



Expression of *CDK7* correlates with molecular subtypes and predicts clinical outcomes in breast cancer

Lin Tang^{1#}, Chengjun Zhu^{2#}, Juan Jin¹, Xin Wang¹, Lingyue Yu³, Xiaoxiang Guan^{1,2}

¹Department of Medical Oncology, Jinling Hospital, Medical School of Nanjing University, Nanjing, China; ²Department of Oncology, The First Affiliated Hospital of Nanjing Medical University, Nanjing, China; ³Center for Clinical Cancer Genetics & Global Health, Department of Medicine, The University of Chicago, Chicago, IL, USA

Contributions: (I) Conception and design: X Guan, L Tang; (II) Administrative support: X Guan; (III) Provision of study materials or patients: C Zhu; (IV) Collection and assembly of data: J Jin, X Wang, L Yu; (V) Data analysis and interpretation: L Tang, C Zhu; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Prof. Xiaoxiang Guan. Department of Medical Oncology, Jinling Hospital, Medical School of Nanjing University, Nanjing 210002, China. Email: xguan@nju.edu.cn.

Background: Cyclin-dependent kinase 7 (CDK7) belongs to CDK family, which is involved in transcriptional activity and cell cycle progression. Recent studies have suggested that CDK7 inhibitor is effective in the treatment of breast cancer, whereas, the efficacy of antitumor response varies significantly in different subtypes of breast cancer. However, the underlying connection between CDK7 and molecular subtypes of breast cancer is currently unknown.

Methods: To further elucidate the role of CDK7 in breast cancer, we investigated the association of *CDK7* expression with different molecular subtypes and clinical outcomes in breast cancer using the OncoPrint, GENT2, UALCAN and Prognoscan database. The correlation between *CDK7* mRNA expression and promoter methylation was established using MEXPRESS datasets. And the *CDK7* expression with different stages and subtypes was also examined by immunohistochemistry (IHC) assay in 140 breast cancer patients.

Results: We found that the expression of *CDK7* was increased in breast cancer. Besides, our data demonstrated that both CDK7 hypomethylation status and copy number variations (CNVs) were highly correlated with mRNA overexpression. Moreover, based on the tissue array of 140 patients and more than 4,000 samples from bc-GenExMiner database, CDK7 protein levels were significantly higher in luminal and human epidermal growth factor receptor 2 positive (HER2+) breast cancer in comparison with triple-negative breast cancer (TNBC) subtype. The expression of *CDK7* was prominently correlated with molecular markers of different subtypes of breast cancer. Furthermore, the results of the univariate analysis indicated that high *CDK7* expression was dramatically associated with poor overall survival (OS) ($P=0.0323$, HR =2.08).

Conclusions: Our results highlighted that elevated *CDK7* expression was correlated with molecular subtypes and acted as a candidate biomarker of poor prognosis in breast cancer.

Keywords: Cyclin-dependent kinase 7 (CDK7); expression; prognosis; molecular subtype; breast cancer

Submitted Sep 17, 2020. Accepted for publication Dec 24, 2020.

doi: 10.21037/tcr-20-2911

View this article at: <http://dx.doi.org/10.21037/tcr-20-2911>

Introduction

Breast cancer is one of the most prominent cancer types that lead to female mortality around the world (1). Great improvement has been achieved in the development of diagnosis and treatment for breast cancer patients. Breast cancer is a genetic and phenotypic heterogeneity disease with multiple distinct subtypes (2). Therefore, more reliable markers need to be pinpointed for early diagnosis and therapeutic strategies for individual patients.

Preclinical research indicated various potential interventional targets related to tumorigenesis including cell cycle process, which was regulated by cyclin-dependent kinases (CDKs) and other regulatory molecules. For example, CDK7 binds to MAT1 and cyclin H to constitute the CDK activating kinase (CAK), which executes its function by phosphorylating other CDKs involved in cell cycle control (3,4). Previous studies reported that CDK7 inhibitors affected transcriptional activity and cell cycle progression in various cancers (5-9).

Based on estrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER2) protein expression status, breast cancers could be classified into luminal A carcinomas (ER+ and/or PR+, HER2-), luminal B carcinomas (ER+ and/or PR+, HER2+), HER2 carcinomas (ER-, PR-, HER2+) and triple-negative carcinomas (ER-, PR-, HER2-) (10). In contrast to ER+ breast cancer, triple-negative breast cancer (TNBC) was reported to be uniquely dependent on CDK7 (7). Another study has shown that HER2+ breast cancers may be more sensitive to CDK7 inhibition THZ1 than ER+ breast cancers (11). To date, recent studies have shown that CDK7 inhibition is effective in breast cancer with significant subtype selectivity. However, little information has been published on the underlying connection between CDK7 and molecular subtypes of breast cancer patients.

Thus, we attempted to elucidate the function of CDK7 in breast cancer using public bioinformatics databases and tissue microarrays process to establish the correlation of *CDK7* expression with promoter methylation status, clinical outcomes and makers of different subtype of breast cancer. Meantime, we further identified that CDK7-related carcinogenic pathways changed at the molecular level in breast cancer. Collectively, our results might further support the application of CDK7 selective inhibitors in the treatment of certain subtypes of breast cancer. We present the following article in accordance with the REMARK reporting checklist (available at <http://dx.doi.org/10.21037/>

tcr-20-2911).

Methods

Public database and bioinformatics analysis

CDK7 mRNA expression in various types of cancer and normal tissues was ascertained by the Oncomine database (<https://www.oncomine.org/resource/main.html>) (12) and GENT2 database (<http://gent2.appex.kr/gent2/>) (13). CDK7 levels and clinical parameters in TCGA (<https://www.cancer.gov/tcga>) were extracted and examined in the UALCAN platform (<http://ualcan.path.uab.edu>) (14). The Prognoscan database (<http://dna00.bio.kyutech.ac.jp/Prognoscan/index.html>) (15) and the Human Protein Atlas database (<https://www.proteinatlas.org>) (16) were used to investigate the prognostic role of CDK7 mRNA in breast cancer. The bc-GenExMiner database (<http://bcgenex.centregauducheau.fr/BC-GEM>) (17) was used to analyze *CDK7* gene on ER, PR, and HER2. The correlation of CDK7 with molecular subtype markers from breast cancers was performed on GEPIA database (<http://gepia.cancer-pku.cn/index.html>) (18). The cBioPortal for Cancer Genomics (<http://www.cbioportal.org/>) (19) was a platform which was applied to visualize and analyze genetic alterations of CDK7. We evaluated the relationship between *CDK7* expression and its methylation and copy number variations (CNVs) using MEXPRESS (<https://mexpress.be>) (20). The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Ethical Committee of the First Affiliated Hospital of Nanjing Medical University (2019-SRFA-197) and individual consent for this retrospective analysis was waived.

