Adverse prognostic and predictive significance of low DNA-dependent protein kinase catalytic subunit (DNA-PKcs) expression in early stage breast cancers

Tarek Abdel-Fatah[1], Arvind Arora[2], Devika Agarwal[3], Paul Moseley[1], Christina Perry[2], Nicola Thompson[1], Andrew Green[4], Emad Rakha[4], Stephen Chan[1], Carlos Caldas[5], Graham Ball[3], Ian O Ellis[4] and Srinivasan Madhusudan [1, 2]*.

[1]Department of Oncology, Nottingham University Hospitals, Nottingham NG51PB, UK.

[2]Academic Unit of Oncology, Division of Cancer and Stem Cells, School of Medicine, University of Nottingham, Nottingham University Hospitals, Nottingham NG51PB, UK.

[3] School of Science and Technology, Nottingham Trent University, Clifton Campus, Nottingham NG11 8NS, UK.

[4]Department of Pathology, Division of Cancer and Stem Cells, School of Medicine, University of Nottingham, Nottingham University Hospitals, Nottingham NG51PB, UK.

[5]Cancer Research UK, Cambridge Research Institute, Li Ka Shing Centre, Robinson Way, Cambridge CB2 0RE, UK.

Running title: DNA-PKcs in breast cancer

Disclosure: The authors have declared no conflicts of interest.

Word count:

* Corresponding author:

Academic Unit of Oncology

Division of Cancer and Stem Cells

School of Medicine

University of Nottingham

Nottingham University Hospitals

Nottingham NG51PB

UK

Telephone: +44(0)115 823 1850

Fax: +44(0)115 823 1849

E-Mail: srinivasan.madhusudan@nottingham.ac.uk

ABSTRACT

Background: DNA-dependent protein kinase catalytic subunit (DNA-PKcs), a serine threonine kinase belonging to the PIKK family (phosphoinositide 3-kinase-like-family of protein kinase), is a critical component of the non-homologous end joining (NHEJ) pathway required for the repair of DNA double strand breaks. DNA-PKcs may be involved in breast cancer pathogenesis.

Methods: We evaluated clinicopathological significance of DNA-PKcs protein expression in 1161 tumours and *DNA-PKcs* mRNA expression in 1950 tumours. We correlated DNA-PKcs to other markers of aggressive phenotypes, DNA repair, apoptosis and cell cycle regulation.

Results: Low DNA-PKcs protein expression was associated with higher tumour grade, higher mitotic index, tumour de-differentiation and tumour type (ps<0.05). Absence of BRCA1, low XRCC1, low SMUG1, low APE1 and low Polβ were also more likely in low DNA-PKcs expressing tumours (ps<0.05). Low DNA-PKcs protein expression was significantly associated with worse breast cancer specific survival (BCCS) in univariate and multivariate analysis (ps<0.01). At the mRNA level, low *DNA-PKcs* was associated with PAM50.Her2 and PAM50.LumA molecular phenotypes (ps<0.01) and poor BCSS. In patients with ER positive tumours who received endocrine therapy, low DNA-PKcs (protein and mRNA) was associated with poor survival. In ER negative patients, low *DNA-PKcs* mRNA remains significantly associated with adverse outcome.

Conclusions: Our study suggests that low DNA-PKcs expression may have prognostic and predictive significance in breast cancers.

INTRODUCTION

DNA-dependent protein kinase catalytic subunit (DNA-PKcs), a serine threonine kinase belonging to the PIKK family (phosphoinositide 3-kinase-like-family of protein kinase), is an essential component of the non-homologous end joining (NHEJ) DNA repair pathway {Kong, 2011 #219;Neal, 2011 #220;Hill, 2010 #221;Meek, 2008 #224;Salles, 2006 #225;Collis, 2005 #226}. NHEJ is central to the repair of DNA double strand breaks (DSBs) that may be generated throughout the cell cycle. DNA-PKcs is recruited to the sites of DSBs by Ku70/Ku80 heterodimer which then leads to its auto-phosphorylation at Ser2056, resulting in a conformational change that promote DSB repair function of DNA-PKcs in cells. In addition, phosphorylation at Thr2609 by ATM (ataxia telangiectasia mutated) or ATR (ataxia telangiectasia-Rad3-related) kinases also regulate biological functions of DNA-PKcs {Kong, 2011 #219;Neal, 2011 #220;Hill, 2010 #221;Meek, 2008 #224;Salles, 2006 #225;Collis, 2005 #226}. Besides an essential role in NHEJ, DNA-PKcs may be involved in cell cycle control, telomere maintenance, transcriptional regulation and p53 dependent apoptosis in cells {Kong, 2011 #219;Neal, 2011 #220;Hill, 2010 #221;Meek, 2008 #224;Salles, 2006 #225;Collis, 2005 #226}.

