Am J Cancer Res 2022;12(2):469-489
www.ajcr.us /ISSN:2156-6976/ajcr0140575

Original Article

Upregulation of Cyclin B2 (*CCNB2*) in breast cancer contributes to the development of lymphovascular invasion

Abrar I Aljohani[1,2], Michael S Toss[1], Khloud A El-Sharawy[3], Sameer Mirza[4], Graham R Ball[5], Andrew R Green[1], Emad A Rakha[1,6,7]

[1]Academic Unit for Translational Medical Sciences, School of Medicine, Nottingham Breast Cancer Research Centre, University of Nottingham Biodiscovery Institute, University Park, Nottingham, UK; [2]Department of Clinical Laboratory Sciences, Faculty of Applied Medical Sciences, Taif University, Taif, Saudi Arabia; [3]Faculty of Science, Damietta University, Damietta, Egypt; [4]Department of Genetics, Cell Biology and Anatomy, University of Nebraska Medical Center, Omaha, USA; [5]The John Van Geest Cancer Research Centre, School of Science and Technology, Nottingham Trent University, Clifton Lane, Nottingham, UK; [6]Histopathology Department, Faculty of Medicine, Menoufia University, Shibïn al-Kawm, Egypt; [7]Department of Histopathology, Nottingham University Hospital NHS Trust, City Hospital Campus, Hucknall Road, Nottingham, UK

Received November 15, 2021; Accepted January 7, 2022; Epub February 15, 2022; Published February 28, 2022

Abstract: Lymphovascular invasion (LVI) is a key step in breast cancer (BC) metastasis. Targeting the molecular drivers of LVI can improve BC patients' management. However, the underlying molecular mechanisms of LVI are complex and interconnected with various carcinogenesis pathways. This study aimed to identify the key regulatory gene associated with LVI and to investigate its mechanisms of action and prognostic significance. Artificial neural network (ANN) was applied to two large transcriptomic datasets of BC with well-characterised LVI status. Cyclin B2 *(CCNB2)* was identified in the top genes associated with LVI positivity. *In vitro* functional assays were carried out to assess the role of *CCNB2* in tumour cell behaviour and their interactions with endothelial cells using a panel of BC cell lines. Large annotated BC cohorts were used to assess the clinical and prognostic role of *CCNB2* at the transcriptomic and protein levels. Knockdown (KD) of *CCNB2* mRNA reduced BC cell migration, inhibited proliferation, blocked the G2/M transition during the cell cycle and increased the number of apoptotic cells. Importantly, KD of *CCNB2* reduced BC cell lines adherence and transmigration across endothelial cell lines. High CCNB2 protein expression was independently associated with LVI positivity in addition to other features of aggressive behaviour, including larger tumour size, higher histological grade, hormonal receptor-negativity, and HER2-positivity, and with shorter survival. We conclude that CCNB2 plays a crucial role in LVI development in BC, implying that CCNB2 could confer a promising therapeutic target to inhibit LVI and reduce metastatic events.

Keywords: Breast cancer, *CCNB2*, prognosis, progression, LVI

Introduction

Breast cancer (BC) is the second most common cause of cancer-related mortality in women. Despite the advances in BC treatment, a considerable proportion of patients develop disease metastasis even within early-stage BC [1]. Lymphovascular invasion (LVI) is a key initial event of the metastatic cascade and serves as a prognostic factor in BC [2]. The lack of targeted therapy to inhibit LVI and prevent metastasis typically stems from the complex underlying molecular mechanisms of LVI, which interact with other carcinogenesis pathways and the subjectivity of assessing LVI in clinical BC cohorts. Thus, deciphering the underlying molecular mechanisms of LVI and the identification of the key genes controlling LVI in BC are warranted.

Machine learning (ML) is an innovative tool that has various applications, including running bioinformatical analysis of large datasets such as proteomic and transcriptomic data. ML depends

on using sets of algorithms to analyse such data to obtain specific results that solve a given complex question [3, 4]. Artificial neural networks (ANNs) are the framework of various ML algorithms to process complex datasets [3]. Previous studies have successfully used ML in discovering novel biomarkers associated with specific clinical outcomes in various cancers [5, 6], including BC [7].

In this study, we hypothesised that the power of ANN in large annotated BC cohorts with well-characterised LVI status can identify genes strongly associated with LVI in BC. As association with LVI can be a secondary event, the direct role of the selected gene should be confirmed both *ex vivo* and *in vitro*. Mechanistic studies were carried out using a panel of cell lines to decipher its role in the tumour cell carcinogenesis processes and the interaction of BC cells with endothelial cells. Clinical samples were used to evaluate its association with the clinicopathological features and outcome.

Materials and methods

ML for the biomarker set enrichment

ANN was used to identify the differentially expressed gene(s) between LVI positivity and negativity in two large BC datasets; the Molecular Taxonomy of Breast Cancer International Consortium (METABRIC; n=1,980) and The Cancer Genome Atlas (TCGA; n=854) [8, 9]. LVI status was previously assessed for 1565 patients, including the Nottingham subset (n=285) from the METABRIC cohort, through histological analysis of H&E-stained paraffin embedded tissues. The LVI status was also defined in the Nottingham subset through the use of immunohistochemistry (IHC) staining for CD31, CD34, and D2-40 in addition to the H&E [10]. The LVI was assessed in the TCGA by the evaluation of histological slides stained with H&E as no IHC marker status was available for these cases. The definition of LVI positive cases was based on the presence of LVI in H&E-stained sections in both cohorts. Positive staining in IHC markers was also used to identify LVI positive cases in the Nottingham subgroup of the METABRIC cohort. The positive LVI in that cohort was based on both H&E and the stain of IHC markers. The ANN-based neural data mining was performed on the genomic expression data obtained from both datasets to locate an enriched set of concordant biomarkers linked to LVI. In this case, we executed a ML strategy grounded in an ANN and integrated with concordance analysis performed across the multiple splits of Monte Carlo data based on the methodology indicated previously [7], which has initially proven to be efficient in eradicating over-fitting and false discovery while enhancing the generalisation of the identified biomarkers. The data were classified into five random groups based on the LVI status, and the ANN algorithm was run separately for each group, with each run consisting of 20 loops. The results were filtered to identify the concordant transcripts with the lowest test errors that were present in multiple loops for each group, and the results for all groups were compared to identify similar transcripts that were present in the various groups. Cyclin B2 (*CCNB2)* was detected in the top ranked genes associated with LVI positivity on both METABRIC and TCGA cohorts, thus we decided to decipher its role in LVI development and BC outcome in both pre-clinical and clinical settings.

In vitro assessment of CCNB2

Pre-clinical studies were carried out to investigate the potential role of *CCNB2* in LVI and other biological functions using the following cells and assays:

BC tumour cells: According to the protein expression levels, *CCNB2* is expressed in the luminal and HER2 enriched cell lines. MCF-7, ZR-75-1 and SK-BR-3 BC cell lines were purchased from the American Type Culture Collection (ATCC, Manassas, VA, USA). MCF-7 and ZR-75-1 were cultured in RPMI 1640 medium with L-glutamine (SH30027.FS; Cytiva, UK) supplemented with 10% Foetal Bovine Serum (FBS). McCoy's 5A medium modified with L-glutamine and sodium bicarbonate liquid (M9309; Sigma, UK) supplemented with 10% FBS was used to culture SK-BR-3.

Endothelial cells: Primary human umbilical vein endothelial cells (HUVECs) and dermal lymphatic endothelial cells (DLECs) were purchased from Promocell (Heidelberg, Germany). HUVECs and DLECs were cultured in endothelial cell growth medium MV2 (C-22022, Promocell, Germany). All cell cultures were performed under a sterile environment in a class II cabi-

net, and cells were incubated with 5% CO_2 at 37°C.

Transfection of CCNB2 using siRNA: To investigate possible functional consequences of *CCNB2* depletion and study its role in BC progression, with emphasising its role in LVI, we used a siRNA-based approach using BC cell lines. Differential expression of *CCNB2* was carried out using western blotting (WB) and cells with high expression of *CCNB2* were chosen.

Using pre-validated silencer select siRNA constructs mainly for designed *CCNB2* or scrambled negative control siRNA (Silencer® Select siRNA, ThermoFisher Scientific), MCF-7, ZR-75-1 and SK-BR-3 were transfected with a mixture of Opti-MEM medium, 25 pmol siRNA and Lipofectamine™ RNAiMAX (13778150; ThermoFisher Scientific, Loughborough, UK) using the forward transfection method according to the manufacturer's protocol. The sequence of *CCNB2*-siRNA is as follows: (5'->3') GGCGAACUGUUUUAGAAGAtt.

Wound healing assay: In the scratching assay to study the effect of *CCNB2* knockdown (KD) on cell migration, the wound repair rate of *CCNB2* KD and control cells was observed by measuring the width of the gap left unhealed. A Culture-Inserts 2 well (Thistle Scientific Ltd, IB-81176), which has a built-in gap, was used to assess the migration ability of the cells according to the manufacture protocol. The wounds were observed by taking images several times via light microscopy (Lecia DMI 3000B, Leica microsystems, Germany) at the following time points: T0, T24, T48 and T72 h. Image J software (1.52 version) was used to measure the wound area and calculate the percentage of wound closure.

Proliferation assay: The effect of *CCNB2* KD on the proliferation of tumour cells was assessed via 3-(4,5-dimethylthiazol-2-yl)-2, 5-diphenyltetrazolium (MTS) assay (Promega, (G3580); CellTitre 96 Aqueous One Solution Cell Proliferation Assay) as described by the manufacturer.

Clonogenic assays: The cells were seeded in six-well plates at 800 cells/well for MCF-7, ZR-75-1 and SK-BR-3. Cells were left in the incubator in the cultured medium for 14 days, following incubation; colonies were washed with phosphate buffered saline (PBS), fixed for 30 min with methanol and stained with crystal violet. Finally, the cloned cells were counted by using a microscope.

Cell cycle and apoptosis assays: The cells were harvested on the fifth day after transfection to detect and analyse the cell cycle distribution via the quantification of DNA content using a Propidium iodide flow cytometry kit (ab139-418, Abcam, UK). Moreover, on the same day, cells were stained for the detection of apoptosis using an Annexin V detection kit (ab14082, Abcam, UK). Samples were analysed using a MACSQuant® analyzer flow cytometer (Miltenyi Biotec, Bergisch Gldbach, Germany) and the data were analysed by FlowJo software (version 14.0.0.0., USA).

Static adhesion assay: Endothelial cells (HUVECs and DLECs) were seeded to be confluent in a 24-well plate. Tumour cell adhesion was assessed following cell labelling with 1 μ*M* Cell Tracker Green CMFDA (Invitrogen, C2925) and incubation at 37°C for 30 min. After labelling, tumour cell media supplemented with serum was used to re-suspend the tumour cells and incubated with endothelial cell monolayers for 35 min at 37°C. Non-adherent cells were washed with tumour cell media, and the adherent tumour cells were counted using a fluorescence microscope (Lecia DMI 3000B, Leica microsystems, Germany) at 10× magnification. In the central area of the well, which was marked manually with grid lines on the bottom of the plate, five fields of view were counted. The results were expressed as the absolute number of cells adhering to the endothelial layer and as the percentage of cells adhering with respect to the control.

Transmigration assay: A confluent endothelial cell monolayer was grown on hanging transwell inserts (Sigma, MCEP24H48). The confluency and integrity of the endothelial cell barrier were shown by preventing lucifer yellow leakage (Invitrogen, L453). After labelling with 1 μ*M* of Cell Tracker Green CMFDA (Invitrogen, C2925), tumour cell transmigration was assessed. After 16 h, transmigration was monitored by counting cells on the bottom of the chamber using a fluorescence microscope (Lecia DMI 3000B, Leica microsystems, Germany).