Tissue microarray, patients, and follow-up

CDK7 was detected by immunohistochemistry (IHC) according to standard methods. The breast cancer tissue microarray (HBre-Duc140Sur-01) was obtained from Shanghai Outdo Biotech (Shanghai, China), and included tissue samples collected from January 2001 to August 2004. The tissue microarray had been previously described (21). The follow-up time was up to July 2013.

Analysis and quantification of staining

Anti-CDK7 antibody (Cell Signaling Technology, USA,

#2916, 1:500) was used to perform IHC staining. *CDK7* expression intensity was divided into: 0 (negative), 1 (weak), 2 (moderate), and 3 (strong). The staining percentage was as follows: 0 (<10% positive cells), 1 (10–25% positive cells), 2 (25–49% positive cells), 3 (50–74% positive cells) and 4 (\geq 75% positive cells). We combined two parameters to calculate the total score. Immunohistochemical staining was performed blindly by two pathologists. While, if the two pathologists disagree, we also consulted a third party for judgment. IHC score ≥ 6 indicated positive staining (22).

Statistical analysis

The Chi-squared and Fisher's exact test were executed to analyze the relationship between *CDK7* expression and clinical pathological characteristics via SPSS 26. We set the significance level for this study to be P values <0.05.

Results

The mRNA levels of *CDK7* in breast cancer

By analyzing the GENT databases and Oncomine databases, we investigated *CDK7* expression in cancer and normal samples. In GENT2 database, *CDK7* levels were significantly upregulated in breast carcinoma (Figure 1A). Oncomine database was used to detect the expression of *CDK7* mRNA in tumors and normal clinical samples, and the mRNA levels of *CDK7* in breast carcinoma were displayed in Figure 1B. Besides, analysis of the Oncomine database revealed that *CDK7* mRNA levels were notably higher in lobular breast carcinoma (LBC), ductal breast carcinoma (DBC) and invasive ductal breast carcinoma (IDBC) (Figure 1C,D,E,F) (10,23).

Analysis of GENT2 and the Oncomine databases indicated that *CDK7* levels were significantly increased in breast carcinoma. Therefore, we next investigated the correlation between *CDK7* levels and clinical parameters of breast carcinoma samples in UALCAN database. According to the subgroup analysis based on the patient's race, age, disease stage, lymph node metastasis status and *TP53* mutation status, *CDK7* expression in breast carcinoma patients was significantly higher than the normal control group. Thus, expression of *CDK7* could be served as a potential diagnostic indicator in breast cancer (Figure 2).

Integrated analysis of *CDK7* DNA methylation and CNVs in breast cancer

DNA methylation was a major epigenetic mechanism that modulated gene expression, which was a vital biological mechanism for tumor occurrence and development (24). Here, we investigated whether *CDK7* mRNA expression was correlated with its methylation status using MEXPRESS. As shown in Figure 3A, the plot indicated that there was a negative correlation between *CDK7* expression and methylation around the CpG island and promoter region. Seven abnormal methylation sites (cg26166102, $r=-0.073$; cg11859953, $r=-0.151$; cg18275539, $r=-0.118$; cg01558324, $r=-0.160$; cg07850415, $r=-0.131$; cg0227054, $r=-0.076$; cg16181978, $r=-0.071$) were observed and the Pearson correlation coefficients were indicated on the right-hand side. Here, DNA methylation of the *CDK7* promoter may further support *CDK7* overexpression in breast carcinoma. In addition to epigenetic mechanism, we speculate that *CDK7* gene mutation may also affect its expression.

Our finding indicated that *CDK7* expression was positively associated with its CNVs changes using MEXPRESS ($r=0.542$, $P<0.0001$). Our results subsequently indicated that CNVs of *CDK7* frequently occurred in breast cancer samples. Meanwhile, the cBioCancer Genomics Portal was accustomed to investigating the frequency of *CDK7* mutations in breast cancer. *CDK7* contained multiple mutations in breast cancer such as amplification, fusion, deep deletion and missense (Figure 3B), among which, missense was the major forms of mutation (Figure 3C).

CDK7 levels predict prognosis in breast cancer

We evaluated whether *CDK7* levels were associated with clinical outcomes among breast cancer. Notably, *CDK7* levels significantly impacted prognosis in breast cancer. The correlation between *CDK7* levels and prognosis of breast cancer patients was displayed in Table 1. Four cohorts (GSE9893, GSE2990, GSE3494-GPL96, GSE4922-GPL96) including 155 cases, 125 cases, 236 cases, and 249 cases at various stages of breast carcinoma showed that high *CDK7* expression was correlated with poor prognosis (Figure 4A,B,C,D,E). To further unveil the prognostic value of *CDK7* in breast carcinoma, we used Human Protein Atlas database ($P=0.0087$) (Figure 4F). Therefore, our results

