Genomic and protein expression analysis reveals Flap endonuclease 1 (FEN1) as a key biomarker in breast and ovarian cancer

Tarek MA Abdel-Fatah[1], Roslin Russell[2,3], Nada Albarakati[4] , David J Maloney[5], Dorjbal Dorjsuren[5], Oscar M Rueda[2,3] , Paul Moseley[1] , Vivek Mohan [4] , Hongmao Sun[5], Rachel Abbotts[4] , Abhik Mukherjee[6] , Devika Agarwal[7] , Jennifer L. Illuzzi[8], Ajit Jadhav[5], Anton Simeonov[5], Graham Ball [7], Stephen Chan[1], Carlos Caldas[2,3], Ian O Ellis[6], David M Wilson III[8] and Srinivasan Madhusudan [1,4*]

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

[2] Department of Oncology, University of Cambridge, Hills Road, Cambridge CB2 2XZ, UK.

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

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

[5]NIH Chemical Genomics Center, National Center for Advancing Translational Sciences, National Institutes of Health, 9800 Medical Center Drive, Rockville, Maryland 20850, USA.

[6] Division of Pathology, School of Molecular Medical Sciences, University of Nottingham, Nottingham University Hospitals, Nottingham NG51PB, UK.

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

[8] Laboratory of Molecular Gerontology, Biomedical Research Center, National Institute on Aging, National Institutes of Health, Baltimore, Maryland 21224-6825, USA.

Running title: FEN1 in breast cancer

Key words: FEN1; breast cancer; prognostic factor;predictive factor; drug target

Conflict of interest: The authors disclose no potential conflicts of interest

Word count: 4563

Tables: 3, **Figures:** 3

*** Corresponding author:**

Dr Srinivasan Madhusudan

Academic Unit of Oncology

Division of Cancer and Stem Cells

School of Medicine

University of Nottingham

Nottingham University Hospitals

Nottingham NG51PB, U.K.

Telephone: +44(0)115 823 1850

Fax: +44(0)115 823 1849

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

ABSTRACT

FEN1 has key roles in Okazaki fragment maturation during replication, long patch base excision repair, rescue of stalled replication forks, maintenance of telomere stability and apoptosis. FEN1 may be dysregulated in breast and ovarian cancers and have clinicopathological significance in patients. We comprehensively investigated FEN1 mRNA expression in multiple cohorts of breast cancer [training set (128), test set (249), external validation (1952)]. FEN1 protein expression was evaluated in 568 oestrogen receptor (ER) negative breast cancers, 894 ER positive breast cancers and 156 ovarian epithelial cancers. FEN1 mRNA overexpression was highly significantly associated with high grade ($p=4.89 \times 10^{-57}$), high mitotic index ($p=5.25 \times 10^{-28}$), pleomorphism ($p=6.31 \times 10^{-19}$), ER negative ($p=9.02 \times 10^{-35}$), PR negative ($p=9.24 \times 10^{-24}$), triple negative phenotype ($p=6.67 \times 10^{-21}$), PAM50.Her2 ($p=5.19 \times 10^{-13}$), PAM50.Basal ($p=2.7 \times 10^{-41}$), PAM50.LumB ($p=1.56 \times 10^{-26}$), integrative molecular cluster 1 (intClust.1) ($p=7.47 \times 10^{-12}$), intClust.5 ($p=4.05 \times 10^{-12}$) and intClust. 10 ($p=7.59 \times 10^{-38}$) breast cancers. FEN1 mRNA overexpression is associated with poor breast cancer specific survival in univariate ($p= 4.4 \times 10^{-16}$) and multivariate analysis ($p= 9.19 \times 10^{-7}$). At the protein level, in ER positive tumours , FEN1 overexpression remains significantly linked to high grade, high mitotic index and pleomorphism (ps<0.01). In ER negative tumours, high FEN1 is significantly associated with pleomorphism, tumour type, lymphovascular invasion, triple negative phenotype, EGFR and HER2 expression (ps<0.05). In ER positive as well as in ER negative tumours, FEN1 protein overexpression is associated with poor survival in univariate and multivariate analysis (ps<0.01). In ovarian epithelial cancers, similarly, FEN1 overexpression is associated with high grade, high stage and poor survival (ps<0.05). We conclude that FEN1 is a promising biomarker in breast and ovarian epithelial cancer.

INTRODUCTION

The flap structure specific endonuclease (FEN1) is critical for processing DNA intermediates generated during DNA long patch base excision repair (LP-BER) and Okazaki fragment maturation during replication. FEN1 is also essential for rescue of stalled replication forks, maintenance of telomere stability and apoptotic fragmentation of DNA (Shen et al., 2005; Zheng et al., 2011). FEN1 belongs to XPG/RAD2 endonuclease family and *FEN1* gene is located at 11q22. FEN1 possesses flap endonuclease, 5' exonuclease and gap-endonuclease activities to accomplish its various biological functions. FEN1 is subjected to post-translational modifications such as acetylation, phosphorylation, sumoylation, methylation and ubiquitylation that regulate nuclease activities as well as protein-protein interactions and sub-cellular compartmentalization (Shen et al., 2005; Zheng et al., 2011).

FEN1 may have a role in carcinogenesis. A tumour suppressor function for FEN1 has been shown in preclinical models (Henneke et al., 2003a; Henneke et al., 2003b; Kucherlapati et al., 2007; Kucherlapati et al., 2002; Wu et al., 2012; Xu et al., 2011). Whereas, *FEN1* homologous knock out in mice is embryonically lethal *FEN1* heterozygous mice are viable (Larsen et al., 2003). A double heterozygous mouse model with a mutation in *FEN1* and adenomatous polyposis coli *(APC)* gene displayed enhanced cancer development and poor survival (Kucherlapati et al., 2007). In addition, a FEN1 E160D mutant mouse model displayed altered DNA repair as well as apoptotic DNA fragmentation and associated with increased mutation frequency and cancer development (Larsen et al., 2008; Zheng et al., 2007). In human studies, polymorphic variants of FEN1 may be associated with increased cancer susceptibility (Liu et al., 2012; Yang et al., 2009). In established tumours, preclinical evidence suggests that FEN1 over expression may promote cancer progression and survival

(Kim, 1998; Kim et al., 2005; Krause et al., 2005; Sato et al., 2003). Proliferating cells consistently over express FEN1 compared to quiescent cells (Kim, 1998). In pro-myelocytic leukemia cells (HL-60), *FEN1* gene expression was shown to be higher during mitotic phase compared to the resting phase of the cell cycle and *FEN1* expression markedly decreased upon induction of terminal differentiation in cells (Kim, 1998). *FEN1* mRNA over expression has also been demonstrated in lung cancer cell lines (Sato et al., 2003) and gastric cancer cell lines (Kim et al., 2005). In human tumours, frequent overexpression of FEN1 has been reported (Singh et al., 2008). In a small cohort of 50 breast tumours, FEN1 was shown to be upregulated in tumours compared to normal tissue in that study (Singh et al., 2008). However, clinicopathological significance of FEN1 upregulation remains unknown in breast and ovarian cancer (Singh et al., 2008).

