A Nomogram Based on CENPP Expression for Survival Prediction in Breast Cancer

Heyan Chen

First Affiliated Hospital of Xi'an Jiaotong University

Shengyu Pu

The First Affiliated Hospital of Xi'an Jiaotong University

Shibo Yu

The First Affiliated Hospital of Xi'an Jiaotong University

Xiaoqin Liao

The First Affiliated Hospital of Xi'an Jiaotong University

Jianjun He

The First Affiliated Hospital of Xi'an Jiaotong University

Huimin Zhang (✉ huimin.zhang@xjtu.edu.cn)

The First Affiliated Hospital of Xi'an Jiaotong University

Research article

Keywords: Breast cancer, CENPP gene, overall survival, nomogram

Posted Date: September 28th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-78191/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Version of Record: A version of this preprint was published at Gland Surgery on June 1st, 2021. See the published version at <https://doi.org/10.21037/gs-21-30>.

Abstract

Introduction: In recent years, it has been found that the expression of 17 centromere proteins (CENPs) is closely related to malignant tumors. This study intends to investigate the prognostic value of CENPs in breast cancer (BC).

Methods: A total of 800 BC patients were included from the TCGA database. The Cox proportional regression models was used to develop a CENPs-related prognostic signature. Furthermore, the mRNA expression and overall survival (OS) of CENPP (centromere protein P) in BC patients with different clinicopathological features was analyzed via GEPIA, bcGenExMiner v4.4 and Kaplan-Meier plotter. Finally, the nomogram was established based on the independent predictors recognized by multivariate Cox regression analysis and further validated by receiver-operating characteristic (ROC) curves and calibration plots internally and externally.

Results: The result shown that age, Her2 status, pathologic_T stage, pathologic_M stage and CENPP expression with independent prognostic values for BC. CENPP was overexpressed in BC tissues and CENPP high expression was associated with better OS. We then established a nomogram based on those independent predictors, and the calibration curve demonstrated good fitness of the nomogram for OS prediction. In the training set, the AUC of 3-year and 5-year survival were 0.757 and 0.797, respectively. In the validation set, the AUC of 3-year and 5-year survival were 0.727 and 0.71.

Conclusion: Our study showed that CENPP is a novel prognostic factor for patients with BC, and the established nomogram can provide valuable information on prognostic prediction for patients with BC.

Introduction

Breast cancer (BC) is the most common malignancy in women worldwide and remains one of the leading causes of cancer death among women all over the world[1]. BC is classified into different molecular subtypes based on the expression of estrogen receptor (ER), progesterone receptor (PR) and Her2 amplification which are also of prognostic value for BC patients. At present, clinical workers combined the above indicators with clinical features to evaluate the prognosis and implement treatment plan for BC patients. However, the prediction performance is not good enough and professionals have been trying to find more biological genes and molecular indicators that can accurately predict the prognosis of BC, which may also be potential therapeutic targets for BC.

Constitutive Centromere-Associated Network (CCAN) underlies the centromere specificity and stability of the kinetochore in mitosis of human cancer cells[2–6]. To date, 17 members of CCAN have been identified in human, including CENPA/C/H/I/K/L/M/N/O/P/Q/R/S/T/U/W/X, and each of them was closely connected and interacted[7]. Centromere protein abnormalities is essential for cancer development[8]. In recent years, previous studies have shown that CENPA/H/U/I/O were associated with breast cancer[9–14], lung cancer[15], bladder cancer[16] and gastric cancer[17], which urged us to investigate the prognostic role of CENPs. Since the expression of CENPC was not found on Xena platform, we excluded it from this study.

In this study, data of 16 CENPs from TCGA database were analyzed and only CENPP was found to be a prognostic factor in the Cox regression analysis. Then, we established a prognostic nomogram based on CENPP expression and clinicopathological features of BC, whose predictive accuracy was further confirmed by internal and external validation.

Materials And Methods

Datasets

We downloaded the mRNA expression profile of CENPs in the TCGA Breast Cancer from the Xena system (<https://xenabrowser.net/datapages/>) for statistical analysis. In this study, we selected 1215 BC samples with raw counts of RNAseq expression data and corresponding clinical information.