For further assessment of *CCNB2* in BC outcome and its association with LVI, we have used large cohorts obtained from BC patients to investigate the role of *CCNB2* at the transcriptomic and proteomic levels.

CCNB2 mRNA expression

Two datasets, the METABRIC [8] and TCGA breast carcinoma [9], were used to assess *CCNB2* mRNA expression and its association with clinicopathological parameters and patient outcome. In the METABRIC, to analyse the extracted mRNA obtained from primary tumour samples, the Illumina Human HT-12 v3 platforms (Illumina, Inc., San Diego, USA) were used. In TCGA, cBioPortal website was used for information about RNASeqV2 data and clinicopathological parameters [11, 12].

CCNB2 protein expression

Tissue samples were derived from the well-characterised Nottingham invasive BC cohort (n=3,173). For each patient, robust clinicopathological profile was readily available, including age at diagnosis, primary tumour size, tumour stage, histological grade, Nottingham Prognostic Index (NPI), and LVI status (Supplementary Table 1). Data on breast cancer specific survival (BCSS) and time to distant metastasis (TTDM) were also available. The BCSS for each patient was calculated in months from the date of the primary surgical treatment to the time of death from BC, while TTDM for each patient was calculated in months from surgery to the occurrence of the first distant metastasis. None of the patients in this study received neoadjuvant therapy as previously described [13].

This cohort had data on oestrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor 2 (HER2) and Ki67 [14-17]. BC molecular subtypes were defined using IHC profiles: luminal A ($ER^+/HER2^-$; Ki67 <10%), luminal B: ($ER^+/HER2^-$; Ki67 ≥10%), HER2-positive class (HER2 positive regard-less of ER or Ki67 status), and TNBC (ER^-, PR^- and $HER2^-$). To further understand the molecular interactions of *CCNB2*, the association with epithelial mesenchymal transition (EMT) related markers, including E-cadherin, N-cadherin, P-cadherin, TGFβ1, and TWIST2, which are available in our database [18, 19], was investigated.

For IHC staining of the primary rabbit monoclonal anti-CCNB2 antibody (ab185622, Abcam, UK), tissue microarrays (TMAs) of the study cohort were prepared using a TMA Grand Master® (3D HISTECH®, Budapest, Hungary) [20, 43]. As per the manufacturer's recommendations, antigen retrieval was performed (citrate buffer pH 6.0 at 1000 W for 20 min using microwave energy). The expression of CCNB2 protein was evaluated by Novocastra Novolink Polymer Detection Systems kit (Code: RE7280-K, Leica, Biosystems, UK), which involved incubating 4 μm sections with CCNB2 antibody (dilution 1:50) in Leica antibody diluent (RE AR9352, Leica, Biosystems, UK) for 60 minutes. Normal kidney tissue was used as a negative control, while normal colon tissue was used as the positive control (**Figure 1A**, **1B**). High resolution digital scanned images of the stained TMAs were obtained by a NanoZoomer scanner (NanoZoomer; Hamamatsu Photonics, Welwyn Garden City, UK) at 20× magnification and viewed using Xplore software (Philips, UK) to assess CCNB2 expression. Immunoreactivity staining was evaluated using a modified histochemical score (H-score) based on a semi-quantitative scoring. The entire field inspection was scored, and the cytoplasmic staining intensity was grouped as follows: score 0= negative, score 1= weak staining, score 2= moderate staining and score 3= strong staining. The percentage of each group was estimated (0-100%). Multiplying the intensity of staining and the percentage of staining provides an H-score, which has a range of 0-300 [13]. The TMAs were scored by two observers, and the interclass correlation coefficient (ICC) test was used to evaluate the concordance rate of the *CCNB2*.

Statistical analysis

GraphPad Prism software version 3.02 and SPSS version 24 (Chicago, IL, USA) were used for statistical analysis. *In vitro* data were represented as the mean ± standard error of the mean (SEM) of three independent experiments conducted in triplicate. A student's *t*-test was performed to determine the significant differences between the control and KD of *CCNB2*.

The association with clinicopathological parameters was evaluated using continuous data on

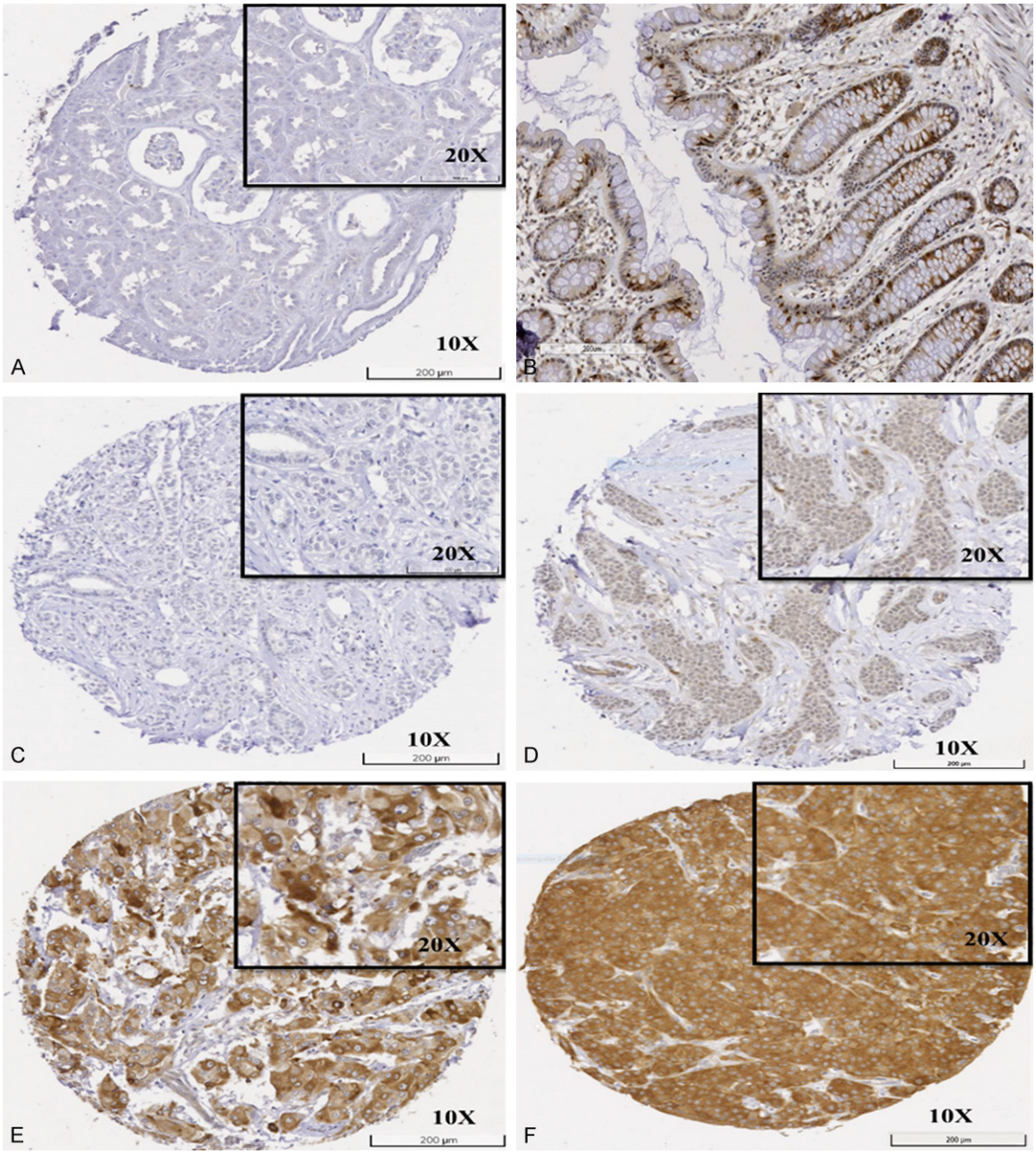


Figure 1. Immunohistochemistry expression of CCNB2 in tissue. A. Negative control of normal kidney tissue; B. Positive control of normal colon tissue; C. Negative expression in invasive breast carcinoma; D. Low expression in invasive breast carcinoma; E. Moderate expression in invasive breast carcinoma; F. High expression in invasive breast carcinoma.

CCNB2 mRNA and protein level. To investigate the differences between three or more groups, one-way analysis of variance (ANOVA) with the post-hoc Tukey multiple comparison test (for parametric data) or Kruskal-Wallis test (for non-parametric distribution) was used. Student t-test (parametric data) or Mann-Whitney test (non-parametric distribution) was used to analyse the differences between two groups. The median was used to categorise the expression of *CCNB2* mRNA and protein, with H-score of 20 was used for protein. The association between dependent and independent variables was analysed using logistic regression analysis. Correlation analysis was performed using the Spearman correlation test. The Log-rank

Table 1. Top 100 concordant genes by ANN from the METABRIC and TCGA breast cancer cohorts

Concordant top 100 genes		
DSCC1	AK124197	ATAD2
CCNE2	TCTN3	ATP2A2
YWHAZ	GPR34	ATP6V1C1
CCNB2	TNRC6A	AURKA
LRCH3	AGAP6	AZIN1
BX116720	AK129555	BBS5
EIF4H	LIN52	BDH2
CCNB1	ACIN1	BRI3BP
AK022140	CLASP1	C1orf91
DPRXP4	AL359560	C1RL
AK130741	PHKA2	CASC5
AK130706	ANAPC10	CCDC130
AK025793	BX093900	CDC25A
BX101409	DA572426	CDC6
BX111162	EBLN2	CDCA5
AL133627	PTBP1	CENPI
AX747098	ZNF493	CKAP5
RRM2	TYK2	CLK1
MLL4	BU624523	CLK4
SPTLC1	PUM1	CPNE3
MLL	DLC1	CSE1L
ERP29	CXorf56	DCAF13
NKTR	DNAH1	DERL1
PML	DR979451	DFNB59
RBM33	ARPC3	DTL
ASXL2	TRPV1	ECHDC2
CD237904	EPG5	ESRP1
CRYGS	AP1S1	FAM83D
BROX	C11orf73	FANK1
SLC25A3	C1R	FBXW4
ALDOB	SOLH	FES
MAT2A	AI655567	FLVCR1
AK129699	ADC	
SRRT	ANKRD10	

and Kaplan-Meier curve tests were used for univariate survival analysis, whilst the Cox Regression model was used for multivariate survival analysis. Statistical significance was defined as *P*-value <0.05.

This study followed the reporting recommendations for tumour markers prognostic studies (REMARK) criteria [21] (Supplementary Table 2).

Results

ANN analysis of BC cohorts to identify LVI associated gene

The result of the analysis was ranked in an order based on the high predictive accuracies and lowest test errors. This resulted in a final model containing the top 100 transcripts that most accurately classified based on LVI status (**Table 1**). *CCNB2* was chosen in this study as it was highly ranked with LVI positivity in both the METABRIC and TCGA cohorts. Other top ranked genes were either associated with LVI positivity in one cohort or associated with LVI negativity. Literature review revealed that *CCNB2* is a key gene controlling cell proliferation and its expression is altered in many cancers [22-24].