We hypothesised that FEN1 may be dysregulated in human breast and ovarian cancer, contributing to the aetiology of the disease. We investigated *FEN1* mRNA as well as FEN1 protein expression in large cohorts of breast and ovarian tumours and correlated to clinicopathological variables and outcome data. In the current study we demonstrate that FEN1 overexpression is associated with aggressive phenotype and poor survival in breast and ovarian cancer. The data provides evidence that FEN1 is a promising biomarker.

MATERIALS AND METHODS

FEN1 gene expression (training set): The study population used was derived from the Nottingham Tenovus Primary Breast Carcinoma Series of women aged 70 years or less, who presented with stage I and II primary operable invasive breast carcinomas. The patient demographics for the training set are summarized in supplementary table 1 of supporting information. Gene expression profiling has been previously described (Chin et al., 2007). Briefly, total RNA was extracted from a series of frozen breast cancers retrieved from Nottingham Hospitals NHS Trust Tumour Bank between 1986 and 1992. RNA integrity and DNA contamination were analysed using Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA). Total RNA was biotin-labelled using the Illumina TotalPrep RNA Amplification kit (Ambion, Austin, TX, USA) according to manufacturer's instructions. Biotin-labelled cRNA (1.5 µg) was used for each hybridisation on Sentrix Human-6 BeadChips (Illumina, San Diego, CA, USA) in accordance with the manufacturer's protocol. Illumina gene expression data containing 47,293 transcripts were analysed and summarised in the Illumina Bead Studio software. Analyses of the probe level data were done using the beadarray Bioconductor package. The expression data are available at the EBI website (http://www.ebi.ac.uk/miamexpress/) with the accession number E-TABM-576.

FEN1 gene expression (Test Set): The Uppsala cohort originally composed of 315 women representing 65% of all breast cancers resected in Uppsala County, Sweden, from January 1, 1987, to December 31, 1989. Demographics are summarized in supplementary Table S2 of supporting information and also described elsewhere (Bergh et al., 1995). Tumour samples were microarray profiled on the Affymetrix U133A&B genechips. Microarray analysis was carried out at the Genome Institute of Singapore. All microarray data are accessible at

National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/). Data can be accessed via series accession number (GSE4922). RNA preparation, microarray hybridization, and data processing were carried out essentially as described (Pawitan et al., 2005). All data were normalized using the global mean method (MAS5), and probe set signal intensities were natural log transformed and scaled by adjusting the mean signal to a target value of log 500.

FEN1 gene expression (external validation): External validation was performed in the METABRIC (Molecular Taxonomy of Breast Cancer International Consortium) cohort. The METABRIC study protocol, detailing the molecular profiling methodology in a cohort of 1980 breast cancer samples is described by Curtis et al (Curtis et al., 2012). 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 et al.). FEN1 expression was investigated in this data set. The Chi-square test was used for testing association between categorical variables and a multivariate Cox model was fitted to the data using as endpoint breast cancer specific death. Recursive partitioning (Hothorn et al., 2006) was used to identify a cut-off in gene expression values such that the resulting subgroups have significantly different survival courses.

FEN1 protein expression in breast cancer: The study was performed in a consecutive series of 1650 patients with primary invasive breast carcinomas who were diagnosed between 1986 and 1999 and entered into the Nottingham Tenovus Primary Breast Carcinoma series.

All patients were treated uniformly in a single institution and have been investigated in a wide range of biomarker studies (Abdel-Fatah et al., 2013b; Abdel-Fatah et al., 2013c; Abdel-Fatah et al., 2014a; Abdel-Fatah et al., 2014b; Sultana et al., 2013). Clinicopathological characteristics of ER negative cohort is summarized in Supplementary Table S8 of supporting information. Patient demographics for ER positive cohort are summarized in Supplementary Table S9 of supporting information. Supplemental treatment data 1 summarizes various adjuvant treatments received by patients in Nottingham Tenovus Primary Breast Carcinoma series.

FEN1 protein expression in ovarian epithelial cancer: Investigation of the expression of FEN1 in ovarian epithelial cancer was carried out on a tissue microarray of 195 consecutive ovarian epithelial cancer cases treated at Nottingham University Hospitals (NUH) between 2000 and 2007. Patients were comprehensively staged as per International Federation of Obstetricians and Gynaecologists (FIGO) Staging System for Ovarian Cancer. Survival was calculated from the operation date until 1st of October 2012 when any remaining survivors were censored. Patient demographics are summarised in summarized in Supplementary Table S14 of supporting information. Platinum resistance was defined as patients who had progression during first-line platinum chemotherapy or relapse within 6 months after treatment. Construction of TMAs and immunohistochemical protocols were similar to those described for breast cancer TMAs previously.

Tissue Microarrays (TMAs) and immunohistochemistry (IHC): Tumours were arrayed in tissue microarrays (TMAs) constructed with 2 replicate 0.6mm cores from the centre and periphery of the tumours. The TMAs were immunohistochemically profiled for FEN1 and other biological antibodies (Supplementary Table S10 of supporting information) as

previously described (Abdel-Fatah et al., 2013b; Abdel-Fatah et al., 2013c; Abdel-Fatah et al., 2014a; Abdel-Fatah et al., 2014b; Sultana et al., 2013). Immunohistochemical staining for FEN1 was performed using the Leica Bond Refine Detection kit according to manufacturer instructions (Leica Microsystems). Pre-treatment of TMA sections was performed with citrate buffer (pH 6.0) antigen for 20 minutes and heated further for 20 minutes in a microwave. TMA sections were then incubated for 15 minutes at room temperature with 1:200 anti-FEN1 rabbit polyclonal antibody (NBP1-67924, Novus Biologicals, Littleton, CO, USA). Also, positive and negative (by omission of the primary antibody and IgG-matched serum) controls were prepared for each set of samples. To validate the use of TMAs for immunophenotyping, full-face sections of 40 cases were stained and protein expression levels of the different 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 immune staining: The tumour cores were evaluated by specialist pathologist blinded to the clinicopathological characteristics of patients. Whole field inspection of the core was scored and intensities of nuclear as well as cytoplasmic 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 as previously described (Abdel-Fatah et al., 2013b; Abdel-Fatah et al., 2013c; Abdel-Fatah et al., 2014a; Abdel-Fatah et al., 2014b; Sultana et al., 2013). Low/negative FEN1 (FEN1-) expression was defined by mean of H-score of ≤100. Not all cores within the TMA were suitable for IHC analysis due to missing cores or absence of tumour cells.

Tumor Marker Prognostic Studies (REMARK) criteria, recommended by McShane et al (McShane et al., 2005), were followed throughout this study. This work was approved by Nottingham Research Ethics Committee.

Statistical analysis: Data analysis was performed using SPSS (SPSS, version 17 Chicago, IL). Where appropriate, Pearson's Chi-square, Fisher's exact, Student's t and ANOVA one way tests were used. Cumulative survival probabilities were estimated using the Kaplan–Meier method, and differences between survival rates were tested for significance using the log-rank test. Multivariate analysis for survival was performed using the Cox proportional hazard model. The proportional hazards assumption was tested using standard log-log plots. Hazard ratios (HR) and 95% confidence intervals (95% CI) were estimated for each variable. All tests were two-sided with a 95% CI and a p value < 0.05 considered significant. For multiple comparisons, p values were adjusted according to Holm-Bonferroni correction method.

