

Keywords: ATM; breast cancer; prognosis; DNA DSB repair

ATM has a major role in the double-strand break repair pathway dysregulation in sporadic breast carcinomas and is an independent prognostic marker at both mRNA and protein levels

S Rondeau¹, S Vacher¹, L De Koning², A Briaux¹, A Schnitzler¹, W Chemlali¹, C Callens¹, R Lidereau¹ and I Bièche^{*,1,3}

¹Pharmacogenomics Unit, Department of Genetics, Institut Curie, 26 rue d'Ulm, Paris 75005, France; ²Department of Translational Research, Institut Curie, 26 rue d'Ulm, Paris 75005, France and ³EA7331, University of Paris Descartes, 4 Avenue de l'Observatoire, Paris 75006, France

Background: Ataxia telangiectasia mutated (ATM) is a kinase that has a central role in the maintenance of genomic integrity by activating cell cycle checkpoints and promoting repair of DNA double-strand breaks (DSB). In breast cancer, a low level of ATM was correlated with poor outcome; however, the molecular mechanism of this downregulation is still unclear.

Methods: We used qRT-PCR assay to quantify mRNA levels of ATM gene in 454 breast tumours from patients with known clinical/pathological status and outcome; reverse phase protein arrays (RPPA) were used to assess the levels of ATM and 14 proteins in 233 breast tumours.

Results: ATM mRNA was associated with poor metastasis-free survival (MFS) ($P=0.00012$) on univariate analysis. ATM mRNA and protein levels were positively correlated ($P=0.00040$). A low level of ATM protein was correlated with poorer MFS ($P=0.000025$). ATM expression at mRNA or protein levels are independent prognostic factors on multivariate analysis ($P=0.00046$ and $P=0.00037$, respectively). The ATM protein level was positively correlated with the levels of six proteins of the DSB repair pathway: H2AX ($P<0.0000001$), XRCC5 ($P<0.0000001$), NBN ($P<0.0000001$), Mre11 ($P=0.0000029$), Rad50 ($P=0.0064$), and TP53BP1 ($P=0.026$), but not with proteins involved in other pathways that are altered in cancer. Low expression of ATM protein was significantly associated with high *miR-203* expression ($P=0.011$).

Conclusion: We confirmed that ATM expression is an independent prognostic marker at both RNA and protein levels. We showed that alteration of ATM is involved in dysregulation of the DSB repair pathway. Finally, *miR-203* may be responsible for downregulation of ATM in breast cancers.

Ataxia telangiectasia mutated (ATM) is a tumour-suppressor gene encoding a serine/threonine kinase that has an essential role in the DNA double-strand break (DSB). ATM transduces a DSB repair signal to downstream effector machinery by phosphorylating critical protein substrates (Hu *et al*, 2010). ATM also has a role in cell cycle arrest, apoptosis

*Correspondence: Dr I Bièche; E-mail: ivan.bieche@curie.fr

Received 8 October 2014; revised 8 January 2015; accepted 12 January 2015; published online 5 March 2015

© 2015 Cancer Research UK. All rights reserved 0007–0920/15

and senescence, in order to prevent genomic instability (Stagni *et al*, 2014).

Germline mutations in *ATM* are responsible for an autosomal recessive disease, ataxia-telangiectasia, characterised by progressive neurodegeneration, telangiectasia, immunodeficiency, thymic and gonadal atrophy, marked predisposition to malignancies and acute sensitivity to ionising radiation (Shiloh and Ziv, 2013).

ATM is also a breast cancer susceptibility gene with low penetrance and low frequency. The evidence for an increased risk of breast cancer is derived from carriers of heterozygous *ATM* mutations (Thompson *et al*, 2005; Renwick *et al*, 2006).

ATM transcript downregulation has been reported in gastric, pancreatic, colorectal, laryngeal and pharyngeal cancers and has been described as a poor prognostic factor (Grabsch *et al*, 2006; Lee *et al*, 2011; Beggs *et al*, 2012; Kim *et al*, 2014).

Genomic alterations of *ATM* gene have only been observed in rare cases (3–5%) of breast tumours (The Cancer Genome Atlas Network, 2012). *ATM* expression has been investigated in breast cancer but rarely in large cohorts. Downregulation of *ATM* has been described at mRNA and protein levels (Ye *et al*, 2007; Tommiska *et al*, 2008; Salimi *et al*, 2012; Bueno *et al*, 2014). *ATM* mRNA underexpression has been correlated with poor prognosis (Ye *et al*, 2007), and *ATM* protein underexpression has recently been described as an independent prognostic factor (Bueno *et al*, 2014). Contradictory associations between *ATM* downregulation and classical clinicopathological parameters have been reported in these studies.

The molecular mechanisms involved in *ATM* deregulation in breast cancers are still unclear. Somatic mutations of the *ATM* gene are found in only about 3% of breast cancers (Cerami *et al*, 2012; Gao *et al*, 2013) and *ATM* epigenetic dysregulation mediated by CpG island methylation has been studied but with discordant results (Vo *et al*, 2004; Treilleux *et al*, 2007; Flanagan *et al*, 2009; Pal *et al*, 2010). *ATM* gene deletion could be one mechanism involved, as loss of the 11q22–q23 region containing *ATM* gene is fairly frequent (The Cancer Genome Atlas Network, 2012). A posttranscriptional *ATM* regulation mechanism mediated by several microRNAs has also been reported in various cancers (Ng *et al*, 2010; Yan *et al*, 2010; Song *et al*, 2011; Lin *et al*, 2012; Bisso *et al*, 2013; Di Francesco *et al*, 2013; Guo *et al*, 2013; Mansour *et al*, 2013; Guo *et al*, 2014; Liang *et al*, 2014; Zhang *et al*, 2014; Zhou *et al*, 2014).

In order to investigate whether variations in the expression of *ATM* gene are associated with classical pathological parameters and outcome, we assessed *ATM* mRNA expression in 454 unilateral invasive breast tumours, using real-time quantitative reverse-transcription PCR (qRT–PCR) and *ATM* protein levels by reverse-phase protein arrays (RPPA) in 233 breast tumours. In addition, we determined the correlation between *ATM* dysregulation and 17 candidate proteins involved in various cancer pathways as well as 12 candidate miRNAs putatively involved in the posttranscriptional regulation of *ATM*.

MATERIAL AND METHODS

Patients and samples. Samples of 454 primary unilateral invasive primary breast tumours excised from women managed at Curie Institute—René Huguenin Hospital (Saint-Cloud, France) from 1978 to 2008 have been analysed. The samples were immediately stored in liquid nitrogen until mRNA and protein extraction. Tumour samples were considered suitable for our study if the proportion of tumour cells exceeded 70%.

