



Published in final edited form as:

*Mol Cell*. 2015 February 19; 57(4): 636–647. doi:10.1016/j.molcel.2015.01.011.

## BRCA1 recruitment to transcriptional pause sites is required for R-loop-driven DNA damage repair

Elodie Hatchi<sup>1,2,\*</sup>, Konstantina Skourti-Stathaki<sup>3</sup>, Steffen Ventz<sup>4,5</sup>, Luca Pinello<sup>4,5</sup>, Angela Yen<sup>6,7</sup>, Kinga Kamieniarz-Gdula<sup>3</sup>, Stoil Dimitrov<sup>1,2</sup>, Shailja Pathania<sup>1,2</sup>, Kristine McKinney<sup>1,2</sup>, Matthew L. Eaton<sup>6,7</sup>, Manolis Kellis<sup>6,7</sup>, Sarah J. Hill<sup>1,2</sup>, Giovanni Parmigiani<sup>4,5</sup>, Nicholas J. Proudfoot<sup>3</sup>, and David M. Livingston<sup>1,2,\*</sup>

<sup>1</sup>Department of Genetics, Harvard Medical School, Boston, MA 02215, USA

<sup>2</sup>Department of Cancer Biology, Dana-Farber Cancer Institute, 450 Brookline Avenue, Boston, MA 02215, USA

<sup>3</sup>Sir William Dunn School of Pathology, University of Oxford, Oxford, OX1 3RE, UK

<sup>4</sup>Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, 450 Brookline Avenue, Boston, MA 02215, USA

<sup>5</sup>Department of Biostatistics, Harvard School of Public Health

<sup>6</sup>Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA

<sup>7</sup>Computer Science and Artificial Intelligence Laboratory (CSAIL), MIT, Cambridge, MA 02139, USA

### SUMMARY

The mechanisms contributing to transcription-associated genomic instability are both complex and incompletely understood. Although R-loops are normal transcriptional intermediates, they are also associated with genomic instability. Here, we show that BRCA1 is recruited to R-loops that form normally over a subset of transcription termination regions. There it mediates the recruitment of a specific, physiological binding partner, senataxin (SETX). Disruption of this complex led to R-loop-driven DNA damage at those loci as reflected by adjacent  $\gamma$ -H2AX accumulation and ssDNA breaks within the untranscribed strand of relevant R-loop structures. Genome-wide analysis revealed widespread BRCA1 binding enrichment at R-loop-rich termination regions (TR) of

© 2015 Published by Elsevier Inc.

\*Correspondence: david\_livingston@dfci.harvard.edu (D.M.L), elodie\_hatchi@dfci.harvard.edu (E.H).

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

#### Author contributions

S.J.H. and D.M.L. originally suggested and first conceived the rationale for this project. E.H. designed the overall project with help from K.S.S., K.K.M., D.M.L. and N.J.P. E.H. performed the experiments and analyzed the experimental data unless otherwise stated. E.H. received advice, reagents and help for the ChIP experiments and the coimmunoprecipitation from K.S.S., K.K.M., S.P. and S.D. K.S.S. performed and analyzed the BRCA1 and SETX DIP and part of BRCA1 and SETX ChIP experiments. A.Y., M.L.E. and M.K. performed the genomic analysis of BRCA1 ChIP-seq data and K.K.G. the meta-analyses correlating RNAPII and DRIP data sets. S.V., L.P. and G.P. performed the global mutational analyses of the breast tumors. E.H. and D.M.L. wrote the manuscript. All authors contributed to the editing of the manuscript.

actively transcribed genes. Strikingly, within some of these genes in BRCA1-null breast tumors, there are specific insertion/deletion mutations located close to R-loop-mediated BRCA1 binding sites within TRs. Thus, BRCA1/SETX complexes support a DNA repair mechanism that addresses R-loop-based DNA damage at transcriptional pause sites.

---

## INTRODUCTION

Germline mutations of the breast cancer susceptibility gene 1 (*BRCA1*) significantly increase the risk of developing breast and ovarian cancers (Tutt and Ashworth, 2002). Since its discovery, researchers have intensively investigated the mechanisms by which full-length BRCA1 (p220) functions as a tumor suppressor.

BRCA1 is engaged in numerous direct and indirect physical interactions with specific partner proteins (Huen et al., 2010). As part of several complexes, BRCA1 contributes to double strand DNA break (DSB) repair by homologous recombination (HR), stalled fork repair, cell cycle checkpoint activation, transcription regulation, heterochromatin maintenance, mitotic spindle formation, RNA splicing control, and estrogen metabolism (Gardini et al., 2014; Gorski et al., 2011; Mullan et al., 2006; Pathania et al., 2011; Savage et al., 2014a; 2014b; Venkitaraman, 2014; Zhu et al., 2011). Many of these BRCA1 properties, and in particular, those that protect genome integrity, likely contribute to its tumor suppressing function (Silver and Livingston, 2012; Tutt and Ashworth, 2002). However, in the absence of a coherent underlying mechanism, there is still no definitive evidence of which specific BRCA1 functions are required for breast and ovarian cancer suppression (Huen et al., 2010; Venkitaraman, 2014). Therefore, identifying new BRCA1 binding partners and their associated functions may yield valuable insights.

Our laboratory has identified senataxin (SETX) as a putative BRCA1 binding partner in a yeast two hybrid and several, independent TAP-MS based genome wide BRCA1/protein interaction screens ((Hill et al., 2014) and unpublished data). Disruption of the *Setx* gene in mice leads to a defect in spermatogenesis, caused by failure of meiotic recombination (Becherel et al., 2013). *Sen1*, the *SETX* yeast homolog, was shown to contribute to the processing of various RNA species and to the distribution of RNA polymerase II (RNAPII) across the genome (Mischo et al., 2011; Steinmetz et al., 2006; Ursic et al., 1997). This likely occurs via its direct interaction with RNAPII and certain RNA processing factors (Suraweera et al., 2009).

While transcription is an essential cellular process, it also represents a potential threat to genome integrity (Kim and Jinks-Robertson, 2012). Several studies indicate that highly transcribed genes exhibit increased rates of mutation and illegitimate recombination (Gaillard et al., 2013). Moreover, a large body of evidence indicates that mutations in certain factors involved at the interface of transcription and RNA processing are associated with genomic instability (Bhatia et al., 2014; Chan et al., 2014; Kleiman and Manley, 1999; Kleiman et al., 2005; Li and Manley, 2006; Stirling et al., 2012). An emerging view is that these mutants contribute to the above-noted phenomena through a common mechanism, which induces the abnormal persistence of co-transcriptional R-loops (three-stranded structures, each consisting of a RNA:DNA hybrid plus the coding strand DNA). Although

R-loops are a naturally occurring consequence of transcription and are essential for diverse cellular events (Skourti-Stathaki and Proudfoot, 2014), they can be potentially deleterious to some cellular functions and compromise genome integrity (Aguilera and García-Muse, 2012; Hamperl and Cimprich, 2014). Indeed, unresolved R-loop structures can expose the displaced, coding ssDNA to nicking and/or other forms of damage (Daniel and Nussenzweig, 2013; Wimberly et al., 2013), as well as impair transcription (Aguilera, 2002; Huertas and Aguilera, 2003) and DNA replication fork progression (Gan et al., 2011; Helmrich et al., 2011).

Interestingly, *SETX* is involved in RNAPII transcription termination and resolves R-loops that form at G-rich transcription pause sites (Skourti-Stathaki et al., 2011). It also associates with processing replication forks and facilitates their progression through RNAPII transcribed genes by displacing R-loops (Alzu et al., 2012). In part through its genetic interaction with DNA repair genes involved in HR, senataxin also protects the genome from transcription-associated instability (Mischo et al., 2011; Ursic et al., 2004). Similarly, *SETX*, by resolving R-loops at sites of transcription and replication collision, is engaged at the interface of replication stress, transcription, and DNA damage (Yüce and West, 2013).

Interestingly, BRCA1-containing complexes restrict DNA damage induced by aberrant transcription or RNA processing via proposed interactions with multiple transcription and RNA processing factors, including RNAPII (Anderson et al., 1998; Bennett et al., 2008; Kawai and Amano, 2012; Kleiman and Manley, 1999; Kleiman et al., 2005; Savage et al., 2014a; Scully et al., 1997).

In view of these associations, we have asked whether BRCA1 plays a significant role in the repair of R-loop associated DNA damage arising at termination sites. We find that BRCA1 and *SETX* form a physiological complex, recruited in a BRCA1-dependent manner to a subset of transcription termination pause sites of highly transcribed genes. There they act to suppress co-transcriptional R-loop-associated DNA damage. Unexpectedly, in breast tumor tissues carrying inherited BRCA1 mutations, insertion/deletion somatic mutations were found in the vicinity of BRCA1-bound gene termination sites where BRCA1 normally engages in the repair of R-loop-associated DNA damage.

## RESULTS

### Identification of BRCA1 as a scaffolding partner for SETX at the $\beta$ -actin transcription termination site

To investigate whether BRCA1 is involved in R-loop-driven DNA damage responses, we first assessed the physiological relevance of a BRCA1 and *SETX* interaction, recently identified in our proteomic screens (Hill et al., 2014) and suggested by others (Becherel et al., 2013). Endogenous BRCA1 and *SETX* co-immunoprecipitation (co-IP) was carried out in HeLa cell extracts, using two, different mono-specific antibodies (Abs) (Figures S1A and S1B). IP of each protein revealed significant co-IP of the other. Irrelevant IgG gave negative results. These results imply the existence of endogenous BRCA1/*SETX* containing complexes in these cells (Figure 1A). Weak or undetectable co-IP followed RNAi-driven BRCA1- and *SETX*-depletion, thereby validating Ab specificity (Figure S1C). Two-way

BRCA1/SETX co-IP was also apparent in primary, diploid human BJ-hTERT fibroblast extracts (Figure 1A, lower panels), suggesting the existence of a physiological interaction between BRCA1 and SETX.