Cell lines and culture: To evaluate the specificity of the FEN1 antibody used in the current study FEN1 deficient and proficient cells were investigated. FEN1 deficient HeLa SilenciX® cells and control FEN1 proficient HeLa SilenciX® cells were purchased from Tebu-Bio (www.tebu-bio.com). SilenciX cells were grown in DMEM medium (with L-Glutamine 580mg/L, 4500 mg/L D19 Glucose, with 110mg/L Sodium Pyruvate) supplemented with 10% FBS, 1% penicillin/streptomycin and 125 µg/ml Hygromycin B. Western blot analysis was performed as described previously (Sultana et al., 2013).

RESULTS

High *FEN1* transcript levels correlate to aggressive biology and adversely impact breast cancer clinical outcomes

We evaluated *FEN1* mRNA expression in multiple cohorts of breast cancer (training set, test set and external validation cohort). Clinicopathological characteristics are summarized in supplementary Tables S1, S2 and S3. In the training set (n=128), 40.6% of tumours had high *FEN1* mRNA expression, which was significantly associated with high grade ($p<0.0001$), high mitotic index ($p<0.0001$), pleomorphism ($p<0.0001$), glandular de-differentiation ($p=0.032$), HER2 overexpression ($p=0.003$), oestrogen receptor (ER) negative ($p=0.001$), progesterone receptor (PgR) negative ($p=0.005$) and triple negative phenotype ($p=0.001$) (Supplementary Table S4). High *FEN1* mRNA expression in tumours was also associated with adverse disease specific survival in patients ($p=0.008$) (Figure 1a). In the test set (n=249), 50.2% of tumours had high *FEN1* mRNA expression, which remained associated with high T-stage ($p=0.007$), lymph node positivity ($p=0.012$), high grade ($p<0.0001$), high molecular grade ($p<0.0001$), mutant p53 ($p<0.0001$) and ER negativity ($p=0.001$) (Supplementary Table S5). High *FEN1* mRNA expression in tumours was associated with adverse disease specific survival in patients ($p=0.00009$) (Figure 1b).

External validation was conducted in a large series of 1952 tumours comprising the METABRIC (Molecular Taxonomy of Breast Cancer International Consortium) cohort. 52.25% (1020/1952) of tumours had high *FEN1* mRNA expression, which was significantly associated with aggressive clinicopathological features (Table 1), including high histological grade ($p<0.0001$), high mitotic index ($p<0.0001$), pleomorphism ($p<0.0001$), glandular de-

differentiation (p=0.006), HER2 overexpression (p<0.0001), absence of hormonal receptors (ER-/PgR-) (p<0.0001), lymph node positivity (0.02), presence of basal like phenotypes (p<0.0001) and triple negative phenotypes (p<0.0001). High *FEN1* mRNA expression was also found to be significantly associated with previously described molecular phenotypes in breast cancer: PAM50.Her2 (p<0.0001), PAM50.Basal (p<0.0001) and PAM50.LumB (p<0.0001) breast tumours. However, PAM50.LumA tumours were more likely to express low levels of *FEN1* mRNA (p<0.0001).

The METABRIC study by joint clustering of copy number and gene expression data has identified 10 novel biological subgroups [labelled integrative clusters (intClust) 1-10] with good, intermediate or poor prognosis (Curtis et al.). We investigated whether *FEN1* mRNA expression would associate with these distinct biological subgroups. High *FEN1* mRNA expression was significantly associated with intClust.1 (p<0.0001), intClust.5 (p<0.0001), intClust.9 (p<0.0001) and intClust.10 (p<0.0001), which had the worst clinical outcome in the METABRIC study (Curtis et al.). Low *FEN1* mRNA expression was associated with intClust.3 (p<0.0001), intClust.4 (p<0.0001), intClust.7 (p=0.003) and intClust.8 (p<0.0001), which had intermediate to good prognosis in the METABRIC study (Curtis et al.). High *FEN1* mRNA expression in tumours was associated with adverse disease specific survival in the whole cohort (p<0.0001) (Figure 1c). In multivariate Cox regression analysis that included other validated prognostic factors, such as lymph node stage, histological grade and tumour size (NPI components), *FEN1* mRNA expression was a powerful independent predictor for clinical outcome (p<0.0001) (Table 2). Moreover, in patients who received adjuvant endocrine therapy (n=1199), high *FEN1* mRNA expression remained significantly associated with adverse disease specific survival (p<0.0001) (Figure 1d). In patients who

received adjuvant chemotherapy (n=413), high *FEN1* mRNA expression was likewise associated with adverse disease specific survival (p=0.019) (Figure 1e).

FEN1 mRNA expression analysis in the training set, test set and in the external validation cohort provides confirmatory evidence that high *FEN1* mRNA expression is associated with adverse clinicopathological features, aggressive molecular phenotypes and poor survival in patients.

FEN1 protein expression is linked to aggressive breast cancer and poor survival

As the multifunctional roles of FEN1 are likely regulated by several mechanisms, including sub-cellular compartmentalization between, for example, the nucleus and cytoplasm/mitochondria, we proceeded to evaluate FEN1 protein expression in independent cohorts of 568 ER negative breast tumours and 894 ER positive breast tumours. Clinicopathological characteristics are summarized in supplementary Tables S6 and S7. Treatment data is summarized in supplementary treatment data 1. We also correlated FEN1 protein expression to other biomarkers of aggressive phenotype (ER, PR, EGFR, CK14, CK5/6, CK17, CK18, HER2) and DNA repair (PARP1, BRCA1, ATM, XRCC1 and TOP2A). Antigens, primary antibodies, clone, source, optimal dilution and scoring system for each immunohistochemical marker are summarized in supplementary Table S8.

We first confirmed the specificity of FEN1 antibody used in the current study. As shown in figure 2a1, FEN1 proficient cell line shows robust FEN1 protein expression whereas FEN1 knockdown cell shows almost complete absence of FEN1 protein expression. We then conducted immunohistochemical investigations. In 568 ER negative tumours (Figure 2a2**),**

we found significant associations between FEN1 expression and pleomorphism (p= 0.012), tumour type (p<0.0001), lymphovascular invasion (p=0.007), progesterone receptor (p<0.0001), EGFR overexpression (p=0.04), HER2 overexpression (p=0.029) and triple negative phenotype (p=0.032). FEN1 expression was also significantly associated with expression of other DNA repair factors, including BRCA1 (p<0.0001), PARP1 (p<0.0001), XRCC1 (p<0.0001) and TOP2A (p<0.0001) (full data is summarized in supplementary Table S9). High nuclear/high cytoplasmic FEN1 expression was associated with poor survival (p=0.003) (Figure 2b). In patients with early stage lymph node negative (low risk) tumours who did not receive adjuvant chemotherapy, high nuclear/high cytoplasmic tumours remained significantly associated with poor survival (p=0.009) (Figure 2c). In patients who received CMF (cyclophosphamide, methotrexate and 5-Fluoruracil) chemotherapy, high nuclear/high cytoplasmic was associated with poor survival (p=0.05) (Figure 2d). The group that received anthracycline adjuvant chemotherapy did not reach significance (p=0.211), although there was a trend toward poor survival in high nuclear/low cytoplasmic tumours (Supplementary Figure S1a). In the multivariate COX model, FEN1 expression is independently associated with breast cancer specific survival (p=0.007), as well as progression free survival (p=0.003) (Table 3).

