Arora, Arvind and Agarwal, Devika and Abdel-Fatah, Tarek M.A. and Lu, Huiming and Croteau, Deborah L. and Moseley, Paul and Aleskandarany, Mohammed A. and Green, Andrew R. and Ball, Graham and Rakha, Emad A. and Chan, Stephen Y.T. and Ellis, Ian O. and Wang, Lisa L. and Zhao, Yongliang and Balajee, Adayabalam S. and Bohr, Vilhelm A. and Madhusudan, Srinivasan (2016) RECQL4 helicase has oncogenic potential in sporadic breast cancers. Journal of Pathology, 238 (4). pp. 495-501. ISSN 1096-9896

RECQL4 helicase has oncogenic potential in sporadic breast cancers

Arvind Arora[1,2*], Devika Agarwal[3*], Tarek MA Abdel-Fatah[2], Huiming Lu[4], Deborah L. Croteau[4], Paul Moseley[2], Mohammed A Aleskandarany[5], Andrew R Green[5], Graham Ball[3], Emad A Rakha[5], Stephen YT Chan[2], Ian O Ellis[5], Lisa L Wang[6], Yongliang Zhao[7], Adayabalam S. Balajee[8], Vilhelm A. Bohr[4] and Srinivasan Madhusudan[1,2*]

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

[2] Department of Oncology, Nottingham University Hospitals, Nottingham, UK.

[3] School of Science and Technology, Nottingham Trent University, Clifton campus, Nottingham UK.

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

[5] Department of Pathology, School of Medicine, University of Nottingham, Nottingham NG51PB, UK.

[6]Texas Children's Cancer Center, Baylor College of Medicine, Houston, Texas 77030, USA

[7]Laboratory of Disease Genomics and Individualized Medicine, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100101, China.

[8] REAC/TS, Oak Ridge Associated Universities , Oak Ridge Institute for Science and Education, 1299 Bethel Valley Road, Building SC-200, Oak Ridge, Tennesse 37830, USA.

* AA and DA equally contributed to the work and should be considered as joint first authors

Running title: RECQL4 in breast cancer

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

Word count: 1497

*** *Correspondence to*:**

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

RECQL4 helicase is a molecular motor that unwinds DNA, a process essential during DNA replication and DNA repair. Germ-line mutations in *RECQL4* cause type II Rothmund-Thomson syndrome (RTS) characterised by a premature aging phenotype and cancer predisposition. *RECQL4* is widely considered as a tumour suppressor, although its role in human breast cancer is largely unknown. As the *RECQL4* gene is localized to chromosome 8q24, a site frequently amplified in sporadic breast cancers, we hypothesised that it may play an oncogenic role in breast tumorigenesis. To address this we analysed large cohorts for gene copy number changes (n=1977), mRNA expression (n=1977) and protein level (n=1902). Breast cancer incidence was also explored in 58 patients with type II RTS. DNA replication dynamics and chemo-sensitivity was evaluated in RECQL4-depleted breast cancer cells *in vitro*. Amplification or gain in gene copy number (30.6%), high level mRNA expression (51%) and high levels of protein (23%) significantly associated with aggressive tumour behaviour including lymph node positivity, larger tumour size, HER2 over expression, ER-negativity, triple negative phenotypes and poor survival. RECQL4 depletion impaired DNA replication rate and increased chemo-sensitivity in cultured breast cancer cells. Thus, although recognised as a "safe guardian of the genome", our data provides compelling evidence that RECQL4 is tumour promoting in established breast cancers.

Key words: RECQL4 helicase; breast cancer; tumour suppressor; oncogene.

INTRODUCTION

DNA helicases are molecular motors that unwind DNA, an essential process required during DNA replication and DNA repair. RecQ Protein-Like 4 (RECQL4) is a key member of the RecQ family of DNA helicases and plays an important role in the maintenance of genomic stability [1-3]. RECQL4 has a role in the initiation of DNA replication, progression of stalled replication forks, telomere maintenance and in repair of DNA double-strand breaks (DSBs) via the homologous recombination (HR) pathway [1-3]. Mutations in the *RECQL4* gene are found in about two-thirds of all cases of Rothmund-Thomson syndrome (RTS), and these patients are designated as having Type II RTS [4, 5]. RTS is characterised by a premature aging and predisposition to cancers, especially lymphomas and osteosarcomas [4, 5]. A tumour suppressor function of RECQL4 has been widely described, although recent evidence also suggests a tumour-promoting role for RECQL4. In preclinical studies, we have recently found overexpression of RECQL4 in prostate cancer cell lines, and depletion of RECQL4 by siRNA or shRNA vectors significantly reduced the growth and survival of metastatic prostate cancer cells [6]. Similarly, in breast cancer cell lines, we have observed overexpression of RECQL4, and found that depletion of RECQL4 promoted apoptosis [7]. Interestingly, the *RECQL4* gene is localized to chromosome 8q24, a site frequently amplified in sporadic breast cancers [8-10]. We therefore hypothesised a tumour-promoting role for RECQL4 in breast cancers.

MATERIALS AND METHODS

RECQL4 gene copy number changes and mRNA levels: The METABRIC (Molecular Taxonomy of Breast Cancer International Consortium) cohort [11] was evaluated for *RECQL4* gene copy number changes and mRNA levels. Patient demographics are summarized in supplementary Table S1 and full methods are discussed in Supplementary Methods online.

RECQL4 protein expression in breast cancer: The study was performed in two cohorts of breast cancers. The first cohort was a consecutive series of 1,650 patients with primary invasive breast carcinomas who were diagnosed between 1986 and 1999. The second cohort was an independent series of 252 ER-α negative invasive breast cancer patients diagnosed and managed at the Nottingham University Hospitals between 1999 and 2007. Immunohistochemical evaluation of RECQL4 is summarized fully in Supplementary methods and Supplementary Table S6. Tumour Marker Prognostic Studies (REMARK) criteria, recommended by McShane et al [12], were followed throughout this study. Ethical approval was obtained from the Nottingham Research Ethics Committee (C202313).

Breast cancer assessment in RTS patients: Incidence of breast cancer was evaluated among a cohort of RTS patients enrolled in a longitudinal clinical research study approved by the Institutional Review Board for Human Subjects Research at Baylor College of Medicine (Houston TX). All subjects or parents provided informed written consent to participate in the study. Clinical information was updated by yearly questionnaires. RTS patients with known mutations in the *RECQL4* gene (Type II RTS) were included in the present study (n=58). Patient demographics are shown in Supplementary Table S7.