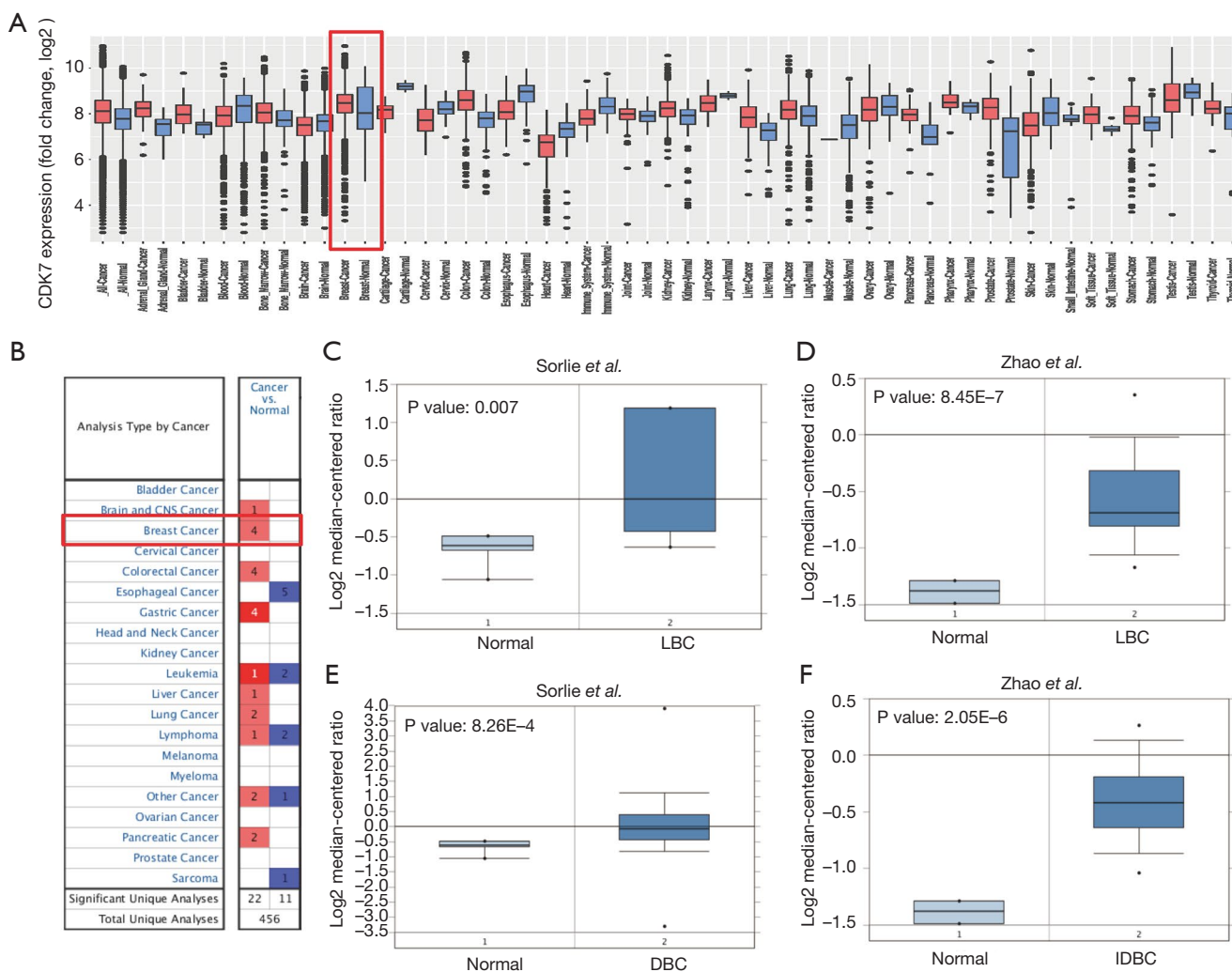


Figure 1 Transcription levels of CDK7 in breast cancer. (A) *CDK7* mRNA levels in normal and different types of cancer was searched in the GENT2 database. (B) The plot showed the number of datasets with highly significant mRNA down-expression (blue) or over-expression (red) of *CDK7* (cancer vs. normal) from the OncoPrint database. Threshold parameters: fold change of 1.5, and P value of 0.001. (C,D,E,F) Box plots compared *CDK7* levels in normal and breast carcinoma using OncoPrint database. The analysis was shown in (C,D) LBC, (E) DBC, (F) IDBC relative to normal breast tissue. *CDK7*, cyclin-dependent kinase 7; LBC, lobular breast carcinoma; DBC, ductal breast carcinoma; IDBC, invasive ductal breast carcinoma.

suggested that high *CDK7* expression could be a risk factor for poor prognosis in breast cancer.

***CDK7* expression is related to molecular subtype and poor prognosis of breast cancer**

We evaluated the clinicopathological features correlated with *CDK7* levels in 140 breast cancer samples by immunohistochemical analysis. *CDK7* protein was mainly existed in the nucleus. Representative photographs of

immunohistochemical staining for breast cancer tissues were presented in *Figure 5A*. As shown in *Table 2*, patients with higher *CDK7* levels appeared to be associated with advanced stages. What's more, the expression of *CDK7* was positively correlated with ER status (P=0.02) and PR status (P=0.011), which suggested that luminal breast cancer might positively express *CDK7*.

Next, we assessed the association between *CDK7* protein levels and molecular subtypes. IHC scores showed that *CDK7* protein was significantly higher in luminal (P=0.0018)

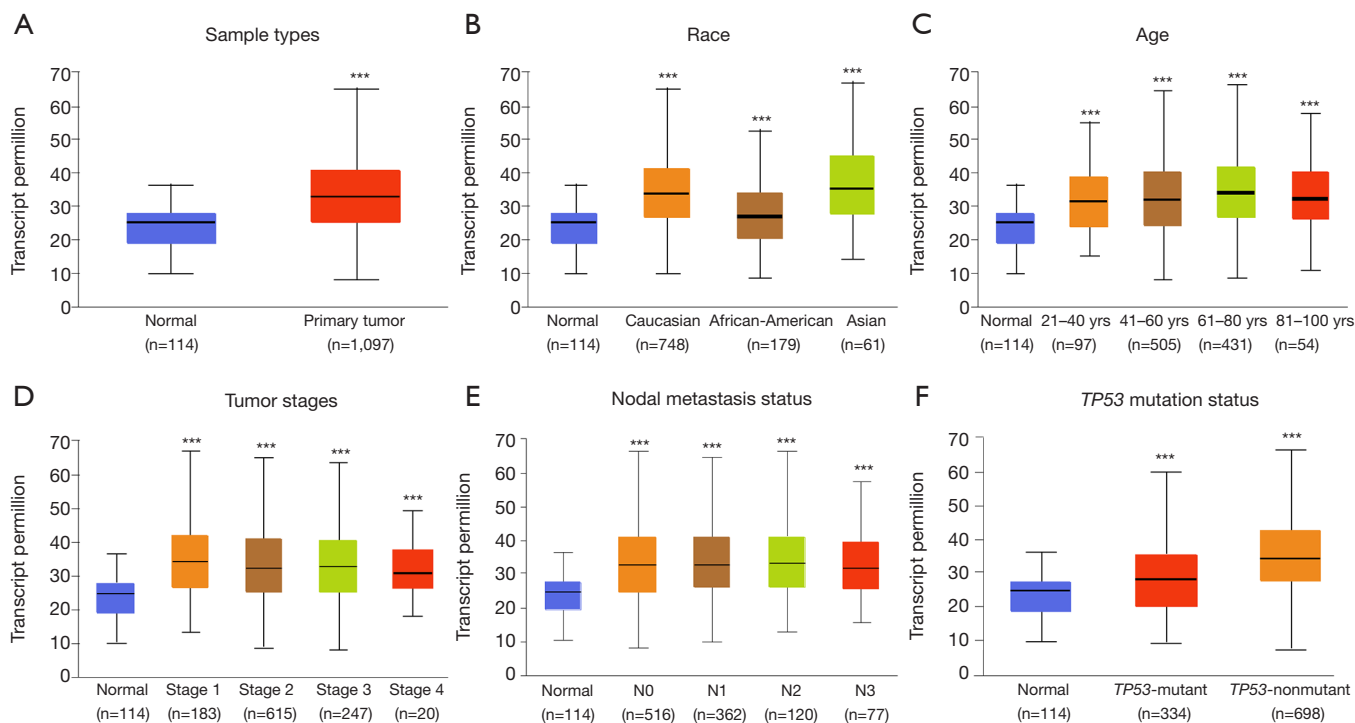


Figure 2 Relationship between *CDK7* expression and clinical parameters in breast cancer. The UALCAN database was used to assess the relationship between *CDK7* expression and clinical parameters such as: (A) sample types, (B) race, (C) age, (D) tumor stages, (E) nodal metastasis status, (F) *TP53* mutation status. ***, $P < 0.001$ versus normal. *CDK7*, cyclin-dependent kinase 7.