Preclinical studies (in vitro investigation of the role of CCNB2 in LVI-related biological processes)

Repression of CCNB2 expression using siRNA: Based on the expression level, MCF-7 (ER^+/PR^+/$HER2^-$), ZR-75-1 (ER^+/$PR^\pm$/$HER2^-$) and SK-BR-3 ($HER2^+$) BC cell lines will be used in this study (**Figure 2A**, **2B**). MCF-7 and ZR-75-1 represent luminal molecular classes, while SK-BR-3 represents $HER2^+$ tumour, which are associated with the highest incidence of LVI [25]. To test the efficacy of *CCNB2* KD, two independent siRNAs targeting *CCNB2* (IDs: s17446 and s17447) were used and compared to a non-targeting scrambled control siRNA (s10873). Hence, the negative control represents the effect of the overexpression of *CCNB2*, whereas the *CCNB2* siRNA represents the downregulation effect of *CCNB2* on the selected BC cell lines. Similar KD was observed with both siRNA targeting (**Figure 2C-H**), and siRNA (s17446) was prioritised for the subsequent functional assessments. Using WB, complete loss of *CCNB2* protein expression in MCF-7, ZR-75-1 and SK-BR-3 (**Figure 2I-K**) was observed when comparing *CCNB2* KD expression relative to GAPDH expression.

Downregulation of CCNB2 suppresses cell migration, proliferation and cell growth: KD of *CCNB2* showed a larger unhealed wound compared to the negative controls (NC) in MCF-7, ZR-75-1 and SK-BR-3 (P=0.0377, P=0.0028 & P=0.0344, respectively, **Figure 3A-C**). This sup-

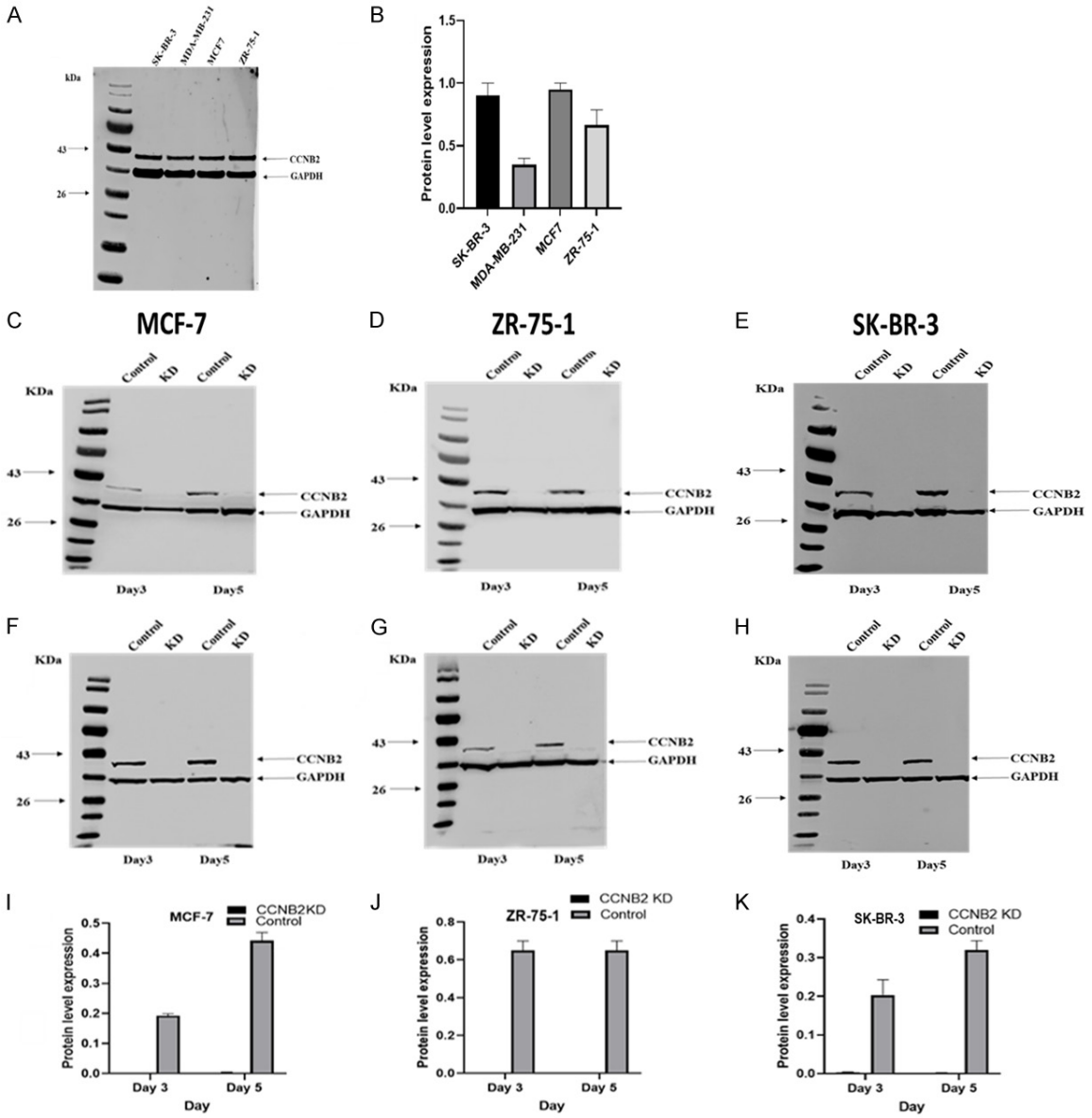


Figure 2. CCNB2 protein expression in breast cancer cell lines (A) Evaluation of differential expression of *CCNB2* in BC cell lines by western blotting (B) quantification of CCNB2 protein level expression in, SKBR-3, MDA-MB-231, MCF-7, and ZR-75-1. Using two pre-validated silencer select siRNAs constructs mainly for *CCNB2* or scrambled negative control siRNA, silencing *CCNB2* using the forward transfection method (IDs: s17447) relative to a non-targeting scrambled control siRNA (s10873) showed complete knockdown in (C) MCF-7 (D) ZR-75-1 and (E) SK-BR-3. Silencing CCNB2 using siRNA (IDs: s17446) relative to a non-targeting scramble control siRNA (s10873) showed complete knockdown in (F) MCF-7, (G) in ZR-75-1 and (H) SK-BR-3. Quantification of CCNB2 protein level expression in the transfected cell and the negative control by densitometry and normalized to GAPDH levels revealed complete loss of CCNB2 protein expression in (I) MCF-7, (J) ZR-75-1 and (K) SK-BR-3. The figures are representative of three or more independent experiments.

ported the role of *CCNB2* in promoting cell migration.

CCNB2 KD significantly reduced the proliferation rate in MCF-7 (**Figure 4A**; P=0.0003 (T48) & P=0.0343 (T72)), ZR-75-1 (**Figure 4B**; P=0.0016 (T24), P<0.0001 (T48) & P=0.0015 (T72)) and SK-BR-3 (**Figure 4C**; P=0.0296 (T24), P<0.0001 (T48) & P=0.0202 (T72)) compared to the control. Additionally, the ability of single cell colony to survive after it was exposed to the transfection agent was significantly

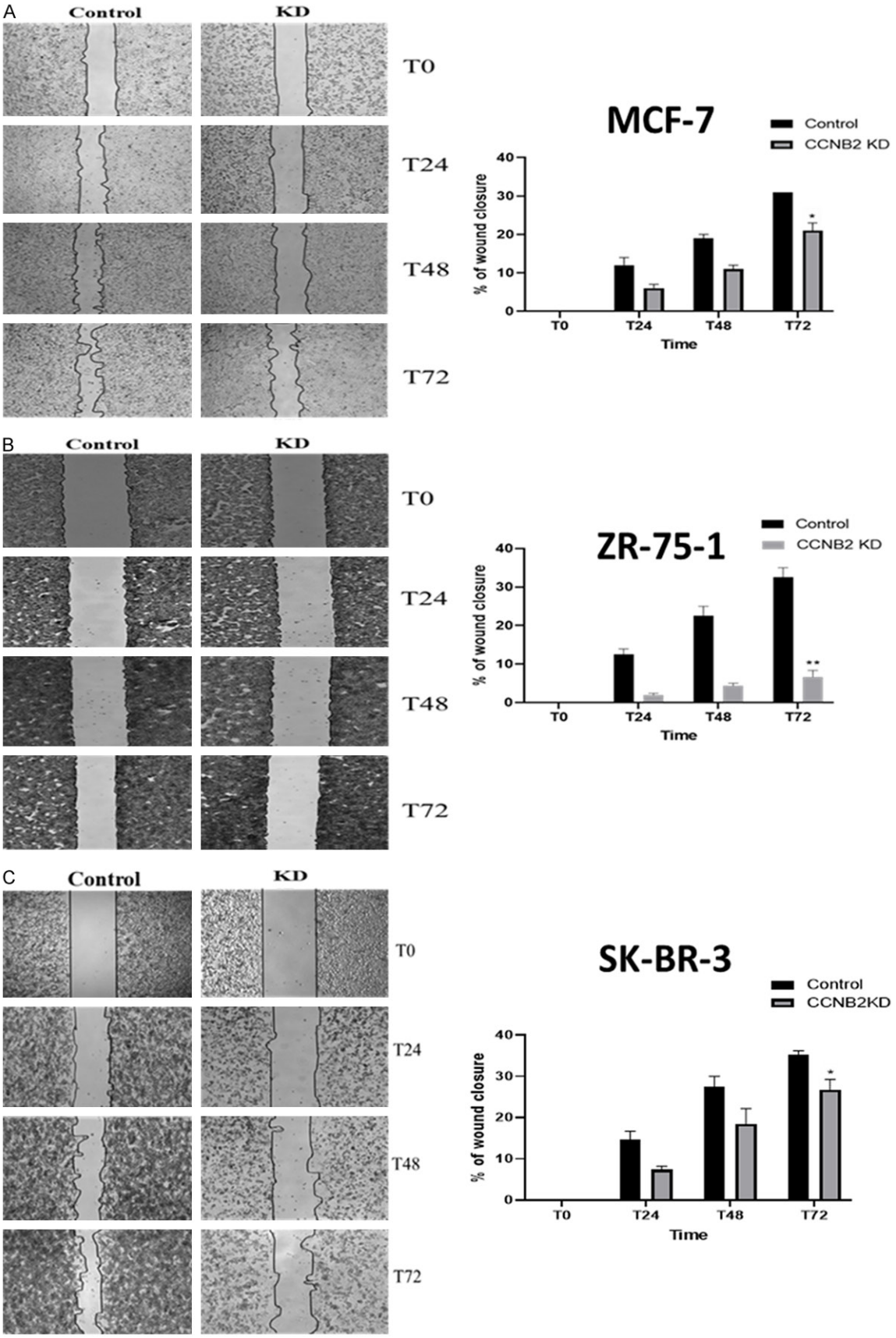

A
Control
KD
T0
T24
T48
T72
MCF-7
Control
CCNB2 KD
% of wound closure
Time
B
ZR-75-1
C
SK-BR-3
CCNB2KD

Figure 3. The effect of knockdown (KD) *CCNB2* by siRNA on cell migration in (A) MCF-7, (B) ZR-75-1 and (C) SK-BR-3 cells. The wound repair rate of *CCNB2* KD and control cells was observed by measuring the width of the gap left unhealed at T0, T24, T48 and T72. (A-C) silencing *CCNB2* was significantly reduced the migration rate in BC cell lines as detected by wound healing assay. Results shown are mean ± standard error of the mean (SEM) of three independent experiments. The *P*-values *<0.05, **<0.01, ***<0.001 and ****<0.0001.

weaker than the control on MCF-7 (**Figure 4D**; P=0.0075), ZR-75-1 (**Figure 4E**; P=0.0034) and SK-BR-3 (**Figure 4F**; P=0.0198).