DNA-PKcs deficiency results in severe combined immunodeficiency (SCID) in mammals {van der Burg, 2009 #218}. The role of DNA-PKcs in carcinogenesis, however, is less well characterised. Single nucleotide polymorphisms in the *DNA-PKcs* gene may be associated with increased risk of bladder {Wang, 2008 #239} and hepatocellular cancer {Long, 2011 #240} development. In lung cancers, similarly, polymorphisms in the *DNA-PKcs* gene appear to confer altered cancer susceptibility {Hu, 2008 #242}. Somatic mutations in the *DNA-PKcs*

gene have been described in breast and pancreatic cancers {Wang, 2008 #248}. DNA-PKcs activity appears to be lower in patients with cancer compared to normal controls, implying a role for impaired DNA-PKcs in cancer risk and development. Interestingly, in established human tumours, the situation is far more complex. In tumours such as nasopharyngeal carcinoma, oesophageal cancer, colorectal cancer and lymphoid malignancies DNA-PKcs is overexpressed and associated with poor prognosis. On the other hand, in gastric and ovarian cancers, low expression of DNA-PKcs is linked to adverse outcomes in patients (reviewed in {Hsu, 2012 #253}). A large body of pre-clinical cancer cell lines or xenograft models based studies suggests that DNA-PKcs down-regulation or inhibition by small molecule inhibitors is associated with sensitization to chemotherapy and radiotherapy {Hsu, 2012 #253;Zhuang, 2011 #254;Du, 2010 #256;Shinohara, 2005 #257}.

The clinicopathological significance of DNA-PKcs expression in breast cancer is unknown. Our hypothesis is that altered DNA-PKcs expression may be associated with aggressive tumour biology and adversely impact upon clinical outcomes. In the current study we have investigated DNA-PKcs protein as well as mRNA expression in large cohorts of primary breast cancers. We demonstrate that low DNA-PKcs expression is associated with an aggressive phenotype and poor clinical outcome in patients.

METHODS

Patients: The study was performed in a consecutive series of 1650 patients with primary invasive breast carcinoma who were diagnosed from 1986 to 1999 and entered into the Nottingham Tenovus Primary Breast Carcinoma series. Patient demographics are summarised in supplementary table S1. This is a well-characterized series of patients with long-term

follow-up that have been investigated in a wide range of biomarker studies. All patients were treated in a uniform way in a single institution with standard surgery (mastectomy or wide local excision) with radiotherapy. Prior to 1989, patients did not receive systemic adjuvant treatment (AT). After 1989, AT was scheduled based on prognostic and predictive factor status, including Nottingham Prognostic Index (NPI), oestrogen receptor-α (ER-α) status, and menopausal status. Patients with NPI scores of <3.4 (low risk) did not receive AT. In pre-menopausal patients with NPI scores of ≥3.4 (high risk), classical Cyclophosphamide, Methotrexate, and 5-Flurouracil (CMF) chemotherapy was given; patients with ER-α positive tumours were also offered endocrine therapy. Postmenopausal patients with NPI scores of ≥3.4 and ER-α positivity were offered endocrine therapy, while ER-α negative patients received classical CMF chemotherapy.

Median follow up was 111 months (range 1 to 233 months). Survival data, including overall survival, disease-free survival (DFS), and development of loco-regional and distant metastases (DM), was maintained on a prospective basis. DFS was defined as the number of months from diagnosis to the occurrence of local recurrence, local lymph node (LN) relapse or DM relapse. Breast cancer specific survival (BCSS) was defined as the number of months from diagnosis to the occurrence of BC related-death. Local recurrence free survival (LRS) was defined the number of months from diagnosis to the occurrence of local recurrence. DM-free survival was defined as the number of months from diagnosis to the occurrence of DM relapse. Survival was censored if the patient was still alive at the time of analysis, lost to follow-up, or died from other causes.

Tumor Marker Prognostic Studies (REMARK) criteria, recommended by McShane et al{McShane, 2005 #261}, were followed throughout this study. Ethical approval was obtained from the Nottingham Research Ethics Committee (C202313).

Construction of Tissue Microarray (TMA): TMAs were constructed. Area-specialised histopathologists identified and marked formalin-fixed paraffin-embedded tissue blocks containing tumour tissue on haematoxylin and eosin stained slides. Two replicate 0.6mm cores from the centre and periphery of the tumours were taken and arrayed into a recipient paraffin block using a tissue puncher/arrayer (Beecher Instruments, Silver Spring, MD, USA). Four micron sections of the tissue array block were cut and placed on Superfrost Plus slides for immunohistochemical staining.

Immunohistochemistry (IHC): The TMAs were immunohistochemically profiled for DNA-PKcs and other biological antibodies (Supplementary Table S2) as previously described {Abdel-Fatah, 2014 #269;Abdel-Fatah, 2014 #270;Abdel-Fatah, 2013 #272}. Immunohistochemical staining was performed using the Thermo Scientific Shandon Sequenza chamber system (REF: 72110017), in combination with the Novolink Max Polymer Detection System (RE7280-K: 1250 tests), and the Leica Bond Primary Antibody Diluent (AR9352), each used according to the manufacturer's instructions (Leica Microsystems). The tissue slides were deparaffinised with xylene and then rehydrated through five decreasing concentrations of alcohol (100%, 90%, 70%, 50% and 30%) for two minutes each. Pre-treatment antigen retrieval was performed on the TMA sections using sodium citrate buffer (pH 6.0), heated for 20 minutes at 95^0C in a microwave (Whirpool JT359 Jet Chef 1000W). The slides were incubated for one hour with the primary mouse monoclonal anti-DNA PKcs antibody clone 3H6 (ab110034, Abcam, Cambridge, UK) at a dilution of 1:1000.

To validate the use of TMAs for immunophenotyping, full-face sections of 40 cases were stained and protein expression levels of DNA-PKcs antibodies were compared. The concordance between TMAs and full-face sections was excellent ($k = 0.8$). Positive and negative (by omission of the primary antibody and IgG-matched serum) controls were included in each run.

