HHS Public Access

Author manuscript

Methods. Author manuscript; available in PMC 2023 August 01.

Published in final edited form as:

Methods. 2022 August ; 204: 263–268. doi:10.1016/j.ymeth.2022.02.010.

Genome-wide investigations on regulatory functions of RECQ1 helicase

Subrata Debnath[a],

Xing Lu[a],

Ashish Lal[b],

Sudha Sharma[a,c,*]

[a]Department of Biochemistry and Molecular Biology, College of Medicine, Howard University, Washington, DC, United States

[b]Regulatory RNAs and Cancer Section, Genetics Branch, National Cancer Institute, National Institutes of Health, Bethesda, MD, United States

[c]National Human Genome Center, College of Medicine, Howard University, Washington, DC, United States

Abstract

DNA helicase *RECQ1* (also known as *RECQL* or *RECQL1*) is a candidate breast cancer susceptibility gene significantly correlated with clinical outcomes of sporadic breast cancer patients. Prior studies have suggested that RECQ1 maintains genomic stability by regulating a wide variety of core cellular functions including DNA replication, DNA damage response, and transcription. However, it is unclear which, if any, of these are the primary functions of RECQ1 as related to its role in suppressing breast cancer. We describe here an unbiased integrative genomics approach that enabled us to discover a previously unknown regulatory role of RECQ1 in promoting Estrogen Receptor alpha (ERα) expression and the expression of specific ERα target genes in ER positive breast cancer cells. We discuss potential future applications of similar experimental strategies in advancing the mechanistic understanding and elucidating specific new details of genome-wide functions of RECQ1 and other RecQ helicases in maintaining genomic stability and preventing cancer.

Keywords

DNA helicase; estrogen receptor; FOXA1; RECQ1; transcriptional regulation; ChIP-seq; G4; R-loops; RNA-DNA hybrid

[*]**Corresponding Author** sudha.sharma@howard.edu (Sudha Sharma).

CRediT authorship contribution statement

Subrata Debnath: Investigation, Writing- Original draft and review; **Xing Lu:** Investigation, Writing- Original draft and review; **Ashish Lal:** Methodology, Writing- Original draft and review; **Sudha Sharma:** Conceptualization, Methodology, Investigation, Writing- original draft, review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

1. Introduction

First discovered in 1994, the *RECQ1* gene (also known as *RECQL* or *RECQL1*) is localized to chromosome 12p12 and encodes the most abundant of the five known homologs of *E. coli* DNA helicase RecQ in humans [1,2]. As is typical of RecQ helicases [3], RECQ1 catalyzes ATP-dependent unwinding of duplex DNA representing a variety of structures intermediates of replication, repair, and recombination *in vitro* [4–8]. In addition to be a structure-specific helicase, RECQ1 also promotes annealing of complementary single-strand DNA in an ATP-independent manner [4]. Biochemical activities of RECQ1 are aided by various protein partners to facilitate its cellular functions [9–17]. Depletion or loss of RECQ1 leads to increased sister chromatid exchanges, chromosomal instability, DNA damage accumulation in cells, and enhances sensitivity to cytotoxic chemotherapy [9,13,18–21]. Notably, knockdown of *RECQ1* in breast cancer cells has a significant effect on gene expression associated with tumorigenesis [22,23]. Another recent study has also implicated RECQ1 in transcriptional regulation [24].

Since the first demonstration in 2007 that RECQ1 is essential for genomic stability [18,19], research from our lab over the past 15 years has contributed significantly to the current understanding of the biochemical and cellular functions of RECQ1 and provided a strong rationale to investigate *RECQ1* mutations in cancer patients, that ultimately resulted in its discovery in 2015 as a candidate breast cancer susceptibility gene by Cybulski *et al.* [25] and Sun *et al.* [26]. Given that germline mutations in *RECQ1* have been associated with increased risk of breast cancer, it is important to determine the mechanisms by which functional loss of RECQ1 promotes breast cancer susceptibility. Additionally, a recent study has identified *RECQ1* mutations as the underlying cause for a new genomic instability disease named as RECON syndrome [27], underscoring the need to investigate RECQ1 function at the molecular level.

Our previous study had revealed that RECQ1 expression predicts clinical outcomes for sporadic breast cancer patients (METABRIC cohort) stratified by estrogen receptor (ER) status, and that low RECQ1 expression negatively influenced survival in ER positive cohorts that received endocrine therapy [28]. To understand the mechanism(s) involved, and to uncover RECQ1's role in cancer and other diseases, we chose to utilize an unbiased, genome-wide approach. Using this approach, we discovered a novel crosstalk between RECQ1 and the master regulatory transcription factor Estrogen Receptor alpha (ERα) [29]. Here, we provide the methodology we used to investigate the role of RECQ1 in gene regulation (Figure 1) and discuss future applications. More broadly, similar approaches can be used to investigate diverse genomic functions of other RecQ homologs in humans.

2. Description of methods

2.1. Identification of RECQ1-regulated genes

2.1.1 Knockdown of *RECQ1* in breast cancer cell lines.—MCF7 (Michigan Cancer Foundation-7) cells (ER positive breast cancer) were reverse transfected in 6-well plates at a density of 3×10^5 cells per well reaching ~70%–80% confluence one day after transfection. Reverse transfection was performed with Allstars negative control

