Am J Cancer Res 2023;13(11):5334-5351
www.ajcr.us /ISSN:2156-6976/ajcr0152335

Original Article

The expression of high mobility group protein 3 (*HMGB3*) in breast cancer with emphasis on its role in lymphovascular invasion

Abrar I Aljohani[1,2], Sami A Alsaeed[2,3], Michael S Toss[2,4], Sara A Raafat[2,5], Andrew R Green[2], Emad A Rakha[2,6,7,8]

[1]Department of Clinical Laboratory Sciences, College of Applied Medical Sciences, Taif University, Taif 21944, Saudi Arabia; [2]Academic Unit for Translational Medical Sciences, School of Medicine, Nottingham Breast Cancer Research Centre, University of Nottingham Biodiscovery Institute, University Park, Nottingham NG7 2RD, United Kingdom; [3]Faculty of Applied Medical Sciences, Northern Border University, Arar 91431-1321, Saudi Arabia; [4]Histopathology Department, Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield S10 2JF, United Kingdom; [5]Histopathology Department, Faculty of Medicine, Mansoura Univeristy, Mansoura 35516, Egypt; [6]Histopathology Department, Faculty of Medicine, Menoufia University, Shibïn Al-Kawm 32521, Egypt; [7]Department of Histopathology, Nottingham University Hospital NHS Trust, City Hospital Campus, Hucknall Road, Nottingham NG5 1PB, United Kingdom; [8]Department of Pathology, Hamad Medical Corporation, Doha 3050, Qatar

Received July 20, 2023; Accepted September 1, 2023; Epub November 15, 2023; Published November 30, 2023

Abstract: Lymphovascular invasion (LVI) is a common phenomenon in breast cancer (BC), and it is correlated to poor outcome. However, the biomarkers that influence the development of LVI remain to be defined. Through rigorous bioinformatics analyses, *high mobility group protein 3 (HMGB3)* was revealed as a driver gene that is associated with the presence of LVI. The purpose of this study was to further investigate the role of *HMGB3* in the pathogenesis of LVI in BC. *In vitro* functional assays were performed to investigate the effect of *HMGB3* silencing on cell proliferation, migration, adherence and transmigration of BC cell lines with dermal lymphatic endothelial cells (DLECs) and human vascular endothelial cells (HUVECs). The correlation of HMGB3 expression with clinicopathological parameters was also assessed at the transcriptomic and the proteomic levels using large BC cohorts with well-characterised LVI status. Silencing *HMGB3* reduced cell proliferation, migration, adherence and transmigration across endothelial cell lines. At the mRNA and protein levels, high HMGB3 expression was significantly correlated with LVI-positivity, higher tumour grade, lymph nodal stage, hormone receptor negativity, HER2 positivity and poor outcome. Moreover, high HMGB3 expression was an independent predictor of shorter breast cancer-specific survival. *HMGB3* plays an oncogenic function and contributes to the development of LVI in BC. Results warrant further investigation as a potential target to inhibit LVI in BC.

Keywords: Breast cancer, *HMGB3*, prognosis, progression, LVI

Introduction

Patients with early-stage breast cancer (BC) have experienced better outcomes as a result of early detection, enhanced diagnostic accuracy and targeted drug therapies [1]. Despite these improvements, metastasis remains the leading cause of BC-related mortality, affecting more than 20% of patients [2]. Various histopathological characteristics, such as tumour size, lymph node status and tumour grade, are strongly correlated with mortality [3, 4]. Lymphovascular invasion (LVI) is considered an early event in the development of tumour metastasis and represents a significant predictor of poor outcome [5]. While the molecular profiles involved in tumour differentiation, such as histological type, grade, and the development of lymph node metastasis, have been well studied [6, 7], the molecular mechanisms underlying LVI, which may serve as potential predictor biomarkers or therapeutic targets, remain unknown. Targeting LVI and its associated genes is a promising approach for inhibiting tumour dissemination in early-stage BC.

High mobility group (HMG) proteins are the second most predominant proteins in the cell, and they are involved in the global assembly of chromatin domains. HMG enhances transcriptional

fine-tuning in response to abrupt environmental changes by interacting with nucleosomes, transcription factors, nucleosome remodelling complexes and histone H1 [8]. The high mobility group box (HMGB) family comprises chromosomal proteins that participate in DNA replication, transcription and repair [9, 10]. Abnormal expression of HMGB is correlated with various cancer hallmarks, including uncontrolled replicative capacity, resistance to apoptosis, tissue invasion and metastasis [8, 11]. The HMGB family includes *HMGB1*, *HMGB2*, *HMGB3* and *HMGB4* [12]. These members have 80% amino acid homology; however, their biological roles in cells are distinct [8, 11]. Several studies have revealed that *HMGB1*, *HMGB2* and *HMGB3* play a significant role in a variety of malignancies, including hepatocellular carcinoma [13], pancreatic cancer [14] and colon cancer [15].

Among the HMGB family members, *HMGB3*, is one of the highly expressed genes associated with LVI positivity in BC as determined in the weighted average difference (WAD) bioinformatics analyses [16]. In normal tissue, HMGB3 expression is high during embryogenesis, with low or no expression in normal adult tissue [17, 18]. High expression of *HMGB3* is strongly associated with the occurrence of numerous tumours and poor prognosis of advanced tumours of the lung [19], bladder [20] and in prostate cancer [21]. To enhance tumour development, *HMGB3* can control the cell cycle and stimulate the proliferation and invasion of cancer cells via the Wnt/beta-catenin, MAPK and other signalling pathways. *HMGB3* can also increase the reactive oxygen species (ROS) formation and tumour cell growth by stimulating the expression of HIF-1α [22]. However, the exact role of *HMGB3* in LVI in BC is unclear as it has not been previously investigated. This study aimed to assess the *in vitro* mechanistic role of *HMGB3* in BC cell lines, with emphasis in its role in LVI development, and to investigate the clinicopathological significance of *HMGB3* at the transcriptomic and proteomic levels using large BC cohorts with long-term follow-up.

Materials and methods

Pre-clinical studies

In vitro studies of HMGB3

Pre-clinical investigations using the following cells and assays were conducted to determine the possible involvement of *HMGB3* in LVI and other biological functions.

Breast cancer cells

The protein expression levels of HMGB3 were used to select BC cell lines (American Type Culture Collection (ATCC, Manassas, VA, USA)) for *in vitro* investigations. The BC cell lines MCF-7 (luminal oestrogen receptor (ER^+)/progesterone receptor (PR^+)/human epidermal growth factor 2 ($HER2^-$)), SK-BR-3 ($HER2^+$, ER^-/PR^-) and MDA-MB-231 (triple negative, ER^-/PR^-/$HER2^-$) were used in this study. MCF-7 and MDA-MB-231 were cultured in Roswell Park Memorial Institute (RPMI 1640) medium with L-glutamine (Cytiva, SH30027.01, UK) supplemented with 10% foetal bovine serum (FBS) (Sigma, F9665, UK), while the SK-BR-3 cell line was grown in McCoy's 5A medium modified with L-glutamine and sodium bicarbonate liquid (Sigma, M9309, UK) supplemented with 10% FBS. Western blot (WB) was performed to detect relative protein expression in all BC cell lines, and the LI-COR Odyssey machine was used for quantification.

Endothelial cells

To investigate the impact of *HMGB3* on LVI, primary human umbilical vein endothelial cells (HUVECs) and dermal lymphatic endothelial cells (DLECs) were used as *in vitro* models of tumour-endothelial interactions. HUVECs and DLECs were purchased from Promocell (C12218, Heidelberg, Germany), and cultured in the endothelial cell growth medium MV2 (Promocell, C-22022, Germany).

All cell lines were tested for *Mycoplasma* monthly, cultured under a sterile condition in a class II cabinet and incubated with 5% CO_2 at 37°C.

Silencing HMGB3 using siRNA

A siRNA-based approach was used to study the potential functional consequences of *HMGB3* knockdown, and its role in BC progression and LVI. Two independent pre-validated Silencer Select siRNA constructs, mainly for *HMGB3,* or scrambled negative control siRNA (Silencer® Select siRNA, AM4611, ThermoFisher Scientific), that did not target *HMGB3*, MCF-7, SK-BR-3 and MDA-MB-231, were transfected using the forward transfection method using Opti-

MEM medium, 25 pmol siRNA, and LipofectamineTM RNAiMAX (13778150; ThermoFisher Scientific, Loughborough, UK). The sequences of *HMGB3*-siRNA were as follows: 5'-GCACCCUGAAACUGUAUCAtt-3' and 5'-CCGAGACAAACCCUUGAUGtt-3'. Similar knockdown was observed on both siRNA targeting *HMGB3*, so siRNA with 5'-GCACCCUGAAACUGUAUCAtt-3' sequence was prioritised for the following *in vitro* studies (Supplementary Figure 1).

MTS assay

The effect of *HMGB3* knockdown on the proliferation of tumour cells was evaluated via the AQueous Non-radioactive Cell Proliferation Assay assay (Promega (G3580); CellTitre 96 Aqueous One Solution Cell Proliferation Assay) according to the manufacturer's protocol.

Colony formation assay

BC cell lines were seeded and grown in culture medium in an incubator for 14 days. Following incubation, colonies were washed with phosphate-buffered saline (PBS), fixed with methanol for 30 min, stained with crystal violet and counted using a microscope.

Wound healing assay

In this assay, the ability of tumour cells to migrate was assessed by measuring the wound repair rate of *HMGB3* knockdown and control at the following time points after transfection: T0h, T24h, and T48h. A Culture-Inserts 2 wells (Thistle Scientific Ltd., IB-81176) with a built-in gap was used according to the manufacture protocol. The wounds were observed by taking images at 10× microscopic magnification several times via light microscopy (Leica Microsystems, Lecia DMI 3000B, Germany). The wound area was measured, and the percentage of wound closure was calculated using Image J software (1.52 version).