Cell line studies: HeLa, MCF7, MDA-MB-231 and BT549 lines were obtained from the ATCC (Manassas, VA, USA). Detailed methodology for Western blotting and imunofluorescence is summarised in Supplementary methods. We generated transient RECQL4 knockdown as well as stable RECQL4 knockdowns in breast cancer cells as

described in Supplementary Methods. Cell numbers were estimated by the MTT assay according to the manufacturer's protocol. To evaluate replication dynamics, DNA fibre assays were performed as described previously [13]. Detailed methodology is also described in Supplementary Methods.

RESULTS AND DISCUSSION

We have recently shown that *RECQL4* gene amplification and elevated levels of RECQL4 expression are common in breast cancer cell lines [7]. Moreover, depletion of RECQL4 not only reduced breast cancer cell proliferation but also impaired tumourigenicity in tumour bearing mice [7]. These data therefore support that RECQL4 may be oncogenic and drive breast tumourigenesis. To test this hypothesis, we conducted a comprehensive clinical study.

***RECQL4* gene amplification or gain in copy number changes in breast cancers:** None of 1970 (0%) tumours had *RECQL4* homozygous deletion, 19/1970 (1%) of tumours had *RECQL4* heterozygous deletion, 1348/1970 (68.4%) of tumours had *RECQL4* neutral gene copy number, 543/1970 (27.6%) of tumours had gain in *RECQL4* copy number and 60/1970 (3%) of tumours had amplification of the *RECQL4* gene. We grouped gain/amplification and homozygous deletion/heterozygous/neutral together. As shown in Supplementary Figure 1A, ER- tumours were more likely to have gain/amplification of *RECQL4* compared to ER+ tumours ($p=0.0003$). Within the various molecular phenotype groups, compared to normal phenotype, tumours that were PAM50.Basal ($p<0.00001$), or PAM50.HER2 ($p<0.00001$) had significantly grreater gain/amplification of *RECQL4* (Figure 1A). Within the ER+ sub-group, PAM50.Luminal B sub-groups had significantly greater gain/amplification of *RECQL4* (Figure 1A) compared to PAM50.Luminal A sub-group ($p<0.00001$) (Figure 1A). As shown in Supplementary Table S5, high stage, grade 3 tumours and lymph node positivity were more common in tumours with gain/amplification of *RECQL4*. As expected, breast cancer specific survival (BCSS) was worse in tumours with gain/amplification of *RECQL4* compared to tumours with neutral changes or loss of *RECQL4* ($p<0.00001$) (Figure 1B).

High levels of *RECQL4* transcripts in breast cancers: 966/1977 (49%) of tumours had low *RECQL4* mRNA levels and 1011/1977 (51%) tumours had high *RECQL4* mRNA levels. ER-

tumours had higher *RECQL4* mRNA levels compared to ER+ tumours ($p<0.0001$) (Supplementary Figure 1B). Within the various molecular phenotype groups, compared to normal phenotype, tumours that were PAM50.Basal ($p<0.00001$), or PAM50.HER2 ($p<0.00001$) had high *RECQL4* mRNA levels (Figure 1C). High levels of *RECQL4* mRNA were highly significantly associated with aggressive clinicopathological features (Table 1) including high histological grade, lymph node positivity, larger tumour size, Nottingham prognostic index (NPI) >3.4, and triple negative phenotype (each, $p<0.001$). High *RECQL4* mRNA level was also found to be significantly associated with previously described molecular phenotypes in breast cancer: Genufu subtype (ER-/HER2-) ($p<0.00001$), Genufu subtype (ER+/HER2-/High proliferation) ($p<0.00001$) and Genufu subtype (HER2 positive) ($p = 0.001$) breast tumours. However, PAM50.Luminal A tumours and Genufu subtype (ER+/HER2-/low proliferation) were more likely to have low levels of *RECQL4* mRNA (each, $p<0.00001$). A high level of *RECQL4* mRNA in the tumour was associated with poor breast cancer specific survival (BCSS) ($p<0.00001$) (Figure 1D). In multivariate Cox regression analysis *RECQL4* mRNA levels remained independently associated with poor BCSS ($p<0.00001$) (Supplementary Table S6).

Mechanistic insights: As shown in Figure 2A, there was a strong correlation between gene copy number changes and mRNA levels ($p<0.00001$). The correlation remains significant across various sub-groups including in ER- ($p<0.0001$), ER+ ($p<0.0001$), PAM50.Basal ($p<0.0001$), PAM50. HER2 ($p<0.0001$), PAM50.Luminal A ($p<0.0001$) and PAM50.Luminal B tumours ($p<0.0001$) (Supplementary Figures S1C and S1D). Taken together the data supports that in a proportion of aggressive tumours, a high mRNA level is due to increased gene copy number.

RECQL4 protein level in breast cancers: The N-terminal region of RECQL4 contains the nuclear as well as mitochondrial targeting sequences and is important for sub-cellular

localisation of RECQL4 [1-3]. In addition, post-translational modification (such as acetylation) of RECQL4 may also alter its sub-cellular localisation [1-3]. As expected, we observed complex sub-cellular localisation of RECQL4 in human breast cancers including exclusively nuclear staining (17.6%), exclusively cytoplasmic staining (23.4%), nuclear-cytoplasmic co-expression (24.8%) or absence of staining (34.2%) (Supplementary Figure S1E). In 20 normal breast tissues, however, we observed exclusively nuclear staining in all samples and no cytoplasmic staining (Supplementary Figure S1E) implying that altered sub-cellular localisation is a feature of cancer and not normal tissue. Tumours with high cytoplasmic/low nuclear RECQL4 levels were significantly associated with high grade, high mitotic index, pleomorphism, NPI>3.4, ER-, and triple negative phenotype (all p values <0.01) (Supplementary Table S7) and poor survival (p=0.042) (Figure 2B). In multivariate analysis the RECQL4 protein level independently influenced survival (p=0.032) (Supplementary Table S8).

Breast cancer incidence in patients with Rothmund-Thomson Syndrome (RTS): Germline mutation in *RECQL4* is causal for two-thirds of patients with Type II RTS, and it has previously been shown that RECQL4 mutation status correlates with risk of developing osteosarcoma [4, 5]. In the largest available cohort of type II RTS patients, we did not observe any increased incidence of breast cancers. The data suggests that either RECQL4 deficiency does not influence breast cancer pathogenesis or that RTS patients have not lived long enough to develop breast cancer.