(*CTL*) siRNA or *RECQ1* siRNAs (SmartPool) using Lipofectamine RNAiMAX Reagent (ThermoFisher) in accordance with the manufacturer's protocol. The final concentration of siRNA in each well was 20 nM. After 48 h of transfection, depletion of RECQ1 and ERα (the protein encoded by *ESR1*) was confirmed by Western blotting from RIPA whole cell lysates. Fifty μg of whole cell lysates were used for immunodetection of RECQ1 and ERα using RECQ1 antibody (Bethyl Laboratories, A300–450A) (1:1000 dilution) and ERα antibody (Santa Cruz Biotechnology, sc542) (1:1000 dilution), respectively. Since we found that ERα protein is significantly reduced in RECQ1-depleted cells, we decided to compare transcriptome-wide changes in *RECQ1* knockdown (KD) with *ESR1* KD in MCF7 cells. To do this, MCF7 cells were similarly transfected with *ESR1* siRNAs (SmartPool). Additionally, 48 h after transfection, total RNA was isolated from each well of a 6-well plate using the RNeasy Plus Micro kit (Qiagen). After determining the quality of the RNA using a Nanodrop spectrophotometer (Thermo Scientific), reverse transcription (RT) followed by quantitative PCR (qPCR) was performed to verify the efficiency of knockdown of *RECQ1* and *ESR1* mRNAs and the expression changes of some genes that were commonly regulated by RECQ1 and ERα 48 h after transfection. For RT, 0.5 μg of total RNA was reverse transcribed using the iScript reverse transcription supermix kit (Bio-Rad) according to the manufacturer's instructions. The resulting cDNA was then subjected to real-time qPCR using iTaq Universal SYBR green Supermix (Bio-Rad). Each qPCR was performed in triplicate, cycled at 95°C for 30 s, followed by 40 cycles of 94°C for 10 s and 60°C for 15 s on the CFX96 real-time PCR system (Bio-Rad). The qPCR data was then normalized to the housekeeping gene *GAPDH,* and differential expression was measured using the $2^{-\Delta\Delta CT}$ method. Nearly 80% knockdown of *RECQ1* mRNA in MCF7 cells was consistently associated with ~50–55% reduction in *ESR1* expression. Specific primers targeting *GAPDH*, *SDHA*, *RECQ1*, *ESR1* and the ERα-transactivated genes *SLC16A1*, *JAK2*, *CAV1* and *TFF1* used in RT-qPCR analysis are available in Lu et al. [29].

2.1.2. RNA isolation.—MCF7 cells were harvested 48 h after transfection with siRNAs as mentioned above and total RNA were isolated using the RNeasy Plus Micro kit (Qiagen). For RNA-seq, 1×10^7 MCF7 cells were reverse transfected with *CTL* siRNA, *ESR1* siRNAs or *RECQ1* siRNAs for 48 h and total RNA was isolated using the RNeasy Plus Micro kit (Qiagen). Knockdown efficiency of *RECQ1* and *ESR1* was verified by RT-qPCR, and samples with at least 70% reduction in *RECQ1* or *ESR1* mRNA levels in *RECQ1* or *ESR1* siRNA transfected cells as compared to *CTL* siRNA transfected cells, respectively, were considered for downstream experiments. RNA integrity was checked with a Bioanalyzer (Agilent), and samples with an RNA integrity number (RIN) of 9.5 were subsequently subjected to mRNA-seq.

2.1.3. mRNA-seq.—To ensure reproducibility, mRNA-seq was performed from biological triplicates. The mRNA-seq samples were pooled and sequenced on HiSeq using the Illumina TruSeq mRNA Prep kit (RS122–2101) and paired-end sequencing was performed. After sequencing, the samples had ~79–101 million pass filter reads with a base call quality of >90% of bases with Q30 and above. Reads 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 was >89%. The mapping statistics were calculated using Picard software. The samples had ~0.88% ribosomal reads. The percent coding base values were between 64 and 66%. The percent untranslated region (UTR) base values were 29 to 31%, and mRNA base values were between 93 and 94% for all samples. Using Picard's MarkDuplicate utility, the library complexity was measured in terms of unique fragments in the mapped reads. The samples had 64 to 70% non-duplicate reads. Read count per gene was calculated by HTSeq (a Python package for analysis of high-throughput sequencing data) using the Gencode annotation database 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 the samples, we used a generalized linear model within DESeq2 that incorporates information from counts and uses negative binomial distribution with a fitted mean and a gene-specific dispersion parameter. DESeq2 used Wald statistics for significance testing and the Benjamini-Hochberg adjustment for multiple corrections.

2.1.4. Pathway analysis.—To identify cellular pathways and biological processes for the genes significantly downregulated upon knockdown of *RECQ1* or *ESR1*, Gene set enrichment analysis (GSEA) was performed for the deregulated genes identified by mRNA-seq from MCF7 cells that were transfected with *CTL* siRNA, *RECQ1* siRNAs or *ESR1* siRNAs (see above). Genes were considered as significantly downregulated if the log2(fold change) was at least 0.5 and $p < 0.05$. GSEA analysis was performed using the GSEA software [30].

2.1.5. RT-qPCR verification.—To validate differential expression of specific genes whose expression changed upon depletion of *RECQ1* or *ESR1*, MCF7 or T47D (another ER positive breast cancer) cells were reverse transfected with *CTL* siRNA, *RECQ1* siRNAs or *ESR1* siRNAs as described above. RT-qPCR was performed from total RNA isolated from the transfected cells as described above. Primer sequences are available in Lu et al. [29].

2.2. Identification of Genome-wide binding sites of RECQ1

2.2.1. RECQ1-ChIP.—To determine whether RECQ1 regulates *ESR1* expression by associating with specific regions on the promoter and/or enhancer of *ESR1*, ChIP (Chromatin Immunoprecipitation) experiments were performed using the ChIP-IT High Sensitivity kit (Active Motif). Ten million MCF7 cells were used for each immunoprecipitation reaction, and duplicate experiments were performed for each antibody. The lysate containing sheared chromatin was incubated overnight at 4°C with 4 μg of anti-RECQ1 (Bethyl Lab, A300–450A) or 4 μg of rabbit IgG as control (Bethyl Lab, P120–101). After the overnight incubation, 30 μl of Protein A Sepharose beads were added to each immunoprecipitation reaction and incubated for 3 h at 4°C. Next, the beads were washed, and DNA bound to the beads was eluted. The washing and elution steps were performed on the ChIP filtration column. To reverse the crosslinks, the eluted ChIP DNA was digested with 5 μl Proteinase K (20 μg/μl) at 55°C for 30 min and then incubated at 80°C for 2 h. The ChIP DNA was then purified using a PCR purification kit according to manufacturer's

instruction and analyzed by qPCR performed in triplicates, with an initial denaturation at 95°C for 3 min, then 40 cycles of denaturation at 95°C for 15 s and annealing and extension at 60°C for 30 s using the CFX96 Real-Time PCR System. The fold enrichment of the target genomic DNA sequences was measured relative to IgG control as fold enrichment = $2^{-(CtIP- CtIgG),}$ where Ct IP and Ct IgG are mean threshold cycles of PCR in triplicates on DNA samples immunoprecipitated with the specific antibody or IgG control, respectively. Melt curve analyses were performed for all qPCR reactions and the presence of a single specific product of expected size was further confirmed by agarose gel electrophoresis. The sequences of primers used for ChIP-qPCR are available in Lu et al [29].