Static adhesion assay

In a 24-well plate, endothelial cells (HUVECs and DLECs) were seeded to confluence. Tumour cell adhesion was determined after cells were labelled with 1 μM Cell Tracker Green CMFDA (Invitrogen, C2925), and incubated for 30 min at 37°C. Following labelling, the tumour cells were resuspended in medium supplemented with serum, and incubated for 35 min at 37°C with endothelial cell monolayers. Non-adherent cells were washed away with tumour cell medium, and adherent tumour cells were counted using a fluorescent microscope (Lecia DMI 3000B, Leica Microsystems, Germany) at a 10× magnification. The findings were represented as the absolute number of cells adhering to the endothelial layer as well as the percentage of cells adhering in comparison to the control.

Transmigration assay

A confluent endothelial cell monolayer was grown on hanging transwell inserts (Sigma, MCEP24H48). Tumour cell transmigration was determined following labelling with 1 μM Cell Tracker Green CMFDA (Invitrogen, C2925). To ensure the confluency and integrity of the endothelial cell barrier, lucifer yellow leakage was used (Invitrogen, L453). After 16 hours, transmigration was observed using a fluorescent microscope (Lecia DMI 3000B, Leica Microsystems, Germany) by counting cells at the bottom of the chamber.

Further clinical studies using large BC cohorts to assess the prognostic and clinicopathological significance of *HMGB3* at the transcriptomic and proteomic levels were performed.

Clinical studies

Study cohorts

In this study, three well characterised BC cohorts were used.

Molecular taxonomy of breast cancer international consortium (METABRIC) cohort: At the transcriptomic level, the METABRIC cohort was used to evaluate the expression of HMGB3. This study enrolled a total of 1980 patients with primary operable invasive BC, and information regarding the validated clinicopathological and transcriptomic data was readily available [23]. The Illumina Totalprep RNA amplification kit (Ambion, Warrington, UK) was used to generate biotin-labelled cRNA from total RNA, which was then hybridised on Illumina Human HT-12 v3 platforms to evaluate mRNA expression.

The cancer genome atlas (TCGA) cohort: mRNA expression of HMGB3 was also evaluated using the TCGA BC cohort (n=854). The cohort was

accessed for RNA-SeqV2-derived mRNA expression. De-identified clinical information for the patients was also accessed, with certain clinicopathological features and outcomes from cBioPortal [24].

Nottingham BC cohort: The expression of HMGB3 at the protein level was assessed using tissue samples from the well-characterised Nottingham invasive BC cohort. A total of 1647 cases were valid for evaluating HMGB3 protein expression from patients who had previously undergone surgery at Nottingham City Hospital. For each patient, a robust clinicopathological profile and outcome data were readily available. These profiles include the patient's age at the time of diagnosis, tumour size and grade, lymph node stage, LVI status, and the Nottingham Prognostic Index (NPI). This cohort has ER, PR, and HER2 data [25-28]. The patients' profiles included also data on BC-specific survival (BCSS) and time to distant metastasis (TTDM). BCSS is defined as the time in months, from the time when the patients underwent surgery to when they died from BC. TTDM is referred to the time in months, from when the patients underwent surgery to when the first distant metastasis occurred. Patient management was uniform and based on tumour features as determined by NPI and hormone receptor status. Patients with an NPI score ≤3.4, representing the excellent prognostic group, underwent no adjuvant therapy, while those with an NPI>3.4 who were ER positive were offered tamoxifen (with or without goserelin [Zoladex] in premenopausal patients). Patients who were ER-negative and fit enough to receive chemotherapy, received cyclophosphamide, methotrexate, and 5-flurouracil (CMF). In this study, no patients received neoadjuvant therapy or Herceptin.

The associations between HMGB3 and the available epithelial-mesenchymal transition (EMT)-related markers, such as E-cadherin, N-cadherin, P-cadherin, TGFβ1, and TWIST2 [29, 30], were investigated.

HMGB3 protein expression

For immunohistochemistry (IHC) staining of the primary rabbit polyclonal anti-HMGB3 antibody (HPA062583, Sigma-Aldrich, UK), HMGB3 protein expression was determined using the Nottingham BC cohort. Tissue microarrays (TMAs) of the study cohort were prepared using a TMA Grand Master® [31]. Antigen retrieval was performed according to the manufacturer's guidelines (citrate buffer, pH 6, at 1000 W for 20 minutes using microwave energy). The expression of the HMGB3 antibody was assessed by staining the TMAs with a Novolink polymer detection systems kit (Code: RE-7280-K, Leica, Biosystems, UK). This involved incubating 4 μm sections with HMGB3 antibody (dilution 1:500) for 60 minutes in Leica antibody diluent (RE AR9352, Leica, Biosystems, UK). As a positive control, ovarian tissue was used, whereas normal kidney tissue was used as a negative control (**Figure 1A** and **1B**). Staining for immunoreactivity was quantified using a modified histochemical score (H-score) based on semi-quantitative scoring. The scoring was done for the entire field, and the nuclear staining intensity was classified as follows: score 0= negative, score 1= weak staining, score 2= moderate staining, and score 3= strong staining. The percentage of each group was calculated (0-100%). The H-score, which ranges from 0 to 300, was calculated by multiplying the intensity of staining by the percentage of staining [32]. Two observers scored the TMAs, and the interclass correlation coefficient (ICC) test was used to assess the concordance of HMGB3.

Statistical analysis

For statistical analysis, GraphPad Prism 3.02 software and SPSS version 24 (Chicago, IL, USA) were used. *In vitro* results were represented as the mean ± standard error of the mean (SEM) of three independent experiments performed in triplicate. The significant differences between the control and silencing *HMGB3* were determined using a student's t-test.

Using continuous data on *HMGB3* mRNA and protein levels, the correlations with clinicopathological characteristics were investigated. 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 to analyse differences between three or more groups. To compare two groups, the Student t-test (parametric data) or Mann-Whitney test (non-parametric distribution) were employed. To categorise the expression of *HMGB3* mRNA, the median was used, with H-score of 50 was used for protein. The Spearman correlation test was

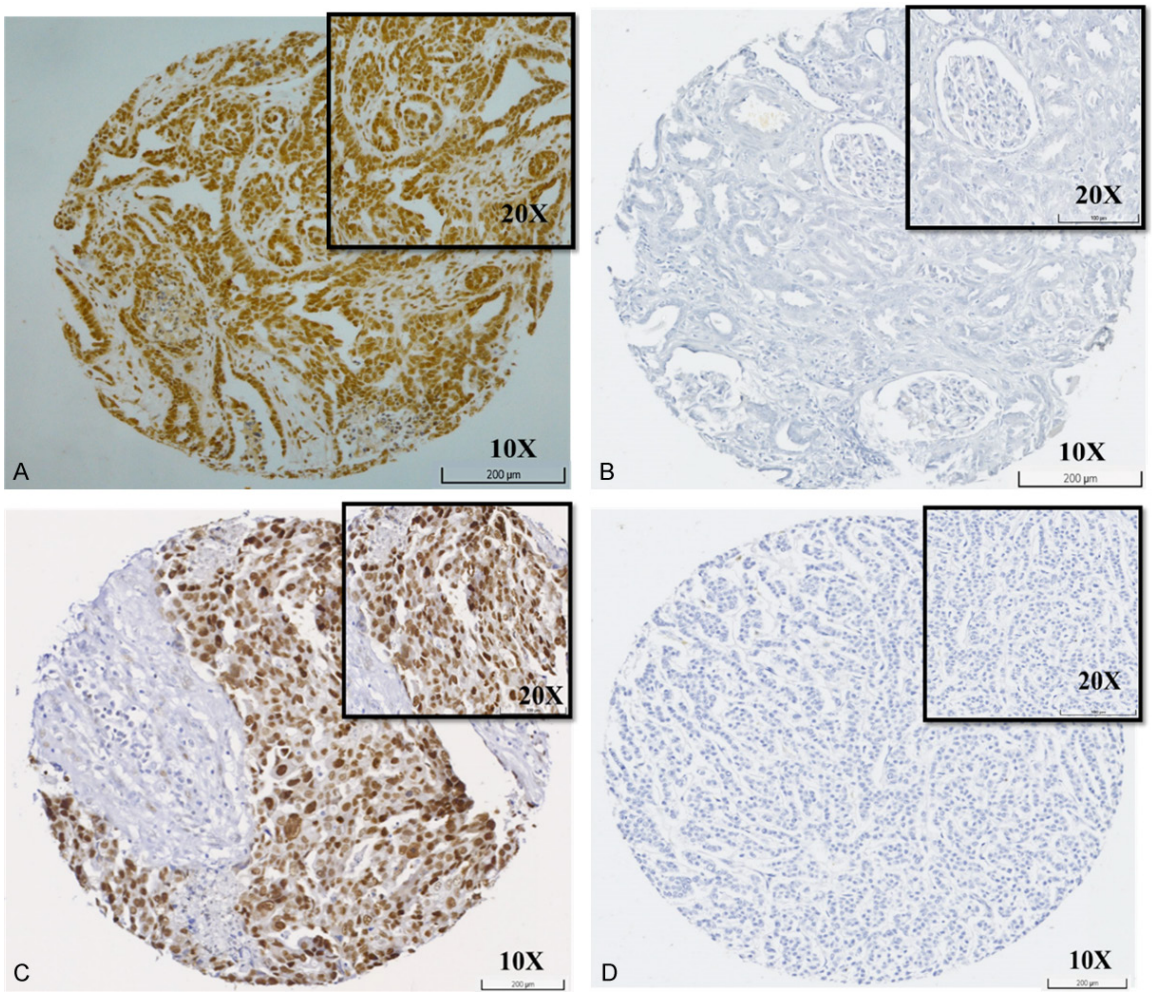


Figure 1. Nuclear expression of HMGB3 protein in invasive breast cancer. (A) Positive control of ovarian tissue stained by HMGB3, (B) Negative control of normal kidney stained by HMGB3, (C) Positive HMGB3 IHC expression and (D) Negative HMGB3 IHC expression. Magnification 10×. Scale bars =200 μm. Inset, magnification 20×. Scale bars =100 μm.

used for correlation analysis. For univariate survival analysis, the Log-rank and Kaplan-Meier curve tests were utilised, whereas for multivariate survival analysis, the Cox Regression model was used. A *p*-value <0.05 was considered statistically significant.