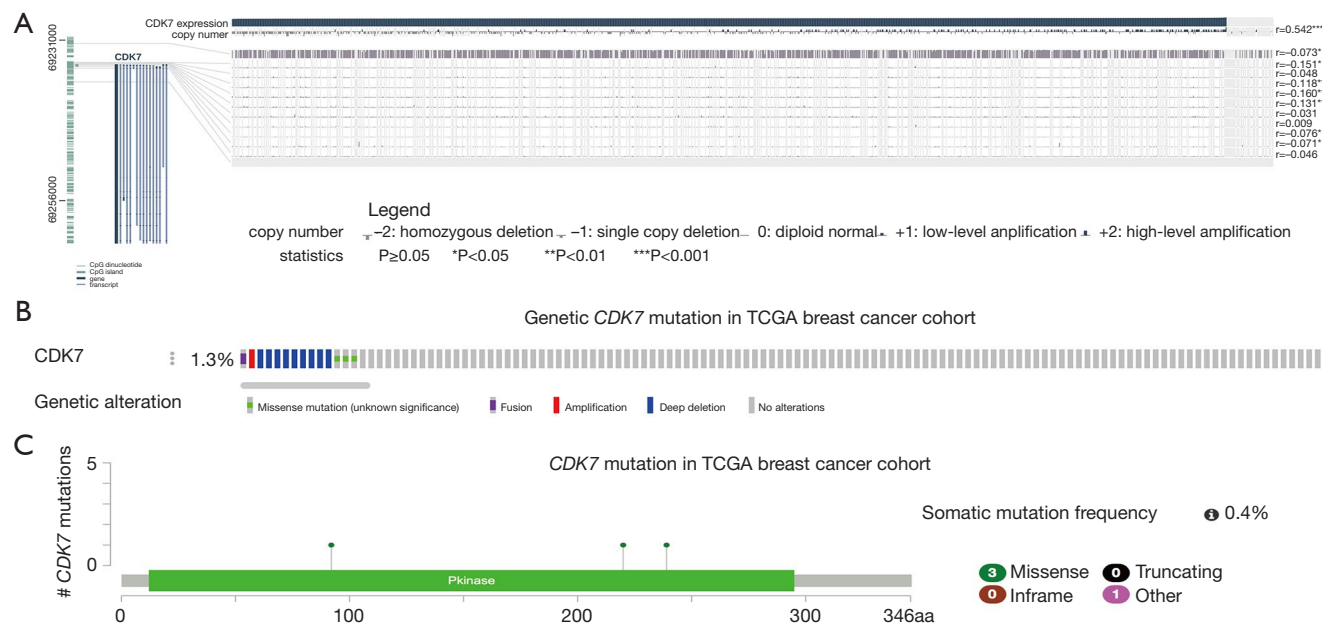


Figure 3 Relationship between *CDK7* levels and its methylation and CNVs in breast cancer. (A) This view showed the relationship between *CDK7* expression and CNVs, as well as methylation around CpG island and promoter region. (B) Genetic *CDK7* mutation in TCGA breast cancer cohort. (C) *CDK7* mutation type in TCGA breast cancer cohort. *CDK7*, cyclin-dependent kinase 7; CNV, copy number variation.

Table 1 Relationship between CDK7 levels and survival trends in breast cancer

Dataset	Endpoint	Number	Probe ID	ln (HR _{high} /HR _{low})	COX P value	ln (HR)	HR (95% CI)
GSE9893	OS	155	3581	0.91	0.008742	0.2	1.22 (1.05–1.41)
GSE2990	DMFS	125	211297_s_at	1.25	0.000905	1.19	3.28 (1.63–6.63)
GSE2990	RFS	125	211297_s_at	0.67	0.044463	0.58	1.78 (1.01–3.12)
GSE3494-GPL96	DSS	236	211297_s_at	1.19	0.020482	0.90	2.46 (1.15–5.26)
GSE4922-GPL96	DFS	249	211297_s_at	0.96	0.046541	0.62	1.86 (1.01–3.44)

CDK7, cyclin-dependent kinase 7; OS, overall survival; DMFS, distant metastasis free survival; RFS, relapse free survival; DSS, disease specific survival; DFS, disease free survival.

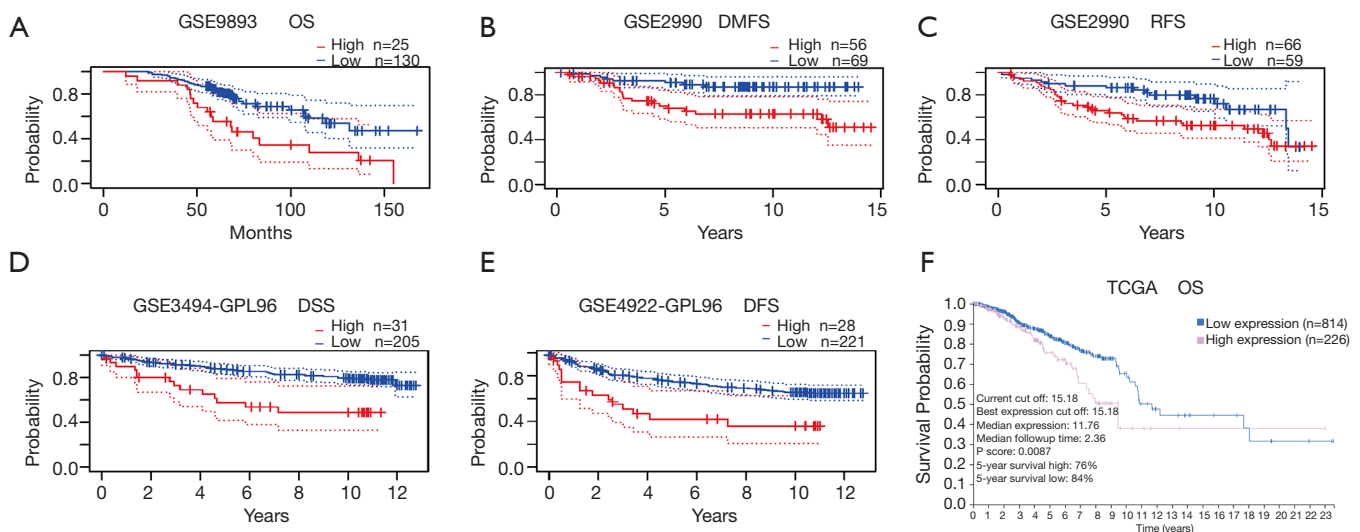


Figure 4 Survival assay of CDK7 in breast cancer patients. Survive curves evaluating the prognostic value of CDK7 mRNA expression in the PrognScan, including: (A) OS (GSE9893 n=155), (B) DMFS (GSE2990 n=125), (C) RFS (GSE2990 n=125), (D) DSS (GSE3494-GPL96 n=236), (E) DFS (GSE4922-GPL96 n=249). (F) Impact of CDK7 level on OS in TCGA breast cancer cohort (n=1,040). CDK7, cyclin-dependent kinase 7; OS, overall survival; DMFS, distant metastasis free survival; RFS, relapse free survival; DSS, disease specific survival; DFS, disease free survival.