The mRNA expression of CENPP in BC patients was analyzed via GEPIA and bcGenExMiner v4.4. GEPIA is a web-based tool to analyze the RNA expression data of 8,587 normal and 9,736 tumor samples from the TCGA and the GTEx projects^[18]. These two databases were used to verify the mRNA expression of CENPP in BC (tumor vs. normal). bcGenExMiner v4.4 is a dataset of published annotated BC transcriptomic. The statistical analyses are divided into three modules: "expression", "prognosis" and "correlation"^[19,20]. The expression module could be utilized to compare the expression of candidate gene under different clinical features, such as receptor status (ER + vs ER-, PR + vs PR-, HER2 + vs HER2- by IHC), nodal status, SBR, age, molecular subtypes and so on. We used bcGenExMiner v4.4 to analyze the relationship between the CENPP gene and hormone receptor status as well as molecular subtypes of BC.

Nomogram construction and validation

Univariable as well as multivariable Cox regression analyses were conducted to screen the predictors for overall survival. Independent prognostic genes as well as clinical predictors were included in the nomogram construction by using the package of rms^[21]. Then, the nomogram was validated internally in the training set and externally in the validation set. The consistency between the predicted and actual 3-year and 5-year survival rates was measured by the receiver-operating characteristic (ROC) curves and calibration plots. The area under the curve (AUC) was calculated, and 0.5 represents a random probability while 1 indicates a complete discrimination. In a perfect calibration model, calibration plots (1000 bootstrap resamples) would fall on a 45° diagonal line.

Statistical Analysis

We conducted all the analysis using SPSS 19.0 (SPSS, Chicago, IL) and R statistical software version 3.5.2 (<https://www.r-project.org/>). All *p* values were two-sided and the level of significance was set at $P < 0.05$.

Results

Clinicopathologic Characteristics of the Patients

The TCGA Breast Cancer (BRCA) dataset from Xena platform cataloged 1215 breast cancer patients. After excluding 415 patients with incomplete relevant information, 800 patients who met the eligible criteria were included, which were divided into a training set ($N = 480$) and a validation set ($N = 320$) randomly at the rate of 3:2. The clinicopathological characteristics of the training set and the validation set were similar (Table 1).

Table 1
Clinicopathological characteristics of the study populations.

cCCharacteristics	Training set (N = 480)		Validation set (N = 320)		P value	cCCharacteristics	Training set (N = 480)		Validation set (N = 320)		P value
	No.	%	No.	%			No.	%	No.	%	
Age					0.896	CENPI					0.541
< 65	334	69.6	220	68.8		Low	204	42.5	143	44.7	
65–75	62	12.9	45	14.1		High	276	57.5	177	55.3	
> 75	84	17.5	55	17.2		CENPK					0.073
Gender					0.784	Low	209	43.5	160	50.0	
Female	475	99.0	316	98.8		High	271	56.5	160	50.0	
Male	5	1.0	4	1.3		CENPL					0.795
ER					0.27	Low	237	49.4	155	48.4	
Negative	115	24.0	66	20.6		High	243	50.6	165	51.6	
Positive	365	76.0	254	79.4		CENPM					0.209
PR					0.602	Low	193	40.2	143	44.7	
Negative	163	34.0	103	32.2		High	287	59.8	177	55.3	
Positive	317	66.0	217	67.8		CENPN					0.118
Her2					0.864	Low	249	51.9	184	57.5	
Negative	370	77.1	245	76.6		High	231	48.1	136	42.5	
Positive	110	22.9	75	23.4		CENPO					0.371
Histological_type					0.418	Low	238	49.6	169	52.8	
IDC	351	73.1	234	73.1		High	242	50.4	151	47.2	
ILC	85	17.7	49	15.3		CENPP					0.885
Other	44	9.2	37	11.6		Low	260	54.2	175	54.7	
Pathologic_T					0.323	High	220	45.8	145	45.3	
T1	113	23.5	84	26.3		CENPQ					0.061
T2	296	61.7	177	55.3		Low	233	48.5	177	55.3	
T3	55	11.5	45	14.1		High	247	51.5	143	44.7	
T4	16	3.3	14	4.4		CENPR					0.285
Pathologic_N					0.98	Low	238	49.6	171	53.4	
N0	226	47.1	152	47.5		High	242	50.4	149	46.6	
N1	157	32.7	106	33.1		CENPS					0.644
N2	57	11.9	38	11.9		Low	238	49.6	164	51.3	