In 894 ER positive breast tumours, we similarly found significant association between FEN1 expression and tumour size (p= 0.004), grade (p<0.0001), pleomorphism (p= 0.0004), tumour type (p<0.0001), tubule formation (p<0.0001) and lymphovascular invasion (p=0.007). FEN1 expression was also associated with other DNA repair factors, such as BRCA1 (p= 0.003), XRCC1 (p<0.0001), ATM (p<0.0001) and TOP2A (p<0.0001) (Full data is summarized in supplementary Table S10). High cytoplasmic/low nuclear FEN1 tumours were associated with poor survival (p=0.00016) in ER positive tumours (Figure 2e). In patients with early

stage lymph node negative (low risk) tumours who did not receive adjuvant tamoxifen, high cytoplasmic/low nuclear FEN1 tumours remained significantly associated with poor survival (p=0.003) (Supplementary Figure S1b). In patients with high risk tumours who did not receive adjuvant tamoxifen, high cytoplasmic/low nuclear FEN1 is associated with poor survival (p=0.026) (Figure 2f). On the other hand, patients with tumours that had low cytoplasmic/low nuclear FEN1 had better survival implying that these tumours could be spared long term adjuvant endocrine therapy. In patients with high risk tumours who received adjuvant tamoxifen, high cytoplasmic/low nuclear FEN1 was associated with poor survival (p=0.003) (Figure 2g). On the other hand, patients with tumours that had high cytoplasmic/high nuclear FEN1 had better survival implying that FEN1 could be predictive biomarker of response to endocrine therapy. In the multivariate COX model, FEN1 expression was independently associated with breast cancer specific survival (p=0.003), as well as progression free survival (p=0.004) (Table 3).

Taken together, the *FEN1* mRNA expression as well as FEN1 protein expression data provides compelling evidence that FEN1 expression is a prognostic and a predictive biomarker in breast cancer.

FEN1 protein expression is linked to aggressive epithelial ovarian cancer and poor survival

We then proceeded to investigate the significance of FEN1 protein expression in 156 ovarian epithelial cancers. Demographics are summarized in supplementary Table S11. Positive nuclear expression of FEN1 was seen in 71/156 (45.5%) tumours, and 85/156 (54.5%) tumours were negative for FEN1 protein expression (Figure 3a). FEN1 nuclear expression was associated with serous cystadenocarcinomas (p=0.05), higher pathological grade (p=

0.009), higher FIGO stage (p=0.046) and larger residual tumour burden following surgery (p=0.034) (full data is summarized in supplementary Tables S12). Positive cytoplasmic expression of FEN1 was seen in 126/156 (80.8%) tumours and 30/156 (19.2%) tumours were negative for FEN1 cytoplasmic expression. FEN1 cytoplasmic expression was significantly associated with serous cystadenocarcinomas (p<0.0001), more likely to be sub-optimally debulked (p=0.002), higher FIGO stage (p=0.025) and larger residual tumour burden following surgery (p=0.005) (full data is summarized in supplementary Tables S13). Investigating nuclear as well cytoplasmic expression together, we found that high cytoplasmic/high nuclear FEN1 tumours had the worst ovarian cancer specific (p=0.006) (Figure 3b) and disease free (p=0.008) (Figure 3c). Evaluating nuclear expression alone or cytoplasmic expression alone, FEN1 over expression remains associated with poor survival in ovarian epithelial cancers (Figures 3d, 3e, Supplementary figures S1c,d). In the multivariate COX model, patients with FEN1 nuclear expression showed 2-fold increase in risk of death (p=0.018) (Supplementary Table S14). The multivariate Cox model was adjusted for CA-125 response, FIGO stage and tumour grade.

Taken together, the data provides evidence that FEN1 overexpression is a promising biomarker in ovarian epithelial cancers.

DISCUSSIONS

This is the largest and the first comprehensive study to evaluate FEN1 in breast and ovarian cancers. In breast cancer, high *FEN1* mRNA is linked to aggressive features such as high grade, high mitotic index, pleomorphism, de-differentiation, PAM50. Her2 and PAM50. basal molecular phenotypes. FEN1 is essential for the repair of oxidative base damage through long-patch base excision repair. The data presented here suggests that high *FEN1* mRNA expression is an adaptive response to oxidative stress that is common in breast cancer cells (Brown and Bicknell, 2001). Although not fully understood, a recent study suggested *FEN1* promoter hypomethylation as a mechanism for *FEN1* mRNA over expression in tumours (Singh et al., 2008). High *FEN1* mRNA seen in tumours with high mitotic index also concurs with previous studies demonstrating FEN1 upregulation in cycling cells (Kim, 1998; Kim et al., 2005; Krause et al., 2005; Sato et al., 2003). In the current study, we have also provided the first evidence that *FEN1* mRNA levels are linked to biologically distinct integrative clusters reported in the METABRIC study (Curtis et al., 2012). High *FEN1* mRNA level was frequent in intClust 10 subgroup which is the most highly genomically instable sub group with basal-like features. Interestingly, low *FEN1* mRNA level was seen in intClust 3 subgroup that is characterised by low genomic instability. Together the data provides evidence that high *FEN1* mRNA could be utilised as a biomarker of genomic instability in human tumours. In addition, high *FEN1* mRNA level is also frequently seen in intClust 5 (HER-2 enriched with worst survival), intClust 9 (8q cis-acting/20qamplified mixed subgroup), and intClust 1 (17q23/20q cis-acting luminal B subgroup) subgroups that also manifest an aggressive phenotype. On the other hand, low *FEN1* mRNA level is linked to intClust 4 (includes both ER-positive and ER-negative cases with a flat copy number landscape and termed the 'CNA-devoid' subgroup with extensive lymphocytic infiltration),

intClust 7 (16p gain/16q loss with higher frequencies of 8q amplification luminal A subgroup) and intClust 8 subgroups (classical 1q gain/16q loss luminal A subgroup) (Curtis et al., 2012). The data implies differential roles for FEN1 in distinct molecular phenotypes of breast cancer. High *FEN1* mRNA is associated with poor survival in univariate as well as in multivariate analyses in the whole cohort which is likely to be related to the aggressive phenotype described previously. As expected, intClust 10, intClust 9, intClust 5 and intClust 1 sub-groups that are associated with high FEN1 levels were also associated with poor prognosis in METABRIC study (Curtis et al., 2012). On the other hand, intClust 3, intClust 4, intClust 7 and intClust 8 that are associated with low FEN1 expression, are associated with good to intermediate prognosis (Curtis et al., 2012). Together, the data provides conclusive evidence that *FEN1* mRNA level has prognostic significance in breast cancer. To investigate if *FEN1* mRNA expression may also have predictive significance, we conducted sub-group analysis in tumours treated with adjuvant chemotherapy or endocrine therapy. In patients treated with endocrine therapy, we provide the first evidence that high *FEN1* mRNA level is associated with poor survival implying resistance to endocrine therapy. The link between FEN1, oestrogen and oestrogen receptors (ER) are beginning to emerge. FEN1 not only interacts directly with ER-α but can also augment the interaction of ER-α with oestrogen response element containing DNA and impact upon estrogen-responsive gene expression in cells (Buterin et al., 2006; Moggs et al., 2005). Our data suggests that *FEN1* mRNA over expression is a novel biomarker for endocrine resistance and is likely related to the role of FEN1 in cell proliferation. We have also demonstrated for the first time that high *FEN1* mRNA level is associated with poor survival in patients who received adjuvant chemotherapy implying resistance to cytotoxic therapy.

