The nucleolar related protein Dyskerin pseudouridine synthase 1 (DKC1) predicts poor prognosis in breast cancer

Khloud A. Elsharawy[1,2], Omar J. Mohammed[1], Mohammed A. Aleskandarany[1], Ayman Hyder[2], Hekmat L. El-Gammal[2], Mohamed I. Abou-Dobara[2], Andrew R. Green[1], Leslie W. Dalton[3] and Emad A. Rakha[1]

[1] Nottingham Breast Cancer Research Centre, Division of Cancer and Stem Cells, School of Medicine, University of Nottingham Biodiscovery Institute, University Park, Nottingham UK.

[2] Faculty of Science, Damietta University, Damietta, Egypt.

[3] Department of Histopathology, South Austin Hospital, Austin, TX USA.

Correspondence:

Professor Emad Rakha

Division of Cancer and Stem Cell,
University of Nottingham,
City Hospital Campus, Hucknall Road,
Nottingham, NG5 1PB, UK
Tel: (44) 0115-9691169
Fax: (44) 0115- 9627768
Email: Emad.rakha@nuh.nhs.uk
emadrakha@yahoo.com

Running Title: Prognostic significance of DKC1 in breast cancer

Keywords: Nucleoli, invasive breast cancer, DKC1, prognosis

ABSTRACT

BACKGROUND: Hypertrophy of the nucleolus is a distinctive cytological feature of malignant cells and corresponds to aggressive behaviour. This study aimed to identify the key gene associated with nucleolar prominence (NP) in breast cancer (BC) and determine its prognostic significance.

METHODS: From The Cancer Genome Atlas (TCGA) cohort, digital whole slide images identified cancers having NP served as label and an information theory algorithm was applied to find which mRNA gene best explained NP. Dyskerin Pseudouridine Synthase 1 (*DKC1*) was identified. *DKC1* expression was assessed using mRNA data of Molecular Taxonomy of Breast Cancer International Consortium (METABRIC, n=1980) and TCGA (n=855). DKC1 protein expression was assessed using immunohistochemistry in Nottingham BC cohort (n=943).

RESULTS: Nuclear and nucleolar expressions of DKC1 protein were significantly associated with higher tumour grade ($p<0.0001$), high nucleolar score ($p<0.001$) and poor Nottingham Prognostic Index ($p<0.0001$). High DKC1 expression was associated with shorter BC specific survival (BCSS). In multivariate analysis, *DKC1* mRNA and protein expressions were independent risk factors for BCSS ($p<0.01$).

CONCLUSION: DKC1 expression is strongly correlated with NP and its overexpression in BC is associated with unfavourable clinicopathological characteristics and poor outcome. This has been a detailed example in the correlation of phenotype with genotype.

INTRODUCTION

Breast cancer (BC) is the most common cancer diagnosed in women worldwide, accounting for approximately 1 in 10 new cancer diagnoses each year and is the second most common cause of death[1,2] due to cancer. BC is a heterogeneous disease with variable morphologies and response to therapy. Some morphological features, especially histological grade, have been well validated to have a strong prognostic value and their assessment helps in prognostic stratification of BC patients for treatment decisions[3].

In the Nottingham cohort, nucleolar prominence (NP) has recently been shown to be a significant predictor for patient outcome as well as of being highly correlated with tumour grade. Since the NP is a distinctive morphological attribute, it is hypothesised to possibly serve as a substitute for the highly subjective pleomorphism component score of the Nottingham BC grading[4]. Consequently, it is deemed imperative to explore the correlations between the nucleolar phenotype and genotype.

The major function of the nucleolus is synthesis and assembly of ribosomes[5] where both are associated with malignant transformation and cancer progression[6,7]. Indeed, ribosome biogenesis depends on the cancer growth rate which is directly related to nucleolar size of malignant cells. Nucleolar size and cell kinetics' parameters are interrelated because of the increasing rate of ribosome biogenesis in proliferating cells[8]. In some solid cancer and haematological malignancies, the ribosome biogenesis rate increases as a consequence of overexpression of the oncogene c-Myc which controls all the steps of ribosome biogenesis[9]. Despite the biological and clinical significance of NP in BC, the key gene associated with NP and its prognostic significance remains to be defined.

Dyskerin Pseudouridine Synthase 1 (DKC1) is a predominantly nucleolar protein encoded by *DKC1* gene and mapped at Xq28[10]. DKC1 is a crucial component of the telomerase complex and is required for normal telomere maintenance and post-transcriptional processing of precursor rRNA. Therefore, DKC1 is necessary for tumour cell progression through mechanisms related to its function in the processing of rRNA precursor[11]. Usually, clinically indolent and slow-growing tumours express lower levels of DKC1 and its inhibition slows or hinders the proliferation in most cell types[11,12]. Through various deprivation of function approaches, emerging evidence suggests that DCK1 may regulate other cellular processes, including IRES-mediated translation, telomere maintenance independent of telomere length regulation, mitosis, transcription, and possibly microRNA processing[13,14]. Upregulation of DKC1 expression has been reported in several human cancers including hepatocellular carcinoma[15], neuroblastoma[16], lymphoma[17], melanoma[18], prostate cancer[19], colorectal cancer[20], and ovarian carcinoma[21].

MATERIALS AND METHODS

Principle of DKC1 selection

We have applied an information theory (IT) approach to The Cancer Genome Atlas (TCGA) breast cancer dataset. The IT approach was used for feature selection to identify the key gene associated with NP which was assessed morphologically in full face invasive BC sections stained with haematoxylin & eosin (H&E) using digital whole slide images (WSI) as explained in our previous study[4]. The TCGA BC cohort was employed since it contains satisfactory whole slide images and mRNA-seq2 data present in 743 cancers. In the IT approach utilised, the nucleolar score served as a label and 20,339 mRNA transcripts served as predictor variables. The IT algorithm is a 'greedy' algorithm and reduces the number of features selected. "Greedy" is the term used in the machine learning community to describe an algorithm which selects the optimal feature at each step and does not alter any choices already made based on findings from future choices[22]. The analysis showed that the attribute which exhibited the highest mutual information (information gain) with NP was *DKC1*. Moreover, the detection of *DKC1* was supported by LASSO regression feature selection. The required functions for LASSO were obtained from R library Glmnet[23]. LASSO regression is also capable of reducing the number of predictors and thereby allowing for a focused study of a few attributes[24]. Therefore, by applying the IT approach and LASSO regression, the selection was limited to a single gene (*DKC1*). This was followed by evaluating *DKC1* mRNA and DKC1 protein expression in large clinically annotated cohorts of BC to evaluate its clinicopathological and prognostic value in invasive BC as described below.