All patients (mean age 61.7 years, range 31–91 years) met the following criteria: primary unilateral nonmetastatic breast carcinoma for which complete clinical, histological and biological data

were available; no radiotherapy or chemotherapy before surgery; and full follow-up at Curie Institute—René Huguenin Hospital.

Treatment consisted of modified radical mastectomy in 291 cases (64.1%) and breast-conserving surgery plus locoregional radiotherapy in 163 cases (35.9%). The patients had a physical examination and routine chest radiography every 3 months for 2 years, then annually. Mammograms were done annually. Adjuvant therapy was administered to 367 patients, consisting of chemotherapy alone in 89 cases, hormone therapy alone in 176 cases and both treatments in 102 cases. The histological type and the number of positive axillary nodes were established at the time of surgery. The malignancy of infiltrating carcinomas was scored according to Scarff–Bloom–Richardson's (SBR) histoprognostic system.

Hormone receptor (HR) (oestrogen receptor (ER α), progesterone receptor (PR)) and human epidermal growth factor receptor 2 (ERBB2) status were determined at the protein level by using biochemical methods (dextran-coated charcoal method, enzyme immunoassay or immunohistochemistry) and confirmed by real-time quantitative RT–PCR assays (Bieche *et al*, 1999, 2001a).

The population was divided into four groups according to HR (ER α and PR) and ERBB2 status, as follows: two luminal subtypes (HR + (ER α + or PR +)/ERBB2 + ($n = 54$)) and (HR + (ER α + or PR +)/ERBB2 – ($n = 289$)); an ERBB2 + subtype (HR – (ER α – and PR –)/ERBB2 + ($n = 43$)) and a triple-negative subtype (HR – (ER α – and PR –)/ERBB2 – ($n = 68$)). Standard prognostic factors are shown in Supplementary Table S1.

The median follow-up is 9.5 years (range 5 months to 33 years); 174 patients had a metastasis.

Ten specimens of adjacent normal breast tissue from breast cancer patients ($n = 4$) or normal breast tissue from women undergoing cosmetic breast surgery ($n = 6$) were used as sources of normal mRNA (Finak *et al*, 2006).

Real-time qRT–PCR for *ATM* expression. Quantitative values were obtained from the cycle number (Ct value) at which the increase in the fluorescence signal associated with exponential growth of PCR products started to be detected by the laser detector of the ABI Prism 7900 sequence detection system (Perkin–Elmer Applied Biosystems, Foster City, CA, USA), using the PE Biosystems analysis software (Perkin Elmer Applied Biosystems) according to the manufacturer's manuals.

The precise amount of total mRNA added to each reaction mix (based on optical density) and its quality (i.e., lack of extensive degradation) are both difficult to assess. Therefore we also quantified transcripts of the *TBP* gene (Genbank accession NM_003194) encoding the TATA box-binding protein (a component of the DNA-binding protein complex TFIID) as an endogenous RNA control and normalised each sample on the basis of its *TBP* content. We selected *TBP* as an endogenous control, because the prevalence of its transcripts is moderate, and because there are no known *TBP* retropseudogenes (retropseudogenes lead to coamplification of contaminating genomic DNA and thus interfere with RT–PCR, despite the use of primers in separate exons) (Bieche *et al*, 1999).

Results are expressed as N -fold differences in *ATM* gene expression relative to the *TBP* gene and termed ' N_{ATM} ' were determined as $N_{ATM} = 2^{\Delta Ct_{sample}}$, where the ΔCt value of the sample was determined by subtracting the Ct value of the *ATM* gene from the Ct value of the *TBP* gene.

The N_{ATM} values of the samples were subsequently normalised such that the median of the N_{ATM} values for the 10 normal breast tissues was 1.

In addition to *TBP*, we previously used a second endogenous control, the *RPLP0* gene, (also known as 36B4) which encodes human acidic ribosomal phosphoprotein P0, in a subgroup of our large series of breast tumour RNAs. Each sample was normalised on the basis of its *TBP* or *RPLP0* content. Similar results were

obtained with the two endogenous controls for various candidate target gene expressions (Bieche *et al*, 2004). Then we only used TBP as an endogenous control.

The primers for *TBP* and *ATM* genes were chosen with the assistance of the Oligo 6.0 program (National Biosciences, Plymouth, MN, USA). We scanned the dbEST and nr databases to confirm the total gene specificity of the nucleotide sequences chosen for the primers and the absence of single-nucleotide polymorphisms. The nucleotide sequences of the primers used were as follows: *ATM*-U (5'-CCAGCTGTGCAGCGAACAAAT-3') and *ATM*-L (5'-TCTAAGCACGTTTCTGCTAACCAGT-3') for *ATM* gene (PCR product of 91 bp), and *TBP*-U (5'-TGCACAGGACCCAAGAGTGA-3') and *TBP*-L (5'-CACATCACAGCTCCCCACCA-3') for *TBP* gene (PCR product of 132 bp). To avoid amplification of contaminating genomic DNA, one of the two primers was placed at the junction between two exons. Agarose gel electrophoresis was used to verify the specificity of PCR amplicons. The conditions of total RNA extraction, cDNA synthesis and PCR were as previously described (Bieche *et al*, 2001b).

miRNAs expression analysis. MicroRNAs were isolated with the RNA extraction procedure used for the protein-coding transcripts (total RNA extraction). Reverse transcription was performed with the QIAGEN miScript Reverse Transcription Kit, according to the manufacturer's protocol (QIAGEN, GmbH, Hilden, Germany). Specific miRNAs were quantified by real-time PCR with the QIAGEN miScript SYBR Green PCR Kit (QIAGEN). The small nucleolar RNU44 was used as an internal control. The expression of *miR-18a*, *miR-26a*, *miR-26b*, *miR-27a*, *miR-100*, *miR-101*, *miR-106a*, *miR-181a2*, *miR-181b1*, *miR-203*, *miR-223* and *miR-421* was subsequently analysed in 30 samples (15 with low level and 15 with high level of ATM protein). The relative expression level of each miRNA, expressed as *N*-fold difference in target miRNA expression relative to RNU44, and termed 'Ntarget', was calculated as follows: $N_{target} = 2^{\Delta Ct}$. The ΔCt value of a given sample was determined by subtracting the Ct value of the target miRNA from the Ct value of RNU44. The Ntarget values of the samples were subsequently normalised so that the median Ntarget value of samples with low level of ATM protein was 1.