In an effort to map BRCA1 and SETX domains that participate in their interaction, we tested multiple, GST-tagged recombinant fragments that contain various BRCA1 and SETX sequences in glutathione S-transferase (GST) binding assays (Scully et al., 1997; Suraweera et al., 2009) (Figure S1D). The results showed that SETX interacts with F6C, which contains a BRCT (BRCA1 C-terminal) motif (Figure 1B). Intriguingly, F6, which also includes F6C, did not interact with SETX, although F6 was sufficiently folded to interact with another BRCT domain binding protein, BACH1 (Figure S1E) (Cantor et al., 2001). Reciprocal SETX mapping revealed that fragment S8 interacted with BRCA1 (Figure 1C). These results imply that specific structural units of both proteins participate in their interaction. A more extensive analysis will be required to obtain a full picture of this process.

In light of the existence of BRCA1/SETX binding, we tested for BRCA1 corecruitment at transcription termination sites associated with R-loops by chromatin immunoprecipitation (ChIP) performed on the endogenous human  $\beta$ -actin gene (Figure 1D) (Skourti-Stathaki et al., 2011). Cross-linked HeLa cell chromatin was harvested before and after siRNA depletion of either BRCA1 or SETX. Depletion efficiency was verified at the protein level (Figure 1E).

Specifically, we found that in mock siRNA-transfected cells BRCA1 binding was significantly enriched (3 to 6 fold) at the transcription termination site (5' pause and pause site probes) when compared to the signals obtained following BRCA1 depletion (Figure 1F). Notably, while BRCA1 depletion suppressed these binding signals, SETX depletion did not. We also confirmed the existence of SETX accumulation at the  $\beta$ -actin transcription termination pause site (Skourti-Stathaki et al., 2011) (Figure 1G). BRCA1 depletion impaired SETX binding to these sites, implying that its recruitment is BRCA1-dependent at this locus. Similar results were obtained when BRCA1 and SETX hairpins were substituted for cognate siRNA species (Figure S2A-B). These findings demonstrate a physiological BRCA1/SETX interaction at the  $\beta$ -actin transcription termination site and suggest that BRCA1 acts as an anchor for SETX therein.

### **R-loop formation at the $\beta$ -actin gene transcription termination site triggers BRCA1/SETX complex recruitment**

To search for a connection between the formation of R-loops and the recruitment of BRCA1/SETX at the  $\beta$ -actin termination site, we first asked whether RNA:DNA hybrids forming within this gene are affected by BRCA1 and SETX depletion. We measured R-loop formation over the  $\beta$ -actin gene, employing DNA:RNA hybrid immunoprecipitation (DIP) of uncrosslinked DNA using a monoclonal Ab specific for RNA:DNA hybrids (S9.6) (Skourti-Stathaki et al., 2011). As shown in Figure 2A, a modest, but reproducible, increase in R-loop abundance was detected within the  $\beta$ -actin gene pause site, following BRCA1 or SETX depletion. While this result suggests that BRCA1 and SETX affect R-loop formation, it is important to note that BRCA1 depletion also elicited a phenotype suggestive of a

transcription termination defect that was previously reported upon SETX depletion. It was manifest by RNAPII accumulation at the pause site, (Skourti-Stathaki et al., 2011) (Figure S2C). The recruitment of other RNA:DNA helicases in the absence of SETX to resolve R-loops over these regions, as well as potential differences in transcription levels after BRCA1 and SETX depletion might explain the lack of a more substantial increase in R-loops levels.

We then tested BRCA1 and SETX chromatin occupancy over the  $\beta$ -actin gene following R-loop suppression. ChIP experiments were performed in HeLa cells before and after over-expression of GFP-tagged RNaseH1 (Figure S3A), which significantly reduced R-loop formation over the  $\beta$ -actin gene (Skourti-Stathaki et al., 2014). In RNaseH1 over-expressing cells, we observed a significant suppression of BRCA1 and SETX binding at the transcription termination site, as compared with control cells (Figure 2B). Moreover, as shown in Figure 2C, RNaseH1 over-expression led to a reduction in the co-IP signal, implying that BRCA1/SETX complex formation at these genomic loci is mediated by R-loops.

Thus, our results suggest that BRCA1/SETX complexes bind at TR sites in response to local R-loop formation and that they also suppress R-loop abundance, in part, by participating in the regulation of transcription and of transcription termination.

### **BRCA1/SETX complexes protect DNA at R-loop-associated termination pause sites from the development of ssDNA breaks**

Since R-loop accumulation can increase the risk of genomic instability by promoting DNA damage (Hamperl and Cimprich, 2014; Skourti-Stathaki and Proudfoot, 2014), we tested whether BRCA1 and/or SETX depletion triggers the generation of DNA damage in the vicinity of persistent R-loops. Specifically,  $\gamma$ H2AX ChIP assays were performed over the  $\beta$ -actin gene. After depletion of either BRCA1 or SETX we observed significantly increased signals of  $\gamma$ H2AX largely restricted to the termination region (Figure 3A). This implied that loss of BRCA1/SETX complexes over the R-loop-associated pause site was contributed to the accumulation of local damage.

Since several studies suggest that the ssDNA of an R-loop is prone to transcription-associated mutations or DNA breaks (Aguilera and García-Muse, 2012; Wimberly et al., 2013), we adapted a ligation-mediated quantitative PCR (LM-qPCR) approach to test whether ssDNA breaks had occurred on the non-template strand following BRCA1 or SETX depletion (Figure 3B). Briefly, we ligated an adaptor to putative ssDNA ends after performing a primer extension reaction over the far 3'-end of the  $\beta$ -actin gene. The goal was to amplify any post-ligation DNA fragments that harbor both the adaptor sequence and a segment of the 3'-end of the  $\beta$ -actin gene.

In the absence of BRCA1 or SETX, significant signals were observed when using 5' pause and pause site primers, thereby reflecting the existence of one or more ssDNA breaks located near the termination region (Figures 3C and S3B). The 3' R-loop-free C region (Figure 2A) was used as a negative control to validate the specificity of the above-noted results. We also tested the anti-sense strand and failed to observe LM-qPCR signals above

background following BRCA1 or SETX depletion (Figure S3C). This suggested that the observed breaks affected only the non-template strand.

We then asked whether RNaseH1 over-expression suppressed the appearance of ssDNA breaks in the  $\beta$ -actin gene in BRCA1- or SETX-depleted cells. The background signals in control region C remained barely detectable, while RNaseH1-expressing BRCA1- or SETX-depleted cells revealed a significant signal decrease within the 5' pause and pause regions (Figures 3C and S3B). Taken together, these data argue that, in the  $\beta$ -actin gene, BRCA1 and SETX are involved in the prevention/repair of ssDNA breaks arising within the coding strand R-loop near their 3'-end binding site(s).

### **BRCA1/SETX complexes form at several R-loop-associated genomic loci and protect cells from the development of ssDNA breaks**

To search for a global effect of BRCA1/SETX function, we first searched for evidence within 2 gene sets. On one hand, we analyzed 2 genes (*ENSA*, *Gemin7*) where transcription termination is also dependent on R-loops (Skourti-Stathaki et al., 2014). In both genes, BRCA1 and SETX binding was enriched at the relevant transcription termination regions as compared to an irrelevant intronic segment (Figures 4A, 4B, S4A and S4B). Although the BRCA1 antibody used in CHIP experiments displayed some background signal, its specificity was confirmed by its sensitivity to BRCA1 depletion. In parallel, we studied the *Akirin1* and *cyclinB1* genes in which transcription termination is regulated by an R-loop-free co-transcriptional cleavage (CoTC) mechanism (Nojima et al., 2013; Skourti-Stathaki et al., 2014; White et al., 2013). BRCA1 and SETX CHIP analysis revealed no specific signals in either intronic region or the CoTC regions of *Akirin1* and *cyclinB1*; reinforcing the hypothesis that BRCA1/SETX recruitment requires R-loop formation (Figures 4A, 4B, S4A and S4B).

We also examined the DNA damage signaling response in these 2 categories of genes by searching for  $\gamma$ H2AX signals over the termination regions using CHIP qPCR. Similar to  $\beta$ -actin, increased  $\gamma$ H2AX signals were clearly visible over the pause element of both *ENSA* and *Gemin7* following BRCA1 or SETX depletion, but not over the control intronic region (Figures 4C and S4C). By contrast, no  $\gamma$ H2AX enrichment was detected at “R-loop-free” CoTC termination regions within *Akirin1* or *cyclinB1*. These results are consistent with our  $\beta$ -actin CHIP data (Figure 3A) and with the presence of local  $\gamma$ H2AX signals in the absence of BRCA1/SETX complexes over *ENSA* and *Gemin7*. Thus, they can be viewed as a reflection of local R-loop-associated DNA damage. These findings imply that BRCA1/SETX complexes participate in the repair or prevention of this form of damage in these two genes bearing R-loop-associated transcription termination regions.