Study cohorts for transcriptomic analysis

The discovery of *DKC1* was by study of TCGA cohort. The TCGA was also used to assess the possible correlation between *DKC1* mRNA expression and the variables recorded in this cohort[25]. The Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) cohort (n=1980) was used to evaluate *DKC1* gene copy number (CN) aberrations and gene expression[26]. Genomic and transcriptomic data for the METABRIC cohort had been obtained using the Affymetrix SNP 6.0 and Illumina HT-12v3 platforms, respectively[26]. The association between *DKC1* mRNA expression, copy number aberrations and clinicopathological parameters, molecular subtypes, and patient outcome was investigated. Breast Cancer Gene Expression Miner online dataset v4.3 (http://bcgenex.centregauducheau.fr/BC-GEM/GEM-requete.php) was also used as external validation of *DKC1* mRNA expression.

Study cohort for protein expression

The Nottingham BC patient cohort was used to evaluate the immunohistochemical (IHC) expression of DKC1. This cohort is a well-characterised large series (n=943) of invasive

BC patients aged ≤70 years and presented at Nottingham City Hospital between 1999 and 2006. The cohort has long-term clinical follow-up and clinicopathological data included patient's age at diagnosis, histological tumour type, tumour grade, tumour size, lymph node status, Nottingham Prognostic Index (NPI), and lymphovascular invasion (LVI). Patient outcome data was obtained including BC-specific survival (BCSS), defined as the time (in months) from the date of primary surgical treatment to the time of death from BC and distant metastasis free survival (DMFS) defined as time (in months) from primary surgical treatment until the first event of distant metastasis. Patients were treated based on tumour features, NPI and hormone receptor status. Endocrine therapy was given to patients who had oestrogen receptor positivity (ER+) tumours with high NPI scores (>3.4), whereas no adjuvant therapy was given to patients with 'good' NPI scores (≤3.4). Premenopausal patients with moderate and poor NPI scores were candidates for chemotherapy, while postmenopausal patients with 'moderate' or 'poor' NPI scores were given hormonal therapy only. Classical treatment of cyclophosphamide, methotrexate and fluorouracil (CMF) was used as a therapy for patients presented with absence of ER expression and clinically fit to receive chemotherapy. None of the patients in the current study cohort received neoadjuvant therapy. Data related to the expression of ER, progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER2) as well as Ki67 were available[27-30]. Molecular subtypes were based on tumour immunohistochemical (IHC) profile and the Elston–Ellis[31] mitotic score as ER+/HER2−; low proliferation (mitotic score 1), ER+/HER2− high proliferation (mitotic scores 2 and 3), HER2-positive class: HER2+ regardless of ER status, TN: ER−, PR−, and HER2−[32]. There was no significant differences in the distribution of the clinicopathological parameters between the Nottingham and the METABRIC cohorts (all correlation coefficients ≥0.948, all $p < 0.0001$)[33] (**Supplementary Table 1**).

DKC1 validation by Western blotting

The antibody specificity of anti-DKC1 antibody (EPR10399, Abcam, UK) was validated using western blotting (WB) performed on cell lysates of a wild and transfected MDA-MB-231 human breast cancer cell line (American Type Culture Collection; Rockville, MD, USA). The forward transfection of siRNA procedure was followed according to *DKC1* siRNAs manufacturer's instructions. In brief, the cells were seeded in 6 well plate at a cell density of 3×10^5 cells per well and incubated overnight in 37^0C 5% CO_2 incubator. The following day, the cells reached about 40% confluence and were transfected with 10nM of three different IDs of *DKC1* siRNA (Cat#:4392420, ThermoFisher Scientific, UK). A transfection with 10nM scrambled siRNA sequence (Cat#:4390843, ThermoFisher Scientific, UK) was carried out in the experiment and considered as a negative control. DKC1 protein expression of untransfected & transfected cells was then determined by the Western

blotting. Briefly, after collecting the cell lysates, a dilution of 1:1000 of the primary antibody and 1:15000 IRDye 800CW Donkey anti-rabbit secondary antibody (LI-COR Biosciences) were applied, and 5% milk /PBS-Tween (0.1%) (Marvel Original Dried Skimmed Milk, Premier Food Groups Ltd., UK) was used for blocking and antibodies incubation. Mouse monoclonal anti-β-actin primary antibody (1:5000) (Sigma-Aldrich, UK) with IRDye 800CW Donkey anti-mouse fluorescent secondary antibody (LI-COR Biosciences) were used to visualise a marker of endogenous control. Visualisation of *DKC1* band was done by using the Odyssey Fc with Image Studio 4.0 (LI-COR Biosciences).

Tissue microarrays and immunohistochemical analysis

Invasive BC tissues were previously arrayed as tissue microarrays (TMA) using the Grand Master® (3D HISTECH®, Budapest, Hungary)[34]. IHC staining was performed on 4 µm TMA thick sections using the Novocastra Novolink™ Polymer Detection Systems kit (Code: RE7280-K, Leica, Biosystems, Newcastle, UK). Antigen retrieval was performed in citrate buffer pH 6.0 using a microwave (Whirlpool JT359 Jet Chef 1000W) for 20 min. Rabbit monoclonal DKC1 was diluted at 1 : 50 in Leica antibody diluent (RE AR9352, Leica, Biosystems, UK) and incubated with the sections for 60 min at room temperature. A negative control was obtained by omitting the incubation with primary antibody while formalin fixed placenta tissue was used as a positive control according to manufacturer's datasheet.

Assessment of DKC1 protein expression

Scanning of TMA stained sections into high-resolution digital images was performed by using a NanoZoomer scanner (NanoZoomer; Hamamatsu Photonics, Welwyn Garden City, UK) at 20x magnification. Scoring of DKC1 nuclear and nucleolar[4] expression was evaluated based on a semi-quantitative scoring using modified histochemical score (H-score), where the intensity of staining was multiplied by the percentage of positive cells in the tissue for each intensity, producing a score ranging between 0 and 300[35]. A score index of 0, 1, 2, and 3 corresponding to negative, weak, moderate, and strong respectively were used for intensity. The percentage (0-100) of positive cells for each intensity was evaluated subjectively. All non-representative cores including folded tissue during processing and staining, cores with only normal breast tissue and cores with invasive tumour < 15% of core surface area were excluded from scoring. All the cores were scored by a trained observer (K. Elsharawy) blinded of histopathological and patient outcome data. Further, to test the inter-observer's reproducibility of the scoring, a subset of TMA cores (10%) was randomly selected and double scored by a second trained observer (M. Aleskandarany). Moreover, for further evaluation of scoring reproducibility, 20% of the cases were double scored by the main observer (K. Elsharawy) after 5 months washout period blind from the first scores.