Evaluation of immunohistochemical staining: The tumour cores were evaluated by specialist pathologists and oncologists blinded to the clinico-pathological characteristics of patients. Whole field inspection of the core was scored and intensities of nuclear staining were grouped as follows: 0 = no staining, 1 = weak staining, 2 = moderate staining, 3 = strong staining. The percentage of each category was estimated (0-100%). H-score (range 0-300) was calculated by multiplying intensity of staining and percentage staining. H-score in a range of 0 – 300 was generated. The median H-score of 100 was taken as the cut-off and low DNA-PKcs expression was classed as H-score of ≤ 200 and >200 was classed as high for DNA-PKcs expression. Not all cores within the TMA were suitable for IHC analysis as some cores were missing or lacked tumour. HER2 expression was assessed according to the new ASCO/CAP guidelines using IHC and fluorescnce in situ hybridisation (FISH) {Wolff, 2013 #395}.

Breast cancer cell lines and Western blot analysis: MCF-7, MDA-MB-231 MDA-MB-436 breast cancer cell lines were purchased from ATCC and grown in RPMI medium supplemented with 10% FBS and 1% penicillin/streptomycin. Western blot analysis was performed as described previously {Sultana, 2013 #276}.

DNA-PKcs **mRNA expression:** The METABRIC (Molecular Taxonomy of Breast Cancer International Consortium) cohort comprising a cohort of 1980 breast cancer samples (Curtis et al.) was evaluated for *DNA-PKcs* mRNA expression. Patient demographics are summarized in supplementary Table S3 of supporting information. ER positive and/or lymphnode negative patients did not receive adjuvant chemotherapy. ER negative and/or lymphnode positive patients received adjuvant chemotherapy. RNA was extracted from fresh frozen tumours and subjected to transcriptional profiling on the Illumina HT-12 v3 platform. The data was pre-processed and normalized as described previously {Curtis, 2012 #404}. *DNA-PKcs* expression was investigated in this data set. The Chi-square test was used for

testing association between categorical variables was fitted to the data using as endpoint breast cancer specific death. X-tile (version 3.6.1, Yale University, USA) was used to identify a cut-off in gene expression values to divide the population in to high/low subgroups prior to Kaplan Meier Survival analysis. Kaplan Meier curves were constructed in SPSS Statistics (Version 21) for DNA-PKcs (probe-ILMN_2253648) and the log score calculated.

Statistical analyses: Data analysis was performed using SPSS (SPSS, version 17 Chicago, IL). Where appropriate, Pearson's Chi-square, Fisher's exact, χ^2 for trend, Student's t and ANOVAs one way tests were performed using SPSS software (SPSS, version16 Chicago, IL). Cumulative survival probabilities were estimated using the Kaplan–Meier method. Differences between survival rates were tested for significance using the log-rank test. Multivariate analysis for survival was performed using the Cox hazard model. The proportional hazards assumption was tested using standard log-log plots. Each variable was assessed in univariate analysis as a continuous and categorical variable and the two models were compared using an appropriate likelihood ratio test. Hazard ratios (HR) and 95% confidence intervals (95% CI) were estimated for each variable. All tests were two-sided with a 95% CI. P values for each test were adjusted with Benjamini and Hochberg multiple P-value adjustment and an adjusted p value of <0.05 was considered significant.

RESULTS

DNA-PKcs protein expression in breast cancer (Whole cohort): We first investigated the expression of DNA-PKcs in a panel of breast cancer cell lines. As shown in Figure 1A, robust expression of DNA-PKcs was evident in MCF-7, MDA-MB-231 and MDA-MB-436 breast cancer cells. We then proceeded to immunohistochemical investigation.

A total of 1161 tumours were suitable for DNA-PKcs protein expression analysis (Table 1). 404/1161 (34.8%) had low DNA-PKcs expression compared to 757/1161 (65.2%) tumours that had high DNA-PKcs expression (Figure 1A). Low DNA-PKcs expression was associated with higher tumour grade (p=0.001), higher mitotic index (p<0.000001), tumour de-differentiation (p=0.008) and tumour type (p=0.031) (Table 1). ER negative tumours were more likely to have low DNA-PKcs expression (p=0.021), as were AR negative tumours (p=0.00001). DNA-PKcs expression was also associated with cytokeratin 18 (CK18) (p=0.007) and cytokeratin 19 (CK19) (p=0.002). Low p21 and low TOPO2A and low p21 was likely in low DNA-PKcs expressing tumours (p<0.05). Absence of BRCA1, low XRCC1, low SMUG1, low APE1, low FEN1 and low Polβ were also more likely associated with low DNA-PKcs expression in tumours (p<0.05).

Low DNA-PKcs expression was significantly associated with worse breast cancer specific survival (BCCS) (p=006) and disease free survival (p<0.0001) (Figure 1B & C). On multivariate cox regression analysis, low DNA-PKcs remains independently associated with BCSS (p<0.001) and DFS (p<0.001), alongside tumour stage, grade and size (Table 4).

Taken together the data suggests that low DNA-PKcs may be associated with aggressive breast cancer. We then proceed to conduct sub-group analysis in ER positive, as well as in ER negative, breast cancers.