Downregulation of CCNB2 arrests cells in S phase and enhances cell apoptosis: KD of *CCNB2* caused S phase arrest in MCF-7, ZR-75-1 and SK-BR-3, inhibiting the transition to the G2/M phase during cell cycle. The percentage of MCF-7 cells in the S phase were significantly increased in the *CCNB2* KD cells (G1 50.5%, S 22%, and G2/M 25.4%) compared with the NC (G1 69.2%, S 10.5%, and G2/M 13.4%) (P=0.032). Similar observations were observed with ZR-75-1 (G1 70.3%, S 15%, and G2/M 14%) compared with control (G1 68.3%, S 8.7%, and G2/M 12.3%) (P=0.9916 in G1 phase, P=0.0248 in S phase, P=0.0309 in G2/M phase). Similarly, the percentage of SK-BR-3 cells in the S phase were significantly increased in the *CCNB2* KD cells (G1 46%, S 35.2%, and G2/M 18.5%) compared with the NC (G1 51%, S 26.5%, and G2/M 15.7%) (P=0.6413 in G1 phase, P=0.0102 in S phase, P=0.0299 in G2/M phase) (**Figure 4G-I**). The number of apoptotic cells in *CCNB2* KD cells increased significantly in comparison with NC in early and late apoptosis as follows: early apoptosis (KD 5.23 vs NC 2.80%); (KD 23.2 vs NC 5%); (KD 4.55 vs NC 0.88%) while in late apoptosis (KD 2.87 vs NC 0.91%); (KD 12.1 vs NC 4.93%); (KD 9.77 vs NC 5.86%) in MCF-7 P=0.0307 (early apoptosis); P=0.0241 (late apoptosis); ZR-75-1 P<0.0001 (early apoptosis); P=0.0132 (late apoptosis) and SK-BR-3 P=0.0035 (early apoptosis); P=0.0100 (late apoptosis) respectively (**Figure 4J-L**).

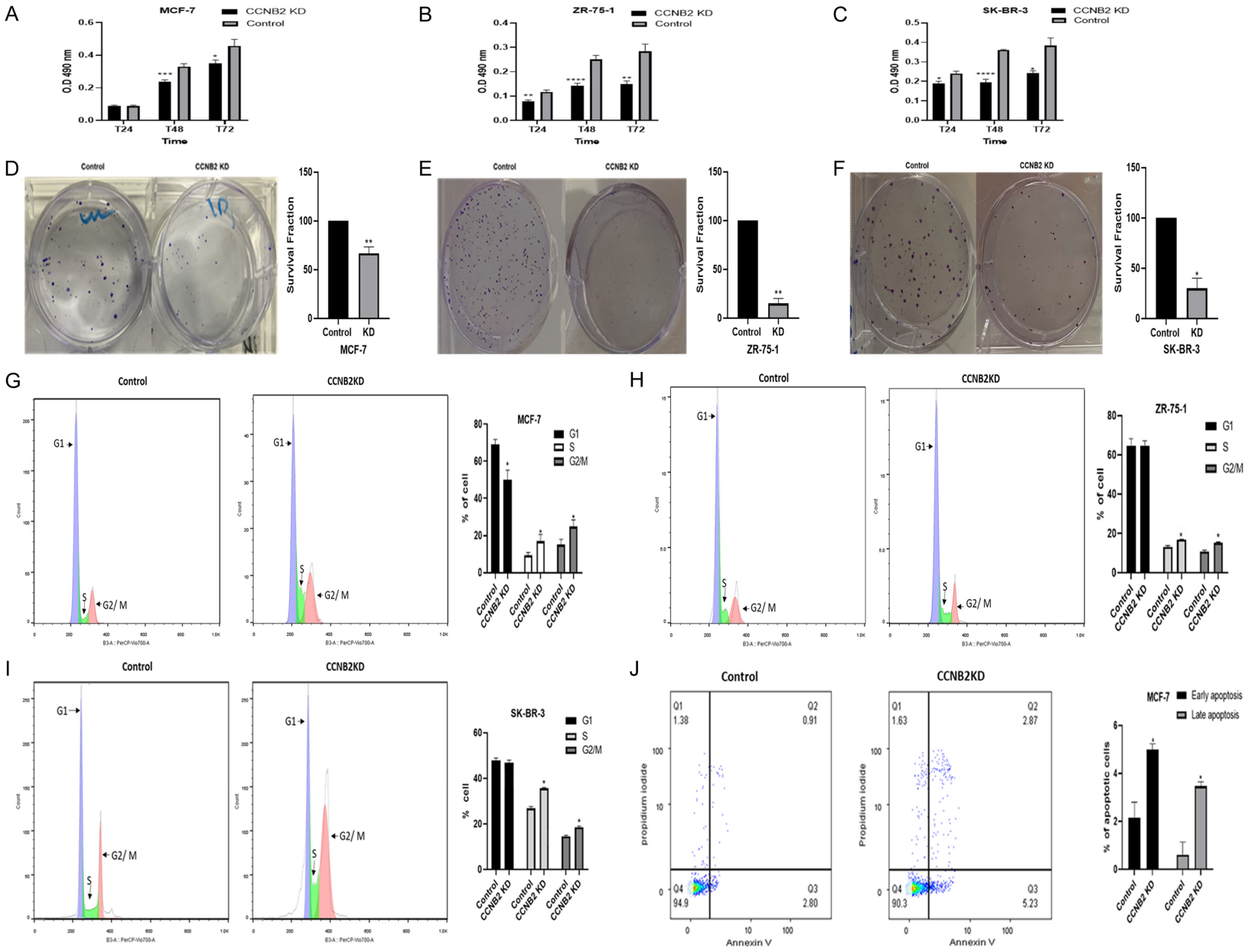
CCNB2 enhances BC cell adhesion and transmigration to the endothelial cells: All MCF-7, ZR-75-1 and SK-BR-3 cells lines showed higher adherence percentage to both HUVECs and DLECs compared to the *CCNB2* transfected cells (MCF-7 both with HUVEC and DLEC P=0.0003), (ZR-75-1 with HUVEC P=0.0298; with DLEC P=0.0026) and (SK-BR-3 with HUVEC P=0.0377; with DLEC P=0.0171) (**Figure 5A-C**). Similar results were observed in the transmigration assay. The number of the tumour cells transmigrated across endothelial cells in the NC group was higher than the KD group ((MCF-7 with HUVEC P=0.0240; with DLEC P=0.0170), (ZR-75-1 with HUVEC P= 0.0471; with DLEC P=0.0002) and (SK-BR-3 with HUVEC P=0.0077; with DLEC P=0.0091)) (**Figure 5D**, **5E**).

The role of CCNB2 in LVI and BC outcome in the clinical BC cohorts

CCNB2 mRNA and protein expression: Univariate and multivariate analysis in the METABRIC & TCGA datasets, demonstrated a significant association between high *CCNB2* mRNA expression and LVI-positivity (P<0.0001). High *CCNB2* mRNA expression was also correlated with features of aggressive tumour behaviour in both datasets. These included high histological grade, large tumour size, hormone receptor negativity (all; P<0.0001), and $HER2^{+}$ (P<0.0001 in METABRIC and P=0.002 in TCGA). In the METABRIC cohort, high expression of *CCNB2* mRNA was significantly associated with the luminal type and HER2 enriched molecular classes (P<0.0001) (**Table 2**). BCSS of BC patients with high *CCNB2* mRNA expression was significantly shorter than those with low expression in the METABRIC cohort, (P<0.0001) (**Figure 6A**) and TCGA dataset (P=0.001) (**Figure 6B**).

CCNB2 protein expression was observed in the cytoplasm of BC cells, with expression levels varying from negative to strong (**Figure 1C-F**). Good concordance was observed between the two scorers in *CCNB2* immuno-scoring (ICC=0.861, *P*<0.0001). High CCNB2 protein expression was observed in 506/1046 (53%) of LVI positive cases.

Similar to the transcriptomic results, high CCNB2 protein expression was significantly associated with clinicopathological parameters characteristic of aggressive behaviour, including higher histological grade, poorer NPI (all P<0.0001), larger tumour size (P=0.032), ER negativity (P=0.001) PR negativity (P=0.007)


A
MCF-7
CCNB2 KD
Control
O.D 490 nm
T24
T48
T72
Time
B
ZR-75-1
C
SK-BR-3
D
Control
CCNB2 KD
Survival Fraction
Control
KD
MCF-7
E
ZR-75-1
F
SK-BR-3
G
CCNB2KD
G1
S
G2/ M
% of cell
H
I
J
propidium iodide
Annexin V
Q1
Q2
Q3
Q4
1.38
0.91
94.9
2.80
1.63
2.87
90.3
5.23
Early apoptosis
Late apoptosis
% of apoptotic cells

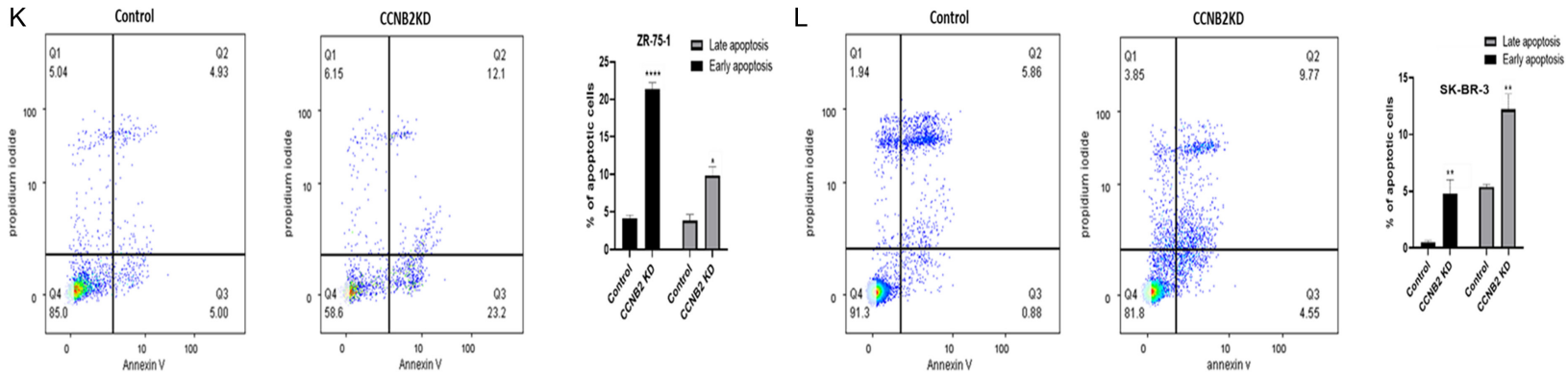
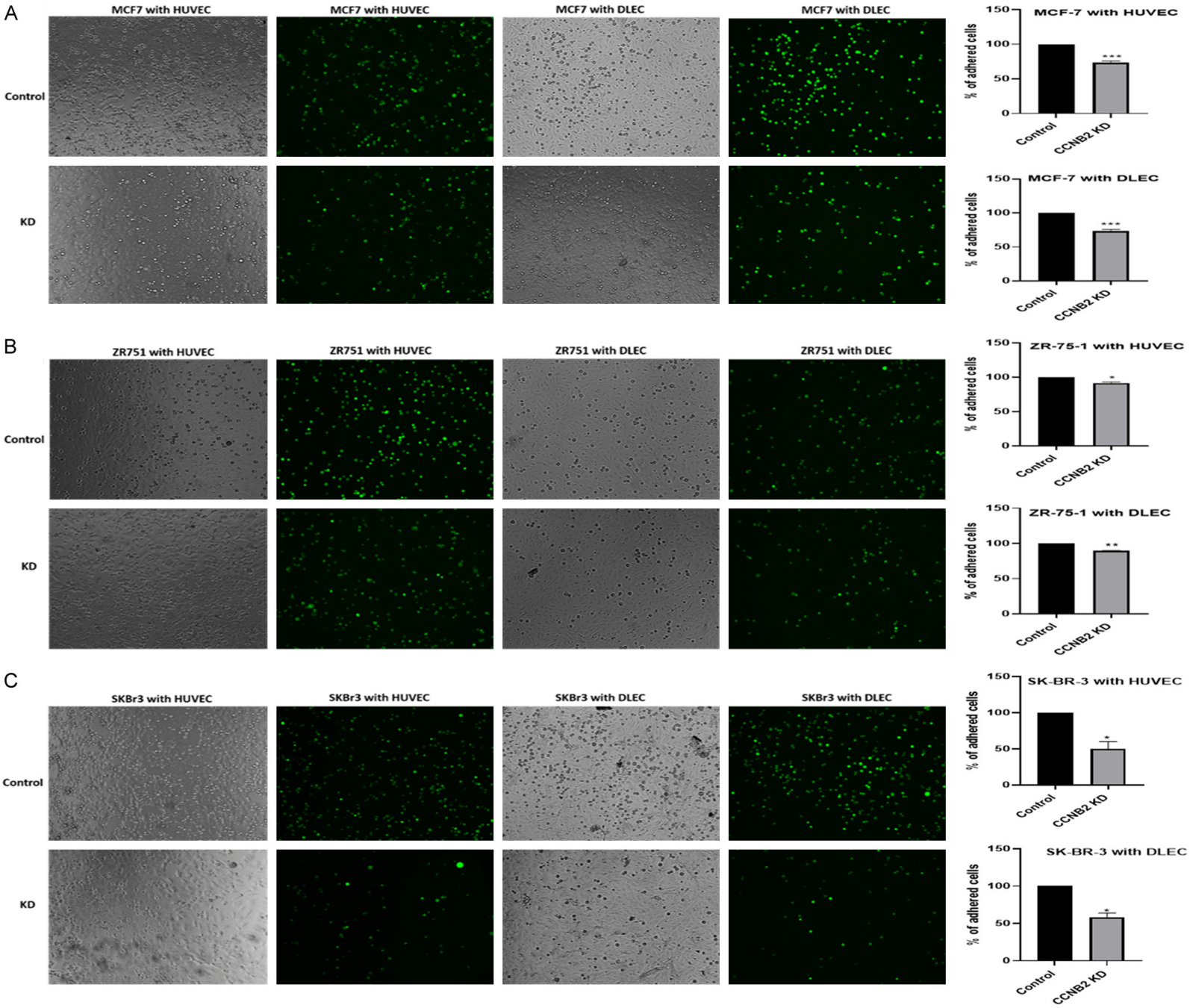