Statistical analysis

IBM-SPSS statistical software 24.0 (SPSS, Chicago, IL, USA) was used in statistical analysis. Inter-observer agreement in DKC1 IHC scoring was assessed using intra-class correlation coefficient (ICC). Dichotomisation of DKC1 proteomic and transcriptomic levels expression was determined based on the prediction of BCSS using X-tile bioinformatics software version 3.6.1 (School of Medicine, Yale University, New Haven, USA)[36]. The H-scores of 110 & 10 were the optimal cut-off values of DKC1 nuclear and nucleolar protein expression. Continuous data of *DKC1* mRNA and DKC1 protein expression were used to assess the association with clinicopathological parameters. Differences between three or more groups were investigated using 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). Student *t*-test (parametric data) or Mann-Whitney test (non-parametric distribution) were used to evaluate the differences between two groups. Spearman's correlation coefficient was calculated to examine the association between continuous variables. Univariate analysis was visualised using Kaplan-Meier curves and significance was assessed by log-rank test. Cox's proportional hazard regression models were built for the multivariate survival analysis to adjust for confounding factors. *P*-values were adjusted by using Bonferroni correction for multiple testing. For all tests, p value < 0.05 was considered as statistically significant. This study followed the reporting recommendations for tumour markers prognostic studies (REMARK) criteria[37].

RESULTS

In this study, we have applied the bespoke bioinformatics tools to identify the key genes associated with NP and this identified *DKC1* as the target gene. Then, DKC1 was investigated at the transcriptomic, genetic and protein levels.

DKC1 mRNA expression and CN aberrations

High *DKC1* mRNA expression (log2 intensity >9.4) was observed in 709/1970 (36%) of the METABRIC cases. In all, 77/1980 (4%) of cases showed *DKC1* CN gain, whereas 115/1980 (6%) showed a CN loss. A significant association was observed between *DKC1* CN variation and *DKC1* mRNA expression (p<0.0001) **(Fig. 1A**).

DKC1 protein expression in BC tissues

Prior to IHC staining, the specificity of the antibody used was validated using WB performed on a *DKC1* siRNA transfected BC cell line. A specific band at the predicted DKC1 molecular weight (58 kDa) was detected for proteins extracted from untransfected cells and those transfected with scrambled siRNA sequences. In addition, the DKC1 band

intensity was significantly reduced with proteins extracted from *DKC1* siRNAs transfected cells, confirming the specificity of the antibody utilised. A single band was observed at β-actin molecular weight (42 kDa) demonstrating the uniformity of loaded protein quantities **(Fig. 2A).**

DKC1 protein expression was observed in the nucleus and nucleoli of invasive BC cells, with expression levels varying from absent to strong (**Fig. 2B, 2C & 2D**). Strong concordance was observed between the two observers in DKC1 immuno-scoring in 10% of the cases (ICC=0.864, $p<0.0001$ for nuclear expression & ICC=0.781, $p<0.001$ for nucleolar expression). Moreover, second scoring of 20% of cases after 5 months washout period confirmed concordance ((ICC=0.822, $p<0.0001$ for nuclear expression & ICC= 0.804, $p<0.0001$ for nucleolar expression). At the optimal DKC1 cut-off values (H-score 110 & 10, respectively), High DKC1 nuclear and nucleolar expression were observed in 574/942 (61%) and 153/942 (16%) of the informative tumours, respectively. There was a significant positive correlation between DKC1 nuclear and nucleolar expression (n=429) (correlation coefficient= 0.143, $p<0.0001$).

Correlation of DKC1 mRNA and protein expression with clinicopathological parameters

High *DKC1* mRNA expression was significantly associated with younger patient age, larger tumour size, higher tumour grade and poorer NPI ($p<0.001$, $p=0.024$, $p<0.0001$ & $p<0.0001$) as shown in **Fig. 1B-E,** respectively. These associations were confirmed using the Breast Cancer Gene-Expression Miner v4.3 (**Supplementary Fig. 1A-C**)

In the TCGA BC dataset, similar associations, as described above, were observed with clinicopathological parameters. In particular, high *DKC1* mRNA expression was significantly associated with high nucleolar score 3 [4] ($p<0.0001$, **Supplementary Table 2**).

High expression of DKC1 protein whether in the nucleus and/or nucleoli was associated with aggressive features of BC including higher tumour grade ($p<0.0001$), larger tumour size ($p=0.04$ only with nucleolar expression), higher mitotic scores ($p<0.0001$), increased nuclear pleomorphism ($p<0.0001$), higher scores of nucleolar prominence ($p<0.001$), poor NPI ($p<0.0001$) and the invasive ductal no special histological type (NST) ($p<0.0001$), **Table 1.**

DKC1 expression and other markers

The correlation of *DKC1* mRNA with other relevant genes was investigated using the METABRIC and TCGA datasets. The genes were chosen based on published information, being either regulatory genes or those that share or support *DKC1* biological function

especially those primarily involved in the ribosomal biogenesis. *DKC1* was positively associated with *GAR1* (p<0.0001), *NOP10* (p<0.001), and *NHP2* (p<0.0001). Moreover, there was a significant association between *DKC1 MKI67* and the regulatory genes *c-Myc* (all p<0.001) (**Supplementary Table 3**). High *DKC1* mRNA expression was associated with those tumours which showed *TP53* mutations (p<0.0001, **Supplementary Table 4**). Also, the statistical analysis showed a significant positive association of high DKC1 protein expression with high Ki67 (χ^2=8.815, p = 0.003).

DKC1 mRNA and protein expression in BC molecular subtypes

At the transcriptomic level in METABRIC cohort, high *DKC1* expression was significantly associated with hormone receptor negative (ER- and PR-), HER2+ tumours and TNBC (all p<0.0001) as shown in **Supplementary Table 4**. Similar results were observed upon analysing the publicly available gene expression data available on the Breast Cancer Gene-Expression Miner v4.3 online platform (**Supplementary Fig. 1D-G**) and TCGA datasets (**Supplementary Table 4**).