DNA-PKcs protein expression in ER positive breast cancer: A total of 835 tumours were suitable for analysis. 272/835 of patients (32.5%) with ER positive breast cancer were classified as having low DNA-PKcs expression. Univariate associations between APE1

expression and clinicopathological variables are shown in Table 2. Low DNA-PKcs expression was associated higher tumour grade (p=0.021), higher mitotic index (p=0.013), tumour de-differentiation (p=0.011) and tumour type (p=0.018). AR- tumours were more likely to have low DNA-PKcs expression (p=0.001). Low DNA-PKcs expression was also associated low TOP2A (p=0.002). DNA repair protein levels associated with low DNA-PKcs expression included absent BRCA1 (p=0.045), XRCC1 (p<0.0001), SMUG1 (p<0.0001), APE1 (p<0.0001) and polβ (p<0.0001) (Table 2). In the ER positive cohort, low DNA-PKcs expression was significantly associated with worse breast cancer specific survival (BCCS) (p=027) (Figure 1C) and disease free survival (p<0.05) (Figure 1D). In ER+/NPI >3.4 high risk breast cancer who received endocrine therapy, low DNA-PKcs expression remains significantly associated with worse breast cancer specific survival (BCCS) (p=048). On the other hand in ER+/NPI >3.4 high risk breast cancer who did not receive endocrine therapy, low DNA-PKcs did not influence BCSS although remained significant for DFS (p=0.044) (Figure 2A and 2B).

DNA-PKcs protein expression in ER negative breast cancer: A total of 311 tumours were suitable for analysis. 122/311 of patients (39.2%) with ER negative breast cancer were classified as having low DNA-PKcs expression. Univariate associations between DNA-PKcs expression and clinicopathological variables are shown in Table 3. Low DNA-PKcs expression was associated higher mitotic index (p=0.002), AR− (p=0.038), low XRCC1 (P=0.003), low APE1 (p<0.0001), low polβ (p=0.017) and low p21 (p=0.01). Low DNA-Pkcs did not influence survival in ER negative tumours (Figure 2C and 2D).

***DNA-PKcs mRNA* expression in breast cancer**: A total of 1975 tumours were suitable for *DNA-PKcs mRNA* expression analysis.   1308/1975 (66.2%) tumours had low *DNA-PKcs mRNA* expression compared to 667/1975 (33.8%) tumours that had high *DNA-PKcs mRNA* expression. As shown in Table 5, low *DNA-PKcs mRNA* was significantly associated with PAM50.Her2 and PAM50.Lum A molecular phenotypes (p<0.01). In contrast high *DNA-PKcs mRNA* was significantly associated with PAM50.basal and PAM50.Lum B (p<0.01).

In ER positive tumours (n=1496), low *DNA-PKcs mRNA* was associated with poor BCSS although with borderline significance (p=0.06). However, in ER positive patients who received adjuvant endocrine therapy (n=1216), low *DNA-PKcs mRNA* was significantly associated with poor BCSS (p=0.009). In ER negative tumours (n=437), low *DNA-PKcs mRNA* was significantly associated with poor BCSS (p=0.041). In addition, in ER negative patients who received adjuvant chemotherapy (n=416), low *DNA-PKcs mRNA* remains significantly associated with poor BCSS (p=0.009)

Taken together the protein and mRNA expression data provides evidence that DNA-PKcs may have prognostic and predictive significance in breast cancer patients.

DISCUSSION

This is the first comprehensive study of DNA-PKcs expression in breast cancer. We found that low protein expression of DNA-PKcs was associated with adverse clinicopathological features including higher tumour grade, higher mitotic index, poorly differentiated and triple negative disease implying that impaired DNA-PKcs may be involved in breast cancer pathogenesis. Low DNA-PKcs was also linked to poor survival in univariate, as well as in

multivariate analysis, suggesting prognostic significance in patients. Interestingly, in ER positive tumours the prognostic significance was retained. Moreover, in patients who received endocrine therapy, low DNA-PKcs/ ER positive tumours were associated with poor breast cancer specific and disease free survival compared to high DNA-PKcs/ ER positive tumours that received endocrine therapy. Similarly at the mRNA level, low *DNA-PKcs* mRNA expression was associated with poor survival in patients who received endocrine therapy. This new observation suggests that DNA-PKcs may predict response to endocrine therapy. Pre-clinically, the link between ER and DNA-PKcs has been reported {Medunjanin, 2010 #415}. DNA-PKcs was shown to phosphorylate and stabilise ER and promote ER dependent transcriptional activity. Moreover, oestrogen was also shown to induce DNA-PKcs expression in breast cancer cells in that study. Together the authors suggested that reduced DNA-PKcs may be expected to attenuate oestrogen/ER induced proliferation signalling in breast cancer cells {Medunjanin, 2010 #415}. However, the data presented in our study contradicts the previous preclinical observation in that low DNA-PKcs was associated with aggressive breast cancer in our study implying that further mechanistic studies are required to investigate fully the functional link between DNA-PKcs and ER in breast cancer. Another possibility for the observed aggressive phenotype in human tumours is that low DNA-PKcs may be associated with, or reflect, genomic instability in cells. This is supported by the finding that low DNA-PKcs was significantly associated with low expression of several DNA repair proteins including BRCA1, XRCC1, SMUG1, APE1 and Polβ. Impaired DNA repair and the associated genomic instability may promote accelerated accumulation of mutations, resulting in a 'mutator phenotype' characterised by aggressive clinical behaviour {Loeb, 2010 #428;Loeb, 2008 #444}. The clinicopathological associations seen in breast cancer in our study also concur with a study of DNA-PKcs expression in 564 gastric cancers {Lee, 2007 #118;Lee, 2005 #149}. Negative DNA-PKcs was found in 20.2% of tumours and