We then investigated FEN1 protein expression immunohistochemicaly in large cohorts of breast cancers. Although strong association between FEN1 mRNA level and high tumour grade as well as high mitotic index was evident, FEN1 protein level analysis revealed a complex association in breast cancer. In the ER positive cohort, grade 3 and higher mitotic index tumours were more likely in low nuclear/high cytoplasmic FEN1 tumours compared to high nuclear/low cytoplasmic FEN1 or high nuclear/high cytoplasmic FEN1 tumours. Surprisingly, grade 3 and high mitotic index tumours were also seen frequently in low nuclear/low cytoplasmic FEN1 tumours in the ER positive cohort. In ER positive tumours cytoplasmic over expression correlated to poor survival. In ER negative tumours, although no significant clinicopathological associations were seen, high nuclear FEN1 was associated with poor survival. A limitation of our study is that it is retrospective and prospective studies will be needed to confirm our observation. Given the complex multifunctional role of FEN1 protein that is likely regulated by sub-cellular compartmentalization and post-translational modification mechanisms, our data suggest that detailed preclinical mechanistic studies will be required to evaluate the roles of FEN1 protein in breast cancer pathogenesis. However, it is important to note that the clinical data presented here is consistent with a recent preclinical study where FEN1 knockdown by siRNA was shown to be associated with reduced cellular proliferation (van Pel et al., 2013). Moreover, treatment with specific FEN1 inhibitors isolated in that study also resulted in reduced proliferation in cells (van Pel et al., 2013). In another preclinical study, FEN1 mRNA depletion by siRNA resulted in increased sensitivity to chemotherapy such as alkylating agents and platinum chemotherapy (Nikolova et al., 2009). Taken together the data suggest that FEN1 mRNA levels are likely to be the best predictors of response to chemotherapy or endocrine therapy in breast cancer.

Interestingly, FEN1 protein expression also linked to other DNA repair factors such as BRCA1, PARP1, XRCC1 and TOP2A implying altered genomic stability in breast tumours. In contrast to ER negative tumours, in ER positive tumours we found an association between high FEN1 and ATM expression. Previous studies indicate a functional link between FEN1 and ER. FEN1 may regulate ER induced transcriptional response by enhancing the interaction of ER with oestrogen response elements- containing DNA **(**Schultz-Norton et al., 2007**)**. Interestingly a recent study suggests that ER may be involved in the regulation of ATM expression (Guo et al., 2013). In light of the preclinical evidence presented above, the clinical data presented here suggest a complex network that may be operating between ER, FEN1 and ATM in breast cancer cells. However, detailed mechanistic studies are required to confirm this hypothesis. In ovarian cancer, similarly, FEN1 expression is linked to aggressive phenotype and poor survival. Recently, we investigated FEN1 in gastric cancers (Abdel-Fatah et al., 2013a). FEN1 protein over expression was associated with high T-stage (p=0.005), lymph node-positive disease (p=0.02) and poor disease specific survival (p=0.006) (Abdel-Fatah et al., 2013a). In another study in prostate cancer, FEN1 protein over expression was associated with aggressive disease (Lam et al., 2006). Taken together the data suggest that FEN1 protein expression has prognostic and predictive significance in cancers.

Our clinical data suggests that FEN1 may be a promising drug target in cancer. Interestingly, a recent study extrapolating yeast genetic interaction data has also identified FEN1 as an attractive anti-cancer target (van Pel et al., 2013). We have recently initiated a FEN1 drug discovery programme. To facilitate the search for novel FEN1 inhibitors, we developed a fluorogenic donor/quencher reporter pair to monitor generation of reaction product in real time(Dorjsuren et al., 2011). A high-throughput screen was recently conducted on 391,275 compounds arrayed as dilution series within a total of 1,407 plates. Primary screening data

has been uploaded to a public database (http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=588795). Detailed *in vitro* and *in vivo* evaluation and validation of novel FEN1 inhibitors is an area of on-going investigation in our laboratory.

In conclusion, the data presented in the current clinical study suggests that FEN1 is promising biomarker in breast and ovarian epithelial cancers.

ACKNOWLEDGEMENTS

Funding: This work was supported in part by an NIH grant to DMWIII (1 R03 MH092154-01) and the Intramural Research Program of the National Institute on Aging, NIH, Molecular Libraries Common Fund Program of the NIH, National Center for Advancing Translational Sciences (USA), University of Nottingham and Nottingham University Hospitals Charity (UK). RA and CP are supported by the Medical Research Council, UK.

Author contributions: TA.F. **,** S.M., and DMWIII provided intellectual input, conceptual framework, designed the study and helped writing the manuscript. T.A.F, P.M., R.A., A.M., D.A., G.B., and V.M. collected protein expression and gene expression data. R.R, O.M.R, C.C. performed FEN1 gene expression analysis in the metabric cohort. N. A and V.M., performed cell based experiments. D.J.M, D.D, H.S., J.L.I., A.J., A.S., and DMWIII performed FEN1 drug discovery screen. S.C. and I.O.E., performed data analysis and helped writing the manuscript. All authors reviewed and approved the final version of the manuscript.

Disclosures: The authors have declared no conflicts of interest.

REFERENCES

Abdel-Fatah, T., Arora, A., Gorguc, I., Abbotts, R., Beebeejaun, S., Storr, S., Mohan, V., Hawkes, C., Soomro, I., Lobo, D. N., Parsons, S. L.,Madhusudan, S., 2013a. Are DNA Repair Factors Promising Biomarkers for Personalized Therapy in Gastric Cancer? Antioxid Redox Signal 18, 2392-8

Abdel-Fatah, T. M., Albarakati, N., Bowell, L., Agarwal, D., Moseley, P., Hawkes, C., Ball, G., Chan, S., Ellis, I. O.,Madhusudan, S., 2013b. Single-strand selective monofunctional uracil-DNA glycosylase (SMUG1) deficiency is linked to aggressive breast cancer and predicts response to adjuvant therapy. Breast Cancer Res Treat 142, 515-527.

Abdel-Fatah, T. M., Perry, C., Dickinson, P., Ball, G., Moseley, P., Madhusudan, S., Ellis, I. O.,Chan, S. Y., 2013c. Bcl2 is an independent prognostic marker of triple negative breast cancer (TNBC) and predicts response to anthracycline combination (ATC) chemotherapy (CT) in adjuvant and neoadjuvant settings. Ann Oncol 24, 2801-2807.

Abdel-Fatah, T. M., Perry, C., Moseley, P., Johnson, K., Arora, A., Chan, S., Ellis, I. O.,Madhusudan, S., 2014a. Clinicopathological significance of human apurinic/apyrimidinic endonuclease 1 (APE1) expression in oestrogen-receptor-positive breast cancer. Breast Cancer Res Treat 143, 411-421.