Depletion of RECQL4 significantly reduced DNA replication rates and increased sensitivity to chemotherapy: DNA synthesis rates were measured using a DNA fibre assay after BrdU incorporation in MCF7, MDA-MB-231 and BT549 breast cancer cell lines plus or minus control siRNA or siRNA to deplete RECQL4 (Supplementary Figure S2). We consistently observed shorter DNA fibre lengths after depletion of RECQL4 in the breast

cancer cells (Figure 2C and 2D). We then generated a stable RECQL4 knock down ER- (MDA-MB- 453) breast cancer cell line (Figure 3A). As shown in Figures 3B- 3D, RECQL4 depleted breast cancer cells were sensitive to treatment with cisplatin, doxorubicin or 5-FU.

In conclusion, we have demonstrated that high copy number, high mRNA levels and high protein levels of RECQL4 is associated with aggressive breast cancers. Although the data suggest to us that RECQL4 has oncogenic potential, it is also possible that RECQL4 overexpression may be a secondary event that may allow cancer cells to maintain high proliferation rate and telomere elongation required for cancer cell survival.

Acknowledgements: We would like to thank Ta-Tara Rideau, Research Coordinator at Texas Children's Cancer Center, Houston, TX for data collection on RTS patients, and Hao Liu, PhD at Baylor College of Medicine, Houston, TX for statistical support. We would like to thank Alfred May for assistance with the immunofluorescence studies of breast cancer cell lines. This work was supported by Funds from the Intramural Program of the National Institutes on Aging, National Institutes of Health, USA. We would also like to thank Roslin Russell and Carlos Caldas at Cancer Research UK, Cambridge Research Institute, Li Ka Shing Centre, Robinson Way, Cambridge CB2 0RE, UK for assisting with the analysis of RECQL4 gene copy number changes.

Author contributions: AA; IHC staining, data analysis, data interpretation, DA: mRNA and gene copy number analysis, data interpretation, TAF: IHC staining, data analysis, data interpretation, HU; siRNA of RECQL4, DNA fibre assay, data analysis, interpretation, DL; siRNA of RECQL4, DNA fibre assay, data analysis, interpretation, PM; IHC staining, MA; RECQL4 staining of normal tissue, data analysis and interpretation, ARG; data analysis and interpretation, GB; mRNA and gene copy number analysis, data interpretation, EAR; TMA construction, data analysis and interpretation, SYT; IHC staining, data analysis, data interpretation, IOE; TMA construction, data analysis and interpretation, LW; data collection and analysis of RTS patients, YZ; shRNA of RECQL4, chemo-sensitization studies, ASB; shRNA of RECQL4, chemo-sensitization studies, VB; siRNA of RECQL4, DNA fibre assay, data analysis, interpretation, SM; study design, data analysis and interpretation.
All authors contributed to writing the manuscript and approved the final version.

REFERENCES

1. Croteau DL, Singh DK, Hoh Ferrarelli L, et al. RECQL4 in genomic instability and aging. *Trends Genet* 2012; **28**: 624-631.
2. Dietschy T, Shevelev I,Stagljar I. The molecular role of the Rothmund-Thomson-, RAPADILINO- and Baller-Gerold-gene product, RECQL4: recent progress. *Cell Mol Life Sci* 2007; **64**: 796-802.
3. Larizza L, Magnani I,Roversi G. Rothmund-Thomson syndrome and RECQL4 defect: splitting and lumping. *Cancer Lett* 2006; **232**: 107-120.
4. Siitonen HA, Sotkasiira J, Biervliet M, et al. The mutation spectrum in RECQL4 diseases. *Eur J Hum Genet* 2009; **17**: 151-158.
5. Wang LL, Gannavarapu A, Kozinetz CA, et al. Association between osteosarcoma and deleterious mutations in the RECQL4 gene in Rothmund-Thomson syndrome. *J Natl Cancer Inst* 2003; **95**: 669-674.
6. Su Y, Meador JA, Calaf GM, et al. Human RecQL4 helicase plays critical roles in prostate carcinogenesis. *Cancer Res* 2010; **70**: 9207-9217.
7. Fang H, Nie L, Chi Z, et al. RecQL4 helicase amplification is involved in human breast tumorigenesis. *PLoS One* 2013; **8**: e69600.
8. Thomassen M, Tan Q,Kruse TA. Gene expression meta-analysis identifies chromosomal regions and candidate genes involved in breast cancer metastasis. *Breast Cancer Res Treat* 2009; **113**: 239-249.
9. Ioannidis P, Mahaira L, Papadopoulou A, et al. 8q24 Copy number gains and expression of the c-myc mRNA stabilizing protein CRD-BP in primary breast carcinomas. *Int J Cancer* 2003; **104**: 54-59.
10. Yokota T, Yoshimoto M, Akiyama F, et al. Frequent multiplication of chromosomal region 8q24.1 associated with aggressive histologic types of breast cancers. *Cancer Lett* 1999; **139**: 7-13.
11. Curtis C, Shah SP, Chin SF, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* 2012; **486**: 346-352.
12. McShane LM, Altman DG, Sauerbrei W, et al. Reporting recommendations for tumor marker prognostic studies (REMARK). *J Natl Cancer Inst* 2005; **97**: 1180-1184.
13. Techer H, Koundrioukoff S, Azar D, et al. Replication dynamics: biases and robustness of DNA fiber analysis. *J Mol Biol* 2013; **425**: 4845-4855.

Table 1: Association between *RECQL4* mRNA expression and clinicopathological variables in the METABRIC cohort.

