

High SURF4 expression is associated with poor prognosis of breast cancer

Jingtong Zhai¹, Jiashu Han², Cong Li¹, Fengzhu Guo¹, Fei Ma¹, Binghe Xu^{1,&}

¹Department of Medical Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China

²4+4 Medical Doctor Program, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100730, China

Correspondence to: Fei Ma, Binghe Xu; email: drmafei@126.com, <https://orcid.org/0000-0001-9432-1902>; xubinghebm@163.com, <https://orcid.org/0000-0002-0234-2747>

Keywords: SURF4, breast cancer, prognostic biomarker, overall survival, relapse-free survival

Received: September 5, 2022

Accepted: November 21, 2022

Published: November 29, 2022

Copyright: © 2022 Zhai et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/3.0/) (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

SURF4 has been suggested as an oncogene in cancer. However, the role of SURF4 in breast cancer has not been demonstrated yet. The data were obtained from TCGA database and 1104 patients were analyzed using bioinformatics analysis. SURF4 is significantly ($P < 0.001$) highly expressed in tumor. High expression of SURF4 was observed in T4, infiltrating ductal carcinoma, ER negative, PR negative, and HER2 positive, female, patients without lymph node metastasis, HER2 overexpression type, and deceased patients. As for characteristics correlated with high expression of SURF4, gender, histological type, molecular subtype, ER, PR, HER2, and vital status exhibited significant differences. The age (HR: 2.317, $P < 0.001$), stage (HR: 2.090, $P < 0.001$), and SURF4 expression (HR: 1.958, $P = 0.005$) exhibited independent prognostic value for overall survival (OS). Patients with high SURF4 expression, higher age, equivocal HER2, higher stages, or positive margin status had shorter OS. The stage (HR: 1.579, $P < 0.001$), and margin status (HR: 1.463, $P = 0.006$) exhibited independent prognostic value for relapse-free survival of breast cancer. High expression of SURF4 was first found in breast cancer. High SURF4 expression was observed in breast cancer tissue and cell. SURF4 promoted the proliferation and migration of 4T1 cells. SURF4 may be a biomarker in diagnosis and prognosis of breast cancer.

INTRODUCTION

Breast cancer is the top malignant tumor that threatens the health of women worldwide with incidence rate of 47.8% and mortality rate of 13.6% [1]. In the past two decades, people have gradually realized that breast cancer is heterogeneous, varying in pathology, genetics, and molecular biology [2, 3]. Breast cancer consists of luminal A, luminal B, human epidermal growth factor receptor 2 (HER2)-enriched, and basal-like molecular subtypes relevant to the clinical practice [4]. Under the guidance of molecular typing, the diagnosis and treatment of breast cancer has gradually entered the era of individualized diagnosis and treatment [5]. The 5-year survival rate and the quality of life of breast cancer patients have been improved [6]. However, some patients

will progress to an advanced stage, and advanced breast cancer is still an incurable disease with a median survival of about 24 to 30 months [7, 8].

Early diagnosis has become the key to save the lives of breast cancer patients. Currently, tissue- and serum-based biomarkers are widely used for early-stage tumor screening and to predict disease progression or recurrence [9]. Commonly used biomarkers for breast cancer include estrogen receptor (ER), progesterone receptor (PR), and HER2 [10]. Since the occurrence and progression are heterogeneous, researchers strive to identify novel biomarkers for optimal breast cancer control [11].

Bioinformatic analysis has contributed much to the finding of novel biomarkers, especially in cancer [12].

Kim et al. have suggested SURF4 as an oncogene in cancer [13]. SURF4 is the human homologous gene of the yeast cargo receptor Erv29p with a molecular size of 30 kD [14, 15]. SURF4 contains 7 transmembrane domains and a double-lysine endoplasmic reticulum carboxyl-terminal sequence, which can interact with the ER Golgi Intermediate compartment-53 and p24 protein [16]. However, the role of SURF4 in breast cancer has not been demonstrated yet.

In this study, high expression of SURF4 was first found in breast cancer. Their relationship was further evaluated using bioinformatics analysis. The diagnostic value and independent predictive value in overall survival (OS) and relapse-free survival (RFS) were evaluated. Besides, the nomogram and Gene set enrichment analysis (GSEA) were conducted. The cell and tissue experiments were carried out for final validation.

RESULTS

Patient characteristics

In total, 1104 patients were analyzed, 1090 of which (98.73%) were females. As shown in Supplementary Table 1, 589 (53.45%) patients were less than 60 years old, and 790 (71.56%) were diagnosed of infiltrating ductal carcinoma. As for molecular subtype, 142 (12.86%) patients were basal-like type, 67 (6.07%) patients were HER2 overexpression type, 422 (38.22%)

patients were luminal A type, 194 (17.57%) patients were luminal B type, and 24 (2.17%) patients were normal-like type. Most patients (56.7%) were in stage II.

High expression of SURF4 in tumor

Compared with normal tissue, SURF4 is significantly ($P < 0.001$) highly expressed in tumor (Figure 1A). Besides, SURF4 is significantly ($P < 0.001$) highly expressed in breast cancer in comparison with paired normal breast tissue (Figure 1B).

Furthermore, the expression of SURF4 grouped by distinctive characteristics were evaluated. As shown in Figure 2A–2H, the expression of SURF4 was highest in T4 ($P = 0.014$), infiltrating ductal carcinoma ($P < 0.001$), ER negative ($P < 0.001$), PR negative ($P < 0.001$), and HER2 positive ($P < 0.001$) patients. Besides, as shown in Figure 3A–3K, high expression of SURF4 was observed in female ($P = 0.026$), patients without lymph node metastasis ($P = 0.037$), HER2 overexpression type ($P < 0.001$), and deceased patients ($P = 0.016$). Other characteristics showed no significantly statistical differences.

Diagnostic value of SURF4 expression for breast cancer

The receiver operating characteristic (ROC) curves were plotted to evaluate the diagnostic value of SURF4

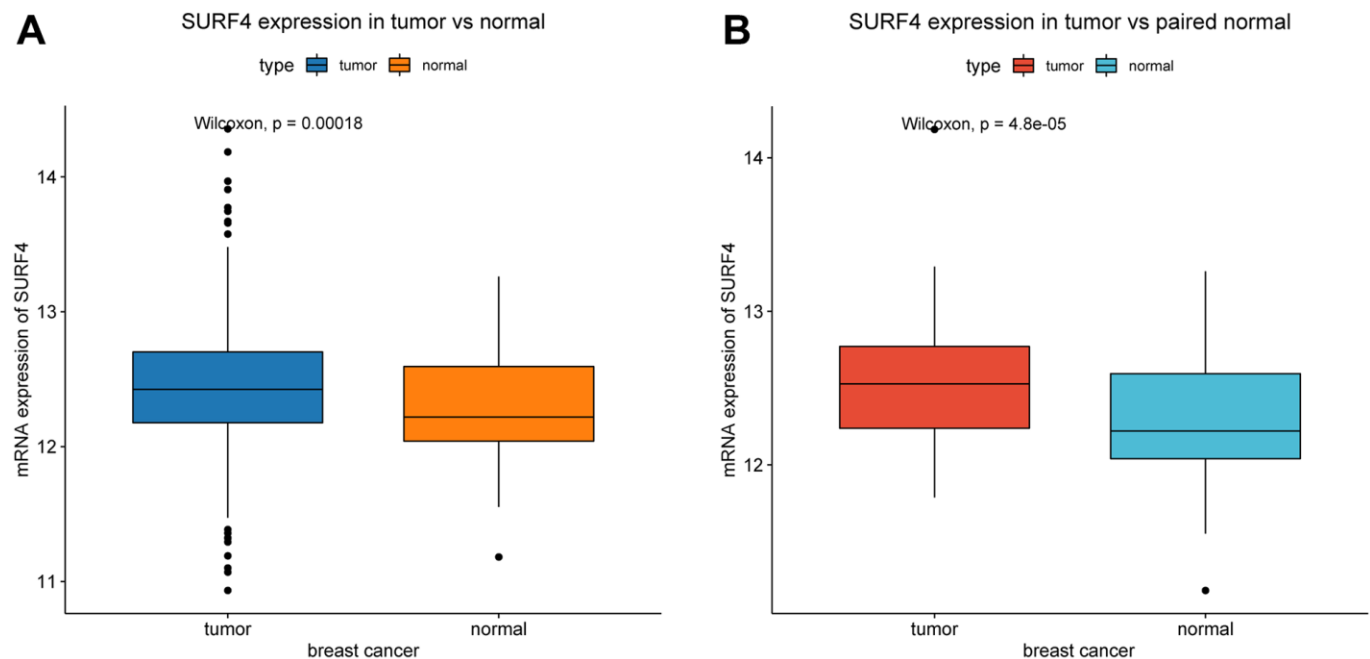


Figure 1. Expression of SURF4 in breast cancer. (A) The expression of SURF4 in tumor and normal tissues. (B) The expression of SURF4 in tumor and paired normal tissues.

expression for breast cancer. As shown in Supplementary Figure 1A, the cut-off value between normal and tumor was 12.241, and the AUC was 0.606. Besides, the AUC for stage I-IV (Supplementary Figure 1B–1E) was 0.581, 0.609, 0.613, and 0.693, respectively.

Characteristics correlated with high expression of SURF4

First, the patients were divided into the low and high SURF4 expression groups according to the cut-off value between normal and tumor obtained from ROC curves. Then, the characteristics correlated with high SURF4 expression were studied. As shown in Supplementary Table 2, gender ($P = 0.032$), histological type ($P < 0.001$), molecular subtype ($P < 0.001$), ER ($P < 0.001$), PR ($P < 0.001$), HER2 ($P < 0.001$), and vital status ($P = 0.002$) exhibited significant differences. However, other characteristics showed no significant differences.

High SURF4 expression is correlated with poor OS

Kaplan–Meier curves were used to evaluate the OS correlated with SURF4 expression. As shown in Figure 4A, high SURF4 expression was significantly correlated with poor OS ($P < 0.001$). Moreover, high SURF4 expression was correlated with poor OS of breast cancer patients in infiltrating ductal carcinoma ($P < 0.001$),

infiltrating lobular carcinoma ($P = 0.018$), luminal B type ($P = 0.019$), normal-like type ($P = 0.016$), ER negative ($P = 0.003$), PR negative ($P < 0.001$), and HER2 negative ($P < 0.001$) by subgroup analysis (Figure 4B–4N).

As shown in Figure 5 and Supplementary Table 3, several important variables were identified by univariate analysis, and confirmed by the subsequent multivariate analysis. The age [hazard ratio (HR): 2.317, 95% confidence interval (CI): 1.452–3.699, $P < 0.001$], stage (HR: 2.090, 95% CI: 1.585–2.755, $P < 0.001$), and SURF4 expression (HR: 1.958, 95% CI: 1.230–3.115, $P = 0.005$) exhibited independent prognostic value for OS of breast cancer.