cCCharacteristics	Training set (N = 480)		Validation set (N = 320)		P value	cCCharacteristics	Training set (N = 480)		Validation set (N = 320)		P value
	No.	%	No.	%			No.	%	No.	%	
N3	40	8.3	24	7.5		High	242	50.4	156	48.8	
Pathologic_M					0.223	CENPT					0.908
M0	421	87.7	269	84.1		Low	262	54.6	176	55.0	
M1	4	0.8	6	1.9		High	218	45.4	144	45.0	
Mx	55	11.5	45	14.1		CENPU					0.385
CENPA					0.164	Low	255	53.1	180	56.3	
Low	210	43.8	156	48.8		High	225	46.9	140	43.8	
High	270	56.3	164	51.3		CENPW					0.523
CENPH					0.078	Low	259	54.0	180	56.3	
Low	229	47.7	173	54.1		High	221	46.0	140	43.8	
High	251	52.3	147	45.9							

Independent Prognostic Factors of BC

Age, Her2 status, pathologic_T stage, pathologic_N stage, pathologic_M stage and CENPP expression were identified as predictive factors for OS of BC in the univariate analysis (Table 2), while except pathologic_N stage, all the other variables were further confirmed as independent predict factors in multivariate analysis (Table 3). The results showed that BC patients of 65–75 years old ($P = 0.015$, $HR = 2.67$ [95% CI: 1.21–5.86]) and > 75 years old ($P < 0.000$, $HR = 3.63$ [95% CI: 1.87–7.03]) had worse OS than those < 65 years old. In addition, Her2 positive patients ($P = 0.027$, $HR = 2.04$ [95% CI: 1.09–3.82]) had worse OS than Her2 negative patients. BC patients with pathologic_T4 stage ($P = 0.003$, $HR = 5.401$ [95% CI: 1.78–16.38]) and pathologic_M1 stage ($P = 0.040$, $HR = 4.45$ [95% CI: 0.07–18.47]) had worse OS compared with pathologic_T1 stage and pathologic_M0 stage, respectively. Additionally, BC patients with expression of CENPP lower than the medium ($P = 0.005$, $HR = 2.35$ [95% CI: 1.30–4.23]) were significantly correlated with worse OS.

Table 2
Univariate analysis of the training set.

Characteristics	Univariate analysis			Characteristics	Univariate analysis		
	HR	95%CI	P value		HR	95%CI	P value
Age				CENPI			
< 65	1			High	1		
65–75	1.85	0.89–3.85	0.100	Low	1.30	0.77–0.80	0.296
> 75	2.87	1.59–5.17	0.000	CENPK			
Gender				High	1		
Female	1			Low	1.11	0.68–1.82	0.670
Male	0.00	0.00–Inf	0.996	CENPL			
ER				High	1		
Negative	1			Low	1.16	0.71–1.90	0.555
Positive	0.62	0.37–1.04	0.072	CENPM			
PR				High	1		
Negative	1			Low	1.40	0.86–2.30	0.177
Positive	0.72	0.44–1.17	0.185	CENPN			
Her2				High	1		
Negative	1			Low	1.26	0.77–2.08	0.356
Positive	2.31	1.32–4.04	0.003	CENPO			
Histological_type				High	1		
IDC	1			Low	1.37	0.83–2.24	0.219
ILC	0.83	0.42–1.65	0.595	CENPP			
Other	1.21	0.56–2.62	0.621	High	1		
Pathologic_T				Low	2.20	1.28–3.80	0.005
T1	1			CENPQ			
T2	1.03	0.55–1.93	0.932	High	1		
T3	1.49	0.68–3.29	0.320	Low	1.29	0.78–2.11	0.320
T4	6.15	2.67–14.16	0.000	CENPR			
Pathologic_N				High	1		
N0	1			Low	1.49	0.9–2.45	0.120
N1	1.22	0.67–2.2	0.520	CENPS			
N2	1.91	0.92–3.95	0.080	High	1		
N3	5.44	2.44–12.16	0.000	Low	1.43	0.85–2.38	0.176

Characteristics	Univariate analysis			Characteristics	Univariate analysis		
	HR	95%CI	P value		HR	95%CI	P value
Pathologic_M				CENPT			
M0	1			High	1		
M1	11.87	3.6-39.15	0.000	Low	0.95	0.58–1.55	0.830
Mx	1.22	0.48–3.06	0.677	CENPU			
CENPA				High	1		
High	1			Low	1.39	0.85–2.30	0.200
Low	1.51	0.92–2.48	0.102	CENPW			
CENPH				High	1		
High	1			Low	1.39	0.84–2.30	0.201
Low	1.51	0.92–2.48	0.106				

Table 3
Multivariate analysis of the training set.

