Genome-wide analysis unveils DNA helicase RECQ1 as a regulator of estrogen response pathway in breast cancer cells

Xing Lu[1], Christophe E. Redon[2], Wei Tang[3], Swetha Parvathaneni[1], Bayan Bokhari[1,4], Subrata Debnath[1], Xiao Ling Li[5], Bruna R. Muys[5], Young Song[6], Lorinc S. Pongor[2], Omar Sheikh[7], Andrew R. Green[7], Srinivasan Madhusudan[7], Ashish Lal[5], Stefan Ambs[3], Javed Khan[6], Mirit I. Aladjem[2*] and Sudha Sharma[1*]

[1]Department of Biochemistry and Molecular Biology, College of Medicine, Howard University, Washington, DC, USA. [2]Developmental Therapeutics Branch, Laboratory of Molecular Pharmacology, Center for Cancer Research (CCR), National Cancer Institute (NCI), National Institutes of Health (NIH), Bethesda, MD, USA. [3]Molecular Epidemiology Section, Laboratory of Human Carcinogenesis, CCR, NCI, NIH, Bethesda, MD, USA. [4]Department of Biochemistry, Faculty of Applied Medical Science, Umm Al-Qura University, Makkah 21421, Saudi Arabia. [5]Regulatory RNAs and Cancer Section, Genetics Branch, CCR, NCI, NIH, Bethesda, MD, USA. [6]Oncogenomics Section, Genetics Branch, CCR, NCI, NIH, Bethesda, MD 20892, USA. [7]Division of Cancer & Stem Cells, Nottingham Breast Cancer Research Centre, University of Nottingham Biodiscovery Institute, Nottingham University Hospitals, Nottingham, NG51PB, UK.

*** Corresponding authors**: sudha.sharma@howard.edu; aladjemm@mail.nih.gov

Abstract

Susceptibility to breast cancer is significantly increased in individuals with germline mutations in *RECQ1*, a gene encoding a DNA helicase essential for genome maintenance. We previously reported that *RECQ1* expression predicts clinical outcomes for sporadic breast cancer patients stratified by estrogen receptor (ER) status. Here, we utilized an unbiased integrative genomics approach to delineate a crosstalk between RECQ1 and ERα, a known master regulatory transcription factor in breast cancer. We found that expression of *ESR1*, the gene encoding ERα, is directly activated by RECQ1. More than 35% of RECQ1 binding sites were co-bound by ERα genome-wide. Mechanistically, RECQ1 cooperates with FOXA1, the pioneer transcription factor for ERα, to enhance chromatin accessibility at the *ESR1*-regulatory regions in a helicase activity-dependent manner. In clinical ERα-positive breast cancers treated with endocrine therapy, high RECQ1 and high FOXA1 co-expressing tumors were associated with better survival. Collectively, these results identify RECQ1 as a novel cofactor for ERα and uncover a previously unknown mechanism by which RECQ1 regulates disease-driving gene expression in ER-positive breast cancer cells.

Introduction

Since the first demonstration that RECQ1 (also known as RECQL or RECQL1) is essential for chromosomal stability (1), emerging biochemical and cellular functions of RECQ1 have provided a strong rationale to investigate the roles of RECQ1 in cancer biology (2-5) ultimately leading to its discovery as a candidate breast cancer susceptibility gene (6-8). However, the underlying mechanisms of RECQ1 in breast cancer biology are not yet understood. Given the inarguable evidences on the diverse functions of RECQ1, unbiased genome-wide approaches could provide deeper mechanistic insights and uncover new molecular functions of RECQ1 in normal development and human disease.

RECQ1 is localized to chromosome 12p12 and encodes a 649 amino acid protein RECQ1, a ubiquitous nuclear enzyme, and the most abundant homolog of the highly conserved RecQ helicase family (9-12). Biochemically, RECQ1 is a DNA-dependent ATPase, binds to single and double-stranded DNA, unwinds DNA duplex in 3'-5' direction and promotes strand annealing (13, 14). Through these multiple catalytic activities, RECQ1 responds to oxidative DNA damage (15), restores productive replication following DNA damage (16-20), participates in DNA double-strand break repair (21), removes chemical alterations to DNA bases via base excision repair pathway (22), and maintains telomeres (23, 24).

DNA repair functions of RECQ1 are mediated through its interactions with critical protein partners including PARP1, RPA, Top3α, MSH2/6, FEN1, and Ku70/80 (25). The essential role of RECQ1 in genome maintenance is underpinned by the fact that RECQ1 depletion in cells results in increased frequency of spontaneous sister chromatid exchanges, chromosomal instability, DNA damage accumulation, and increased sensitivity to cytotoxic chemotherapy (12, 26, 27). In addition to its DNA repair functions, and similar to its homologs in *Neurospora* (28) and Rat (29),

human RECQ1 also seems to have gene regulatory functions (30). The knockdown of RECQ1 in breast cancer cells has a significant effect on gene expression associated with tumorigenesis (31); however, the molecular mechanism of gene regulation by RECQ1 remains to be elucidated.

Both RECQ1 catalytic functions as well as expression levels are related to breast cancer. Whole-genome sequencing efforts revealed that rare, recurrent *RECQ1* mutations in the catalytic domain increase the risk of breast cancer by 5-fold among unselected cases from Poland and by 16-fold among higher-risk cases in Quebec (7). The association of *RECQ1* mutations with breast cancer was further confirmed in a Chinese population, suggesting that *RECQ1* mutations are not limited to specific populations (8). Subsequently, we conducted an evaluation of RECQ1 mRNA and protein expression in the large METABRIC cohort of sporadic breast cancer patients (n=1977) providing the first clinical evidence that altered RECQ1 expression is significantly associated with patient survival (32). In this dataset, high RECQ1 protein levels significantly ($p=0.021$) correlates with better survival in ER-positive tumors that received endocrine therapy indicating a mechanistic link of RECQ1 and estrogen response pathway (32).

Estrogen receptor alpha (ERα) is a member of the nuclear receptor family of transcription factors and a master regulator of tumor biology in two-thirds of all human breast cancers (33). ERα expression is both necessary and sufficient to predict the responsiveness to anti-estrogen in a high proportion of breast tumors, and low ERα expression level is generally associated with a poor prognosis (34). Understanding the mechanisms that regulate ERα activity in breast cancer is critical to understand how ERα mediates gene transcription and what occurs during endocrine resistance (35-39).

Given the clinical significance of ERα and the significant correlation between RECQ1 expression levels in breast tumors with clinical outcome in ERα-positive disease, we sought to

determine whether RECQ1 regulates ERα signaling, contributing to breast cancer and response to therapy. Here, we describe a previously unrecognized regulation of ERα signaling by RECQ1 and uncover the mechanisms of RECQ1 target gene regulation in breast cancer cells.

Results

RECQ1 depletion leads to downregulation of gene sets associated with estrogen response

Our recently published results show that low RECQ1 expression in the METABRIC dataset correlated with poor survival among ER-positive patients who received endocrine therapy (32). In our follow-up investigation, we also found that knockdown of *RECQ1* in a breast cancer cell line results in reduced ERα protein levels (32). These initial observations prompted us to investigate a potential role of RECQ1 in ERα signaling. To begin this investigation, we used an unbiased approach by determining global changes in gene expression that occur upon RECQ1 depletion by performing RNA-seq in biological triplicates following *RECQ1* knockdown in MCF7 cells grown in regular complete medium. We also performed RNA-seq from MCF7 cells after transient knockdown of *ESR1,* the gene encoding ERα. This would allow us to compare the transcriptomes regulated by RECQ1 and ERα. As expected, we observed a strong decrease in *RECQ1* and *ESR1* mRNA levels upon knockdown of *RECQ1* and *ESR1*, respectively (Fig. 1A). Interestingly, *ESR1* mRNA levels were also decreased when *RECQ1* was depleted but *RECQ1* mRNA levels did not change upon knockdown of *ESR1* (Fig. 1A). We therefore examined the effect of siRNA-mediated knockdown of *RECQ1* on ERα expression in MCF7 and T47D breast cancer cell lines that are widely used models for ER-positive breast cancer. In both cell lines, knockdown of *RECQ1* significantly reduced *ESR1* mRNA and ERα protein levels (Fig. S1A-D). These experiments were performed using SMARTpool of siRNAs against *RECQ1*. We next confirmed the role of RECQ1

as a positive regulator of ERα expression in MCF7 cells by using 4 individual siRNAs targeting *RECQ1* (Fig. S2A) and in MCF7 RECQ1-KO cells generated using CRISPR/Cas9 (Fig.s S2B and C). Because ERα mediates estrogen-stimulated cell proliferation (40), we next examined the effect of *RECQ1* knockdown on cell proliferation in response to estrogen treatment. Unlike control cells where proliferation significantly increased upon estrogen treatment, *RECQ1* knockdown cells did not show an increase in cell proliferation upon estrogen treatment in MCF7 and T47D cells (Fig. S1E and S1F) indicating crosstalk between RECQ1 and estrogen response pathway.

At the genome-wide level, 291 genes were significantly downregulated, and 119 genes were upregulated ($0.5<\log2 \text{ fold change}<-0.5$ and $p<0.05$) upon knockdown of *RECQ1* (Fig. 1B and 1C; Table S2). Using these statistical cut-offs, 1355 genes were downregulated and 1117 were upregulated upon *ESR1* knockdown (Fig. 1D and Table S3). Gene set enrichment analyses (GSEA) revealed the downregulation of many early and late estrogen-response genes as most significant upon *RECQ1* knockdown (Fig. 1E), and this pattern significantly overlapped with the expected enrichment of these pathways upon *ESR1* knockdown (Fig. 1F). Given the observed similarity between the enrichment of gene sets that were downregulated following knockdown of *RECQ1* or *ESR1*, we next looked at the intersection of genes whose expression was significantly decreased upon knockdown of *RECQ1* or *ESR1* in our RNA-seq. We found a subset of genes that were commonly downregulated upon *RECQ1* or *ESR1* knockdown (Fig. 1G).

However, a majority of genes that were downregulated upon the *ESR1* knockdown were not downregulated upon *RECQ1* knockdown (Fig. 1G). This could be because RECQ1 may be regulating a subset of ERα targets and/or due to experimental variation in the extent of *ESR1* downregulation upon *RECQ1* knockdown. Importantly, as we had seen by RT-qPCR (Fig. S1B and D), we observed a significant decrease in *ESR1* mRNA levels (Fold change = 0.75, p=0.0028)

in our RNA-seq following *RECQ1* knockdown (Table S4). Of the 291 genes downregulated upon *RECQ1* knockdown, 73 genes were also downregulated upon knockdown of *ESR1* (Fig. 1G). Of these 73 genes, 26 were early and late estrogen-response genes (Table S5). As shown in the IGV snapshot for 7 of these 26 genes (Fig. 2A), namely, *ASTL, CAV1, JAK2, MYB, OLFLM3, PLAC1*, *SLC16A1,* and *TFF1*, we observed a significant decrease in the expression of these genes upon knockdown of *RECQ1* or *ESR1*. We validated the downregulation of select genes by RT-qPCR following knockdown of *RECQ1* or *ESR1* in both MCF7 and T47D cells, indicating that these genes are not cell-type specific targets of RECQ1 and ERα (Fig. 2B and C). This result together with the observed significant enrichment of estrogen response genes upon *RECQ1* knockdown indicates that RECQ1 has a direct or indirect role in regulating genes involved in estrogen signaling.