The guidelines for reporting recommendations for tumour marker prognostic studies (REMARK) were followed in this study [33] (Supplementary Table 1).

Results

Pre-clinical studies

Impacts of HMGB3 silencing in BC cells

The expression of the HMGB3 protein was determined in four BC cell lines, including HER2-enriched (SK-BR-3), luminal (MCF-7 and ZR-75-1), and triple-negative (MDA-MB-231) cells (Supplementary Figure 1A). The highest levels were seen in SK-BR-3, MDA-MB-231 and MCF-7, and the lowest ones in ZR-75-1 which was excluded from the subsequent experiments (Supplementary Figure 1B).

When *HMGB3* knockdown expression was compared to β-actin expression in MCF-7, SK-BR-3, and MDA-MB-231 at days 3, 5 and 7 post-transfection, complete reduction of HMGB3 protein expression was seen (Supplementary Figure 1C-N).

Down regulation of *HMGB3* resulted in a considerable decrease in BC proliferation rate, which was demonstrated in MCF-7 (P=0.0120 at T24h, P=0.0247 at T48h and P=0.0160 at

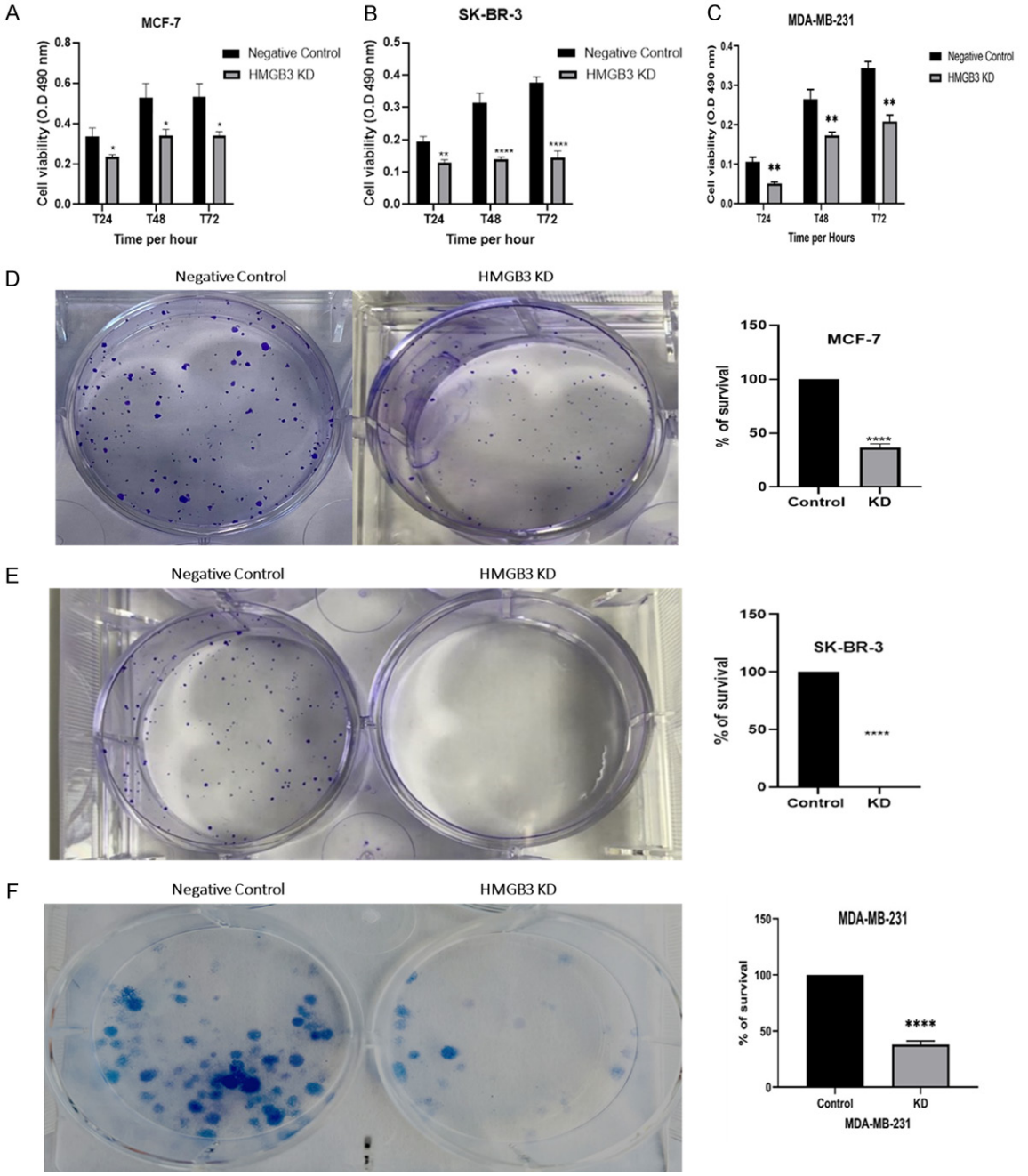


Figure 2. The effect of *HMGB3* knockdown (KD) by siRNA on cell proliferation in (A) MCF-7, (B) SK-BR-3 and (C) MD-MB-231 cells. (A-C) Cell proliferation was significantly reduced after KD in BC cell lines as detected by the MTS assay. *HMGB3* siRNA transfection reduced the ability of BC cell lines to colonise in (D) MCF-7, (E) SK-BR-3 and (F) MD-MB-231 as detected by the colony formation assay. Results are presented as mean ± standard error of the mean (SEM) of three independent experiments. The *p*-values are * ≤0.05, ** ≤0.01, *** ≤0.001 and **** ≤0.0001.

T72h), SK-BR-3 (P=0.0027 at T24h, P<0.0001 at T48h and P<0.0001 at T72h), and MDA-MB-231 (P=0.0014 at T24h, P=0.0092 at T48h and P=0.0060 at T72h), as compared to the control (**Figure 2A-C**).

In addition, a clonogenic experiments were performed to examine how *HMGB3* knockdown affected cell growth and survival. The capacity of a single cell colony to survive after being transfected was much lower than the control in

MCF-7, SK-BR-3, and MDA-MB-231 (all P< 0.0001) (**Figure 2D-F**). The knockdown of *HMGB3* resulted in a significantly larger unhealed wound in comparison to the negative controls in MCF-7, SK-BR-3, and MDA-MB-231 (P<0.0001, P=0.0040, and P=0.0002, respectively) (**Figure 3A-C**).

The role of HMGB3 in the interaction between BC cells and endothelial cells

There was a higher adherence to HUVECs in the *HMGB3*-untransfected cells than *HMGB3*-transfected cells in MCF-7 (P=0.0002), SK-BR-3 (P=0.0068), and MDA-MB-231 (P<0.0001); higher adherence to DLECs than *HMGB3*-transfected cells was observed in cell lines MCF-7 (P=0.0205), SK-BR-3 (P=0.0006), and MDA-MB-231 (P=0.0001) (**Figure 4A-C**).

Tumour cell transmigration through HUVECs was higher in the control than in the knockdown group as demonstrated in MCF-7 (P=0.0208), SK-BR-3 (P=0.0260), and MDA-MB-231 (P<0.0001); transmigration through DLECs was higher in the control than in the knockdown group in MCF-7 (P=0.0307), SK-BR-3 (P=0.0006), and MDA-MB-231 (P=0.0008) (**Figure 5A** and **5B**).

The role of HMGB3 in LVI and BC outcome in the clinical BC cohorts

HMGB3 mRNA expression: The METABRIC and TCGA cohorts were used to assess *HMGB3* mRNA expression. High *HMGB3* mRNA expression was observed in 901/1980 (46%) of the METABRIC BC cases, and in 427/854 (50%) of the TCGA cases. In both transcriptomic datasets, high expression of *HMGB3* mRNA was significantly associated with LVI positivity, and high tumour grade (all P<0.0001). In the METABRIC cohort, high expression of *HMGB3* mRNA was significantly correlated with larger tumour size (P=0.001), poor NPI (P<0.0001), and higher LN stage (P=0.048; **Table 1**).

In both METABRIC and TCGA datasets, ER^-, PR^-, and $HER2^+$ tumours all demonstrated significantly high expression of *HMGB3* mRNA (all P<0.0001) (**Table 1**). Analysis of the METABRIC cohort regarding the intrinsic (PAM50) subtypes showed that high *HMGB3* mRNA expression was associated with $HER2^+$, luminal B, basal-like, luminal A, and normal-like subtypes in descending order (all P<0.0001) (**Table 1**).

Survival analysis of *HMGB3* mRNA showed that high expression was associated with shorter BCSS in both the METABRIC and TCGA cohorts (P<0.0001 and P=0.003, respectively) (**Figure 6A** and **6B**).