Characteristics	Multivariate analysis		
	HR	95%CI	P value
Age			
< 65	1		
65–75	2.67	1.21–5.86	0.015
> 75	3.63	1.87–7.03	0.000
Her2			
Negative	1		
Positive	2.04	1.09–3.82	0.027
Pathologic_T			
T1	1		
T2	1.12	0.57–2.21	0.736
T3	2.13	0.91–5.04	0.083
T4	5.40	1.78–16.38	0.003
Pathologic_N			
N0	1		
N1	1.17	0.63–2.17	0.620
N2	1.10	0.47–2.58	0.820
N3	1.48	0.55–3.98	0.437
Pathologic_M			
M0	1		
M1	4.45	0.07–18.47	0.040
Mx	0.61	0.22–1.67	0.332
CENPP			
High	1		
Low	2.35	1.30–4.23	0.005

The expression pattern of CENPP in BC

Since CENPP expression was identified as the only independent prognostic factor in CENPs for BC patients, we analyzed its expression pattern through GEPIA and found that CENPP was overexpressed in BC tissues compared with normal tissues (Fig. 1a). Data from bcGenExMiner v4.4 showed that the CENPP mRNA level in Luminal A subtypes ranked the highest among all 5 subtypes classified by PAM50 (Fig. 1b). Further analysis showed that the expression of CENPP was higher in ER + or PR + tumors, whereas lower in Her2 + tumors (Fig. 1c-e). In order to investigate the correlation between CENPP expression and clinicopathological features, patients in the training set was divided into CENPP high and CENPP

low groups by the median expression of CENPP. CENPP high expression group had higher ER + and PR + ratio and had a lower death rate compared to CENPP low expression group (Table 4).

Table 4

Correlation between clinical characteristics and expression of CENPP in the training set.

Characteristics	High expression (N = 220)	Low expression (N = 260)	P
	No.(%)	No.(%)	value
Age			0.619
< 65	150 (68.2)	184 (70.8)	
> 75	32 (14.5)	30 (11.5)	
65–75	38 (17.3)	46 (17.7)	
Gender			0.837
Female	218 (99.1)	256 (98.5)	
Male	2 (0.9)	4 (1.5)	
ER			< 0.001
Negative	33 (15.0)	82 (31.5)	
Positive	187 (85.0)	178 (68.5)	
PR			< 0.001
Negative	50 (22.7)	113 (43.5)	
Positive	170 (77.3)	147 (56.5)	
Her2			0.676
Negative	172 (78.2)	198 (76.2)	
Positive	48 (21.8)	62 (23.8)	
Histological_type			0.08
IDC	151 (68.6)	200 (76.9)	
ILC	48 (21.8)	37 (14.2)	
Other	21 (9.5)	23 (8.8)	
Pathologic_T			0.189
T1	54 (24.5)	59 (22.7)	
T2	132 (60.0)	164 (63.1)	
T3	30 (13.6)	25 (9.6)	
T4	4 (1.8)	12 (4.6)	
Pathologic_N			0.548
N0	105 (47.7)	121 (46.5)	
N1	75 (34.1)	82 (31.5)	
N2	26 (11.8)	31 (11.9)	
N3	14 (6.4)	26 (10.0)	

Characteristics	High expression (N = 220)	Low expression (N = 260)	P value
	No.(%)	No.(%)	
Pathologic_M			0.179
M0	193 (87.7)	228 (87.7)	
M1	0 (0.0)	4 (1.5)	
Mx	27 (12.3)	28 (10.8)	
Status			0.003
Alive	202 (91.8)	214 (82.3)	
Dead	18 (8.2)	46 (17.7)	

Prognostic values of CENPP in BC

Kaplan-Meier plotter showed that higher CENPP expression was associated with better OS in BC patients (Fig. 2a). Regarding histological types, we concluded that higher expression of CENPP was associated with better OS in IDC or ILC (Fig. 2b-c). In addition, CENPP high expression indicated better OS in BC with ER + or PR+, whereas no significantly correlation with prognosis in ER- or PR- tumors (Fig. 2d-g). Moreover, regardless of Her2 status, higher expression of CENPP was associated with better OS (Fig. 2h-i).