RECQ1 ChIP-seq reveals significant genome-wide colocalization of RECQ1 and ERα

To determine if RECQ1 could play a direct role in the regulation of gene expression in ER-positive breast cancer cells, we performed ChIP-seq in duplicates for RECQ1 and for ERα in MCF7 cells grown in regular complete medium using a previously described method (41). Data from the replicates were pooled and peaks were called using model-based analysis for ChIP-seq (MACS). This resulted in the identification of thousands of RECQ1 binding events (Table S6) distributed predominantly in intergenic regions and introns (Fig. 3A, Fig. S3A, and Table S7). More than half (~59%) of the RECQ1 peaks colocalized with transcription start sites (Fig. S3B). We next validated our RECQ1 ChIP-seq data for 11 RECQ1-bound genes. We observed ~8-34-fold enrichment of RECQ1 at the target sites by ChIP-qPCR from MCF7 cells (Fig. 3B); knockdown of *RECQ1* in MCF7 cells resulted in decreased expression of a majority of these genes (Fig. S3C).

Moreover, RECQ1 ChIP-seq peaks showed the most significant enrichment of estrogen response genes (Fig. 3C). Among the genes bound by RECQ1, a subset changed in expression upon RECQ1 knockdown (Table S8).

Given the central role of ERα as a transcription factor in ER-positive breast cancer and our data showing that RECQ1 enhances basal ERα expression and some ERα targets, we next performed ERα ChIP-seq in MCF7 cells. ChIP-seq for ERα revealed thousands of ChIP-seq peaks distributed predominantly in intergenic regions and introns (Fig. 3D, Fig. S4A, and Table S9). The enrichment of the ERα binding events compared to the genome was strongest in the promoter regions (Fig. S4B and Table S7). When we compared the genome-wide binding events of RECQ1 with that of ERα, we found that ~35% (2423 out of a total of 7761 peaks) of RECQ1 peaks colocalize with ERα peaks (Fig. 3E); ~12% of total ERα bound sites overlapped with RECQ1 peaks (Fig. 3E). IGV snapshots for select genes that were bound by RECQ1 (Fig. 3B) are shown in Fig. S4C. Moreover, when we compared the RECQ1 peaks with ERα peaks in promoters and gene bodies, we found that ~59% (1853 out of a total of 3137 peaks) of RECQ1 peaks colocalize with ERα peaks and ~36% of ERα peaks colocalize with RECQ1 peaks (Fig. 3F and Table S10). Although our study is the first to report genome-wide RECQ1 binding sites via ChIP-seq, comparison of our ERα ChIP-seq data with 9 previously published ERα ChIP-seq from MCF7 cells that had ~22,000 to 78,000 peaks, showed ~75-95% colocalization (Fig. 3G). Moreover, comparing our RECQ1 ChIP-seq data with ChIP-seq data for a number of transcription factors including our ERα ChIP-seq showed that RECQ1 genome-wide binding positively correlates with ERα, FOXA1, GATA3, TCF7L2 and ZNF217 (Fig. 3H).

To better understand the relationship between RECQ1 and its co-localized factors with associated chromatin features, we performed k-means clustering of ChIP-seq peaks into 3 clusters

based on H3K4me1 (enhancer marker) and H3K4me3 (promoter marker) signals (Fig. 4A). Peaks in cluster 1 correspond to strong promoters with high H3K4me3 signal and low H3K4me1 signal, peaks in cluster 2 correspond to strong enhancers with high H3K4me1 signal and low H3K4me3 signal, whereas peaks in cluster 3 correspond to weak and/or inactive enhancers and promoters. K-means clustering of the ChIP-seq data revealed that in addition to active promoter regions, RECQ1 binding sites are associated with active enhancers similar to FOXA1 and GATA3 binding (ChIP-Seq data from ENCODE) (Fig. 4A).

The colocalization of RECQ1 and ERα at the genome-wide level, combined with the significant enrichment of estrogen response pathway in the genes downregulated upon RNA-seq overlap between the gene expression changes that occur upon knockdown of RECQ1, indicated a potential role of RECQ1 as a coregulator for ERα signaling. Therefore, we next compared RECQ1 ChIP-seq peaks with a previously published RNA Pol II (POLR2A) binding over a time-course of estrogen treatment in MCF7 cells (42) (Fig. 4B and C). The correlation of RECQ1-peaks with estradiol-stimulated POLR2A occupancy onto chromatin is rapidly and significantly increased, as early as 5 min after treatment, and returns to an unstimulated level at 6 hours after estradiol treatment indicating potentially dynamic recruitment of RECQ1 to the estrogen-responsive gene promoters (Fig. 4B). Similar analysis with the ERα ChIP-seq peaks revealed a stronger correlation with estrogen-stimulated POLR2A recruitment than RECQ1 ChIP-seq peaks, however, the trend was similar (Fig. 4C). The correlation of RECQ1 ChIP-seq with estrogen-stimulated POLR2A recruitment was comparable with the publicly available ChIP-seq datasets for FOXA1 and GATA3, the two proteins known to be critical in tethering ERα to the DNA (33) (Fig. 4C). In comparison to these well-established regulators of estrogen response, publicly available ChIP-seq datasets of REST and PAX8 showed significantly reduced correlation with POLR2A occupancy

(Fig. 4C). Taken together, these analyses suggest that the genome-wide binding of RECQ1 is correlated with estrogen-induced transcriptional dynamics in MCF7 cells and implicate RECQ1 in the estrogen response pathway.

RECQ1 associates with FOXA1 at the *ESR1* locus to enhances *ESR1* transcription

ERα is known to regulate its own expression by binding at the *ESR1* gene locus (43). Therefore, in this study we decided to focus on a single RECQ1-bound gene locus and determining the molecular mechanism by which RECQ1 activates *ESR1* transcription. RECQ1 ChIP-seq compared to input DNA (Fig. 5A) or compared to IgG control (Table S6), and validation using ChIP-qPCR (Fig. 5B and 5C) demonstrated recruitment of RECQ1 at the *ESR1* promoter and enhancer regions. RNA Pol II ChIP from control and RECQ1 knockdown MCF7 cells showed reduced occupancy of RNA Pol II at the *ESR1* promoter and enhancer regions and exon 1 (+284) located near the transcription start site that has been previously shown to govern ERα expression in MCF7 cells (43, 44) (Fig. 5D). These results indicate that RECQ1 regulates ERα expression directly at the transcriptional level as reflected by the reduced RNA Pol II occupancy at regulatory regions of *ESR1* upon *RECQ1* knockdown. However, in co-IP experiments, we could not detect ERα in RECQ1 IPs (Fig. 5E).

Because FOXA1 is known to regulate *ESR1* transcription (45, 46), we next sought to determine if RECQ1 regulates *ESR1* expression by associating with FOXA1. Indeed, in co-IP experiments, we pulled down FOXA1 in RECQ1 IPs indicating that RECQ1 associates with FOXA1 (Fig. 5E). The RECQ1-FOXA1 interaction was not disrupted in the presence of Benzonase, an enzyme that degrades DNA and RNA, indicating that RECQ1 forms a protein-protein complex with FOXA1 (Fig. 5E). This observation is consistent with the reported

association of FOXA1 with RECQ1 in chromatin and chromatin-free complexes (47). We next performed sequential ChIP to determine if RECQ1, FOXA1, and ERα are bound at the same regulatory regions of *ESR1*. We found significant enrichment of *ESR1* enhancer 1 and 2 in reChIP of RECQ1 ChIP material with FOXA1 or ERα antibodies (Fig. 5F) indicating that these proteins work together to regulate the *ESR1* gene.

To further investigate the role of FOXA1 in binding of RECQ1 to the *ESR1* enhancer regions, we performed ChIP-qPCR for RECQ1 after knockdown of FOXA1 in MCF7 cells with siRNAs. Although in cells transfected with CTL siRNA, RECQ1 was enriched ~20-fold at the promoter and enhancer region of *ESR1*, knockdown of FOXA1 resulted in complete loss of binding of RECQ1 at these regions (Fig. 5G) suggesting that FOXA1 could recruit RECQ1 to these regions that are known to play important roles in *ESR1* transcription. As expected, FOXA1 was more enriched at the *ESR1* enhancer (~140-fold) as compared to the *ESR1* promoter (~30-fold) and this binding of FOXA1 to the ESR1 regulatory regions was dramatically reduced upon FOXA1 knockdown (Fig. 5G). Immunoblotting verified substantial decrease in FOXA1 expression in the FOXA1 siRNA-transfected cells (Fig. 5H); there was no change on RECQ1 mRNA or protein levels upon *FOXA1* knockdown (Fig. 5H and S5) suggesting that the observed loss of RECQ1 binding at the *ESR1* promoter and enhancer upon knockdown of FOXA1 was not due to decreased RECQ1 expression.

Helicase activity of RECQ1 is required for FOXA1-mediated *ESR1* transcription

Given that FOXA1 is a known pioneer factor that can bind compact chromatin (39, 45, 48), we next assessed whether RECQ1 helicase activity might contribute to chromatin accessibility at *ESR1* regulatory regions in MCF7 cells and hence facilitate *ESR1* expression.

FAIRE (Formaldehyde Assisted Isolation of Regulatory Elements) is a method that assays for nucleosome-free regions of the genome with open chromatin; and higher values for enrichment correspond to more accessible DNA (49). We first performed FAIRE followed by qPCR for enrichment of *ESR1* promoter and enhancer regions to determine the effect of loss of endogenous RECQ1 on chromatin accessibility at these specific *ESR1* regulatory regions (Fig. 6A). Loss of endogeneous RECQ1 induced a decrease in chromatin accessibility at *ESR1* promoter and enhancer regions as assessed by FAIRE-qPCR in RECQ1-WT and RECQ1-KO MCF7 cells (Fig. 6A). Consistent with the reduced chromatin accessibility at these *ESR1* regulatory regions, RECQ1-KO cells transfected with an empty vector showed reduced *ESR1* expression as compared to RECQ1-WT cells transfected with an empty vector (Fig. 6B). Thus, the pattern of accessibility changes at the *ESR1* regulatory regions in RECQ1-WT and RECQ1-KO is comparable to the *ESR1* mRNA expression in these isogenic cells (Fig. 6A and B). Overexpression of wild-type RECQ1 in RECQ1-KO cells restored *ESR1* expression to the RECQ1-WT level whereas overexpression of a well-characterized helicase-dead RECQ1-K119R variant failed to restore *ESR1* expression in RECQ1-KO cells (Fig. 6B).