Figure 4. The effect of knockdown (KD) *CCNB2* by siRNA on cell proliferation in (A) MCF-7, (B) ZR-75-1 and (C) SK-BR-3 cells. (A-C) Cell proliferation was significantly reduced after KD in BC cell lines as detected by MTS assay. *CCNB2* siRNA transfection reduced the ability of BC cell lines to colonise in (D) MCF-7, (E) ZR-75-1 and (F) SK-BR-3 as detected by colony formation assay. The effect of knockdown (KD) *CCNB2* by siRNA on cell cycle and apoptosis. Data of BC cells showed arrest in S phase of the cell cycle of (G) MCF-7, (H) ZR-75-1 cells and (I) SK-BR-3. (G-I) Silencing *CCNB2* all BC cell lines increased the S phase as detected by flow cytometry. Data of BC cells showed KD of *CCNB2* enhanced apoptosis in (J) MCF-7, (K) ZR-75-1 cells and (L) SK-BR-3. (J-L) More apoptotic cells in the transfected cells were observed as detected by flow cytometry. Results shown are mean ± standard error of the mean (SEM) of three independent experiments. The *P*-values *<0.05, **<0.01, ***<0.001 and ****<0.0001.


A
MCF7 with HUVEC
MCF7 with HUVEC
MCF7 with DLEC
MCF7 with DLEC
Control
KD
MCF-7 with HUVEC
MCF-7 with DLEC
% of adhered cells
150
100
50
0
***
Control
CCNB2 KD
B
ZR751 with HUVEC
ZR751 with HUVEC
ZR751 with DLEC
ZR751 with DLEC
Control
KD
ZR-75-1 with HUVEC
ZR-75-1 with DLEC
*
**
C
SKBr3 with HUVEC
SKBr3 with HUVEC
SKBr3 with DLEC
SKBr3 with DLEC
Control
KD
SK-BR-3 with HUVEC
SK-BR-3 with DLEC

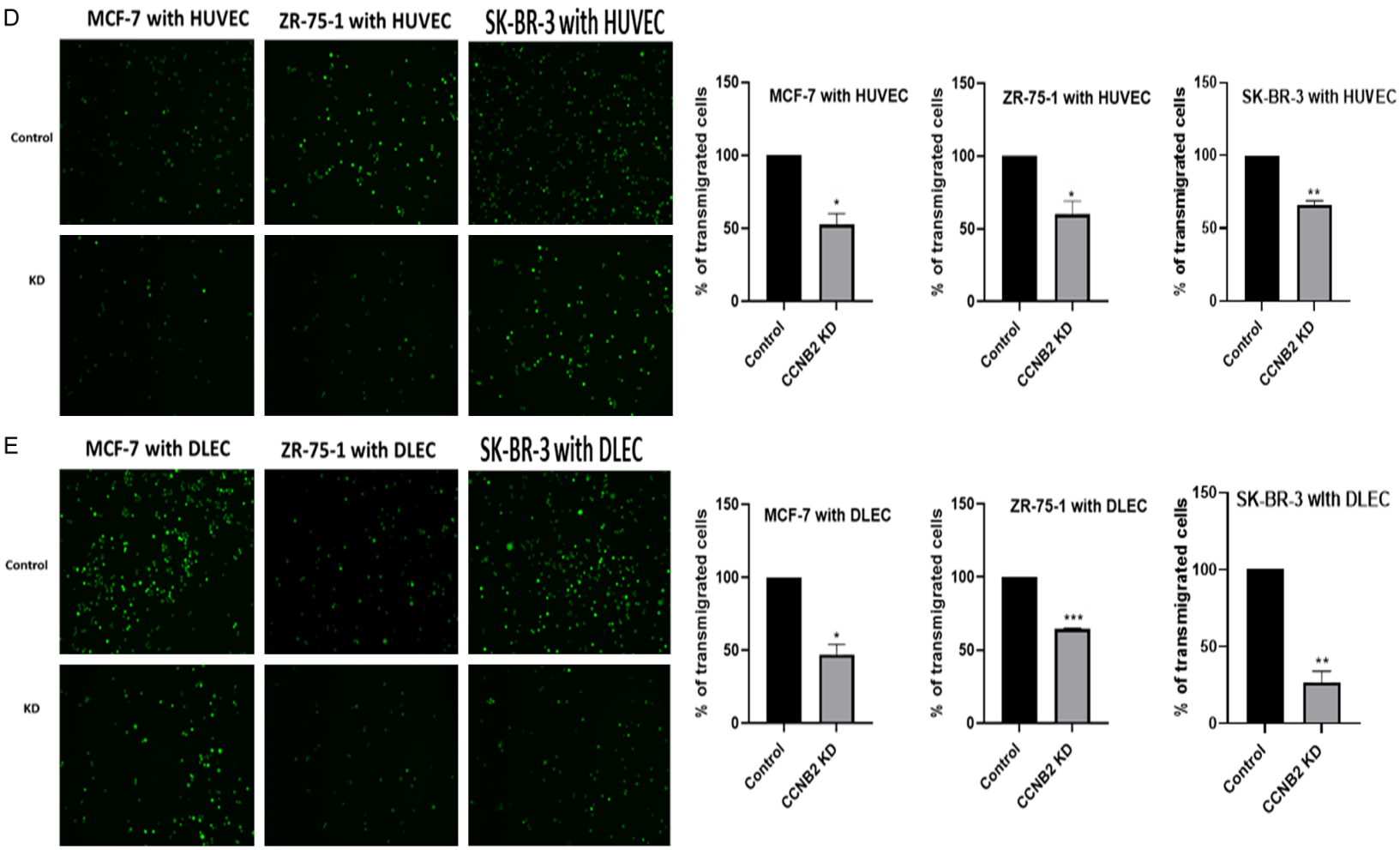


Figure 5. Representative photomicrographs of tumour cell adhesion across vascular and lymphatic endothelial cells (HUVECs, DLECs) (A) MCF-7, (B) ZR-75-1 cells and (C) SK-BR-3. (A-C) Silencing *CCNB2* decreased the number of all BC cell lines adhered with HUVECs and DLECs. Representative photomicrographs of tumour cell (MCF-7, ZR-75-1 and SK-BR-3) transmigration across (D) HUVECs and (E) DLECs. (D, E) The number of tumour cells transmigrated across HUVECs and DLECs was higher in the control group compared to the KD group. Results shown are mean ± standard error of the mean (SEM) of three independent experiments. The *P*-values *<0.05, **<0.01, ***<0.001 and ****<0.0001.

Table 2. Statistical associations between *CCNB2* mRNA expression and clinic-pathological parameters in the METABRIC (n=1980) and TCGA (n=854) breast carcinoma datasets

Parameters	*CCNB2* mRNA (METABRIC)			*CCNB2* mRNA (TCGA)		
	Number (%)	Mean Rank	*P*-value	Number (%)	Mean Rank	*P*-value
Patient Age (year)						
≤50	424 (21.4)	8.25	**<0.0001**	231 (27)	466.3	**0.005**
>50	1556 (78.6)	7.95		623 (73)	413.1	
Tumour Size						
≤2 cm	622 (31.7)	7.58	**<0.0001**	239 (28)	362.5	**<0.0001**
>2 cm	1338 (68.3)	8.09		615 (72)	452.8	
Tumour Grade						
1	170 (9)	7.16	**<0.0001**	89 (11)	198.5	**<0.0001**
2	770 (40.6)	7.66		375 (46)	308.3	
3	952 (50.4)	8.50		352 (43)	568.4	
Nodal Status						
Negative	1035 (52.5)	7.90	**<0.0001**	426 (51)	418.5	0.439
Positive	938 (47.5)	8.14		423 (49)	431.5	
Lympho-vascular Invasion						
Negative	930 (59.4)	7.93	**<0.0001**	559 (65)	392.5	**<0.0001**
Positive	635 (40.6)	8.10		295 (35)	493.8	
Oestrogen Receptor						
Negative	474 (23.9)	8.77	**<0.0001**	185 (22)	625.7	**<0.0001**
Positive	1506 (76.1)	7.78		639 (78)	350.8	
Progesterone Receptor						
Negative	940 (47.4)	8.53	**<0.0001**	272 (33)	552.1	**<0.0001**
Positive	1040 (52.6)	7.71		546 (67)	338.4	
HER2 Status						
Negative	1733 (87.5)	7.94	**<0.0001**	567 (81)	339.1	**0.002**
Positive	247 (12.5)	8.52		133 (19)	399.1	
Immunohistochemistry subtypes						
$ER^+/HER2^-$ Low Proliferation	368 (36.9)	8.74	**<0.0001**	Not available		
$ER^+/HER2^-$ High Proliferation	368 (36.9)	8.37				
Triple Negative	151 (15.1)	7.19				
$HER2^+$	110 (11.1)	8.52				

Significant *P* values are in bold.

and HER2 positivity (P<0.0001) (**Table 3**). Positive LVI status was associated with high *CCNB2* expression, which was an independent of other clinical parameters (P=0.038) (**Table 4**).

Survival analysis revealed that higher CCNB2 protein expression was associated with shorter BCSS (P=0.042, **Figure 6C**) and TTDM (P=0.039, **Figure 6D**). When we stratified the cohort into molecular subtypes, high expression of *CCNB2* was significantly associated with poor outcome in the ER positive (luminal A) tumours (P=0.007, **Figure 6E**) and HER2 positive tumours (P=0.028, **Figure 6H**) but not in the TNBC (P=0.219, **Figure 6G**). Multivariate survival analysis revealed that high expression of *CCNB2* is an independent marker of shorter survival (P=0.045) regardless of the other variables, including tumour size, histological grade, lymph nodal stage and LVI (**Table 5**).