We then asked whether BRCA1/SETX function at R-loop-forming termination sites is a genome-wide phenomenon. First, comet assays were performed to search for widespread DNA breaks as a consequence of BRCA1 or SETX depletion. As predicted, under alkaline conditions, which detect both ssDNA breaks and DSBs, we observed a significant increase in comet tail lengths in both BRCA1- and SETX-depleted cells. Moreover, RNaseH1 overexpression in both settings decreased the abundance of cells with comets as well as their length (Figure 4D-4E). Similar results were observed when the cells were treated with a low



dose of  $\alpha$ -amanitin that acts to inhibit RNAPII elongation. These results confirm that R-loop structures, as by-products of transcription, create numerous fragile genomic sequences prone to BRCA1/SETX depletion-associated damage. Although we also observed an increase in comet tail lengths in BRCA1- and SETX-depleted cells assayed in neutral buffer conditions (likely reflecting DSB), these DNA damage phenotypes did not revert after overexpression of RNaseH1 or exposure to  $\alpha$ -amanitin (Figure S4D). Thus, R-loop dependent DNA damage following BRCA1/SETX depletion is mainly composed of ssDNA breaks.

### **BRCA1 is globally associated with transcription termination sites of highly transcribed genes**

In light of the possibility that R-loop-associated termination regions are characteristics of larger numbers of genes, we performed a meta-analysis from several deep sequencing datasets to search for genomic co-localization of BRCA1 peaks at transcription termination regions. We used the ENCODE BRCA1 ChIP-seq data set to identify candidate genes regulated by this mechanism (Consortium, 2012). For each gene, a candidate transcription termination region (TR) was defined as a 4kb segment downstream of the transcription termination site (TTS)(Figure 5A).

Initially, 764 distinct BRCA1 peaks significantly overlapping a genomic TR were identified. Since BRCA1 is also associated with a subset of promoters (McKinney, K. et al, unpublished) (Gardini et al., 2014), we further filtered these data to exclude any peaks overlapping promoter regions and their associated transcripts. In this filtered gene set, 196 BRCA1 peaks, each associated with a TR, corresponded to 184 genes that are referred to here as BRCA1 TR genes (Figure 5B and Table S1).

Since the DNA damage that developed over R-loop-associated TRs upon BRCA1/SETX depletion is proposed to be R-loop-, and, therefore, transcription-dependent, we analyzed the expression levels of these BRCA1 TR genes. RNA-seq data were used to analyze HeLa cell transcriptomes available from ENCODE. The expression profiles of genes whose filtered TRs did or did not overlap BRCA1 ChIP-seq peaks were compared. The data showed that BRCA1 TR genes clearly displayed higher mRNA levels ( $p=1.1E-7$ ), when compared to genes whose termination sites are not associated with a BRCA1 binding site (Figures 5C, S5 and Table S1). These results indicate that BRCA1 TR genes are more highly expressed than those lacking BRCA1 TR binding peaks.

To explore further whether BRCA1 binding at TR is associated with transcription termination, we searched for evidence of the co-localization at termination sites of BRCA1 peaks and RNAPII paused at the 3' end of the relevant genes (RNAPII S2P) (Davidson et al., 2014). Here we used ENCODE RNAPII CTD Ser2-P ChIP-seq data (Consortium, 2012). This analysis showed that more than 70% of BRCA1 peaks within TRs overlap paused RNAPII (Figure 5D and Table S1), suggesting that BRCA1 is actively engaged in transcription-associated events at termination sites. To investigate whether BRCA1 TR are associated with R-loops, we compared them with the RNA:DNA IP database (DRIP-seq) generated with cells before and after RNaseH1 treatment (DRIP+RH) (Ginno et al., 2012). We found that BRCA1 TR peaks show a significant enrichment of DRIP signal as compared

to DRIP+RH (Figure 5E), implying the existence of R-loop formation at the BRCA1-associated TR loci.

These results suggest that in mammalian cells BRCA1 binds to TR (associated with paused RNAPII) in a substantial subset of highly transcribed genes, whose pause sites reflect a strong tendency to form R-loops. Taken together, our findings indicate that BRCA1 and SETX participate in the prevention/repair of ssDNA damage occurring at specific regions (TR) in response to the transcription-associated formation of R-loops in physiological conditions. Conceivably, these genes require BRCA1/SETX-dependent DNA damage surveillance to ensure regulation of transcription (Huppert et al., 2008; Skourti-Stathaki and Proudfoot, 2014).

### **BRCA1 mutant breast cancers reveal genomic alterations at BRCA1-associated termination sites**

We next asked whether a defect in the newly detected BRCA1/SETX function could participate in the pathogenesis of BRCA1-deficient breast cancer tumors. Recent large-scale DNA sequencing screens performed in various human cancers have shed light on the nature and location of associated somatic mutations (Alexandrov et al., 2013; Pleasance et al., 2010). We performed a mutation analysis in the 184 BRCA1 TR genes, using the complete catalog of somatic mutations obtained from the whole genome sequencing of 21 breast cancers, a subset of which (n=5) were BRCA1 mutant (Nik-Zainal et al., 2012). This catalog includes single base mutations (SBM), insertions and deletions (indels), and chromosomal rearrangements. For each patient the genome of healthy/normal mammary tissue was also sequenced, and any mutations therein were deleted from the patient mutation analysis. Since R-loops can be long, we first included promoter regions in this search. More specifically, we first defined for each BRCA1 TR gene a whole gene region of special interest from TSS -1250bp to TTS +5kb and searched for BRCA1-specific mutations.

We divided the 21 patient data sets into 3 breast cancer types: 12 sporadic (i.e. tumors WT/WT for both BRCA1 and BRCA2), 5 BRCA1, and 4 BRCA2 mutant tumors. Figure 6A compares each of the mutant groups to WT, in search of a quantitative measure of any difference in mutation rate (i.e. effect size) between any 2 groups. The effect size was computed as a standardized difference in mutation counts when all selected genes are considered together. It was obtained separately for each type of mutation analyzed. The results clearly indicate that, within the BRCA1 TR gene regions of interest, there was significant enrichment for indels only in the mutant BRCA1 tumors ( $z=2.875$ ,  $p=0.009$ ), despite the fact that more indels were detected throughout the entire genomes of BRCA2 than BRCA1 breast cancers (Nik-Zainal et al., 2012). By contrast, these genomic loci were enriched for SBM in the BRCA2 tumors ( $z=2.144$ ,  $p=0.021$ ). Of note, no significant enrichment for rearrangements was observed. These results indicate that, within BRCA1 TR genes, distinct mutational patterns exist that separate BRCA1 and BRCA2 tumors. This suggests that any biochemical defects that the absence of these proteins elicits may, at least in part, be different.

In parallel, we also performed a global mutation analysis among the “R-loop free” CoTC-regulated genes (“negative” genes) (Nojima et al., 2013) where BRCA1 is not recruited



(Figure 4A). We failed to detect any significant change in the species of mutations detected in either BRCA1 or BRCA2 null tumors (Figure S6).

To further investigate the mutational profile at the 3'-ends of the BRCA1 TR genes, we focused on a narrower region defined as: TTS +/-4kb. Strikingly, the only significant enrichment observed was for indels in BRCA1-mutant tumors ( $z=3.464$ ,  $p=0.002$ ) (Figure 6B). Precise mapping of the indels in the BRCA1 TR genes in the BRCA1-null tumors indicated that 3 out of 6 were located ~300-400bp from the BRCA1 TR peak; 2 were located further away (last intron and ~1.8kb downstream of the BRCA1 peak); and another was close to the promoter region (Figures 6C-6E and Table S2). The absence of SBM enrichment in the 3'-end regions of the TR genes in BRCA2 mutant carriers suggests that most of it accumulates at 5'-ends and within the gene body. Of note, none of the few mutations detected in the CoTC genes was located in 3'-end regions.

Overall, these results show that BRCA1-deficient tumors are significantly enriched for indels in BRCA1 TR genes compared to BRCA2 and sporadic breast cancers. A significant fraction of the BRCA1 tumor-associated indels lie in the vicinity of BRCA1 binding TR loci, supporting the view that BRCA1 plays a critical role in preventing/repairing R-loop-mediated damage in the vicinity of R-loop-associated TR.

## DISCUSSION

In this study, a combination of biochemical, molecular, and genetic data provide evidence that a newly identified BRCA1/SETX complex is required to restrain the development of R-loop-mediated DNA damage at specific genomic regions. In particular, we have delineated a sequence of events that is predicted to occur over a substantial subset of R-loop-associated transcriptional pause sites and, when impaired, results in the BRCA1 TR-associated indel mutations observed in BRCA1<sup>-/-</sup> breast cancer genomes.

Under physiological conditions, R-loops play a role in transcription termination at G-rich pause sites, like those in *β-actin*, *Gemin7* and *ENSA* (Skourti-Stathaki et al., 2014). They are also present at promoter regions where there is a region of GC-skew that insures the protection of vicinal CpG island sequences against DNA methylation (Ginno et al., 2012). Though these structures represent R-loops associated with distinct biological functions, they are, nonetheless, still predicted to induce DNA damage (for review see (Skourti-Stathaki and Proudfoot, 2014)).