We next determined if RECQ1 helicase may cooperate with FOXA1 in modulating DNA accessibility at *ESR1* promoter and enhancer regions to restore *ESR1* expression in RECQ1-KO cells. To test this, we performed FAIRE-qPCR assays in RECQ1-KO MCF7 cells overexpressing FOXA1 alone or in combination with wild-type RECQ1 or the helicase dead RECQ1-K119R (Fig. 6C). Western blot analysis of transfected cells validated comparable expression of wild-type and K119R RECQ1 proteins in RECQ1-KO MCF7 cells (Fig. 6D). As compared to RECQ1-KO cells transfected with an empty vector, overexpression of FOXA1 in RECQ1-KO cells increased FAIRE-enrichment at *ESR1* promoter (~2-fold) and enhancer (~6-fold) (Fig. 6C). Co-transfection

of FOXA1 and wild-type RECQ1 expressing constructs resulted in further increase in FAIRE-enrichment at *ESR1* promoter (~3.5-fold) and enhancer (~9-fold) as compared to RECQ1-KO cells transfected with an empty vector (Fig. 6C). In contrast, co-transfection with RECQ1-K119R resulted in FAIRE-enrichment at *ESR1*-promoter (~1.5-fold) and enhancer (~4-fold) indicating that the helicase activitity of RECQ1 is essential to cooperate with FOXA1 to enhance chromatin accessibility at *ESR1* regulatory regions (Fig. 6C). Consistent with enhanced chromatin accessibility as indicated by FAIRE enrichment at *ESR1* regulatory regions, co-transfection of FOXA1 and wild-type RECQ1 expressing constructs increased *ESR1* and *TFF1* expression in RECQ1-KO cells (Fig. 6E) wheras cotransfection with RECQ1-K119R had minimal effect.

These results indicate that RECQ1 acts as a partner for FOXOA1 in its pioneer activity. MCF7 cells lacking RECQ1 have reduced accessibility at the *ESR1* promoter as well as at *ESR1* enhancer where FOXOA1 has an impact on accessibility that is dependent on the helicase activity of RECQ1. This is consistent with the role of FOXOA1 as a pioneer factor which are thought to act at the enhancers to initiate chromatin opening (50). Promoters for the most part maintain an open chromatin configuration (51). Therefore, our data suggests a more complex relationship of RECQ1 with regulatory chromatin. Collectively, these data provide evidence to support our hypothesis that RECQ1 helicase cooperates with FOXA1 to enhance chromatin accessibility and enhance *ESR1* expression.

RECQ1-FOXA1 co-expression and survival outcomes in clinical ERα positive breast cancers

Pre-clinical evidence presented thus far suggests that RECQ1-FOXA1 interaction may influence ERα positive breast cancer pathogenesis and impact survival outcomes in patients who received endocrine therapy. We have investigated the clinical significance of RECQ1 (32) or FOXA1 (52)

in a large cohort of breast cancers. We observed nuclear staining only for RECQ1 and FOXA1. Here, we investigated the clinicopathological significance of RECQ1/FOXA1 protein co-expression in ERα positive breast cancers (n=1406). Patient demographics are summarized in Tables S11 and S12. All patients received tamoxifen adjuvant endocrine therapy. As shown in Fig. 6D, breast cancer-specific survival (BCSS) was better in patients whose tumor had high RECQ1/high FOXA1 co-expression compared to those with low RECQ1/low FOXA1 co-expression ($p=0.009$). Tumors with low RECQ1/low FOXA1 co-expression were associated with higher tumor stage, high tumor grade, de-differentiation, pleomorphism, higher mitotic index, and high-risk Nottingham prognostic index (NPI >3.4) as compared to tumors with high RECQ1-high FOXA1 co-expression (all p values <0.05, Table S13). These data suggest that RECQ1/FOXA1 co-expression in ERα positive breast cancers has clinicopathological and prognostic significance.

Based on our data we propose a working model on the transcriptional circuitry linking RECQ1 to ERα and ultimately the biological response in ER-positive breast cancer cells (Fig. 7). According to this model, RECQ1 promotes transcription of *ESR1,* the gene encoding ERα. By regulating ERα levels, and as a cofactor for ERα in cooperation with FOXA1 and other yet to be identified factors, RECQ1 binding to promoter or enhancer regions regulates expression of a subset of ERα target genes.

Discussion

Previous work has exhaustively described various roles of RECQ1 helicase in DNA repair and its requirement for genome maintenance. Here, we report an unexpected role of RECQ1 in the regulation of ERα signaling in breast cancer cells in the absence of estrogen. We demonstrate that RECQ1 directly regulates the expression of *ESR1,* a well-known therapeutic target that has been

shown to determine clinical outcomes in breast cancer (33, 37). Using RNA-Seq, we found that *ESR1* and several of its downstream target genes are significantly downregulated in RECQ1 knockdown MCF7 cells. ChIP-seq analysis revealed a significant overlap of RECQ1 genomic binding sites with ERα, including binding to the *ESR1*. Mechanistically, RECQ1 is recruited to the regulatory regions of the *ESR1* gene through its interaction with FOXA1 and cooperates with FOXA1 to facilitate chromatin accessibility and promote *ESR1* expression. In addition, RECQ1 levels are significantly associated with clinical outcomes in ER-positive breast cancer patients, specifically in those receiving tamoxifen treatments. Taken together, our findings identify a novel mechanism of RECQ1 as an upstream regulator of ERα and strengthen the clinical relevance of the RECQ1 helicase in breast cancer.

Transcriptional regulation by RECQ1, either by associating with the promoter of target genes or through cooperating with other transcriptional regulators, has not been studied before. Our findings begin to address this paucity using unbiased genome-wide approaches. The analysis of the transcriptome upon depletion of RECQ1 in ER-positive MCF7 cells allowed us to unveil the unanticipated function of RECQ1 as a regulator of *ESR1* expression and a modulator of ERα-dependent gene expression. Indeed, we found that the siRECQ1 downregulated transcriptome is significantly enriched in early and late-estrogen response genes. A comparison with siESR1 downregulated genes indicated a subset of genes commonly regulated by RECQ1 and ERα, however, a majority of differential gene expression in siRECQ1 cells are affected through an ERα-independent mechanism. ChIP-seq in MCF7 cells demonstrated significant enrichment of RECQ1 binding at a subset of estrogen-responsive genes. Comparative analysis of genome-wide binding of RECQ1 and ERα in MCF7 cells revealed that >35% of RECQ1-binding peaks overlap with bona fide genomic ERα binding peaks independent of its ligand, estrogen. We note that a majority

of ERα binding peaks were not enriched for RECQ1, thus indicating that RECQ1 may not be a general regulator of ERα-driven transcription or RECQ1-mediated gene expression change may not require stable association of RECQ1. However, the fact that a third of the total RECQ1-binding peaks were co-bound with ERα genome-wide and that ~60% of the RECQ1-binding peaks were co-bound by ERα at the promoters and gene bodies, suggests a mechanistic crosstalk between these two proteins for a wide range of genomic functions in breast cancer cells. The data presented in this study focused on a single RECQ1-bound gene locus provides mechanistic inisghts on the role of RECQ1 in activating *ESR1* transcription. Future studies utilizing a more global approach will establish mechanisms by which RECQ1 regulates the transcription of specific genes important for genome stability and cancer.

ERα requires cofactors to assist with DNA binding and transcriptional regulation (53), and FOXA1 acts as a pioneer factor for ERα by modulating chromatin structure and promoter-enhancer interactions (46, 54). Our results suggest a novel role of RECQ1 in facilitating chromatin accessibility through a possible interaction with FOXA1 and in a helicase-activity dependent manner. Through its robust DNA-binding and unwinding activities, RECQ1 helicase cooperates with FOXA1 to promote *ESR1* expression as we have shown. RECQ1 in association with FOXA1 and other yet to be identified cofactors may enhance the accessibility of transcriptional machinery and promote genomic binding of transcription factors such as ERα, and its co-regulators to control physiologically relevant gene expression programs (Fig. 7). Given that RECQ1 controls the transcription of *ESR1*, it may be difficult to separate out the changes in ERα-induced genes for which RECQ1 acts as a co-activator, versus those due to the fact that ERα levels decrease upon RECQ1 loss. Germline mutations in RECQ1 that increase breast cancer risk inactivate its helicase activity (8), and thereby are also expected to impair its ability to regulate ERα expression and

signaling in ER-positive breast cancer. Further studies using global approaches such as ATAC-Seq will determine how the displacement of nucleosomes by the pioneer factor FOXA1 is impacted by dysfunctional RECQ1 and affects FOX1A and ERα binding and activity genome-wide.

The implication of RECQ1 helicase activity in facilitating local chromatin states to modulate ERα occupancy at specific genomic loci is conceptually novel. However, transcriptional regulation by RECQ1 may also employ helicase-independent mechanisms through enhancer-binding and protein interactions. Whether RECQ1 acts by participation in a complex with ERα, or by operating in a complex with its key co-regulators such as FOXA1 (47) and other proteins remains to be elucidated in future studies. A critical interactor of RECQ1, PARP1 has been shown to interact with breast cancer related histone-modifying enzymes and regulate gene expression in breast cancer cells (55, 56). Results of genomic analyses conducted in the MCF7 human breast cancer cell line have localized RECQ1 at gene promoters as well as transcription start sites and intronic regions in the genome. Our integrative genomics approach provides important new, information about the chromatin properties (e.g. histone marks dictating chromatin state) of RECQ1 binding sites. Whether binding of RECQ1 to these sites modify histones and thereby regulates chromatin openness is unknown. Additional factors and chromatin features that provide specificity to RECQ1 to be recruited at these subsets of genomic sites are unknown at this time.

What may be the role of RECQ1 in estrogen-dependent gene regulation in ERα-positive breast cancers? One key unifying feature of our RNA-Seq analysis upon RECQ1 knockdown and RECQ1 ChIP-seq analysis is the identification of estrogen response as the top enriched pathway. A functional role of RECQ1 in regulating a subset of estrogen response genes that underlie proliferation of breast cancer cells is evident by significantly compromised stimulation of

proliferation of RECQ1-knockdown MCF7 and T47D cells upon estrogen treatment. Comparison of RECQ1 ChIP-seq peaks with RNA POLII occupancy after estrogen treatment shows a significant correlation with the dynamic transcriptional changes and suggests a role for RECQ1 as a mediator of early transcriptional response in estrogen signaling. This is consistent with our findings that RECQ1 cooperates with FOXA1 which is essential to reprogram the genome-wide occupancy of the ERα and mediate transcriptional response in estrogen signaling (57). However, we do not yet know if RECQ1 is required for the implementation of an estrogen-regulated transcriptional program in an ERα-dependent manner.