HMGB3 protein expression: HMGB3 protein expression ranged from negative to strong in the nucleus of invasive BC cells (**Figure 1C** and **1D**). High HMGB3 protein expression was observed in 787/1647 (47.8%) invasive BC cases. There was a high degree of concordance between the TMAs scored by both scorers in HMGB3 immunoscoring (ICC=0.8, P<0.0001).

Similar to the transcriptomics results, high HMGB3 protein expression was significantly associated with aggressive tumour features, including LVI positivity, younger age, higher tumour grade, poor NPI, high mitotic count (all P<0.0001), high pleomorphism (P=0.005), high tubular formation (P=0.001), and higher LN stage (P=0.049) (**Table 2**). The correlation of HMGB3 protein expression with IHC subtypes was comparable to the mRNA findings, with HMGB3 protein expression being highest in $HER2^+$ tumours, followed by TN tumours. Within the $ER^+/HER2^-$ group, the high-proliferation class had considerably higher HMGB3 expression than the low-proliferation class (P<0.0001) (**Table 2**). At the protein level, there was a significant association between the high expression of HMGB3 and shorter BCSS and TTDM (all P<0.0001) (**Figure 6C** and **6D**). In all cohorts, multivariate Cox regression analysis showed that higher HMGB3 expression predicted poor BCSS independent of the tumour size and grade, LN stage, and LVI (HR 1.4; 95% CI=1.1-1.7; P=0.004 for METABRIC cohort, HR 1.7; 95% CI=1.1-2.8; P=0.028 for TCGA cohort, and HR 1.4; 95% CI=1.1-1.7; P=0.003 for Nottingham cohort) (**Table 3**).

A positive linear correlation between HMGB3 protein and mRNA expression was observed in the Nottingham subset (n=288) of the METABRIC cohort based on the Spearman's rank correlation coefficient (r=0.2, P=0.016).

HMGB3 expression and LVI related biomarkers: To further assess the role of HMGB3 in BC and its relationships with other genes involved in various LVI processes, the METABRIC and TCGA datasets were analysed for correlations between HMGB3 mRNA expression and other genes related to invasion, EMT and adhesion. A

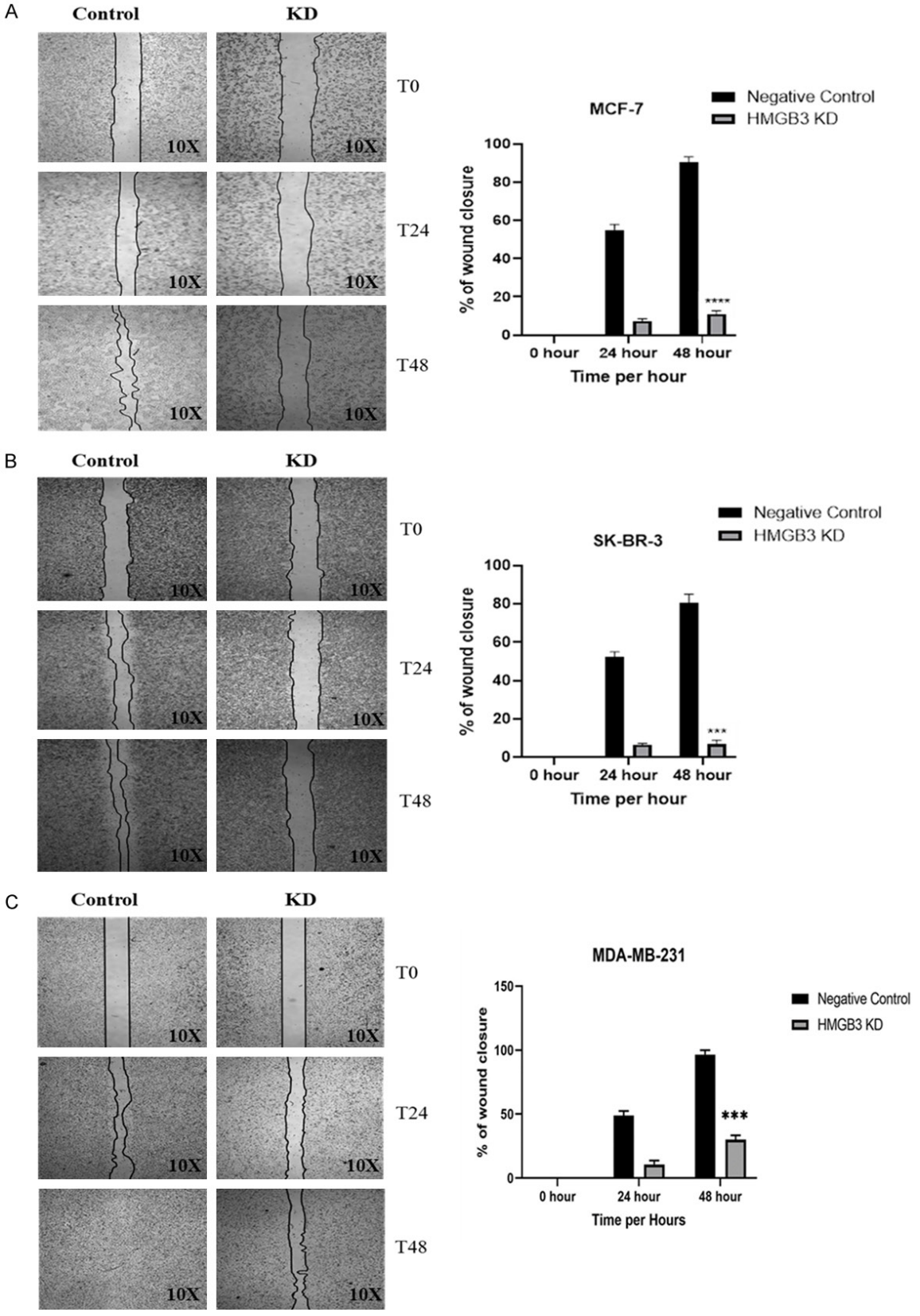


Figure 3. The effect of *HMGB3* knockdown (KD) by siRNA on cell migration in (A) MCF-7, (B) SK-BR-3 and (C) MD-MB-231 cells. The wound repair rate of *HMGB3* KD and control cells was observed by measuring the width of the gap left unhealed at T0h, T24h and T48h. (A-C) Silencing *HMGB3* significantly reduced the migration rate in BC cell lines as detected by the wound healing assay. Results are presented as mean ± standard error of the mean (SEM) of three independent experiments. The *p*-values are * ≤0.05, ** ≤0.01, *** ≤0.001 and **** ≤0.0001.

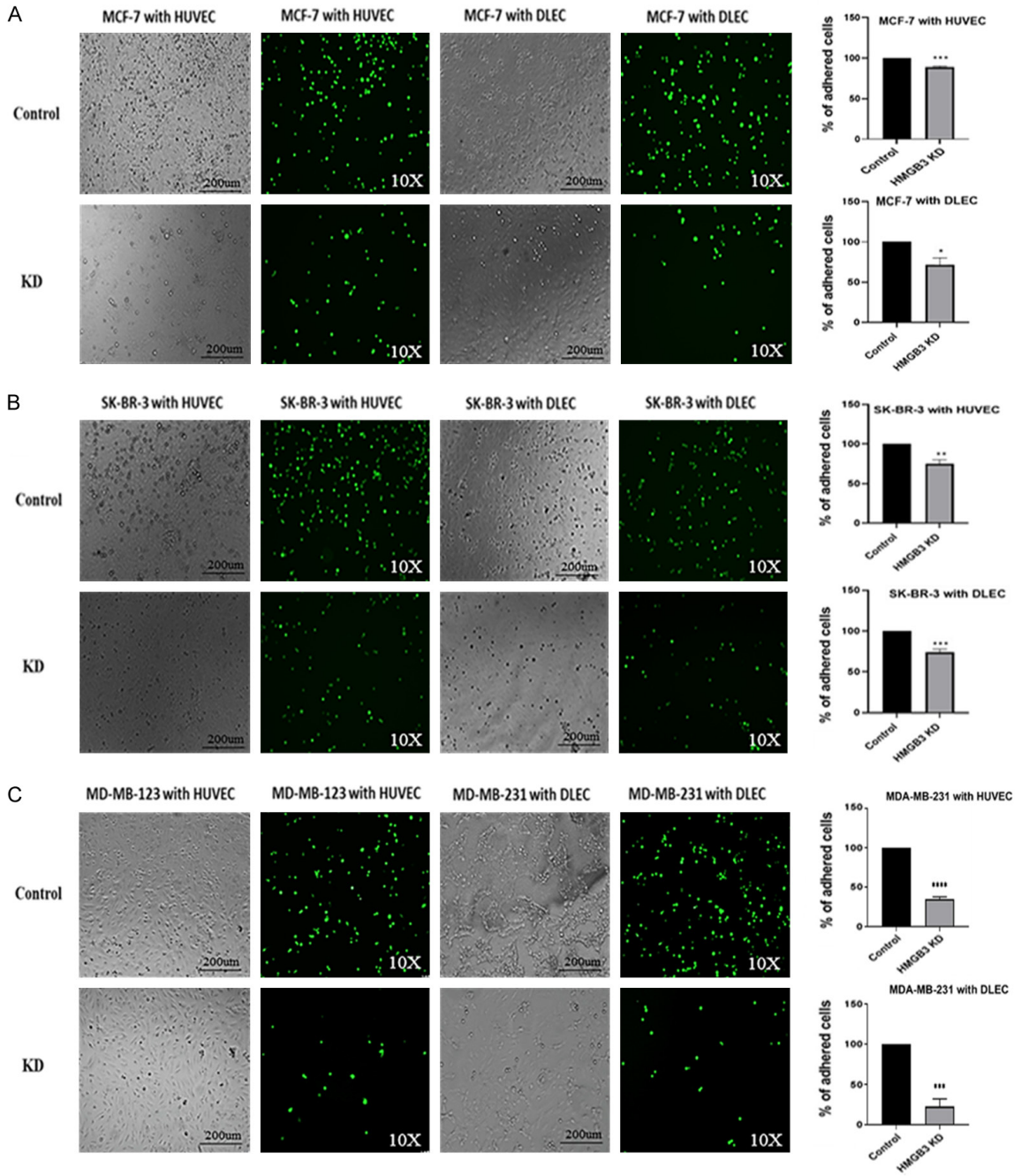


Figure 4. Representative photomicrographs of tumour cell adhesion across vascular and lymphatic endothelial cells (HUVECs and DLECs). (A) MCF-7, (B) SK-BR-3 and (C) MD-MB-231. (A-C) Silencing *HMGB3* decreased the number of all BC cells that adhered with HUVECs and DLECs. Results are presented as mean ± standard error of the mean (SEM) of three independent experiments. The *p*-values are * ≤0.05, ** ≤0.01, *** ≤0.001 and **** ≤0.0001. Magnification 10×. Scale bars =200 μm.