How BRCA1 recognizes R-loop structures remains to be determined. Current knowledge focuses mainly upon the factors that regulate their formation (Chan et al., 2014; Hamperl and Cimprich, 2014; Skourti-Stathaki and Proudfoot, 2014). However, little is known about potential direct interaction between specific proteins and R-loops. Several possibilities exist. For example, there could be direct recognition of a RNA:DNA hybrid, e.g. as occurs with the hybrid-binding domain (HBD) of RNaseH1 (Cerritelli and Crouch, 2009), and/or of the non-template ssDNA. For BRCA1, it is also possible that the branched structure of an R-loop resembles the DNA flaps, branched DNA, and/or four-way junctions for which BRCA1 exhibits an intrinsic affinity (Paull et al., 2001). Alternatively, BRCA1 recruitment to R-

loops could be mediated by other protein complexes, that normally interact with these special DNA structures. These could be components of the RNAPII holoenzyme, splicing machinery, and/or specific chromatin remodeling complexes that are associated with R-loops (Bochar et al., 2000; Boulé and Zakian, 2007; Savage et al., 2014a; Scully et al., 1997; Yarden and Brody, 1999).

Results presented here show that a deficiency in BRCA1/SETX function results in unrepaired ssDNA breaks on the non-template strand at certain R-loop-associated TRs.  $\gamma$ H2AX accumulation was observed by ChIP, although the extent of its genomic presence appears to be more restricted than observed at sites of DSB (Paull et al., 2000). It is tempting to speculate that the restricted presence of  $\gamma$ H2AX signals is a reflection of the presence of ssDNA breaks as the source of DNA damage and/or that  $\gamma$ H2AX spreading is antagonized by high levels of transcription (Iacovoni et al., 2010).

The above considerations aside, how DNA damage occurs at such sites remains unknown. Multiple molecular mechanisms may foster or contribute to the DNA damage associated with R-loops. Superhelical stress and G4 and/or flap endonucleases could generate local ssDNA breaks, some of which could devolve into DSB (Hamperl and Cimprich, 2014). In addition, genome-wide sequencing of human cancers suggests that non-random, clustered mutations may be concentrated in long ssDNA regions some of which might form R-loop structures (Roberts et al., 2012). Indeed, during transcription, possibly due to the relative chemical susceptibility to damage of ssDNA, there is an increase in the mutation rate associated with the activity of editing enzymes like AID (Activation-Induced cytosine Deaminase) or APOBEC proteins (Apolipoprotein B mRNA-Editing Catalytic Polypeptide) (Alexandrov et al., 2013; Beale et al., 2004; Chan et al., 2012). Recently, the latter have been suggested to play a role in the mutational processes that affect breast cancer genomes (Burns et al., 2013; Nik-Zainal et al., 2012; 2014).

Taken together, our findings show that BRCA1 contributes to the control and/or repair of R-loop-mediated DNA damage at specific sites, which is, potentially, a significant contributor to the maintenance of genomic stability. Similarly, the implications of a proposed role for BRCA2 in Transcription-Associated Recombination (TAR) and in the processing of R-loops in partnership with RNA processing factors suggest that these structures are a source of cancer-related instability (Bhatia et al., 2014; Gallardo et al., 2003; Huertas and Aguilera, 2003; Savolainen and Helleday, 2009).

*BRCA1*- and *BRCA2*-mutant cancers exhibit differences in histopathology, gene expression profiles, and clinical course, even though they share similarities in their marked levels of genomic instability (Nik-Zainal et al., 2012). By contrast, our observations imply that, at certain transcription termination regions, the mutational signatures of *BRCA1*- and *BRCA2*-null tumors are different. Conceivably, these differences arise from the possibility that *BRCA1* and 2 respond to/interact with R-loops differently and even in different contexts. Additional whole genome sequencing from greater numbers of *BRCA* mutant cancers would be required to address these possibilities in the future. Further studies will also be needed to determine whether such *BRCA1/BRCA2* mutational differences contribute-directly or indirectly- to the biological differences between *BRCA1* and *BRCA2* breast cancers.

## Experimental Procedures

For detailed experimental procedures, see Supplemental Information.

**Co-immunoprecipitation (co-IP) analysis**—Whole cell extracts (WCE) were prepared as previously described (Tardat et al., 2010), except that the lysis mixture was sonicated with a Sonic Dismembrator for 15s at an amplitude of 20% (Fisher Scientific, Model 120). 200µg of WCE were incubated for 2hr at 4°C with anti-BRCA1#1 (SG11, mouse monoclonal), anti-BRCA1#2 (MS110, mouse monoclonal), anti-SETX#1 (A301-105A, Bethyl), anti-SETX#2 (A301-104A, Bethyl), or control IgG. Immune complexes were collected with protein A/G magnetic beads (Dynabeads, Invitrogen) for 15min at 4°C and washed with IP buffer (with 1mM DTT) and increasing concentrations of KCl (50/100/150mM KCl). Bound proteins were eluted in LDS sample buffer 1X (Invitrogen) and analyzed by immunoblotting.

## qPCR ChIP and DIP experiments

ChIP experiments on BRCA1, SETX and  $\gamma$ H2AX were performed using a modified version of N.J.P laboratory's protocol (Skourti-Stathaki et al., 2011), as detailed in the Extended Experimental Procedures. DNA:RNA hybrid precipitation (DIP) analysis was performed with the specific monoclonal RNA:DNA hybrid Ab (S9.6) as described in (Skourti-Stathaki et al., 2011). Sequences of the DNA primers are listed in Extended Experimental Procedures.

## Mutational analyses of breast cancers

Detailed description of the computational and statistical analyses is available in Extended Experimental Procedures.

The integrative Genomics Viewer (IGV) genome browser was used to visualize the indels and the BRCA1 and RNAPIISer2P ChIP-seq profiles across the termination regions of the BRCA1 TR genes (Thorvaldsdóttir et al., 2013).

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements

We thank all Livingston laboratory members for support, technical advice, and helpful discussions. E.H and other Livingston members were supported by grants from the National Cancer Institute (NCI) - Mechanisms of Breast Development and Carcinogenesis (2P01CA80111-16) and BRCA1 Function in Post Damage Foci (5R01CA136512-05). K.K.G was supported by Marie Curie IEF. KSS and NJP were supported by a Programme grant from the Wellcome Trust and G.P by NCI 5P30 CA006516-46.

## REFERENCES

- Aguilera A. The connection between transcription and genomic instability. *Embo J.* 2002; 21:195–201. [PubMed: 11823412]
- Aguilera A, García-Muse T. R Loops: From Transcription Byproducts to Threats to Genome Stability. *Molecular Cell.* 2012; 46:115–124. [PubMed: 22541554]

