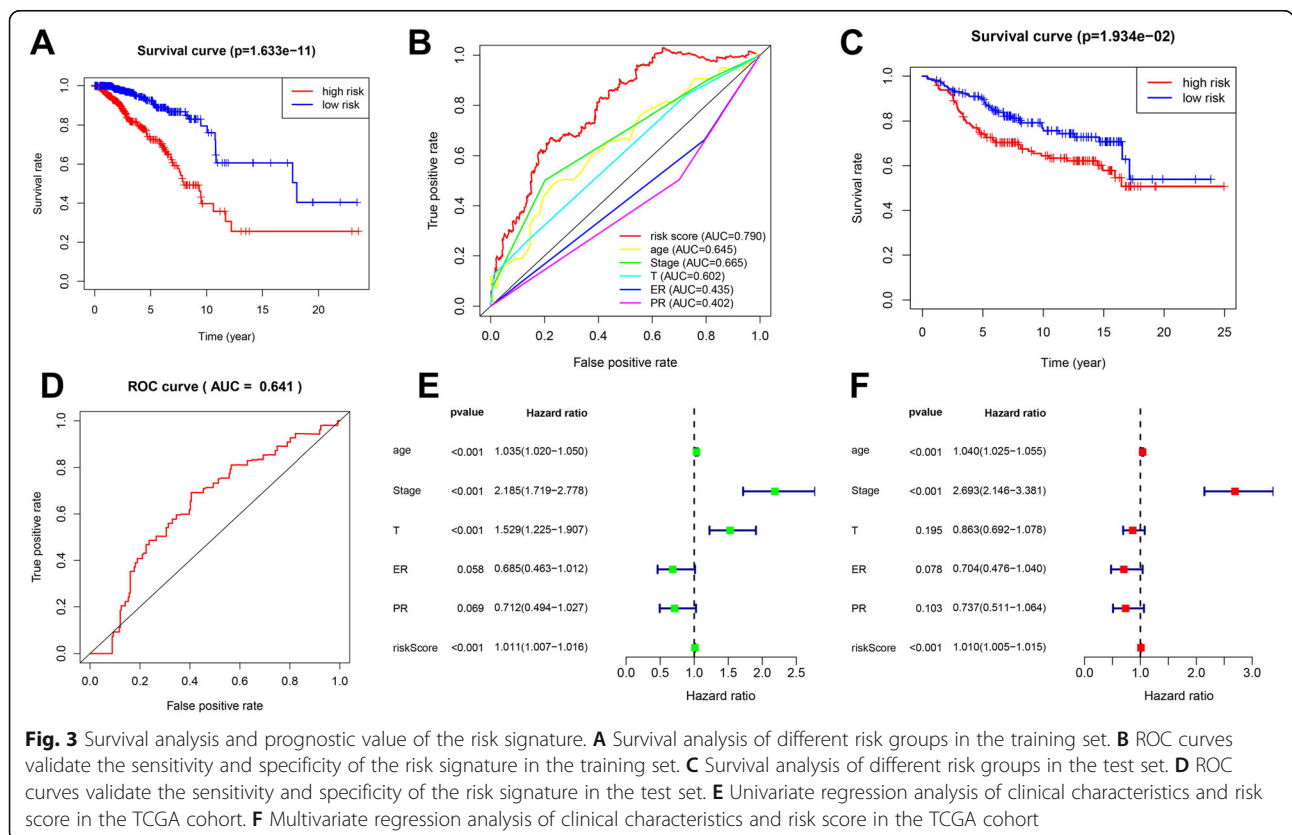
Survival curves were used to analyze the relationship between risk groups and OS of breast cancer patients. In the training set, the overall survival rate in the high-risk group was lower than in the low-risk group ($P < 0.05$) (Fig. 3A). The area under curve (AUC) of the risk score