Correlation between CCNB2 expression and other related biomarkers: To further evaluate the role of *CCNB2* in BC and its interaction with other genes related to the various LVI related process, we interrogated the METABRIC and TCGA datasets for the correlation between

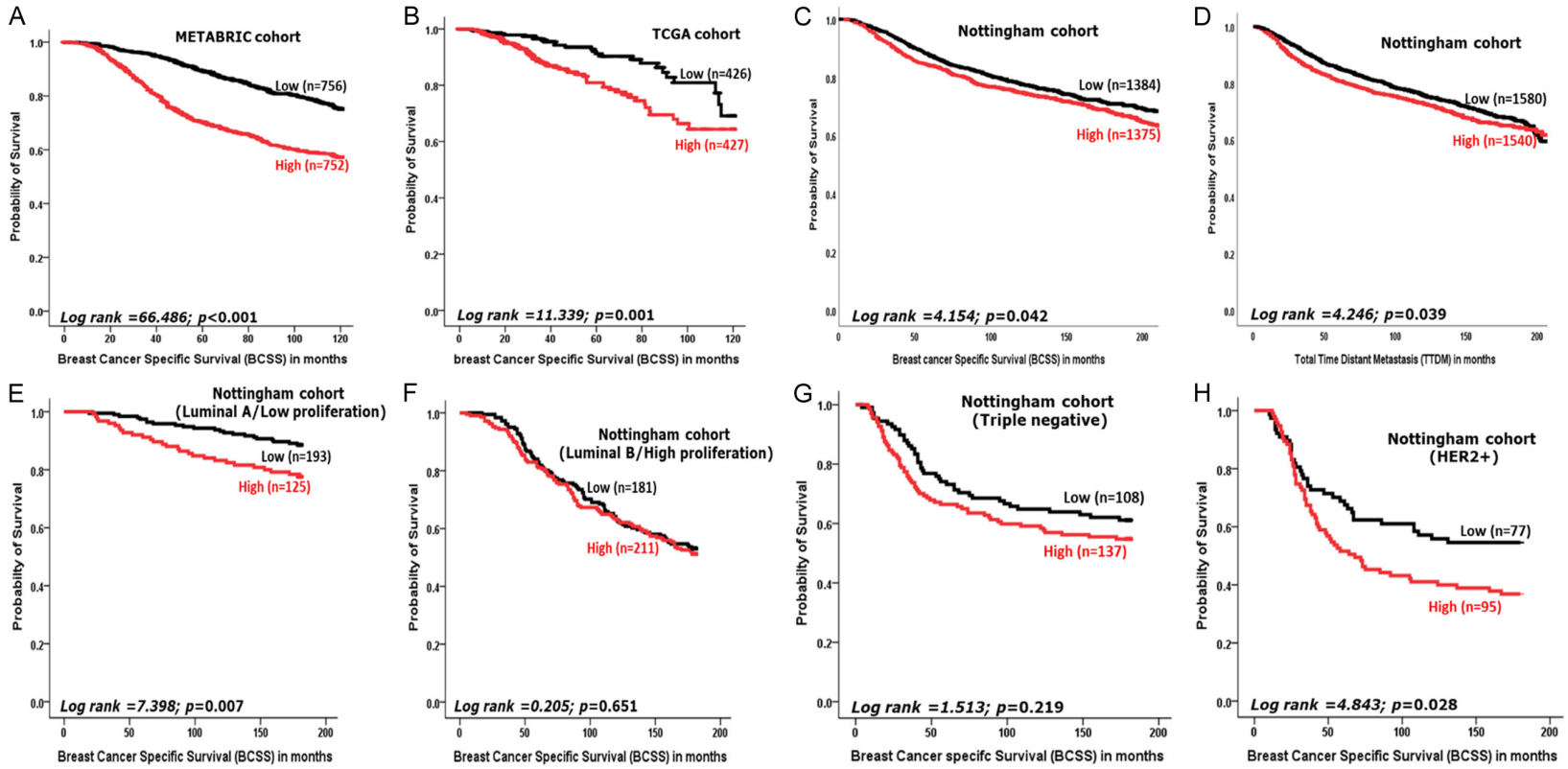


Figure 6. Kaplan-Meier survival plots showing the association between *CCNB2* (A) mRNA expression (METABRIC) (B) mRNA expression (TCGA) (C) protein expression in the Nottingham cohort and breast cancer specific survival (BCSS) (D) the protein expression and total time distant metastasis (TTDM) (E) protein expression in luminal A and BCSS (F) protein expression in luminal B and BCSS (G) protein expression in triple negative and BCSS and (H) protein expression in $HER2^{+}$ and BCSS.

Table 3. Statistical associations between CCNB2 protein expression and the clinic-pathological factors in Nottingham breast cancer cohort (n=3178)

Nottingham Breast Cancer Cohort			
Parameters	CCNB2 protein		
	Number %	Mean Rank	*P*-value
Patient Age (year)			
≤50	1042 (33)	1672.3	**<0.0001**
>50	2131 (67)	1545.3	
Tumour Size			
≤2 cm	1768 (56)	1555.6	**0.032**
>2 cm	1403 (44)	1624.3	
Tumour Grade			
1	543 (17)	1314.0	**<0.0001**
2	1131 (36)	1503.9	
3	1496 (47)	1745.7	
Mitosis			
1	1284(42)	1361.82	**<0.0001**
2	598 (19)	1597.05	
3	1211 (39)	1718.63	
Pleomorphism			
1	58 (1.8)	1213.87	**<0.0001**
2	1084 (35)	1368.07	
3	1949 (36.2)	1654.48	
Tubular Formation			
1	205 (6.6)	1282.29	**<0.0001**
2	984 (31.7)	1510.29	
3	1906 (61.7)	1596.05	
Lympho-vascular Invasion			
Negative	2205 (70)	1549.6	**0.004**
Positive	954 (30)	1650.1	
Axillary Nodal Stage			
1	1989 (63)	1563.5	0.136
2	903 (28)	1607.3	
3	277 (9)	1666.6	
Nottingham Prognostic Index			
Good	1041 (33)	1416.0	**<0.0001**
Moderate	1632 (51)	1659.3	
Poor	496 (16)	1695.0	
Oestrogen Receptor			
Negative	704 (22)	1677.0	**0.001**
Positive	2453 (78)	1550.9	
Progesterone Receptor			
Negative	1256 (41)	1585.5	**0.007**
Positive	1813 (59)	1500.0	
HER2 Status			
Negative	2677 (87)	1513.8	**<0.0001**
Positive	402 (13)	1714.3	
Triple Negative			
Negative	2630 (84)	1538.7	**0.007**
Positive	483 (16)	1656.5	

Significant *P* values are in bold.

CCNB2 mRNA and other genes involved in invasion, EMT and adhesion. This showed a correlation between *CCNB2* and many genes involved the EMT and stromal degradations biomarkers, including *E-cadherin*, *N-cadherin*, *P-cadherin*, and *GSK3B* (all *P*<0.0001). Moreover, there was a significant correlation between *CCNB2* and matrix metalloproteinase (MMPs), including *MMP1*, *MMP7*, *MMP9*, *MMP12* and *MMP20* (all P<0.0001) (**Table 6**). At the protein level, there was a positive correlation between *CCNB2* with E-cadherin, (P=0.0001), N-cadherin (P=0.048), P-cadherin, (P= 0.030), TGFβ1 (P=0.008) and TWIST2 (P=0.004), which are already available in our cohort (**Table 6**).

Using the GeneMANIA database (https://genemania.org/), a network was created to illustrate the molecular interactions between *CCNB2* and EMT-related markers. The network revealed that *CCNB2* genetically interacted with some EMT-related markers such as *CTTNB1* and was co-expressed with *NFKB1*, which has an interaction with other EMT-related markers including *E-cadherin, N-cadherin*, *P-cadherin*, and *TWIST1* (**Figure 7**).

Discussion

Metastasis is the main cause of BC-related mortality [26]. Thus, early detection of metastatic potential in BC patients can be critical in reducing cancer-related mortality. LVI is an early event in the development of metastasis and a potent prognostic factor [27]. LVI in BC, similar to other cancers, has attracted attention not only for its prognostic role but also for being a potential therapeutic target. However, the complex molecular mechanism of LVI and its overlap with other carcinogenesis pathways contribute to the difficulty in identifying the key driver gene of LVI that can be targeted. In our study, we identified a strong association between LVI and

Table 4. Binary regression for predictors of CCNB2 protein expression in Nottingham breast cancer cohort and other variables

Nottingham Breast Cancer Cohort				
Parameters	Hazard ratio (HR)	95% confident interval (CI)		Significance *P*-value
		Lower	Upper	
CCNB2 Protein Expression	1.190	1.010	1.402	**0.038**
Tumour Size	1.889	1.598	2.232	**<0.001**
Lymph Nodal Stage	2.793	2.465	3.166	**<0.001**

Significant *P* values are in bold.

Table 5. Multivariate Cox regression for predictors of breast cancer specific survival (BCSS) and CCNB2 protein expression in Nottingham breast cancer cohort

Nottingham Breast Cancer Cohort				
Parameters	Hazard ratio (HR)	95% confident interval (CI)		Significance *P*-value
		Lower	Upper	
CCNB2 Protein Expression	1.001	1.000	1.002	**0.045**
Tumour Size	1.511	1.311	1.743	**<0.001**
Lymph Nodal Stage	1.700	1.540	1.876	**<0.001**
Tumour Grade	1.533	1.377	1.708	**<0.001**
Lymphovascular Invasion Status	1.491	1.291	1.721	**<0.001**

Significant *P* values are in bold.

Table 6. Correlations of *CCNB2* expression with mRNA and protein expression of epithelial mesenchymal transition (EMT) and matrix metalloproteinases (MMPs) related genes

Gene names	METABRIC cohort		TCGA cohort		Nottingham cohort	
	Correlation value	*P* value	Correlation value	*P* value	Correlation value	*P* value
EMT related genes						
E-cadherin	-0.125	**<0.0001**	-0.140	**<0.0001**	-0.108	**<0.0001**
N-cadherin	0.151	**<0.0001**	0.133	**<0.0001**	0.095	**0.048**
P-cadherin	0.322	**<0.0001**	0.280	**<0.0001**	0.261	**0.030**
TGFB1	0.025	0.260	0.252	**<0.0001**	0.080	**0.008**
TWIST1	0.045	**0.045**	0.131	**<0.0001**	**Not available**	
TWIST2	0.270	**<0.0001**	0.214	**<0.0001**	0.087	**0.004**
ZEB1	0.281	**<0.0001**	0.420	**<0.0001**	**Not available**	
ZEB2	0.143	**<0.0001**	0.214	**<0.0001**		
NFKB1	0.176	**<0.0001**	0.254	**<0.0001**		
GSK3B	0.330	**<0.0001**	0.129	**<0.0001**		
CTNNB1	0.073	**0.001**	0.172	**<0.0001**		
MMPs related genes						
MMP1	0.328	**<0.0001**	0.428	**<0.0001**	**Not available**	
MMP7	0.134	**<0.0001**	0.190	**<0.0001**		
MMP9	0.308	**<0.0001**	0.227	**<0.0001**		
MMP11	0.051	**0.022**	0.064	**0.060**		
MMP12	0.382	**<0.0001**	0.350	**<0.0001**		
MMP15	0.199	**<0.0001**	0.104	**0.002**		
MMP20	0.092	**<0.0001**	0.134	**<0.0001**		
MMP25	0.165	**<0.0001**	0.073	**0.034**		

Significant *P* values are in bold.