VARIABLE	RECQL4 mRNA levels		P Value
	Low	High	
	N(%)	N (%)	
A) Pathological Parameters			
Lymph node involvement			
Negative	537 (55.9%)	498 (49.3%)	**0.003**
Positive (1-3)	129 (13.4%)	185 (18.3%)	
Positive (>3)	295 (30.7%)	327 (32.4%)	
Grade			
G1	137 (15.0%)	32 (3.3%)	**$2.4x10^{-47}$**
G2	466 (51.2%)	304 (31.1%)	
G3	308 (33.8%)	642 (65.6%)	
Tumour size (cm)			
T 1a+b(1.0)	58 (6.0%)	34 (3.4%)	**$1.3x10^{-5}$**
T 1c(>1.0-2.0)	413 (43.1%)	353 (35.4%)	
T2 (>2.0-5)	450 (46.9%)	551 (55.2%)	
T3 (>5)	38 (4.0%)	60 (6.0%)	
NPI			
≤ 3.4	298 (35.3%)	120 (12.4%)	**$7.2x10^{-27}$**
>3.4	603 (66.9%)	849 (87.6%)	
HER2 overexpression (No)	883 (91.4%)	849 (84.0%)	**$5.3x10^{-7}$**
(Yes)	83 (8.6%)	162 (16.0%)	
Triple negative (No)	866 (89.6)	794 (78.5)	**$1.6x10^{-11}$**
(Yes)	100 (10.4)	217 (21.5)	
ER (Negative)	156 (16.1%)	314 (31.1%)	**$7.0x10^{-15}$**
(Positive)	810 (83.9%)	697 (68.9%)	
PgR (Negative)	359 (37.2%)	577 (57.1%)	**$7.8x10^{-19}$**
(Positive)	607 (62.8%)	434 (42.9%)	
Genefu subtype			
ER-/HER2 negative	47 (9.8%)	103 (20.0%)	**$6.0x10^{-6}$**
ER+/HER2 negative/high proliferation	97 (20.2%)	269 (52.3%)	**$9.2x10^{-26}$**
ER+/HER2 negative/low proliferation	299 (62.3%)	69 (13.4%)	**$3.1x10^{-57}$**
HER2 positive	37 (7.7%)	73 (14.2%)	**0.001**
PAM50 subtype			
PAM50.HER2	66 (8.1%)	172 (17.9%)	**$1.9x10^{-9}$**
PAM50.Basal	90 (11.1%)	240 (25.0%)	**$7.6x10^{-14}$**
PAM50.LumA	515 (63.5%)	200 (20.8%)	**$2.1x10^{-74}$**
PAM50.LumB	140 (17.3%)	349 (36.3%)	**$3.9x10^{-19}$**

Bold = Statistically significant;HER2: human epidermal growth factor receptor 2; ER: oestrogen receptor; PgR: progesterone receptor; Triple negative: ER-/PgR-/HER2- . High proliferation = high Ki67 index, Low proliferation = low Ki67 index.

Figure legends

Figure 1. *RECQL4* copy number, mRNA and protein levels in breast cancer. **A.** *RECQL4* gene copy number changes in PAM50. Molecular phenotypes. **B**. Kaplan Meier curves showing BCSS (breast cancer specific survival) in the whole cohort. **C**. *RECQL4* mRNA levels in PAM50. Molecular phenotypes. **D**. Kaplan Meier curves showing BCSS (breast cancer specific survival) in the whole cohort.

Figure 2. A. Correlation between *RECQL4* gene copy number and mRNA levels in the whole cohort. **B.** Kaplan Meier curves showing BCSS based on RECQL4 protein levels in the whole cohort. **C.** Effects of RECQL4 knockdown using siRNA on DNA synthesis assessed by DNA fibre assay (see Supplementary Methods for details). **D.** In MCF7 and MDA-MB-231 cells DNA fibre lengths were reduced by around 50% after RECQL4 knockdown. Compared with the other two lines, BT549 cells showed shorter fibre tracks to begin with, but like the other cells, RECQL4 depletion further reduced the rate of synthesis of DNA

Figure 3. RECQL4 depletion and chemosensitivity to chemotherapeutic drugs in MDA-MB453 cells. RECQL4 knockdown by adenovirus-mediated shRNA (**A**) and treatment with cisplatin (**B**), doxorubicin (**C**) or 5-FU (**D**). Cell survival was measured by the 3-(4,5-dimethylthiazol-2-yl)-2,5- diphenyltetrazolium bromide (MTT) assay following the manufacturer's instructions (Molecular Probes). Absorbance values at 540 nm were read on a Spectra Max 250 spectrophotometer (Molecular Devices). All MTT assays include 10 duplicated wells for each time-point of each cell line. The data was represented as mean ± SD from three independent experiments. *, $P<0.05$; **, $P<0.01$.