2.2.2. ChIP-seq.—Ten million MCF7 cells were harvested for each ChIP-seq reaction, and duplicate experiments were performed for each antibody. For ChIP-seq, crosslinked chromatin from MCF7 cell lysates was fragmented by sonication using a Bioruptor (Diagnode) at setting 9 with 30 cycles of 30 s on and 60 s off to generate fragments <200 bp long. After reversing the crosslinks as mentioned above, the size of DNA fragments was confirmed by agarose gel electrophoresis. The lysate containing sheared DNA fragments was then divided into two parts. The first part was subjected to ChIP using the anti-RECQ1, anti-ERα (Abcam, ab32063), and IgG antibodies, as described above for ChIP-qPCR. The second part was used as the corresponding input control. After purification of the material bound to the beads using the ChIP-IT High Sensitivity kit (Active Motif), the ChIP 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 ~30–64 million pass filter reads with a base call quality of above ~90% of bases with Q30.

2.2.3. ChIP-seq peak calling: To identify ChIP-seq peaks, 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, q-value cutoff = 1.00e-02 and model fold = [5, 50].

2.2.4. Heatmap and Colocalization analysis: *BAMscale* was used to generate normalized sequencing coverage for the RECQ1 and ERα ChIP-seq data [31]. The heatmap for ChIP-seq peaks colocalization was created using the *rtracklayer* (version1.48.0) [32] and *ComplexHeatmap* [33] (version 2.4.3) packages in R (version 4.0.2, script available at: https://github.com/ncbi/BAMscale) using the RECQ1 called ChIP-seq peaks (n=7,023) with the RECQ1 and ERα ChIP-seq data. ChIP-seq identified thousands of RECQ1 binding events distributed predominantly in intergenic regions and introns, and majority of the RECQ1 peaks colocalized with transcription start sites in MCF7 cells. When we compared the genome-wide binding events of RECQ1 with that of ERα, we found that ~35% of RECQ1 peaks colocalize with ERα peaks. When we compared the RECQ1 peaks with ERα peaks in promoters and gene bodies, we found that ~59% of RECQ1 peaks colocalize with ERα peaks and ~36% of ERα peaks colocalize with RECQ1 peaks. Our study is the first to report genome-wide RECQ1 binding sites via ChIP-seq. When we compared our ERα ChIP-seq data with 9 previously published ERα ChIP-seq from MCF7 cells (GSM1967545,

GSM2257828, GSM986061, GSM1019125, GSM1187116, GSM1643947, GSM1643949, GSM2752596 and GSM2752590), there was ~75–95% colocalization, thus ERα ChIP-seq also served as a positive control for ChIP-seq experiments. We also performed colocalization of RECQ1 ChIP-seq along with FOXA1 (ENCFF255FPM), GATA3 (ENCFF477GZL), H3K27ac (ENCFF411FCW), H3K4me1 (ENCFF983TTS), H3K4me3 (ENCFF862CKA) and H3K9me3 (ENCFF688REP) ChIP-seq data that was downloaded from ENCODE. The ChIP-seq peaks were clustered using k-means clustering, setting the centers to 3, and using the histone markers H3K4me1 (a marker of enhancers) and H3K4me3 (a marker of promoters) for cluster identification. The RECQ1 ChIP-seq peaks were extended to 5 kb upstream or downstream of the peak when creating the heatmaps.

To explore how genome-wide binding of RECQ1 correlates with estrogen response and the binding of transcriptional factors known to be involved in estrogen signaling, colocalization analyses were performed between the binding sites of POL2RA (RNA polymerase II) from cells collected over a time-course of 0 to 320 min after estradiol treatment [34] and ChIP-seq data for RECQ1 (two biological replicates), ERα (two biological replicates, this study), PAX8, REST, FOXA1 and GATA3. Publicly available ChIP-seq bed files were obtained from Cistrome (http://cistrome.org) [35] 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). We used the genome inspector program of the genomatix genome analyzer 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). The average of colocalization values for each anchor sets were normalized using the time point t=0 for estradiol treatment. The 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 the genomic sequence for which the experimental evidence for transcription initiation was available. The percent colocalization was determined as the fraction of RECQ1 ChIP-seq peaks within the 5 kb of the TSSs. 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.

3. Discussion

RECQ1 is a candidate breast cancer susceptibility gene significantly correlated with clinical outcomes of sporadic breast cancer patients. However, the mechanisms by which RECQ1 helicase acts to suppress breast cancer development, progression and therapeutic response are poorly understood. An unbiased integrative genomics approach allowed us to discover RECQ1 as a co-factor of ERα, the major driving transcription factor in the mammary gland development as well as a master regulator of tumor development and progression and the primary therapeutic target in luminal breast cancer which accounts for ~70% of