Regarding the association with the intrinsic PAM50 subtypes[38], high expression of *DKC1* mRNA was observed in basal-like, HER2+ and Luminal B tumours (**Fig. 1F**, p<0.0001). These findings were confirmed using the Breast Cancer Gene-Expression Miner v4.3 (**Supplementary Fig. 1H**). In the SCMGENE subtypes, high expression of *DKC1* mRNA was observed in the ER-/HER2- cases followed by ER+/HER2- high proliferation class (p<0.0001, **Fig. 1G**).

DKC1 nuclear and nucleolar protein expression was associated with negative ER status (p=0.04 & p<0.0001 respectively). Moreover, DKC1 nucleolar protein showed a significant correlation within HER2+ and triple negative (TN) tumours (both p<0.0001), **Table 2.**

There was a higher protein expression of DKC1 (nuclear & nucleolar) in the ER+ high proliferative tumours than in the other molecular subtypes (p<0.0001) as shown in **Table 1.**

Correlation of DKC1 mRNA and protein expression with patient outcome

In METABRIC cohort, high *DKC1* mRNA expression was associated with poor BCSS in all cases (HR = 1.5, 95%CI = 1.3–1.8; p<0.0001). Moreover, *DKC1* mRNA expression was predictive of BCSS only in luminal B cases (HR = 1.5, 95%CI = 1.1–2.1; p=0.015) as shown in **Supplementary Fig. 2A-E**. The relationship between high *DKC1* mRNA expression and poor patient outcome in ER+ disease, but not ER- disease, was shown using the TCGA cohort (**Supplementary Fig. 3**).

Both high DKC1 nuclear and nucleolar protein expressions, when assessed individually, were associated with poor outcome (HR = 2.5, 95%CI = 1.7–3.7; $p<0.0001$ & HR = 1.5, 95%CI = 1.1–2.2; $p=0.038$ respectively) **Fig. 3A, B**. When the analysis was limited to molecular subtypes, high expression of DKC1 nuclear protein, was significantly associated with poor outcome in ER+ high proliferation tumours (HR = 4.4, 95%CI = 1.6–12.3; $p=0.002$), HER2+ tumours (HR = 2.6, 95%CI = 1.1–6.7; $p=0.039$) and TNBC (HR = 1.5, 95%CI = 1.1–6.2; $p=0.035$) **Fig. 3D-G**. However, no significant association of DKC1 nucleolar protein expression was identified with outcome in BC subtypes ($p>0.05$).

For further analysis, combinatorial DKC1 protein expression groups were created [i.e. low nuclear/low nucleolar, high nuclear/low nucleolar expression, low nuclear/high nucleolar and high nuclear/high nucleolar]. A significant difference in patient survival was observed between these four groups, where tumours with low nuclear and low nucleolar DKC1 expression showed the best outcome, whereas the tumours with high nuclear and high nucleolar expression showed the worst outcome ($p<0.0001$) **Fig.3C**.

The multivariate Cox-proportional models, including other prognostic covariates such as tumour size, grade and nodal stage, showed that DKC1 nuclear, combinatorial protein expression and *DKC1* mRNA in the METABRIC dataset were independent predictors for poor prognosis in whole cases ($p=0.001$, HR 2.037, 95% CI= 1.373–3.023, $p=0.003$, HR 2.746, 95% CI= 1.484–5.083 & $p=0.006$, HR 1.316, 95% CI =1.097–1.579) as shown in **Table 3 & Supplementary Table 5**.

In addition, high DKC1 nuclear protein expression was significantly associated with shorter distant metastases-free survival (DMFS) (HR = 2.1, 95%CI = 1.5–2.9; $p<0.0001$). Likewise, combinatorial protein expression was associated with shorter DMFS (HR = 1.3, 95%CI = 1.1–1.4; $p=0.001$) (**Supplementary Fig. 4**).

DISCUSSION

In malignant tumours, the number and size of nucleoli are usually an indication of the rate of ribosome production which is regarded as a major metabolic requisite for cell growth and proliferation[39,40]. Nucleolar function and size are directly related to cell doubling time in cancer cells and quantitative morphometric evaluation of nucleolar size was considered as a cytological parameter of the tumour cells proliferation rate[4,41]. In this study, we assessed NP in the TCGA breast cancer dataset as previously described [4], and used two greedy algorithms information theory and validated it using LASSO regression test[22-24] to identify genes driving NP. This demonstrated that out of the 20,339 genes investigated, *DKC1* was the top differentially expressed gene. Then DKC1 expression was evaluated at the proteomic, transcriptomic and genomic levels in large cohorts of invasive BC.

There were significant associations between high *DKC1* mRNA in TCGA breast cancer dataset and *DKC1* protein expression in Nottingham cohort with high nucleolar scoring[4]. These findings supported our hypothesis that DKC1 plays a role in the nucleoli appearance and size likely through its mechanism in ribosomal biogenesis.

Our results also showed positive correlations between nuclear and nucleolar DKC1 protein expression in the breast tumour cells. It was reported that newly synthesised DKC1 initially localizes to the nucleoplasm, followed by consecutive translocation to the nucleoli and the nuclear Cajal bodies. Usually, co-localisation of DKC1 on the nuclear Cajal bodies occurred only when it had already accumulated in the nucleoli[42]. DKC1 is involved in the pseudouridination and processing of small spliceosomal RNAs through its binding to H/ACA small Cajal body RNAs[43].

The current study confirms the significant association between the high expression of DKC1, at both protein and mRNA levels, and clinicopathological parameters characteristics of poor prognosis and with shorter survival; findings which are in line with other studies[15,44,45]. Some studies have also confirmed that *DKC1* overexpression is involved in tumourigenic processes and has prognostic value in numerous types of cancer[19,21,46]. The association between *DKC1* mRNA and shorter survival was identified in both METABRIC and TCGA cohorts. Moreover, our analysis revealed that the prognostic significance of DKC1 protein and mRNA in BC was independent of other variables, demonstrating its potential clinical relevance in improving survival rate prediction.