RPPA. Reverse phase protein array (RPPA) technology is a miniaturised 'antigen-down' or 'dot-blot' immunoassay suitable for quantifying the relative, semi-quantitative or quantitative abundance of total protein levels and posttranslational modifications across a variety of biological samples (Akbani *et al*, 2014).

Samples were disrupted in Laemmli buffer (50 mM Tris pH=6.8, 2% SDS, 5% glycerol, 2 mM DTT, 2.5 mM EDTA, 2.5 mM EGTA, 1 × HALT Phosphatase inhibitor (Perbio 78420), Protease inhibitor cocktail complete MINI EDTA-free (Roche 1836170, 1 tablet/10 ml), 4 mM Na₃VO₄ and 20 mM NaF) qsp 5 ml H₂O, using a TissueLyser (Qiagen) and two 5-mm stainless beads per sample. Extracts were then boiled for 10 min at 100 °C, passed through a fine needle to reduce viscosity and centrifuged 15 min at 13 000 r.p.m. The supernatant was harvested and stored at -80 °C. Protein concentration was determined (Pierce BCA Reducing Agent Compatible Kit, Pierce, Rockford, IL, USA, ref 23252). Samples were deposited onto nitrocellulose-covered slides (Fast slides, Maine Manufacturing, Sanford, ME, USA) using a dedicated arrayer (2470 arrayer, Aushon Biosystems, Billerica, MA, USA). Five serial dilutions, ranging from 1000 to 62.5 μg ml⁻¹, and two technical replicates per dilution were printed for each sample. Arrays were labelled with specific antibodies recognising studied proteins (see Supplementary Table S2 for the list of antibody references) or without primary antibody (negative control), using an Autostainer Plus (Dako, Glostrup, Denmark). Briefly, slides were incubated with avidin, biotin and peroxidase blocking reagents (Dako) before saturation with TBS containing 0.1% Tween-20 and 5% BSA (TBST-BSA). Slides were then probed

overnight at 4 °C with primary antibodies diluted in TBST-BSA. After washes with TBST, arrays were probed with horseradish peroxidase-coupled secondary antibodies (Jackson ImmunoResearch Laboratories, Newmarket, UK) diluted in TBST-BSA for 1 h at RT. To amplify the signal, slides were incubated with Bio-Rad Amplification Reagent (Bio-Rad Laboratories, Hercules, CA, USA) for 15 min at RT. The arrays were washed with TBST, probed with Alexa647-Streptavidin (Molecular Probes, Life Technologies, Grand Island, NY, USA) diluted in TBST-BSA for 1 h at RT and washed again in TBST. For staining of total proteins, arrays were incubated 15 min in 7% acetic acid and 10% methanol, rinsed twice in water, incubated 10 min in Sypro Ruby (Life Technologies, Grand Island, NY, USA) and rinsed again. The processed slides were dried by centrifugation and scanned using a GenePix 4000B microarray scanner (Molecular Devices, Sunnyvale, CA, USA). Spot intensity was determined with the MicroVigene software (VigeneTech Inc., Carlisle, MA, USA). All primary antibodies used in RPPA have been previously tested by western blotting to assess their specificity for the protein of interest.

Raw data were normalised using Normacurve (Troncale *et al*, 2012), which normalises for fluorescent background per spot, a total protein stain and potential spatial bias on the slide. Next, each RPPA slide was median centred and scaled (divided by median absolute deviation). We then corrected for remaining sample loading effects individually for each array by correcting the dependency of the data for individual arrays on the median value of each sample over all arrays using a linear regression.

Statistical analysis. The distributions of ATM mRNA and protein, other target protein and miRNA levels were characterised by their median values and ranges. Relationships between ATM mRNA and protein levels and other target genes and miRNA levels, and between mRNA and protein levels and clinical parameters, were identified using nonparametric tests, the Kruskal-Wallis *H* test (relationship between one quantitative parameter and ≥2 qualitative parameters) and the Spearman rank correlation test (relation between two quantitative parameters). Differences were considered significant at confidence levels >95% (*P*<0.05).

Metastasis-free survival (MFS) was determined as the interval between initial diagnosis and detection of the first metastasis.

To visualise the efficacy of a molecular marker (ATM mRNA and protein levels) to discriminate between two populations (patients who developed/did not develop metastases) in the absence of an arbitrary cutoff value, data were summarised in a ROC (receiver operating characteristic) curve (Hanley and McNeil, 1982). The AUC (area under curve) was calculated as a single measure to discriminate efficacy. Survival distributions were estimated by the Kaplan-Meier method, and the significance of differences between survival rates were ascertained with the log-rank test.

The Cox proportional hazards regression model was used to assess prognostic significance, and the results are presented as hazard ratios and 95% confidence intervals (CIs).

RESULTS

Relationship between ATM mRNA expression in breast tumours and classical clinicopathological parameters and patient outcome. ATM mRNA expression level was assessed in the whole series of 454 samples (Supplementary Table S1) and in 10 normal breast tissue samples.

ATM mRNA expression showed no significant difference between the two groups of normal breast tissue controls: the four adjacent normal breast tissues from breast cancer patients vs the six normal breast tissues from women undergoing cosmetic breast surgery.

ATM mRNA expression level was weakly associated with three classical predictive factors: age (*P*=0.042), SBR histological grade

Table 1. Characteristics of the 454 breast tumours relative to ATM mRNA expression

	Total population (%)	ATM mRNA expression relative to normal breast	P ^a
Total	454 (100.0)	0.98 (0.21–4.99) ^b	
Age, years			
≤ 50	97 (21.4)	0.88 (0.28–2.42)	0.042
> 50	357 (78.6)	1.00 (0.21–4.99)	
SBR histological grade^{c,d}			
I	58 (13.0)	1.15 (0.38–1.96)	0.015
II	228 (51.2)	0.98 (0.21–4.99)	
III	159 (35.7)	0.91 (0.24–2.60)	
Lymph node status^e			
0	119 (26.3)	1.06 (0.29–4.99)	0.33 (NS)
1–3	237 (52.3)	0.98 (0.21–3.18)	
> 3	97 (21.4)	0.92 (0.24–2.43)	
Macroscopic tumour size^f			
≤ 25 mm	222 (49.8)	1.02 (0.21–4.99)	0.022
> 25 mm	224 (50.2)	0.90 (0.24–3.18)	
ERα status			
Negative	116 (25.6)	1.0 (0.28–4.99)	0.11 (NS)
Positive	338 (74.4)	0.96 (0.21–3.18)	
PR status			
Negative	192 (42.3)	0.95 (0.21–4.99)	0.58 (NS)
Positive	262 (57.7)	0.99 (0.24–3.18)	
ERBB2 status			
Negative	357 (78.6)	0.99 (0.21–4.99)	0.82 (NS)
Positive	97 (21.4)	0.94 (0.28–3.18)	
Molecular subtypes			
HR – ERBB2 –	68 (15.0)	1.02 (0.40–4.99)	0.25 (NS)
HR – ERBB2 +	43 (9.5)	1.06 (0.28–2.63)	
HR + ERBB2 –	289 (63.7)	0.98 (0.21–2.43)	
HR + ERBB2 +	54 (11.9)	0.87 (0.41–3.18)	