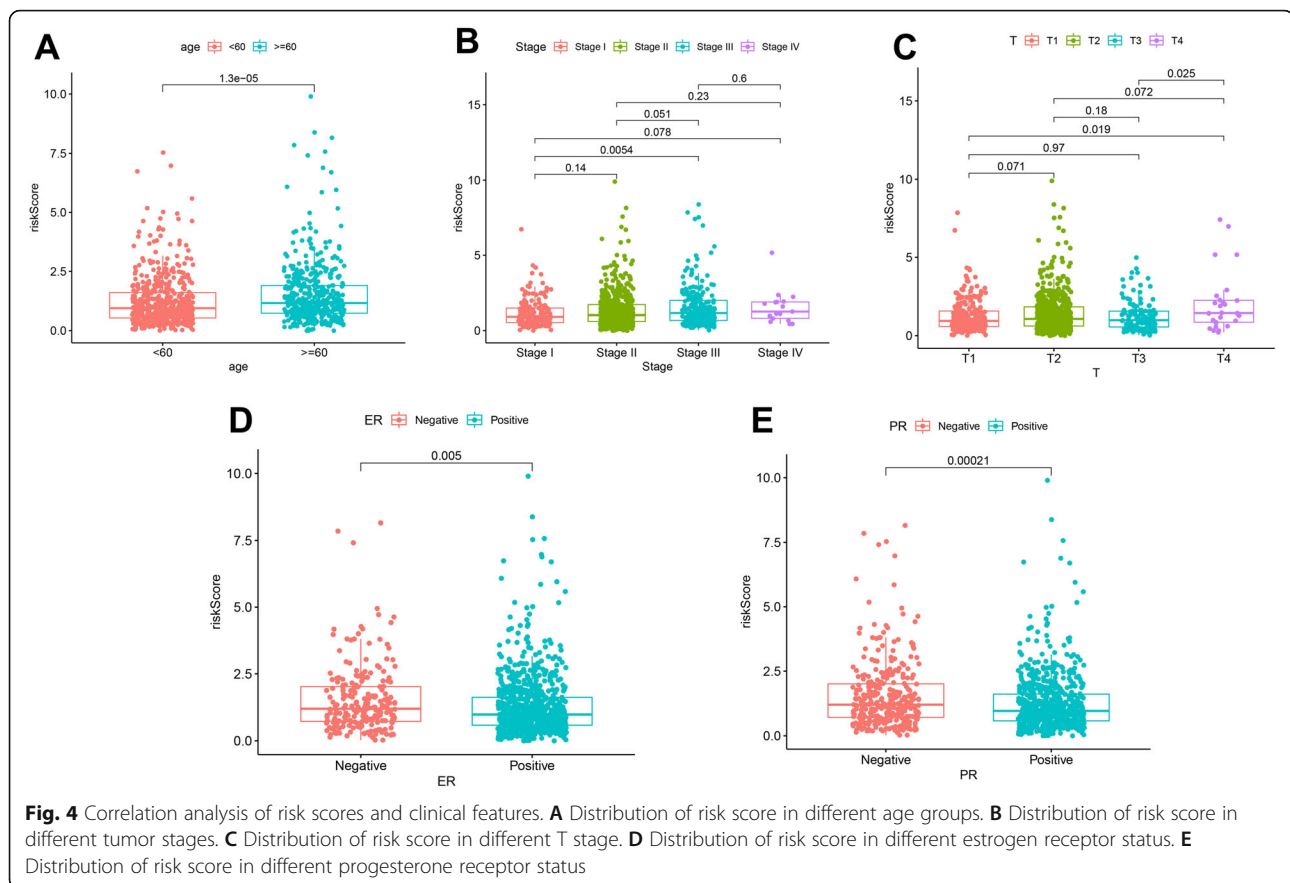
(AUC 0.790) was higher than other clinical features, suggesting the high precision of the risk signature for predicting patient survival (Fig. 3B). The results of the test set were consistent with the training set. The survival of patients in different risk groups differed significantly ($P < 0.05$), and the risk score could predict the survival of breast cancer patients well (AUC 0.641) (Fig. 3C, D). Univariate and multifactorial regression analyses suggested that risk score could be an independent risk factor affecting the prognosis of breast cancer patients ($P < 0.05$) (Fig. 3E, F).

We analyzed the level of risk scores in different clinical characteristics separately. The results suggest that people older than 60 years, advanced tumor stage, and estrogen and progesterone receptor-negative patients have higher risk scores ($P < 0.05$), which may represent a poorer prognosis (Fig. 4). This group belongs to the high-risk group of the risk signature and requires more attention.

Transcription factor regulatory networks and functional enrichment analysis

To explore the possible mechanisms by which cancer driver genes affect breast cancer progression and prognosis, we first tried to explore the co-expression network of transcription factors and cancer driver genes. Through Pearson’s correlation coefficient analysis ($|R^2| > 0.3$ and $P < 0.05$), 206 transcription factors were associated with





cancer driver gene expression, and DDX3X, JAK1, EGR2, IKZF3, and CCR7 were the five most enriched cancer driver genes (Fig. 5A). A series of twelve cancer driver gene expressions were identified to be associated with transcription factors, of which CCR7, BRD7, DDX3X, and UBE2A were upregulated in breast cancer tissues, and the others were downregulated in breast cancer tissues. Next, all molecules were substituted into Cytoscape to map the regulatory network (Fig. 5B).