High SURF4 expression is correlated with poor RFS

As shown in Figure 6A, high SURF4 expression was significantly correlated with poor RFS ($P = 0.005$). Moreover, high SURF4 expression was correlated with poor RFS of breast cancer patients in infiltrating ductal carcinoma ($P = 0.018$), ER negative ($P = 0.019$), PR negative ($P = 0.008$) by subgroup analysis (Figure 6B–6N).

As shown in Figure 7 and Supplementary Table 4, several important variables were identified by univariate analysis, and confirmed by the subsequent multivariate

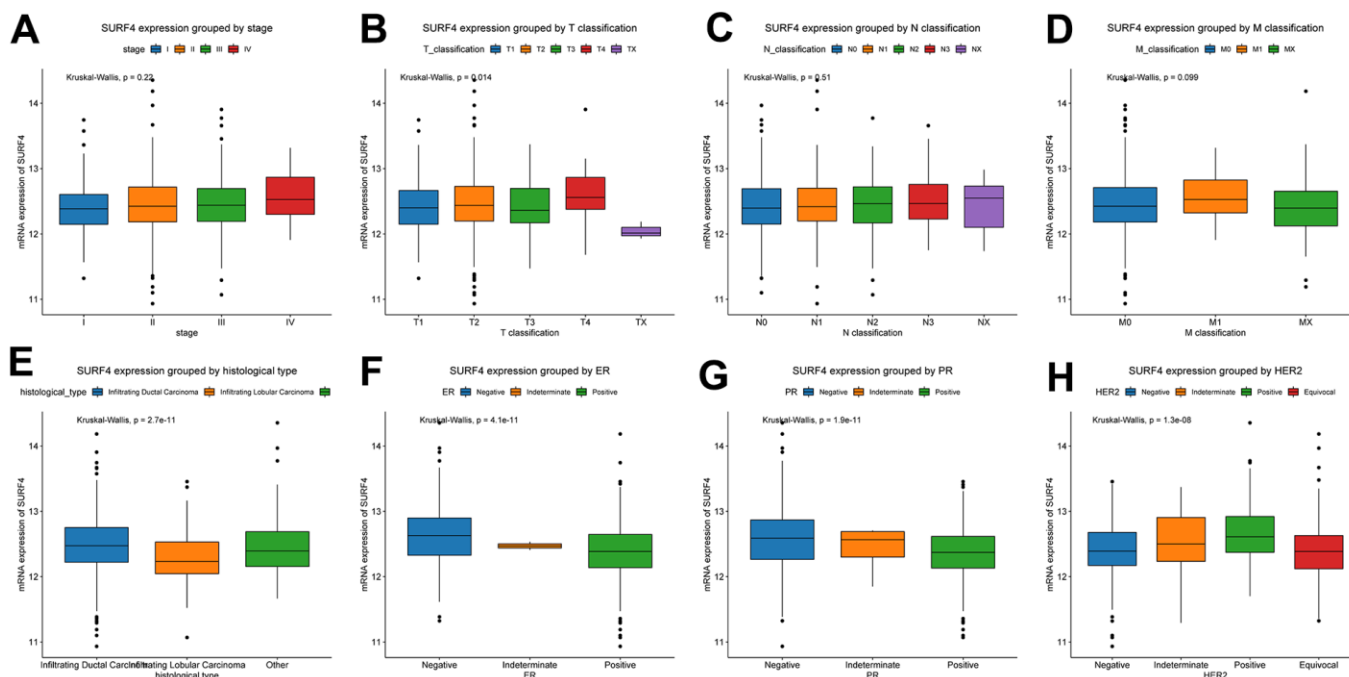


Figure 2. Expression of SURF4 in breast cancer and its relationship with clinicopathological parameters. Expression of SURF4 grouped by (A) stage, (B) T classification, (C) N classification, (D) M classification, (E) histological type, (F) ER, (G) PR, and (H) HER2.

analysis. The stage (HR: 1.579, 95% CI: 1.266-1.970, $P < 0.001$), and margin status (HR: 1.463, 95% CI: 1.114-1.921, $P = 0.006$) exhibited independent prognostic value for RFS of breast cancer.

Predictive value of SURF4 for survival

The nomograms of SURF4 were generated to predicate survival of breast cancer patients. Patients with high SURF4 expression showed shorter OS (Figure 8). Besides, the patients with higher age, equivocal HER2, higher stages, or positive margin status also had shorter OS.

SURF4 related signaling pathways by GSEA

As shown in Supplementary Table 5, the GSEA analysis was further performed to evaluate the SURF4 related

signaling pathways. The top 5 signaling pathways were shown in Figure 9A. The results showed that SURF4 may influence the development of breast cancer by controlling chemokine signaling pathway, Th17 cell differentiation, primary immunodeficiency, etc.

Expression of SURF4 in TIMER database

The expressions of SURF4 in different cancers were further studied using TIMER database (Figure 9B). Significant high SURF4 expression was found in bladder urothelial carcinoma, etc.

High SURF4 expression in breast cancer tissue and cell

As shown in Figure 10A, SURF4 expression was significantly higher in tumor than adjacent normal tissue

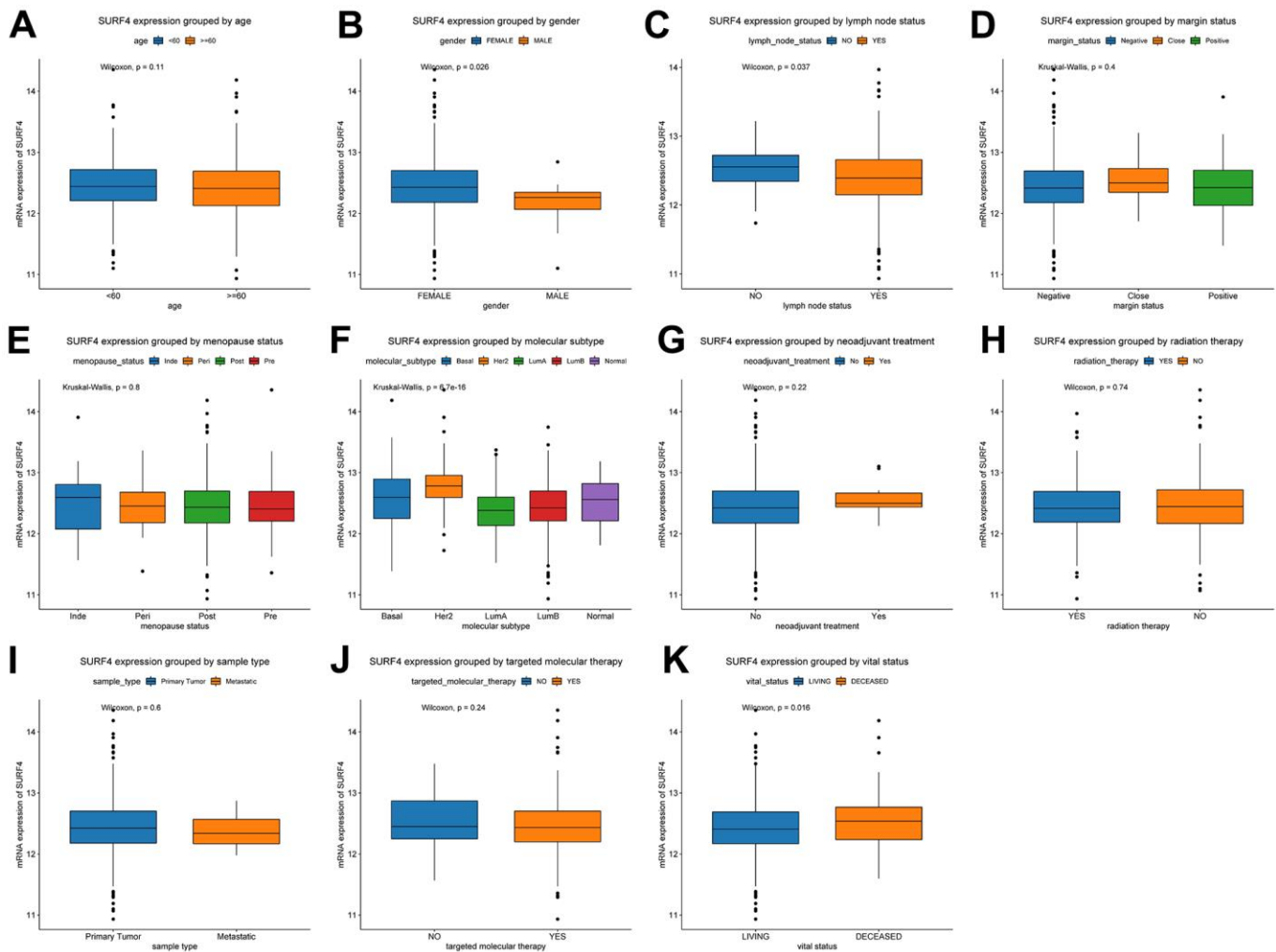


Figure 3. Expression of SURF4 in breast cancer and its relationship with clinicopathological parameters. Expression of SURF4 grouped by (A) age, (B) gender, (C) lymph node status, (D) margin status, (E) menopause status, (F) molecular subtype, (G) neoadjuvant treatment, (H) radiation therapy, (I) sample type, (J) targeted molecular therapy, and (K) vital status.

($P < 0.001$). As shown in Figure 10B, significant high SURF4 expression ($P < 0.01$) was observed in MCF7 (Human luminal A type breast cancer cell), BT474 (Human luminal B type breast cancer cell), SKBR3 (Human HER2 overexpression type breast cancer cell), MDAMB231 and 4T1 (Human triple-negative breast cancer cell). Meanwhile, HaCaT (Human

immortalized epidermal cell) and MCF10A (Human normal breast cell) showed much lower SURF4 expression. Importantly, 4T1 cell line showed the highest SURF4 expression and used subsequently for *in vitro* experiment. The high SURF4 expression in breast cancer tissue was further validated by immunohistochemistry (IHC) (Figure 10C).