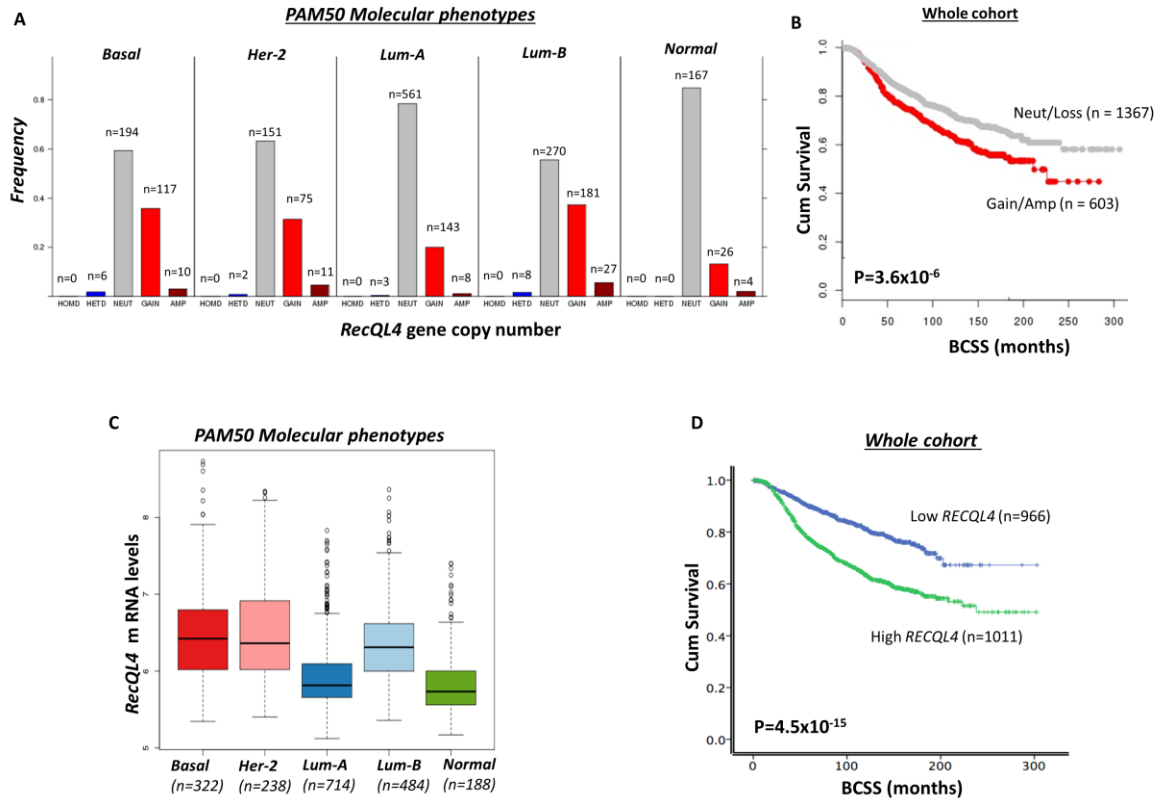


Figure 1

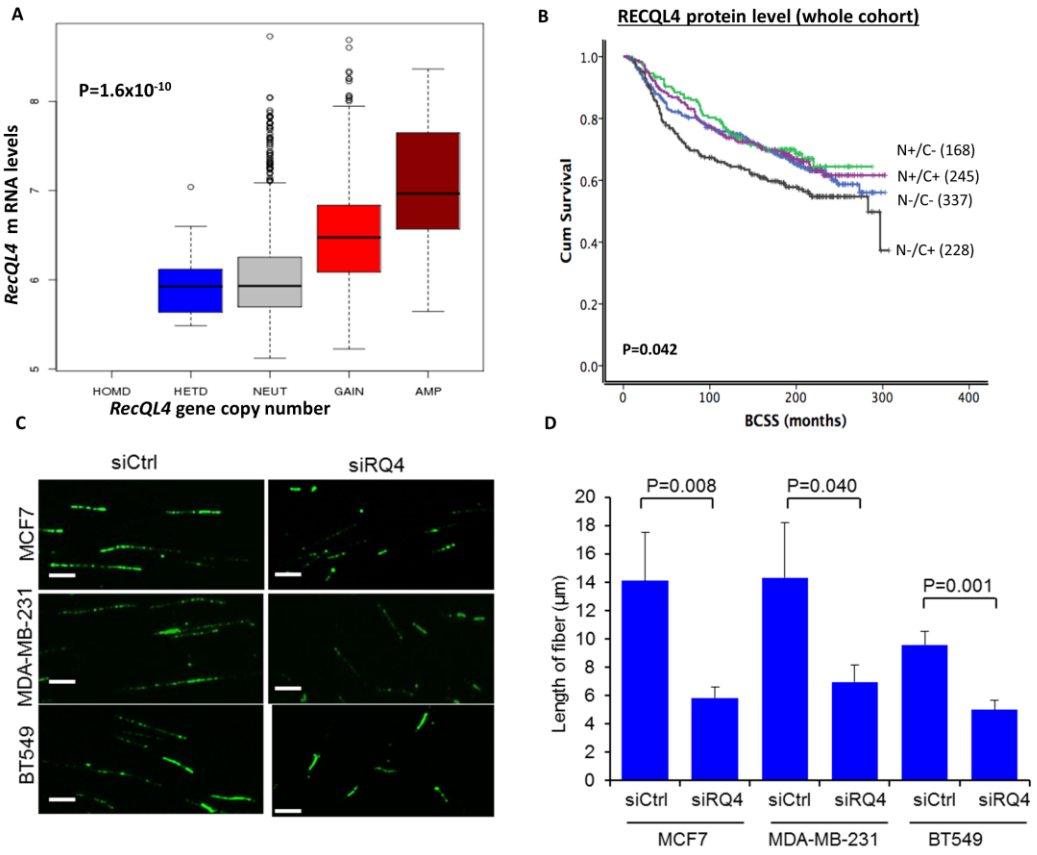
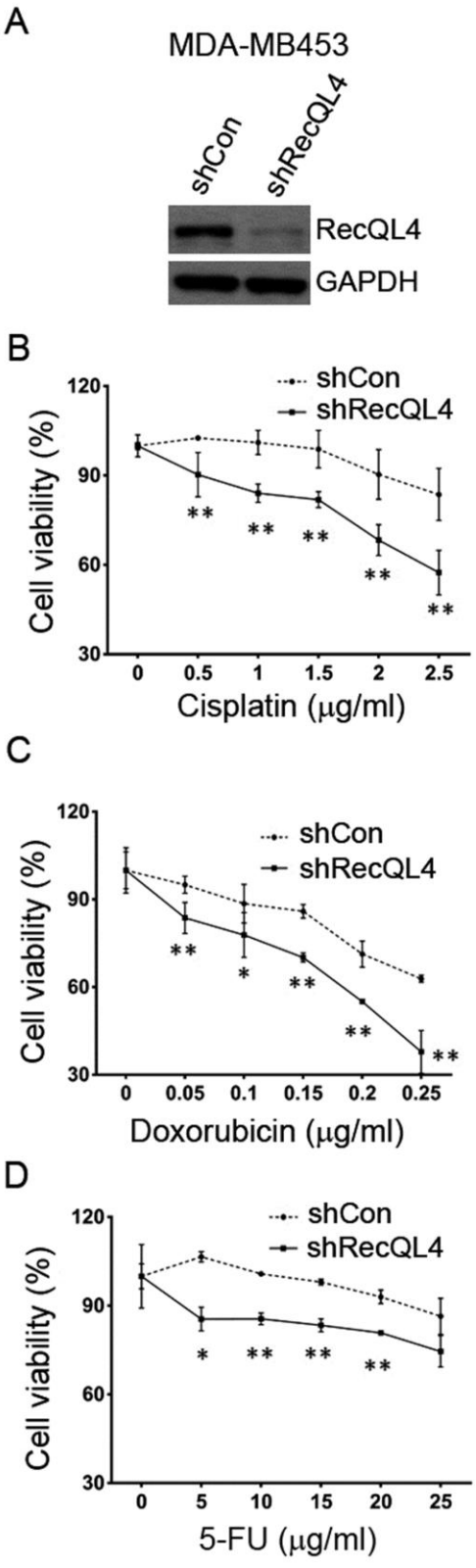


Figure 2


A
MDA-MB453
shCon
shRecQL4
RecQL4
GAPDH
B
shCon
shRecQL4
Cell viability (%)
120
90
60
30
0
0.5
1
1.5
2
2.5
Cisplatin (μg/ml)
C
shCon
shRecQL4
Cell viability (%)
120
90
60
30
0
0.05
0.1
0.15
0.2
0.25
Doxorubicin (μg/ml)
D
shCon
shRecQL4
Cell viability (%)
120
90
60
30
0
5
10
15
20
25
5-FU (μg/ml)

Supplementary Table S1: Clinicopathological characteristics in the METABRIC external validation cohort

Variables	N (%)
Age at diagnosis [Median (range)]	61.8 (21.93-96.29)
Tumour size [Median (range)]	23 (1, 182)
NPI [Median (95% CI)]	4.04 (3.99-4.09)
Survival [Median (Months, 95% Cl)]	149 (141-159)
Lymph nodes status 0 1 2 3 >3	 1012 336 170 112 316
ER status Positive Negative	 1485 437
PAM50 subtype Basal HER2 Luminal A Luminal B Normal Not classified	 322 238 714 484 188 6
Adjuvant systemic therapy (AT)	
No AT	290
Hormone therapy (HT)	1014
Chemotherapy	226
Hormone + chemotherapy	192