Abbreviations: ATM = ataxia telangiectasia mutated; ER = oestrogen receptor; ERBB2 = human epidermal growth factor receptor 2; HR = hormone receptor; NS = not significant; PR = progesterone receptor; SBR = Scarff–Bloom–Richardson. The bold values are statistically significant ($P < 0.05$).

^aKruskal–Wallis H test.
^bMedian (range).
^cScarff–Bloom–Richardson classification.
^dInformation available for 445 patients.
^eInformation available for 453 patients.
^fInformation available for 446 patients.

($P = 0.015$), and macroscopic tumour size ($P = 0.022$) (Table 1). Using HR (ERα and PR) and ERBB2 status, we subdivided our total population ($n = 454$) into four subgroups: HR + /ERBB2 + ($n = 54$), HR + /ERBB2 – ($n = 289$), HR – /ERBB2 + ($n = 43$), and HR – /ERBB2 – ($n = 68$). No correlation was observed between the subgroups and ATM mRNA levels (Table 1). We also tested the possible relationship between ATM and Ki67 expression, also determined for the same tumours. No correlation was observed between ATM and Ki67 mRNA levels (Spearman test: $r = -0.055$, $P = 0.24$), suggesting no link between ATM expression and cell proliferation.

We assessed the impact of variations of ATM mRNA levels on patient outcome by studying survival curves. AUC analyses were performed to identify a cut point, which divides the cohort into relevant ATM expression subgroups. Compared with tumours expressing high ATM mRNA levels (> 1.30 , $n = 123$, 27.1%), tumours with low ATM mRNA expression (≤ 1.30 , $n = 331$, 72.9%) were significantly associated with poor MFS ($P = 0.00012$; Figure 1A). Patients with lower ATM expression had a 5-year MFS and a 10-year MFS of $70.4 \pm 2.5\%$ and $59.0 \pm 2.8\%$, respectively, and patients with higher ATM expression had a 5-year MFS and a 10-year MFS of $85.7 \pm 3.2\%$ and $78.7 \pm 3.9\%$, respectively.

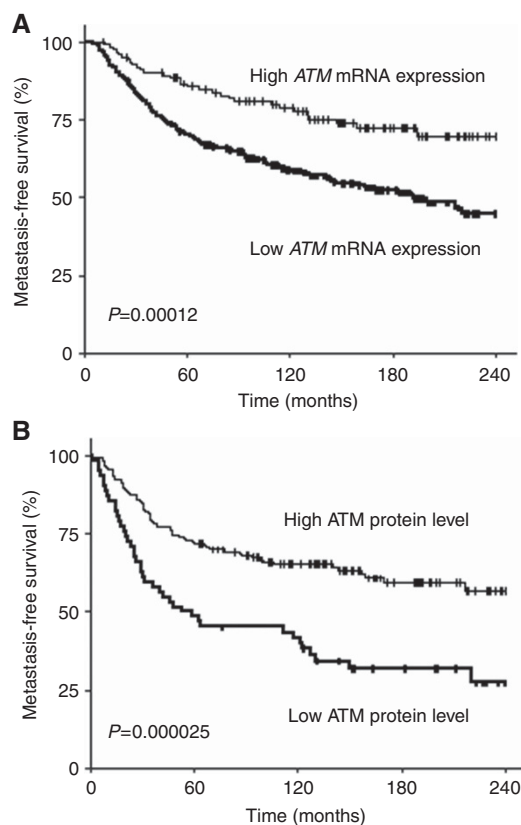


Figure 1. Survival curves of two patient groups according to ATM mRNA and protein levels. (A) MFS curves according to ATM mRNA level in 454 breast cancers. (B) MFS curves according to ATM protein level in 233 breast cancers.

Multivariate analysis using a Cox proportional hazards model assessed the predictive value for MFS of the parameters found to be significant on univariate analysis, that is, SBR histological grade, lymph node status, macroscopic tumour size, PR status (Supplementary Table S1) and ATM mRNA level (Figure 1A). The prognostic significance of lymph node status ($P = 0.000023$), macroscopic tumour size ($P = 0.0026$), PR status ($P = 0.004$) and ATM mRNA level ($P = 0.00046$) persisted (Table 2A). This analysis confirmed that the ATM mRNA level is an independent prognostic factor.

As ATM is a minor susceptibility gene for breast cancer, we wondered whether ATM alteration was observed more frequently among young patients in our breast cancer series. We observed a significant difference of ATM mRNA level between patients age < 37 years ($n = 11$; median = 0.55) and breast cancer patients aged > 37 years ($n = 443$; median = 0.98) ($P = 0.020$), and a less significant difference was observed at the 50-year cutoff ($P = 0.042$; Table 1).

Relationship between ATM protein level in breast tumours and classical clinicopathological parameters and patient outcome. We assessed the ATM protein level using RPPA in 233 tumour samples, including 144 of the 454 samples on which the mRNA expression level was also determined. A positive correlation was observed between ATM mRNA level and ATM protein level (Spearman test: $r = +0.297$, $P = 0.00040$), supporting our data concerning ATM mRNA levels and also suggesting that dysregulation of ATM in breast cancer is in part transcriptional.

The characteristics of the 233 tumours tested by RPPA are shown in Supplementary Table S3. ERBB2 status was available for

Table 2A. Results of Cox multivariate analysis: multivariate analysis of MFS in 454 patients (RNA study)

Characteristics	HR	95% CI	P-value ^a
Lymph node status			
0	1		0.000023
1–3	1.63	1.30–2.03	
>3	2.64	1.69–4.14	
Macroscopic tumour size			
≤25 mm	1		0.0026
>25 mm	1.60	1.18–2.16	
SBR histological grade			
I	1		(0.30) NS
II	1.14	0.89–1.44	
III	1.29	0.80–2.08	
PR status			
Negative	1		0.004
Positive	0.64	0.47–0.87	
ATM mRNA level			
≤1.30	1		0.00046
>1.30	0.49	0.33–0.73	

Abbreviations: CI = confidence interval; HR = hazard ratio; MFS = metastasis-free survival; NS = not significant. The bold values are statistically significant ($P < 0.05$).
^aMultivariate Cox analysis.