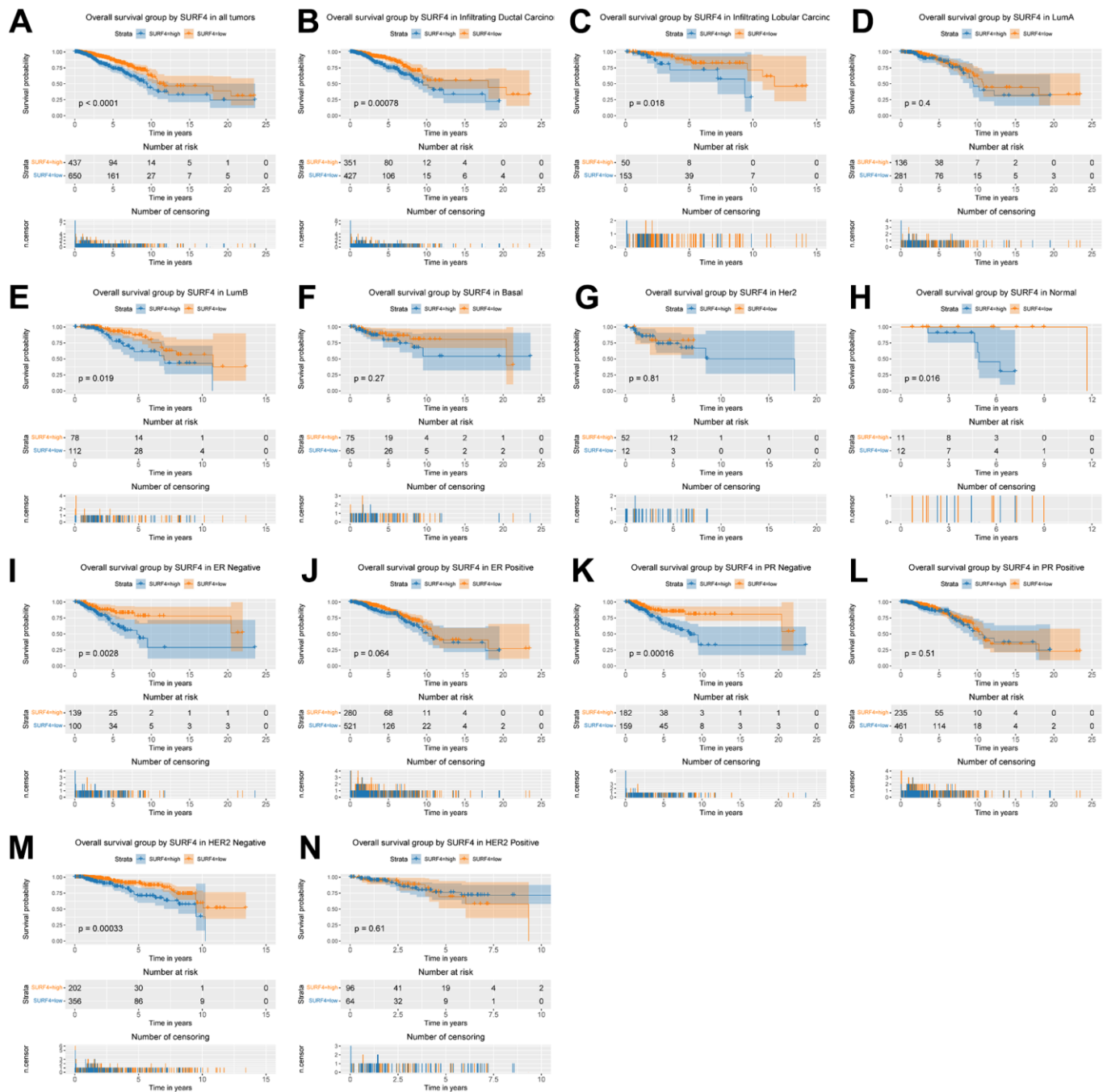


Figure 4. Kaplan-Meier curve for OS group by SURF4. (A) SURF4 in all tumors. (B–H) Subgroup analysis according to typing in infiltrating ductal carcinoma, infiltrating lobular carcinoma, Luminal A, Luminal B, basal-like, HER2 overexpression, normal-like. (I–N) Subgroup analysis according to status in ER negative, ER positive, PR negative, PR positive, HER2 negative, and HER2 positive.

SURF4 promoted the proliferation and migration of 4T1 cells

The successful over-expression (O-SURF4) and knockdown (si-SURF4) of SURF4 was shown in Figure 11A. As shown in Figure 11B, si-SURF4 group showed decreased cell proliferation ($P < 0.01$), while O-SURF4 group showed increased cell proliferation ($P < 0.01$). The results of colony formation (Figure 11C) and living/dead cell staining (Figure 11D, 11E) were consistent. Si-SURF4 group showed fewer colonies ($P < 0.05$) and more dead cells ($P < 0.01$), while O-SURF4

group showed more colonies ($P < 0.05$) and fewer dead cells ($P < 0.05$). Finally, the migration distance was shorter in si-SURF4 group ($P < 0.05$) and longer in O-SURF4 group (Figure 11F, 11G, $P < 0.01$). The results indicated that SURF4 promoted the proliferation and migration of 4T1 cells.

DISCUSSION

In 2020, the incidence of breast cancer reached 11.7%, surpassing lung cancer for the first time to become the world's most commonly diagnosed cancer. Continuous

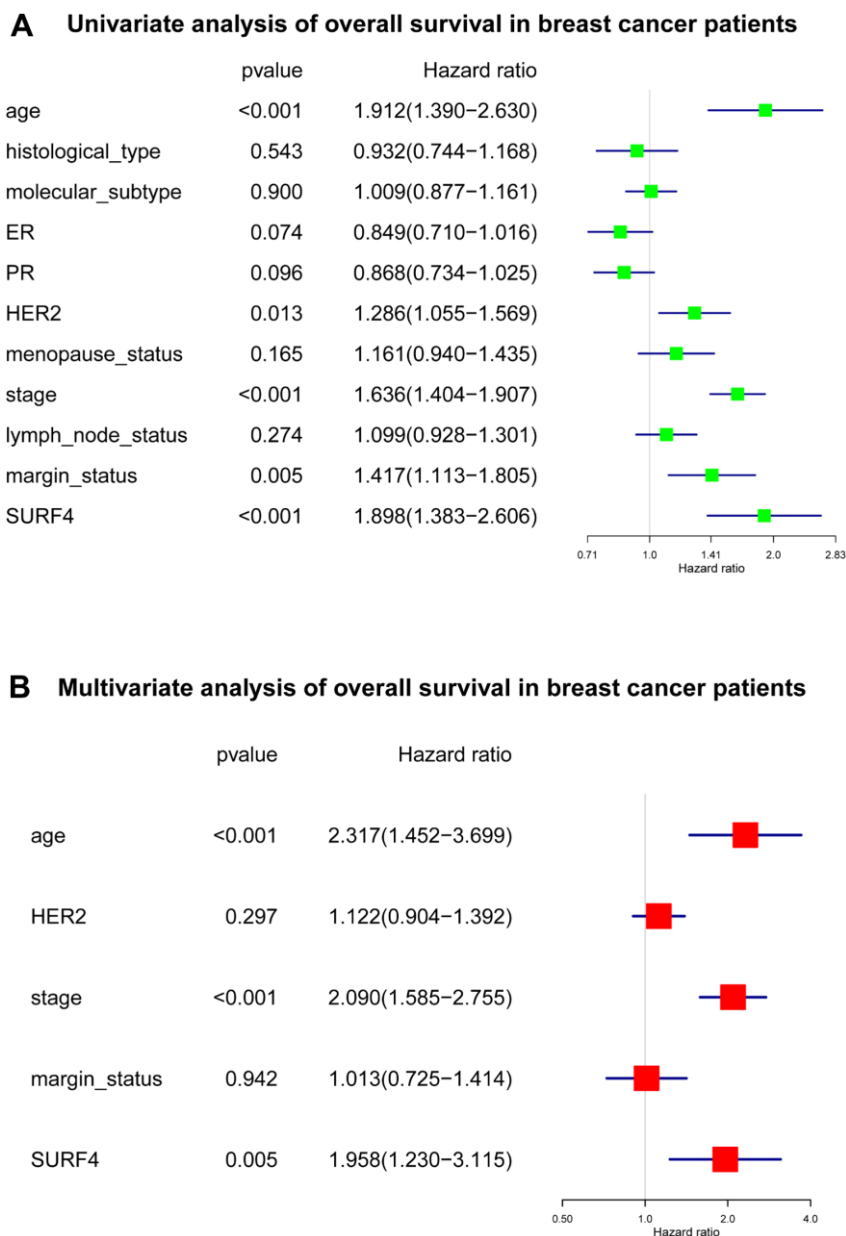


Figure 5. Forest plot of Cox regression analysis about SURF4 and OS. (A) Univariate analysis of OS in breast cancer patients. **(B)** Multivariate analysis of OS in breast cancer patients.

exploration of earlier diagnosis methods has become the most urgent task [17]. Breast cancer is a very heterogeneous disease [18]. It has become widely accepted that breast cancer can be classified according to molecular markers. In 2001, Peru et al. applied complementary DNA microarray technology to detect postoperative specimens, and divided breast cancer

into different subtypes with different clinical prognosis, namely luminal A, luminal B, basal-like, HER2 overexpression, and normal-like type [19].

Diverse types of breast cancer have different biological characteristics and treatment methods. Hormone receptor-positive luminal A and luminal B have the