and HER2+ ($P=0.0194$), compared with CDK7 protein levels in TNBC (Figure 5B). Also, CDK7 levels were significantly different between luminal *vs.* not luminal ($P=0.0437$), and TNBC *vs.* not TNBC ($P=0.0039$) (Figure 5C,D). However, CDK7 levels were not associated with HER2 status ($P=0.0523$) (Figure 5E). Samples of HER2 positive status were excluded in the next analysis because CDK7 levels were significantly higher in HER2 positive samples, which could improve CDK7 expression in ER- groups. CDK7 protein levels were significantly positively correlated with ER status after excluding HER2-positive cases ($P=0.0179$) (Figure 5F).

More importantly, we investigated the relationship

between CDK7 levels and clinical outcome by extracting survival data. The results of the univariate analysis indicated that CDK7 overexpression was remarkably correlated with shorter OS in breast cancer patients ($P=0.0323$, HR =2.08) (Figure 5G).

In bc-GenExMiner database, CDK7 expression was enriched in luminal and HER2+ compared with basal-like subtype (Figure 6A). ER and PR status were both positively associated with CDK7 expression (Figure 6B,C). In addition, there were no obvious difference of CDK7 levels between HER2+ and HER2- group (Figure 6D). Moreover, compared with TNBC and basal-like breast cancer, CDK7 levels were apparently upregulated in breast cancer

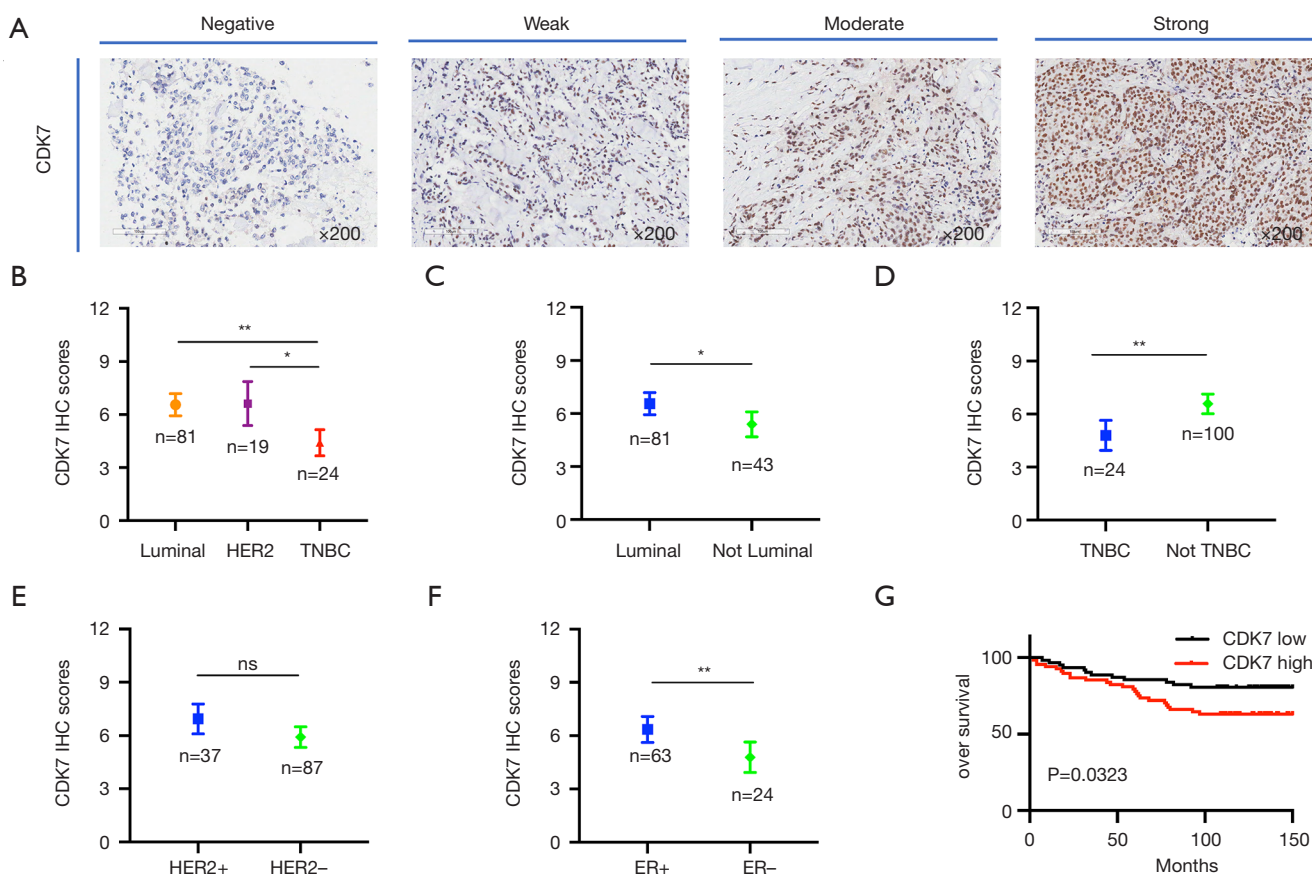


Figure 5 CDK7 levels correlated with molecular subtypes. (A) CDK7 staining was shown as representative photographs. The staining results were negative (0), weak (1+), moderate (2+), and strong (3+). (B) IHC scores indicated that CDK7 levels were higher in luminal and HER2+ compared with TNBC subtype. (C) IHC scores indicated that CDK7 levels were higher in luminal compared with not luminal subtype. (D) IHC scores indicated that CDK7 levels were lower in TNBC compared with not TNBC subtype. (E) IHC scores indicated that CDK7 levels were not significantly associated with HER2 status at whole level. (F) CDK7 levels were higher in ER+ than ER- breast cancer at the absence of HER2 enriched cases. (G) Kaplan Meier survival curves compared the high and low expression of CDK7 in breast cancer patients. *, $P < 0.05$; **, $P < 0.005$. IHC, immunohistochemistry; TNBC, triple-negative breast cancer.

subtypes other than TNBC and basal-like breast cancer (Figure 6E,F).