A clear majority of the breast cancer patients express ERα in their tumors making the tumors amenable to endocrine therapies for ERα positive breast cancer patients (33, 35). We have previously shown that RECQ1 deficiency was not only associated with aggressive breast cancer phenotypes but in ERα-positive tumors that received endocrine therapy, low RECQ1 expression was also linked with poor survival (32). In another clinical study, high FOXA1 expression was associated with ERα-positive tumors, smaller tumor size, lower histological grade, and better survival (52). This is further supported by a recent study suggesting that high FOXA1 may predicting late recurrence in patients with ER-positive breast cancer (58). Although the mechanisms of endocrine resistance in breast cancer is complex (36), studies of response to selective estrogen receptor modulators (SERMs) like tamoxifen indicate that negative (co-repressors) and positive (co-activators) co-regulators of ERα may influence the balance of agonistic versus antagonistic activities of tamoxifen and determine endocrine sensitivity or resistance (36). Favorable survival outcomes in high RECQ1/high FOXA1 expressing ERα-positive tumors indicate that functional interaction between ERα, RECQ1, and FOXA1 may contribute to the anti-cancer activity of tamoxifen. However, a limitation of the clinical study is

that it is retrospective and further prospective validation will be required to confirm our initial observations.

In conclusion, we have demonstrated that RECQ1 can modulate a subset of ER-driven gene expression by regulating the expression of *ESR1* in a helicase activity-dependent and FOXA1-assisted manner. Cooperation of RECQ1 and FOXA1 is clinically relevant in response to endocrine therapy in ER-positive breast cancer. Together, this study provides the first mechanism by which RECQ1 could alter the progression and therapeutic response of ER-positive breast cancer. The notion that RECQ1 has a specific and direct role in gene expression control is novel and of potential significance to cancer biology. *RECQ1* is a breast cancer susceptibility gene significantly correlated with clinical outcomes of sporadic breast cancer patients. Given that ERα is the major driving transcription factor in the mammary gland development as well as breast cancer initiation and progression (59), evaluating the impact of breast cancer risk-associated *RECQ1* variants on ER signaling may uncover the mechanisms by which RECQ1 helicase acts to suppress breast cancer development and progression. Elucidating the role of RECQ1 in coordinating genomic stability with transcriptomic networks could potentially predict cancer risk, achieve early diagnosis, track the prognosis of tumor fate, and ultimately provide valuable targets for novel therapeutic approaches.

Data Availability

RNA-seq data upon RECQ1 and ESR1 knockdown in MCF7 cells have been deposited in GEO: The accession number is GSE152323 and the reviewer token is qvepomgmvvobfel. ChIP-seq data for RECQ1 and ERα have been deposited in GEO: The accession number is GSE153286 and the Reviewer token to download the data: obilweoodlwhfwf

Conflict of interest statement. None declared

Acknowledgements

We thank the NCI sequencing core facility for performing RNA-seq. This work was supported by the National Institute of General Medical Sciences of the National Institutes of Health (under Award Number SC1GM093999); National Science Foundation (under Award Number NSF1832163), National Institute on Aging (under Award Number 1R25 AG047843) to (S.S.) and the Intramural Research Program (A.L., S.A., J.K, and M.I.A.) of the National Cancer Institute (NCI), Center for Cancer Research (CCR), NIH. B.B. was supported by a graduate student fellowship from Umm Al-Qura University, Saudi Arabia.

Materials and Methods

Cell culture, transfection, and treatment

MCF7 and T47D human breast cancer cell lines were purchased from the American Type Culture Collection (ATCC). All cells were grown as monolayers in Dulbecco's modified Eagle's medium (DMEM, Thermo Fisher) supplemented with 1% (vol/vol) sodium pyruvate, 10% (vol/vol) fetal bovine serum (FBS) and 10 IU/ml penicillin/streptomycin (Life Technologies) and maintained at 37°C, 5% CO2 in a humidified incubator. On-Target plus SMARTpool small interfering RNAs (siRNAs) against *RECQ1, ESR1, FOXA1,* and control siRNAs were purchased from Dharmacon. All siRNA transfections were performed by reverse transfection at a final concentration of 20 nM of the siRNA using Lipofectamine RNAiMAX (Invitrogen) as instructed by the manufacturer. Isogenic MCF7 RECQ1 knockout (RECQ1-KO) and its wild type control (RECQ1-WT) cells were generated using the CRISPR/Cas9 system (19).

Cell proliferation assays

Cell Counting Kit-8 (CCK-8, Dojindo Molecular Technologies) was used to determine the effect of Estrogen (17β-Estradiol (E2, Sigma)) on RECQ1 knockdown MCF7 and T47D cell proliferation. Forty-eight hours after siRNA transfection, MCF7 or T47D cells were plated in triplicates in a 96-well plate (8×10^3 cells /well) and incubated with phenol-red free DMEM containing 5% charcoal-stripped FBS (Gibco) for 24 hr at 37°C in 5% CO_2. The next day, the cells were treated with 10 nM E2 or 0.1% ethanol and incubated for an additional 48 hr at 37°C in 5% CO_2. Cell proliferation was measured by adding 10 µl of CCK-8 reagent to each well containing 100 µl of growth medium. The plates were incubated at 37ºC, and absorbance was measured at 450 nM every hour for 4 hr. The relative cell proliferation was calculated by normalizing the absorbance values to the untreated condition in each cell type.

RNA isolation and RT-qPCR

Total RNA from MCF7 and T47D cells was extracted using the TRIzol reagent (Invitrogen) according to the manufacturer's instructions. A total of 0.5 µg of RNA was used for reverse transcription using the iScript Reverse Transcription Supermix kit (Bio-Rad) according to the manufacturer's instructions. The cDNA was subjected to real-time quantitative PCR using iTaq Universal SYBR Green Supermix (Bio-Rad) in triplicate. Reactions were cycled at 95°C for 30 s; followed by 40 cycles of 94°C for 10 s and 60°C for 15s with fluorescence data collection during the anneal/extension step on the CFX96 Real-Time PCR System (Bio-Rad). The relative transcript levels were normalized to the housekeeping gene *GAPDH* and differential expression measured using the $2^{-\Delta\Delta CT}$ method (60). The housekeeping gene *SDHA* served as a negative control in RT-qPCR experiments. The sequence of primers are detailed in Supplementary Table S1.

RNA-seq

For RNA-seq, ~10 million MCF7 cells transfected with CTL siRNA, *ESR1* siRNAs (smartpool) or *RECQ1* siRNAs (smartpool) for 48 hr were harvested and total RNA was isolated with the use of the RNeasy plus micro kit (Qiagen). RNA integrity was checked with a Bioanalyzer (Agilent), and only samples with an RNA integrity number (RIN) of >9.5 were subsequently subjected to mRNA-seq. The mRNA-seq samples were pooled and sequenced on HiSeq using Illumina TruSeq mRNA Prep Kit RS-122-2101 and paired-end sequencing. The samples had ~79 to 101 million pass filter reads with a base call quality of above ~90% of bases with Q30 and above. Reads of the samples were trimmed for adapters and low-quality bases using Trimmomatic software before alignment with the reference genome (Human - hg19) and the annotated transcripts using STAR. The average mapping rate of all samples was ~95%. Unique alignment is above 89%. The mapping statistics are calculated using Picard software. The samples had ~0.88% ribosomal reads. Percent coding bases were between 64-66%. Percent UTR bases are 29-31%, and mRNA bases were between 93-94% for all the samples. Library complexity was measured in terms of unique fragments in the mapped reads using Picard's MarkDuplicate utility. The samples have 64-70% non-duplicate reads.

Read count per gene was calculated by HTSeq under the annotation of Gencode and normalized by size factor implemented in the *DESeq2* package. Regularized-logarithm transformation (rlog) values of gene expression were used to perform hierarchical clustering and principal component analysis. To assess differential gene expression between different conditions (e.g., constructs vs. mocks), we used a generalized linear model within *DESeq2* that incorporates information from counts and uses negative binomial distribution with fitted mean and a gene-

specific dispersion parameter. *DESeq2* used Wald statistics for significance testing and the Benjamini-Hochberg adjustment for multiple corrections.

Immunoblotting

MCF7 and T47D cells were harvested after washing with phosphate-buffered saline (PBS) and whole cell lysates were prepared using radioimmunoprecipitation assay (RIPA) buffer containing protease inhibitor cocktail (Roche). The bicinchoninic acid (BCA) protein quantification kit (Thermo Scientific) was used to perform protein quantification. A total of 50 µg of the whole cell lysate was mixed with Laemmli Buffer (BioRad), boiled at 95°C for 5 min, and subjected to sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) followed by transfer to a polyvinylidene difluoride membrane (PVDF, Life Technologies). The membrane and primary antibodies, anti-RECQ1 (Bethyl lab), anti-ERα (Abcam), anti-FOXA1(Abcam), and anti-GAPDH (Cell Signaling), were incubated on a rotating platform overnight at 4°C and followed by secondary antibodies conjugated to HRP (Jackson Immuno Research Laboratories). Chemiluminescent HRP substrate Immobilon Western kit (Invitrogen) was used to develop the immunoblots.

Co-Immunoprecipitation

Whole cell lysate of MCF7 cells prepared in RIPA buffer (1 mg total protein) was incubated with Dynabeads Protein A coupled with an antibody against human RECQ1 (Bethyl Lab) or normal rabbit IgG (Vector Labs) for over-night at 4°C in presence or absence of benzonase (Sigma, 50U/ml). Following four times washes with 1x PBS, the immunocomplexes were eluted with 2x SDS-sample buffer by boiling at 95°C for 5 min and resolved by 10% SDS-PAGE followed by

Western blot detection using specific antibodies against RECQ1 (Bethyl Lab), FOXA1(Abcam) and ERα (Santa Cruz Biotechnology).