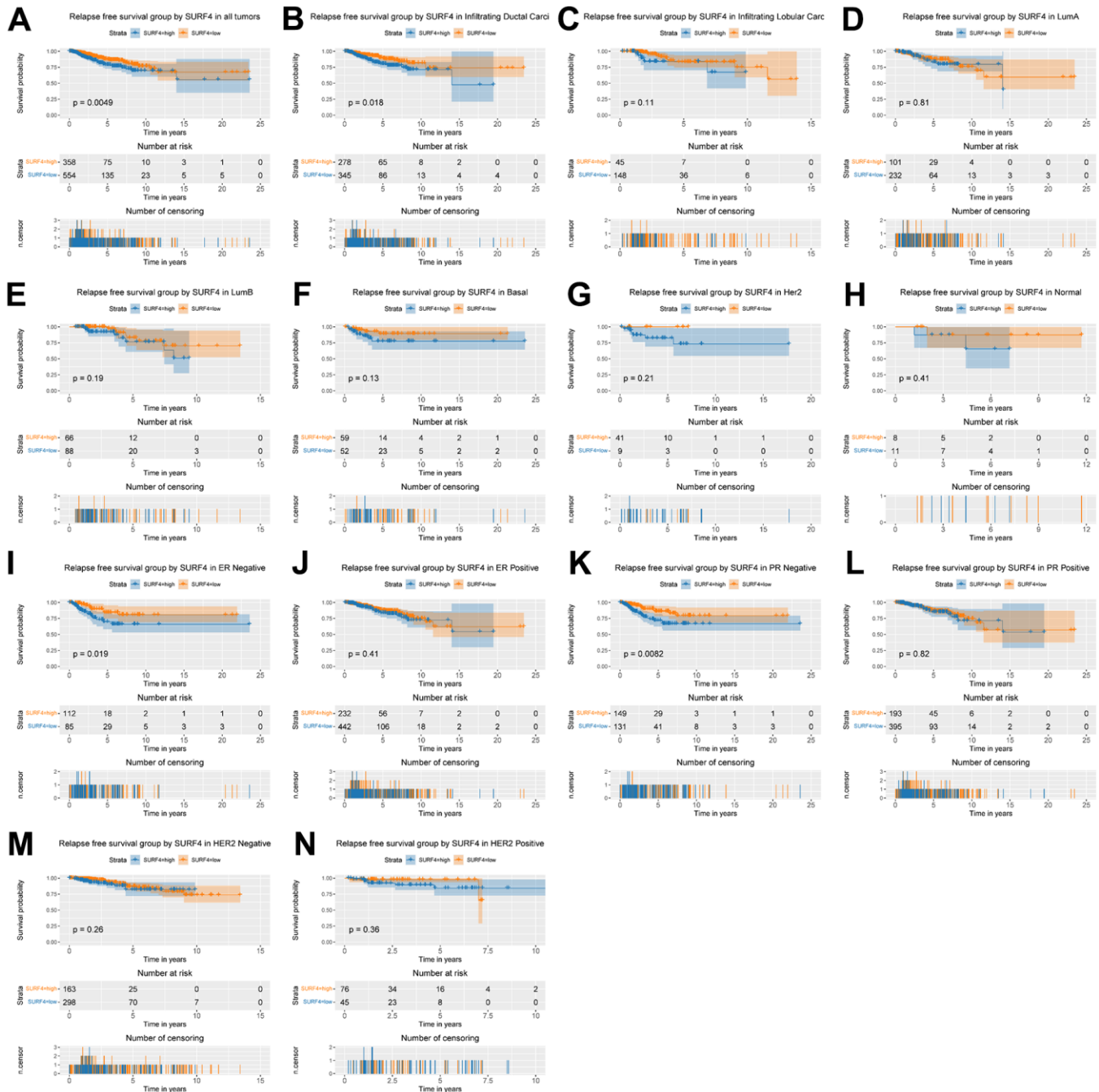


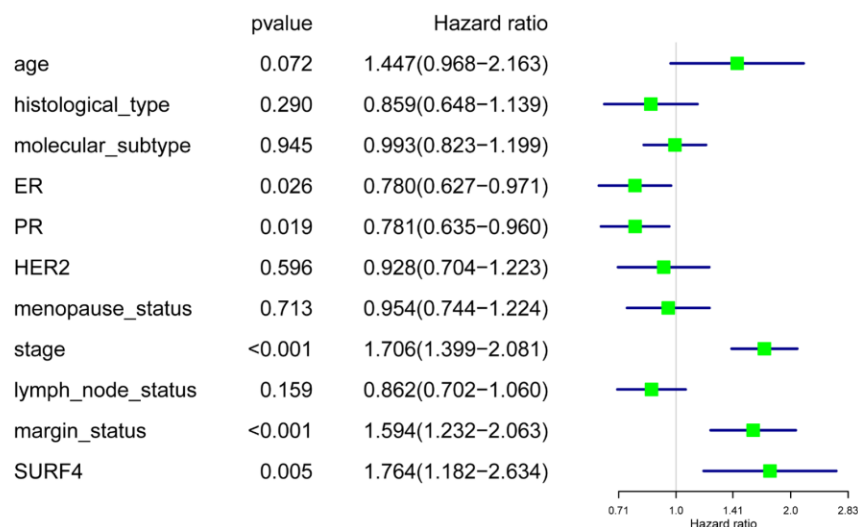
Figure 6. Kaplan-Meier curve for RFS group by SURF4. (A) SURF4 in all tumors. (B–H) Subgroup analysis according to typing in infiltrating ductal carcinoma, infiltrating lobular carcinoma, Luminal A, Luminal B, basal-like, HER2 overexpression, normal-like. (I–N) Subgroup analysis according to status in ER negative, ER positive, PR negative, PR positive, HER2 negative, and HER2 positive.

highest proportions in breast cancer, and can be treated with endoprostheses with relatively good prognosis [20]. The 5-year selective ER modulator tamoxifen therapy can reduce the recurrence rate by about 50% and the mortality by about 30% [21]. Compared with luminal A type, luminal B type has the following characteristics: lower expression levels of estrogen or estrogen-related genes, low or no expression of PR, higher grade of tumor and higher expression levels of proliferation-related genes, growth factor receptor pathways such as IGF-1R and PI3K/AKT/mTOR are easily activated, etc. [22]. HER2 overexpression type has poor biological behavior

and high recurrence rate, but its treatment has made significant progress with the application of targeted drugs such as trastuzumab, pertuzumab, and lapatinib, etc. [5, 23]. The basal-like type lacking hormone receptors and overexpressing HER2 is often considered triple-negative breast cancer, which can only be treated with chemotherapy, and is more common in patients with *BRCA1* mutations or African ancestry [24, 25].

Our results indicated that SURF4 was significantly ($P < 0.001$) highly expressed in tumor. High expression of SURF4 was observed in T4, infiltrating ductal

A Univariate analysis of relapse free survival in breast cancer patients



B Multivariate analysis of relapse free survival in breast cancer patients

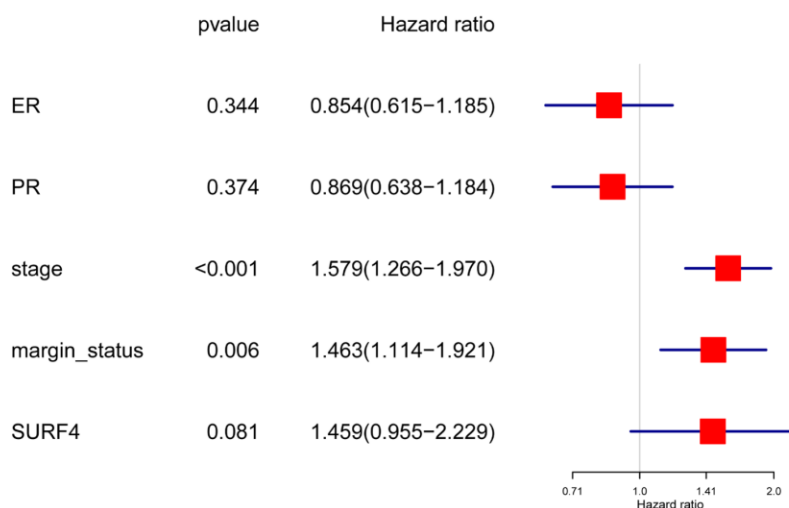


Figure 7. Forest plot of Cox regression analysis about SURF4 and RFS. (A) Univariate analysis of RFS in breast cancer patients. **(B)** Multivariate analysis of RFS in breast cancer patients.

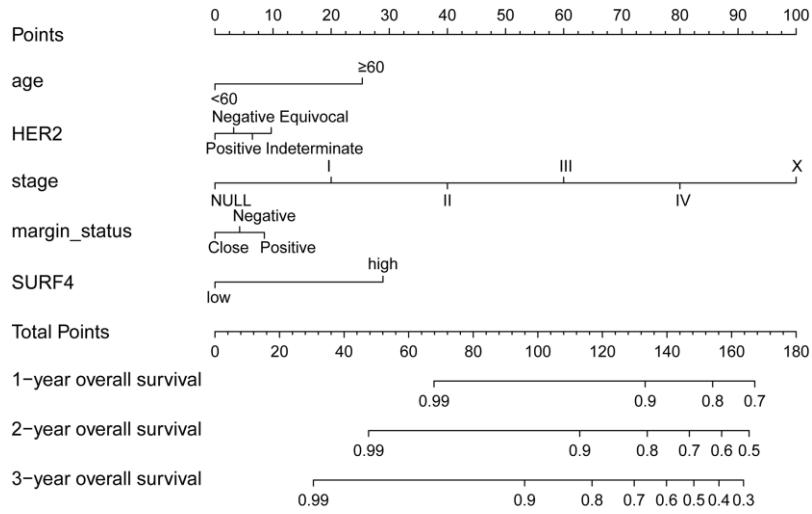


Figure 8. The nomogram about SURF4 and OS. 1-, 3- and 5-year related survival probabilities were obtained by drawing a line straight down to the risk axis.

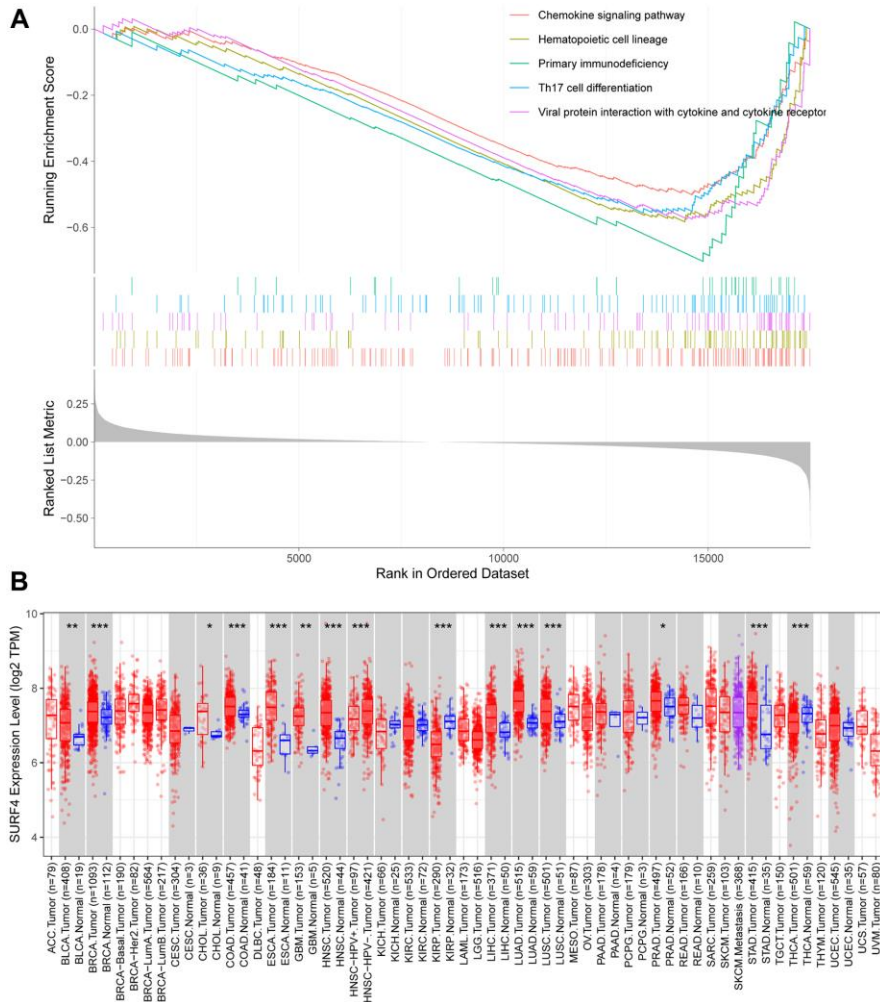


Figure 9. Expression of SURF4 by enrichment analysis and in TIMER database. (A) Top 5 enrichments with enriched high expression of SURF4. (B) Expression of SURF4 in diverse types of human cancers in the TIMER database.

carcinoma, ER negative, PR negative, HER2 positive, female, patients without lymph node metastasis, HER2 overexpression type, and deceased patients. The diagnostic ability increased with stage increased as the AUC for stage I-IV was 0.581, 0.609, 0.613, and 0.693, respectively. The age, stage, and SURF4 expression exhibited independent prognostic value for OS of breast cancer. Patients with high SURF4 expression, higher age, equivocal HER2, higher stages, or positive margin status had shorter OS. The stage and margin status exhibited independent prognostic value for RFS of breast cancer.

As is known, Ki-67 is a nuclear antigen associated with proliferating cells, and currently a widely used tumor cell proliferation activity marker [26]. Ki-67 expression level is significantly higher in malignant tumors [27]. Ki-67 has been used as one of the proliferation markers in early breast cancer [28]. Relevant studies have shown that changes in the expression level of Ki-67 can be used as a sensitivity indicator to predict the efficacy of neoadjuvant endocrine therapy [29]. Neoadjuvant endocrine therapy trials IMPACT and Z1031 showed a survival benefit after inhibition of Ki-67 expression [30, 31].