Abdel-Fatah, T. M., Russell, R., Agarwal, D., Moseley, P., Abayomi, M. A., Perry, C., Albarakati, N., Ball, G., Chan, S., Caldas, C., Ellis, I. O.,Madhusudan, S., 2014b. DNA polymerase beta deficiency is linked to aggressive breast cancer: A comprehensive analysis of gene copy number, mRNA and protein expression in multiple cohorts. Mol Oncol S1574-7891, 00002-00007.

Bergh, J., Norberg, T., Sjogren, S., Lindgren, A.,Holmberg, L., 1995. Complete sequencing of the p53 gene provides prognostic information in breast cancer patients, particularly in relation to adjuvant systemic therapy and radiotherapy. Nat Med 1, 1029-1034.

Brown, N. S.,Bicknell, R., 2001. Hypoxia and oxidative stress in breast cancer. Oxidative stress: its effects on the growth, metastatic potential and response to therapy of breast cancer. Breast Cancer Res 3, 323-327.

Buterin, T., Koch, C.,Naegeli, H., 2006. Convergent transcriptional profiles induced by endogenous estrogen and distinct xenoestrogens in breast cancer cells. Carcinogenesis 27, 1567-1578.

Chin, S. F., Teschendorff, A. E., Marioni, J. C., Wang, Y., Barbosa-Morais, N. L., Thorne, N. P., Costa, J. L., Pinder, S. E., van de Wiel, M. A., Green, A. R., Ellis, I. O., Porter, P. L., Tavare, S., Brenton, J. D., Ylstra, B.,Caldas, C., 2007. High-resolution aCGH and expression profiling identifies a novel genomic subtype of ER negative breast cancer. Genome Biol 8, R215.

Curtis, C., Shah, S. P., Chin, S. F., Turashvili, G., Rueda, O. M., Dunning, M. J., Speed, D., Lynch, A. G., Samarajiwa, S., Yuan, Y., Graf, S., Ha, G., Haffari, G., Bashashati, A., Russell,

R., McKinney, S., Langerod, A., Green, A., Provenzano, E., Wishart, G., Pinder, S., Watson, P., Markowetz, F., Murphy, L., Ellis, I., Purushotham, A., Borresen-Dale, A. L., Brenton, J. D., Tavare, S., Caldas, C.,Aparicio, S., 2012. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature 486, 346-352.

Dorjsuren, D., Kim, D., Maloney, D. J., Wilson, D. M., 3rd,Simeonov, A., 2011. Complementary non-radioactive assays for investigation of human flap endonuclease 1 activity. Nucleic Acids Res 39, e11.

Henneke, G., Friedrich-Heineken, E.,Hubscher, U., 2003a. Flap endonuclease 1: a novel tumour suppresser protein. Trends Biochem Sci 28, 384-390.

Henneke, G., Koundrioukoff, S.,Hubscher, U., 2003b. Phosphorylation of human Fen1 by cyclin-dependent kinase modulates its role in replication fork regulation. Oncogene 22, 4301-4313.

Hothorn, T., Hornik, K.,Zeileis, A., 2006. Unbiased Recursive Partitioning: A Conditional Inference Framework. J Comput Graph Stat 15 651-674.

Kim, I. S., 1998. Down-regulation of human FEN-1 gene expression during differentiation of promyelocytic leukemia cells. Exp Mol Med 30, 252-256.

Kim, J. M., Sohn, H. Y., Yoon, S. Y., Oh, J. H., Yang, J. O., Kim, J. H., Song, K. S., Rho, S. M., Yoo, H. S., Kim, Y. S., Kim, J. G.,Kim, N. S., 2005. Identification of gastric cancer-related genes using a cDNA microarray containing novel expressed sequence tags expressed in gastric cancer cells. Clin Cancer Res 11, 473-482.

Krause, A., Combaret, V., Iacono, I., Lacroix, B., Compagnon, C., Bergeron, C., Valsesia-Wittmann, S., Leissner, P., Mougin, B.,Puisieux, A., 2005. Genome-wide analysis of gene expression in neuroblastomas detected by mass screening. Cancer Lett 225, 111-120.

Kucherlapati, M., Nguyen, A., Kuraguchi, M., Yang, K., Fan, K., Bronson, R., Wei, K., Lipkin, M., Edelmann, W.,Kucherlapati, R., 2007. Tumor progression in Apc(1638N) mice with Exo1 and Fen1 deficiencies. Oncogene 26, 6297-6306.

Kucherlapati, M., Yang, K., Kuraguchi, M., Zhao, J., Lia, M., Heyer, J., Kane, M. F., Fan, K., Russell, R., Brown, A. M., Kneitz, B., Edelmann, W., Kolodner, R. D., Lipkin, M.,Kucherlapati, R., 2002. Haploinsufficiency of Flap endonuclease (Fen1) leads to rapid tumor progression. Proc Natl Acad Sci U S A 99, 9924-9929.

Lam, J. S., Seligson, D. B., Yu, H., Li, A., Eeva, M., Pantuck, A. J., Zeng, G., Horvath, S.,Belldegrun, A. S., 2006. Flap endonuclease 1 is overexpressed in prostate cancer and is associated with a high Gleason score. BJU Int 98, 445-451.

Larsen, E., Gran, C., Saether, B. E., Seeberg, E.,Klungland, A., 2003. Proliferation failure and gamma radiation sensitivity of Fen1 null mutant mice at the blastocyst stage. Mol Cell Biol 23, 5346-5353.

Larsen, E., Kleppa, L., Meza, T. J., Meza-Zepeda, L. A., Rada, C., Castellanos, C. G., Lien, G. F., Nesse, G. J., Neuberger, M. S., Laerdahl, J. K., William Doughty, R.,Klungland, A.,

2008. Early-onset lymphoma and extensive embryonic apoptosis in two domain-specific Fen1 mice mutants. Cancer Res 68, 4571-4579.

Liu, L., Zhou, C., Zhou, L., Peng, L., Li, D., Zhang, X., Zhou, M., Kuang, P., Yuan, Q., Song, X.,Yang, M., 2012. Functional FEN1 genetic variants contribute to risk of hepatocellular carcinoma, esophageal cancer, gastric cancer and colorectal cancer. Carcinogenesis 33, 119-123.

McShane, L. M., Altman, D. G., Sauerbrei, W., Taube, S. E., Gion, M.,Clark, G. M., 2005. Reporting recommendations for tumor marker prognostic studies (REMARK). J Natl Cancer Inst 97, 1180-1184.

Moggs, J. G., Murphy, T. C., Lim, F. L., Moore, D. J., Stuckey, R., Antrobus, K., Kimber, I.,Orphanides, G., 2005. Anti-proliferative effect of estrogen in breast cancer cells that re-express ERalpha is mediated by aberrant regulation of cell cycle genes. J Mol Endocrinol 34, 535-551.

Nikolova, T., Christmann, M.,Kaina, B., 2009. FEN1 is overexpressed in testis, lung and brain tumors. Anticancer Res 29, 2453-2459.