associated with aggressive tumour and poor survival. Interestingly, frameshift mutation of (A)$_{10}$ mononucleotide repeats in the *DNA-PKcs* gene was demonstrated in 24.3% of gastric cancers which was associated with higher risk of lymph node metastasis in that study {Lee, 2007 #118;Lee, 2005 #149}. Whether a similar mechanism also operates in breast cancer is currently unknown and could be an area for future investigation. Moreover, a previous study conducted extensive profiling of DNA-PKcs mRNA and protein levels across normal and cancerous tissues {Moll, 1999 #567}. The authors demonstrated low DNA-PKcs protein was consistently low in breast cancer but no clinicopathological association data was reported. In normal tissues although there no significant changes at mRNA level, considerable variation were evident at the protein level in that study implying that DNA-PKcs protein expression was likely regulated at the post-transcriptional or post- translational level in that study{Moll, 1999 #567}. Although the data presented in our study suggest that DNA-PKcs may regulated at the transcriptional level in breast cancer, whether additional mechanisms also operate is unknown. High resolution deep sequencing of the DNA-PKcs gene in a cohort of breast cancer could provide further insight. In ovarian cancers, similarly, low DNA-PKcs was associated with stage and lymph node metastasis {Shao, 2007 #568}. In contrast, in nasopharyngeal carcinoma, oesophageal cancer, colorectal cancer and lymphoid malignancies DNA-PKcs overexpression was associated with poor prognosis {Hsu, 2012 #572}. Taken together, the data suggest a complex role for DNA-Pkcs in different solid tumours.

A surprising finding in the current study was that low *DNA-PKcs* mRNA in ER negative tumours (including in patients who received adjuvant chemotherapy) was associated with poor survival and at the protein level, although there was trend, it did not reach significance. This is in contrast to several pre-clinical reports that provide evidence that DNA-PKcs gene knock down or inhibition by small molecule inhibitors result in increased sensitivity to

chemotherapy and radiotherapy {Hsu, 2012 #253;Zhuang, 2011 #254;Du, 2010 #256;Shinohara, 2005 #257}. Although intriguing, the data concurs with a recent report in a small cohort of breast cancers where high DNA-PKcs expression was associated with significant benefit from radiotherapy response in breast cancer patients {Soderlund Leifler, 2010 #573}. The authors concluded that the role of proficient DNA-PKcs in p53 dependent pro-apoptotic response to cytotoxic therapy may have contributed to the observed effect{Soderlund Leifler, 2010 #573}. However, in our study we did not observe any associations with p53 expression in tumours. Further mechanistic studies are required to clarify the role of DNA-PKcs in chemotherapy and radiotherapy response in human tumours.

In conclusion we have shown that altered DNA-PKcs expression may influence pathogenesis and have prognostic and predictive significance in breast cancer.

Table 1. DNA-PKcs protein expression in breast cancer (whole cohort).

VARIABLE	DNA-PKcs PROTEIN EXPRESSION		P- value
	Low n=404	High n= 757	
A) Pathological Parameters			
Tumour Size			
<1cm	32 (7.8%)	80 (10.5%)	0.268
>1-2cm	198 (48.5%)	380 (50.1%)	
>2-5cm	166 (40.7%)	284 (37.4%)	
>5cm	12 (2.9%)	15 (2.0%)	
Tumour Stage			
1	262 (63.9%)	450 (59.2%)	0.130
2	118 (28.8%)	230 (30.3%)	
3	30 (7.3%)	80 (10.5%)	
Tumour Grade			
G1	57 (14.0%)	120 (15.8%)	**0.001**
G2	108 (26.5%)	271 (35.7%)	
G3	243 (59.6%)	368 (48.5%)	
Mitotic Index			
M1 (low; mitoses < 10)	110 (27.2%)	275 (36.4%)	**9.0×10^{-6}**
M2 (medium; mitoses 10-18)	61 (15.1%)	157 (20.8%)	
M3 (high; mitosis >18)	233 (57.7%)	324 (42.9%)	
Tubule Formation			
1 (>75% of definite tubule)	16 (4.0%)	40 (5.3%)	**0.008**
2 (10%-75% definite tubule)	111 (27.5%)	268 (35.4%)	
3 (<10% definite tubule)	277 (68.6%)	448 (59.3%)	
Pleomorphism			
1 (small-regular uniform)	6 (1.5%)	20 (2.6%)	0.403
2 (Moderate variation)	145 (35.9%)	278 (36.8%)	
3 (Marked variation)	253 (62.6%)	458 (60.6%)	
Tumour Type			
IDC-NST	232 (64.3%)	382 (54.8%)	**0.031**
Tubular Carcinoma	54 (15.0%)	136 (20.8%)	
Medullary Carcinoma	12 (3.3%)	12 (1.8%)	
ILC	25 (6.9%)	65 (9.9%)	
Others	38 (10.5%)	59 (9.0%)	
Lymphovascular Invasion			
No	270 (66.5%)	489 (65.0%)	0.614
Yes	136 (33.5%)	263 (35.0%)	
B) Aggressive phenotype			