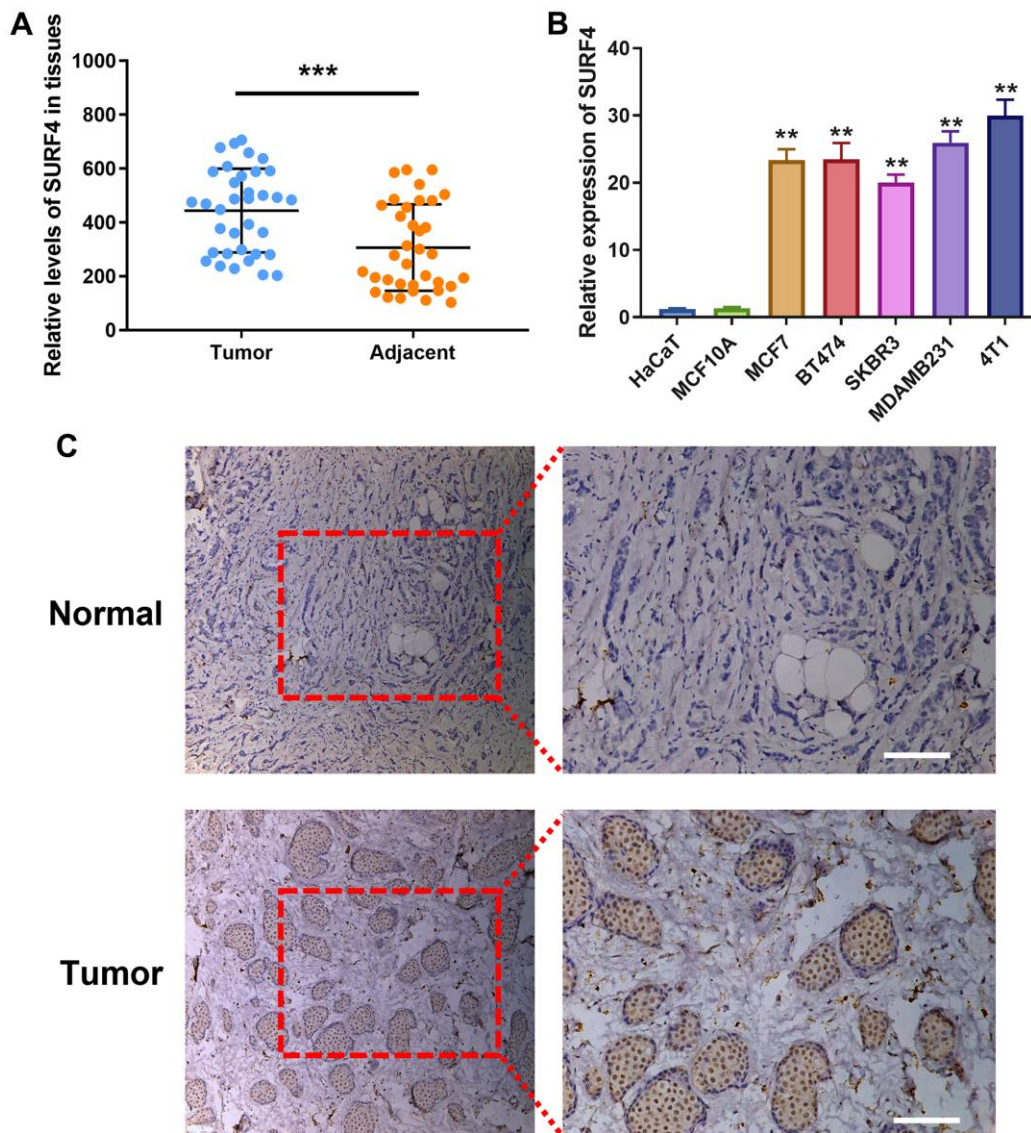


Figure 10. Expression of SURF4 in human breast tissues and cell lines. (A) SURF4 expression in tumor and adjacent normal tissue by qRT-PCR (N = 35). (B) SURF4 expression in HaCaT (Human immortalized epidermal cell), MCF10A (Human normal breast cell), MCF7 (Human luminal A type breast cancer cell), BT474 (Human luminal B type breast cancer cell), SKBR3 (Human HER2 overexpression type breast cancer cell), MDAMB231 and 4T1 (Human triple-negative breast cancer cell) by qRT-PCR in triplicate. (C) SURF4 expression in normal and tumor tissue by IHC. The tissues were from single patient. Scale bar = 100 μ m. **P < 0.01; ***P < 0.001.

Neoadjuvant endocrine therapy has been used as the primary endpoint, but its value in predicting the efficacy of neoadjuvant chemotherapy remains controversial [32].

Limited studies have showed the role of SURF4 in cancer. Yue et al. found SURF4 possessed the ability for maintaining stemness of ovarian cancer, and may serve as a potential target [33]. Kim et al. reported that SURF4 could induce cellular transformation and cell migration *in vitro* and has oncogenic transformation ability *in vivo* [13]. Our study reported a novel biomarker SURF4 in breast cancer. High SURF4 expression was confirmed in breast cancer tissue and cells. SURF4 expression was significantly higher in tumor. The high SURF4 expression in breast cancer tissue was further validated by IHC. Besides, significant high SURF4 expression ($P < 0.01$) was observed in MCF7 (Human luminal A type breast cancer cell), BT474 (Human luminal B type breast cancer cell), SKBR3 (Human HER2

overexpression type breast cancer cell), MDAMB231 and 4T1 (Human triple-negative breast cancer cell). The *in vitro* experiments suggested SURF4 promoted the proliferation and migration of 4T1 cells. Combing the results of GSEA analysis, SURF4 may influence the development of breast cancer by controlling chemokine signaling pathway, etc. However, the underlying mechanisms need to be studied in the future. Also, the *in vivo* experiments can improve the evidence power of the findings.

In conclusion, high expression of SURF4 was first found in breast cancer. SURF4 expression exhibited independent prognostic value for OS, and patients with high SURF4 expression had shorter OS. SURF4 promoted the proliferation and migration of 4T1 cells. SURF4 may be a biomarker to play a role in diagnosis and prognosis of breast cancer. Our findings may indicate SURF4 as a novel therapeutic target for treatment of breast cancer.

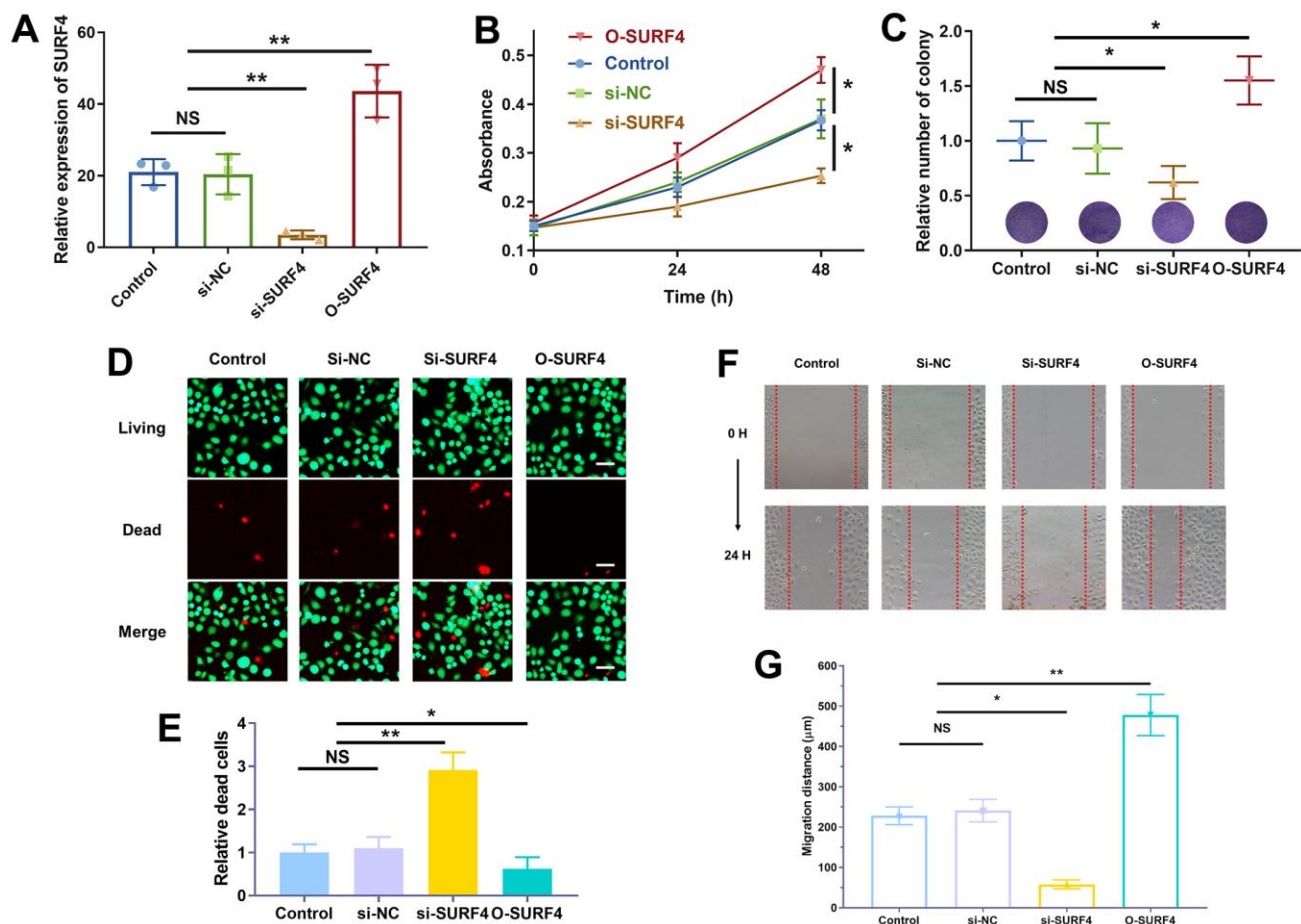


Figure 11. SURF4 promoted cell proliferation and migration of breast cancer cell 4T1. (A) Efficiency of plasmid transfection by qRT-PCR. (B) The cell viability by CCK8 assay. (C) Cell proliferation ability by colony formation assay. (D, E) Cell proliferation ability by living/dead staining assay. Scale bar = 50 μm. (F, G) Cell migration ability by wound healing assay. The experiments were repeated for 3 times. 40 x under light microscopy. NS, no significance; * $P < 0.05$; ** $P < 0.01$.

MATERIALS AND METHODS

Data processing and comparison

The files of mRNA expression and associated clinical data were acquired from the The Cancer Genome Atlas (TCGA) database [34]. Non-parametric rank sum tests were used to evaluate the SURF4 mRNA expression. Wilcoxon rank sum tests and Kruskal-Wallis tests were used to compare two and multiple subgroups, respectively. Chi-square tests along with Fisher's exact tests were used to evaluate the characteristics correlated with SURF4 expression.