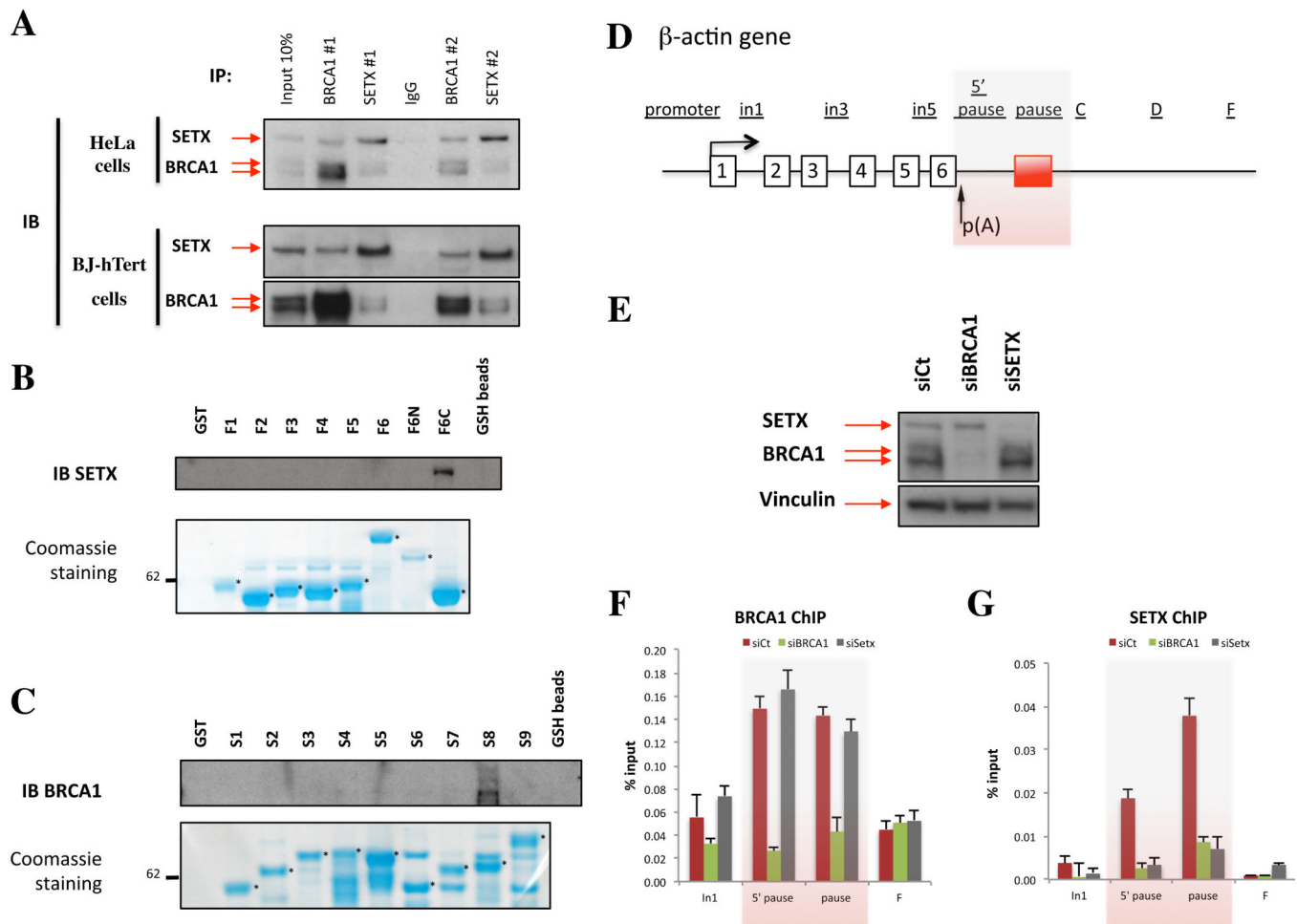
- Aguilera A, Gómez-González B. Genome instability: a mechanistic view of its causes and consequences. *Nat Rev Genet.* 2008; 9:204–217. [PubMed: 18227811]
- Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SAJR, Behjati S, Biankin AV, Bignell GR, Bolli N, Borg Å, Borresen-Dale A-L, et al. Signatures of mutational processes in human cancer. *Nature.* 2013; 500:415–421. [PubMed: 23945592]
- Alzu A, Bermejo R, Begnis M, Lucca C, Piccini D, Carotenuto W, Saponaro M, Brambati A, Cocito A, Foiani M, et al. Senataxin Associates with Replication Forks to Protect Fork Integrity across RNA-Polymerase-II-Transcribed Genes. *Cell.* 2012; 151:835–846. [PubMed: 23141540]
- Anderson SF, Schlegel BP, Nakajima T, Wolpin ES, Parvin JD. BRCA1 protein is linked to the RNA polymerase II holoenzyme complex via RNA helicase A. *Nat Genet.* 1998; 19:254–256. [PubMed: 9662397]
- Beale RCL, Petersen-Mahrt SK, Watt IN, Harris RS, Rada C, Neuberger MS. Comparison of the differential context-dependence of DNA deamination by APOBEC enzymes: correlation with mutation spectra in vivo. *Journal of Molecular Biology.* 2004; 337:585–596. [PubMed: 15019779]
- Becherel OJ, Yeo AJ, Stellati A, Heng EYH, Luff J, Suraweera AM, Woods R, Fleming J, Carrie D, McKinney K, et al. Senataxin plays an essential role with DNA damage response proteins in meiotic recombination and gene silencing. *PLoS Genet.* 2013; 9:e1003435. [PubMed: 23593030]
- Bennett CB, Westmoreland TJ, Verrier CS, Blanchette CAB, Sabin TL, Phatnani HP, Mishina YV, Huper G, Selim AL, Madison ER, et al. Yeast screens identify the RNA polymerase II CTD and SPT5 as relevant targets of BRCA1 interaction. *PLoS ONE.* 2008; 3:e1448. [PubMed: 18197258]
- Bhatia V, Barroso SI, García-Rubio ML, Tumini E, Herrera-Moyano E, Aguilera A. BRCA2 prevents R-loop accumulation and associates with TREX-2 mRNA export factor PCID2. *Nature.* 2014; 511:362–365. [PubMed: 24896180]
- Bochar DA, Wang L, Beniya H, Kinev A, Xue Y, Lane WS, Wang W, Kashanchi F, Shiekhhattar R. BRCA1 is associated with a human SWI/SNF- related complex: linking chromatin remodeling to breast cancer. *Cell.* 2000; 102:257–265. [PubMed: 10943845]
- Boulé J-B, Zakian VA. The yeast Pif1p DNA helicase preferentially unwinds RNA DNA substrates. *Nucleic Acids Research.* 2007; 35:5809–5818. [PubMed: 17720711]
- Burns MB, Lackey L, Carpenter MA, Rathore A, Land AM, Leonard B, Refsland EW, Kotandeniya D, Tretyakova N, Nikas JB, et al. APOBEC3B is an enzymatic source of mutation in breast cancer. *Nature.* 2013; 494:366–370. [PubMed: 23389445]
- Cantor SB, Bell DW, Ganesan S, Kass EM, Drapkin R, Grossman S, Wahrer DC, Sgroi DC, Lane WS, Haber DA, et al. BACH1, a novel helicase-like protein, interacts directly with BRCA1 and contributes to its DNA repair function. *Cell.* 2001; 105:149–160. [PubMed: 11301010]
- Cerritelli SM, Crouch RJ. Ribonuclease H: the enzymes in eukaryotes. *Febs J.* 2009; 276:1494–1505. [PubMed: 19228196]
- Chan K, Sterling JF, Roberts SA, Bhagwat AS, Resnick MA, Gordenin DA. Base Damage within Single-Strand DNA Underlies In Vivo Hypermutability Induced by a Ubiquitous Environmental Agent. *PLoS Genet.* 2012; 8:e1003149. [PubMed: 23271983]
- Chan YA, Hieter P, Stirling PC. Mechanisms of genome instability induced by RNA-processing defects. *Trends Genet.* 2014; 30:245–253. [PubMed: 24794811]
- Consortium TEP. An integrated encyclopedia of DNA elements in the human genome. *Nature.* 2012; 489:57–74. [PubMed: 22955616]
- Daniel JA, Nussenzweig A. The AID-Induced DNA Damage Response in Chromatin. *Molecular Cell.* 2013; 50:309–321. [PubMed: 23664375]
- Davidson L, Muniz L, West S. 3' end formation of pre-mRNA and phosphorylation of Ser2 on the RNA polymerase II CTD are reciprocally coupled in human cells. *Genes Dev.* 2014; 28:342–356. [PubMed: 24478330]
- Gaillard H, Herrera-Moyano E, Aguilera A. Transcription-associated genome instability. *Chem. Rev.* 2013; 113:8638–8661. [PubMed: 23597121]
- Gallardo M, Luna R, Erdjument-Bromage H, Tempst P, Aguilera A. Nab2p and the Thp1p-Sac3p complex functionally interact at the interface between transcription and mRNA metabolism. *J. Biol. Chem.* 2003; 278:24225–24232. [PubMed: 12702719]

- Gan W, Guan Z, Liu J, Gui T, Shen K, Manley JL, Li X. R-loop-mediated genomic instability is caused by impairment of replication fork progression. *Genes Dev.* 2011; 25:2041–2056. [PubMed: 21979917]
- Gardini A, Baillat D, Cesaroni M, Shiekhhattar R. Genome-wide analysis reveals a role for BRCA1 and PALB2 in transcriptional co-activation. *Embo J.* 2014; 33:890–905. [PubMed: 24591564]
- Ginno PA, Lott PL, Christensen HC, Korf I, Chédin F. R-Loop Formation Is a Distinctive Characteristic of Unmethylated Human CpG Island Promoters. *Molecular Cell.* 2012; 45:814–825. [PubMed: 22387027]
- Gorski JJ, Savage KI, Mulligan JM, McDade SS, Blayney JK, Ge Z, Harkin DP. Profiling of the BRCA1 transcriptome through microarray and ChIP-chip analysis. *Nucleic Acids Research.* 2011; 39:9536–9548. [PubMed: 21880590]
- Hamperl S, Cimprich KA. The contribution of co-transcriptional RNA:DNA hybrid structures to DNA damage and genome instability. *DNA Repair.* 2014; 19:84–94. [PubMed: 24746923]
- Helmrich A, Ballarino M, Tora L. Collisions between replication and transcription complexes cause common fragile site instability at the longest human genes. *Molecular Cell.* 2011; 44:966–977. [PubMed: 22195969]
- Hill SJ, Rolland T, Adelmant G, Xia X, Owen MS, Dricot A, Zack TI, Sahni N, Jacob Y, Hao T, et al. Systematic screening reveals a role for BRCA1 in the response to transcription-associated DNA damage. *Genes Dev.* 2014; 28:1957–1975. [PubMed: 25184681]
- Huen MSY, Sy SMH, Chen J. BRCA1 and its toolbox for the maintenance of genome integrity. *Nat. Rev. Mol. Cell Biol.* 2010; 11:138–148. [PubMed: 20029420]
- Huertas P, Aguilera A. Cotranscriptionally Formed DNA:RNA Hybrids Mediate Transcription Elongation Impairment and Transcription-Associated Recombination. *Molecular Cell.* 2003; 12:711–721. [PubMed: 14527416]
- Huppert JL, Bugaut A, Kumari S, Balasubramanian S. G-quadruplexes: the beginning and end of UTRs. *Nucleic Acids Research.* 2008; 36:6260–6268. [PubMed: 18832370]
- Iacovoni JS, Caron P, Lassadi I, Nicolas E, Massip L, Trouche D, Legube GEL. High-resolution profiling of  $\gamma$ H2AX around DNA double strand breaks in the mammalian genome. *Embo J.* 2010; 29:1446–1457. [PubMed: 20360682]
- Kawai S, Amano A. BRCA1 regulates microRNA biogenesis via the DROSHA microprocessor complex. *The Journal of Cell Biology.* 2012; 197:201–208. [PubMed: 22492723]
- Kim N, Jinks-Robertson S. Transcription as a source of genome instability. *Nat Rev Genet.* 2012; 13:204–214. [PubMed: 22330764]
- Kleiman FE, Manley JL. Functional interaction of BRCA1-associated BARD1 with polyadenylation factor CstF-50. *Science.* 1999; 285:1576–1579. [PubMed: 10477523]
- Kleiman FE, Wu-Baer F, Fonseca D, Kaneko S, Baer R, Manley JL. BRCA1/BARD1 inhibition of mRNA 3' processing involves targeted degradation of RNA polymerase II. *Genes Dev.* 2005; 19:1227–1237. [PubMed: 15905410]
- Li X, Manley JL. Cotranscriptional processes and their influence on genome stability. *Genes Dev.* 2006; 20:1838–1847. [PubMed: 16847344]
- Mischo HE, Gómez-González B, Grzechnik P, Rondón AG, Wei W, Steinmetz L, Aguilera A, Proudfoot NJ. Yeast Sen1 Helicase Protects the Genome from Transcription-Associated Instability. *Molecular Cell.* 2011; 41:21–32. [PubMed: 21211720]
- Mullan PB, Quinn JE, Harkin DP. The role of BRCA1 in transcriptional regulation and cell cycle control. *Oncogene.* 2006; 25:5854–5863. [PubMed: 16998500]
- Nik-Zainal S, Alexandrov LB, Wedge DC, Van Loo P, Greenman CD, Raine K, Jones D, Hinton J, Marshall J, Stebbings LA, et al. Mutational Processes Molding the Genomes of 21 Breast Cancers. *Cell.* 2012; 149:979–993. [PubMed: 22608084]
- Nik-Zainal S, Wedge DC, Alexandrov LB, Petljak M, Butler AP, Bolli N, Davies HR, Knappskog S, Martin S, Papaemmanuil E, et al. Association of a germline copy number polymorphism of APOBEC3A and APOBEC3B with burden of putative APOBEC-dependent mutations in breast cancer. *Nat Genet.* 2014; 46:487–491. [PubMed: 24728294]