Supplementary Table S2: Clinicopathological characteristics of Nottingham Tenovus series

Variable	n*	Cases	(%)
Menopausal status	1650		
Pre-menopausal		612	(37.0)
postmenopausal		1038	(63.0)
Tumour Grade (NGS)	1650		
G1		306	(18.5)
G2		531	(32.2)
G3		813	(49.3)
Lymph node involvement	1650		
Negative		1056	(64.0)
Positive (1-3 nodes)		486	(29.5)
Positive (>3 nodes)		108	(6.5)
Tumour size (cm)	1650		
T1 a + b (≤1.0)		187	(11.0)
T1 c (>1.0 -2.0)		868	(53.0)
T2 (>2.0-5)		579	(35.0)
T3 (>5)		16	(1.0)
Tumour type	1650		
IDC-NST		941	(57)
Tubular		349	(21)
ILC		160	(10)
Medullary (typical/atypical)		41	(2.5)
Others		159	(9.5)
NPI subgroups	1650		
Excellent prognosis (2.08-2.40)	Low risk	207	(12.5)
Good prognosis (2.42-3.40)		331	(20.1)
Moderate I prognosis (3.42 to 4.4)	High risk	488	(29.6)
Moderate II prognosis (4.42 to 5.4)		395	(23.9)
Poor prognosis (5.42 to 6.4)		170	(10.3)
Very poor prognosis (6.5–6.8)		59	(3.6)

Survival at 20 years	1650		
Alive and well		1055	(64.0)
Dead from disease		468	(28.4)
Dead from other causes		127	(7.6)
Adjuvant systemic therapy (AT)			
No AT		665	(42.0)
Hormone therapy (HT)		642	(41.0)
Chemotherapy		307	(20.0)
Hormone + chemotherapy		46	(3.0)

* Number of cases for which data were available.

NPI; Nottingham prognostic index.

Supplemental Table S3: Clinicopathological characteristics of ER- cohort

Variable	n*	Cases	(%)
Menopausal status	252		
Pre-menopausal		122	(48.5)
postmenopausal		130	(51.5)
Tumour Grade (NGS)	252		
G1		1	(0.3)
G2		27	(10.6)
G3		224	(89.1)
Lymph node involvement	252		
Negative		121	(48)
Positive (1-3 nodes)		86	(34)
Positive (>3 nodes)		45	(18)
Tumour size (cm)	252		
T1 a + b (≤1.0)		28	(11)
T1 c (>1.0 -2.0)		106	(42)
T2 (>2.0-5)		103	(41)
T3 (>5)		15	(6)
Tumour type	252		
IDC-NST		224	(89.0)
Tubular		5	(2.0)
ILC		8	(3.0)
Medullary (typical/atypical)		5	(2.0)
Others		0	(4.0)
NPI subgroups	252		
Excellent prognosis (2.08-2.40)	Low risk	0	(0.0)
Good prognosis (2.42-3.40)		0	(0.0)
Moderate I prognosis (3.42 to 4.4)	High risk	111	(44.0)
Moderate II prognosis (4.42 to 5.4)		81	(32.0)
Poor prognosis (5.42 to 6.4)		38	(15.0)
Very poor prognosis (6.5–6.8)		22	(9.0)

Survival at 5 years	252		
Alive and well		176	(70.0)
Dead from disease		73	(29.0)
Dead from other causes		3	(1.0)

* Number of cases for which data were available.

NPI; Nottingham prognostic index.

Supplementary Table S4: RTS patient's demographics.

Rothmund-Thomson Syndrome (Type II)			
	All n=58	Female n=24	Male n=34
Age, median (range)	17.5 yrs (10 mos - 51 yrs)	13 yrs (10 mos - 40 yrs)	18.5 yrs (10 mos - 51 yrs)
<40 years old, n (%)	55 (94.8%)	23 (95.8%)	32 (94.1%)
Follow-up, median (range)	12 yrs (10 mos - 42 yrs)	10 yrs (10 mos - 37 yrs)	16 yrs (10 mos - 42 yrs)

Supplementary Table S5: *RECQL4* gene copy number alterations in the METABRIC cohort (n=1980)

	RECQL4 Gene Copy Number changes				
	HETD N(%)	**NEUT N(%)**	**GAIN N(%)**	**AMP N(%)**	**P value**
Stage					
1	5 (35.7)	369 (36.9)	121(30.9)	6(14.0)	**0.036**
2	8 (57.1)	556 (55.6)	225(57.5)	31(72.1)	
3	1 (7.2)	69 (6.9)	41(10.5)	6(13.9)	
4	0 (0.0)	6 (0.6)	4(1.1)	0(0.0)	
Grade					
G1	0 (0.0)	143 (11.2)	25 (4.7)	1(1.8)	**0.0001**
G2	8 (44.4)	576 (45.0)	160(30.4)	21(37.5)	
G3	10(55.6)	562 (43.8)	342(64.9)	34(60.7)	
Lymph node involvement					
Negative	11(57.9)	739 (54.8)	268(49.4)	22(36.7)	**0.01**
Positive	8(42.1)	609(45.2)	275(50.6)	38(63.3)	
PAM50 subtype					
PAM50.Her2	2(10.5)	151(12.8)	75(14.5)	11(19.6)	**0.00001**
PAM50.Basal	6(31.6)	194(16.5)	117(22.7)	10(17.9)	
PAM50.LumA	3(15.8)	561(47.7)	143(27.7)	8(14.3)	
PAM50.LumB	8(42.1)	270(23.0)	181(35.1)	27(48.2)	

Supplementary Table S6: Multivariate analysis in the METABRIC cohort confirms that *RECQL4* mRNA over expression is a powerful independent prognostic factor.

	P-Value	**HR**	**95% CI for HR**	
			Lower	**Upper**
Breast Cancer Specific Survival				
RECQL4 Expression	**0.000015**	1.366	1.185	1.573
NPI	**0.00016**	1.319	1.142	1.524
Tumour Grade				
G1	**0.039**	0.893	0.538	1.481
G2	0.246	0.992	0.629	1.563
G3	0.178	1.357	0.734	2.509
LN involvement				
LN (1-3)	**0.00046**	1.980	1.328	2.950
LN(>3)	0.312	1.277	0.973	1.678

Bold: Statistically significant; HR: Hazard Ratio; CI: Confidence interval; LN: Lymph node; NPI: Nottingham Prognostic Index.