GO and KEGG analysis suggest possible mechanisms by which cancer driver genes impact breast cancer prognosis. GO analysis showed that the functional sets of immune response, T cell activity, and immune-related receptors were significantly enriched. KEGG analysis suggests that B cell receptor signaling pathway, T cell receptor signaling pathway, and PD-1 checkpoint pathway were significantly enriched. These results suggest a strong association between cancer driver genes and the immune status, which supports our further studies on the risk signature and tumor microenvironment.

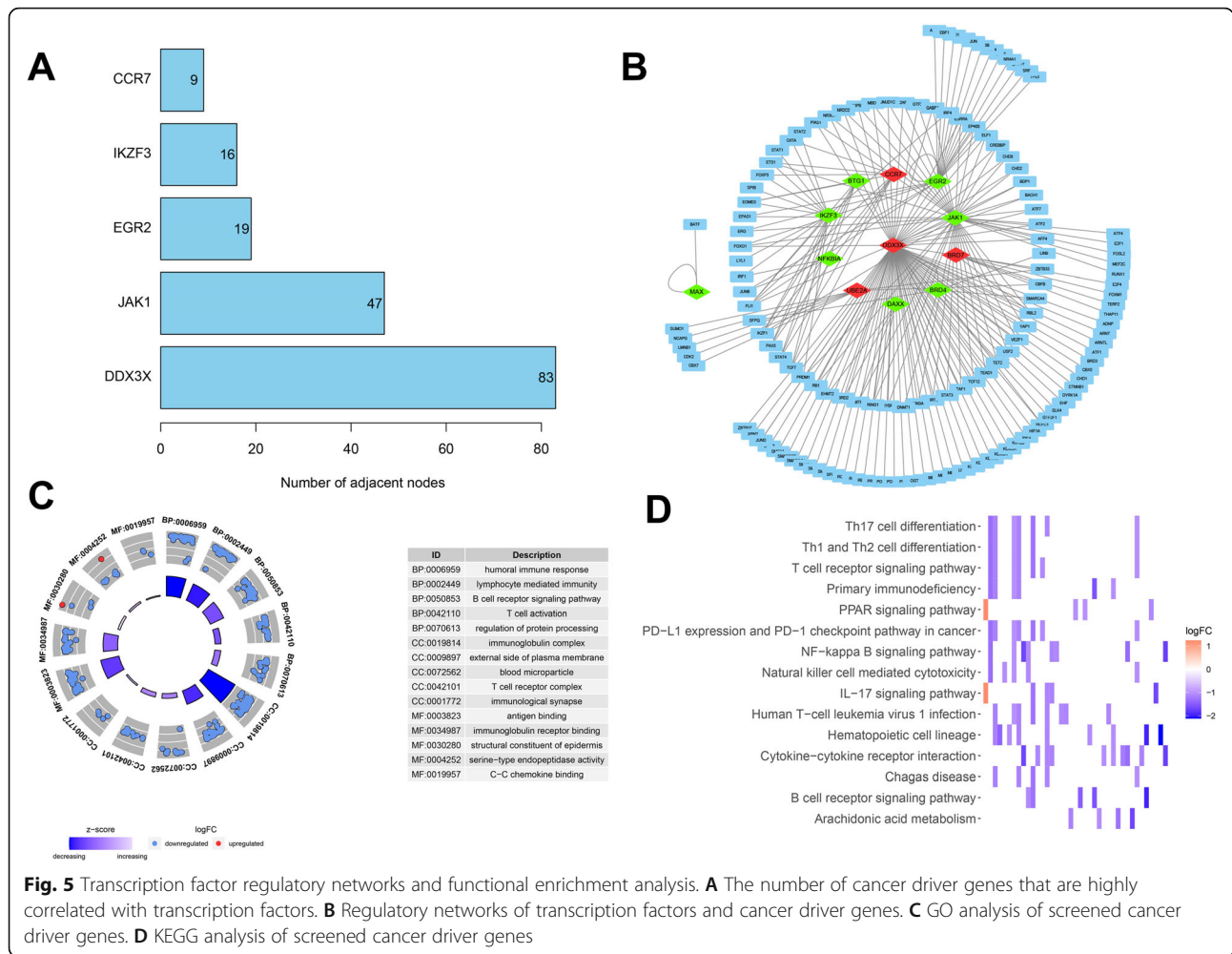
Correlation of the risk signature with tumor microenvironment

Alterations in the tumor microenvironment may be an essential way in which cancer driver genes affect breast

cancer. We analyzed the correlation between PD-1 levels, tumor microenvironment, and risk groups. The results suggested that PD-1 expression levels, stromal scores, and immune scores were higher in the low-risk group compared to the high-risk group ($P < 0.05$) (Fig. 6A-D). We likewise analyzed the correlation of immune cells and immune pathways with the risk signature, and the results suggested that most immune function scores differed significantly in the risk groups ($P < 0.05$) (Fig. 6E, F).