significant weak to moderate positive linear correlation was observed between HMGB3 and N-cadherin, P-cadherin, GSK3B, TWIST1, TWIST2, ZEB1, ZEB2, NFKB1, and CTNNB1, while the correlation was negative with E-cadherin (all P<0.05). Moreover, there was a weak to moderate positive linear correlation between HMGB3 and the expression of various MMPs, including MMP1, MMP7, MMP9, MMP12, MMP15, and MMP25 (all P<0.05) (**Table 4**).

The correlation between HMGB3 and EMT-related markers, such as E-cadherin, N-cadherin, and P-cadherin was further studied at the pro-

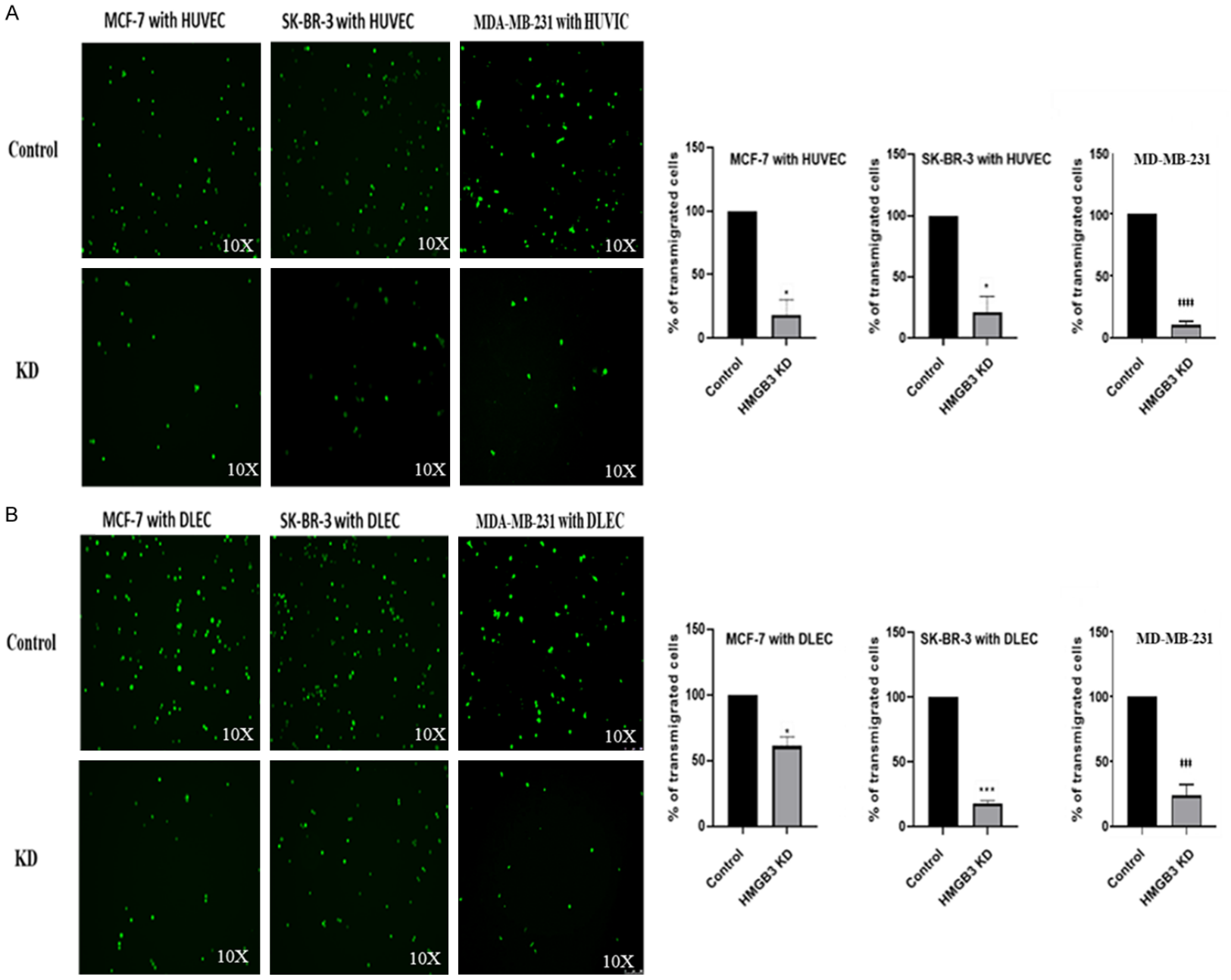

A
MCF-7 with HUVEC
SK-BR-3 with HUVEC
MDA-MB-231 with HUVIC
Control
KD
10X
% of transmigrated cells
MCF-7 with HUVEC
SK-BR-3 with HUVEC
MD-MB-231
Control
HMGB3 KD
B
MCF-7 with DLEC
SK-BR-3 with DLEC
MDA-MB-231 with DLEC
MCF-7 with DLEC
SK-BR-3 with DLEC
MD-MB-231

Figure 5. Representative photomicrographs of tumour cell (MCF-7, SK-BR-3 and MD-MB-231) transmigration across (A) HUVECs and (B) DLECs. (A, B) The number of tumour cells that transmigrated across HUVECs and DLECs was higher in the control group than in the KD group. Results are presented as mean ± standard error of the mean (SEM) of three independent experiments. The *p*-values are * ≤0.05, ** ≤0.01, *** ≤0.001 and **** ≤0.0001.

Table 1. Statistical associations between *HMGB3* mRNA expression and clinicopathological parameters in the METABRIC (n=1980) and TCGA (n=854) breast carcinoma datasets

Parameters	*HMGB3* mRNA (METABRIC)			*HMGB3* mRNA (TCGA)		
	Number (%)	Mean Rank	*P* value	Number (%)	Mean Rank	*P* value
Patient Age (year)						
≤50	424 (21.4)	7.58	0.695	231 (27)	435.0	0.587
>50	1556 (78.6)	7.56		623 (73)	424.1	
Lymphovascular Invasion (LVI)						
Negative	930 (59)	7.46	**<0.0001**	559 (65)	403.1	**<0.0001**
Positive	635 (41)	7.65		295 (35)	437.7	
Tumour Size						
≤2 cm	622 (31.7)	7.42	**0.001**	239 (28)	411.8	0.245
>2 cm	1338 (68.3)	7.62		615 (72)	433.6	
Tumour Grade						
I	170 (9.0)	7.09	**<0.0001**	89 (11)	267.2	**<0.0001**
II	770 (40.6)	7.43		375 (46)	338.9	
III	952 (50.3)	7.77		352 (43)	518.4	
Lymph node stage				Not available		
I (Negative nodes)	1035 (52.5)	7.51	**0.048**			
II (1-3 positive nodes)	622 (31.5)	7.63				
III (>3 positive nodes)	316 (16.0)	7.62				
Nottingham prognostic index (NPI) groups				Not available		
Good	680 (34.3)	7.35	**<0.0001**			
Moderate	1101 (55.6)	7.67				
Poor	199 (10.1)	7.78				
Oestrogen Receptor (ER)						
Negative	474 (23.9)	7.82	**<0.0001**	185 (22)	580.9	**<0.0001**
Positive	1506 (76.1)	7.49		639 (78)	363.7	
Progesterone Receptor (PR)						
Negative	940 (47.4)	7.65	**<0.0001**	272 (33)	510.4	**<0.0001**
Positive	1040 (52.6)	7.49		546 (67)	359.2	
Human epidermal growth factor receptor 2 (HER2)						
Negative	1733 (87.5)	7.48	**<0.0001**	567 (81)	339.9	**0.004**
Positive	247 (12.5)	8.18		133 (19)	395.6	
PAM50 Subtypes				Not available		
Luminal A	718 (36.4)	7.31	**<0.0001**			
Luminal B	488 (24.7)	7.81				
$HER2^+$ enriched	240 (12.1)	8.15				
Basal like	329 (16.7)	7.65				
Normal like	199 (10.1)	7.04				

P values in bold are statistically significant.

tein level using the Nottingham BC cohort. There was a significant weak negative correlation between HMGB3 and E-cadherin (P=0.001), while the correlation was positive with N-cadherin (P=0.023), and P-cadherin (P=0.002) (**Table 4**).

Disscussion

HMGB3 is a multifunctional protein that performs a variety of functions in many cellular compartments, and has been identified as a critical regulator of tumour development [34].