Supplementary Table S7. RECQL4 nuclear and cytoplasmic co-expression and breast cancer.

VARIABLE	RECQL4(Nuclear & Cytoplasmic) Protein Co-expression				P- value
	Rn-/RC- N (%)	Rn+/RC- N (%)	Rn-/Rc+ N (%)	Rn+/Rc+	
A) Pathological Parameters					
Tumour Size					
≤1cm	31 (9.1)	17 (9.6)	25 (10.6)	25 (10.0)	0.927
>1-2cm	171 (50.0)	94 (52.8)	110 (46.6)	132 (53.0)	
>2-5cm	133 (38.9)	62 (34.8)	96 (40.7)	88 (35.3)	
>5cm	7 (2.0)	5 (2.8)	5 (2.1)	4 (1.6)	
Tumour Stage					
1	209 (60.8)	109 (61.2)	142 (60.4)	165 (66.0)	0.785
2	102 (29.7)	53 (29.8)	73 (31.1)	69 (27.6)	
3	33 (9.6)	16 (9.0)	20 (8.5)	16 (6.4)	
Tumour Grade					
G1	55 (16.0)	26 (14.6)	35 (14.8)	42 (16.9)	**0.001**
G2	101 (29.4)	83 (46.6)	62 (26.3)	82 (32.9)	
G3	187 (54.5)	69 (38.8)	139 (58.9)	125 (50.2)	
Mitotic Index					
M1 (low; mitoses < 10)	100 (29.9)	88 (49.4)	58 (25.6)	315 (32.1)	**$1.4X10^{-8}$**
M2 (medium; mitoses 10-18)	69 (20.6)	27 (15.2)	33 (14.5)	196 (20.0)	
M3 (high; mitosis >18)	166 (49.6)	63 (35.4)	136 (59.9)	471 (48.0)	
Tubule Formation					
1 (>75% definite tubule)	20(6.0)	8 (4.5)	10 (4.4)	54 (5.5)	0.282
2 (10%-75% definite tubule)	102 (30.4)	60 (33.7)	78 (34.4)	336 (34.2)	
3 (<10% definite tubule)	213 (63.6)	110 (61.8)	139 (61.2)	592 (60.3)	
Pleomorphism					
1 (small-regular uniform)	11 (3.3)	2 (1.1)	1 (0.4)	5 (2.1)	**0.014**
2 (Moderate variation)	120 (35.9)	83 (46.9)	74 (32.7)	87 (36.0)	
3 (Marked variation)	203 (60.8)	92 (52.0)	151 (66.8)	150 (62.0)	
Tumour Type					
IDC-NST	221 (65.8)	90 (50.8)	143 (62.2)	152 (62.6)	**$6.7x10^{-7}$**
Tubular Carcinoma	62 (18.5)	35 (19.8)	49 (21.3)	59 (24.3)	
Medullary Carcinoma	9 (2.7)	2 (1.1)	11 (4.8)	4 (1.6)	
ILC	19 (5.7)	35 (19.8)	10 (4.3)	12 (4.9)	
Others	6(61.8)	3 (1.7)	1 (0.4)	2 (0.8)	
Mixed NST/Lobular/Special Type	19 (5.7)	12 (6.8)	16 (7.0)	14 (5.8)	
Lymph Node involvement					
Negative	180 (59.6)	102 (60.4)	114 (57.3)	145 (63.9)	0.530
Positive (1-3)	97 (32.1)	55 (32.5)	71 (35.7)	73 (32.2)	
Positive (>3)	25 (8.3)	12(7.1)	14 (7.0)	9 (4.0)	
B) Aggressive Phenotype					
Her2 overexpression					
No	284 (84.0)	151 (86.8)	191 (82.3)	203 (83.2)	0.662
Yes	54 (16.0)	23 (13.2)	41 (17.7)	41(16.8)	

Triple Negative Phenotype No Yes	294 (85.2) 51 (14.8)	162 (91.0) 16 (9.0)	184 (78.0) 52 (22.0)	198 (79.2) 52 (20.8)	**0.001**
NPI ≤3.4 >3.4	97 (29.8) 229 (70.2)	67 (39.6) 102 (60.4)	54 (24.0) 171 (76.0)	75 (31.1) 166 (68.9)	**0.010**
C) Hormone Receptors					
ER Negative Positive	105 (31.4) 229 (68.6)	30 (17.0) 146 (83.0)	75 (32.8) 154 (67.2)	54 (22.5) 186 (77.5)	**$3.4X10^{-4}$**
PgR Negative Positive	162 (49.4) 166 (50.6)	65 (38.5) 104 (61.5)	108 (47.2) 121 (52.8)	84(36.2) 148 (63.8)	**0.006**