Discussion

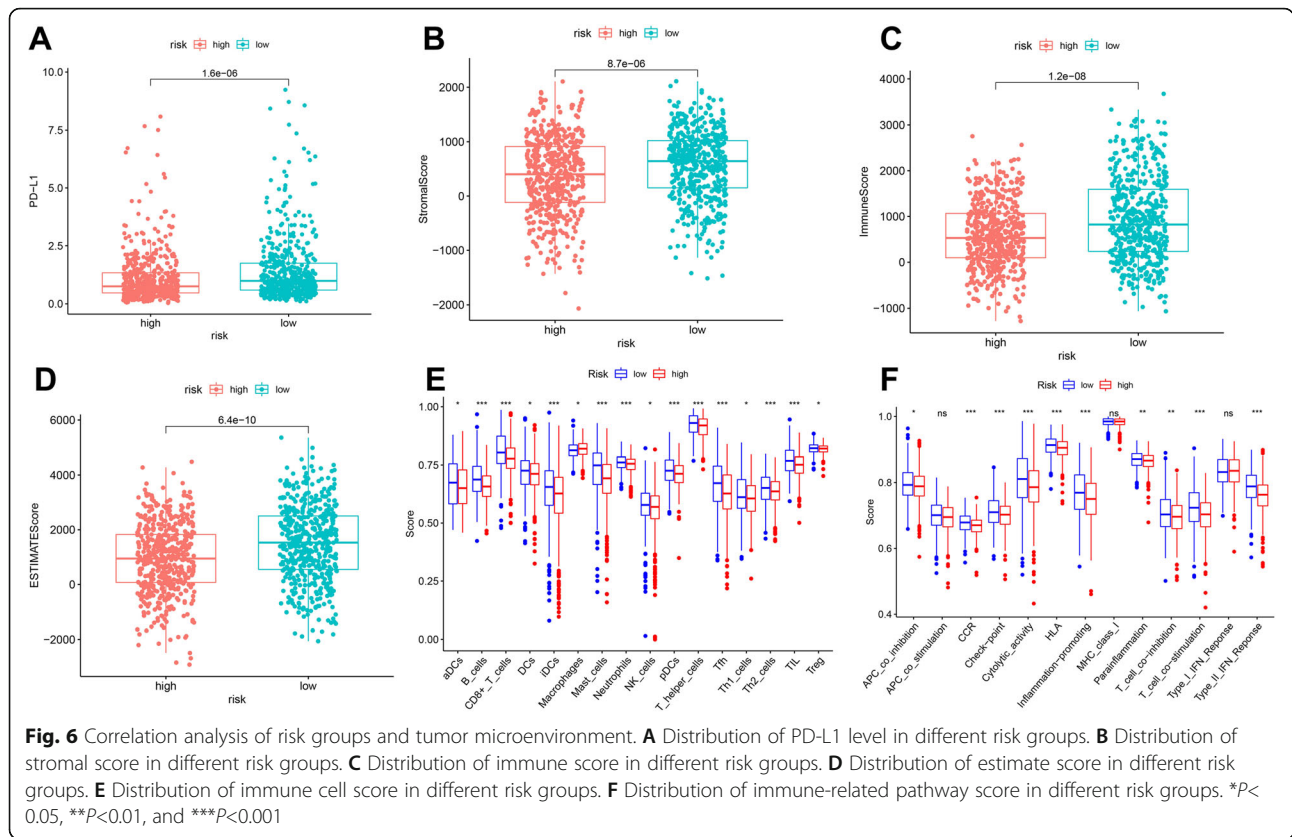
In the precision medicine era, the mutation and expression characteristics of breast cancer patients' tumor genomes are playing an increasingly critical role in the choice of therapeutic strategies [23]. Tumor heterogeneity is the main reason for the different responses to the treatment of breast cancer patients [24, 25]. Different molecular typing is used to explain breast cancer heterogeneity. For example, estrogen and progesterone receptor status can indicate whether a patient is suitable for endocrine therapy [26]. In contrast, traditional tumor markers are gradually showing their limitations, while novel markers at the molecular level, such as mRNA, miRNA, and DNA methylation, are being intensively investigated [27–30]. Tumor-infiltrating lymphocytes (TIL) are a heterogeneous



population of lymphocytes [31]. Triple-negative breast cancers have the highest degree of TIL infiltration, followed by HER2+ breast cancers [32]. The KEYNOTE-086 study found that abundant mesenchymal TIL was closely associated with better efficacy of pembrolizumab [33]. The combined consideration of PD-L1 levels in tumor cells and immune cells predicts the possible benefit of receiving immunotherapy in breast cancer patients [34]. Molecular targeted therapy, immunotherapy, and novel molecular biomarkers have broadened the treatment opportunities for breast cancer, improving patients' quality of life and prolonging survival time.