When BC molecular subtypes were considered, the significant association between DKC1 protein and poor patient outcome was observed in the ER+ high proliferation (i.e. luminal B), HER2+ and TNBC classes whereas the high mRNA expression was only limited to the luminal B subtype. The most common type of BC constituting nearly 55–70% is the ER+/luminal tumour, and those tumours are variable in terms of recurrence, mortality

rates and disease prognosis[47,48]. These observations further endorse DKC1 functions in playing crucial roles in tumour growth and progression

DKC1 performs two fundamental functions for cell proliferation. Firstly, DKC1 is a component of the H/ACA small nucleolar ribonucleoprotein particles (snoRNPs) involved the pseudouridylation of ribosomal RNA (rRNA) molecules and necessary for their processing. Secondly, it is required for telomerase activity by stabilising the telomerase RNA component[11]. The faster the rate of cell proliferation, the higher the demand for protein production which is compatible with increased rRNA synthesis[49,50]. It has been reported that any dysregulation of DKC1 levels results in defects of ribosome biogenesis and a reduction of rRNA pseudouridylation which in turn hinders the normal ribosome rRNA processing rate[51]. For instance, Montanaro et al. have demonstrated that reduced *DKC1* gene expression by specific RNA interference in BC cell lines resulted in a reduction of rRNA pseudouridylation which subsequently effected the survival of proliferating cells[52]. The role of DKC1 in mitosis was also confirmed[11], where dyskerin was identified as one of seventy genes which correlated with the development of aneuploidy. Alawi et al. have demonstrated that dyskerin expression peaks during G_2/M and loss of dyskerin function has a widely disruptive effect on mitosis and triggers the spindle-assembly checkpoint[13]. Our findings showed that high DKC1 expression was significantly associated with proliferation as assessed by Ki67 labelling index, which was also observed in other studies in BC[52] hepatocellular carcinoma[15] and prostate cancer[19]; confirming that DKC1 is critical for mitotic progression and proliferation in these cancers.

DKC1 is the direct and conserved transcriptional target of c-Myc[53], which explains the strong correlation between its up-regulation and active cell proliferation[54]. In our study, we observed a significant positive association between *DKC1* and *c-Myc* in mRNA expression. Previous studies have demonstrated that tumour oncogene *c-Myc* controls the transcription of *DKC1* gene in addition to other proteins which are required for rRNA processing[9,55]. *TP53* mutations were also highly prevalent in breast tumours with high *DKC1* mRNA expression in METABRIC. There is mounting evidence that the usual increase of ribosome biogenesis (one of the main functions of *DKC1*) in cancer cells is the consequence of frequent alterations of two major tumour suppressors*, TP53* and retinoblastoma (RB) genes[55]. In addition, tumours with altered *p53* and/or retinoblastoma protein *pRb* functions are characterised by significantly larger/more conspicuous nucleoli than tumours with normally functioning *p53* and *pRb*[56].

We further investigated the association of *DKC1* expression with other H/ACA ribonucleoproteins, *NHP2*, *NOP10*, and *GAR1* which play important roles in disease progression. *DKC1*, *NHP2*, and *NOP10* form a core trimer that directly binds to H/ACA RNAs. The three proteins are interdependent with each other for stability and also regulate

constancy of the bound RNAs[57]. *GAR1* binds only to *DKC1* and is needed for a proper functioning of the H/ACA RNPs, but its absence does not reduce the stability of the rRNA[58]. These findings confirmed the significant positive correlation between H/ACA ribonucleoproteins and *DKC1* in BC[43]. Alterations in *DKC1* expression will potentially disrupt the biogenesis of H/ACA pathway and consequently affect ribosome synthesis and impair cell proliferation.

In the last decade, there have been a few attempts to construct *DKC1* inhibitors. However, one in silico study successfully determined a small molecule inhibitor (Pyrazofurin) that exerted an ability to weaken the pharmacological and physiological activities of DKC1 through inhibiting its function in pseudouridylation of rRNA. Although Pyrazofurin failed to progress phase II clinical trials, however, its chemical structure should continue to be exploited as a pharmacokinetic model to develop a potent, effective and safe *DKC1* inhibitor that may eventually be used for BC highly expressing *DKC1*[59].

A few limitations of this study findings are worth mentioning. On one hand, the semi-quantitative H-score method used to evaluate the immunohistochemical protein expression in the Nottingham cohort, might be regarded as having substantial subjectivity. This was addressed by double scoring a subset of the cancers to ensure the reproducibility and liability of the procedure. On the other hand, only one representative TMA core from each tumour tissue was arrayed and scored instead of considering replicates to express the tumour heterogeneity. This was due to the limited tissue resources in our biobank. However, to overcome the issue, 20 full face sections of randomly selected breast cancer cases were stained, prior to TMA application, with DKC1 antibody to assess the staining homogeneity and to evaluate the pertinence of using tissue microarrays (TMAs). These showed homogeneously distributed DKC1 expression, deeming the use of TMA to assess DKC1 an appropriate tissue platform, to mitigate the limited resources as well as testing the hypothesis a large BC cohort. Finally, the 'weak' but significant correlation between nuclear and nucleolar expression of DKC1 might also be regarded as a weakness point in the study. This might be due to the subjectivity of the method which has been used in determining NP in our previous study[4].

CONCLUSION

This study reveals a significant correlation between the morphological features of NP and an underlying molecular and protein description (DKC1). The importance of DKC1 was demonstrated in three independent datasets where each dataset contributed to the description of DKC1 from different perspectives. DKC1 is significantly associated with high nucleolar score and with poor prognostic characteristics and poor patients' outcome. Overexpression of DKC1 appears to play a role in the proliferation and progression of the aggressive BC subtypes including the luminal B, TNBC, and HER2 molecular subtypes.

449 Findings here encourage further investigation of DKC1 as it might relate to guiding
450 targeted therapies and to evaluate its role in response to chemotherapy.

451

ADDITIONAL INFORMATION

ACKNOWLEDGMENTS

We thank Breast Cancer Now and PathLAKE project for supporting this study. These new Centres are supported by the Data to Early Diagnosis and Precision Medicine strand of the government's Industrial Strategy Challenge, managed and delivered by UK Research and Innovation (UKRI).

AUTHOR CONTRIBUTION

KE: scored all the cases, took the lead in writing the manuscript, data analysis and interpretation. **OM & MA:** helped in double scoring, data interpretation and reviewing the article**. AH, HG & MIA:** contributed in data analysis, interpretation, writing and reviewing the article. **AG & L D:** contributed in data analysis**,** study design and reviewing the article. **ER:** conceived and planned the presented idea, data interpretation and reviewing the article. All authors reviewed and approved the final version of the manuscript. All authors declared their contribution(s) to the study

ETHICAL APPROVAL AND CONSENT TO PARTICIPATE

This work obtained ethics approval by the North West – Greater Manchester Central Research Ethics Committee under the title; Nottingham Health Science Biobank (NHSB), reference number 15/NW/0685. All patients included were consented to participate in the study and to use their materials in research. All samples from Nottingham used in this study were pseudo-anonymized and stored in compliance with the UK Human Tissue Act. The study was performed in accordance with the Declaration of Helsinki.