Table 2B. Results of Cox multivariate analysis: multivariate analysis of MFS in 233 patients (protein study)

Characteristics	HR	95% CI	P-value
Lymph node status			
0	1		0.00031
1–3	1.63	1.26–2.12	
>3	2.66	1.58–4.49	
Macroscopic tumour size			
≤25 mm	1		0.0063
>25 mm	1.76	1.16–2.66	
ATM protein level			
≤0.81	1		0.00037
>0.81	0.50	0.34–0.73	

Abbreviations: CI = confidence interval; HR = hazard ratio; MFS = metastasis-free survival; NS = not significant. The bold values are statistically significant ($P < 0.05$).

only 152 of these old samples, and these samples therefore could not be subdivided into HR/ERBB2 subgroups. No significant correlation was observed between ATM protein level and classical prognostic factors (Supplementary Table S4). AUC analyses identified a cut point that divided the cohort into relevant ATM protein level subgroups. The log-rank test showed a highly significant difference in MFS between patients with low (≤ 0.81 , $n = 62$, 26.6%) and high (> 0.81 , $n = 171$, 73.4%) ATM protein levels ($P = 0.000025$; Figure 1B). Patients with lower ATM expression had a poorer prognosis (5-year MFS: $48.4 \pm 6.3\%$; 10-year MFS: $41.7 \pm 6.3\%$) than those with higher ATM expression (5-year MFS: $71.8 \pm 3.5\%$; 10-year MFS: $65.2 \pm 3.7\%$). Multivariate analysis showed that ATM protein level, like the ATM mRNA level, was an independent prognostic factor ($P = 0.00037$; Table 2B).

Relationship between ATM protein level and other proteins of the DSB repair pathway. In order to study the possible implication of ATM in various well-known cancer pathways, we also used RPPA to study 13 additional proteins involved in the DSB repair pathway (H2AX, XRCC5, NBN, Mre11, Rad50 and TP53BP1), the epithelial–mesenchymal transition pathway (E-cadherin and Vimentin), two oncogenic pathways (p44/42-MAPK, P-p44/42-MAPK, AKT, P-AKT, mTOR, P-mTOR, MEK1/2 and P-MEK1/2) and cell proliferation (Ki67).

Significant correlations were observed only with proteins of the DSB repair pathway but not with proteins of the other classically dysregulated cancer pathways, as ATM was positively correlated with H2AX ($P < 0.0000001$), XRCC5 ($P < 0.0000001$), NBN ($P < 0.0000001$), Mre11 ($P = 0.0000029$), Rad50 ($P = 0.0064$) and TP53BP1 ($P = 0.026$) (Table 3).

As for ATM, the prognostic values of the DSB repair pathway proteins were also evaluated in this series of 233 breast tumour proteins (Table 3). No prognostic marker more discriminant than ATM ($P = 0.000025$) was identified, highlighting the preponderant role of ATM in this pathway.

Relationship between ATM protein level and the expression status of 12 miRNAs. To investigate whether dysregulation of the ATM gene was triggered by miRNAs, we tested the possible negative correlation between ATM protein level and 12 candidate

miRNAs levels described in the literature as posttranscriptional regulators of the ATM gene (Ng *et al*, 2010; Yan *et al*, 2010; Song *et al*, 2011; Lin *et al*, 2012; Bisso *et al*, 2013; Di Francesco *et al*, 2013; Guo *et al*, 2013, 2014; Mansour *et al*, 2013; Liang *et al*, 2014; Zhang *et al*, 2014; Zhou *et al*, 2014). *MiR-18a*, *miR-26a*, *miR-26b*, *miR-27a*, *miR-100*, *miR-101*, *miR-106a*, *miR-181a2*, *miR-181b1*, *miR-203*, *miR-223* and *miR-421* levels were analysed in 15 low ATM-expressing (marked low levels of both mRNA and protein) and 15 high ATM-expressing (marked high levels of both mRNA and protein) breast tumours (Table 4). A negative correlation was observed only between ATM and *miR-203* expression status, as the median *miR-203* value was 0.41 in high ATM-expressing breast tumours compared with the normalised value of 1 in low ATM-expressing breast tumours ($P = 0.011$). It should be noted that three miRNAs (*miR-26a*, *miR-26b* and *miR-101*) showed a positive correlation with the ATM protein level.

DISCUSSION

In breast cancer, ATM downregulation has been described at both the mRNA and protein levels and has been associated with poor outcome. In order to confirm the results of previous studies in our large cohort of breast cancers, we tested 10 normal breast tissue RNAs and 454 unilateral invasive primary breast tumour RNAs, using the qRT-PCR method.

A significant correlation was observed between a low level of ATM mRNA and histopathological grade III tumours ($P = 0.015$) and high macroscopic tumour size ($P = 0.022$), suggesting that a low level of ATM has a role in breast tumour aggressiveness and confirming the tumour-suppressor role of ATM.

It should be noted that ATM gene is a breast cancer susceptibility gene associated with moderate risk (Thompson *et al*, 2005; Renwick *et al*, 2006). We observed a higher frequency of ATM mRNA dysregulation in young breast cancer patients, supporting the idea that ATM dysregulation is a major molecular step in tumorigenesis in young breast cancer patients.

Survival analysis revealed that patients with low ATM mRNA levels had shorter MFS ($P = 0.00012$), as previously described (Ye *et al*, 2007). Multivariate analysis showed that ATM mRNA status was an independent prognostic factor ($P = 0.00046$). The positive correlation ($P = 0.00040$) observed between ATM mRNA and ATM protein levels assessed by RPPA in 233 breast tumours indicated that regulation of ATM is in part transcriptional. A low level of ATM protein was correlated with poor outcome ($P = 0.000025$). ATM protein level was also an independent prognostic factor ($P = 0.00037$), as described in a recent study (Bueno *et al*, 2014).