- Nojima T, Dienstbier M, Murphy S, Proudfoot NJ, Dye MJ. Definition of RNA Polymerase II CoTC Terminator Elements in the Human Genome. *Cell Reports*. 2013; 3:1080–1092. [PubMed: 23562152]
- Pathania S, Nguyen J, Hill SJ, Scully R, Adelmant GO, Marto JA, Feunteun J, Livingston DM. BRCA1 Is Required for Postreplication Repair after UV-Induced DNA Damage. *Molecular Cell*. 2011; 44:235–251. [PubMed: 21963239]
- Paull TT, Cortez D, Bowers B, Elledge SJ, Gellert M. Direct DNA binding by Brca1. *Proceedings of the National Academy of Sciences of the United States of America*. 2001; 98:6086–6091. [PubMed: 11353843]
- Paull TT, Rogakou EP, Yamazaki V, Kirchgessner CU, Gellert M, Bonner WM. A critical role for histone H2AX in recruitment of repair factors to nuclear foci after DNA damage. *Current Biology*. 2000; 10:886–895. [PubMed: 10959836]
- Pleasance ED, Cheetham RK, Stephens PJ, McBride DJ, Humphray SJ, Greenman CD, Varela I, Lin M-L, Ordóñez GR, Bignell GR, et al. A comprehensive catalogue of somatic mutations from a human cancer genome. *Nature*. 2010; 463:191–196. [PubMed: 20016485]
- Roberts SA, Sterling J, Thompson C, Harris S, Mav D, Shah R, Klimczak LJ, Kryukov GV, Malc E, Mieczkowski PA, et al. Clustered Mutations in Yeast and in Human Cancers Can Arise from Damaged Long Single-Strand DNA Regions. *Molecular Cell*. 2012; 46:424–435. [PubMed: 22607975]
- Savage KI, Gorski JJ, Barros EM, Irwin GW, Manti L, Powell AJ, Pellagatti A, Lukashchuk N, McCance DJ, McCluggage WG, et al. Identification of a BRCA1-mRNA splicing complex required for efficient DNA repair and maintenance of genomic stability. *Molecular Cell*. 2014a; 54:445–459. [PubMed: 24746700]
- Savage KI, Matchett KB, Barros EM, Cooper KM, Irwin GW, Gorski JJ, Orr KS, Vohhodina J, Kavanagh JN, Madden AF, et al. BRCA1 Deficiency Exacerbates Estrogen-Induced DNA Damage and Genomic Instability. *Cancer Res*. 2014b; 74:2773–2784. [PubMed: 24638981]
- Savolainen L, Helleday T. Transcription-associated recombination is independent of XRCC2 and mechanistically separate from homology-directed DNA double-strand break repair. *Nucleic Acids Research*. 2009; 37:405–412. [PubMed: 19043071]
- Scully R, Anderson SF, Chao DM, Wei W, Ye L, Young RA, Livingston DM, Parvin JD. BRCA1 is a component of the RNA polymerase II holoenzyme. *Proceedings of the National Academy of Sciences of the United States of America*. 1997; 94:5605–5610. [PubMed: 9159119]
- Silver DP, Livingston DM. Mechanisms of BRCA1 Tumor Suppression. *Cancer Discovery*. 2012; 2:679–684. [PubMed: 22843421]
- Skourti-Stathaki K, Proudfoot NJ. A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression. *Genes Dev*. 2014; 28:1384–1396. [PubMed: 24990962]
- Skourti-Stathaki K, Kamieniarz-Gdula K, Proudfoot NJ. R-loops induce repressive chromatin marks over mammalian gene terminators. *Nature*. 2014
- Skourti-Stathaki K, Proudfoot NJ, Gromak N. Human Senataxin Resolves RNA/DNA Hybrids Formed at Transcriptional Pause Sites to Promote Xrn2-Dependent Termination. *Molecular Cell*. 2011; 42:794–805. [PubMed: 21700224]
- Steinmetz EJ, Warren CL, Kuehner JN, Panbehi B, Ansari AZ, Brow DA. Genome-Wide Distribution of Yeast RNA Polymerase II and Its Control by Sen1 Helicase. *Molecular Cell*. 2006; 24:735–746. [PubMed: 17157256]
- Stirling PC, Chan YA, Minaker SW, Aristizabal MJ, Barrett I, Sipahimalani P, Kobor MS, Hieter P. R-loop-mediated genome instability in mRNA cleavage and polyadenylation mutants. *Genes Dev*. 2012; 26:163–175. [PubMed: 22279048]
- Suraweera A, Lim Y, Woods R, Birrell GW, Nasim T, Becherel OJ, Lavin MF. Functional role for senataxin, defective in ataxia oculomotor apraxia type 2, in transcriptional regulation. 2009; 18:3384–3396.
- Tardat M, Brustel J, Kirsh O, Lefevbre C, Callanan M, Sardet C, Julien E. The histone H4 Lys 20 methyltransferase PR-Set7 regulates replication origins in mammalian cells. *Nature Cell Biology*. 2010; 12:1086–1093.



- Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinformatics*. 2013; 14:178–192. [PubMed: 22517427]
- Tutt A, Ashworth A. The relationship between the roles of BRCA genes in DNA repair and cancer predisposition. *Trends Mol Med*. 2002; 8:571–576. [PubMed: 12470990]
- Ursic D, Himmel KL, Gurley KA, Webb F, Culbertson MR. The yeast SEN1 gene is required for the processing of diverse RNA classes. *Nucleic Acids Research*. 1997; 25:4778–4785. [PubMed: 9365256]
- Ursic D, Chinchilla K, Finkel JS, Culbertson MR. Multiple protein/protein and protein/RNA interactions suggest roles for yeast DNA/RNA helicase Sen1p in transcription, transcription-coupled DNA repair and RNA processing. *Nucleic Acids Research*. 2004; 32:2441–2452. [PubMed: 15121901]
- Venkitaraman AR. Cancer suppression by the chromosome custodians, BRCA1 and BRCA2. *Science*. 2014; 343:1470–1475. [PubMed: 24675954]
- White E, Kamieniarz-Gdula K, Dye MJ, Proudfoot NJ. AT-rich sequence elements promote nascent transcript cleavage leading to RNA polymerase II termination. *Nucleic Acids Research*. 2013; 41:1797–1806. [PubMed: 23258704]
- Wimberly H, Shee C, Thornton PC, Sivaramakrishnan P, Rosenberg SM, Hastings PJ. R-loops and nicks initiate DNA breakage and genome instability in non-growing *Escherichia coli*. *Nat Comms*. 2013; 4:2115.
- Yarden RI, Brody LC. BRCA1 interacts with components of the histone deacetylase complex. *Proceedings of the National Academy of Sciences of the United States of America*. 1999; 96:4983–4988. [PubMed: 10220405]
- Yüce Ö, West SC. Senataxin, defective in the neurodegenerative disorder ataxia with oculomotor apraxia 2, lies at the interface of transcription and the DNA damage response. *Molecular and Cellular Biology*. 2013; 33:406–417. [PubMed: 23149945]
- Zhu Q, Pao GM, Huynh AM, Suh H, Tonnu N, Nederlof PM, Gage FH, Verma IM. BRCA1 tumour suppression occurs via heterochromatin-mediated silencing. *Nature*. 2011; 477:179–184. [PubMed: 21901007]



**Figure 1. BRCA1 interacts with SETX and is required for SETX recruitment to the R-loop-associated termination pause region of the human  $\beta$ -actin gene**

(A) Co-immunoprecipitation (IP) of endogenous BRCA1 and SETX in HeLa cell (upper panel) and in BJ-hTert human fibroblast extracts (lower panels), using antibodies against BRCA1 (BRCA1 #1/#2) or SETX (SETX #1/#2). IgG, negative control.

(B) *In vitro* interaction assay using recombinant GST-BRCA1 fragments performed in HeLa cells. GSH bead-bound proteins were eluted and analyzed by immunoblot. Bottom: Relative abundance of each recombinant affinity purified fusion protein (marked with a star) visualized by SDS-PAGE after Coomassie staining. Additional bands are read-through and/or degradation products of recombinant proteins.

(C) Same as in (B) but with recombinant GST-SETX fragments.

(D) Schematic representation of the human  $\beta$ -actin gene: exons are numbered, and the red box reflects the existence of a transcription pause element. Location of primers used for the ChIP experiments are depicted above the diagram.