DATA AVAILABILITY

The authors confirm the data that has been used in this work is available on reasonable request.

COMPETING INTERESTS

The authors declare that they have no competing of interest.

FUNDING

This research was supported and funded by Damietta University and Egyptian Ministry of Higher education and Scientific Research.

REFERENCES

1. Parada H, Jr., Sun X, Tse CK, Olshan AF, Troester MA. Lifestyle Patterns and Survival Following Breast Cancer in the Carolina Breast Cancer Study. *Epidemiology (Cambridge, Mass)* 2019; **30**(1): 83-92; e-pub ahead of print 2018/10/10; doi 10.1097/ede.0000000000000933.
2. Andisha NM, McMillan DC, Gujam FJA, Roseweir A, Edwards J. The relationship between phosphorylation status of focal adhesion kinases, molecular subtypes, tumour microenvironment and survival in patients with primary operable ductal breast cancer. *Cellular Signalling* 2019; doi https://doi.org/10.1016/j.cellsig.2019.04.006.
3. Rakha EA, El-Sayed ME, Lee AH, Elston CW, Grainge MJ, Hodi Z *et al.* Prognostic significance of Nottingham histologic grade in invasive breast carcinoma. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2008; **26**(19): 3153-3158; e-pub ahead of print 2008/05/21; doi 10.1200/jco.2007.15.5986.
4. Elsharawy KA, Toss MS, Abuelmaaty SR, Ball G, Green AR, Aleskandarany MA *et al.* Prognostic Significance of Nucleolar Assessment in Invasive Breast Cancer. *Histopathology* 2019; e-pub ahead of print 2019/11/19; doi 10.1111/his.14036.
5. Lam YW, Trinkle-Mulcahy L, Lamond AI. The nucleolus. *Journal of Cell Science* 2005; **118**(7): 1335; doi 10.1242/jcs.01736.
6. Montanaro L, Trere D, Derenzini M. Nucleolus, ribosomes, and cancer. *The American journal of pathology* 2008; **173**(2): 301-310; e-pub ahead of print 2008/06/28; doi 10.2353/ajpath.2008.070752.
7. Bywater MJ, Poortinga G, Sanij E, Hein N, Peck A, Cullinane C *et al.* Inhibition of RNA polymerase I as a therapeutic strategy to promote cancer-specific activation of p53. *Cancer cell* 2012; **22**(1): 51-65; e-pub ahead of print 2012/07/14; doi 10.1016/j.ccr.2012.05.019.
8. Derenzini M, Trerè D, Pession A, Govoni M, Sirri V, Chieco P. Nucleolar size indicates the rapidity of cell proliferation in cancer tissues. *The Journal of pathology* 2000; **191**(2): 181-186; doi 10.1002/(sici)1096-9896(200006)191:2<181::Aid-path607>3.0.Co;2-v.
9. van Riggelen J, Yetil A, Felsher DW. MYC as a regulator of ribosome biogenesis and protein synthesis. *Nature Reviews Cancer* 2010; **10**(4): 301-309; doi 10.1038/nrc2819.
10. Heiss NS, Knight SW, Vulliamy TJ, Klauck SM, Wiemann S, Mason PJ *et al.* X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions. *Nature genetics* 1998; **19**(1): 32-38; e-pub ahead of print 1998/05/20; doi 10.1038/ng0598-32.
11. Alawi F, Lin P. Dyskerin is required for tumor cell growth through mechanisms that are independent of its role in telomerase and only partially related to its function in precursor rRNA processing. *Mol Carcinog* 2011; **50**(5): 334-345; e-pub ahead of print 2010/12/10; doi 10.1002/mc.20715.
12. Lin P, Mobasher ME, Alawi F. Acute dyskerin depletion triggers cellular senescence and renders osteosarcoma cells resistant to genotoxic stress-induced apoptosis. *Biochemical and biophysical research communications* 2014; **446**(4): 1268-1275; e-pub ahead of print 2014/04/03; doi 10.1016/j.bbrc.2014.03.114.
13. Alawi F, Lin P. Dyskerin Localizes to the Mitotic Apparatus and Is Required for Orderly Mitosis in Human Cells. *PLOS ONE* 2013; **8**(11): e80805; doi 10.1371/journal.pone.0080805.
14. Fong YW, Ho JJ, Inouye C, Tjian R. The dyskerin ribonucleoprotein complex as an OCT4/SOX2 coactivator in embryonic stem cells. *eLife* 2014; **3**; e-pub ahead of print 2014/11/20; doi 10.7554/eLife.03573.
15. Liu B, Zhang J, Huang C, Liu H. Dyskerin Overexpression in Human Hepatocellular Carcinoma Is Associated with Advanced Clinical Stage and Poor Patient Prognosis. *PLOS ONE* 2012; **7**(8): e43147; doi 10.1371/journal.pone.0043147.