Chromatin immunoprecipitation (ChIP) and re-ChIP

ChIP-IT High Sensitivity kit (Active Motif) was used to determine the associations of RECQ1, Pol II, and FOXA1 with the genomic regions on the promoter and enhancer of ERS1gene. MCF7 cells transfected with CTL siRNA or RECQ1 siRNAs, and RECQ1-KO or RECQ1-WT cells were grown at a density of 1×10^7 per 15 cm dish and subjected to fixation, chromatin sonication and immunoprecipitated using 4 μg of anti-RECQ1 (Bethyl Lab), anti-Pol II (Abcam), anti-FOXA1 (Abcam), or the same amount of rabbit IgG, followed by DNA purification according to manufacturer's instruction. The immunoprecipitated fraction was analyzed by qPCR using iTaq Universal SYBR Green Supermix (Bio-Rad) with technical triplicates, with an initial denaturation at 95°C for 3 min; then 40 cycles of denaturation at 95°C for 15 s and annealing/extension at 60°C for 30 s using the CFX96 Real-Time PCR System (Bio-Rad). Fold enrichment of the targeted genomic sequences were calculated over IgG as: fold enrichment = $2^{-(\mathrm{CtIP}-\mathrm{CtIgG}),}$ where CtIP and CtIgG are mean threshold cycles of PCR in triplicates on DNA samples immunoprecipitated with the specific antibody or IgG control, respectively. Melt curve analyses and agarose gel electrophoresis was used to confirm the presence of a single specific product after the qPCR. The sequences of primers are listed in Table S1.

For ChIP and re-ChIP experiments, cell lysates were collected as described above and incubated overnight incubation with the first antibody, either anti-RECQ1 (Bethyl Laboratories, Inc.) or anti-Pol II (Abcam) and Protein G Sepharose beads. Immuno-complexes were washed and followed by elution using 10 mM DTT at 37ºC for 30 min. The eluted DNA were diluted 50-fold

with lysis buffer and incubated with the second antibody, anti-Pol II or anti-RECQ1, for 3 hr at 4°C. Protein G Sepharose beads were then added and incubated overnight with rotation at 4°C. The following day, samples were processed according to the protocol of ChIP-IT High Sensitivity kit (Active Motif), and subsequent qPCR analysis was carried out as described above.

ChIP-seq

Chromatin from MCF7 cells was fragmented by sonication using a Bioruptor (Diagnode) at 9 strength outcome (30 cycles, 30 seconds on plus 60 seconds off) to generate fragments less than 200 bp long and confirmed by agarose gel electrophoresis. Sheared DNA fragments were divided into two parts. One part was immunoprecipitated with specific RECQ1 and ERα antibodies as described above for ChIP-qPCR and the second part was used as the corresponding input. After purification using the ChIP-IT High Sensitivity kit (Active Motif), immunoprecipitated DNA and input DNA were converted into sequencing libraries using the TruSeq ChIP Sample Preparation kit (Illumina), which were then sequenced in single-end 75-bp sequencing using a NextSeq 500 system (Illumina). The samples had ~30 to 64 million pass filter reads with a base call quality of above ~90% of bases with Q30 and above.

Peak calling: RECQ1 and ERα ChIP-seq peaks were called against IgG controls by using the MACS2 algorithm from the Genomatix genome analyzer (https://www.genomatix.de/solutions/genomatix-genome-analyzer.html), with the following parameters: broad region calling on, band width = 300, qvalue cutoff = 1.00e-02 and model fold = [5, 50].

Heatmap and Colocalization analysis: The normalized sequencing coverage for the RECQ1 and ERα ChIP-seq data was created using *BAMscale* (61). Colocalization of signal heatmap was created using the *rtracklayer* (version1.48.0) (62) and *ComplexHeatmap* (63) (version 2.4.3) packages in R (version 4.0.2, script available at: https://github.com/ncbi/BAMscale) using the RECQ1 called peaks (n=7,023) with the RECQ1 and ERα ChIP-seq data, along with FOXA1 (ENCFF255FPM), GATA3 (ENCFF477GZL), H3K27ac (ENCFF411FCW), H3K4me1 (ENCFF983TTS), H3K4me3 (ENCFF862CKA) and H3K9me3 (ENCFF688REP) data downloaded from ENCODE. Peaks were clustered using k-means clustering, setting the centers to 3, and using the H3K4me1 (enhancer mark) and H3K4me3 (promoter) histone marks for cluster identification. The RECQ1 peaks were extended by 5 kb upstream and downstream when creating the heatmaps.

Colocalization analyses between RECQ1 and other ChIP-seq bed files or transcription start regions were performed using the genome inspector program of the genomatix genome analyzer. Colocalization analysis were performed between the binding sites of POL2RA from cells collected over a time-course of 0 to 320 min after estradiol treatment (42) and binding sites of RECQ1 (duplicates, this study), ERα (duplicates, this study), PAX8, REST, FOXA1 and GATA3. Publicly available ChIPseq bed files were obtained from Cistrome ((http://cistrome.org/) (64) and ENCODE: POLR2A (GSM1091921 [control], GSM1091915 [10 min], GSM1091916 [20 min], GSM1091917 [40 min], GSM1091918 [80 min], GSM1091919 [160 min], GSM1091920 [320 min]; PAX8 (GSM2828671, GSM2828670), REST (GSM1010891), FOXA1 (ENCFF255FPM, GSM1534737, GSM3092505, GSM798437, GSM798436) and GATA3 (ENCFF477GZL, GSM986068, GSM1241752, GSM720423). The genome inspector program was used to quantify colocalization in 5 kb windows between anchor sets for RECQ1 (n=2 datasets), ERα (n=2), PAX8

(n=2), REST (n=1), FOXA1 (n=5) and GATA3 (n=4) and each of the partner sets POLR2A binding sites for 0, 10, 20, 40, 60, 160 and 320 min (n=1 dataset each). Averages of colocalization values for each anchor sets were normalized using the time point t=0 for estradiol treatment. Colocalization analysis between the RECQ1 binding sites (7,023 peaks) and the transcriptional start regions (TSRs, 291,570 regions) was also performed. As per the genomatix genome annotation and analysis project, TSRs are defined as regions of genomic sequence for which experimental evidence for transcription initiation is available. The extent of colocalization (%) was measured as the fraction of RECQ1 peaks within the 5-kb window of the TSRs.

Validation of ChIP-seq data: 10 genes were selected for validation of ChIP-seq data by ChIP-qPCR. The selection of genes is based on three conditions. (1) The genomic sequence of the gene was commonly bound by RECQ1 and Pol II, (2) at least one binding site of RECQ1 or Pol II was located in the promoter region of the gene, and (3) the functional relationship of the gene with *ERS1* had been previously reported.

Formaldehyde-assisted isolation of regulatory elements (FAIRE) and FAIRE-qPCR

FAIRE and FAIRE_qPCR was carried out based on a protocol described by Rodríguez-Gil et al (49). Briefly, 5×10^6 of MCF7 cells were cross-linked with 1% formaldehyde for 10 min at room temperature and then quenched with 125mM glycine. Fixed cells were lysed in SDS lysis buffer and chromatin-bound DNA sheared by sonication to obtain fragments around 200-300bp length, representing 1-2 nucleosomes. Each sample was divided into two aliquots. One aliquot was treated with proteinase K at 37°C for 4 hr and followed by incubation at 65°C for 6 hr to reverse the crosslink, and another left untreated. 1% of untreated aliquot was used as an input. All samples

were then subjected to three consecutive phenol-chloroform extractions extracted by phenol/chloroform. The purified DNAs were subjected to qPCR analyses and primers used were the same as those used in ChIP qPCR, encompassing the promoter and enhancer of *ESR1*. The promoter of *GAPDH*, an actively transcribed housekeeping gene, served as positive control, and a heterochromatin region on chromosome 12 was used as a negative control. The relative enrichment for each amplicon was calculated using the comparative *Ct* method such that a ratio is calculated for the signal from the FAIRE sample relative to the signal from input control DNA (65).

Clinical study of RECQ1 and FOXA1 expression in ERα-positive human breast cancers

Patients: The study was performed in a consecutive series of patients with ERα $^{+}$ primary invasive breast carcinomas who were diagnosed between 1986 and 1999 and entered into the Nottingham Tenovus Primary Breast Carcinoma series. Patient demographics are summarized in Supplementary Table S11. This is a well-characterized series of patients with long-term follow-up that has been investigated in a wide range of biomarker studies (32). All patients were treated uniformly in a single institution with standard surgery (mastectomy or wide local excision), followed by Radiotherapy. Prior to 1989, patients did not receive systemic adjuvant treatment (AT). After 1989, AT was scheduled based on prognostic and predictive factor status, including Nottingham Prognostic Index (NPI), ERα status, and menopausal status. Patients with NPI scores of <3.4 (low risk) did not receive AT. In premenopausal patients with NPI scores of ≥3.4 (high risk), classical Cyclophosphamide, Methotrexate, and 5-Fluorouracil (CMF) chemotherapy was given; patients with ER-α positive tumors were also offered endocrine therapy. Postmenopausal patients with NPI scores of ≥3.4 and ERα positivity were offered endocrine therapy. Median follow up was 111 months (range 1 to 233 months). Survival data, including breast cancer-specific survival (BCSS), was maintained on a prospective basis. Breast cancer-specific survival (BCSS)

was defined as the number of months from diagnosis to the occurrence of BC related-death. Survival was censored if the patient was still alive at the time of analysis, lost to follow-up, or died from other causes. Tumor Marker Prognostic Studies (REMARK) criteria, recommended by McShane et al (66), were followed throughout this study. Ethical approval was obtained from the Nottingham Research Ethics Committee (C202313).

Tissue Microarrays (TMAs) and immunohistochemistry (IHC): Tumors were arrayed in tissue microarrays (TMAs) constructed with 0.6mm cores sampled from the periphery of the tumors. The TMAs were immunohistochemically profiled as described previously for RECQ1 (32), FOXA1 (52), and other biological antibodies (Supplementary Table S12). Immunohistochemical staining was performed using the Thermo Scientific Shandon Sequenza chamber system (REF: 72110017), in combination with the Novolink Max Polymer Detection System (RE7280-K: 1250 tests), and the Leica Bond Primary Antibody Diluent (AR9352), each used according to the manufacturer's instructions (Leica Microsystems). Leica Autostainer XL machine was used to dewax and rehydrate the slides. Pre-treatment antigen retrieval was performed on the TMA sections using sodium citrate buffer (pH 6.0) and heated for 20 min at 95^0C in a microwave (Whirpool JT359 Jet Chef 1000W). A set of slides were incubated for 60 min with the primary anti-RECQ1 antibody (Bethyl Laboratories, catalog no. A300-450A) at a dilution of 1:1000 respectively. Negative and positive (by omission of the primary antibody and IgG-matched serum) controls were included in each run. The negative control ensured that all the staining was produced from the specific interaction between antibody and antigen.

Mouse monoclonal antibody to FOXA1 (clone 2F83, ab40868; Abcam, Cambridge, UK) was optimized at a working dilution of 1:2000 using full-face sections of mouse fetal lung tissue

as a positive control tissue. Immunohistochemical staining of FOXA1 was carried out using a Dako Cytomation Techmate500 plus (Dako Cytomation, Cambridge, UK) automatic immune stainer with a linked streptavidin-biotin technique following the manufacturer's instructions after microwave antigen retrieval in citrate buffer (pH 6.0). Negative controls were performed by omitting the primary antibody. Sections were counterstained in hematoxylin and coverslipped using DPX mounting medium.