Pawitan, Y., Bjohle, J., Amler, L., Borg, A. L., Egyhazi, S., Hall, P., Han, X., Holmberg, L., Huang, F., Klaar, S., Liu, E. T., Miller, L., Nordgren, H., Ploner, A., Sandelin, K., Shaw, P. M., Smeds, J., Skoog, L., Wedren, S.,Bergh, J., 2005. Gene expression profiling spares early breast cancer patients from adjuvant therapy: derived and validated in two population-based cohorts. Breast Cancer Res 7, R953-964.

Sato, M., Girard, L., Sekine, I., Sunaga, N., Ramirez, R. D., Kamibayashi, C.,Minna, J. D., 2003. Increased expression and no mutation of the Flap endonuclease (FEN1) gene in human lung cancer. Oncogene 22, 7243-7246.

Shen, B., Singh, P., Liu, R., Qiu, J., Zheng, L., Finger, L. D.,Alas, S., 2005. Multiple but dissectible functions of FEN-1 nucleases in nucleic acid processing, genome stability and diseases. Bioessays 27, 717-729.

Singh, P., Yang, M., Dai, H., Yu, D., Huang, Q., Tan, W., Kernstine, K. H., Lin, D.,Shen, B., 2008. Overexpression and hypomethylation of flap endonuclease 1 gene in breast and other cancers. Mol Cancer Res 6, 1710-1717.

Sultana, R., Abdel-Fatah, T., Abbotts, R., Hawkes, C., Albarakati, N., Seedhouse, C., Ball, G., Chan, S., Rakha, E. A., Ellis, I. O.,Madhusudan, S., 2013. Targeting XRCC1 deficiency in breast cancer for personalized therapy. Cancer Res 73, 1621-1634.

van Pel, D. M., Barrett, I. J., Shimizu, Y., Sajesh, B. V., Guppy, B. J., Pfeifer, T., McManus, K. J.,Hieter, P., 2013. An Evolutionarily Conserved Synthetic Lethal Interaction Network Identifies FEN1 as a Broad-Spectrum Target for Anticancer Therapeutic Development. PLoS Genet 9, e1003254.

Wu, Z., Lin, Y., Xu, H., Dai, H., Zhou, M., Tsao, S., Zheng, L.,Shen, B., 2012. High risk of benzo[alpha]pyrene-induced lung cancer in E160D FEN1 mutant mice. Mutat Res 731, 85-91.

Xu, H., Zheng, L., Dai, H., Zhou, M., Hua, Y.,Shen, B., 2011. Chemical-induced cancer incidence and underlying mechanisms in Fen1 mutant mice. Oncogene 30, 1072-1081.

Yang, M., Guo, H., Wu, C., He, Y., Yu, D., Zhou, L., Wang, F., Xu, J., Tan, W., Wang, G., Shen, B., Yuan, J., Wu, T.,Lin, D., 2009. Functional FEN1 polymorphisms are associated with DNA damage levels and lung cancer risk. Hum Mutat 30, 1320-1328.

Zheng, L., Dai, H., Zhou, M., Li, M., Singh, P., Qiu, J., Tsark, W., Huang, Q., Kernstine, K., Zhang, X., Lin, D.,Shen, B., 2007. Fen1 mutations result in autoimmunity, chronic inflammation and cancers. Nat Med 13, 812-819.

Zheng, L., Jia, J., Finger, L. D., Guo, Z., Zer, C.,Shen, B., 2011. Functional regulation of FEN1 nuclease and its link to cancer. Nucleic Acids Res 39, 781-794.

Table 1: Association between *FEN1* mRNA expression and clinico-pathologic variables

Variable	***FEN1* mRNA Expression**		X^2 ***Adjusted p value***
	Low **N=932 (47.75%)**	**High** **N= 1020 (52.25%)**	
A) Pathological Parameters			
Tumour Size			
T1 a + b (≤1.0)	5 (0.26%)	3 (0.15%)	0.629
T1 c (>1.0 -2.0)	1 (0.05%)	2 (0.1%)	0.938
T2 (>2.0-5)	10 (0.51%)	4 (0.2%)	0.131
T3 (>5)	909 (46.57%)	999 (51.18%)	0.649
Lymph node stage			
Negative	520 (26.64%)	492 (25.2%)	**0.000991**
Positive (1-3 nodes)	225 (11.53%)	281 (14.4%)	0.0961
Positive (>3 nodes)	184 (9.43%)	244 (12.5%)	**0.0297**
Grade**			
G1	136 (6.97%)	28 (1.43%)	**9.38×10^{-21}**
G2	471 (24.13%)	293 (15.01%)	**9.64×10^{-23}**
G3	278 (14.24%)	673 (34.48%)	**4.89×10^{-57}**
Tumour Types			
IDC-NST	11 (0.56%)	6 (0.31%)	0.245
Tubular	24 (1.23%)	2 (0.1%)	**1.18×10^{-5}**
ILC	2 (0.1%)	2 (0.1%)	0.681
Medullary	2 (0.1%)	23 (1.18%)	**0.000143**
Others	754 (38.63%)	818 (41.91%)	0.737
Mitotic Index			
M1 (low; mitoses < 10)	591 (30.28%)	410 (21%)	**1.89×10^{-24}**
M2 (medium; mitoses 10-18)	158 (8.09%)	220 (11.27%)	**0.0117**
M3 (high; mitosis >18)	40 (2.05%)	216 (11.07%)	**5.25×10^{-28}**
Pleomorphism			
1 (small-regular uniform)	12 (0.61%)	5 (0.26%)	0.099
2 (Moderate variation)	388 (19.88%)	208 (10.66%)	**4.15×10^{-24}**
3 (Marked variation)	390 (19.98%)	633 (32.43%)	**6.31×10^{-19}**
Tubule formation			
1 (>75% of definite tubule)	51 (2.61%)	6 (0.31%)	**3.69×10^{-10}**
2 (10%-75% definite tubule)	243 (12.45%)	111 (5.69%)	**5.55×10^{-18}**
3 (<10% definite tubule)	496 (25.41%)	729 (37.35%)	**1.19×10^{-16}**
Molecular phenotypes			
Her2 overexpression (No)	864 (44.26%)	846 (43.34%)	**9.88×10^{-11}**
(Yes)	68 (3.48%)	174 (8.91%)	**9.88×10^{-11}**
Triple negative (No)	859 (44.01%)	780 (39.96%)	**6.67×10^{-21}**
(Yes)	73 (3.74%)	240 (12.3%)	**6.67×10^{-21}**