16. Westermann F, Henrich K-O, Wei JS, Lutz W, Fischer M, König R *et al.* High Skp2 Expression Characterizes High-Risk Neuroblastomas Independent of MYCN Status. *Clinical Cancer Research* 2007; **13**(16): 4695-4703; doi 10.1158/1078-0432.Ccr-06-2818.
17. Piva R, Pellegrino E, Mattioli M, Agnelli L, Lombardi L, Boccalatte F *et al.* Functional validation of the anaplastic lymphoma kinase signature identifies CEBPB and Bcl2A1 as critical target genes. *The Journal of Clinical Investigation* 2006; **116**(12): 3171-3182; doi 10.1172/JCI29401.
18. McDonald SL, Edington HD, Kirkwood JM, Becker D. Expression Analysis of Genes identified by Molecular Profiling of VGP Melanomas and MGP Melanoma-Positive Lymph Nodes. *Cancer Biology & Therapy* 2004; **3**(1): 110-120; doi 10.4161/cbt.3.1.662.
19. Sieron P, Hader C, Hatina J, Engers R, Wlazlinski A, Müller M *et al.* DKC1 overexpression associated with prostate cancer progression. *British Journal Of Cancer* (Molecular Diagnostics) 2009; **101**: 1410; doi 10.1038/sj.bjc.6605299
20. Witkowska A, Gumprecht J, Glogowska-Ligus J, Wystrychowski G, Owczarek A, Stachowicz M *et al.* Expression profile of significant immortalization genes in colon cancer. *International journal of molecular medicine* 2010; **25**(3): 321-329; e-pub ahead of print 2010/02/04; doi 10.3892/ijmm_00000348.
21. Schaner ME, Ross DT, Ciaravino G, Sorlie T, Troyanskaya O, Diehn M *et al.* Gene expression patterns in ovarian carcinomas. *Molecular biology of the cell* 2003; **14**(11): 4376-4386; e-pub ahead of print 2003/09/05; doi 10.1091/mbc.e03-05-0279.
22. Brown G, Pocock A, Zhao M-J, Luján M. Conditional likelihood maximisation: a unifying framework for information theoretic feature selection. *J Mach Learn Res* 2012; **13**(null): 27–66.
23. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of statistical software* 2010; **33**(1): 1-22; e-pub ahead of print 2010/09/03.
24. Tibshirani R, Bien J, Friedman J, Hastie T, Simon N, Taylor J *et al.* Strong rules for discarding predictors in lasso-type problems. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 2012; **74**(2): 245-266; doi 10.1111/j.1467-9868.2011.01004.x.
25. Ciriello G, Gatza ML, Beck AH, Wilkerson MD, Rhie SK, Pastore A *et al.* Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer. *Cell* 2015; **163**(2): 506-519; e-pub ahead of print 2015/10/10; doi 10.1016/j.cell.2015.09.033.
26. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ *et al.* The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* 2012; **486**(7403): 346-352; e-pub ahead of print 2012/04/24; doi 10.1038/nature10983.
27. Aleskandarany MA, Abduljabbar R, Ashankyty I, Elmouna A, Jerjees D, Ali S *et al.* Prognostic significance of androgen receptor expression in invasive breast cancer: transcriptomic and protein expression analysis. *Breast cancer research and treatment* 2016; **159**(2): 215-227; e-pub ahead of print 2016/08/16; doi 10.1007/s10549-016-3934-5.
28. Rakha EA, Agarwal D, Green AR, Ashankyty I, Ellis IO, Ball G *et al.* Prognostic stratification of oestrogen receptor-positive HER2-negative lymph node-negative class of breast cancer. *Histopathology* 2017; **70**(4): 622-631; e-pub ahead of print 2016/10/27; doi 10.1111/his.13108.
29. Rakha EA, Elsheikh SE, Aleskandarany MA, Habashi HO, Green AR, Powe DG *et al.* Triple-negative breast cancer: distinguishing between basal and nonbasal subtypes. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2009; **15**(7): 2302-2310; e-pub ahead of print 2009/03/26; doi 10.1158/1078-0432.Ccr-08-2132.

30. Muftah AA, Aleskandarany MA, Al-Kaabi MM, Sonbul SN, Diez-Rodriguez M, Nolan CC *et al.* Ki67 expression in invasive breast cancer: the use of tissue microarrays compared with whole tissue sections. *Breast cancer research and treatment* 2017; **164**(2): 341-348; e-pub ahead of print 2017/05/10; doi 10.1007/s10549-017-4270-0.
31. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology* 1991; **19**(5): 403-410; e-pub ahead of print 1991/11/01; doi 10.1111/j.1365-2559.1991.tb00229.x.
32. Senkus E, Kyriakides S, Ohno S, Penault-Llorca F, Poortmans P, Rutgers E *et al.* Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals of oncology : official journal of the European Society for Medical Oncology* 2015; **26 Suppl 5**: v8-30; e-pub ahead of print 2015/09/01; doi 10.1093/annonc/mdv298.
33. El Ansari R, Craze ML, Miligy I, Diez-Rodriguez M, Nolan CC, Ellis IO *et al.* The amino acid transporter SLC7A5 confers a poor prognosis in the highly proliferative breast cancer subtypes and is a key therapeutic target in luminal B tumours. *Breast Cancer Research* 2018; **20**(1): 21; doi 10.1186/s13058-018-0946-6.
34. Abd El-Rehim DM, Ball G, Pinder SE, Rakha E, Paish C, Robertson JF *et al.* 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. *International journal of cancer* 2005; **116**(3): 340-350; e-pub ahead of print 2005/04/09; doi 10.1002/ijc.21004.
35. McCarty KS, Jr., McCarty KS, Sr. Histochemical approaches to steroid receptor analyses. *Seminars in diagnostic pathology* 1984; **1**(4): 297-308; e-pub ahead of print 1984/11/01.
36. Camp RL, Dolled-Filhart M, Rimm DL. X-Tile. *A New Bio-Informatics Tool for Biomarker Assessment and Outcome-Based Cut-Point Optimization* 2004; **10**(21): 7252-7259; doi 10.1158/1078-0432.Ccr-04-0713.
37. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM *et al.* REporting recommendations for tumour MARKer prognostic studies (REMARK). *British journal of cancer* 2005; **93**(4): 387-391; doi 10.1038/sj.bjc.6602678.
38. Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T *et al.* Supervised risk predictor of breast cancer based on intrinsic subtypes. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2009; **27**(8): 1160-1167; e-pub ahead of print 2009/02/11; doi 10.1200/jco.2008.18.1370.
39. Fischer AH, Bardarov S, Jr., Jiang Z. Molecular aspects of diagnostic nucleolar and nuclear envelope changes in prostate cancer. *Journal of cellular biochemistry* 2004; **91**(1): 170-184; e-pub ahead of print 2003/12/23; doi 10.1002/jcb.10735.
40. Hernandez-Verdun D. The nucleolus: a model for the organization of nuclear functions. *Histochemistry and cell biology* 2006; **126**(2): 135-148; e-pub ahead of print 2006/07/13; doi 10.1007/s00418-006-0212-3.
41. Derenzini M, Trere D, Pession A, Montanaro L, Sirri V, Ochs RL. Nucleolar function and size in cancer cells. *The American journal of pathology* 1998; **152**(5): 1291-1297; e-pub ahead of print 1998/05/20.
42. Heiss NS, Girod A, Salowsky R, Wiemann S, Pepperkok R, Poustka A. Dyskerin Localizes to the Nucleolus and Its Mislocalization Is Unlikely to Play a Role in the Pathogenesis of Dyskeratosis Congenita. *Human Molecular Genetics* 1999; **8**(13): 2515-2524; doi 10.1093/hmg/8.13.2515.
43. Ye K. H/ACA guide RNAs, proteins and complexes. *Current opinion in structural biology* 2007; **17**(3): 287-292; e-pub ahead of print 2007/06/19; doi 10.1016/j.sbi.2007.05.012.
44. Zhang M, Pan Y, Jiang R, Hou P, Shan H, Chen F *et al.* DKC1 serves as a potential prognostic biomarker for human clear cell renal cell carcinoma and promotes its proliferation, migration and invasion via the NFkappaB pathway. *Oncology reports* 2018; **40**(2): 968-978; e-pub ahead of print 2018/06/15; doi 10.3892/or.2018.6484.