all breast cancer cases [36]. We determined the genomic landscape of endogenous RECQ1 binding and compared it with the ERα ChIP-seq to show that one third of RECQ1-binding sites are co-bound with ERα genome-wide in MCF7 cells. Transcriptome profiling by RNA-seq revealed estrogen response as the most significantly enriched pathway in RECQ1 KD MCF7 cells. Integrating the results from these unbiased experiments as well as taking advantage of publicly available ChIP-seq and RNA-seq datasets, followed by functional analysis including cell-based assays, we found that RECQ1 is an important determinant of ERα expression and cooperates with ERα by binding to the promoters and/or enhancers of a subset of ERα cistrome, thereby promoting the expression of specific ERα target genes in breast cancer cells [29]. We are currently investigating how estrogen treatment modulates RECQ1 landscape in ERα positive cells, and the subset of genes whose expression is commonly regulated by ERα and RECQ1 binding under basal condition and upon estrogen stimulation. Experiments using CRISPR/Cas9 edited isogenic MCF7 cell lines reconstituted with wild-type or helicase deficient breast cancer risk-associated RECQ1 point mutants will examine if the ERα mediated transcriptional regulation and signaling is dependent on RECQ1 helicase activity. Overall, this initial elucidation of the RECQ1 as a new component of the regulatory machinery underlying ERα transcriptional program provides mechanistic insight to RECQ1's role in breast cancer. A similar experimental approach can be used to identify additional factors that cooperate with RECQ1 for regulating disease-driving gene expression changes in breast cancer (Figure 2a).

Although Figure 2 depicts putative roles of RECQ1, the described approach could be broadly applicable to investigate the molecular function(s) of other human RecQ homologs [3]. In doing so, it is important to consider which cell lines to perform ChIP-seq and RNA-seq experiments depending upon the goal of the study and expression levels of the protein of interest. We recommend careful evaluation of available antibodies to determine the specificity and immunoprecipitation efficiency for a successful ChIP-seq. Specificity of siRNAs for target gene knockdown is a critical factor in RNA-seq analyses to minimize off-target effects. When available, an isogenic knockout cell line can serve as an excellent negative control for ChIP-seq, and for transcriptome profiling by RNA-seq. Bioinformatic analysis of genome-wide binding sites reveals enriched pathways with statistical significance. Functional relevance of genomic binding can be assessed by the expression of target gene(s) in knockdown cells. Combining the ChIP-seq data with RNA-Seq in control and knockdown or isogenic knockout cells can identify functionally and biologically relevant genomic targets RECQ1 and related RecQ helicases. Overlap of statistically significant enriched pathways in these two independent and unbiased analyses will suggest biological relevance that should be corroborated by functional assays.

Genome-wide chromatin accessibility is important for global epigenetic control of gene expression [37]. Indeed, MCF7 cells lacking RECQ1 have reduced chromatin accessibility at the *ESR1* regulatory regions indicating a role in facilitating local chromatin states at specific genomic loci [29]. Through its robust DNA-binding/unwinding activities, RECQ1 may facilitate, at a subset of genomic sites, genome accessibility of transcriptional machinery and transcription factors to control physiologically relevant gene expression programs (Figure 2b). In our initial study, using a locus-centric approach by Formaldehyde-assisted isolation of regulatory elements (FAIRE)-qPCR, we reported that RECQ1 helicase cooperates

with FOXA1, the known pioneer factor of ERα, to enhance chromatin accessibility and promote *ESR1* expression [29]. Profiling the genomic targets of RECQ1 by ChIP-seq and integrating that with chromatin accessibility data from high throughput sequencing methods [35,38] such as assay for transposase-accessible chromatin with high-throughput sequencing (ATAC-seq) [39], DNase I hypersensitive site sequencing (DNase-seq) [40], or FAIRE-seq [41] in RECQ1 proficient and deficient cells can provide important new information about the chromatin features (e.g. histone marks dictating chromatin state) of RECQ1 binding sites genome-wide, and lead to further identification of factors that provide specificity to RECQ1 to be recruited at the specific subsets of genomic sites.

Gene regulatory function of RECQ1 may also be attributed to its ability to recognize specific DNA structures genome wide. We had reported that the promoters of genes downregulated upon RECQ1 silencing in triple negative MDA-MB-231 breast cancer cells were significantly enriched for G-quadruplex (G4) motifs with potential to form G4 structures [22]. However, the functional significance of RECQ1 binding to G4 motifs *in vivo* [22] and G4 DNA structures *in vitro* [5] has been unclear until very recently. Liu *et al.* have now demonstrated that the endogenous RECQ1 can unfold intramolecular G4 DNA [42]. Thus, RECQ1 may modulate gene expression by binding to the G4 motifs and regulating the *in vivo* stability of G4 structures in the regulatory chromatin (Figure 2c, left panel). ChIP-seq of RECQ1 and the G4-specific antibody BG4 [43,44], in the presence and absence of G4 stabilizers [45], and integrating that with the RNA-seq in RECQ1 proficient and deficient cells should determine RECQ1 associated G4 structures genome-wide and how that affects transcriptomic changes.

Transcriptional regulation by G4 DNA depends on the composition of G4 motif sequence, its location with respect to transcription start sites, and whether it is present in the template or non-template strand [46]. Intramolecular G4 DNA on the non-template strand could interfere with transcription by impeding the reannealing of two DNA strands behind the RNA polymerase complex favoring the formation and stabilization of RNA:DNA hybrids, the so called R-loops [45,47,48]. Given that recombinant RECQ1 can bind DNA-RNA hybrids *in vitro* and the availability of a variety of assays to identify endogenous R-loops [49–52], it should be possible to use genome-wide approaches to determine whether endogenous RECQ1 has a role at R-loops mediated transcriptional outcomes (Figure 2c, right panel).

Biological functions of RECQ1 helicase are expected to be dictated by interacting protein partners, post-translational modifications, and local chromatin features to facilitate specific signaling and regulatory pathways for genome homeostasis. Mechanistic investigation of this is likely to have therapeutic implications. For example, RECQ1 as a key cofactor that promotes ERα and its target gene expression could have significance as a potential molecular target in endocrine therapy for breast cancer. With the availability of high throughput techniques to profile DNA-protein interactions and integrating those with epigenomic and transcriptomic changes genome-wide, an important next step in understanding the molecular role of RECQ1 in breast cancer is to gain a more complete understanding of the mechanisms of genome-wide gene regulation by RECQ1 in breast cancer cells, including the involvement of cooperating transcriptional partners.