Construction and validation of the nomogram for OS

Based on the results of the multivariate analysis, a nomogram was performed with independent prognostic predictors for BC including age, Her2, pathologic_T stage, pathologic_M stage and CENPP expression (Fig. 3). Different variables of each patient pointed to a score according to the top scale, then we could sum up all scores to get a total score. Based on the total score of the bottom scale, the 3-year and 5-year survival probabilities of BC could be evaluated. Next, the ROC curve was performed to validate the nomogram in the training set (Fig. 4a-b) and in the validation set (Fig. 4c-d). In the training set, the AUC of 3 – year survival and 5-year survival were 0.757 and 0.797, respectively. In the validation set, the AUC of 3 – year survival and 5 – year survival were 0.727 and 0.71, respectively. The calibration plot (Fig. 5a-d) suggested that the nomogram was well calibrated. These results suggested that this nomogram displayed good accuracy in predicting both 3-year and 5-year overall survival for patients with BC.

Discussion

In this study, data of CENPs expression and clinicopathological features of BC patients were downloaded and analyzed from the TCGA database, which aimed to discover more biological genes and molecular indexes to accurately predict the prognosis of BC. Based on the Cox regression analysis, we identified that BC patients with older age, Her2 positive, advanced T stage, advanced M stage, or low expression of CENPP were accompanied by worse OS. Next, high expression of CENPP was proved to be associated with better OS in ER + breast cancer or PR + breast cancer, whereas regardless of Her2 status, high expression of CENPP was associated with better OS. Finally, we constructed a nomogram on the basis of CENPP expression as well as other independent predictors. The 3-year and 5-year of AUCs in the training set were 0.757 and 0.797, and that of AUCs in the validation set were 0.727 and 0.71.

Cell mitosis is the process of transferring genetic information from the parent cell to the daughter cell. In the process of mitosis, CENPs not only provided energy for the separation of sister chromatids, but also served as a genomic information monitoring function. Once this process loses normal regulation or makes mistakes, it may induce the

occurrence of malignant tumors[22, 23]. CENPs play a pivotal role in maintaining normal mitosis in cells. In this study, CENPP was identified as the only prognosis-related gene in CENPs for patients with BC. Some studies reported that CENPP was associated with mixed uterine carcinosarcoma and among its related pathways are mitotic metaphase, anaphase and signaling by G protein coupled receptor (GPCR)^[24–26]. However, its role in BC is unknown. To our knowledge, this is the first study to investigate the prognostic value of CENPP in BC.

Hormone receptor status plays a key role in the formation and development of BC. ER and PR status as an important biological indicator of choosing treatment schemes have been widely recognized and accepted in BC patients. It has been clinically confirmed that ER + is different from ER- in predicting the clinical efficacy and survival of BC. Compared with ER- BC, the tumor differentiation of ER + BC is better, the invasiveness is lower, and the long-term survival rate is higher. According to the data of the American Registry of Cancer Research, 20% of patients with ER + breast cancer were PR-[27]. Studies have shown that ER + PR- is a more invasive subtype of ER + breast cancer[28, 29]. The overall survival and disease-free survival of ER + PR- breast cancer was lower than that of ER + PR + breast cancer. Purdie C A et al. believed that PR was an independent predictor of early breast cancer prognosis^[30]. Our study found that the mRNA level of CENPP was positively associated with ER status and PR status, and higher mRNA expression of CENPP indicated better OS in BC with ER + or PR+, which suggests a close relationship between CENPP and hormone receptor pathway and the underlying mechanism requires further investigation. In addition, recent studies have shown that the overexpression of Her2 did not only indicate invasiveness and poor prognosis in BC, but also predict the sensitivity of BC to systemic treatment. The result of our study showed that no matter the Her2 status was positive or negative, high expression of CENPP was associated with better OS which indicate an inconsequential association between CENPP and HER2 pathway. Collectively, these findings suggested that CENPP was an effective prognostic predictor of BC and might also be a potential target for HR + BC.