The specific expression of cancer driver genes is a major contributor to the formation of the tumor micro-environment, which further leads to various biological processes such as uncontrolled tumor cell proliferation, invasion, metastasis, recurrence, and drug resistance [21, 35]. Martínez's study analyzed the gene expression of over 20,000 tumor samples and mapped cancer driver gene profiles [9]. Algorithms and identification of cancer

driver genes have been intensively studied [36, 37]; however, the clinical value of newly identified cancer driver genes has been relatively under-explored. Aberrant expression and mutations of cancer driver genes are very common in breast cancer patients, approaching 50% [12]. Comprehensive analysis of the clinical significance of these genes may lead to the discovery of new clinical biomarkers. We studied cancer driver gene expression profiles of more than 1000 breast cancer samples and identified cancer driver genes associated with breast cancer progression and prognosis. For example, Rizeq's study found that activation of the C-C chemokine receptor 7 (CCR7)-related complex increased tumor cell proliferation and migration [38]. Vahidi's study concluded that CD8-positive memory T cells in tumor-draining lymph nodes are associated with CCR7 [39]. We found the specific expression and prognostic value of CCR7 in breast cancer. Both Niu [40] and our study confirmed the oncogenic role of bromodomain-containing 7 (BRD7) molecule in breast cancer. In



general, these studies suggest the involvement of cancer driver genes in complex tumor biological behavior and clinically important roles. Based on the identification of the cancer driver genes, we further constructed a thirteen-mRNA prognostic signature with high specificity and sensitivity of prediction. The survival of patients in different risk groups differed significantly, and the risk score was an independent risk factor in predicting breast cancer prognosis.

Several clinical trials have demonstrated favorable results of PD-1/PD-L1 antibodies in the treatment of breast cancer [34, 41, 42]. Targeting PD-1-related pathways effectively blocks immune surveillance and escape and activates the immune response of T cells against tumor cells [43, 44]. The stromal environment, immune cell level, and composition are crucial in the immunotherapy process. Cancer driver genes can indirectly alter the tumor microenvironment through multiple regulatory networks [21], but fewer studies have been conducted. Raskov's study [45] found that mutations in some cancer driver genes could increase accessibility for DNA targeting chemotherapeutics and reduce cytotoxic drug resistance in colorectal cancer. In our study, KEGG analysis revealed significant enrichment of immune-related pathways such as T cell activation and PD-1-

related pathways, suggesting possible interactions between cancer driver genes and immunity. We found that the survival time was longer in the low-risk group compared to the high-risk group, who also had higher PD-1 expression, immune-related scores, and tumor microenvironment scores. Our findings suggest that cancer driver genes may be involved in the body's immune response and have an enhancing effect on the effectiveness of immunotherapy.

Conclusion

In summary, we analyzed the expression profiles of 568 cancer driver genes in breast cancer and identified driver genes associated with breast cancer prognosis. We further constructed a risk signature to predict breast cancer prognosis, and both the training and external test sets performed sensitive predictive efficacy. Cancer driver genes may affect the biological behavior of breast cancer by mediating transcriptional processes and alterations in the immune and tumor microenvironment. Our study confirms the critical role of cancer driver genes in breast cancer, and it is expected to supply a reference for the prognostic stratification and treatment strategy of breast cancer.

Acknowledgements

We thank our anonymous reviewers for their valuable comments on this manuscript, which have led to much many improvements to the article.

Authors' contributions

XWD and GL are responsible for writing and submitting the papers; JL is responsible for the data collection and analysis; CYZ and QL are responsible for the production of pictures; TSC and HW are responsible for the ideas and guidance. The authors read and approved the final manuscript.

Funding

This study was supported by two funding from the Shanghai Municipal Health Commission (No. 202040180, No. 202040187).

Availability of data and materials

All data was obtained from The Cancer Genome Atlas (<https://portal.gdc.cancer.gov/>) and Gene Expression Omnibus (<https://www.ncbi.nlm.nih.gov/geo/>) databases.