Understanding the genomic elements of RECQ1 binding sites (e.g. potential secondary structures, histone marks dictating chromatin state) may also elucidate the genomic instability in RECQ1-deficiency. Unbiased integrative functional genomics approaches as described here can reveal underlying mechanisms by which RECQ1 performs its critical functions in genome maintenance and how the genetic or functional loss of RECQ1 results in cancer predisposition.

Acknowledgements

This work was funded 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 to Sudha Sharma and the Intramural Research Program of the National Institutes of Health, National Cancer Institute, and Center for Cancer Research to Ashish Lal. We also acknowledge support from the National Institute on Aging under Award Number 1R25 AG047843 and infrastructure support from the National Institute of Minority Health Disparities under award number G12MD007597.

References

[1]. KL Puranam, Blackshear PJ. Cloning and characterization of RECQL, a potential human homologue of the Escherichia coli DNA helicase RecQ, J.Biol.Chem. 269 (1994) 29838–29845. [PubMed: 7961977]

[2]. Puranam KL, Kennington E, Sait SN, Shows TB, Rochelle JM, Seldin MF, et al. Chromosomal localization of the gene encoding the human DNA helicase RECQL and its mouse homologue, Genomics. 26 (1995) 595–598. [PubMed: 7607686]

[3]. Croteau DL, Popuri V, Opresko PL, Bohr VA. Human RecQ helicases in DNA repair, recombination, and replication, Annu.Rev.Biochem. 83 (2014) 519–552. [PubMed: 24606147]

[4]. Sharma S, Sommers JA, Choudhary S, Faulkner JK, Cui S, Andreoli L, et al. Biochemical analysis of the DNA unwinding and strand annealing activities catalyzed by human RECQ1, J.Biol.Chem. 280 (2005) 28072–28084. [PubMed: 15899892]

[5]. Popuri V, Bachrati CZ, Muzzolini L, Mosedale G, Costantini S, Giacomini E, et al. The Human RecQ helicases, BLM and RECQ1, display distinct DNA substrate specificities, J.Biol.Chem. 283 (2008) 17766–17776. [PubMed: 18448429]

[6]. Popuri V, Croteau DL, Brosh RM Jr, Bohr VA. RECQ1 is required for cellular resistance to replication stress and catalyzes strand exchange on stalled replication fork structures, Cell.Cycle. 11 (2012) 4252–4265. [PubMed: 23095637]

[7]. Popuri V, Hsu J, Khadka P, Horvath K, Liu Y, Croteau DL, et al. Human RECQL1 participates in telomere maintenance, Nucleic Acids Res. 42 (2014) 5671–5688. [PubMed: 24623817]

[8]. Bugreev DV, Brosh RM Jr, Mazin AV. RECQ1 possesses DNA branch migration activity, J.Biol.Chem. 283 (2008) 20231–20242. [PubMed: 18495662]

[9]. Debnath S, Sharma S. RECQ1 Helicase in Genomic Stability and Cancer, Genes (Basel). 11 (2020) E622. doi: 10.3390/genes11060622. [PubMed: 32517021]

[10]. Cui S, Klima R, Ochem A, Arosio D, Falaschi A, Vindigni A. Characterization of the DNA-unwinding activity of human RECQ1, a helicase specifically stimulated by human replication protein A, J.Biol.Chem. 278 (2003) 1424–1432. [PubMed: 12419808]

[11]. Doherty KM, Sharma S, Uzdilla LA, Wilson TM, Cui S, Vindigni A, et al. RECQ1 helicase interacts with human mismatch repair factors that regulate genetic recombination, J.Biol.Chem. 280 (2005) 28085–28094. [PubMed: 15886194]

[12]. Sharma S, Phatak P, Stortchevoi A, Jasin M, Larocque JR. RECQ1 plays a distinct role in cellular response to oxidative DNA damage, DNA Repair (Amst). 11 (2012) 537–549. [PubMed: 22542292]

[13]. Berti M, Chaudhuri A Ray, Thangavel S, Gomathinayagam S, Kenig S, Vujanovic M, et al. Human RECQ1 promotes restart of replication forks reversed by DNA topoisomerase I inhibition, Nat.Struct.Mol.Biol. 20 (2013) 347–354. [PubMed: 23396353]

[14]. Parvathaneni S, Stortchevoi A, Sommers JA, Brosh RM Jr, Sharma S. Human RECQ1 interacts with Ku70/80 and modulates DNA end-joining of double-strand breaks, PLoS One. 8 (2013) e62481. [PubMed: 23650516]

[15]. Sami F, Lu X, Parvathaneni S, Roy R, Gary RK, Sharma S. RECQ1 interacts with FEN-1 and promotes binding of FEN-1 to telomeric chromatin, Biochem.J. 468 (2015) 227–244. [PubMed: 25774876]

[16]. Banerjee T, Sommers JA, Huang J, Seidman MM, Brosh RM Jr. Catalytic Strand Separation by RECQ1 Is Required for RPA-Mediated Response to Replication Stress, Curr.Biol. 25 (2015) 2830–2838. [PubMed: 26455304]

[17]. Woodrick J, Gupta S, Camacho S, Parvathaneni S, Choudhury S, Cheema A, et al. A new sub-pathway of long-patch base excision repair involving 5' gap formation, EMBO J. 36 (2017) 1605–1622. [PubMed: 28373211]

[18]. Sharma S, Brosh RM Jr. Human RECQ1 is a DNA damage responsive protein required for genotoxic stress resistance and suppression of sister chromatid exchanges, PLoS One. 2 (2007) e1297. [PubMed: 18074021]

[19]. Sharma S, Stumpo DJ, Balajee AS, Bock CB, Lansdorp PM, Brosh RM Jr, et al. RECQL, a member of the RecQ family of DNA helicases, suppresses chromosomal instability, Mol.Cell.Biol. 27 (2007) 1784–1794. [PubMed: 17158923]