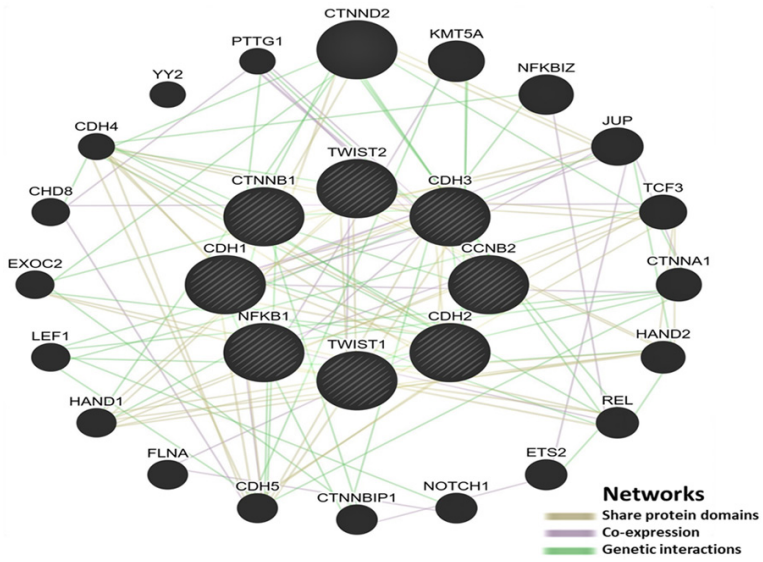


Figure 7. A schematic network illustrating the molecular interactions between *CCNB2* and EMT-related markers generated by GeneMANIA database.

CCNB2 utilising large cohorts of BC and ANN algorithms, which is a powerful technique with low-test errors. Likewise, recent studies using different bioinformatics and co-expression analyses confirmed the association of *CCNB2* with human cancer progression [22, 28, 29].

Cyclins are crucial elements of the cell-division cycle, and various studies, including Stamatakos et al., [30] have discovered that defects in their functionality can lead to carcinogenesis. Although evidence is still emerging, recent studies have revealed more information regarding the mechanistic pathways through which cyclins can influence the oncogenic potential of cells [22, 23]. Cyclins are indispensable core cell cycle regulators modulating the activation of cyclin dependent kinases (CDKs) complexes that propel cells through the cell cycle. Many of the cyclins are also identified as established oncogenes in several human tumours [30]. There are several different cyclins that are active in different parts of the cell cycle and that contribute to various other functions. The *CCNB2* gene, which is a member of the cyclin B family, has a role in the G2/M transition through cell division control (CDC2) activation, and its inhibition leads to cell cycle arrest [31, 32]. *CCNB2* is a critical part of cell cycle regulation. It is one of the central protein kinases, and when activated, it becomes a master regulator during M-phase transition, phosphorylation, and the activation of other protein kinases [33, 34]. The expression of *CCNB2* is normally tightly regulated [22]. However, high expression of *CCNB2* has been found in human tumours, including lung [24], liver [22], bladder [35] and BC [23]. *CCNB2* may act as an oncogene and could be a potential biomarker for unfavourable outcomes [23, 36]. This suggests that *CCNB2* may also be involved in BC progression. The role of *CCNB2* in LVI is unclear, and to our knowledge, this is the first study to report on the role of *CCNB2* in LVI.

In our preclinical studies, we have revealed that knocking down *CCNB2* significantly inhibits cell proliferation and reduces cell growth. As expected, silencing *CCNB2* arrested the S phase of the cell cycle, which inhibited the cells from entering the G2/M transition phase during the cell cycle in BC cell lines. Similarly, knocking down *CCNB2* promoted cell apoptosis. The failure to remove the dead cells via the inhibition of cell apoptosis causes the malignant proliferation of cancer cells [37]. A study showed that this occurs through the interference of proteolysis and the obstruction of normal intracellular trafficking [35]. In Lei et al.'s investigation of its role in bladder cancer, *CCNB2*'s tumourigenic function was linked to the dysregulation and abhorrent expression of cell-cycle associated proteins [35]. Li et al. investigated the effect of *CCNB2* in liver cancer and arrived at a similar view, concluding that *CCNB2* may promote cell apoptosis by causing S phase arrest and instigating the functional failure of the G2/M DNA/genetic damage checkpoint during the cell division cycle [22]. Checkpoints are important in repairing damaged DNA, as well as maintaining genomic integrity. Defects in this domain can lead to mutations and oncogenesis [38]. It is essential for the primary tumour cells to proliferate to invade the surrounding tissue and thus establish metastasis cascades, which can lead to the degradation of the basement membrane and LVI. Proliferation continues until the invasion of vascular or lymphatic channels occur. At

this stage, tumour cells can evade apoptosis [39]. Therefore, *CCNB2* is one of the essential regulators in cell proliferation, and a necessary prerequisite step in the development of LVI and, eventually, distant metastasis.

Although the role of *CCNB2* in the cell cycle and proliferation is well known, its role in driving LVI in BC is not characterised. Our results revealed that the silencing of *CCNB2* suppressed the migration of the cells and reduced the number of tumour cells adhered and transmigrated across the vascular and lymphatic endothelial cell lines. This was supported by the positive correlation between EMT and MMPs related markers. N-cadherin is linked to the EMT, which plays a vital role in the invasion and intravasation into the bloodstream and in the degradation of the extracellular matrix (ECM) caused by the production of proteases. Thus, an invasion into the stroma is a result of the loss of the linkage between the epithelial of the BC cell and other epithelial cells by the up-regulation of N-cadherin [40]. Moreover, P-cadherin, an important molecule that can activate integrin molecules, allows cancer cells to attach to ECM and activates the invasion of cancer cells [41], which could explain the role of *CCNB2* in LVI as it has a positive correlation with P-cadherin. To invade the tissue, tumour cells must penetrate the tissue boundaries by breaking down the ECM through MMPs and the urokinase plasminogen activator (uPA) system. Inhibition of uPA halts the invasion and expression of MMP9 [42]. The LVI tumour microenvironment is strongly correlated with MMPs, specifically MMP9 and MMP1, expression. These expressions are responsible for cancer cell intravasation and metastasis in BC [25]. Thus, our *in vitro* assessment indicated the contribution of the upregulation of *CCNB2* expression in migration, adhesion and transmigration through endothelial cell lines, which induced the LVI process. This is also consistent with a previous study of bladder cancer that reported that the silencing of *CCNB2* can restrain migration and invasion leading to inhibition of tumour metastasis and prolonging survival time [35].

Our study has explored the prognostic significance of *CCNB2* expression using multiple well-annotated BC cohorts. Overexpression is associated with aggressive clinicopathological BC features, including LVI positivity, high histological grade, large tumour size, hormonal receptor negativity and HER2 positivity, and worse clinical outcomes.

Because *CCNB2* is a cell cycle regulator, the cell cycle could be one of the mechanisms contributing to LVI in BC. However, the molecular mechanisms that cause BC tumour cells to invade and disseminate in vascular spaces are still largely unknown and require more research. Cancer treatment strategies that induce cell cycle arrest in cancer cells are effective [43]. This suggests that *CCNB2* is a key gene in BC, particularly given our findings regarding lymphatic invasion, which is a prerequisite for metastasis cascade. Thus, *CCNB2* may represent a potential target for LVI in BC patients.

Our study provides evidence that *CCNB2* is a key modulator of LVI in BC and could confer as a potential therapeutic target to suppress LVI and improve patient's outcome. However, *in vitro* and/or *in vivo* experiments investigating the differential gene expression/pathways observed in *CCNB2* knockdown cells are necessary.

Acknowledgements

We would like to acknowledge Dr. Omar Mohammed for his help and encouragement during this study. Abrar Aljohani is supported and funded by Taif University; Kingdom of Saudi Arabia. The authors are part of the PathLAKE digital pathology consortium. These new Centres are supported by a £ 50 m investment from the Data to Early Diagnosis and Precision Medicine strand of the government's Industrial Strategy Challenge Fund, managed and delivered by UK Research and Innovation (UKRI).

Informed consent was obtained from all individuals prior to surgery to use their tissue materials in research.

Disclosure of conflict of interest

None.

Address correspondence to: Emad A Rakha, Department of Histopathology, Nottingham University Hospital NHS Trust, City Hospital Campus, Hucknall Road, Nottingham, UK. Tel: +44-0115-9691169; Fax: +44-0115-9627768; E-mail: emad.rakha@nottingham.ac.uk; emad.rakha@nuh.nhs.uk

References

[1] Ganz PA and Goodwin PJ. Breast cancer survivorship: where are we today? Adv Exp Med Biol 2015; 862: 1-8.

[2] Rakha EA, Martin S, Lee AH, Morgan D, Pharoah PD, Hodi Z, Macmillan D and Ellis IO. The prognostic significance of lymphovascular invasion in invasive breast carcinoma. Cancer 2012; 118: 3670-3680.

[3] Zafeiris D, Rutella S and Ball GR. An artificial neural network integrated pipeline for biomarker discovery using alzheimer's disease as a case study. Comput Struct Biotechnol J 2018; 16: 77-87.

[4] Nevin L; PLOS Medicine Editors. Advancing the beneficial use of machine learning in health care and medicine: toward a community understanding. PLoS Med 2018; 15: e1002708.

[5] Swan AL, Stekel DJ, Hodgman C, Allaway D, Alqahtani MH, Mobasheri A and Bacardit J. A machine learning heuristic to identify biologically relevant and minimal biomarker panels from omics data. BMC Genomics 2015; 16 Suppl 1: S2.

[6] Libbrecht MW and Noble WS. Machine learning applications in genetics and genomics. Nat Rev Genet 2015; 16: 321-332.

[7] Abdel-Fatah TMA, Agarwal D, Liu DX, Russell R, Rueda OM, Liu K, Xu B, Moseley PM, Green AR, Pockley AG, Rees RC, Caldas C, Ellis IO, Ball GR and Chan SYT. SPAG5 as a prognostic biomarker and chemotherapy sensitivity predictor in breast cancer: a retrospective, integrated genomic, transcriptomic, and protein analysis. Lancet Oncol 2016; 17: 1004-1018.

[8] Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, Speed D, Lynch AG, Samarajiwa S, Yuan Y, Gräf S, Ha G, Haffari G, Bashashati A, Russell R, McKinney S; METABRIC Group; Langerød A, Green A, Provenzano E, Wishart G, Pinder S, Watson P, Markowetz F, Murphy L, Ellis I, Purushotham A, Børresen-Dale AL, Brenton JD, Tavaré S, Caldas C and Aparicio S. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature 2012; 486: 346.

[9] Ciriello G, Gatza ML, Beck AH, Wilkerson MD, Rhie SK, Pastore A, Zhang H, McLellan M, Yau C, Kandoth C, Bowlby R, Shen H, Hayat S, Fieldhouse R, Lester SC, Tse GM, Factor RE, Collins LC, Allison KH, Chen YY, Jensen K, Johnson NB, Oesterreich S, Mills GB, Cherniack AD, Robertson G, Benz C, Sander C, Laird PW, Hoadley KA and King TA; TCGA Research NetworkPerou CM. Comprehensive molecular portraits of invasive lobular breast cancer. Cell 2015; 163: 506-519.

[10] Mohammed RA, Martin SG, Mahmmod AM, Macmillan RD, Green AR, Paish EC and Ellis IO. Objective assessment of lymphatic and blood vascular invasion in lymph node-negative breast carcinoma: findings from a large case series with long-term follow-up. J Pathol 2011; 223: 358-365.

[11] Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, Antipin Y, Reva B, Goldberg AP, Sander C and Schultz N. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov 2012; 2: 401-404.

[12] Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E, Cerami E, Sander C and Schultz N. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal 2013; 6: pl1.