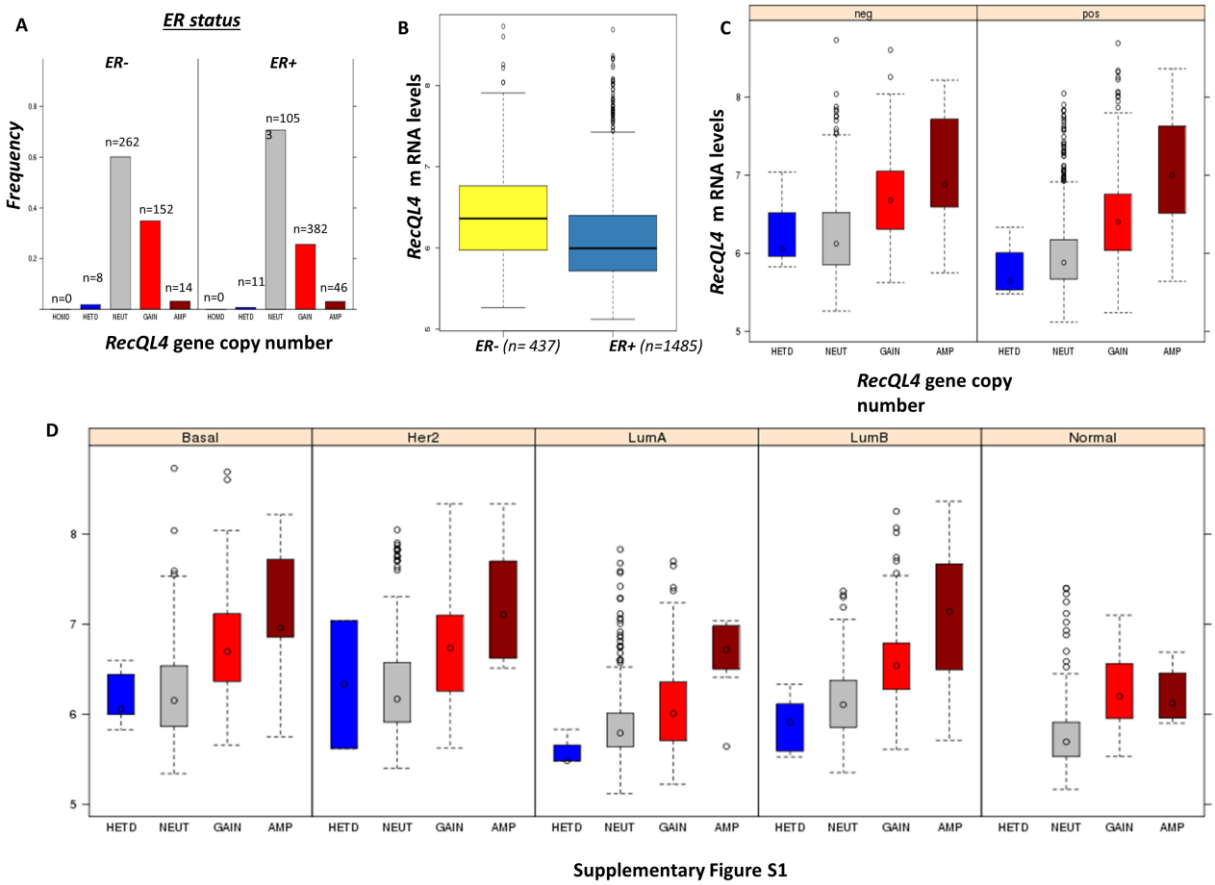
Supplementary Table S8: RECQL4 protein expression and survival – Multivariate Analysis

	P value	Exp (B)	95% CI of Exp (B)	
			Lower	Upper
Breast cancer specific Survival				
RECQL4 (Nuclear)	**0.032**	0.753	0.581	0.976
RECQL4 (Cyto)	0.352	1.126	0.877	1.446
Tumour Grade	**3.0×10^{-6}**	1.718	1.370	2.155
Lymph Node **involvement**	**5.6×10^{-12}**	1.980	1.630	2.404
Tumour Size	0.061	1.211	0.991	1.478
ER Status	0.698	1.061	0.788	1.428
HER2 Status	**0.001**	1.686	1.246	2.280
Endocrine Therapy	0.072	1.239	0.981	1.564
Chemotherapy	0.410	1.153	0.821	1.620

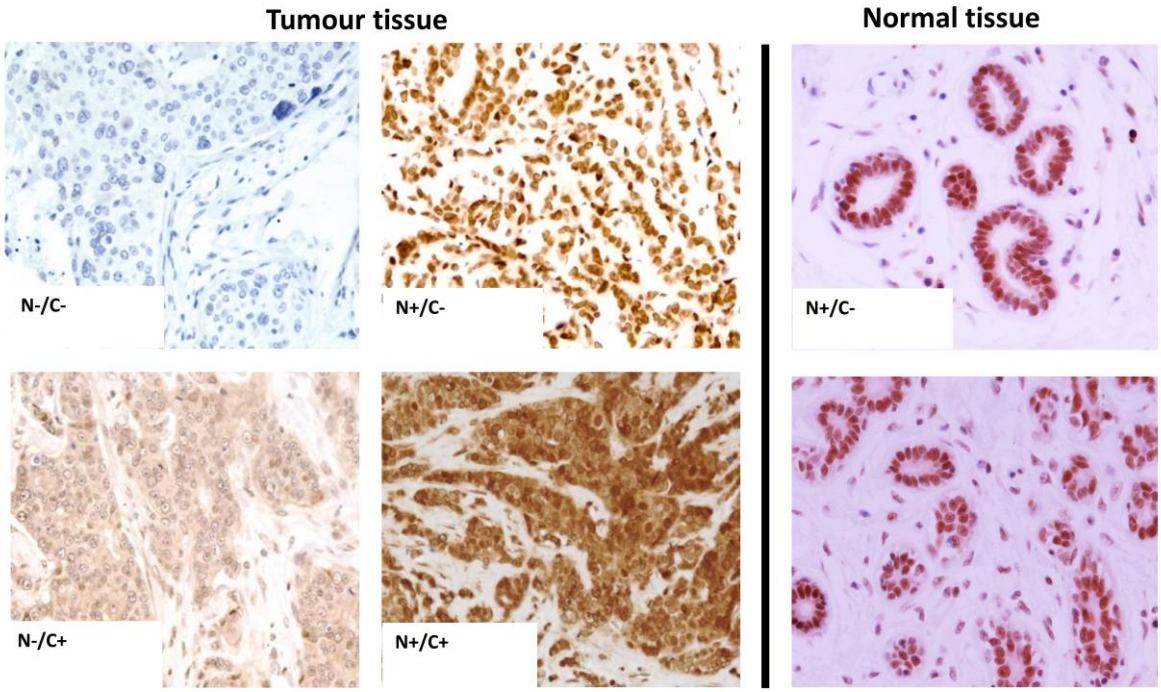
Supplementary Figure legends

Supplementary Figure S1. A, *RECQL4* gene copy number changes in ER+ and ER- breast cancers. B, *RECQL4* mRNA expression in ER+ and ER- breast cancers. C, Correlation between *RECQL4* gene copy numbers and mRNA expression based on ER status (**C**) and PAM50. D, Molecular phenotypes. E, Photomicrographs showing RECQL4 protein expression and sub-cellular localisation in tumour tissue and normal breast tissue. N= nuclear staining, C= cytoplasmic staining, '+' = positive for staining, '-' = negative for staining].

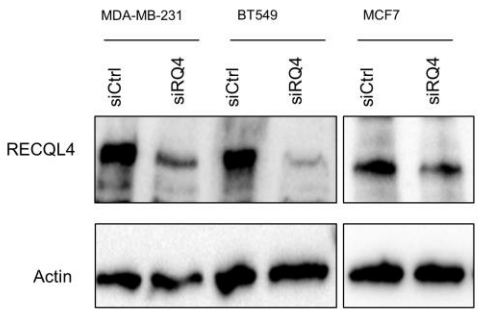
Supplementary Figure S2. RECQL4 knockdown using siRNA.



Supplementary Figure S1



Supplementary Figure S1E



Supplementary Figure S2