Declarations**Ethics approval and consent to participate**

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Jinzhou Medical University, Jinzhou, China. ²Department of Oncology, Seventh People's Hospital of Shanghai University of Traditional Chinese Medicine, Shanghai 200000, China. ³Department of Pharmacy, Seventh People's Hospital of Shanghai University of Traditional Chinese Medicine, Shanghai, China. ⁴Department of Central Laboratory, Seventh People's Hospital of Shanghai University of Traditional Chinese Medicine, Shanghai, China. ⁵Office of Academic Research, Seventh People's Hospital of Shanghai University of Traditional Chinese Medicine, Shanghai, China.

Received: 1 July 2021 Accepted: 31 August 2021

Published online: 10 September 2021

References

- Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2021. *CA Cancer J Clin.* 2021;71(1):7–33. <https://doi.org/10.3322/caac.21654>.
- Guo Y, Mao X, Qiao Z, Chen B, Jin F. A novel promoter CpG-based signature for long-term survival prediction of breast cancer patients. *Front Oncol.* 2020;10:579692. <https://doi.org/10.3389/fonc.2020.579692>.
- Liang F, Qu H, Lin Q, Yang Y, Ruan X, Zhang B, et al. Molecular biomarkers screened by next-generation RNA sequencing for non-sentinel lymph node status prediction in breast cancer patients with metastatic sentinel lymph nodes. *World J Surg Oncol.* 2015;13(1):258. <https://doi.org/10.1186/s12957-015-0642-z>.
- Prejac J, Dedić Plavetić N, Gotovac Jerčić K, Borovečki F. A first report of a rare TP53 variant associated with Li-Fraumeni syndrome manifesting as invasive breast cancer and malignant solitary fibrous tumor. *World J Surg Oncol.* 2021;19(1):254. <https://doi.org/10.1186/s12957-021-02370-8>.
- Dieci MV, Miglietta F, Guarneri V. Immune infiltrates in breast cancer: recent updates and clinical implications. *Cells.* 2021;10(2):223.
- Jia R, Li Z, Liang W, Ji Y, Weng Y, Liang Y, et al. Identification of key genes unique to the luminal a and basal-like breast cancer subtypes via bioinformatic analysis. *World J Surg Oncol.* 2020;18(1):268. <https://doi.org/10.1186/s12957-020-02042-z>.
- Yuan Q, Zheng L, Liao Y, Wu G. Overexpression of CCNE1 confers a poorer prognosis in triple-negative breast cancer identified by bioinformatic analysis. *World J Surg Oncol.* 2021;19(1):86. <https://doi.org/10.1186/s12957-021-02200-x>.
- Liu C, Li Y, Wei M, Zhao L, Yu Y, Li G. Identification of a novel glycolysis-related gene signature that can predict the survival of patients with lung adenocarcinoma. *Cell Cycle.* 2019;18(5):568–79. <https://doi.org/10.1080/15384101.2019.1578146>.
- Martínez-Jiménez F, Muiños F, Sentís I, Deu-Pons J, Reyes-Salazar I, Arnedo-Pac C, et al. A compendium of mutational cancer driver genes. *Nat Rev Cancer.* 2020;20(10):555–72. <https://doi.org/10.1038/s41568-020-0290-x>.
- Bossi D, Cicalese A, Dellino GI, Luzi L, Riva L, D'Alesio C, et al. In vivo genetic screens of patient-derived tumors revealed unexpected frailty of the transformed phenotype. *Cancer Discov.* 2016;6(6):650–63. <https://doi.org/10.1158/2159-8290.CD-15-1200>.
- Zhao S, Liu J, Nanga P, Liu Y, Cicek AE, Knoblauch N, et al. Detailed modeling of positive selection improves detection of cancer driver genes. *Nat Commun.* 2019;10(1):3399. <https://doi.org/10.1038/s41467-019-11284-9>.
- Zhang G, Wang Y, Chen B, Guo L, Cao L, Ren C, et al. Characterization of frequently mutated cancer genes in Chinese breast tumors: a comparison of Chinese and TCGA cohorts. *Ann Transl Med.* 2019;7(8):179. <https://doi.org/10.21037/atm.2019.04.23>.