Table 3. Relationship between ATM protein level and other proteins involved in various well-known signalling pathways that are altered in cancer

Protein	Correlation with ATM protein level ^a	Prognostic value ^b
DSB repair pathway		
H2AX	0.428 ^c <0.000001 ^d	0.0036
XRCC5	0.426 <0.000001	0.083
NBN	0.408 <0.000001	0.033
Mre11	0.317 0.0000029	0.045
Rad50	0.178 0.0064	0.032
TP53BP1	0.145 0.026	0.039
EMT		
Vimentin	0.019 0.78 (NS)	
E-cadherin	0.009 0.89 (NS)	
Cell proliferation		
Ki67	0.023 0.72 (NS)	
MAPK pathway		
p44/42-MAPK	-0.151 0.02	
P-p44/42-MAPK	-0.066 0.32 NS	
MEK1/2	0.101 0.12 (NS)	
P-MEK1/2	0.013 0.83 (NS)	
AKT/mTOR pathway		
AKT	0.048 0.47 (NS)	
P-AKT	0.055 0.41 (NS)	
mTOR	-0.012 0.85 (NS)	
P-mTOR	-0.03 0.65 (NS)	

Abbreviations: ATM = ataxia telangiectasia mutated; DSB = double-strand break; EMT = epithelial-mesenchymal transition; MAPK = mitogen-activated protein kinase; mTOR = mammalian target of rapamycin; NS = not significant. The bold values are statistically significant ($P < 0.05$).
^aSpearman rank correlation test.
^bLog-rank test.
^cSpearman correlation coefficient.
^dP-value of Spearman rank correlation test.

These results are in agreement with studies describing down-regulation of ATM mRNA (Ye *et al*, 2007; Salimi *et al*, 2012; Bueno *et al*, 2014) or protein (Tommiska *et al*, 2008; Bueno *et al*, 2014) in breast cancer. We also demonstrated the independent prognostic value of ATM mRNA level.

The mechanisms of dysregulation of ATM in breast cancers have not been fully elucidated. Posttranscriptional regulation of ATM mediated by miRNAs has been described in many cancers, including breast cancers. In order to investigate this hypothesis, we selected 12 miRNAs with luciferase studies from the literature, demonstrating their activity on ATM mRNA (Ng *et al*, 2010; Yan *et al*, 2010; Song *et al*, 2011; Lin *et al*, 2012; Bisso *et al*, 2013; Di Francesco *et al*, 2013; Guo *et al*, 2013; Mansour *et al*, 2013; Guo *et al*, 2014; Liang *et al*, 2014; Zhang *et al*, 2014; Zhou *et al*, 2014). A negative correlation was observed only between miR-203 level and ATM protein level ($P = 0.011$). Regulation of ATM by miR-203 has only been described in colorectal cancer. Our results

Table 4. Relationships between ATM protein level and 12 candidate miRNAs

miRNA	Breast cancer with low level of ATM expression (n = 15)	Breast cancer with high level of ATM expression (n = 15)	P-value ^a	ROC—AUC ^b
miR-203	1 (0.19–7.89) ^c	0.41 (0.05–3.96)	0.011	0.773
miR-421	1 (0.34–7.39)	0.57 (0.33–6.76)	0.056 (NS)	0.704
miR-27A	1 (0.33–3.27)	0.54 (0.24–2.25)	0.078 (NS)	0.689
miR-181b1	1 (0.25–3.88)	0.61 (0.24–2.35)	0.32 (NS)	0.607
miR-181a2	1 (0.33–3.06)	0.78 (0.25–2.89)	0.35 (NS)	0.600
miR-18A	1 (0.11–13.60)	0.73 (0.26–3.72)	0.44 (NS)	0.582
miR-106a	1 (0.24–8.97)	1.01 (0.29–2.53)	0.8 (NS)	0.527
miR-100	1 (0.19–30.36)	1.43 (0.35–3.72)	0.31 (NS)	0.391
miR-223	1 (0.25–24.92)	1.94 (0.15–14.60)	0.14 (NS)	0.340
miR-101	1 (0.12–73.62)	3.42 (1.48–12.82)	0.01	0.224
miR-26b	1 (0.32–42.05)	3.05 (0.75–9.44)	0.0095	0.222
miR-26a	1 (0.35–22.40)	3.56 (1.30–12.45)	0.0045	0.196

Abbreviations: ATM = ataxia telangiectasia mutated; miRNA = microRNA; NS = not significant. The bold values are statistically significant ($P < 0.05$).
^aKruskal–Wallis H test.
^bROC (receiver operating characteristics)—AUC (area under curve) analysis.
^cMedian (range) miRNA levels.

suggest that miR-203 could be partly involved in dysregulation of ATM in breast cancers. Further investigations are necessary to demonstrate the interaction between miR-203 and ATM protein level.

ATM is widely known to be a central player in cell proliferation and DNA-damage response (DDR). RPPA assay was used to study correlations between ATM protein level and other proteins of the DSB repair pathway, cell proliferation pathway and other pathways (EMT, MAPK and AKT/mTOR). Our results showed a relationship between ATM protein dysregulation and six other proteins of the DSB repair pathway, that is, H2AX, XRCC5, NBN, Mre11, Rad50 and TP53BP1. Conversely, no correlation was observed between ATM expression and proliferation, as demonstrated by Ki67 at mRNA and protein levels. These results confirm that ATM dysregulation has a major role in DSB repair pathway defect but not in cell proliferation. Our results suggest that, in addition to the well-known fact that reduction of ATM levels acts by altering the formation of DSB repair foci, ATM expression seems also linked to expression levels of several major actors of the DSB repair pathway. This DNA repair defect may be associated with homologous recombination DNA repair defect called ‘BRCAness’, particularly concerning the sensitivity to PARP inhibitors. PARP inhibitors have been shown to be highly lethal to tumour cells with deficiencies of DDR factors, such as BRCA1 or BRCA2. Their activity is based on the principle of synthetic lethality, which consists of targeting two separate molecular pathways that are nonlethal when disrupted individually but are lethal when inhibited simultaneously. The sensitivity of cancers with ATM downregulation to PARP inhibitors has been studied in breast and other cancers, with promising results (Williamson *et al*, 2010; Gilardini Montani *et al*, 2013; Kubota *et al*, 2014).

In conclusion, our results show that ATM is an outstanding prognostic marker in breast cancer and that ATM downregulation (both at the transcriptional and posttranscriptional levels) leads to dysregulation of the DSB repair pathway, which is probably involved in breast tumorigenesis. This could predict a favourable response to treatment with PARP inhibitors.