Evaluation of immune staining: Whole field inspection of the core was scored, and intensities of nuclear staining were grouped as follows: 0 = no staining, 1 = weak staining, 2 = moderate staining, 3 = strong staining. The percentage of each category was estimated (0-100%). H-score (range 0-300) was calculated by multiplying the intensity of staining and percentage staining. RECQ1 expression was categorized based on the frequency histogram distributions. The tumor cores were evaluated by two scorers (AA and MA) and the concordance between the two scorers was excellent ($k = 0.79$). Xtile (Version 3.6.1) was used to identify a cut-off in protein expression values such that the resulting subgroups had significantly different survival courses. An H score of ≥ 215 was taken as the cut-off for high RECQ1 level (32). Not all cores within the TMA were suitable for IHC assessments as some cores were missing or containing inadequate invasive cancer (<15% tumor). The FOXA1 H-score cut-off point for determining positive and negative staining was chosen as the median of the H-score of the informative cases (H-score P10) (52). HER2 scoring was performed using the manufacturer recommendations (Ventana Medical System, AZ, USA). Breast carcinomas that were considered positive for HER2 protein overexpression met threshold criteria for the intensity and pattern of membrane staining (2+ or greater on a scale of 0 to 3+) and for the percentage of positive tumor cells (>10%).

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. 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.

1. Sharma, S., D. J. Stumpo, A. S. Balajee, C. B. Bock, P. M. Lansdorp, R. M. Brosh Jr, and P. J. Blackshear. 2007. RECQL, a member of the RecQ family of DNA helicases, suppresses chromosomal instability. Mol. Cell. Biol. 27:1784-1794. doi: MCB.01620-06 [pii].

2. Sharma, S. 2014. An appraisal of RECQ1 expression in cancer progression. Front. Genet. 5:426. doi: 10.3389/fgene.2014.00426 [doi].

3. Mendoza-Maldonado, R., V. Faoro, S. Bajpai, M. Berti, F. Odreman, M. Vindigni, T. Ius, A. Ghasemian, S. Bonin, M. Skrap, G. Stanta, and A. Vindigni. 2011. The human RECQ1 helicase is highly expressed in glioblastoma and plays an important role in tumor cell proliferation. Mol. Cancer. 10:83-4598-10-83. doi: 10.1186/1476-4598-10-83 [doi].

4. Li, D., H. Liu, L. Jiao, D. Z. Chang, G. Beinart, R. A. Wolff, D. B. Evans, M. M. Hassan, and J. L. Abbruzzese. 2006. Significant effect of homologous recombination DNA repair gene polymorphisms on pancreatic cancer survival. Cancer Res. 66:3323-3330. doi: 66/6/3323 [pii].

5. Viziteu, E., B. Klein, J. Basbous, Y. L. Lin, C. Hirtz, C. Gourzones, L. Tiers, A. Bruyer, L. Vincent, C. Grandmougin, A. Seckinger, H. Goldschmidt, A. Constantinou, P. Pasero, D. Hose, and J. Moreaux. 2017. RECQ1 helicase is involved in replication stress survival and drug resistance in multiple myeloma. Leukemia. 31:2104-2113. doi: 10.1038/leu.2017.54 [doi].

6. Bowden, A. R., and M. Tischkowitz. 2019. Clinical implications of germline mutations in breast cancer genes: RECQL. Breast Cancer Res. Treat. 174:553-560. doi: 10.1007/s10549-018-05096-6 [doi].

7. Cybulski, C., J. Carrot-Zhang, W. Kluzniak, B. Rivera, A. Kashyap, D. Wokolorczyk, S. Giroux, J. Nadaf, N. Hamel, S. Zhang, T. Huzarski, J. Gronwald, T. Byrski, M. Szwiec, A. Jakubowska, H. Rudnicka, M. Lener, B. Masojc, P. N. Tonin, F. Rousseau, B. Gorski, T. Debniak, J. Majewski, J. Lubinski, W. D. Foulkes, S. A. Narod, and M. R. Akbari. 2015. Germline RECQL mutations are associated with breast cancer susceptibility. Nat. Genet. 47:643-646. doi: 10.1038/ng.3284 [doi].

8. Sun, J., Y. Wang, Y. Xia, Y. Xu, T. Ouyang, J. Li, T. Wang, Z. Fan, T. Fan, B. Lin, H. Lou, and Y. Xie. 2015. Mutations in RECQL Gene Are Associated with Predisposition to Breast Cancer. PLoS Genet. 11:e1005228. doi: 10.1371/journal.pgen.1005228 [doi].

9. Sharma, S., K. M. Doherty, and R. M. Brosh Jr. 2006. Mechanisms of RecQ helicases in pathways of DNA metabolism and maintenance of genomic stability. Biochem. J. 398:319-337. doi: BJ20060450 [pii].

10. Chu, W. K., and I. D. Hickson. 2009. RecQ helicases: multifunctional genome caretakers. Nat. Rev. Cancer. 9:644-654. doi: 10.1038/nrc2682 [doi].

11. Brosh, R. M.,Jr. 2013. DNA helicases involved in DNA repair and their roles in cancer. Nat. Rev. Cancer. 13:542-558. doi: 10.1038/nrc3560 [doi].

12. Croteau, D. L., V. Popuri, P. L. Opresko, and V. A. Bohr. 2014. Human RecQ helicases in DNA repair, recombination, and replication. Annu. Rev. Biochem. 83:519-552. doi: 10.1146/annurev-biochem-060713-035428 [doi].

13. Sharma, S., J. A. Sommers, S. Choudhary, J. K. Faulkner, S. Cui, L. Andreoli, L. Muzzolini, A. Vindigni, and R. M. Brosh Jr. 2005. Biochemical analysis of the DNA unwinding and strand annealing activities catalyzed by human RECQ1. J. Biol. Chem. 280:28072-28084. doi: M500264200 [pii].

14. Pike, A. C., S. Gomathinayagam, P. Swuec, M. Berti, Y. Zhang, C. Schnecke, F. Marino, F. von Delft, L. Renault, A. Costa, O. Gileadi, and A. Vindigni. 2015. Human RECQ1 helicase-driven DNA unwinding, annealing, and branch migration: insights from DNA complex structures. Proc. Natl. Acad. Sci. U. S. A. 112:4286-4291. doi: 10.1073/pnas.1417594112 [doi].

15. Sharma, S., P. Phatak, A. Stortchevoi, M. Jasin, and J. R. Larocque. 2012. RECQ1 plays a distinct role in cellular response to oxidative DNA damage. DNA Repair (Amst). 11:537-549. doi: 10.1016/j.dnarep.2012.04.003 [doi].

16. Popuri, V., D. L. Croteau, R. M. Brosh Jr, and V. A. Bohr. 2012. RECQ1 is required for cellular resistance to replication stress and catalyzes strand exchange on stalled replication fork structures. Cell. Cycle. 11:4252-4265. doi: 10.4161/cc.22581 [doi].

17. Berti, M., A. Ray Chaudhuri, S. Thangavel, S. Gomathinayagam, S. Kenig, M. Vujanovic, F. Odreman, T. Glatter, S. Graziano, R. Mendoza-Maldonado, F. Marino, B. Lucic, V. Biasin, M. Gstaiger, R. Aebersold, J. M. Sidorova, R. J. Monnat Jr, M. Lopes, and A. Vindigni. 2013. Human RECQ1 promotes restart of replication forks reversed by DNA topoisomerase I inhibition. Nat. Struct. Mol. Biol. 20:347-354. doi: 10.1038/nsmb.2501 [doi].

18. Lu, X., S. Parvathaneni, T. Hara, A. Lal, and S. Sharma. 2013. Replication stress induces specific enrichment of RECQ1 at common fragile sites FRA3B and FRA16D. Mol. Cancer. 12:29-4598-12-29. doi: 10.1186/1476-4598-12-29 [doi].

19. Parvathaneni, S., and S. Sharma. 2019. The DNA repair helicase RECQ1 has a checkpoint-dependent role in mediating DNA damage responses induced by gemcitabine. J. Biol. Chem. 294:15330-15345. doi: 10.1074/jbc.RA119.008420 [doi].

20. Banerjee, T., J. A. Sommers, J. Huang, M. M. Seidman, and R. M. Brosh Jr. 2015. Catalytic Strand Separation by RECQ1 Is Required for RPA-Mediated Response to Replication Stress. Curr. Biol. 25:2830-2838. doi: 10.1016/j.cub.2015.09.026 [doi].

21. Parvathaneni, S., A. Stortchevoi, J. A. Sommers, R. M. Brosh Jr, and S. Sharma. 2013. Human RECQ1 interacts with Ku70/80 and modulates DNA end-joining of double-strand breaks. PLoS One. 8:e62481. doi: 10.1371/journal.pone.0062481 [doi].

22. Woodrick, J., S. Gupta, S. Camacho, S. Parvathaneni, S. Choudhury, A. Cheema, Y. Bai, P. Khatkar, H. V. Erkizan, F. Sami, Y. Su, O. D. Schärer, S. Sharma, and R. Roy. 2017. A new sub-pathway of long-patch base excision repair involving 5' gap formation. Embo J. 36:1605-1622. doi: 10.15252/embj.201694920 [doi].

23. Sami, F., X. Lu, S. Parvathaneni, R. Roy, R. K. Gary, and S. Sharma. 2015. RECQ1 interacts with FEN-1 and promotes binding of FEN-1 to telomeric chromatin. Biochem. J. 468:227-244. doi: 10.1042/BJ20141021 [doi].

24. Popuri, V., J. Hsu, P. Khadka, K. Horvath, Y. Liu, D. L. Croteau, and V. A. Bohr. 2014. Human RECQL1 participates in telomere maintenance. Nucleic Acids Res. 42:5671-5688. doi: 10.1093/nar/gku200 [doi].

25. Debnath, S., and S. Sharma. 2020. RECQ1 Helicase in Genomic Stability and Cancer. Genes (Basel). 11:E622. doi: 10.3390/genes11060622. doi: E622 [pii].

26. Sharma, S., and R. M. Brosh Jr. 2007. Human RECQ1 is a DNA damage responsive protein required for genotoxic stress resistance and suppression of sister chromatid exchanges. PLoS One. 2:e1297. doi: 10.1371/journal.pone.0001297 [doi].

27. Sami, F., and S. Sharma. 2013. Probing Genome Maintenance Functions of human RECQ1. Comput. Struct. Biotechnol. J. 6:e201303014. doi: 10.5936/csbj.201303014 [doi].