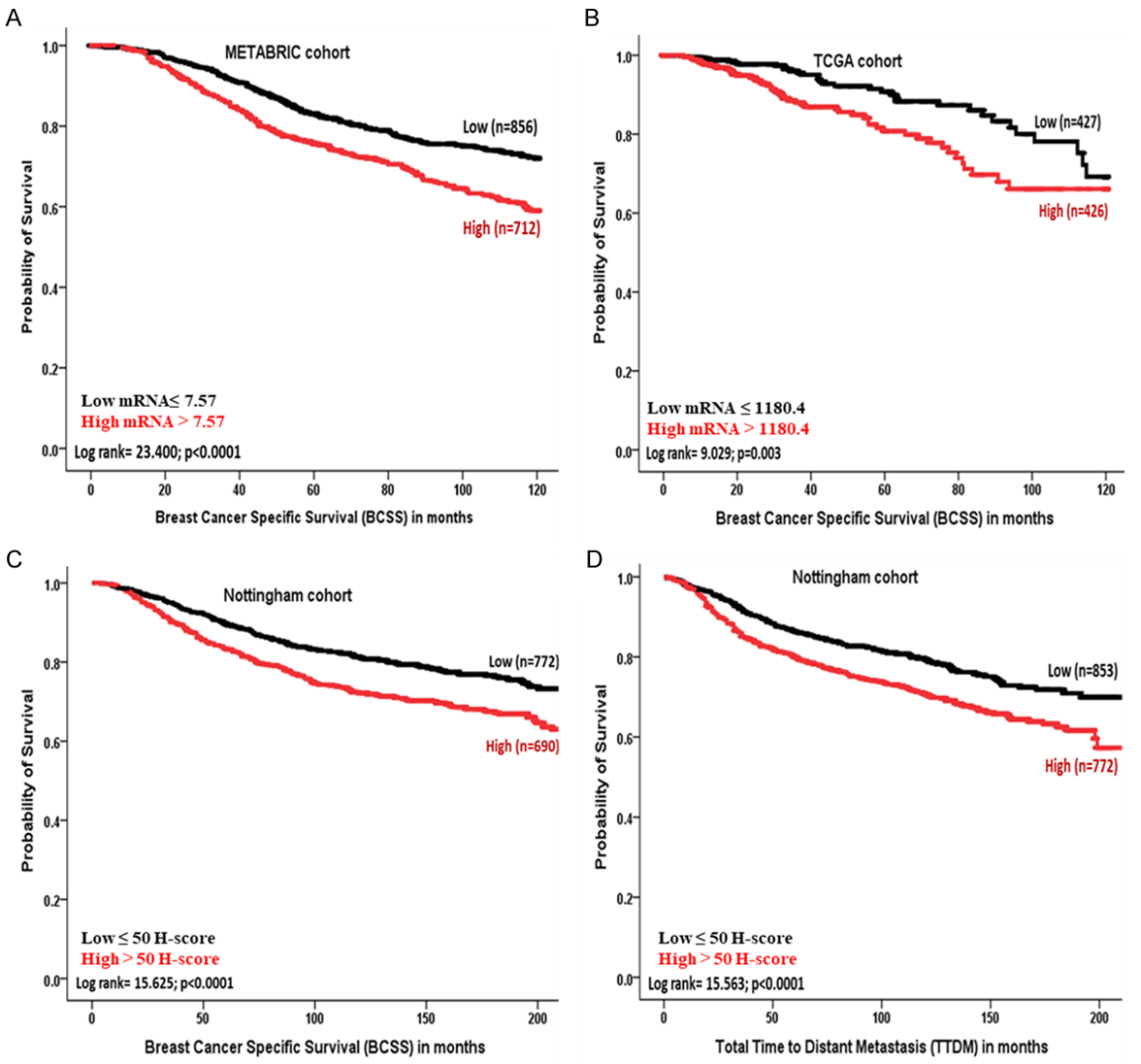


Figure 6. Kaplan-Meier survival plots showing the association between *HMGB3* mRNA expression and breast cancer-specific survival (BCSS) in (A) whole cohort (METABRIC) and (B) whole cohort (TCGA). Associations are also shown in the Nottingham cohort between HMGB3 protein expression and (C) BCSS and (D) total time to distant metastasis (TTDM).

Recently, aberrant *HMGB3* was identified as a pro-carcinogen, promoting tumour development, proliferation, invasion, and metastasis in a variety of tumour types, including gastric [35], lung [36], colorectal [12], and urinary bladder tumours [20]. Although *HMGB3* has previously been associated with cancer cell proliferation and migration [12], no studies have been performed to date to investigate the potential role of HMGB3 in the development of LVI.

HMGB3 was determined as a key gene associated with LVI positivity through stringent bioinformatics analysis [16]. Similarly, subsequent research employing bioinformatic and co-expression analyses revealed a link between *HMGB3* and the development of human cancers, including gastric cancer [8]. This study aimed to evaluate the *in vitro* mechanistic role of *HMGB3* in BC cell lines with emphasis on the role in LVI development, and to investigate the clinicopathological significance of *HMGB3* at the transcriptomic and proteomic levels using large BC cohorts with long-term clinical follow-up.

The conducted pre-clinical experiments showed that silencing *HMGB3* suppressed cell proliferation and growth. Tumour cells must proliferate and evade apoptosis to penetrate surrounding

Table 2. Statistical associations between HMGB3 protein expression and the clinicopathological factors in the Nottingham breast cancer (BC) cohort (n=1647)

Parameters	HMGB3 protein		
	Number %	Mean Rank	*P* value
Patient Age (year)			
≤50	574 (35)	882.3	**<0.0001**
>50	1070 (65)	790.4	
Lymphovascular Invasion (LVI)			**<0.0001**
Negative	1096 (67)	790.1	
Positive	541 (33)	877.6	
Tumour Size			0.668
≤2 cm	914 (56)	815.5	
>2 cm	725 (44)	825.6	
Tumour Grade			**<0.0001**
I	206 (13)	697.9	
II	584 (52)	779.4	
III	854 (35)	882.0	
Mitosis Scores			**<0.0001**
I	622 (38.3)	715.66	
II	312 (19.2)	838.60	
III	686 (42.5)	883.71	
Pleomorphism Scores			**0.005**
I	21 (1.3)	674.48	
II	465 (28.7)	759.01	
III	1135 (70)	834.83	
Tubular formation			**0.001**
I	83 (5.1)	619.11	
II	472 (29.1)	814.69	
III	1066 (65.8)	824.31	
Lymph Node Stage			**0.049**
I (Negative nodes)	975 (60)	799.5	
II (1-3 positive nodes)	500 (31)	843.9	
III (>3 positive nodes)	167 (9)	882.7	
Nottingham Prognostic Index (NPI) groups			**<0.0001**
Good	462 (28)	733.2	
Moderate	880 (54)	853.1	
Poor	295 (18)	851.8	
Oestrogen Receptor (ER)			**0.001**
Negative	386 (23)	895.1	
Positive	1262 (77)	802.9	
Progesterone Receptor (PR)			0.108
Negative	673 (41)	834.7	
Positive	951 (59)	796.9	
HER2 status			**<0.0001**
Negative	1402 (86)	84.5	
Positive	225 (14)	997.9	
Triple Negative phenotype			0.185
No	1359 (83)	811.1	
Yes	276 (17)	852.2	

Immunohistochemistry Subtypes			**<0.0001**
$ER^+/HER2^-$ Low Proliferation	528 (34)	655.6	
$ER^+/HER2^-$ High Proliferation	522 (33.7)	802.2	
Triple Negative	275 (17.7)	809.2	
$HER2^+$	225 (14.6)	953.5	

P values in bold are statistically significant.

Table 3. Multivariate Cox regression for predictors of breast cancer-specific survival (BCSS) and *HMGB3* mRNA expression in the METABRIC and TCGA cohorts and protein expression in the Nottingham BC cohort

METABRIC Cohort				
Parameters	Hazard ratio (HR)	95% confidence interval (CI)		Significance *P* value
		Lower	Upper	
HMGB3 mRNA Expression	1.391	1.113	1.739	**0.004**
Tumour Size	1.474	1.126	1.928	**0.005**
Lymph Nodal Stage	1.969	1.538	2.520	**<0.0001**
Tumour Grade	1.456	1.192	1.779	**<0.0001**
Lymphovascular Invasion (LVI)	1.469	1.159	1.862	**0.001**
TCGA Cohort				
Parameters	Hazard ratio (HR)	95% confidence interval (CI)		Significance *P* value
		Lower	Upper	
HMGB3 mRNA Expression	1.728	1.061	2.816	**0.028**
Tumour Size	1.523	0.873	2.656	0.138
Lymph Nodal	1.196	0.712	2.010	0.498
Tumour Grade	1.092	0.761	1.567	0.634
Lymphovascular Invasion (LVI)	1.878	1.149	3.070	**0.012**
Nottingham BC Cohort				
Parameters	Hazard ratio (HR)	95% confidence interval (CI)		Significance *P* value
		Lower	Upper	
HMGB3 Protein Expression	1.350	1.104	1.650	**0.003**
Tumour Size	1.383	1.116	1.714	**0.003**
Lymph Nodal Stage	1.708	1.478	1.973	**<0.0001**
Tumour Grade	1.649	1.375	1.977	**<0.0001**
Lymphovascular Invasion (LVI)	1.620	1.307	2.009	**<0.0001**

P values in bold are statistically significant.

tissue and develop metastatic cascades; proliferation continues until tumour cells invade the vascular or lymphatic channels. *In vitro* functional assays showed that silencing *HMGB3* in BC cell lines reduced migration. This finding supports a study that evaluated the expression of β-catenin, a major WNT pathway protein, to determine that *HMGB3* stimulated colorectal cancer migration via the WNT/beta-catenin pathway. *HMGB3* can increase the expression of β-catenin, and the downstream genes *c-Myc* and *MMP7* [12].