[20]. Viziteu E, Klein B, Basbous J, Lin YL, Hirtz C, Gourzones C, et al. RECQ1 helicase is involved in replication stress survival and drug resistance in multiple myeloma, Leukemia. 31 (2017) 2104–2113. [PubMed: 28186131]

[21]. Parvathaneni S, Sharma S. The DNA repair helicase RECQ1 has a checkpoint-dependent role in mediating DNA damage responses induced by gemcitabine, J.Biol.Chem. 294 (2019) 15330–15345. [PubMed: 31444271]

[22]. Li XL, Lu X, Parvathaneni S, Bilke S, Zhang H, Thangavel S, et al. Identification of RECQ1-regulated transcriptome uncovers a role of RECQ1 in regulation of cancer cell migration and invasion, Cell.Cycle. 13 (2014) 2431–2445. [PubMed: 25483193]

[23]. Lu X, Parvathaneni S, Li XL, Lal A, Sharma S. Transcriptome guided identification of novel functions of RECQ1 helicase, Methods. 108 (2016) 111–117. [PubMed: 27102625]

[24]. Bakr A, Hey J, Sigismondo G, Liu CS, Sadik A, Goyal A, et al. ID3 promotes homologous recombination via non-transcriptional and transcriptional mechanisms and its loss confers sensitivity to PARP inhibition, Nucleic Acids Res. 49 (2021) 11666–11689. [PubMed: 34718742]

[25]. Cybulski C, Carrot-Zhang J, Kluzniak W, Rivera B, Kashyap A, Wokolorczyk D, et al. Germline RECQL mutations are associated with breast cancer susceptibility, Nat.Genet. 47 (2015) 643–646. [PubMed: 25915596]

[26]. Sun J, Wang Y, Xia Y, Xu Y, Ouyang T, Li J, et al. Mutations in RECQL Gene Are Associated with Predisposition to Breast Cancer, PLoS Genet. 11 (2015) e1005228. [PubMed: 25945795]

[27]. Abu-Libdeh B, Jhujh SS, Dhar S, Sommers JA, Datta A, Longo GM, et al. RECON syndrome is a genome instability disorder caused by mutations in the DNA helicase RECQL1, J.Clin.Invest. (2022).

[28]. Arora A, Parvathaneni S, Aleskandarany MA, Agarwal D, Ali R, Abdel-Fatah T, et al. Clinicopathological and Functional Significance of RECQL1 Helicase in Sporadic Breast Cancers, Mol.Cancer.Ther. 16 (2017) 239–250. [PubMed: 27837030]

[29]. Lu X, Redon CE, Tang W, Parvathaneni S, Bokhari B, Debnath S, et al. Genome-Wide Analysis Unveils DNA Helicase RECQ1 as a Regulator of Estrogen Response Pathway in Breast Cancer Cells, Mol.Cell.Biol. 41 (2021) e00515–20. doi: 10.1128/MCB.00515-20. Print 2021 Mar 24. [PubMed: 33468559]

[30]. Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J, et al. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes, Nat.Genet. 34 (2003) 267–273. [PubMed: 12808457]

[31]. Pongor LS, Gross JM, Vera Alvarez R, Murai J, Jang SM, Zhang H, et al. BAMscale: quantification of next-generation sequencing peaks and generation of scaled coverage tracks, Epigenetics Chromatin. 13 (2020) 21–020-00343-x.

[32]. Lawrence M, Gentleman R, Carey V. rtracklayer: an R package for interfacing with genome browsers, Bioinformatics. 25 (2009) 1841–1842. [PubMed: 19468054]

[33]. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data, Bioinformatics. 32 (2016) 2847–2849. [PubMed: 27207943]

[34]. wa Maina C, Honkela A, Matarese F, Grote K, Stunnenberg HG, Reid G, et al. Inference of RNA polymerase II transcription dynamics from chromatin immunoprecipitation time course data, PLoS Comput.Biol. 10 (2014) e1003598.

[35]. Zheng R, Wan C, Mei S, Qin Q, Wu Q, Sun H, et al. Cistrome Data Browser: expanded datasets and new tools for gene regulatory analysis, Nucleic Acids Res. 47 (2019) D729–D735. [PubMed: 30462313]

[36]. Chi D, Singhal H, Li L, Xiao T, Liu W, Pun M, et al. Estrogen receptor signaling is reprogrammed during breast tumorigenesis, Proc.Natl.Acad.Sci.U.S.A. 116 (2019) 11437–11443. [PubMed: 31110002]

[37]. Klemm SL, Shipony Z, Greenleaf WJ. Chromatin accessibility and the regulatory epigenome, Nat.Rev.Genet. 20 (2019) 207–220. [PubMed: 30675018]

[38]. Klein DC, Hainer SJ. Genomic methods in profiling DNA accessibility and factor localization, Chromosome Res. 28 (2020) 69–85. [PubMed: 31776829]

[39]. Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide, Curr.Protoc.Mol.Biol. 109 (2015) 21.29.1–21.29.9.

[40]. Song L, Crawford GE. DNase-seq: a high-resolution technique for mapping active gene regulatory elements across the genome from mammalian cells, Cold Spring Harb Protoc. 2010 (2010) pdb.prot5384.

[41]. Bianco S, Rodrigue S, Murphy BD, Gévry N. Global Mapping of Open Chromatin Regulatory Elements by Formaldehyde-Assisted Isolation of Regulatory Elements Followed by Sequencing (FAIRE-seq), Methods Mol.Biol. 1334 (2015) 261–272. [PubMed: 26404156]

[42]. Liu NN, Song ZY, Guo HL, Yin H, Chen WF, Dai YX, et al. Endogenous Bos taurus RECQL is predominantly monomeric and more active than oligomers, Cell.Rep. 36 (2021) 109688.