ACKNOWLEDGEMENTS

We thank the staff of Institut Curie—René Huguenin Hospital for their assistance in specimen collection and patient care. We thank Aurélie Barbet, Caroline Lecerf and Floriane Bard for performing the RPPA experiments. This work was supported by the Comité Départemental des Hauts-de-Seine de la Ligue Nationale Contre le Cancer and the Association pour la Recherche en Cancérologie de Saint-Cloud (ARCS). The RPPA platform is supported by the Cancéropôle Ile-de-France.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- Akbani R, Becker KF, Carragher N, Goldstein T, de Koning L, Korf U, Liotta L, Mills GB, Nishizuka SS, Pawlak M, Petricoin 3rd EF, Pollard HB, Serrels B, Zhu J (2014) Realizing the promise of reverse phase protein arrays for clinical, translational, and basic research: a workshop report: the RPPA (reverse phase protein array) society. *Mol Cell Proteomics* **13**: 1625–1643.
- Beggs AD, Domingo E, McGregor M, Presz M, Johnstone E, Midgley R, Kerr D, Oukrif D, Novelli M, Abulafi M, Hodgson SV, Fadhil W, Ilyas M, Tomlinson IP (2012) Loss of expression of the double strand break repair protein ATM is associated with worse prognosis in colorectal cancer and loss of Ku70 expression is associated with CIN. *Oncotarget* **3**: 1348–1355.
- Bieche I, Onody P, Laurendeau I, Olivi M, Vidaud D, Lidereau R, Vidaud M (1999) Real-time reverse transcription-PCR assay for future management of ERBB2-based clinical applications. *Clin Chem* **45**: 1148–1156.
- Bieche I, Parfait B, Laurendeau I, Girault I, Vidaud M, Lidereau R (2001a) Quantification of estrogen receptor alpha and beta expression in sporadic breast cancer. *Oncogene* **20**: 8109–8115.
- Bieche I, Parfait B, Le Doussal G, Olivi M, Rio MC, Lidereau R, Vidaud M (2001b) Identification of CGA as a novel estrogen receptor-responsive gene in breast cancer: an outstanding candidate marker to predict the response to endocrine therapy. *Cancer Res* **61**: 1652–1658.
- Bieche I, Tozlu S, Onody P, Driouch K, Vidaud M, Lidereau R (2004) Expression of PEA3/E1AF/ETV4, an Ets-related transcription factor, in breast tumors: evidence of MMP2, NRG1 and CGB as major transactivated target genes. *Carcinogenesis* **25**: 405–411.
- Bisso A, Faleschini M, Zampa F, Capaci V, De Santa J, Santarpia L, Piazza S, Cappelletti V, Daidone M, Agami R, Del Sal G (2013) Oncogenic miR-181a/b affect the DNA damage response in aggressive breast cancer. *Cell Cycle* **12**: 1679–1687.
- Bueno RC, Canevari RA, Villacis RA, Domingues MA, Caldeira JR, Rocha RM, Drigo SA, Rogatto SR (2014) ATM down-regulation is associated with poor prognosis in sporadic breast carcinomas. *Ann Oncol* **25**: 69–75.
- Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, Antipin Y, Reva B, Goldberg AP, Sander C, Schultz N (2012) The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* **2**: 401–404.
- Di Francesco A, De Pitta C, Moret F, Barbieri V, Celotti L, Mognato M (2013) The DNA-damage response to gamma-radiation is affected by miR-27a in A549 cells. *Int J Mol Sci* **14**: 17881–17896.
- Finak G, Sadekova S, Pepin F, Hallett M, Meterisian S, Halwani F, Khetani K, Souleimanova M, Zabolotny B, Omeroglu A, Park M (2006) Gene expression signatures of morphologically normal breast tissue identify basal-like tumors. *Breast Cancer Res* **8**: R58.
- Flanagan JM, Munoz-Alegre M, Henderson S, Tang T, Sun P, Johnson N, Fletcher O, Dos Santos Silva I, Peto J, Boshoff C, Narod S, Petronis A (2009) Gene-body hypermethylation of ATM in peripheral blood DNA of bilateral breast cancer patients. *Hum Mol Genet* **18**: 1332–1342.
- Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E, Cerami E, Sander C, Schultz N (2013) Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* **6**: p11.
- Gilardini Montani MS, Prodosmo A, Stagni V, Merli D, Monteonofrio L, Gatti V, Gentileschi MP, Barila D, Soddu S (2013) ATM-depletion in breast cancer cells confers sensitivity to PARP inhibition. *J Exp Clin Cancer Res* **32**: 95.
- Grabsch H, Dattani M, Barker L, Maughan N, Maude K, Hansen O, Gabbert HE, Quirke P, Mueller W (2006) Expression of DNA double-strand break repair proteins ATM and BRCA1 predicts survival in colorectal cancer. *Clin Cancer Res* **12**: 1494–1500.
- Guo P, Lan J, Ge J, Nie Q, Guo L, Qiu Y, Mao Q (2014) MiR-26a enhances the radiosensitivity of glioblastoma multiforme cells through targeting of ataxia-telangiectasia mutated. *Exp Cell Res* **320**: 200–208.
- Guo X, Yang C, Qian X, Lei T, Li Y, Shen H, Fu L, Xu B (2013) Estrogen receptor alpha regulates ATM Expression through miRNAs in breast cancer. *Clin Cancer Res* **19**: 4994–5002.
- Hanley JA, McNeil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* **143**: 29–36.
- Hu H, Du L, Nagabayashi G, Seeger RC, Gatti RA (2010) ATM is down-regulated by N-Myc-regulated microRNA-421. *Proc Natl Acad Sci USA* **107**: 1506–1511.
- Kim H, Saka B, Knight S, Borges M, Childs E, Klein A, Wolfgang C, Herman J, Adsay VN, Hruban RH, Goggins M (2014) Having pancreatic cancer with tumoral loss of ATM and normal TP53 protein expression is associated with a poorer prognosis. *Clin Cancer Res* **20**: 1865–1872.
- Kubota E, Williamson CT, Ye R, Elegbede A, Peterson L, Lees-Miller SP, Bebb DG (2014) Low ATM protein expression and depletion of p53 correlates with olaparib sensitivity in gastric cancer cell lines. *Cell Cycle* **13**: 2129–2137.
- Lee KW, Tsai YS, Chiang FY, Huang JL, Ho KY, Yang YH, Kuo WR, Chen MK, Lin CS (2011) Lower ataxia telangiectasia mutated (ATM) mRNA expression is correlated with poor outcome of laryngeal and pharyngeal cancer patients. *Ann Oncol* **22**: 1088–1093.
- Liang L, Zhu J, Zaorsky NG, Deng Y, Wu X, Liu Y, Liu F, Cai G, Gu W, Shen L, Zhang Z (2014) MicroRNA-223 enhances radiation sensitivity of U87MG cells in vitro and in vivo by targeting ataxia telangiectasia mutated. *Int J Radiat Oncol Biol Phys* **88**: 955–960.
- Lin F, Li R, Pan ZX, Zhou B, Yu de B, Wang XG, Ma XS, Han J, Shen M, Liu HL (2012) miR-26b promotes granulosa cell apoptosis by targeting ATM during follicular atresia in porcine ovary. *PLoS One* **7**: e38640.
- Mansour WY, Bogdanova NV, Kasten-Pisula U, Rieckmann T, Kocher S, Borgmann K, Baumann M, Krause M, Petersen C, Hu H, Gatti RA, Dikomey E, Dork T, Dahm-Daphi J (2013) Aberrant overexpression of miR-421 downregulates ATM and leads to a pronounced DSB repair defect and clinical hypersensitivity in SKX squamous cell carcinoma. *Radiother Oncol* **106**: 147–154.
- Ng WL, Yan D, Zhang X, Mo YY, Wang Y (2010) Over-expression of miR-100 is responsible for the low-expression of ATM in the human glioma cell line: M059J. *DNA Repair (Amst)* **9**: 1170–1175.
- Pal R, Srivastava N, Chopra R, Gochhait S, Gupta P, Prakash N, Agarwal G, Bamezai RN (2010) Investigation of DNA damage response and apoptotic gene methylation pattern in sporadic breast tumors using high throughput quantitative DNA methylation analysis technology. *Mol Cancer* **9**: 303.
- Renwick A, Thompson D, Seal S, Kelly P, Chagtai T, Ahmed M, North B, Jayatilake H, Barfoot R, Spanova K, McGuffog L, Evans DG, Eccles D, Easton DF, Stratton MR, Rahman N (2006) ATM mutations that cause ataxia-telangiectasia are breast cancer susceptibility alleles. *Nat Genet* **38**: 873–875.
- Salimi M, Mozdarani H, Majidzadeh K (2012) Expression pattern of ATM and cyclin D1 in ductal carcinoma, normal adjacent and normal breast tissues of Iranian breast cancer patients. *Med Oncol* **29**: 1502–1509.
- Shiloh Y, Ziv Y (2013) The ATM protein kinase: regulating the cellular response to genotoxic stress, and more. *Nat Rev Mol Cell Biol* **14**: 197–210.
- Song L, Lin C, Wu Z, Gong H, Zeng Y, Wu J, Li M, Li J (2011) miR-18a impairs DNA damage response through downregulation of ataxia telangiectasia mutated (ATM) kinase. *PLoS One* **6**: e25454.
- Stagni V, Oropallo V, Fianco G, Antonelli M, Cina I, Barila D (2014) Tug of war between survival and death: exploring ATM function in cancer. *Int J Mol Sci* **15**: 5388–5409.
- The Cancer Genome Atlas Network (2012) Comprehensive molecular portraits of human breast tumors. *Nature* **490**: 61–70.