28. Cogoni, C., and G. Macino. 1999. Posttranscriptional gene silencing in Neurospora by a RecQ DNA helicase. Science. 286:2342-2344. doi: 8089 [pii].

29. Lau, N. C., A. G. Seto, J. Kim, S. Kuramochi-Miyagawa, T. Nakano, D. P. Bartel, and R. E. Kingston. 2006. Characterization of the piRNA complex from rat testes. Science. 313:363-367. doi: 1130164 [pii].

30. Lu, X., S. Parvathaneni, X. L. Li, A. Lal, and S. Sharma. 2016. Transcriptome guided identification of novel functions of RECQ1 helicase. Methods. 108:111-117. doi: 10.1016/j.ymeth.2016.04.018 [doi].

31. Li, X. L., X. Lu, S. Parvathaneni, S. Bilke, H. Zhang, S. Thangavel, A. Vindigni, T. Hara, Y. Zhu, P. S. Meltzer, A. Lal, and S. Sharma. 2014. Identification of RECQ1-regulated transcriptome uncovers a role of RECQ1 in regulation of cancer cell migration and invasion. Cell. Cycle. 13:2431-2445. doi: 10.4161/cc.29419 [doi].

32. Arora, A., S. Parvathaneni, M. A. Aleskandarany, D. Agarwal, R. Ali, T. Abdel-Fatah, A. R. Green, G. R. Ball, E. A. Rakha, I. O. Ellis, S. Sharma, and S. Madhusudan. 2017. Clinicopathological and Functional Significance of RECQL1 Helicase in Sporadic Breast Cancers. Mol. Cancer. Ther. 16:239-250. doi: 10.1158/1535-7163.MCT-16-0290 [doi].

33. Carroll, J. S. 2016. Mechanisms of oestrogen receptor (ER) gene regulation in breast cancer. Eur. J. Endocrinol. 175:R41-9. doi: 10.1530/EJE-16-0124 [doi].

34. Clarke, R., M. C. Liu, K. B. Bouker, Z. Gu, R. Y. Lee, Y. Zhu, T. C. Skaar, B. Gomez, K. O'Brien, Y. Wang, and L. A. Hilakivi-Clarke. 2003. Antiestrogen resistance in breast cancer and the role of estrogen receptor signaling. Oncogene. 22:7316-7339. doi: 1206937 [pii].

35. Musgrove, E. A., and R. L. Sutherland. 2009. Biological determinants of endocrine resistance in breast cancer. Nat. Rev. Cancer. 9:631-643. doi: 10.1038/nrc2713 [doi].

36. Osborne, C. K., and R. Schiff. 2011. Mechanisms of endocrine resistance in breast cancer. Annu. Rev. Med. 62:233-247. doi: 10.1146/annurev-med-070909-182917 [doi].

37. Zhao, L., S. Zhou, and J. Å Gustafsson. 2019. Nuclear Receptors: Recent Drug Discovery for Cancer Therapies. Endocr. Rev. 40:1207-1249. doi: 10.1210/er.2018-00222 [doi].

38. Carroll, J. S., and M. Brown. 2006. Estrogen receptor target gene: an evolving concept. Mol. Endocrinol. 20:1707-1714. doi: me.2005-0334 [pii].

39. Lupien, M., and M. Brown. 2009. Cistromics of hormone-dependent cancer. Endocr. Relat. Cancer. 16:381-389. doi: 10.1677/ERC-09-0038 [doi].

40. Ciocca, D. R., and M. A. Fanelli. 1997. Estrogen receptors and cell proliferation in breast cancer. Trends Endocrinol. Metab. 8:313-321. doi: S1043-2760(97)00122-7 [pii].

41. Holmes, K. A., G. D. Brown, and J. S. Carroll. 2016. Chromatin Immunoprecipitation-Sequencing (ChIP-seq) for Mapping of Estrogen Receptor-Chromatin Interactions in Breast Cancer. Methods Mol. Biol. 1366:79-98. doi: 10.1007/978-1-4939-3127-9_8 [doi].

42. wa Maina, C., A. Honkela, F. Matarese, K. Grote, H. G. Stunnenberg, G. Reid, N. D. Lawrence, and M. Rattray. 2014. Inference of RNA polymerase II transcription dynamics from chromatin immunoprecipitation time course data. PLoS Comput. Biol. 10:e1003598. doi: 10.1371/journal.pcbi.1003598 [doi].

43. Ellison-Zelski, S. J., N. M. Solodin, and E. T. Alarid. 2009. Repression of ESR1 through actions of estrogen receptor alpha and Sin3A at the proximal promoter. Mol. Cell. Biol. 29:4949-4958. doi: 10.1128/MCB.00383-09 [doi].

44. Powers, G. L., S. J. Ellison-Zelski, A. J. Casa, A. V. Lee, and E. T. Alarid. 2010. Proteasome inhibition represses ERalpha gene expression in ER+ cells: a new link between proteasome

activity and estrogen signaling in breast cancer. Oncogene. 29:1509-1518. doi: 10.1038/onc.2009.434 [doi].

45. Carroll, J. S., X. S. Liu, A. S. Brodsky, W. Li, C. A. Meyer, A. J. Szary, J. Eeckhoute, W. Shao, E. V. Hestermann, T. R. Geistlinger, E. A. Fox, P. A. Silver, and M. Brown. 2005. Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. Cell. 122:33-43. doi: S0092-8674(05)00453-8 [pii].

46. Jozwik, K. M., and J. S. Carroll. 2012. Pioneer factors in hormone-dependent cancers. Nat. Rev. Cancer. 12:381-385. doi: 10.1038/nrc3263 [doi].

47. Li, X., W. Wang, J. Wang, A. Malovannaya, Y. Xi, W. Li, R. Guerra, D. H. Hawke, J. Qin, and J. Chen. 2015. Proteomic analyses reveal distinct chromatin-associated and soluble transcription factor complexes. Mol. Syst. Biol. 11:775. doi: 10.15252/msb.20145504 [doi].

48. Glont, S. E., I. Chernukhin, and J. S. Carroll. 2019. Comprehensive Genomic Analysis Reveals that the Pioneering Function of FOXA1 Is Independent of Hormonal Signaling. Cell. Rep. 26:2558-2565.e3. doi: S2211-1247(19)30207-4 [pii].

49. Rodríguez-Gil, A., T. Riedlinger, O. Ritter, V. V. Saul, and M. L. Schmitz. 2018. Formaldehyde-assisted Isolation of Regulatory Elements to Measure Chromatin Accessibility in Mammalian Cells. J. Vis. Exp. (134):57272. doi:10.3791/57272. doi: 10.3791/57272 [doi].

50. Iwafuchi-Doi, M., G. Donahue, A. Kakumanu, J. A. Watts, S. Mahony, B. F. Pugh, D. Lee, K. H. Kaestner, and K. S. Zaret. 2016. The Pioneer Transcription Factor FoxA Maintains an

Accessible Nucleosome Configuration at Enhancers for Tissue-Specific Gene Activation. Mol. Cell. 62:79-91. doi: S1097-2765(16)00179-9 [pii].

51. Kang, S. H., C. M. Kiefer, and T. P. Yang. 2003. Role of the promoter in maintaining transcriptionally active chromatin structure and DNA methylation patterns in vivo. Mol. Cell. Biol. 23:4150-4161. doi: 1813 [pii].

52. Habashy, H. O., D. G. Powe, E. A. Rakha, G. Ball, C. Paish, J. Gee, R. I. Nicholson, and I. O. Ellis. 2008. Forkhead-box A1 (FOXA1) expression in breast cancer and its prognostic significance. Eur. J. Cancer. 44:1541-1551. doi: 10.1016/j.ejca.2008.04.020 [doi].

53. Green, K. A., and J. S. Carroll. 2007. Oestrogen-receptor-mediated transcription and the influence of co-factors and chromatin state. Nat. Rev. Cancer. 7:713-722. doi: 10.1038/nrc2211 [doi].

54. Voss, T. C., R. L. Schiltz, M. H. Sung, P. M. Yen, J. A. Stamatoyannopoulos, S. C. Biddie, T. A. Johnson, T. B. Miranda, S. John, and G. L. Hager. 2011. Dynamic exchange at regulatory elements during chromatin remodeling underlies assisted loading mechanism. Cell. 146:544-554. doi: 10.1016/j.cell.2011.07.006 [doi].

55. Krishnakumar, R., and W. L. Kraus. 2010. PARP-1 regulates chromatin structure and transcription through a KDM5B-dependent pathway. Mol. Cell. 39:736-749. doi: 10.1016/j.molcel.2010.08.014 [doi].

56. Schiewer, M. J., and K. E. Knudsen. 2014. Transcriptional roles of PARP1 in cancer. Mol. Cancer. Res. 12:1069-1080. doi: 10.1158/1541-7786.MCR-13-0672 [doi].

57. Hurtado, A., K. A. Holmes, C. S. Ross-Innes, D. Schmidt, and J. S. Carroll. 2011. FOXA1 is a key determinant of estrogen receptor function and endocrine response. Nat. Genet. 43:27-33. doi: 10.1038/ng.730 [doi].

58. Horimoto, Y., N. Sasahara, R. Sasaki, M. T. Hlaing, A. Sakaguchi, H. Saeki, A. Arakawa, T. Himuro, and M. Saito. 2020. High FOXA1 protein expression might predict late recurrence in patients with estrogen-positive and HER2-negative breast cancer. Breast Cancer Res. Treat. 183:41-48. doi: 10.1007/s10549-020-05751-x [doi].

59. Chi, D., H. Singhal, L. Li, T. Xiao, W. Liu, M. Pun, R. Jeselsohn, H. He, E. Lim, R. Vadhi, P. Rao, H. Long, J. Garber, and M. Brown. 2019. Estrogen receptor signaling is reprogrammed during breast tumorigenesis. Proc. Natl. Acad. Sci. U. S. A. 116:11437-11443. doi: 10.1073/pnas.1819155116 [doi].

60. Livak, K. J., and T. D. Schmittgen. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 25:402-408. doi: S1046-2023(01)91262-9 [pii].

61. Pongor, L. S., J. M. Gross, R. Vera Alvarez, J. Murai, S. M. Jang, H. Zhang, C. Redon, H. Fu, S. Y. Huang, B. Thakur, A. Baris, L. Marino-Ramirez, D. Landsman, M. I. Aladjem, and Y. Pommier. 2020. BAMscale: quantification of next-generation sequencing peaks and generation of scaled coverage tracks. Epigenetics Chromatin. 13:21-020-00343-x. doi: 10.1186/s13072-020-00343-x [doi].

62. Lawrence, M., R. Gentleman, and V. Carey. 2009. rtracklayer: an R package for interfacing with genome browsers. Bioinformatics. 25:1841-1842. doi: 10.1093/bioinformatics/btp328 [doi].