The malignant features and mechanism of action of *HMGB3* in LVI remain unknown. Importantly, in this study, *HMGB3* knockdown reduced adhesion and transmigration across endothelial cell lines in MCF-7, SK-BR-3, and MDA-MB-231; the positive associations between EMT- and MMP-related markers support these results. Although the association between HMGB3 and these biomarkers ranged from weak to moderate correlation, it was statistically significant which indicates that these markers are contributing to the same oncogenic pathway in the context of LVI process. N-cadherin is associated with EMT, which is required for cell invasion and intravasation into the bloodstream and for extracellular matrix (ECM) destruction induced by protease synthe-

Table 4. Correlations of HMGB3 expression with mRNA and protein expression of epithelial-mesenchymal transition (EMT) and matrix metalloproteinase- (MMP-) 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	-0.252	**<0.0001**	-0.197	**<0.0001**	-0.128	**0.001**
N-cadherin	0.118	**<0.0001**	0.169	**<0.0001**	0.099	**0.023**
P-cadherin	0.064	**0.005**	0.196	**<0.0001**	0.125	**0.002**
TGFβ1	0.001	0.947	0.216	**<0.0001**	0.052	0.827
TWIST1	0.149	**<0.0001**	0.149	**<0.0001**	Not available	
TWIST2	0.282	**<0.0001**	0.2173	**<0.0001**	0.062	0.186
ZEB1	0.202	**<0.0001**	0.347	**<0.0001**	Not available	
ZEB2	0.231	**<0.0001**	0.197	**<0.0001**		
NFKB1	0.114	**<0.0001**	0.150	**<0.0001**		
GSK3B	0.286	**<0.0001**	0.121	**<0.0001**		
CTNNB1	0.138	**<0.0001**	0.074	**0.032**		
MMPs related genes						
MMP1	0.188	**<0.0001**	0.382	**<0.0001**	Not available	
MMP7	0.056	**0.013**	0.104	**0.002**		
MMP9	0.133	**<0.0001**	0.184	**<0.0001**		
MMP11	0.138	**<0.0001**	0.016	0.640		
MMP12	0.137	**<0.0001**	0.273	**<0.0001**		
MMP15	0.315	**<0.0001**	0.214	**<0.0001**		
MMP20	0.023	0.299	0.031	0.367		
MMP25	0.051	**0.023**	0.069	**0.043**		

P values in bold are statistically significant.

sis. Increased N-cadherin expression thus results in the loss of the connection between the epithelium of BC cells and other epithelial cells, leading to invasion into the stroma [37]. Additionally, P-cadherin, a critical protein that may activate integrin molecules, enables cancer cells to adhere to the ECM and initiates invasion [38]. The microenvironment of tumours exhibiting LVI is closely associated with the expression of MMPs, notably MMP9 and MMP1. The expressions of these MMPs are involved in the intravasation and metastasis of BC cells [39]. This finding also corroborates a previous study reporting that, by limiting the activity of MMP9, downregulation of *HMGB3* expression decreased the invasion of gastric cancer cells [22]. Although the *in vitro* studies revealed that overexpression of *HMGB3* aided migration, adhesion and transmigration via endothelial cell lines, triggering the LVI process in the MCF-7, SK-BR-3, and MDA-MB-231 cell lines, more mechanistic investigations into how *HMGB3* triggers LVI and *in vivo* animal experiments are warranted.

This study evaluated the clinicopathological and prognostic significance of *HMGB3* expressions using large, and swell-annotated BC cohorts. High expression of *HMGB3* at the mRNA and protein levels was associated with aggressive BC features, including LVI positivity, higher tumour grade, poor NPI, ER/PR negativity, and HER2 positivity. High *HMGB3* expression at the mRNA and protein levels was significantly associated with poor BCSS and TTDM, and this association was independent of other prognostic factors, which is consistent with previous studies [15, 36]. A weak correlation between the protein and mRNA expressions was observed which could be explained by the rate at which mRNA is translated into protein, which is often referred to as 'translational efficiency'.

Translational efficiency has a significant influence on both mRNA and protein levels [40]. The subjectivity of the H-score method in the interpretation of expression in IHC staining sections is another potential reason [41], in addition to the use of whole tissue, including different cell types, in the METABRIC cases.

In conclusion, our findings suggest that *HMGB3* has an oncogenic role in BC, and that it is involved in the pathogenesis of LVI. This study demonstrated the impact of *HMGB3* silencing on several processes of tumour development related to LVI, including migration, adhesion and transmigration. Although the findings of this study suggest that *HMGB3* is a critical gene in LVI, especially in light of its impact on lymphatic invasion, and is a precursor for the metastatic cascade, more research is required to further substantiate these findings.

Acknowledgements

Abrar I 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 £50m 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 NG5 1PB, United Kingdom. Tel: +44-0115-9691169; Fax: +44-0115-9627768; E-mail: emad.rakha@nottingham.ac.uk; Emad.rakha@nuh.nhs.uk

References

[1] Marshall DC, Webb TE, Hall RA, Salciccioli JD, Ali R and Maruthappu M. Trends in UK regional cancer mortality 1991-2007. Br J Cancer 2016; 114: 340-347.

[2] 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.

[3] Wo JY, Chen K, Neville BA, Lin NU and Punglia RS. Effect of very small tumor size on cancer-specific mortality in node-positive breast cancer. J Clin Oncol 2011; 29: 2619-2627.

[4] Rakha EA, Reis-Filho JS, Baehner F, Dabbs DJ, Decker T, Eusebi V, Fox SB, Ichihara S, Jacquemier J, Lakhani SR, Palacios J, Richardson AL, Schnitt SJ, Schmitt FC, Tan PH, Tse GM, Badve S and Ellis IO. Breast cancer prognostic classification in the molecular era: the role of histological grade. Breast Cancer Res 2010; 12: 207.

[5] 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.

[6] Sotiriou C, Wirapati P, Loi S, Harris A, Fox S, Smeds J, Nordgren H, Farmer P, Praz V, Haibe-Kains B, Desmedt C, Larsimont D, Cardoso F, Peterse H, Nuyten D, Buyse M, Van de Vijver MJ, Bergh J, Piccart M and Delorenzi M. Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. J Natl Cancer Inst 2006; 98: 262-272.

[7] Mobadersany P, Yousefi S, Amgad M, Gutman DA, Barnholtz-Sloan JS, Velázquez Vega JE, Brat DJ and Cooper LAD. Predicting cancer outcomes from histology and genomics using convolutional networks. Proc Natl Acad Sci U S A 2018; 115: E2970-E2979.

[8] Fang J, Ge X, Xu W, Xie J, Qin Z, Shi L, Yin W, Bian M and Wang H. Bioinformatics analysis of the prognosis and biological significance of HMGB1, HMGB2, and HMGB3 in gastric cancer. J Cell Physiol 2020; 235: 3438-3446.

[9] Agresti A and Bianchi ME. HMGB proteins and gene expression. Curr Opin Genet Dev 2003; 13: 170-178.

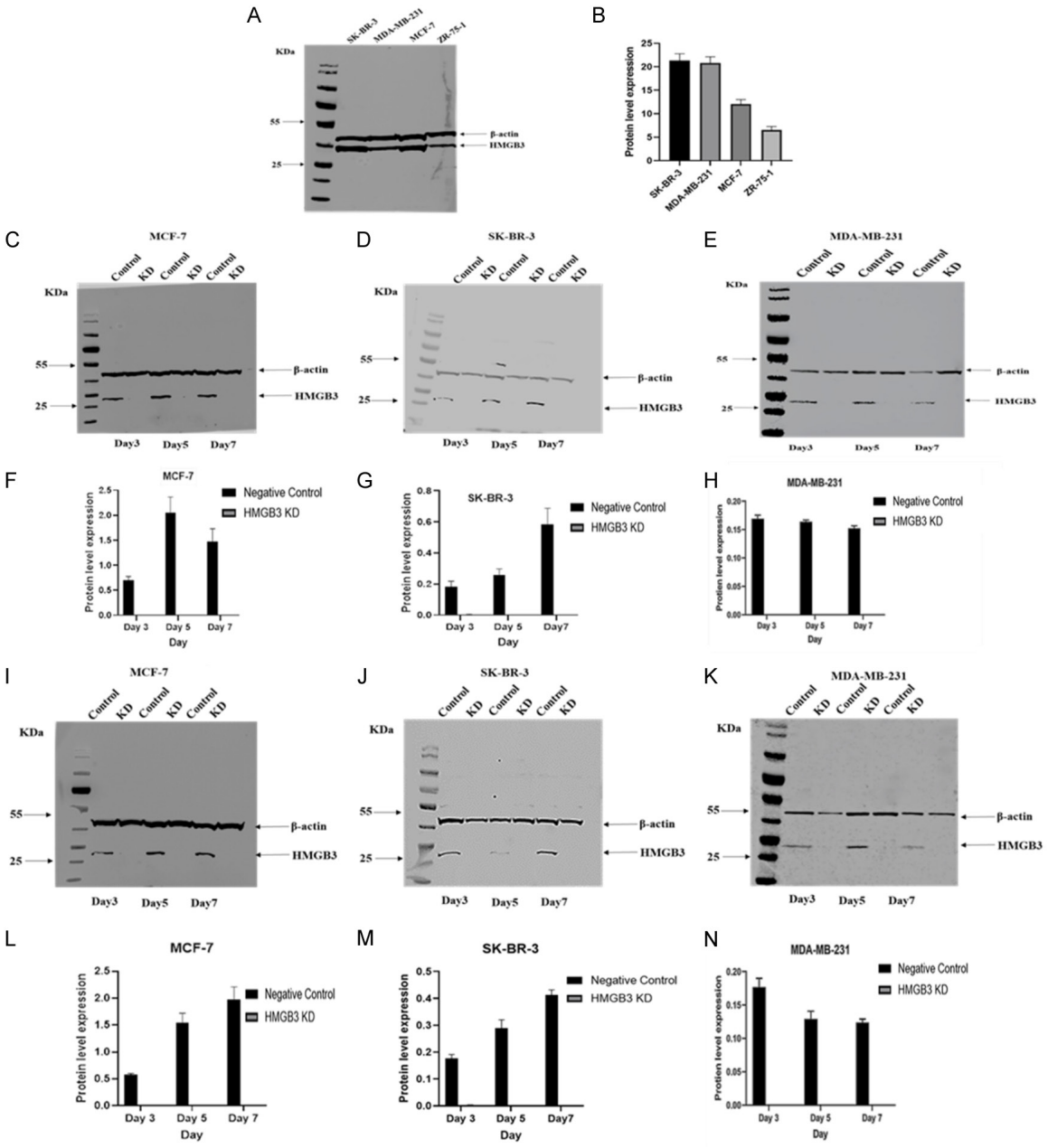
[10] Štros M. HMGB proteins: interactions with DNA and chromatin. Biochim Biophys Acta 2010; 1799: 101-113.