- Thompson D, Duedal S, Kirner J, McGuffog L, Last J, Reiman A, Byrd P, Taylor M, Easton DF (2005) Cancer risks and mortality in heterozygous ATM mutation carriers. *J Natl Cancer Inst* **97**: 813–822.
- Tommiska J, Bartkova J, Heinonen M, Hautala L, Kilpivaara O, Eerola H, Aittomaki K, Hofstetter B, Lukas J, von Smitten K, Blomqvist C, Ristimaki A, Heikkila P, Bartek J, Nevanlinna H (2008) The DNA damage signalling kinase ATM is aberrantly reduced or lost in BRCA1/BRCA2-deficient and ER/PR/ERBB2-triple-negative breast cancer. *Oncogene* **27**: 2501–2506.
- Treilleux I, Chapot B, Goddard S, Pisani P, Angele S, Hall J (2007) The molecular causes of low ATM protein expression in breast carcinoma; promoter methylation and levels of the catalytic subunit of DNA-dependent protein kinase. *Histopathology* **51**: 63–69.
- Troncale S, Barbet A, Coulibaly L, Henry E, He B, Barillot E, Dubois T, Hupe P, de Koning L (2012) NormaCurve: a SuperCurve-based method that simultaneously quantifies and normalizes reverse phase protein array data. *PLoS One* **7**: e38686.
- Vo QN, Kim WJ, Cvitanovic L, Boudreau DA, Ginzinger DG, Brown KD (2004) The ATM gene is a target for epigenetic silencing in locally advanced breast cancer. *Oncogene* **23**: 9432–9437.
- Williamson CT, Muzik H, Turhan AG, Zamo A, O'Connor MJ, Bebb DG, Lees-Miller SP (2010) ATM deficiency sensitizes mantle cell lymphoma cells to poly(ADP-ribose) polymerase-1 inhibitors. *Mol Cancer Ther* **9**: 347–357.
- Yan D, Ng WL, Zhang X, Wang P, Zhang Z, Mo YY, Mao H, Hao C, Olson JJ, Curran WJ, Wang Y (2010) Targeting DNA-PKcs and ATM with miR-101 sensitizes tumors to radiation. *PLoS One* **5**: e11397.
- Ye C, Cai Q, Dai Q, Shu XO, Shin A, Gao YT, Zheng W (2007) Expression patterns of the ATM gene in mammary tissues and their associations with breast cancer survival. *Cancer* **109**: 1729–1735.
- Zhang X, Nie Y, Li X, Wu G, Huang Q, Cao J, Du Y, Li J, Deng R, Huang D, Chen B, Li S, Wei B (2014) MicroRNA-181a functions as an oncomir in gastric cancer by targeting the tumour suppressor gene ATM. *Pathol Oncol Res*; e-pub ahead of print 16 February 2014.
- Zhou Y, Wan G, Spizzo R, Ivan C, Mathur R, Hu X, Ye X, Lu J, Fan F, Xia L, Calin GA, Ellis LM, Lu X (2014) miR-203 induces oxaliplatin resistance in colorectal cancer cells by negatively regulating ATM kinase. *Mol Oncol* **8**: 83–92.

This work is published under the standard license to publish agreement. After 12 months the work will become freely available and the license terms will switch to a Creative Commons Attribution-NonCommercial-Share Alike 4.0 Unported License.

Supplementary Information accompanies this paper on British Journal of Cancer website (<http://www.nature.com/bjc>)