Currently, the nomogram has been developed and shown to be more accurate in predicting prognosis in some cancers than the conventional staging systems^[31–33]. This study attempted to establish a prognostic nomogram of BC and to determine whether the model can accurately predict survival of patients with BC. Age, Her2 status, pathologic_T stage, pathologic_M stage and CENPP expression were identified as predictive factors for OS of BC in the multivariate Cox analysis. As we known, these variables including age, Her2 status, pathologic_T stage, pathologic_M stage were the conventional prognosis predictors in BC. This is the first study to set up a nomogram based on the CENPP expression and conventional prognosis predictors to predict OS in patients with BC. Through validated internally in the training set and externally in the validation set, the nomogram performed well in predicting survival, and its prediction was supported by the ROC curves and the calibration curves. When a patient is diagnosed with BC and has obtained the above clinicopathological results, we can predict his clinical prognosis according to her own features and CENPP expression level. If the predicted prognosis is poor, intensive treatment might be recommended in the hope of gaining a better outcome.

There're some limitations of this study. Firstly, this was a retrospective study, all the data of this study were obtained from publicly available databases. More prospective studies are needed to further confirm our conclusions. Secondly, the demographic and clinical information provided by the TCGA database were not complete. For example, the database lacked detailed treatment records like surgery, marital status and insurance status information. Different surgical approaches, marital status and insurance status may influenced outcomes of BC patients. Despite this limitation, these researches provided more deep insights into possible molecular mechanisms of prognosis for BC.

Conclusions

In this study, we identified that higher expression of CENPP is associated with better prognosis, and established a prognostic nomogram with good performance based on CENPP expression and clinicopathological features. Our study

provided a novel method for clinical evaluation and a potential biomarker/target for BC which will be further validated in a prospective study and investigated in basic researches.

Declarations

Acknowledgement

We thank The Cancer Genome Atlas (TCGA) Database for sharing the large amount of data. The authors also thank Natural Science Foundation of Shaanxi Province, China (No. 2018JQ8004) for funding to do this work.

Funding

This work was supported by grants from the Natural Science Foundation of Shaanxi Province, China (No. 2018JQ8004).

Authors contribution statement

Conceptualization, Heyan Chen and Huimin Zhang; Data curation, Heyan Chen, Shengyu Pu and Xiaoqin Liao; Formal analysis, Heyan Chen, Shengyu Pu, Xiaoqin Liao, Jianjun He and Huimin Zhang; Funding acquisition, Huimin Zhang; Methodology, Heyan Chen, Shibo Yu and Huimin Zhang; Software, Shibo Yu; Supervision, Jianjun He; Writing – original draft, Heyan Chen; Writing – review & editing, Huimin Zhang.

Conflict of interest

The authors declare no conflict of interest.

References

1. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *Cancer J Clin.* 2015;65(2):87–108.
2. Chittori S, Hong J, Saunders H, Feng H, Ghirlando R, Kelly AE, Bai Y, Subramaniam S. Structural mechanisms of centromeric nucleosome recognition by the kinetochore protein CENP-N. *Science.* 2018;359(6373):339–43.
3. Hoischen C, Yavas S, Wohland T, Diekmann S. CENP-C/H/I/K/M/T/W/N/L and hMis12 but not CENP-S/X participate in complex formation in the nucleoplasm of living human interphase cells outside centromeres. *PLoS One.* 2018;13(3):e0192572.
4. Lera RF, Norman RX, Dumont M, Dennee A, Martin-Koob J, Fachinetti D, Burkard ME. Plk1 protects kinetochore-centromere architecture against microtubule pulling forces. *EMBO Rep.* 2019;20(10):e48711.
5. Logsdon GA, Gambogi CW, Liskovych MA, Barrey EJ, Larionov V, Miga KH, Heun P, Black BE. Human Artificial Chromosomes that Bypass Centromeric DNA. *Cell.* 2019;178(3):624–39 e619.
6. Nechemia-Arbely Y, Miga KH, Shoshani O, Aslanian A, McMahon MA, Lee AY, Fachinetti D, Yates JR 3rd, Ren B, Cleveland DW. DNA replication acts as an error correction mechanism to maintain centromere identity by restricting CENP-A to centromeres. *Nat Cell Biol.* 2019;21(6):743–54.
7. McKinley KL, Sekulic N, Guo LY, Tsinman T, Black BE, Cheeseman IM. The CENP-L-N Complex Forms a Critical Node in an Integrated Meshwork of Interactions at the Centromere-Kinetochore Interface. *Mol Cell.* 2015;60(6):886–98.
8. Beh TT, Kalitsis P. The Role of Centromere Defects in Cancer. *Prog Mol Subcell Biol.* 2017;56:541–54.
9. Komatsu M, Yoshimaru T, Matsuo T, Kiyotani K, Miyoshi Y, Tanahashi T, Rokutan K, Yamaguchi R, Saito A, Imoto S, et al. Molecular features of triple negative breast cancer cells by genome-wide gene expression profiling analysis. *Int J Oncol.* 2013;42(2):478–506.