[43]. Biffi G, Tannahill D, McCafferty J, Balasubramanian S. Quantitative visualization of DNA G-quadruplex structures in human cells, Nat.Chem. 5 (2013) 182–186. [PubMed: 23422559]

[44]. Liano D, Chowdhury S, Di Antonio M. Cockayne Syndrome B Protein Selectively Resolves and Interact with Intermolecular DNA G-Quadruplex Structures, J.Am.Chem.Soc. 143 (2021) 20988–21002. [PubMed: 34855372]

[45]. De Magis A, Manzo SG, Russo M, Marinello J, Morigi R, Sordet O, et al. DNA damage and genome instability by G-quadruplex ligands are mediated by R loops in human cancer cells, Proc.Natl.Acad.Sci.U.S.A. 116 (2019) 816–825. [PubMed: 30591567]

[46]. Kim N. The Interplay between G-quadruplex and Transcription, Curr.Med.Chem. 26 (2019) 2898–2917. [PubMed: 29284393]

[47]. Belotserkovskii BP, Soo Shin JH, Hanawalt PC. Strong transcription blockage mediated by R-loop formation within a G-rich homopurine-homopyrimidine sequence localized in the vicinity of the promoter, Nucleic Acids Res. 45 (2017) 6589–6599. [PubMed: 28498974]

[48]. Miglietta G, Russo M, Capranico G. G-quadruplex-R-loop interactions and the mechanism of anticancer G-quadruplex binders, Nucleic Acids Res. 48 (2020) 11942–11957. [PubMed: 33137181]

[49]. Brázda V, Havlík J, Kolomazník J, Trenz O, Šťastný J. R-Loop Tracker: Web Access-Based Tool for R-Loop Detection and Analysis in Genomic DNA Sequences, Int.J.Mol.Sci. 22 (2021) 12857. doi: 10.3390/ijms222312857. [PubMed: 34884661]

[50]. Sanz LA, Castillo-Guzman D, Chédin F. Mapping R-Loops and RNA:DNA Hybrids with S9.6-Based Immunoprecipitation Methods, J.Vis.Exp. (174). doi (2021) 10.3791/62455.

[51]. Ramirez P, Crouch RJ, Cheung VG, Grunseich C. R-Loop Analysis by Dot-Blot, J.Vis.Exp. (167):10.3791/62069. doi (2021) 10.3791/62069.

[52]. Rassoulzadegan M, Sharifi-Zarchi A, Kianmehr L. DNA-RNA Hybrid (R-Loop): From a Unified Picture of the Mammalian Telomere to the Genome-Wide Profile, Cells. 10 (2021) 1556. doi: 10.3390/cells10061556. [PubMed: 34205454]

Highlights

- Experimental design to investigate genome-wide binding of endogenous RECQ1
- Experimental design to investigate transcriptome-wide changes in RECQ1 deficiency
- Integrative genomics approach to identify potential mechanisms of transcriptional regulation by RECQ1 helicase

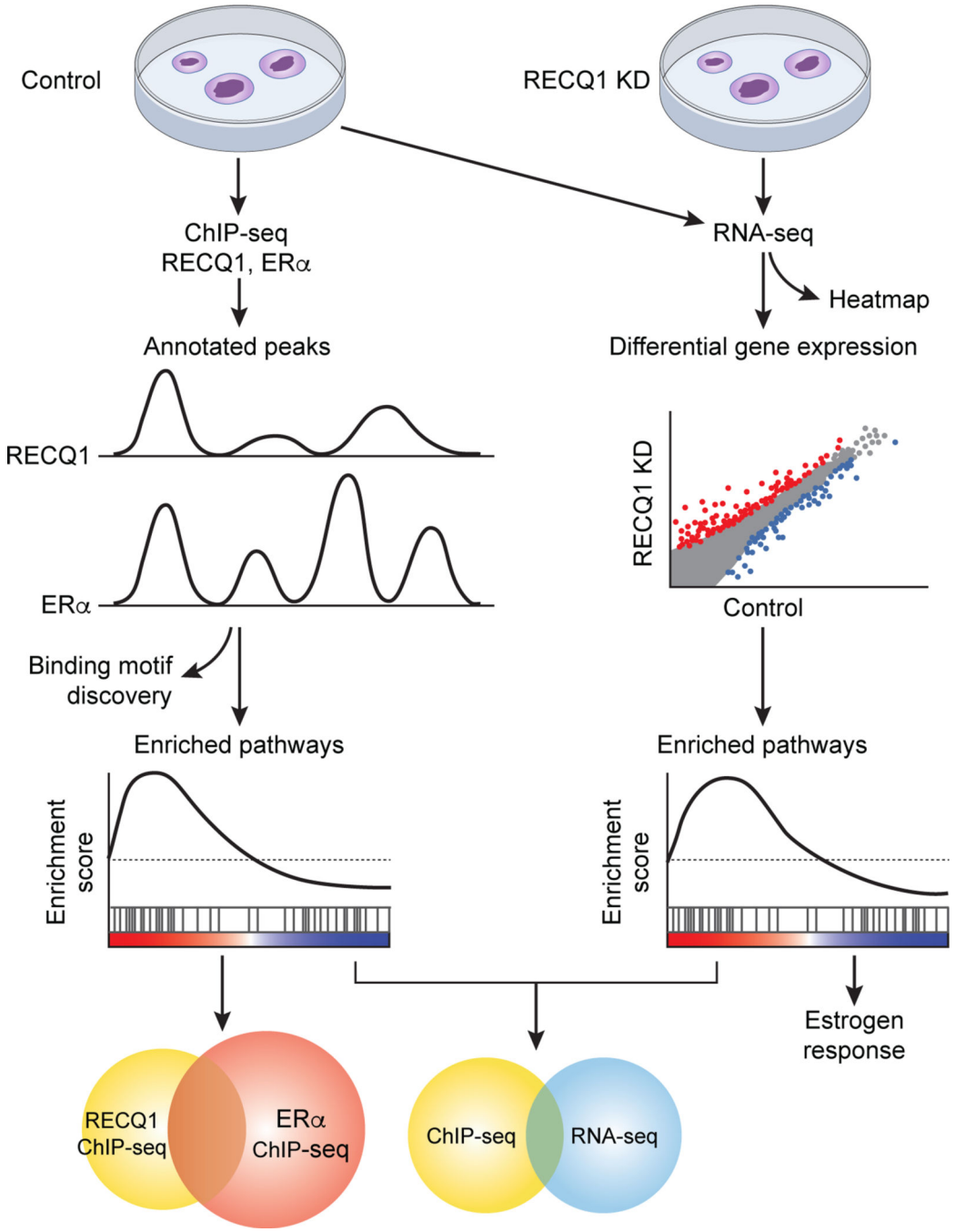

Figure 1. Schematic of experimental approach utilized to identify RECQ1-ERα crosstalk.