(E) Immunoblot that reflects the efficiency of the siRNA-mediated depletion of BRCA1 and SETX. Top panel, SETX and BRCA1; lower panel, Vinculin used as a loading control.

(F) ChIP analyses performed on the  $\beta$ -actin gene from mock-treated (red bars), BRCA1- (green bars) or SETX-depleted cells (grey bars) using BRCA1 Ab #3. Average ChIP values  $\pm$  SD from three independent experiments are shown.

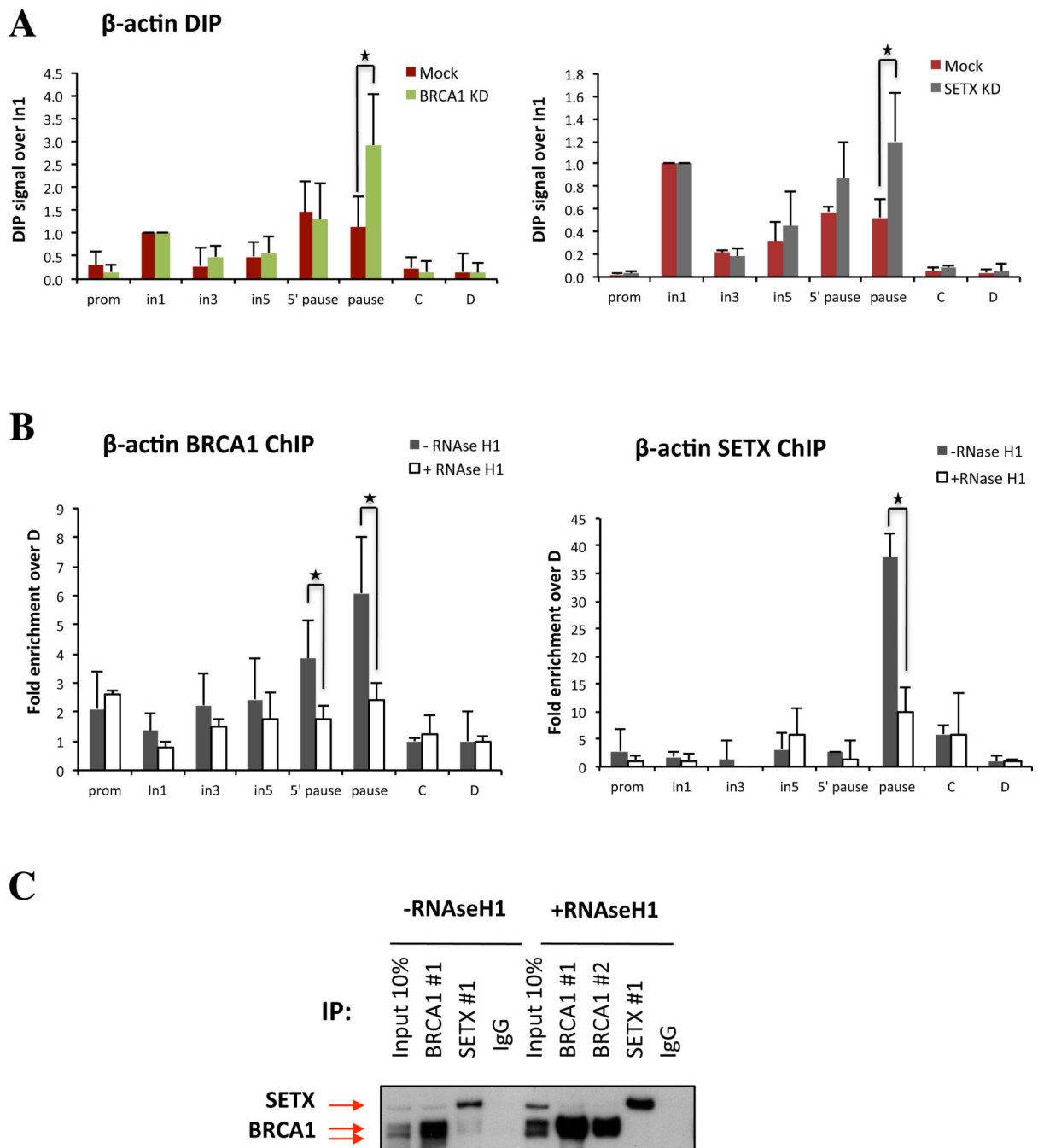
(G) Same ChIP experiments as in (F) but with SETX Ab #1.  
See also Figures S1 and S2.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



**Figure 2. BRCA1 recruitment to the  $\beta$ -actin termination pause site is mediated by R-loops**  
 (A) RNA:DNA immunoprecipitation (DIP) analyses performed in HeLa from mock-treated or BRCA1-depleted cells (BRCA1 KD) (left) and mock-treated or SETX-depleted cells (SETX KD) (right).  
 (B) BRCA1 and SETX ChIP experiment performed in control conditions (-RNaseH1) or with RNaseH1 over-expression (+RNaseH1).  
 Average DIP and ChIP values  $\pm$  SD from three, independent experiments are shown.  
 $p^* < 0.05$ .

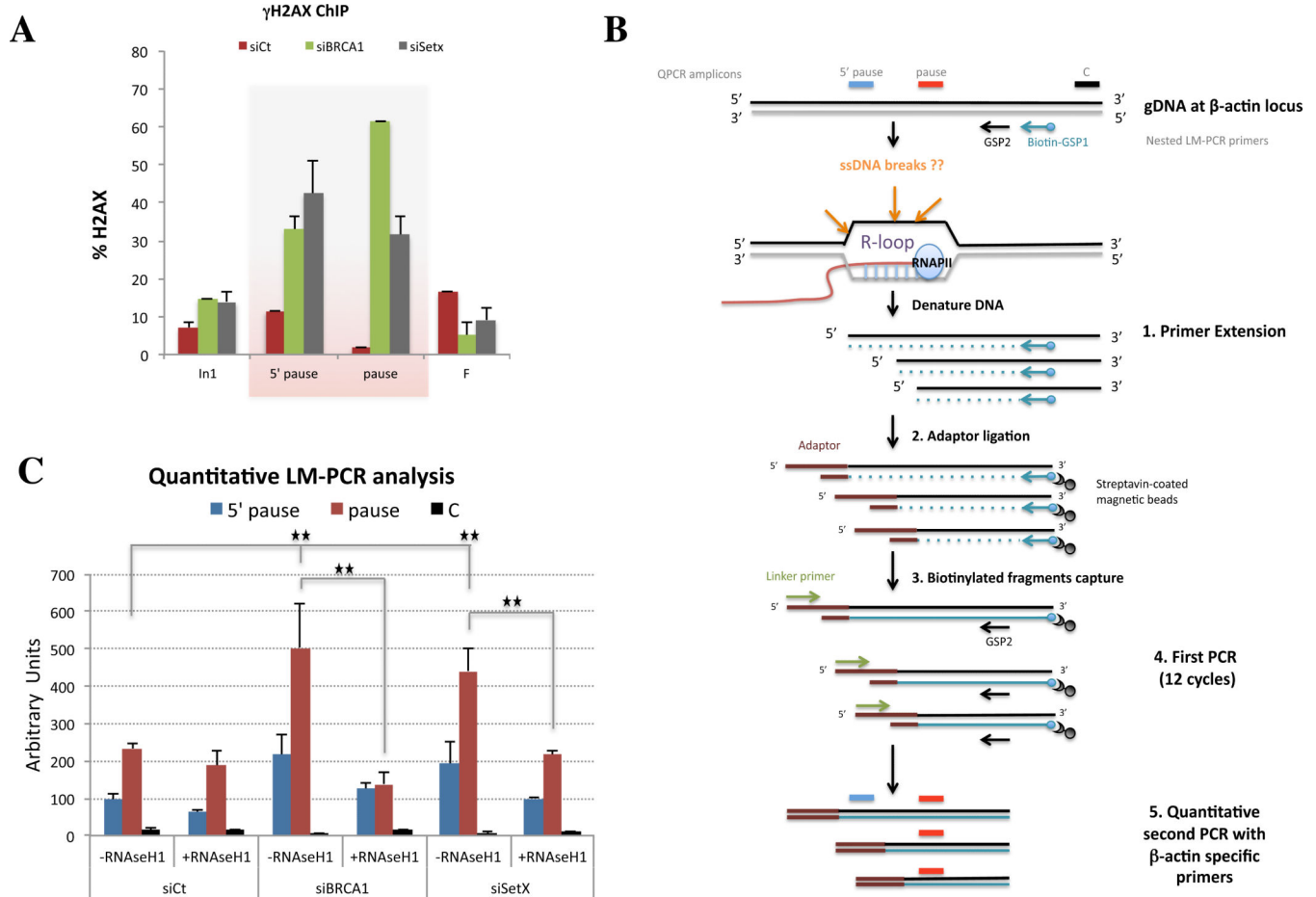
(C) BRCA1 and SETX co-IP experiments performed in HeLa cells +/- RNaseH1 expression. IgG, negative control.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