Diagnostic value evaluation

The pROC program was used for visualization of ROC curves, which were plotted to evaluate the diagnostic value of SURF4 [35]. According to the cut-off value obtained from ROC curves between normal and tumor, the patients were further divided into the low and high SURF4 expression groups. The area under the ROC curves (AUC) was also calculated.

Survival evaluation and nomogram plotting

The Kaplan-Meier curves were used to analyze the OS and RFS by the R survival package [36]. To investigate the independent predictive ability of SURF4 in breast cancer, univariate and multivariate Cox analysis were carried out. According to SURF4 expression, the patients with breast cancer were divided into groups. Based on different SURF4 expression, the age, HER2, stage, margin status, and 1,3,5-year OS were compared.

GSEA analysis and TIMER database mining

To investigate the relation between SURF4 expression and enriched signaling pathways, the TCGA database was first searched, and then GSEA analysis was performed. The TIMER database was used for study of SURF4 expression in different cancers.

Sample collection

Breast cancer and adjacent tissues were obtained from 35 subjects, and placed in liquid nitrogen immediately after resection. The study was approved by the institutional ethical committee and conformed to the Declaration of Helsinki.

Cell culture and plasmid transfection

The cells were cultured in 1640 medium containing 10% fetal bovine serum, and transfected with the si-NC (negative control), si-SURF4 (small-interfering RNAs

against SURF4) and O-SURF4(overexpressed SURF4) plasmids. All the plasmids were purchased from Genepharma (Shanghai, China).

Real-time quantitative PCR

Total RNA was extracted using TRIzol (Invitrogen, USA). 1 mL of isolated RNA was used for the reverse transcription. The real-time quantitative PCR (RT-qPCR) procedure was then completed. The experiments were repeated for 3 times. The primers are as follows: SURF4 (Forward: 5'-CCTTTAAGGCTTGGCCTACG-3'; Reverse: 5'-GGGCCAGGTTCTCATCAAA-3'), and β -actin (Forward: 5'-GGAGCGAGATCCCTCCAA AAT-3'; Reverse: 5'-GGCTGTTGTCATACTTCTCAT GG-3').

Immunohistochemistry staining

The tissues were from single patient. The immunohistochemistry (IHC) staining was performed according to the manual instructions (#13079, Cell Signaling Technology, MA, USA). Anti-SURF4 primary antibody (ab133369, Abcam company, Shanghai, China) and corresponding secondary antibody (rabbit) were used. The fluorescence microscope was used for imaging.

Cell proliferation and migration assay

After a 24-hour culture period with different plasmids, the absorbance at 450 nm was measured after the addition of 10 μ L of CCK-8 reagent (CK04, Dojindo company, Beijing, China) [37]. After being scraped to create a 1-mm gap, the cells treated with different plasmids were grown for 48 hour, and images were obtained at 0 and 48 hours to record the migration distance [38]. The experiments were repeated for 3 times.

Statistical analysis

R3.5.1 was used to perform bioinformatics analysis [39]. The Kaplan-Meier curve was used to analyze the survival rate [40]. To investigate the independent predictive capability of SURF4 in breast cancer, univariate and multivariate Cox analyses were used. It was statistically significant at $P < 0.05$.

Abbreviations

ER: estrogen receptor; PR: progesterone receptor; HER2: human epidermal growth factor receptor 2; TCGA: The Cancer Genome Atlas; GSEA: Gene set enrichment analysis; OS: overall survival; RFS: relapse-free survival; HR: hazard ratio; CI: confidence interval; ROC: receiver operating characteristic; AUC: area under the ROC curves.

AUTHOR CONTRIBUTIONS

J.Z. and B.X. designed this study and had full access to all of data in the study; J.Z. and F.M. extracted the data; J.H., C.L., and F.G. analyzed and interpreted the data; J.Z. and J.H. performed the experiments; J.Z. and F.M. wrote the paper. All authors reviewed the manuscript.

CONFLICTS OF INTEREST

The authors have no conflicts of interest.

ETHICAL STATEMENT AND CONSENT

The study was approved by the Ethics Committee of Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College. The written consent of all subjects or their legal representatives has been obtained.

FUNDING

This study was funded by the National Key Research and Development Program of China (2021YFF1201300).

Editorial note

[&]This corresponding author has a verified history of publications using a personal email address for correspondence.

REFERENCES

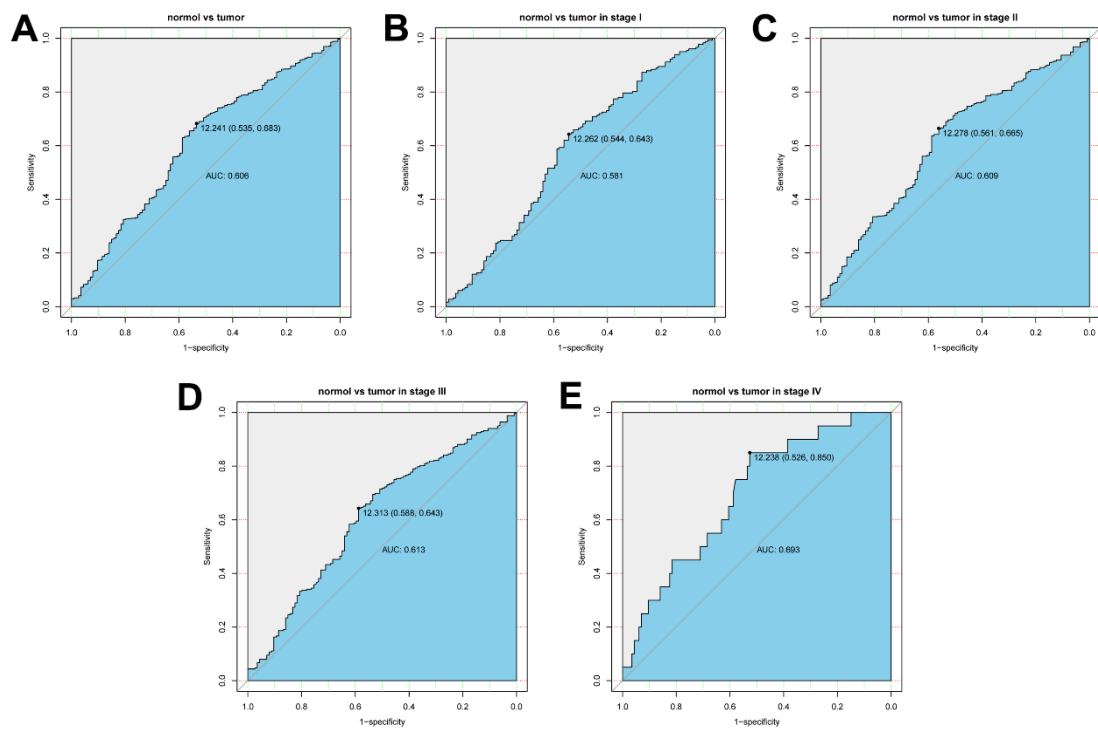
1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin.* 2021; 71:209–49. <https://doi.org/10.3322/caac.21660> PMID:33538338
2. Sørli T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matese JC, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci USA.* 2001; 98:10869–74. <https://doi.org/10.1073/pnas.191367098> PMID:11553815
3. Stephens PJ, Tarpey PS, Davies H, Van Loo P, Greenman C, Wedge DC, Nik-Zainal S, Martin S, Varela I, Bignell GR, Yates LR, Papaemmanuil E, Beare D, et al, and Oslo Breast Cancer Consortium (OSBREAC). The landscape of cancer genes and mutational processes in breast cancer. *Nature.* 2012; 486:400–4. <https://doi.org/10.1038/nature11017> PMID:22722201
4. Harbeck N, Gnant M. Breast cancer. *Lancet.* 2017; 389:1134–50. [https://doi.org/10.1016/S0140-6736\(16\)31891-8](https://doi.org/10.1016/S0140-6736(16)31891-8) PMID:27865536
5. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature.* 2012; 490:61–70. <https://doi.org/10.1038/nature11412> PMID:23000897
6. Allemani C, Matsuda T, Di Carlo V, Harewood R, Matz M, Nikšić M, Bonaventure A, Valkov M, Johnson CJ, Estève J, Ogunbiyi OJ, Azevedo E Silva G, Chen WQ, et al, and CONCORD Working Group. Global surveillance of trends in cancer survival 2000-14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. *Lancet.* 2018; 391:1023–75. [https://doi.org/10.1016/S0140-6736\(17\)33326-3](https://doi.org/10.1016/S0140-6736(17)33326-3) PMID:29395269
7. Cardoso F, Spence D, Mertz S, Corneliusen-James D, Sabelko K, Gralow J, Cardoso MJ, Peccatori F, Paonessa D, Benares A, Sakurai N, Beishon M, Barker SJ, Mayer M. Global analysis of advanced/metastatic breast cancer: Decade report (2005-2015). *Breast.* 2018; 39:131–8. <https://doi.org/10.1016/j.breast.2018.03.002> PMID:29679849
8. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. 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. <https://doi.org/10.3322/caac.21492> PMID:30207593
9. Kwong GA, Ghosh S, Gamboa L, Patriotis C, Srivastava S, Bhatia SN. Synthetic biomarkers: a twenty-first century path to early cancer detection. *Nat Rev Cancer.* 2021; 21:655–68. <https://doi.org/10.1038/s41568-021-00389-3> PMID:34489588
10. Loibl S, Poortmans P, Morrow M, Denkert C, Curigliano G. Breast cancer. *Lancet.* 2021; 397:1750–69. [https://doi.org/10.1016/S0140-6736\(20\)32381-3](https://doi.org/10.1016/S0140-6736(20)32381-3) PMID:33812473
11. Loke SY, Lee ASG. The future of blood-based biomarkers for the early detection of breast cancer. *Eur J Cancer.* 2018; 92:54–68. <https://doi.org/10.1016/j.ejca.2017.12.025> PMID:29413690