63. Gu, Z., R. Eils, and M. Schlesner. 2016. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics. 32:2847-2849. doi: 10.1093/bioinformatics/btw313 [doi].

64. Zheng, R., C. Wan, S. Mei, Q. Qin, Q. Wu, H. Sun, C. H. Chen, M. Brown, X. Zhang, C. A. Meyer, and X. S. Liu. 2019. Cistrome Data Browser: expanded datasets and new tools for gene regulatory analysis. Nucleic Acids Res. 47:D729-D735. doi: 10.1093/nar/gky1094 [doi].

65. Simon, J. M., P. G. Giresi, I. J. Davis, and J. D. Lieb. 2012. Using formaldehyde-assisted isolation of regulatory elements (FAIRE) to isolate active regulatory DNA. Nat. Protoc. 7:256-267. doi: 10.1038/nprot.2011.444 [doi].

66. McShane, L. M., D. G. Altman, W. Sauerbrei, S. E. Taube, M. Gion, G. M. Clark, and Statistics Subcommittee of the NCI-EORTC Working Group on Cancer Diagnostics. 2005. Reporting recommendations for tumor marker prognostic studies (REMARK). J. Natl. Cancer Inst. 97:1180-1184. doi: 97/16/1180 [pii].

Figure legends

Figure 1. Genome-wide *RECQ1*-regulated transcripts are significantly enriched for estrogen-responsive genes.

(A, B) RNA-seq was performed from total RNA isolated from MCF7 cells transfected for 48 hr with CTL, *RECQ1,* or *ESR1* siRNAs. IGV snapshot for the *RECQ1* and *ESR1* loci are shown (A) and the number of genes differentially expressed upon *RECQ1* knockdown is shown in panel B.

(C, D) Volcano plots from the RNA-seq data showing significant differential gene expression upon knockdown of *RECQ1* or *ESR1* in MCF7 cells. Red and green dots correspond to genes up- or down-regulated, respectively. **(E, F)** GSEA analysis was performed from the RNA-seq data upon *RECQ1* or *ESR1* knockdown in MCF7 cells. Data shows significant enrichment of estrogen response early and late genes in genes downregulated upon *RECQ1* (E) or *ESR1* (F) knockdown in MCF7 cells. **(G)** Venn diagram for the comparison of the number of genes downregulated upon knockdown of *RECQ1* or *ESR1* in the RNA-seq from MCF7 cells.

Figure 2. A subset of ERα target genes are downregulated upon *RECQ1* knockdown.

(A) IGV snapshot is shown for select genes downregulated in the RNA-seq data from MCF7 cells upon knockdown of *RECQ1* or *ESR1*. **(B, C)** RT-qPCR was performed for *ESR1*, *RECQ1* and a subset of ERα targets downregulated upon knockdown of *RECQ1* or *ESR1* in MCF7 (B) and T47D cells (C). The housekeeping gene *SDHA* was used as negative control. Error bars represent standard deviation from 3 independent experiments.

Figure 3. Genome-wide identification of RECQ1 binding sites shows significant co-occupancy of RECQ1 and ERα-bound regions.

(A) Genomic distribution of RECQ1 binding sites identified by ChIP-seq (rep #2) in MCF7 cells is shown. Genomic distribution from both replicates is shown in Fig. S2. **(B)** Validation of a subset of RECQ1 targets in MCF7 cells using RECQ1 ChIP-qPCR is shown. B48 primer targets lamin B2 origin of DNA replication and was used as a positive control since RECQ1 is known to bind to the lamin B2 origin in unperturbed cells (18). The B13 primers target a region 5 kb away from the origin and was used a negative control. **(C)** The molecular signature database of GSEA

(https://www.gsea-msigdb.org/gsea/msigdb/annotate.jsp) was used for regions of the genome bound by RECQ1 as identified by ChIP-seq. The graph shows significant enrichment of estrogen response genes together with other pathways. **(D)** Genomic distribution of ERα ChIP-seq (rep #2) is shown. Data from both replicates is shown in Fig. S4. **(E, F)** Comparison of RECQ1 and ERα binding sites genome-wide (E) and in the promoters and gene bodies (F) shows significant overlap. **(G)** Comparison of our ERα ChIP-seq data with 9 previously published ERα ChIP-seq from MCF7 cells shows a strong overlap between our ERα ChIP-seq and previously published ERα ChIP-seq. **(H)** Correlation between our ERα ChIP-seq (ER1 and ER2), RECQ1 ChIP-seq, and previously published ChIP-seq for select transcription factors.

Figure 4. Chromatin features of genome-wide RECQ1 binding.

(A) Signal distribution of RECQ1, ERα, FOXA1, GATA3, H3K27ac, H3K4me1, H3K4me3 and H3K9me3 ChIP-seq data over RECQ1 peaks. The RECQ1 peaks were clustered into three groups based on H3K4me1 and H3K4me3 signal using K-means clustering. Peaks in cluster 1 correspond to strong promoters with high H3K4me3 signal and low H3K4me1 signal. Peaks in cluster 2 correspond to strong enhancers with high H3K4me1 signal and low H3K4me3 signal, with very strong RECQ1, ERα, FOXA1 and GATA3 occupancy. Cluster 3 correspond to weak and/o inactive enhancers and promoters. **(B)** Colocalization of the RECQ1 binding sites (average of 2 replicates) with the RNA PolII (POL2RA) binding sites. POL2RA binding sites were obtained over a time-course of 0, 5, 10 and 20 min of estradiol treatment. The histogram x axis extends 5-kb upstream and 5-kb downstream from the center of the RECQ1 peaks. The extent of colocalization (%) is measured as the fraction of RECQ1 peaks within the 5-kb window of the POLR2A peaks. (see the Materials and Methods section for details). **(C)** Graph depicting

correlations between the binding sites of RECQ1, ERα, PAX8, REST, FOXA1 and GATA3 with the binding sites of POL2RA from cells collected over a time-course of 0 to 320 min after estradiol treatment. Correlations were normalized using the time point t=0. Correlations between peaks were calculated over 5-kb windows. (see the Materials and Methods section for details).

Figure 5. RECQ1 binding to the *ESR1* locus is FOXA1-dependent.

(A) IGV snapshot of RECQ1 and ERα ChIP-seq peaks is shown for the *ESR1* locus. **(B)** Location of primers used to amplify promoter and enhancer regions of *ESR1* is shown. **(C, D)** ChIP-qPCR shows strong enrichment of RECQ1 at specific promoter and enhancer regions of *ESR1* (C). MCF7 cells were transfected with control (siCTL) siRNA or RECQ1 siRNAs (siRECQ1) for 48 hr and ChIP-qPCR was performed using a control IgG antibody or RECQ1 antibody. Data shows that *RECQ1* knockdown in MCF7 cells results in reduced RNA Pol II occupancy at the specific promoter and enhancer regions (D). **(E)** RECQ1 associates with FOXA1 but not ERα protein as assessed by immunoblotting following IP from MCF7 whole-cell extracts of MCF7 cells using a control IgG antibody or RECQ1 antibody. The RECQ-FOXA1 interaction is not DNA- or RNA-dependent because it is not sensitive to Benzonase. **(F)** ChIP-re-ChIP assays show the colocalization of RECQ1, FOXA1, and ERα at *ESR1* regulatory regions. ChIP-reChIP assays in MCF7 cells with RECQ1 as first ChIP, followed by a reChIP with either IgG or antibody against FOXA1 or ERα. ReChIP DNA was quantified by qPCR at *ESR1* enhancer 1 and enhancer 2. **(G)** ChIP-qPCR was performed from MCF7 cells transfected with siCTL or FOXA1 siRNAs (siFOXA1) using a FOXA1 antibody or RECQ1 antibody. Data shows that binding of RECQ1 to a region in the *ESR1* promoter and enhancer is abolished upon knockdown of *FOXA1* with siRNAs in MCF7 cells. As expected, binding of FOXA1 to these regions was lost upon knockdown of

FOXA1 in siRNAs in MCF7 cells. **(H)** Immunoblotting was performed from MCF7 whole-cell lysates prepared after transfection of MCF7 cells with a control (CTL) or FOXA1 siRNAs for 48 hr. The immunonlotting was used to confirm knockdown of FOXA1 in the ChIP-qPCR experiment in panel (F); *FOXA1* knockdown did not affect RECQ1 protein levels. GAPDH was used as loading control.

Figure 6. RECQ1 helicase cooperates with FOXA1 to regulate *ESR1* expression.

(A) FAIRE assays were performed from MCF7 RECQ1-WT or isogenic MCF7 RECQ1-KO cells and the changes in chromatin assembly at an *ESR1* promoter (ESR1_proA) region and an *ESR1* enhancer (*ESR1*_enh1) was assessed by FAIRE-qPCR. The prmoter regions of GAPDH and a hetrochromatin region in Chromosome 12 were used as positive and negative controls for open chromatin, respectively. **(B)** RT-qPCR assays for *ESR1* and the housekeeping gene *SDHA* were performed from MCF7 RECQ1-WT transfected with empty vector or a vector that expresses WT RECQ1 (RECQ1-WT) for 48 hr. In parallel, RT-qPCR assays for *ESR1* and the housekeeping gene *SDHA* were performed from MCF7 RECQ1-KO cells transfected with empty vector or a vector that expresses WT RECQ1 (RECQ1-WT) or helicase-dead RECQ1 mutant (RECQ1-K119R) for 48 hr. Fold change refers to fold change in gene expression normalized to GAPDH. **(C-E)** MCF7 RECQ1-KO cells were transfected with empty vector, FOXA1 expressing vector (FOXA1), or FOXA1 expressing vector in combination with RECQ1-expressing vectors that were RECQ1 WT (RECQ1-WT) or helicase-dead RECQ1 mutant (RECQ1-K119R). Changes in chromatin accessibility at *ESR1* regulatory regions in these cells were determined by FAIRE-qPCR (C). An increase in FAIRE-enrichments by FOXA1 was further enhanced by cotransfection with a helicase active wildtype RECQ1. Relative enrichment of FAIRE signal normalized to input chromatin is

shown for the promoter of *GAPDH* (positive control), a heterochromatin region on Chr. 12 (negative control), and the *ESR1* promoterA and enhancer 1. In these experiments, the effect on FOXA1 and RECQ1 protein was determined by immunoblotting **(D)** using GAPDH as loading control. The effect on *ESR1* or *TFF1* expression was determined by RT-qPCR 48 hr after transfection. *GAPDH* was used as loading control (E). **(E)** Kaplan Meier curves for RECQ1/FOXA1 co-expression and breast cancer-specific survival (BCSS) in ERα^{+} breast cancers.

Figure 7. Proposed model on the role of RECQ1 in regulating the expression of ERα and a subset of ERα target genes.