10. Liao WT, Feng Y, Li ML, Liu GL, Li MZ, Zeng MS, Song LB. Overexpression of centromere protein H is significantly associated with breast cancer progression and overall patient survival. *Chin J Cancer*. 2011;30(9):627–37.
11. Lin SY, Lv YB, Mao GX, Chen XJ, Peng F. The effect of centromere protein U silencing by lentiviral mediated RNA interference on the proliferation and apoptosis of breast cancer. *Oncol Lett*. 2018;16(5):6721–8.
12. McGovern SL, Qi Y, Pusztai L, Symmans WF, Buchholz TA. Centromere protein-A, an essential centromere protein, is a prognostic marker for relapse in estrogen receptor-positive breast cancer. *Breast Cancer Res*. 2012;14(3):R72.
13. Rajput AB, Hu N, Varma S, Chen CH, Ding K, Park PC, Chapman JA, Sengupta SK, Madarnas Y, Elliott BE, et al. Immunohistochemical Assessment of Expression of Centromere Protein-A (CENPA) in Human Invasive Breast Cancer. *Cancers (Basel)*. 2011;3(4):4212–27.
14. Thangavelu PU, Lin CY, Vaidyanathan S, Nguyen THM, Dray E, Duijf PHG. Overexpression of the E2F target gene CENPI promotes chromosome instability and predicts poor prognosis in estrogen receptor-positive breast cancer. *Oncotarget*. 2017;8(37):62167–82.
15. Wang X, Chen D, Gao J, Long H, Zha H, Zhang A, Shu C, Zhou L, Yang F, Zhu B, et al. Centromere protein U expression promotes non-small-cell lung cancer cell proliferation through FOXM1 and predicts poor survival. *Cancer Manag Res*. 2018;10:6971–84.
16. Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, Li B, Liu XS. TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. *Cancer Res*. 2017;77(21):e108–10.
17. Cao Y, Xiong J, Li Z, Zhang G, Tu Y, Wang L, Jie Z. CENPO expression regulates gastric cancer cell proliferation and is associated with poor patient prognosis. *Mol Med Rep*. 2019;20(4):3661–70.
18. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res*. 2017;45(W1):W98–102.
19. Jezequel P, Campone M, Gouraud W, Guerin-Charbonnel C, Leux C, Ricolleau G, Campion L. bc-GenExMiner: an easy-to-use online platform for gene prognostic analyses in breast cancer. *Breast Cancer Res Treat*. 2012;131(3):765–75.
20. Jezequel P, Frenel JS, Campion L, Guerin-Charbonnel C, Gouraud W, Ricolleau G, Campone M: **bc-GenExMiner 3.0: new mining module computes breast cancer gene expression correlation analyses**. *Database (Oxford)* 2013, 2013:bas060.
21. Nunez E, Steyerberg EW, Nunez J. [Regression modeling strategies]. *Rev Esp Cardiol*. 2011;64(6):501–7.
22. Beaver JA, Gustin JP, Yi KH, Rajpurohit A, Thomas M, Gilbert SF, Rosen DM, Ho Park B, Lauring J. PIK3CA and AKT1 mutations have distinct effects on sensitivity to targeted pathway inhibitors in an isogenic luminal breast cancer model system. *Clin Cancer Res*. 2013;19(19):5413–22.
23. Lee H, Kim SJ, Jung KH, Son MK, Yan HH, Hong S, Hong SS. A novel imidazopyridine PI3K inhibitor with anticancer activity in non-small cell lung cancer cells. *Oncol Rep*. 2013;30(2):863–9.
24. Eskat A, Deng W, Hofmeister A, Rudolphi S, Emmerth S, Hellwig D, Ulbricht T, Doring V, Bancroft JM, McAinsh AD, et al. Step-wise assembly, maturation and dynamic behavior of the human CENP-P/O/R/Q/U kinetochore sub-complex. *PLoS One*. 2012;7(9):e44717.
25. Foltz DR, Jansen LE, Black BE, Bailey AO, Yates JR 3rd, Cleveland DW. The human CENP-A centromeric nucleosome-associated complex. *Nat Cell Biol*. 2006;8(5):458–69.
26. Okada M, Cheeseman IM, Hori T, Okawa K, McLeod IX, Yates JR 3rd, Desai A, Fukagawa T. The CENP-H-I complex is required for the efficient incorporation of newly synthesized CENP-A into centromeres. *Nat Cell Biol*. 2006;8(5):446–57.
27. Dunnwald LK, Rossing MA, Li CI. Hormone receptor status, tumor characteristics, and prognosis: a prospective cohort of breast cancer patients. *Breast Cancer Res*. 2007;9(1):R6.