12. Zou J, Wang E. Cancer Biomarker Discovery for Precision Medicine: New Progress. *Curr Med Chem*. 2019; 26:7655–71.
<https://doi.org/10.2174/0929867325666180718164712> PMID:30027846
13. Kim J, Hong CM, Park SM, Shin DH, Kim JY, Kwon SM, Kim JH, Kim CD, Lim DS, Lee D. SURF4 has oncogenic potential in NIH3T3 cells. *Biochem Biophys Res Commun*. 2018; 502:43–7.
<https://doi.org/10.1016/j.bbrc.2018.05.116> PMID:29777698
14. Emmer BT, Lascuna PJ, Tang VT, Kotnik EN, Saunders TL, Khoriaty R, Ginsburg D. Murine Surf4 is essential for early embryonic development. *PLoS One*. 2020; 15:e0227450.
<https://doi.org/10.1371/journal.pone.0227450> PMID:31978056
15. Reeves JE, Fried M. The surf-4 gene encodes a novel 30 kDa integral membrane protein. *Mol Membr Biol*. 1995; 12:201–8.
<https://doi.org/10.3109/09687689509027508> PMID:7540914
16. Mitrovic S, Ben-Tekaya H, Koegler E, Gruenberg J, Hauri HP. The cargo receptors Surf4, endoplasmic reticulum-Golgi intermediate compartment (ERGIC)-53, and p25 are required to maintain the architecture of ERGIC and Golgi. *Mol Biol Cell*. 2008; 19:1976–90.
<https://doi.org/10.1091/mbc.e07-10-0989> PMID:18287528
17. Cardoso F, Kyriakides S, Ohno S, Penault-Llorca F, Poortmans P, Rubio IT, Zackrisson S, Senkus E, and ESMO Guidelines Committee. Electronic address: clinicalguidelines@esmo.org. Early breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up†. *Ann Oncol*. 2019; 30:1194–220.
<https://doi.org/10.1093/annonc/mdz173> PMID:31161190
18. Kravchenko J, Akushevich I, Seewaldt VL, Abernethy AP, Lyerly HK. Breast cancer as heterogeneous disease: contributing factors and carcinogenesis mechanisms. *Breast Cancer Res Treat*. 2011; 128:483–93.
<https://doi.org/10.1007/s10549-011-1347-z> PMID:21225455
19. Perou CM, Sørlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge O, Pergamenschikov A, Williams C, et al. Molecular portraits of human breast tumours. *Nature*. 2000; 406:747–52.
<https://doi.org/10.1038/35021093> PMID:10963602
20. Ignatiadis M, Sotiriou C. Luminal breast cancer: from biology to treatment. *Nat Rev Clin Oncol*. 2013; 10:494–506.
<https://doi.org/10.1038/nrclinonc.2013.124> PMID:23881035
21. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Aromatase inhibitors versus tamoxifen in early breast cancer: patient-level meta-analysis of the randomised trials. *Lancet*. 2015; 386:1341–52.
[https://doi.org/10.1016/S0140-6736\(15\)61074-1](https://doi.org/10.1016/S0140-6736(15)61074-1) PMID:26211827
22. Creighton CJ. The molecular profile of luminal B breast cancer. *Biologics*. 2012; 6:289–97.
<https://doi.org/10.2147/BTT.S29923> PMID:22956860
23. LaBoy C, Siziopikou KP, Rosen L, Blanco LZ Jr, Pincus JL. Clinicopathologic features of unexpectedly HER2 positive breast carcinomas: An institutional experience. *Pathol Res Pract*. 2021; 222:153441.
<https://doi.org/10.1016/j.prp.2021.153441> PMID:33857853
24. Foulkes WD, Stefansson IM, Chappuis PO, Bégin LR, Goffin JR, Wong N, Trudel M, Akslen LA. Germline BRCA1 mutations and a basal epithelial phenotype in breast cancer. *J Natl Cancer Inst*. 2003; 95:1482–5.
<https://doi.org/10.1093/jnci/djg050> PMID:14519755
25. Carey LA, Perou CM, Livasy CA, Dressler LG, Cowan D, Conway K, Karaca G, Troester MA, Tse CK, Edmiston S, Deming SL, Geradts J, Cheang MCU, et al. Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. *JAMA*. 2006; 295:2492–502.
<https://doi.org/10.1001/jama.295.21.2492> PMID:16757721
26. Sun X, Kaufman PD. Ki-67: more than a proliferation marker. *Chromosoma*. 2018; 127:175–86.
<https://doi.org/10.1007/s00412-018-0659-8> PMID:29322240
27. Menon SS, Guruvayoorappan C, Sakthivel KM, Rasmi RR. Ki-67 protein as a tumour proliferation marker. *Clin Chim Acta*. 2019; 491:39–45.
<https://doi.org/10.1016/j.cca.2019.01.011> PMID:30653951
28. Urruticoechea A, Smith IE, Dowsett M. Proliferation marker Ki-67 in early breast cancer. *J Clin Oncol*. 2005; 23:7212–20.
<https://doi.org/10.1200/JCO.2005.07.501> PMID:16192605
29. Goncalves R, DeSchryver K, Ma C, Tao Y, Hoog J, Cheang M, Crouch E, Dahiya N, Sanati S, Barnes M, Sarian LOZ, Olson J, Allred DC, Ellis MJ. Development of a Ki-67-based clinical trial assay for neoadjuvant endocrine therapy response monitoring in breast cancer. *Breast Cancer Res Treat*. 2017; 165:355–64.
<https://doi.org/10.1007/s10549-017-4329-y>

- PMID:[28612227](#)
30. Dowsett M, Smith IE, Ebbs SR, Dixon JM, Skene A, Griffith C, Boeddinghaus I, Salter J, Detre S, Hills M, Ashley S, Francis S, Walsh G, and IMPACT Trialists. Short-term changes in Ki-67 during neoadjuvant treatment of primary breast cancer with anastrozole or tamoxifen alone or combined correlate with recurrence-free survival. *Clin Cancer Res*. 2005; 11:951s–8s. PMID:[15701892](#)
 31. Ellis MJ, Suman VJ, Hoog J, Lin L, Snider J, Prat A, Parker JS, Luo J, DeSchryver K, Allred DC, Esserman LJ, Unzeitig GW, Margenthaler J, et al. Randomized phase II neoadjuvant comparison between letrozole, anastrozole, and exemestane for postmenopausal women with estrogen receptor-rich stage 2 to 3 breast cancer: clinical and biomarker outcomes and predictive value of the baseline PAM50-based intrinsic subtype--ACOSOG Z1031. *J Clin Oncol*. 2011; 29:2342–9. <https://doi.org/10.1200/JCO.2010.31.6950> PMID:[21555689](#)
 32. Dowsett M, Nielsen TO, A'Hern R, Bartlett J, Coombes RC, Cuzick J, Ellis M, Henry NL, Hugh JC, Lively T, McShane L, Paik S, Penault-Llorca F, et al, and International Ki-67 in Breast Cancer Working Group. Assessment of Ki67 in breast cancer: recommendations from the International Ki67 in Breast Cancer working group. *J Natl Cancer Inst*. 2011; 103:1656–64. <https://doi.org/10.1093/jnci/djr393> PMID:[21960707](#)
 33. Yue Y, Xia L, Xu S, Wang C, Wang X, Lu W, Xie X. SURF4 maintains stem-like properties via BIRC3 in ovarian cancer cells. *J Gynecol Oncol*. 2020; 31:e46. <https://doi.org/10.3802/jgo.2020.31.e46> PMID:[32026660](#)
 34. Cai H, Jiao Y, Li Y, Yang Z, He M, Liu Y. Low CYP24A1 mRNA expression and its role in prognosis of breast cancer. *Sci Rep*. 2019; 9:13714. <https://doi.org/10.1038/s41598-019-50214-z> PMID:[31548577](#)
 35. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Müller M. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. 2011; 12:77. <https://doi.org/10.1186/1471-2105-12-77> PMID:[21414208](#)
 36. Liu N, Zhou Y, Lee JJ. IPDfromKM: reconstruct individual patient data from published Kaplan-Meier survival curves. *BMC Med Res Methodol*. 2021; 21:111. <https://doi.org/10.1186/s12874-021-01308-8> PMID:[34074267](#)
 37. Cai H, Dai X, Guo X, Zhang L, Cao K, Yan F, Ji B, Liu Y. Ataxia telangiectasia mutated inhibitor-loaded copper sulfide nanoparticles for low-temperature photothermal therapy of hepatocellular carcinoma. *Acta Biomater*. 2021; 127:276–86. <https://doi.org/10.1016/j.actbio.2021.03.051> PMID:[33812073](#)
 38. Cai H, Wang R, Guo X, Song M, Yan F, Ji B, Liu Y. Combining Gemcitabine-Loaded Macrophage-like Nanoparticles and Erlotinib for Pancreatic Cancer Therapy. *Mol Pharm*. 2021; 18:2495–506. <https://doi.org/10.1021/acs.molpharmaceut.0c01225> PMID:[34078087](#)
 39. Heinemann J. Cluster Analysis of Untargeted Metabolomic Experiments. *Methods Mol Biol*. 2019; 1859:275–85. https://doi.org/10.1007/978-1-4939-8757-3_16 PMID:[30421236](#)
 40. Ranstam J, Cook JA. Kaplan-Meier curve. *Br J Surg*. 2017; 104:442. <https://doi.org/10.1002/bjs.10238> PMID:[28199017](#)

SUPPLEMENTARY MATERIALS

Supplementary Figure



Supplementary Figure 1. ROC curve to assess the diagnostic value of SURF4. (A) ROC curve of SURF4 in normal vs. tumor. **(B–E)** ROC curve of SURF4 in stage I, stage II, stage III, and stage IV.

Supplementary Tables

Supplementary Table 1. Patient characteristics.

Characteristics	Numbers of cases
Age	
<60	589(53.45)
>=60	513(46.55)
Gender	
NA	2(0.18)
Female	1090(98.73)
Male	12(1.09)
Histological type	
NA	3(0.27)
Infiltrating Ductal Carcinoma	790(71.56)
Infiltrating Lobular Carcinoma	204(18.48)
Other	107(9.69)
Molecular subtype	
NA	255(23.10)
Basal	142(12.86)
HER2	67(6.07)
LuminalA	422(38.22)
LuminalB	194(17.57)
Normal	24(2.17)
ER	
NA	50(4.53)
Indeterminate	2(0.18)
Negative	239(21.65)
Positive	813(73.64)
PR	
NA	51(4.62)
Indeterminate	4(0.36)
Negative	345(31.25)
Positive	704(63.77)
HER2	
NA	183(16.58)
Equivocal	180(16.30)
Indeterminate	12(1.09)
Negative	565(51.18)
Positive	164(14.86)
Menopause status	
NA	93(8.42)
Inde	34(3.08)
Peri	40(3.62)
Post	706(63.95)
Pre	231(20.92)
T classification	
NA	2(0.18)
T1	281(25.45)
T2	640(57.97)
T3	138(12.5)
T4	40(3.62)
TX	3(0.27)
N classification	
NA	2(0.18)

N0	516(46.74)
N1	367(33.24)
N2	120(10.87)
N3	79(7.16)
NX	20(1.81)
M classification	
NA	2(0.18)
M0	917(83.06)
M1	22(1.99)
MX	163(14.76)
Stage	
NA	10(0.91)
I	182(16.49)
II	626(56.70)
III	252(22.83)
IV	20(1.81)
X	14(1.27)
Lymph node status	
NA	379(34.33)
No	28(2.54)
Yes	697(63.13)
Margin status	
NA	72(6.52)
Close	31(2.81)
Negative	922(83.51)
Positive	79(7.16)
Vital status	
NA	2(0.18)
Deceased	155(14.04)
Living	947(85.78)
Radiation therapy	
NA	102(9.24)
No	445(40.31)
Yes	557(50.45)
Neoadjuvant treatment	
NA	3(0.27)
No	1088(98.55)
Yes	13(1.18)
Targeted molecular therapy	
NA	525(47.55)
No	46(4.17)
Yes	533(48.28)
Sample type	
Metastatic	7(0.63)
Primary Tumor	1097(99.37)
SURF4	
high	445(40.31)
low	659(59.69)

Note: ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2; NA, not available; X represents uncertain.