Her2 overexpression No Yes	351 (88.2%) 47 (11.8%)	657 (88.2%) 88 (11.8%)	0.999
Triple Negative Phenotype No Yes	305 (77.8%) 87 (22.2%)	614 (82.5%) 130 (17.5%)	0.054
Basal Like Phenotype No Yes	329 (85.(%) 54 (14.1%)	637 (87.5%) 91 (12.5%)	0.452
Cytokeratin 6 (CK6) Negative Positive	298 (83.9%) 57 (16.1%)	529 (83.7%) 103 (16.3%)	0.921
Cytokeratin 14 (CK14) Negative Positive	297 (84.1%) 56 (15.9%)	552 (88.3%) 73 (11.7%)	0.063
Cytokeratin 18 (CK18) Negative Positive	45 (14.2%) 272 (85.8%)	50 (8.4%) 544 (91.6%)	**0.007**
Cytokeratin 19 (CK19) Negative Positive	32 (9.1%) 319 (90.9%)	27 (4.3%) 602 (95.7%)	**0.002**
C) Hormone receptors			
ER Negative Positive	127 (31.8%) 272 (68.2%)	192 (25.4%) 563 (74.6%)	**0.021**
PgR Negative Positive	176 (46.8%) 200 (53.2%)	294 (41.4%) 416 (58.6%)	0.087
AR Negative Positive	154 (47.4%) 171 (52.6%)	190 (31.6%) 412 (68.4%)	**0.0001**
D) DNA Repair			
BRCA1 Absent Normal	67 (24.3%) 209 (75.7%)	99 (18.1%) 447 (81.9%)	**0.038**

XRCC1 Low High	84 (27.4%) 223 (72.6%)	56 (10.1%) 497 (89.9%)	**$1.0x10^{-6}$**
FEN1 Low High	239 (80.5%) 58 (19.5%)	360 (69.6%) 157 (30.4%)	**0.001**
SMUG1 Low High	128 (46.7%) 146 (53.3%)	161 (32.4%) 336 (67.6%)	**$8.4x10^{-5}$**
APE1 Low High	261 (72.7%) 98 (27.3%)	264 (39.5%) 404 (60.5%)	**$1.0x10^{-6}$**
PolB Low High	201 (55.4%) 162 (44.6%)	196 (28.6%) 490 (71.4%)	**$1.0x10^{-6}$**
ATM Absent Present	138 (54.3%) 116 (45.7%)	246 (54.1%) 209 (45.9%)	0.946
E) Cell cycle/apoptosis regulators			
P16 Low High	247 (79.9%) 62 (20.1%)	460 (89.7%) 53 (10.3%)	**$9.8x10^{-5}$**
P21 Low High	198 (62.7%) 118 (37.3%)	287 (53.8%) 246 (46.2%)	**0.012**
MIB1 Low High	97 (29.7%) 230 (70.3%)	218 (34.1%) 422 (65.9%)	0.167
P53 Low expression High expression	256 (77.3%) 75 (22.7%)	490 (78.7%) 133 (21.3%)	0.641
Bcl-2 Negative Positive	233 (34.6%) 441 (65.4%)	378 (36.0%) 673 (64.0%)	0.207

TOP2A			
Low	156 (52.5%)	229 (41.6%)	**0.002**
Overexpression	141 (47.5%)	321 (58.4%)	

Bold= statistically significant; BRCA1: Breast cancer 1, early onset; HER2: human epidermal growth factor receptor 2; ER: oestrogen receptor; PgR: progesterone receptor; CK: cytokeratin; Basal-like: ER-, HER2- and positive expression of either CK5/6, CK14 or EGFR; Triple negative: ER-/PgR-/HER2-

Table 2. DNA-PKcs and ER positive breast cancer

VARIABLE	DNA-PKcs PROTEIN EXPRESSION		P- value
	Low n= 272	High n= 563	
A) Pathological Parameters			
Tumour Size <1cm >1-2cm >2-5cm >5cm	23 (8.5%) 137 (50.7%) 105 (38.9%) 5 (1.9%)	65 (11.6%) 293 (52.1%) 195 (34.7%) 9 (1.6%)	0.453
Tumour Stage 1 2 3	167 (61.4%) 87 (32.0%) 18 (6.6%)	338 (60.1%) 173 (30.8%) 51 (9.1%)	0.479
Tumour Grade G1 G2 G3	51 (18.9%) 96 (35.6%) 123 (45.6%)	111 (19.8%) 249 (44.3%) 202 (35.9%)	**0.021**
Mitotic Index M1 (low; mitoses < 10) M2 (medium; mitoses 10-18) M3 (high; mitosis >18)	104 (38.7%) 53 (19.7%) 112 (41.6%)	255 (45.5%) 131 (23.4%) 175 (31.2%)	**0.013**
Tubule Formation 1 (>75% of definite tubule) 2 (10%-75% definite tubule) 3 (<10% definite tubule)	14 (5.2%) 87 (32.3%) 168 (62.5%)	36 (6.4%) 237 (42.2%) 288 (51.3%)	**0.011**
Pleomorphism 1 (small-regular uniform) 2 (Moderate variation) 3 (Marked variation)	6 (2.2%) 131 (48.7%) 132 (49.1%)	17 (3.0%) 261 (46.5%) 283 (50.4%)	0.715
Tumour Type IDC-NST Tubular Carcinoma Medullary Carcinoma ILC Others	142 (56.8%) 47 (18.8%) 3 (1.2%) 22 (8.8%) 36 (14.4%)	248 (50.4%) 131 (26.6%) 1 (0.2%) 59 (12.0%) 53 (10.8%)	**0.018**
Lymphovascular Invasion No Yes	177 (65.6%) 93 (34.4%)	362 (64.9%) 196 (35.1%)	0.847
B) Aggressive phenotype			