- Krøigård AB, Larsen MJ, Lænkholm AV, Knoop AS, Jensen JD, Bak M, et al. Identification of metastasis driver genes by massive parallel sequencing of successive steps of breast cancer progression. *PLoS One.* 2018;13(1):e0189887. <https://doi.org/10.1371/journal.pone.0189887>.
- Cui Y, Guo G. Immunomodulatory function of the tumor suppressor p53 in host immune response and the tumor microenvironment. *Int J Mol Sci.* 2016;17(11):1942.
- Yaghoobi H, Azizi H, Oskooei VK, Taheri M, Ghafouri-Fard S. Assessment of expression of interferon γ (IFN- γ) gene and its antisense (IFNG-AS1) in breast cancer. *World J Surg Oncol.* 2018;16(1):211. <https://doi.org/10.1186/s12957-018-1508-1>.
- Kim J. In silico analysis of differentially expressed genesets in metastatic breast cancer identifies potential prognostic biomarkers. *World J Surg Oncol.* 2021;19(1):188. <https://doi.org/10.1186/s12957-021-02301-7>.
- Zhou X, Xiao C, Han T, Qiu S, Wang M, Chu J, et al. Prognostic biomarkers related to breast cancer recurrence identified based on Logit model analysis. *World J Surg Oncol.* 2020;18(1):254. <https://doi.org/10.1186/s12957-020-02026-z>.
- Bianchi JJ, Zhao X, Mays JC, Davoli T. Not all cancers are created equal: tissue specificity in cancer genes and pathways. *Curr Opin Cell Biol.* 2020;63:135–43. <https://doi.org/10.1016/j.ceb.2020.01.005>.
- Liu WJ, Du Y, Wen R, Yang M, Xu J. Drug resistance to targeted therapeutic strategies in non-small cell lung cancer. *Pharmacol Ther.* 2020;206:107438. <https://doi.org/10.1016/j.pharmthera.2019.107438>.
- Hwang KT, Kim BH, Oh S, Park SY, Jung J, Kim J, et al. Prognostic role of KRAS mRNA expression in breast cancer. *J Breast Cancer.* 2019;22(4):548–61. <https://doi.org/10.4048/jbc.2019.22.e55>.
- Wang G, Wang Q, Liang N, Xue H, Yang T, Chen X, et al. Oncogenic driver genes and tumor microenvironment determine the type of liver cancer. *Cell Death Dis.* 2020;11(5):313. <https://doi.org/10.1038/s41419-020-2509-x>.
- Li G, Du X, Wu X, et al. Large-scale transcriptome analysis identified a novel cancer driver genes signature for predicting the prognostic of patients with hepatocellular carcinoma. *Front Pharmacol.* 2021;12:638622. <https://doi.org/10.3389/fphar.2021.638622>.
- Zhu C, Hu H, Li J, Wang J, Wang K, Sun J. Identification of key differentially expressed genes and gene mutations in breast ductal carcinoma in situ using RNA-seq analysis. *World J Surg Oncol.* 2020;18(1):52. <https://doi.org/10.1186/s12957-020-01820-z>.
- Hanna WM, Rüschoff J, Bilous M, Coudry RA, Dowsett M, Osamura RY, et al. HER2 in situ hybridization in breast cancer: clinical implications of polysomy 17 and genetic heterogeneity. *Mod Pathol.* 2014;27(1):4–18. <https://doi.org/10.1038/modpathol.2013.103>.
- Liu X, Jin G, Qian J, Yang H, Tang H, Meng X, et al. Digital gene expression profiling analysis and its application in the identification of genes associated with improved response to neoadjuvant chemotherapy in breast cancer. *World J Surg Oncol.* 2018;16(1):82. <https://doi.org/10.1186/s12957-018-1380-z>.
- Bianchini G, Balko JM, Mayer IA, Sanders ME, Gianni L. Triple-negative breast cancer: challenges and opportunities of a heterogeneous disease. *Nat Rev Clin Oncol.* 2016;13(11):674–90. <https://doi.org/10.1038/nrclinonc.2016.66>.
- Mao XH, Ye Q, Zhang GB, Jiang JY, Zhao HY, Shao YF, et al. Identification of differentially methylated genes as diagnostic and prognostic biomarkers of breast cancer. *World J Surg Oncol.* 2021;19(1):29. <https://doi.org/10.1186/s12957-021-02124-6>.