28. Canello G, Maisonneuve P, Rotmensz N, Viale G, Mastropasqua MG, Pruneri G, Montagna E, Iorfida M, Mazza M, Balduzzi A, et al. Progesterone receptor loss identifies Luminal B breast cancer subgroups at higher risk of relapse. *Ann Oncol*. 2013;24(3):661–8.
29. Prat A, Cheang MC, Martin M, Parker JS, Carrasco E, Caballero R, Tyldesley S, Gelmon K, Bernard PS, Nielsen TO, et al. Prognostic significance of progesterone receptor-positive tumor cells within immunohistochemically defined luminal A breast cancer. *J Clin Oncol*. 2013;31(2):203–9.
30. Purdie CA, Quinlan P, Jordan LB, Ashfield A, Ogston S, Dewar JA, Thompson AM. Progesterone receptor expression is an independent prognostic variable in early breast cancer: a population-based study. *Br J Cancer*. 2014;110(3):565–72.
31. International Bladder Cancer. Nomogram C, Bochner BH, Kattan MW, Vora KC. Postoperative nomogram predicting risk of recurrence after radical cystectomy for bladder cancer. *J Clin Oncol* 2006, 24(24):3967–72.
32. Karakiewicz PI, Briganti A, Chun FK, Trinh QD, Perrotte P, Ficarra V, Cindolo L, De la Taille A, Tostain J, Mulders PF, et al. Multi-institutional validation of a new renal cancer-specific survival nomogram. *J Clin Oncol*. 2007;25(11):1316–22.
33. Touijer K, Scardino PT. Nomograms for staging, prognosis, and predicting treatment outcomes. *Cancer*. 2009;115(13 Suppl):3107–11.

Figures

Figure 1

The expression level and overall survival of CENPP in BC. (a) The expression level of CENPP between BC tissues and normal breast in GEPIA database. (b-e) The expression level of CENPP in different molecular subtypes (b), ER status (c), PR status (d) and HER2 status (e) (bcGenExMiner v4.4 database). ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor.

Figure 2

Kaplan-Meier survival curve of overall survival between the high-expression group and low-expression group of CENPP in BC. In all patients (a), IDC patients (b), ILC patients (c), ER+ patients (d), ER- patients (e), PR+ patients (f), PR- patients (g), Her2+ patients (h) and Her2- patients (i).

Figure 3

A nomogram to predict the prognosis of breast cancer.

Figure 4

Validation of the nomogram by the ROC curves. The ROC curve for predicting patient survival at 3 years (a) and 5 years (b) in the training set, and at 3 years (c) and 5 years (d) in the validation set. ROC, receiver-operating characteristic; AUC, area under the curve. TP, true positive rate; FP, false positive rate.

Figure 5

Validation of the nomogram by the calibration curves. The calibration curve for predicting patient survival at 3 years (a) and 5 years (b) in the training set, and at 3 years (c) and 5years (d) in the validation set. Nomogram-predicted probability of overall survival is plotted on the x-axis; actual overall survival is plotted on the y-axis.