45. von Stedingk K, Koster J, Piqueras M, Noguera R, Navarro S, Påhlman S *et al.* snoRNPs Regulate Telomerase Activity in Neuroblastoma and Are Associated with Poor Prognosis. *Translational oncology* 2013; **6**(4): 447-457; doi 10.1593/tlo.13112.
46. Poncet D, Belleville A, t'kint de Roodenbeke C, Roborel de Climens A, Ben Simon E, Merle-Beral H *et al.* Changes in the expression of telomere maintenance genes suggest global telomere dysfunction in B-chronic lymphocytic leukemia. *Blood* 2008; **111**(4): 2388-2391; e-pub ahead of print 2007/12/14; doi 10.1182/blood-2007-09-111245.
47. Dawson SJ, Rueda OM, Aparicio S, Caldas C. A new genome-driven integrated classification of breast cancer and its implications. *The EMBO journal* 2013; **32**(5): 617-628; e-pub ahead of print 2013/02/12; doi 10.1038/emboj.2013.19.
48. Rakha EA, El-Sayed ME, Green AR, Paish EC, Powe DG, Gee J *et al.* Biologic and clinical characteristics of breast cancer with single hormone receptor positive phenotype. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2007; **25**(30): 4772-4778; e-pub ahead of print 2007/09/19; doi 10.1200/jco.2007.12.2747.
49. Mochizuki Y, He J, Kulkarni S, Bessler M, Mason PJ. Mouse dyskerin mutations affect accumulation of telomerase RNA and small nucleolar RNA, telomerase activity, and ribosomal RNA processing. *Proceedings of the National Academy of Sciences of the United States of America* 2004; **101**(29): 10756-10761; e-pub ahead of print 2004/07/09; doi 10.1073/pnas.0402560101.
50. Filipowicz W, Pogacic V. Biogenesis of small nucleolar ribonucleoproteins. *Current opinion in cell biology* 2002; **14**(3): 319-327; e-pub ahead of print 2002/06/18; doi 10.1016/s0955-0674(02)00334-4.
51. Ruggero D, Grisendi S, Piazza F, Rego E, Mari F, Rao PH *et al.* Dyskeratosis congenita and cancer in mice deficient in ribosomal RNA modification. *Science (New York, NY)* 2003; **299**(5604): 259-262; e-pub ahead of print 2003/01/11; doi 10.1126/science.1079447.
52. Montanaro L, Brigotti M, Clohessy J, Barbieri S, Ceccarelli C, Santini D *et al.* Dyskerin expression influences the level of ribosomal RNA pseudo-uridylation and telomerase RNA component in human breast cancer. *The Journal of pathology* 2006; **210**(1): 10-18; doi 10.1002/path.2023.
53. Alawi F, Lee MN. DKC1 is a direct and conserved transcriptional target of c-MYC. *Biochemical and biophysical research communications* 2007; **362**(4): 893-898; e-pub ahead of print 2007/09/08; doi 10.1016/j.bbrc.2007.08.071.
54. Alawi F, Lin P, Ziober B, Patel R. Correlation of dyskerin expression with active proliferation independent of telomerase. *Head & neck* 2011; **33**(7): 1041-1051; e-pub ahead of print 2011/06/16; doi 10.1002/hed.21579.
55. Penzo M, Montanaro L, Treré D, Derenzini M. The Ribosome Biogenesis-Cancer Connection. *Cells* 2019; **8**(1): 55; doi 10.3390/cells8010055.
56. Trere D, Ceccarelli C, Montanaro L, Tosti E, Derenzini M. Nucleolar size and activity are related to pRb and p53 status in human breast cancer. *The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society* 2004; **52**(12): 1601-1607; e-pub ahead of print 2004/11/24; doi 10.1369/jhc.4A6454.2004.
57. Grozdanov PN, Roy S, Kittur N, Meier UT. SHQ1 is required prior to NAF1 for assembly of H/ACA small nucleolar and telomerase RNPs. *RNA (New York, NY)* 2009; **15**(6): 1188-1197; e-pub ahead of print 2009/04/21; doi 10.1261/rna.1532109.
58. Darzacq X, Kittur N, Roy S, Shav-Tal Y, Singer RH, Meier UT. Stepwise RNP assembly at the site of H/ACA RNA transcription in human cells. *The Journal of cell biology* 2006; **173**(2): 207-218; e-pub ahead of print 2006/04/17; doi 10.1083/jcb.200601105.
59. Rocchi L, Barbosa AJ, Onofrillo C, Del Rio A, Montanaro L. Inhibition of human dyskerin as a new approach to target ribosome biogenesis. *PLoS One* 2014; **9**(7): e101971; e-pub ahead of print 2014/07/11; doi 10.1371/journal.pone.0101971.

Figure legends

Figure 1: *DKC1* mRNA expression and its association with copy number variations, clinicopathological parameters and molecular subtypes **(A)** *DKC1* and gene copy number variations **(B)** *DKC1* and patient age. **(C)** *DKC1* and tumour size. **(D)** *DKC1* and tumour grade **(E)** *DKC1* and Nottingham prognostic index. **(F)** *DKC1* and PAM50 BC subtypes **(G)** *DKC1* and SMCGENE subtypes in the METABRIC cohort using one-way analysis of variance with the post-hoc Tukey test.

Figure 2: **(A)** A representative Western blotting for DKC1 expression in cell lysate of MDA-MB-231 breast cancer cell line with the lanes (from left to right) of untransfected, transfected with scrambled siRNA and transfected with three different IDs ([ID1 represents ID s4110], [ID2 represents ID s4111] and [ID3 represents ID s4112]) of *DKC1* siRNA **(B)** Negative DKC1 IHC expression **(C)** Positive DKC1 IHC nuclear expression in invasive breast cancer TMA cores and **(D)** Positive DKC1 IHC nucleolar expression.

Figure 3: DKC1 protein expression and breast-cancer-specific survival (BCSS) **(A)** DKC1 nuclear expression and BCSS **(B)** DKC1 nucleolar expression and BCSS **(C)** combinatorial DKC1 protein expression and BCSS **(D)** DKC1 and BCSS in oestrogen receptor (ER) + low proliferation tumours **(E)** DKC1 and BCSS in (ER) + high proliferation tumours **(F)** DKC1 and BCSS human epidermal growth factor receptor 2 positive (HER2+) tumours **(G)** DKC1 and BCSS of triple negative tumours in the studied cohort.