[11] Guo ZS, Liu Z, Bartlett DL, Tang D and Lotze MT. Life after death: targeting high mobility group box 1 in emergent cancer therapies. Am J Cancer Res 2013; 3: 1-20.

[12] Zhang Z, Chang Y, Zhang J, Lu Y, Zheng L, Hu Y, Zhang F, Li X, Zhang W and Li X. HMGB3 promotes growth and migration in colorectal cancer by regulating WNT/β-catenin pathway. PLoS One 2017; 12: e0179741.

[13] Zheng W, Yang J, Dong Z, Wang L, Fang M, Wu W, Yao D and Yao M. High mobility group box 3 as an emerging biomarker in diagnosis and prognosis of hepatocellular carcinoma. Cancer Manag Res 2018; 10: 5979-5989.
[14] Takeda T, Izumi H, Kitada S, Uramoto H, Tasaki T, Zhi L, Guo X, Kawatsu Y, Kimura T, Horie S, Nabeshima A, Noguchi H, Wang KY, Sasaguri Y, Kohno K and Yamada S. The combination of a nuclear HMGB1-positive and HMGB2-negative expression is potentially associated with a shortened survival in patients with pancreatic ductal adenocarcinoma. Tumour Biol 2014; 35: 10555-10569.
[15] Ueda M, Takahashi Y, Shinden Y, Sakimura S, Hirata H, Uchi R, Takano Y, Kurashige J, Iguchi T, Eguchi H, Sugimachi K, Yamamoto H, Doki Y, Mori M and Mimori K. Prognostic significance of high mobility group box 1 (HMGB1) expression in patients with colorectal cancer. Anticancer Res 2014; 34: 5357-5362.
[16] Kurozumi S, Joseph C, Sonbul S, Alsaeed S, Kariri Y, Aljohani A, Raafat S, Alsaleem M, Ogden A, Johnston SJ, Aleskandarany MA, Fujii T, Shirabe K, Caldas C, Ashankyty I, Dalton L, Ellis IO, Desmedt C, Green AR, Mongan NP and Rakha EA. A key genomic subtype associated with lymphovascular invasion in invasive breast cancer. Br J Cancer 2019; 120: 1129-1136.
[17] Guo S, Wang Y, Gao Y, Zhang Y, Chen M, Xu M, Hu L, Jing Y, Jing F, Li C, Wang Q and Zhu Z. Knockdown of high mobility group-box 3 (HMGB3) expression inhibits proliferation, reduces migration, and affects chemosensitivity in gastric cancer cells. Med Sci Monit 2016; 22: 3951-3960.
[18] Wang K, Yin YH, Yang ZQ, Yu HF, Wang YS, Guo B and Yue ZP. Hmgb3 induces the differentiation of uterine stromal cells through targeting Ptn. Reprod Sci 2019; 26: 891-899.
[19] Xie Y, Wang L and Yang D. CircEPSTI1 promotes the progression of non-small cell lung cancer through miR-145/HMGB3 axis. Cancer Manag Res 2020; 12: 6827-6836.
[20] Li M, Cai Y, Zhao H, Xu Z, Sun Q, Luo M, Gu L, Meng M, Han X and Sun H. Overexpression of HMGB3 protein promotes cell proliferation, migration and is associated with poor prognosis in urinary bladder cancer patients. Tumour Biol 2015; 36: 4785-4792.
[21] Yamada Y, Nishikawa R, Kato M, Okato A, Arai T, Kojima S, Yamazaki K, Naya Y, Ichikawa T and Seki N. Regulation of HMGB3 by antitumor miR-205-5p inhibits cancer cell aggressiveness and is involved in prostate cancer pathogenesis. J Hum Genet 2018; 63: 195-205.
[22] Wen B, Wei YT and Zhao K. The role of high mobility group protein B3 (HMGB3) in tumor proliferation and drug resistance. Mol Cell Biochem 2021; 476: 1729-1739.
[23] 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: 346-352.
[24] 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.
[25] 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.
[26] Mahmoud SM, Paish EC, Powe DG, Macmillan RD, Lee AH, 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.
[27] 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.
[28] 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.
[29] Abd El-Rehim DM, Ball G, Pinder SE, Rakha E, Paish C, Robertson JF, 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.
[30] 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.
[31] 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 cere-

visiae-like 1 (SEC14L1) is a prognostic factor in breast cancer associated with lymphovascular invasion. Mod Pathol 2018; 31: 1675-1682.
[32] 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.
[33] 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 MARKer prognostic studies (REMARK). Br J Cancer 2005; 93: 387-391.
[34] Zheng WJ, Yao M, Fang M, Wang L, Dong ZZ and Yao DF. Abnormal expression of HMGB-3 is significantly associated with malignant transformation of hepatocytes. World J Gastroenterol 2018; 24: 3650-3662.
[35] Gong Y, Cao Y, Song L, Zhou J, Wang C and Wu B. HMGB3 characterization in gastric cancer. Genet Mol Res 2013; 12: 6032-6039.
[36] Song N, Liu B, Wu JL, Zhang RF, Duan L, He WS and Zhang CM. Prognostic value of HMGB3 expression in patients with non-small cell lung cancer. Tumour Biol 2013; 34: 2599-2603.
[37] Mrozik KM, Blaschuk OW, Cheong CM, Zannettino ACW and Vandyke K. N-cadherin in cancer metastasis, its emerging role in haematological malignancies and potential as a therapeutic target in cancer. BMC Cancer 2018; 18: 939.
[38] Ribeiro AS and Paredes J. P-cadherin linking breast cancer stem cells and invasion: a promising marker to identify an "intermediate/metastable" EMT state. Front Oncol 2015; 4: 371.
[39] Aleskandarany MA, Sonbul SN, Mukherjee A and Rakha EA. Molecular mechanisms underlying lymphovascular invasion in invasive breast cancer. Pathobiology 2015; 82: 113-123.
[40] Maier T, Güell M and Serrano L. Correlation of mRNA and protein in complex biological samples. FEBS Lett 2009; 583: 3966-3973.
[41] Rizzardi AE, Johnson AT, Vogel RI, Pambuccian SE, Henriksen J, Skubitz AP, Metzger GJ and Schmechel SC. Quantitative comparison of immunohistochemical staining measured by digital image analysis versus pathologist visual scoring. Diagn Pathol 2012; 7: 42.



Supplementary Figure 1. HMGB3 protein expression in breast cancer cell lines. (A) Evaluation of the differential expression of HMGB3 in BC cell lines by western blotting and (B) quantification of HMGB3 protein expression level in SK-BR-3, MDA-MB-231, MCF-7 and ZR-75-1. Using pre-validated Silencer Select siRNA constructs, mainly for *HMGB3* (5'-GCACCCUGAAACUGUAUCAtt-3'), or scrambled negative control siRNA, silencing *HMGB3* using the forward transfection method relative to a non-targeting scrambled control siRNA showed complete knockdown in (C) MCF-7, (D) SK-BR-3 and (E) MDA-MB-231. Quantification of HMGB3 protein expression level in the transfected cells and negative control by densitometry and normalization to β-actin levels revealed complete loss of HMGB3 protein expression in (F) MCF-7, (G) SK-BR-3 and (H) MDA-MB-231. Similar knockdown results was observed on siRNA targeting *HMGB3* with sequence 5'-CCGAGACAAACCCUUGAUGtt-3' (I-N). Each bar represents the mean ± standard error of the mean (SEM) of three or more independent experiments. β-actin was included as a loading control.

Supplementary Table 1. 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.	3-4
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.	7
3	Describe treatments received and how chosen (e.g., randomized or rule-based).	7
Specimen characteristics		
4	Describe type of biological material used (including control samples) and methods of preservation and storage.	4-7
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.	4-8
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.	4-8
7	Precisely define all clinical endpoints examined.	4-8
8	List all candidate variables initially examined or considered for inclusion in models.	4-8
9	Give rationale for sample size; if the study was designed to detect a specified effect size, give the target power and effect size.	4-8
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.	8
11	Clarify how marker values were handled in the analyses; if relevant, describe methods used for cutpoint determination.	8
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.	10-11
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.	10-11
Analysis and presentation		
14	Show the relation of the marker to standard prognostic variables.	11
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.	11
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.	11
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.	11
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.	11-13
20	Discuss implications for future research and clinical value.	13