[13] Aljohani AI, Joseph C, Kurozumi S, Mohammed OJ, Miligy IM, Green AR and Rakha EA. Myxovirus resistance 1 (MX1) is an independent predictor of poor outcome in invasive breast cancer. Breast Cancer Res Treat 2020; 181: 541-551.

[14] Abd El-Rehim DM, Pinder SE, Paish CE, Bell JA, Rampaul RS, Blamey RW, Robertson JF, Nicholson RI and Ellis IO. Expression and co-expression of the members of the epidermal growth factor receptor (EGFR) family in invasive breast carcinoma. Br J Cancer 2004; 91: 1532-1542.

[15] Mahmoud SMA, Paish EC, Powe DG, Macmillan RD, Lee AHS, Ellis IO and Green AR. An evaluation of the clinical significance of FOXP3+ infiltrating cells in human breast cancer. Breast Cancer Res Treat 2011; 127: 99-108.

[16] Mahmoud SM, Lee AH, Paish EC, Macmillan RD, Ellis IO and Green AR. Tumour-infiltrating macrophages and clinical outcome in breast cancer. J Clin Pathol 2012; 65: 159-163.

[17] Rakha EA, Elsheikh SE, Aleskandarany MA, Habashi HO, Green AR, Powe DG, El-Sayed ME, Benhasouna A, Brunet JS, Akslen LA, Evans AJ, Blamey R, Reis-Filho JS, Foulkes WD and Ellis IO. Triple-negative breast cancer: distinguishing between basal and nonbasal subtypes. Clin Cancer Res 2009; 15: 2302-2310.

[18] Abd El-Rehim DM, Ball G, Pinder SE, Rakha E, Paish C, Robertson JFR, Macmillan D, Blamey RW and Ellis IO. High-throughput protein expression analysis using tissue microarray technology of a large well-characterised series identifies biologically distinct classes of breast cancer confirming recent cDNA expression analyses. Int J Cancer 2005; 116: 340-350.

[19] Rakha EA, Abd El Rehim D, Pinder SE, Lewis SA and Ellis IO. E-cadherin expression in invasive non-lobular carcinoma of the breast and its prognostic significance. Histopathology 2005; 46: 685-693.
[20] Sonbul SN, Aleskandarany MA, Kurozumi S, Joseph C, Toss MS, Diez-Rodriguez M, Nolan CC, Mukherjee A, Martin S, Caldas C, Ellis IO, Green AR and Rakha EA. Saccharomyces cerevisiae-like 1 (SEC14L1) is a prognostic factor in breast cancer associated with lymphovascular invasion. Mod Pathol 2018; 31: 1675-1682.
[21] McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M and Clark GM; Statistics Subcommittee of the NCI-EORTC Working Group on Cancer Diagnostics. Reporting recommendations for tumour maker prognostic studies (REMARK). Br J Cancer 2005; 93: 387-391.
[22] Li R, Jiang X, Zhang Y, Wang S, Chen X, Yu X, Ma J and Huang X. Cyclin B2 overexpression in human hepatocellular carcinoma is associated with poor prognosis. Arch Med Res 2019; 50: 10-17.
[23] Shubbar E, Kovács A, Hajizadeh S, Parris TZ, Nemes S, Gunnarsdóttir K, Einbeigi Z, Karlsson P and Helou K. Elevated cyclin B2 expression in invasive breast carcinoma is associated with unfavorable clinical outcome. BMC Cancer 2013; 13: 1.
[24] Qian X, Song X, He Y, Yang Z, Sun T, Wang J, Zhu G, Xing W and You C. CCNB2 overexpression is a poor prognostic biomarker in Chinese NSCLC patients. Biomed Pharmacother 2015; 74: 222-227.
[25] Aleskandarany MA, Sonbul SN, Mukherjee A and Rakha EA. Molecular mechanisms underlying lymphovascular invasion in invasive breast cancer. Pathobiology 2015; 82: 113-123.
[26] Liedtke C, Mazouni C, Hess KR, André F, Tordai A, Mejia JA, Symmans WF, Gonzalez-Angulo AM, Hennessy B, Green M, Cristofanilli M, Hortobagyi GN and Pusztai L. Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. J Clin Oncol 2008; 26: 1275-1281.
[27] Debled M, De Mascarel I, Brouste V, Mauriac L and Macgrogan G. Re: population-based study of peritumoral lymphovascular invasion and outcome among patients with operable breast cancer. J Natl Cancer Inst 2010; 102: 275-276.
[28] Tang J, Kong D, Cui Q, Wang K, Zhang D, Gong Y and Wu G. Prognostic genes of breast cancer identified by gene co-expression network analysis. Front Oncol 2018; 8: 374.
[29] Peng R, Wang Y, Mao L, Fang F and Guan H. Identification of core genes involved in the metastasis of clear cell renal cell carcinoma. Cancer Manag Res 2020; 12: 13437-13449.
[30] Stamatakos M, Palla V, Karaiskos I, Xiromeritis K, Alexiou I, Pateras I and Kontzoglou K. Cell cyclins: triggering elements of cancer or not? World J Surg Onc 2010; 8: 111.
[31] Petri ET, Errico A, Escobedo L, Hunt T and Basavappa R. The crystal structure of human cyclin B. Cell Cycle 2007; 6: 1342-1349.
[32] Malumbres M and Barbacid M. Cell cycle, CDKs and cancer: a changing paradigm. Nat Rev Cancer 2009; 9: 153-166.
[33] Bellanger S, de Gramont A and Sobczak-Thépot J. Cyclin B2 suppresses mitotic failure and DNA re-replication in human somatic cells knocked down for both cyclins B1 and B2. Oncogene 2007; 26: 7175-7184.
[34] Wu T, Zhang X, Huang X, Yang Y and Hua X. Regulation of cyclin B2 expression and cell cycle G2/M transition by menin. J Biol Chem 2010; 285: 18291-18300.
[35] Lei CY, Wang W, Zhu YT, Fang WY and Tan WL. The decrease of cyclin B2 expression inhibits invasion and metastasis of bladder cancer. Urol Oncol 2016; 34: 237, e1-10.
[36] Bolognese F, Wasner M, Dohna CL, Gurtner A, Ronchi A, Muller H, Manni I, Mossner J, Piaggio G, Mantovani R and Engeland K. The cyclin B2 promoter depends on NF-Y, a trimer whose CCAAT-binding activity is cell-cycle regulated. Oncogene 1999; 18: 1845-1853.
[37] Tang D, Lotze MT, Kang R and Zeh HJ. Apoptosis promotes early tumorigenesis. Oncogene 2011; 30: 1851-1854.
[38] Ghelli Luserna di Rora' A, Iacobucci I and Martinelli G. The cell cycle checkpoint inhibitors in the treatment of leukemias. J Hematol Oncol 2017; 10: 77.
[39] Hunter KW, Crawford NP and Alsarraj J. Mechanisms of metastasis. Breast Cancer Res 2008; 10 Suppl 1: S2.
[40] Kozlowski J, Kozlowska A and Kocki J. Breast cancer metastasis-insight into selected molecular mechanisms of the phenomenon. Postepy Hig Med Dosw (Online) 2015; 69: 447-451.
[41] Vieira AF and Paredes J. P-cadherin and the journey to cancer metastasis. Mol Cancer 2015; 14: 178.
[42] Weigelt B, Peterse JL and van't Veer LJ. Breast cancer metastasis: markers and models. Nat Rev Cancer 2005; 5: 591.
[43] Han YH, Mun JG, Jeon HD, Kee JY and Hong SH. Betulin inhibits lung metastasis by inducing cell cycle arrest, autophagy, and apoptosis of metastatic colorectal cancer cells. Nutrients 2020; 12: 66.

Supplementary Table 1. Clinic-pathological features of Nottingham primary series of invasive breast cancer (n=3173)

Clinic-pathological features	N (%)
Age	
<50 years	1042 (32.8)
≥50 years	2131 (67.2)
Tumour Size	
<2 cm	1768 (55.8)
≥2 cm	1403 (44.2)
Grade	
Grade-I	543 (17.1)
Grade-II	1131 (35.7)
Grade-III	1496 (47.2)
Nottingham Prognostic Index (NPI)	
Good prognostic group (GPG)	1041 (32.8)
Moderate prognostic group (MPG)	1632 (51.5)
Poor prognostic group (PPG)	496 (15.7)
Stage	
Stage-I	1989 (62.8)
Stage-II	903 (28.5)
Stage-III	277 (8.7)
Lympho-vascular Invasion (LVI)	
Negative	2205 (69.8)
Positive	954 (30.2)
Estrogen Receptor Status (IHC)	
Negative	704 (22.3)
Positive	2453 (77.7)
Progesterone Receptor Status (IHC)	
Negative	1256 (40.9)
Positive	1813 (59.1)
HER2 (IHC)	
Negative	2677 (86.9)
Positive	402 (13.1)
Triple-Negative (IHC)	
No	2630 (84.5)
Yes	483 (15.5)

Supplementary Table 2. PREMARK criteria for the study

Item to be reported		Page no.
INTRODUCTION		
1	State the marker examined, the study objectives, and any pre-specified hypotheses.	3
MATERIALS AND METHODS		
	Patients	
2	Describe the characteristics (e.g., disease stage or co-morbidities) of the study patients, including their source and inclusion and exclusion criteria.	5-6
3	Describe treatments received and how chosen (e.g., randomized or rule-based).	5-6
	Specimen characteristics	
4	Describe type of biological material used (including control samples) and methods of preservation and storage.	6
	Assay methods	
5	Specify the assay method used and provide (or reference) a detailed protocol, including specific reagents or kits used, quality control procedures, reproducibility assessments, quantitation methods, and scoring and reporting protocols. Specify whether and how assays were performed blinded to the study endpoint.	5-6
	Study design	
6	State the method of case selection, including whether prospective or retrospective and whether stratification or matching (e.g., by stage of disease or age) was used. Specify the time period from which cases were taken, the end of the follow-up period, and the median follow-up time.	6
7	Precisely define all clinical endpoints examined.	6
8	List all candidate variables initially examined or considered for inclusion in models.	6
9	Give rationale for sample size; if the study was designed to detect a specified effect size, give the target power and effect size.	6
	Statistical analysis methods	
10	Specify all statistical methods, including details of any variable selection procedures and other model-building issues, how model assumptions were verified, and how missing data were handled.	6
11	Clarify how marker values were handled in the analyses; if relevant, describe methods used for cutpoint determination.	6
RESULTS		
	Data	
12	Describe the flow of patients through the study, including the number of patients included in each stage of the analysis (a diagram may be helpful) and reasons for dropout. Specifically, both overall and for each subgroup extensively examined report the numbers of patients and the number of events.	8
13	Report distributions of basic demographic characteristics (at least age and sex), standard (disease-specific) prognostic variables, and tumor marker, including numbers of missing values.	8
	Analysis and presentation	
14	Show the relation of the marker to standard prognostic variables.	8-9
15	Present univariable analyses showing the relation between the marker and outcome, with the estimated effect (e.g., hazard ratio and survival probability). Preferably provide similar analyses for all other variables being analyzed. For the effect of a tumor marker on a time-to-event outcome, a Kaplan-Meier plot is recommended.	8
16	For key multivariable analyses, report estimated effects (e.g., hazard ratio) with confidence intervals for the marker and, at least for the final model, all other variables in the model.	8
17	Among reported results, provide estimated effects with confidence intervals from an analysis in which the marker and standard prognostic variables are included, regardless of their statistical significance.	8
18	If done, report results of further investigations, such as checking assumptions, sensitivity analyses, and internal validation.	
DISCUSSION		
19	Interpret the results in the context of the pre-specified hypotheses and other relevant studies; include a discussion of limitations of the study.	9-10
20	Discuss implications for future research and clinical value.	10