28. Ghafouri-Fard S, Oskooei VK, Azari I, Taheri M. Suppressor of cytokine signaling (SOCS) genes are downregulated in breast cancer. *World J Surg Oncol*. 2018;16(1):226. <https://doi.org/10.1186/s12957-018-1529-9>.
29. Mohamadizadeh-Hanjani Z, Shahbazi S, Geranpayeh L. Investigation of the SPAG5 gene expression and amplification related to the NuMA mRNA levels in breast ductal carcinoma. *World J Surg Oncol*. 2020;18(1):225. <https://doi.org/10.1186/s12957-020-02001-8>.
30. Yu YZ, Mu Q, Ren Q, Xie LJ, Wang QT, Wang CP. miR-381-3p suppresses breast cancer progression by inhibition of epithelial-mesenchymal transition. *World J Surg Oncol*. 2021;19(1):230.
31. Criscitiello C, Bagnardi V, Pruneri G, Vingiani A, Esposito A, Rotmensz N, et al. Prognostic value of tumour-infiltrating lymphocytes in small HER2-positive breast cancer. *Eur J Cancer*. 2017;87:164–71. <https://doi.org/10.1016/j.ejca.2017.10.011>.
32. Stanton SE, Disis ML. Clinical significance of tumor-infiltrating lymphocytes in breast cancer. *J Immunother Cancer*. 2016;4(1):59. <https://doi.org/10.1186/s40425-016-0165-6>.
33. Adams S, Loi S, Toppmeyer D, Cescon DW, de Laurentiis M, Nanda R, et al. Pembrolizumab monotherapy for previously untreated, PD-L1-positive, metastatic triple-negative breast cancer: cohort B of the phase II KEYNOTE-086 study. *Ann Oncol*. 2019;30(3):405–11. <https://doi.org/10.1093/annonc/mdy518>.
34. Schmid P, Rugo HS, Adams S, Schneeweiss A, Barrios CH, Iwata H, et al. Atezolizumab plus nab-paclitaxel as first-line treatment for unresectable, locally advanced or metastatic triple-negative breast cancer (IMpassion130): updated efficacy results from a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Oncol*. 2020;21(1):44–59. [https://doi.org/10.1016/S1470-2045\(19\)30689-8](https://doi.org/10.1016/S1470-2045(19)30689-8).
35. Matter MS, Marquardt JU, Andersen JB, Quintavalle C, Korokhov N, Stauffer JK, et al. Oncogenic driver genes and the inflammatory microenvironment dictate liver tumor phenotype. *Hepatology*. 2016;63(6):1888–99. <https://doi.org/10.1002/hep.28487>.
36. Guo WF, Zhang SW, Zeng T, Li Y, Gao J, Chen L. A novel network control model for identifying personalized driver genes in cancer. *PLoS Comput Biol*. 2019;15(11):e1007520. <https://doi.org/10.1371/journal.pcbi.1007520>.
37. Guo WF, Zhang SW, Zeng T, Akutsu T, Chen L. Network control principles for identifying personalized driver genes in cancer. *Brief Bioinform*. 2020; 21(5):1641–62. <https://doi.org/10.1093/bib/bbz089>.
38. Rizeq B, Malki M. The role of CCL21/CCR7 chemokine axis in breast cancer progression. *Cancers (Basel)*. 2020;12(4):1036.
39. Vahidi Y, Bagheri M, Ghaderi A, Faghih Z. CD8-positive memory T cells in tumor-draining lymph nodes of patients with breast cancer. *BMC Cancer*. 2020;20(1):257. <https://doi.org/10.1186/s12885-020-6714-x>.
40. Niu W, Luo Y, Zhou Y, Li M, Wu C, Duan Y, et al. BRD7 suppresses invasion and metastasis in breast cancer by negatively regulating YB1-induced epithelial-mesenchymal transition. *J Exp Clin Cancer Res*. 2020;39(1):30. <https://doi.org/10.1186/s13046-019-1493-4>.
41. Schmid P, Cortes J, Pusztai L, McArthur H, Kümmel S, Bergh J, et al. Pembrolizumab for early triple-negative breast cancer. *N Engl J Med*. 2020; 382(9):810–21. <https://doi.org/10.1056/NEJMoa1910549>.
42. Li X, Lian Z, Wang S, Xing L, Yu J. Interactions between EGFR and PD-1/PD-L1 pathway: implications for treatment of NSCLC. *Cancer Lett*. 2018;418:1–9. <https://doi.org/10.1016/j.canlet.2018.01.005>.
43. Cabo M, Offringa R, Zitvogel L, Kroemer G, Muntasell A, Galluzzi L. Trial Watch: Immunostimulatory monoclonal antibodies for oncological indications. *Oncoimmunology*. 2017;6(12):e1371896. <https://doi.org/10.1080/2162402X.2017.1371896>.
44. Topalian SL, Drake CG, Pardoll DM. Immune checkpoint blockade: a common denominator approach to cancer therapy. *Cancer Cell*. 2015;27(4): 450–61. <https://doi.org/10.1016/j.ccell.2015.03.001>.
45. Raskov H, Søby JH, Troelsen J, Bojesen RD, Gøgenur I. Driver gene mutations and epigenetics in colorectal cancer. *Ann Surg*. 2020;271(1):75–85. <https://doi.org/10.1097/SLA.0000000000003393>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