Her2 overexpression No Yes	249 (93.6%) 17 (6.4%)	511 (92.6%) 41 (7.4%)	0.588
Cytokeratin 6 (CK6) Negative Positive	229 (94.2%) 14 (5.8%)	454 (94.4%) 27 (5.6%)	0.935
Cytokeratin 14 (CK14) Negative Positive	220 (90.9%) 22 (9.1%)	444 (93.5%) 31 (6.5%)	0.215
Cytokeratin 18 (CK18) Negative Positive	10 (4.5%) 214 (95.5%)	8 (1.8%) 444 (98.2%)	**0.041**
Cytokeratin 19 (CK19) Negative Positive	11 (4.6%) 229 (95.4%)	16 (3.3%) 462 (96.7%)	0.411
C) Hormone receptors			
PgR Negative Positive	59 (23.1%) 196 (76.9%)	117 (22.0%) 415 (78.0%)	0.718
AR Negative Positive	74 (32.9%) 151 (67.1%)	96 (20.8%) 366 (79.2%)	**0.001**
D) DNA Repair			
BRCA1 Absent Normal	38 (19.3%) 159 (80.7%)	54 (13.1%) 359 (86.9%)	**0.045**
XRCC1 Low High	46 (22.3%) 160 (77.7%)	29 (7.0%) 388 (93.0%)	**1.0×10^{-6}**
FEN1 Low High	163 (81.1%) 38 (18.9%)	265 (68.1%) 124 (31.9%)	**0.001**
SMUG1 Low High	78 (42.9%) 104 (57.1%)	97 (26.4%) 271 (73.6%)	**9.3×10^{-5}**

APE1 Low High	169 (71.0%) 69 (29.0%)	186 (36.9%) 318 (63.1%)	**1.0×10^{-6}**
PolB Low High	122 (50.8%) 118 (49.2%)	115 (22.5%) 397 (77.5%)	**1.0×10^{-6}**
ATM Absent Present	84 (49.7%) 85 (50.3%)	174 (51.2%) 166 (48.8%)	0.754
E) Cell cycle/apoptosis regulators			
P16 Low High	190 (92.2%) 16 (7.8%)	368 (98.1%) 7 (1.9%)	**4.8×10^{-4}**
P21 Low High	114 (53.5%) 99 (46.5%)	193 (48.9%) 202 (51.1%)	0.273
MIB1 Low High	84 (38.0%) 137 (62.0%)	196 (40.8%) 284 (59.2%)	0.478
P53 Low expression High expression	202 (87.8%) 28 (12.2%)	412 (87.3%) 60 (12.7%)	0.840
Bcl-2 Negative Positive	59 (22.8%) 200 (77.2%)	110 (21.6%) 400 (78.4%)	0.701
TOP2A Low Overexpression	106 (55.5%) 85 (44.5%)	171 (41.9%) 237 (58.1%)	**0.002**

Bold= statistically significant; BRCA1: Breast cancer 1, early onset; HER2: human epidermal growth factor receptor 2; ER: oestrogen receptor; PgR: progesterone receptor; CK: cytokeratin; Basal-like: ER-, HER2- and positive expression of either CK5/6, CK14 or EGFR; Triple negative: ER-/PgR-/HER2-

Table 3. DNA-PKcs and ER negative breast cancer

VARIABLE	DNA-PK PROTEIN EXPRESSION		P- value
	Low n= 122	High n= 189	
A) Pathological Parameters			
Tumour Size <1cm >1-2cm >2-5cm >5cm	8 (6.3%) 55 (43.3%) 58 (45.7%) 6 (4.7%)	13 (6.8%) 83 (43.7%) 88 (46.3%) 6 (3.2%)	0.911
Tumour Stage 1 2 3	85 (66.9%) 30 (23.6%) 12 (9.4%)	107 (56.0%) 55 (28.8%) 29 (15.2%)	0.122
Tumour Grade G1 G2 G3	2 (1.6%) 8 (6.3%) 117 (92.1%)	6 (3.2%) 19 (10.0%) 165 (86.8%)	0.330
Mitotic Index M1 (low; mitoses < 10) M2 (medium; mitoses 10-18) M3 (high; mitosis >18)	3 (2.4%) 6 (4.8%) 117 (92.9%)	16 (8.5%) 25 (13.2%) 148 (78.3%)	**0.002**
Tubule Formation 1 (>75% of definite tubule) 2 (10%-75% definite tubule) 3 (<10% definite tubule)	1 (0.8%) 19 (15.1%) 106 (84.1%)	2 (1.1%) 30 (15.9%) 157 (83.1%)	0.953
Pleomorphism 1 (small-regular uniform) 2 (Moderate variation) 3 (Marked variation)	0 (0.0%) 9 (7.1%) 117 (92.9%)	1 (0.5%) 15 (7.9%) 173 (91.5%)	0.690
Tumour Type IDC-NST Tubular Carcinoma Medullary Carcinoma ILC Others	87 (82.1%) 6 (5.7%) 9 (8.5%) 3 (2.8%) 1 (0.9%)	132 (84.1%) 3 (1.9%) 11 (7.0%) 5 (3.2%) 6 (3.8%)	0.307