Comparing the results from RECQ1 and ERα ChIP-seq revealed significant co-occupancy of these proteins in MCF7 cells genome-wide. Transcriptomic analyses in control versus RECQ1 KD cells revealed estrogen response as the major RECQ1-regulated pathway in MCF7 cells. Integrating the results of RECQ1 ChIP-seq and RNA-seq identified a subset of RECQ1-regulated genes in MCF7 cells. We utilized siRNA knockdown for transcriptome analysis in our original study but the functions of RECQ1 (or other RecQ protein) can

also be investigated following shRNA-mediated knockdown or CRISPR-Cas9-mediated knockout in cell lines.

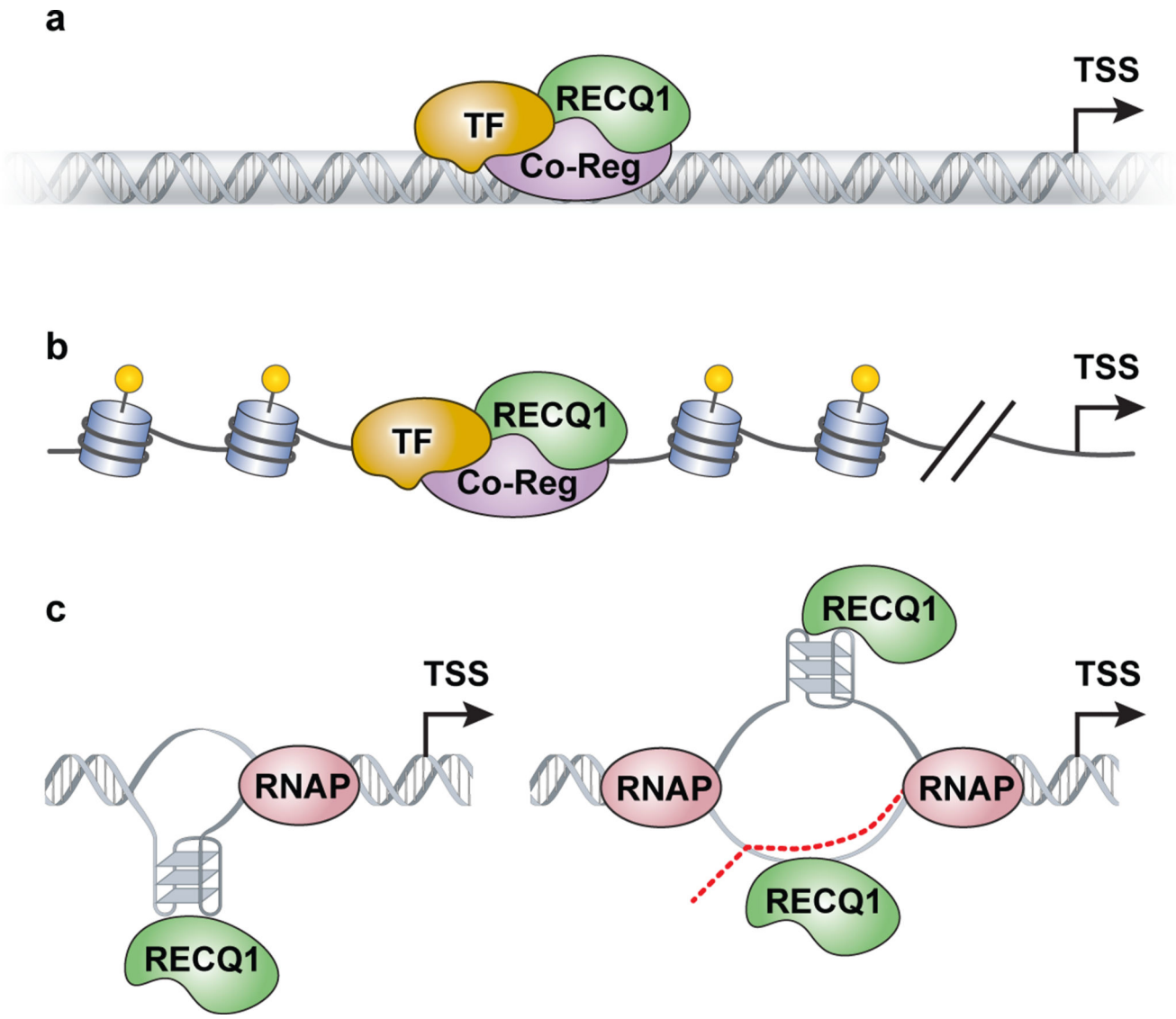

Figure 2. Suggested applications of genome-wide approaches to investigate mechanisms of transcriptional regulation by RECQ1.

Depicted here are the putative roles of RECQ1 but this could be broadly applicable to all RecQ helicase homologs in humans. (a) Analogous to its crosstalk with ERα and FOXA1, RECQ1 may cooperate with other transcription factors (TF) and co-regulators (Co-reg) for regulating the expression of a subset of their target genes, (b) Catalytic activities of RECQ1 may facilitate chromatin accessibility of transcription factors and promote gene expression, and (c) Structure-specific DNA binding of RECQ1 and its ability to unwind G4 structures (left) and/or R-loop (right) may alter transcriptional outcomes. RNA polymerase II is denoted as RNAP. TSS denotes transcriptional start site.