Correlation between CDK7 and luminal, HER2+, and basal-like subtype markers

We evaluated whether there was any association between CDK7 levels and luminal, HER2+, and basal-like subtype markers from breast cancer using the GEPIA web tools. Correlation analysis showed that CDK7 levels were positively correlated with luminal subtype markers of estrogen receptor (ESR1) ($R=0.37, P < 0.001$), progesterone receptor (PGR) ($R=0.25, P < 0.001$), endothelial transcription

factor 3 (GATA3) ($R=0.42, P < 0.001$), and were negatively correlated with basal-like subtype marker epidermal growth factor receptor (EGFR) ($R=-0.091, P < 0.001$), keratin 5 (KRT5) ($R=-0.11, P < 0.001$), but were not correlated with HER2+ subtype markers of epidermal growth factor receptor-2 (ERBB2) ($P=0.43$) (Figure 7).

Discussion

Altogether, our analysis of the Oncomine database and UALCAN showed that CDK7 mRNA was significantly upregulated among breast cancer. Our study verified the correlation between CDK7 expression and methylation

Table 2 Relationship between *CDK7* expression and clinicopathologic characteristics of breast cancer samples

Variables	CDK7 ^a , n (%)		Total, n (%)	P value
	Low	High		
Age				0.369
<50 years	32 (47.8)	26 (40.0)	58 (43.9)	
≥50 years	35 (52.2)	39 (60.0)	74 (56.1)	
Pathology grade ^b				0.145
Grade 1	8 (12.3)	3 (4.7)	11 (8.5)	
Grade 2	54 (83.1)	54 (84.4)	108 (83.7)	
Grade 3	3 (4.6)	7 (10.9)	10 (7.8)	
T stage				0.001
1	22 (32.8)	5 (7.7)	27 (20.5)	
2	37 (55.2)	54 (83.1)	91 (68.9)	
3	8 (11.9)	6 (9.2)	14 (10.6)	
Nodal status ^c				0.747
0	23 (35.4)	25 (38.5)	48 (36.9)	
1	20 (30.8)	19 (29.2)	39 (30.0)	
2	16 (24.6)	18 (27.7)	34 (26.2)	
3	6 (9.2)	3 (4.6)	9 (6.9)	
ER status ^d				0.020
Positive	35 (55.6)	46 (75.4)	81 (65.3)	
Negative	28 (44.4)	15 (24.6)	43 (34.7)	
PR status ^e				0.011
Positive	27 (42.9)	40 (65.6)	67 (54.1)	
Negative	36 (57.1)	21 (34.4)	57 (45.9)	
HER2 status ^f				0.754
Positive	18 (28.6)	19 (31.1)	37 (29.8)	
Negative	45 (71.4)	42 (68.9)	87 (70.2)	
Subtypes				0.047
Luminal	35 (55.6)	46 (75.4)	81 (65.3)	
HER2-positive	11 (17.5)	8 (13.1)	19 (15.3)	
TNBC	17 (26.9)	7 (11.5)	24 (19.4)	

^a, eight samples staining missed; ^b, two and one samples of pathology grade status of CDK7 low and high expression group were missed, respectively; ^c, two samples of nodal status of *CDK7* low expression group were missed; ^d, four and four samples of ER status of CDK7 low and high expression group were missed, respectively; ^e, four and four samples of PR status of CDK7 low and high expression group were missed, respectively; ^f, four and four samples of HER2 status of CDK7 low and high expression group were missed, respectively. CDK7, cyclin-dependent kinase 7; ER, estrogen receptor; PR, progesterone receptor; TNBC, triple-negative breast cancer.

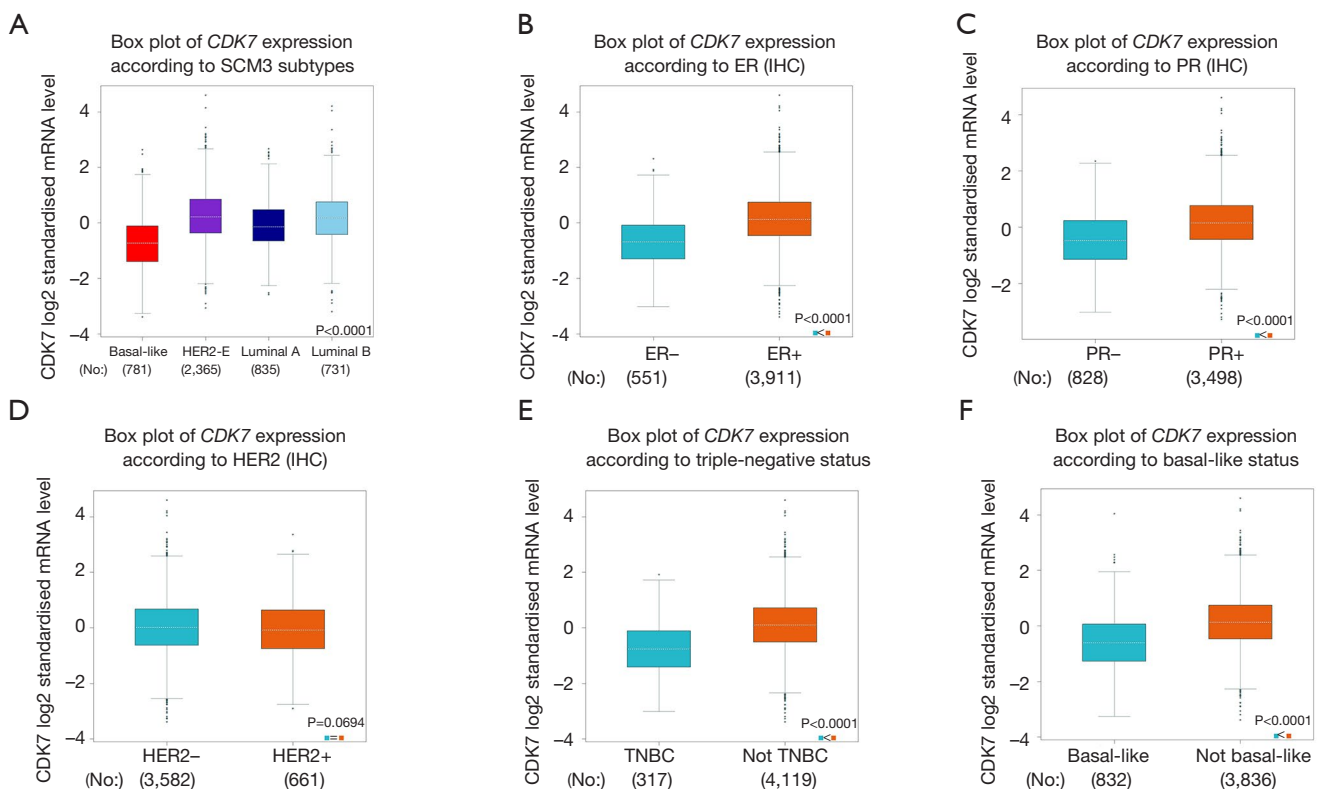


Figure 6 CDK7 levels in different groups were analyzed according to clinical parameters by the bc-GenExMiner database. The box plot showed as: (A) subtypes, (B) ER, (C) PR, (D) HER2, (E) triple-negative status, and (F) basal-like status. CDK7, cyclin-dependent kinase 7; ER, estrogen receptor; PR, progesterone receptor; IHC, immunohistochemistry.