Supplementary Table 2. Characteristics correlated with SURF4 expression.

Parameter	Variable	N	SURF4 mRNA expression				χ^2	P value
			high	%	low	%		
Age	<60	589	247	(55.76)	342	(51.90)	1.586	0.210
	>=60	513	196	(44.24)	317	(48.10)		
Gender	Female	1090	442	(99.77)	648	(98.33)	5.125	0.032
	Male	12	1	(0.23)	11	(1.67)		
Histological type	Infiltrating Ductal Carcinoma	790	355	(80.32)	435	(66.01)	28.903	<0.001
	Infiltrating Lobular Carcinoma	204	51	(11.54)	153	(23.22)		
	Other	107	36	(8.14)	71	(10.77)		
Molecular subtype	Basal	142	76	(21.23)	66	(13.44)	65.143	<0.001
	HER22	67	54	(15.08)	13	(2.65)		
	LuminalA	422	137	(38.27)	285	(58.04)		
	LumialB	194	79	(22.07)	115	(23.42)		
ER	Normal	24	12	(3.35)	12	(2.44)	41.049	<0.001
	Indeterminate	2	1	(0.24)	1	(0.16)		
	Negative	239	139	(32.71)	100	(15.90)		
	Positive	813	285	(67.06)	528	(83.94)		
PR	Indeterminate	4	2	(0.47)	2	(0.32)	38.503	<0.001
	Negative	345	185	(43.63)	160	(25.44)		
	Positive	704	237	(55.90)	467	(74.24)		
HER2	Equivocal	180	63	(16.94)	117	(21.31)	32.137	<0.001
	Indeterminate	12	6	(1.61)	6	(1.09)		
	Negative	565	205	(55.11)	360	(65.57)		
Menopause status	Positive	164	98	(26.34)	66	(12.02)	3.592	0.311
	Inde	34	19	(4.65)	15	(2.49)		
	Peri	40	17	(4.16)	23	(3.82)		
	Post	706	281	(68.70)	425	(70.60)		
T classification	Pre	231	92	(22.49)	139	(23.09)	7.354	0.133
	T1	281	103	(23.25)	178	(27.01)		
	T2	640	268	(60.50)	372	(56.45)		
	T3	138	51	(11.51)	87	(13.20)		
	T4	40	21	(4.74)	19	(2.88)		
	TX	3	0	(0.00)	3	(0.46)		
	N0	516	199	(44.92)	317	(48.10)		
N1	367	145	(32.73)	222	(33.69)			
N2	120	55	(12.42)	65	(9.86)			
N3	79	34	(7.67)	45	(6.83)			
M classification	NX	20	10	(2.26)	10	(1.52)	1.710	0.414
	M0	917	372	(83.97)	545	(82.70)		
	M1	22	11	(2.48)	11	(1.67)		
Stage	MX	163	60	(13.54)	103	(15.63)	4.441	0.335
	I	182	62	(14.12)	120	(18.32)		
	II	626	253	(57.63)	373	(56.95)		
	III	252	108	(24.60)	144	(21.98)		
	IV	20	10	(2.28)	10	(1.53)		
Lymph node status	X	14	6	(1.37)	8	(1.22)	3.084	0.106
	No	28	15	(5.47)	13	(2.88)		
	Yes	697	259	(94.53)	438	(97.12)		
Margin status	Close	31	14	(3.42)	17	(2.73)	0.680	0.692
	Negative	922	366	(89.49)	556	(89.25)		
	Positive	79	29	(7.09)	50	(8.03)		
Vital status	Deceased	155	81	(18.28)	74	(11.23)	10.910	0.002
	Living	947	362	(81.72)	585	(88.77)		

Radiation therapy	No	445	190	(47.03)	255	(42.64)	1.880	0.186
	Yes	557	214	(52.97)	343	(57.36)		
Neoadjuvant treatment	No	1088	437	(98.65)	651	(98.94)	0.192	0.785
	Yes	13	6	(1.35)	7	(1.06)		
Targeted molecular therapy	No	46	22	(9.09)	24	(7.12)	0.747	0.454
	Yes	533	220	(90.91)	313	(92.88)		
Sample type	Metastatic	7	2	(0.45)	5	(0.76)	0.403	0.707
	Primary Tumor	1097	443	(99.55)	654	(99.24)		

Note: ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2; NA, not available; X represents uncertain.

Supplementary Table 3. Univariate and multivariate analysis for overall survival.

Parameters	Univariate analysis			Multivariate analysis		
	HR	95% CI	P value	HR	95% CI	P value
Age	1.91	1.39-2.63	<0.001	2.32	1.45-3.70	<0.001
ER	0.85	0.71-1.02	0.074			
HER2	1.29	1.05-1.57	0.013	1.12	0.90-1.39	0.297
Histological type	0.93	0.74-1.17	0.543			
Lymph node status	1.10	0.93-1.30	0.274			
Margin status	1.42	1.11-1.81	0.005	1.01	0.72-1.41	0.942
Menopause status	1.16	0.94-1.43	0.165			
Molecular subtype	1.01	0.88-1.16	0.901			
PR	0.87	0.73-1.03	0.096			
Stage	1.64	1.40-1.91	<0.001	2.09	1.59-2.75	<0.001
SURF4	1.90	1.38-2.61	<0.001	1.96	1.23-3.11	0.005

Note: ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2; HR, hazard ratio; CI, confidence interval.

Supplementary Table 4. Univariate and multivariate analysis for relapse-free survival.

Parameters	Univariate analysis			Multivariate analysis		
	HR	95% CI	P value	HR	95% CI	P value
Age	1.45	0.97-2.16	0.072			
ER	0.78	0.63-0.97	0.026	0.85	0.61-1.18	0.344
HER2	0.93	0.70-1.22	0.596			
Histological type	0.86	0.65-1.14	0.290			
Lymph node status	0.86	0.70-1.06	0.159			
Margin status	1.59	1.23-2.06	<0.001	1.46	1.11-1.92	0.006
Menopause status	0.95	0.74-1.22	0.713			
Molecular subtype	0.99	0.82-1.20	0.945			
PR	0.78	0.64-0.96	0.019	0.87	0.64-1.18	0.374
Stage	1.71	1.40-2.08	<0.001	1.58	1.27-1.97	<0.001
SURF4	1.76	1.18-2.63	0.005	1.46	0.95-2.23	0.081

Note: ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2; HR, hazard ratio; CI, confidence interval.

Supplementary Table 5. Enriched signaling pathways by GSEA analysis.

Description	Set size	Enrichment score	NES	P value	P adjusted value	Q value
Chemokine signaling pathway	188	-0.5014	-1.9249	1.9E-07	6.4E-05	5.3E-05
Hematopoietic cell lineage	94	-0.5834	-2.0432	5.9E-07	9.9E-05	8.1E-05
Viral protein interaction with cytokine and cytokine receptor	95	-0.5749	-2.0224	1.4E-06	0.00016	0.00013
Th17 cell differentiation	106	-0.5569	-1.9797	2.8E-06	0.00024	0.00019
Primary immunodeficiency	36	-0.703	-2.0913	3.9E-06	0.00026	0.00021
Cytokine-cytokine receptor interaction	282	-0.4153	-1.6693	1.1E-05	0.00056	0.00046
Cell cycle	119	0.47578	1.88178	1.2E-05	0.00056	0.00046
Intestinal immune network for IgA production	45	-0.647	-2.003	1.6E-05	0.00058	0.00048
B cell receptor signaling pathway	80	-0.5554	-1.9029	1.6E-05	0.00058	0.00048
Osteoclast differentiation	127	-0.503	-1.8477	1.7E-05	0.00058	0.00048
Systemic lupus erythematosus	50	-0.6302	-1.9953	2.4E-05	0.00072	0.00059
Inflammatory bowel disease	63	-0.5732	-1.8773	4.3E-05	0.00111	0.00092
Glutamatergic synapse	111	-0.5066	-1.8239	4.4E-05	0.00111	0.00092
Allograft rejection	35	-0.6609	-1.9448	0.00011	0.00254	0.0021
Taurine and hypotaurine metabolism	16	-0.7963	-1.9652	0.00011	0.00254	0.0021
Amphetamine addiction	69	-0.5521	-1.8401	0.00013	0.00279	0.0023
MicroRNAs in cancer	160	0.40191	1.6424	0.0002	0.004	0.0033
Autoimmune thyroid disease	50	-0.5818	-1.8419	0.00025	0.00461	0.0038
Staphylococcus aureus infection	87	-0.5019	-1.743	0.0003	0.00535	0.00441
Cocaine addiction	49	-0.5858	-1.8508	0.00034	0.00564	0.00465
Asthma	28	-0.677	-1.8853	0.00041	0.00644	0.00532
Signaling pathways regulating pluripotency of stem cells	140	0.39619	1.60953	0.00047	0.00706	0.00582
Cell adhesion molecules	147	-0.4349	-1.6294	0.0006	0.00868	0.00716
Malaria	47	-0.5783	-1.8155	0.00066	0.00895	0.00739
Nucleocytoplasmic transport	101	0.42755	1.66382	0.00067	0.00895	0.00739
Type I diabetes mellitus	41	-0.584	-1.7902	0.00073	0.00939	0.00775
Tuberculosis	177	-0.4053	-1.5455	0.00083	0.01024	0.00845
Graft-versus-host disease	38	-0.6079	-1.8309	0.00113	0.01345	0.0111
Th1 and Th2 cell differentiation	90	-0.4707	-1.6373	0.00121	0.01385	0.01143
Basal cell carcinoma	63	0.4654	1.64309	0.00143	0.01588	0.0131
T cell receptor signaling pathway	101	-0.4443	-1.5744	0.00239	0.02571	0.02121
Insulin resistance	105	-0.4493	-1.5999	0.00308	0.03207	0.02646
Viral myocarditis	57	-0.5092	-1.6451	0.00345	0.03485	0.02876
Protein digestion and absorption	100	0.39887	1.54667	0.00382	0.03706	0.03058
Tight junction	163	-0.3965	-1.5008	0.00398	0.03706	0.03058
Small cell lung cancer	92	0.40964	1.56086	0.00401	0.03706	0.03058
Rheumatoid arthritis	88	-0.4545	-1.5792	0.00414	0.03723	0.03072
Lipid and atherosclerosis	213	-0.37	-1.4395	0.005	0.04378	0.03612
NF-kappa B signaling pathway	99	-0.4281	-1.5101	0.00538	0.04497	0.0371
Long-term potentiation	67	-0.4755	-1.5762	0.0054	0.04497	0.0371
Bladder cancer	40	0.51751	1.68382	0.00555	0.04504	0.03716
NOD-like receptor signaling pathway	175	-0.3746	-1.431	0.00592	0.04683	0.03864
DNA replication	36	0.53172	1.68595	0.00605	0.04683	0.03864