Basal like	(No)	900 (46.11%)	918 (47.03%)	**1.69×10^{-8}**
	(Yes)	32 (1.64%)	102 (5.23%)	**1.69×10^{-8}**
ER	(Negative)	95 (4.87%)	342 (17.52%)	**9.02×10^{-35}**
	(Positive)	818 (41.91%)	667 (34.17%)	**1.03×10^{-30}**
PgR	(Negative)	329 (16.85%)	593 (30.38%)	**9.24×10^{-24}**
	(Positive)	603 (30.89%)	427 (21.88%)	**9.24×10^{-24}**
PAM50.Her2		61 (3.12%)	177 (9.07%)	**5.19×10^{-13}**
PAM50.Basal		43 (2.2%)	279 (14.29%)	**2.7×10^{-41}**
PAM50.LumA		548 (28.07%)	166 (8.5%)	**3.78×10^{-84}**
PAM50.LumB		129 (6.61%)	355 (18.19%)	**1.56×10^{-26}**
intClust.1		27 (1.38%)	112 (5.74%)	**7.47×10^{-12}**
intClust.2		29 (1.49%)	42 (2.15%)	0.287
intClust.3		217 (11.12%)	70 (3.59%)	**2.74×10^{-24}**
intClust.4		227 (11.63%)	104 (5.33%)	**1.37×10^{-16}**
intClust.5		43 (2.2%)	142 (7.27%)	**4.05×10^{-12}**
intClust.6		42 (2.15%)	43 (2.2%)	0.839
intClust.7		110 (5.64%)	79 (4.05%)	**0.00316**
intClust.8		185 (9.48%)	110 (5.64%)	**3.35×10^{-8}**
intClust.9		36 (1.84%)	110 (5.64%)	**1.06×10^{-8}**
intClust.10		16 (0.82%)	208 (10.66%)	**7.59×10^{-38}**

* Statistically significant; **: grade as defined by NGS; 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: Multivariate analysis using Cox regression analysis in the METABRIC cohort confirms that *FEN1* mRNA over expression is a powerful independent prognostic factor

Variable	**BCSS**		**OS**	
	HR	***p***	**HR**	***p***
***FEN1* mRNA over expression**	1.25	**9.19×10^{-7}**	1.14	**4.39×10^{-4}**
Tumour size	1.01	**3.59×10^{-7}**	1.01	**2.28×10^{-8}**
Grade		**0.01**		0.08
G1	1.0		1.0	
G2	1.64		1.23	
G3	1.90		1.31	
Lymph node		**1.46×10^{-4}**		**2.01×10^{-3}**
Negative	1.0		1.0	
Positive (1-3 nodes)	1.57		1.30	
Positive (>3 nodes)	3.30		2.31	

** Statistically significant

BCSS; Breast cancer specific survival, OS; overall survival, HR; hazard ratio, CI; confident interval

Table 3: Multivariate survival analysis using Cox regression for Nottingham breast cancer cohort.

Clinico-pathological variables	ER Negative cohort, Breast cancer specific Survival at 10 years		ER Negative cohort, Progression Free Survival at 10 years		ER Positive cohort, Breast cancer specific Survival at 10 years		ER Positive cohort, Progression Free Survival at 10 years	
	HR (95% CI)	***P* value**	**HR (95% CI)**	***P* value**	**HR (95% CI)**	***P* value**	**HR (95% CI)**	***P* value**
FEN1 protein expression		**0.007***		**0.003***		**0.003***		**0.004***
Nuc - / Cyt -	1		1		1		1	
Nuc - / Cyt +	1.314 (0.85 - 2.031)		1.617 (1.088 - 2.403)		1.53 (1.163 – 2.013)		1.75 (1.245 – 2.459)	
Nuc + / Cyt -	1.894 (1.239 - 2.896)		1.851 (1.254 - 2.731)		1.247 (0.828 – 1.883)		1.08 (0.605 – 1.929)	
Nuc + / Cyt +	1.958 (1.194 - 3.213)		1.817 (1.130 - 2.920)		0.842 (0.551 – 1.287)		0.947 (0.543 – 1.653)	
XRCC1 protein expression (Continuous)	0.497 (0.347 - 0.713)	**<0.0001***	0.502 (0.36 – 0.701)	**<0.0001***	0.537 (0.384 – 0.753)	**<0.0001***	0.418 (0.282 – 0.619)	**<0.0001***
Tumour size (Continuous)	1.051 (1.011- 1.092)	**0.012***	1.042 (1.001 – 1.083)	0.43	1.101 (0.993 – 1.221)	0.069	1.052 (0.914 – 1.211)	0.480
Lymph node stage		**<0.0001***		**<0.0001***		**<0.0001***		**0.001***
Negative	1		1		1		1	
Positive (1-3 nodes)	1.13 (0.81-1.60)		1.157 (0.807 – 1.685)		1.666 (1.26 – 2.203)		1.848 (1.303 – 2.622)	
Positive (>3 nodes)	2.23 (1.48-3.37)		5.286 (3.698 – 7.557)		3.346 (2.198 – 5.094)		4.331 (2.654 – 7.068)	
Chemotherapy		**0.029***		0.832		0.384		0.676
No	1		1		1		1	
Yes	0.029 (0.64 - 0.976)		0.832 (0.687 – 1.007)		1.147 (0.842 – 1.562)		1.085 (0.739 – 1.594)	
Tumour grade		0.477		0.821		0.07		**<0.0001***
Grade 1(low)	1		1		1		1	
Grade 2 (intermediate)	2.989 (0.379 - 23.579)		0.977 (0.282 – 3.382)		1.296 (0.903 – 1.86)		1.826 (1.036 – 3.22)	
Grade 3 (high)	3.329 (0.451 - 24.603)		0.841 (0.26 – 2.723)		1.56 (1.064 – 2.288)		3.424 (1.952 – 6.008)	
Lymphovascular invasion		0.705		0.635		**0.033***		**0.001***
No	1		1		1		1	
Yes	0. 941 (0.688 - 1.288)		1.073 (0.803 – 1.432)		1.324 (1.022 – 1.715)		1.71 (1.239 – 2.359)	
Her2 expression (Continuous)	1.107 (0.769 - 1.594)	0.585	1.061 (0.761 – 1.479)	0.726	1.229 (0.775 – 1.947)	0.381	1.549 (0.925 -2.593)	0.096

* statistically significant

Figure legends

Figure 1. *FEN1* gene expression in breast cancer. Kaplan Meier curves showing breast cancer specific survival in the (a) training set, (b) test set, (c) external validation (METABRIC) cohort, (d) METABRIC cohort patients receiving endocrine therapy, and (e) METABRIC cohort patients receiving chemotherapy.

Figure 2. FEN1 protein expression in breast cancer. (a) 1.Western blot showing specificity of FEN1 antibody. 2. Microphotographs of FEN1 protein expression in breast cancer tissue (magnification x 200). Kaplan Meier curves showing breast cancer specific survival in the (b) ER negative (-) breast cancer (whole cohort), (c) ER negative (-) breast cancer patients who received no chemotherapy. (d) ER negative (-) breast cancer patients who received CMF chemotherapy, (e) ER positive (+) breast cancer patients (whole cohort), (f) high risk ER positive (+) breast cancer patients who received no endocrine therapy and (g) high risk ER positive (+) breast cancer patients who received endocrine therapy. N= nuclear expression, C= cytoplasmic expression, '-'= negative expression, '+' = positive expression. (h)

Figure 3. FEN1 protein expression in ovarian cancer. (a) Microphotographs of FEN1 protein expression in ovarian cancer tissue (magnification x 200). Investigating nuclear and cytoplasmic expression, Kaplan Meier curves showing cancer specific survival in epithelial ovarian cancer (b), disease free survival (c). (d) Investigating nuclear expression of FEN1 alone, Kaplan Meier curves showing cancer specific survival. (e) Investigating cytoplasmic expression of FEN1 alone, Kaplan Meier curves showing disease free survival (DFS) survival in epithelial ovarian cancer patients.