status or copy number amplification, which may lead to an increase in *CDK7* expression levels. Consistently, *CDK7* overexpression was significantly associated with disease stages, nodal metastasis status and *TP53* mutation status, which indicated that high expression of *CDK7* could be involved in breast carcinoma progression. Accordingly, high *CDK7* levels were significantly associated with shorter OS.

Based on tissue microarray and database information, our analysis showed that *CDK7* protein levels were significantly higher in luminal and HER2+ breast cancer compared with TNBC. We found the *CDK7* levels were positively associated with luminal subtype makers. Although the potential mechanism of the relationship between *CDK7* expression and ER remained unclear, previous studies had revealed that *CDK7* was closely related to ER phosphorylation at Ser118, and *in vitro* studies had confirmed that *CDK7* can phosphorylate this site (25). This may provide a potential explanation for the association between *CDK7* and ER protein levels in breast carcinoma.

The result supports the design of *CDK7* inhibitor in treating ER+ breast cancer, either being administered alone or in coordination with hormone therapy, which was expected to be a promising strategy to overcome resistance to endocrine therapy. These results indicated that *CDK7* was not only a biomarker with diagnostic and prognostic value, but also a potential therapeutic target.

For the moment, a great deal of selective small molecule *CDK7* inhibitors have been developed. Preclinical studies have found that *CDK7* inhibitors can be effective for a variety of solid tumors, including breast cancer (7,8,26-28). Based on these previous studies, currently four *CDK7* inhibitors, CT7001, SY-1365, SY-5609 and LY3405105, have now progressed to phase I/II clinical trials to test monotherapy as well as combination treatments in breast cancer. Our research demonstrated that luminal subtype has the highest expression of *CDK7*. However, the sensitivity of *CDK7* inhibitors to different subtypes of breast cancer did not depend on the expression levels. It had been reported that it was TNBC not ER+

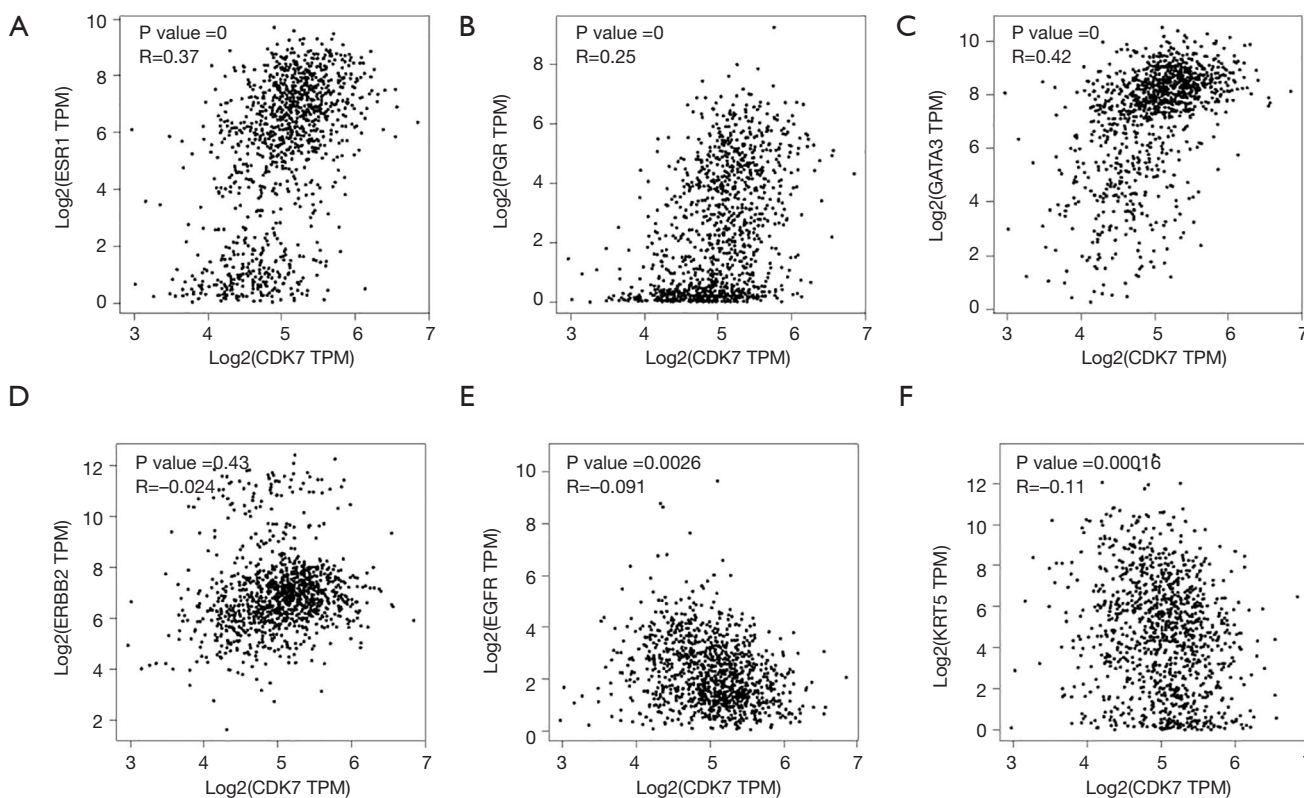


Figure 7 Correlations between CDK7 levels and markers of different subtypes of breast cancer. Correlation analysis of TCGA exhibited that *CDK7* mRNA was associated with (A) ESR1, (B) PGR, (C) GATA3, (D) ERBB2, (E) EGFR, and (F) KRT5. CDK7, cyclin-dependent kinase 7.

breast cancer that was sensitive to CDK7 inhibitor THZ1 (7). Mechanistically, it may inhibit TNBC-related genes regulated by CDK7 mediated super-enhancers to increase the sensitivity of THZ1. Furthermore, recent results have also suggested that in HER2+ breast cancer CDK7 inhibitor may have a strong response to combined inhibition of CDK7 and HER2 (11). The result suggested that CDK7 inhibition can be an additional therapeutic option in refractory breast cancer. This may explain that although CDK7 inhibitor THZ1 was not selective for breast cancer subtypes, the sensitivity of different subtypes varies greatly. In other words, THZ1 sensitivity did not correlate with *CDK7* expression levels.