Lymphovascular Invasion No Yes	84 (66.7%) 42 (33.3%)	122 (65.2%) 65 (34.8%)	0.794
B) Aggressive phenotype			
Her2 overexpression No Yes	94 (75.8%) 30 (24.2%)	140 (74.9%) 47 (25.1%)	0.851
Triple Negative Phenotype No Yes	31 (26.3%) 87 (73.7%)	49 (27.4%) 130 (72.6%)	0.834
Basal Like Phenotype No Yes	55 (50.5%) 54 (49.5%)	73 (44.5%) 91 (55.5%)	0.335
Cytokeratin 6 (CK6) Negative Positive	66 (61.7%) 41 (38.3%)	73 (49.0%) 76 (51.0%)	**0.044**
Cytokeratin 14 (CK14) Negative Positive	75 (70.1%) 32 (29.9%)	107 (71.8%) 42 (28.2%)	0.765
Cytokeratin 18 (CK18) Negative Positive	34 (37.0%) 58 (63.0%)	42 (30.0%) 98 (70.0%)	0.269
Cytokeratin 19 (CK19) Negative Positive	21 (19.8%) 85 (80.2%)	11 (7.3%) 139 (92.7%)	**0.003**
C) Hormone receptors			
PgR Negative Positive	115 (98.3%) 2 (1.7%)	177 (100%) 0 (0.0%)	0.081
AR Negative Positive	79 (79.8%) 20 (20.2%)	94 (67.6%) 45 (32.4%)	**0.038**
D) DNA Repair			

BRCA1 Absent Normal	29 (36.7%) 50 (63.3%)	45 (34.1%) 87 (65.9%)	0.700
XRCC1 Low High	36 (38.3%) 58 (61.7%)	27 (20.5%) 105 (79.5%)	**0.003**
FEN1 Low High	69 (78.4%) 19 (21.6%)	94 (75.2%) 31 (24.8%)	0.586
SMUG1 Low High	48 (57.1%) 36 (42.9%)	64 (50.8%) 62 (49.2%)	0.366
APE1 Low High	84 (75.7%) 27 (24.3%)	75 (47.5%) 83 (52.5%)	**4.0×10^{-6}**
PolB Low High	69 (61.9%) 43 (38.4%)	79 (47.0%) 89 (53.0%)	**0.017**
ATM Absent Present	54 (64.3%) 30 (35.7%)	71 (62.3%) 43 (37.7%)	0.773
E) Cell cycle/apoptosis regulators			
P16 Low High	53 (54.6%) 44 (45.4%)	90(66.7%) 45 (33.3%)	0.063
P21 Low High	81 (82.7%) 17 (17.3%)	90 (67.7%) 43 (32.3%)	**0.010**
MIB1 Low High	11 (11.0%) 89 (89.0%)	22 (14.0%) 135 (86.0%)	0.482
P53 Low expression High expression	53 (53.0%) 47 (47.0%)	75 (50.7%) 73 (49.3%)	0.719

Bcl-2 Negative Positive	84 (75.7%) 27 (24.3%)	122 (76.7%) 37 (23.3%)	0.841
TOP2A Low Overexpression	46 (45.5%) 55 (54.5%)	58 (42.3%) 79 (57.7%)	0.622

Bold= statistically significant; BRCA1: Breast cancer 1, early onset; HER2: human epidermal growth factor receptor 2; ER: oestrogen receptor; PgR: progesterone receptor; CK: cytokeratin; Basal-like: ER-, HER2- and positive expression of either CK5/6, CK14 or EGFR; Triple negative: ER-/PgR-/HER2-

Table 4. Multivariate analysis in breast cancer.

	P Value	Exp (B)	95% CI of Exp(B)	
			Lower	Upper
Disease Free Survival				
DNA-PKcs (low protein)	**6.7×10^{-5}**	0.886	0.834	0.940
Stage	**1.0×10^{-6}**	1.728	1.515	1.971
Grade	**2.9×10^{-4}**	1.270	1.116	1.446
Size	**4.0×10^{-3}**	1.130	1.039	1.230
Breast Cancer Specific Survival				
DNA-PKcs (low protein)	**4.0×10^{-3}**	0.901	0.839	0.968
Stage	**1.0×10^{-6}**	1.936	1.661	2.256
Grade	**1.0×10^{-6}**	2.005	1.672	2.404
Size	**5.0×10^{-3}**	1.157	1.045	1.281

Table 5: Association between DNA-PKcs mRNA expression and molecular phenotypes.

PAM50 subtype	DNA-PKc mRNA Expression		X^2 Adjusted p value
	Low N=1308 (66.23%)	**High** N= 667 (33.77%)	
PAM50.Her2	166(14.70%)	72(11.20%)	**0.040**
PAM50.Basal	156(13.80%)	174(27.10%)	$\mathbf{4.25 \times 10^{-12}}$
PAM50.LumA	545(48.20%)	169(26.40%)	$\mathbf{1.98 \times 10^{-19}}$
PAM50.LumB	263(23.30%)	226(35.30%)	$\mathbf{5.94 \times 10^{-08}}$