Conclusions

In summary, the present study suggested that *CDK7* was overexpressed in different subtypes of breast cancer and

correlated with methylation of the *CDK7* promoter. *CDK7* could be used as a promising biomarker for predicting prognosis of breast carcinoma. However, further experiments and clinical trials are required to support the application of CDK7 selective inhibitors in different breast cancer subtypes.

Acknowledgments

Funding: This work was supported by National Natural Science Foundation of China (No. 81773102), and Key International Cooperation of National Natural Science Foundation of China (No. 81920108029).

Footnote

Reporting Checklist: The authors have completed the REMARK reporting checklist. Available at <http://dx.doi.org/10.21037/tcr-20-2911>

[org/10.21037/tcr-20-2911](https://doi.org/10.21037/tcr-20-2911)

Data Sharing Statement: Available at <http://dx.doi.org/10.21037/tcr-20-2911>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <http://dx.doi.org/10.21037/tcr-20-2911>). 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). The study was approved by the Ethical Committee of the First Affiliated Hospital of Nanjing Medical University (2019-SRFA-197) and individual consent for this retrospective analysis was 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. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018;68:394-424.
2. Beca F, Polyak K. Intratumor heterogeneity in breast cancer. *Adv Exp Med Biol* 2016;882:169-89.
3. Fisher RP. The CDK network: linking cycles of cell division and gene expression. *Genes Cancer* 2012;3:731-8.
4. Yankulov KY, Bentley DL. Regulation of CDK7 substrate specificity by MAT1 and TFIIF. *EMBO J* 1997;16:1638-46.
5. Chipumuro E, Marco E, Christensen CL, et al. CDK7 inhibition suppresses super-enhancer-linked oncogenic transcription in MYCN-driven cancer. *Cell* 2014;159:1126-39.
6. Christensen CL, Kwiatkowski N, Abraham BJ, et al. Targeting transcriptional addictions in small cell lung cancer with a covalent CDK7 inhibitor. *Cancer Cell* 2014;26:909-22.
7. Wang Y, Zhang T, Kwiatkowski N, et al. CDK7-dependent transcriptional addiction in triple-negative breast cancer. *Cell* 2015;163:174-86.
8. Zhang Z, Peng H, Wang X, et al. Preclinical efficacy and molecular mechanism of targeting CDK7-dependent transcriptional addiction in ovarian cancer. *Mol Cancer Ther* 2017;16:1739-50.
9. Greenall SA, Lim YC, Mitchell CB, et al. Cyclin-dependent kinase 7 is a therapeutic target in high-grade glioma. *Oncogenesis* 2017;6:e336.
10. Sorlie T, Tibshirani R, Parker J, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci U S A* 2003;100:8418-23.
11. Sun B, Mason S, Wilson RC, et al. Inhibition of the transcriptional kinase CDK7 overcomes therapeutic resistance in HER2-positive breast cancers. *Oncogene* 2020;39:50-63.
12. Rhodes DR, Yu J, Shanker K, et al. ONCOMINE: a cancer microarray database and integrated data-mining platform. *Neoplasia* 2004;6:1-6.
13. Park SJ, Yoon BH, Kim SK, et al. GENT2: an updated gene expression database for normal and tumor tissues. *BMC Med Genomics* 2019;12:101.
14. Chandrashekar DS, Bashel B, Balasubramanya SAH, et al. UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses. *Neoplasia* 2017;19:649-58.
15. Mizuno H, Kitada K, Nakai K, et al. PrognScan: a new database for meta-analysis of the prognostic value of genes. *BMC Med Genomics* 2009;2:18.
16. Pontén F, Schwenk JM, Asplund A, et al. The Human Protein Atlas as a proteomic resource for biomarker discovery. *J Intern Med* 2011;270:428-46.
17. Jézéquel P, Frénel JS, Campion L, et al. bc-GenExMiner 3.0: new mining module computes breast cancer gene expression correlation analyses. *Database (Oxford)* 2013;2013:bas060.
18. Tang Z, Li C, Kang B, et al. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res* 2017;45:W98-102.
19. Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* 2012;2:401-4.
20. Koch A, De Meyer T, Jeschke J, et al. MEXPRESS:

- visualizing expression, DNA methylation and clinical TCGA data. *BMC Genomics* 2015;16:636.
21. Tang L, Jin J, Xu K, et al. SOX9 interacts with FOXC1 to activate MYC and regulate CDK7 inhibitor sensitivity in triple-negative breast cancer. *Oncogenesis* 2020;9:47.
 22. Huang T, Ding X, Xu G, et al. CDK7 inhibitor THZ1 inhibits MCL1 synthesis and drives cholangiocarcinoma apoptosis in combination with BCL2/BCL-XL inhibitor ABT-263. *Cell Death Dis* 2019;10:602.
 23. Zhao H, Langerod A, Ji Y, et al. Different gene expression patterns in invasive lobular and ductal carcinomas of the breast. *Mol Biol Cell* 2004;15:2523-36.
 24. Bird A. DNA methylation patterns and epigenetic memory. *Genes Dev* 2002;16:6-21.
 25. Chen D, Riedl T, Washbrook E, et al. Activation of estrogen receptor alpha by S118 phosphorylation involves a ligand-dependent interaction with TFIID and participation of CDK7. *Mol Cell* 2000;6:127-37.
 26. Patel H, Periyasamy M, Sava GP, et al. ICEC0942, an orally bioavailable selective inhibitor of CDK7 for cancer treatment. *Mol Cancer Ther* 2018;17:1156-66.
 27. Wang BY, Liu QY, Cao J, et al. Selective CDK7 inhibition with BS-181 suppresses cell proliferation and induces cell cycle arrest and apoptosis in gastric cancer. *Drug Des Devel Ther* 2016;10:1181-9.
 28. Hu S, Marineau JJ, Rajagopal N, et al. Discovery and characterization of SY-1365, a selective, covalent inhibitor of CDK7. *Cancer Res* 2019;79:3479-91.

Cite this article as: Tang L, Zhu C, Jin J, Wang X, Yu L, Guan X. Expression of CDK7 correlates with molecular subtypes and predicts clinical outcomes in breast cancer. *Transl Cancer Res* 2021;10(2):669-680. doi: 10.21037/tcr-20-2911