**Figure 3. BRCA1/SETX complex at the  $\beta$ -actin pause site protects cells from R-loop-driven ssDNA breaks**

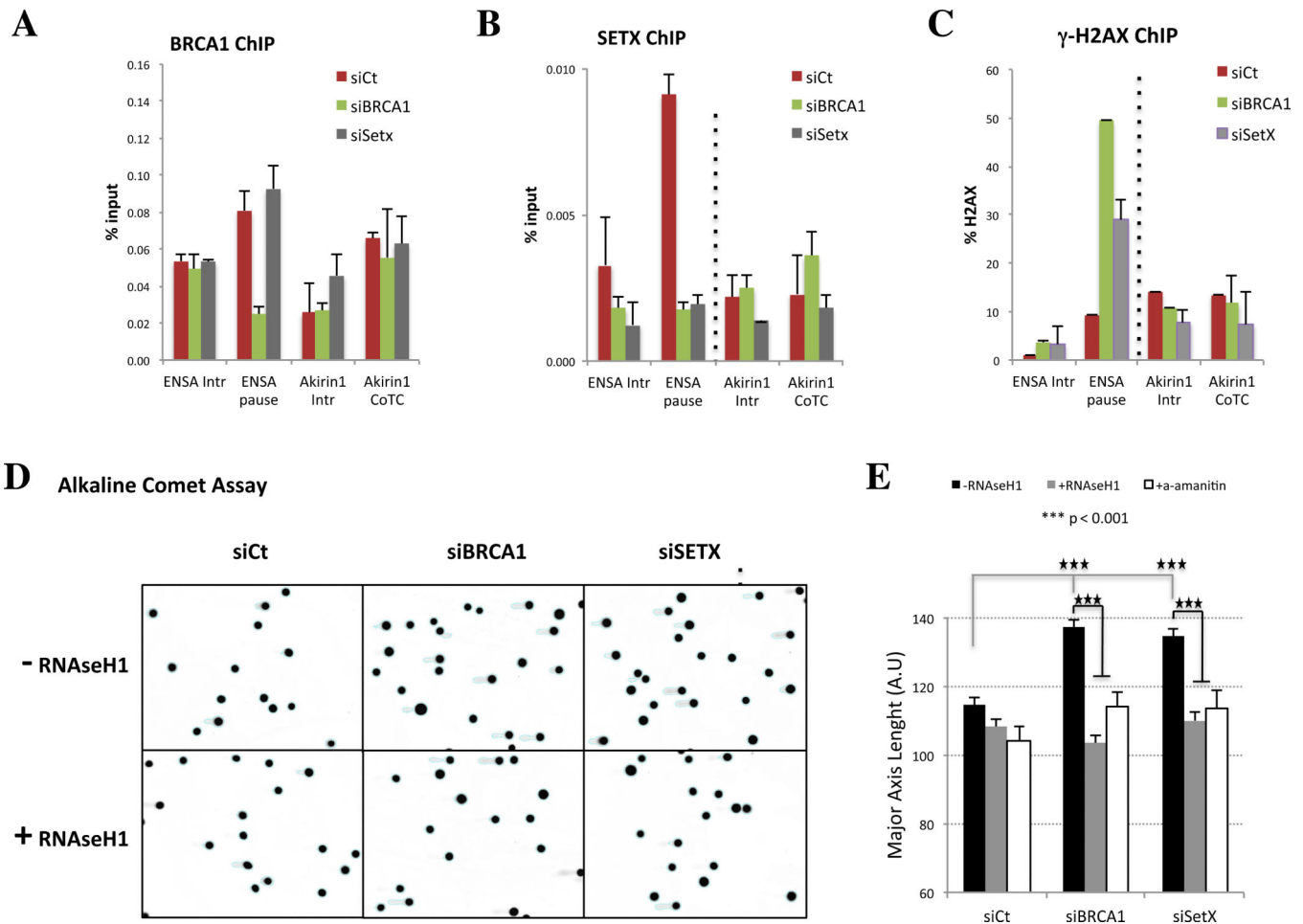
(A) ChIP analysis performed on the  $\beta$ -actin gene as in Figure 1 using  $\gamma$ H2AX and total H2AX antibodies. Histograms represent the proportion of total H2AX phosphorylated on Ser139 (i.e.  $\gamma$ H2AX). Average ChIP values  $\pm$  SD from three independent experiments are shown.

(B) LM-PCR strategy used to identify R-loop-associated ssDNA breaks on the coding strand. See Extended Experimental Procedures.

(C) Quantitative detection of ssDNA after LM-PCR performed on the  $\beta$ -actin gene before and after BRCA1 or SETX knockdown and with or without ectopic RNaseH1 expression. QPCR values are average  $\pm$  SD from three independent experiments. \* $P$ <0.05, \*\* $P$ <0.007 by one-tailed Student's  $t$ -test.

See also Figures S3.





**Figure 4. DNA damage arising in absence of BRCA1/SETX complexes at termination pause sites is R-loop dependent**

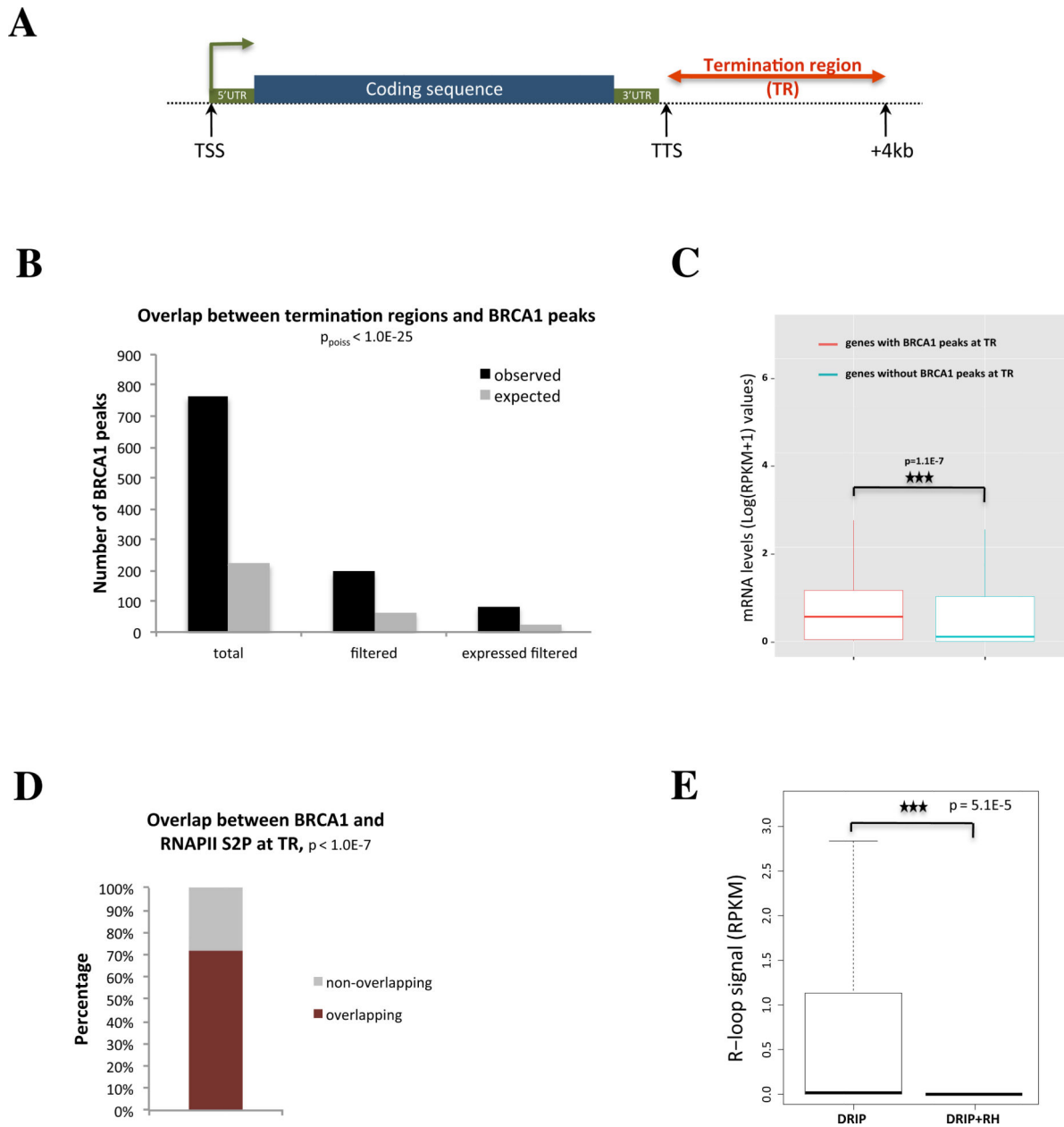
(A and B) BRCA1 (A) and SETX (B) ChIP analyses performed on the ENSA and Akirin1 genes in HeLa cells transfected with siCt, siBRCA1 or siSETX. ENSA and Akirin1 transcription termination is regulated by R-loops and CoTC sequences, respectively. Intronic regions (Intr) were studied as controls. Average ChIP values  $\pm$  SD from three independent experiments are shown.

(C)  $\gamma$ H2AX ChIP experiments performed as in (A) and analyzed as in Figure 3A. Average ChIP values  $\pm$  SD from three independent experiments are shown.

(D) Representative pictures of comet assays performed under alkaline conditions in HeLa cells transfected with siCt, siBRCA1 or siSETX in the absence (-RNaseH1) and presence (+RNaseH1) of ectopic RNaseH1 expression.

(E) Quantitative analysis of comet tail lengths for each condition showed in (D). Average tail lengths  $\pm$  SEM from three independent experiments are shown. \*\*\* $P < 0.001$  by two-tailed Student's *t*-test.

See also Figure S4.



**Figure 5. BRCA1 binds the transcription termination region of a substantial subset of actively expressed mammalian genes**

(A) Diagram of candidate BRCA1 TR binding regions. Putative termination regions (TR) were defined as segments extending from the TTS to TTS+4kb.

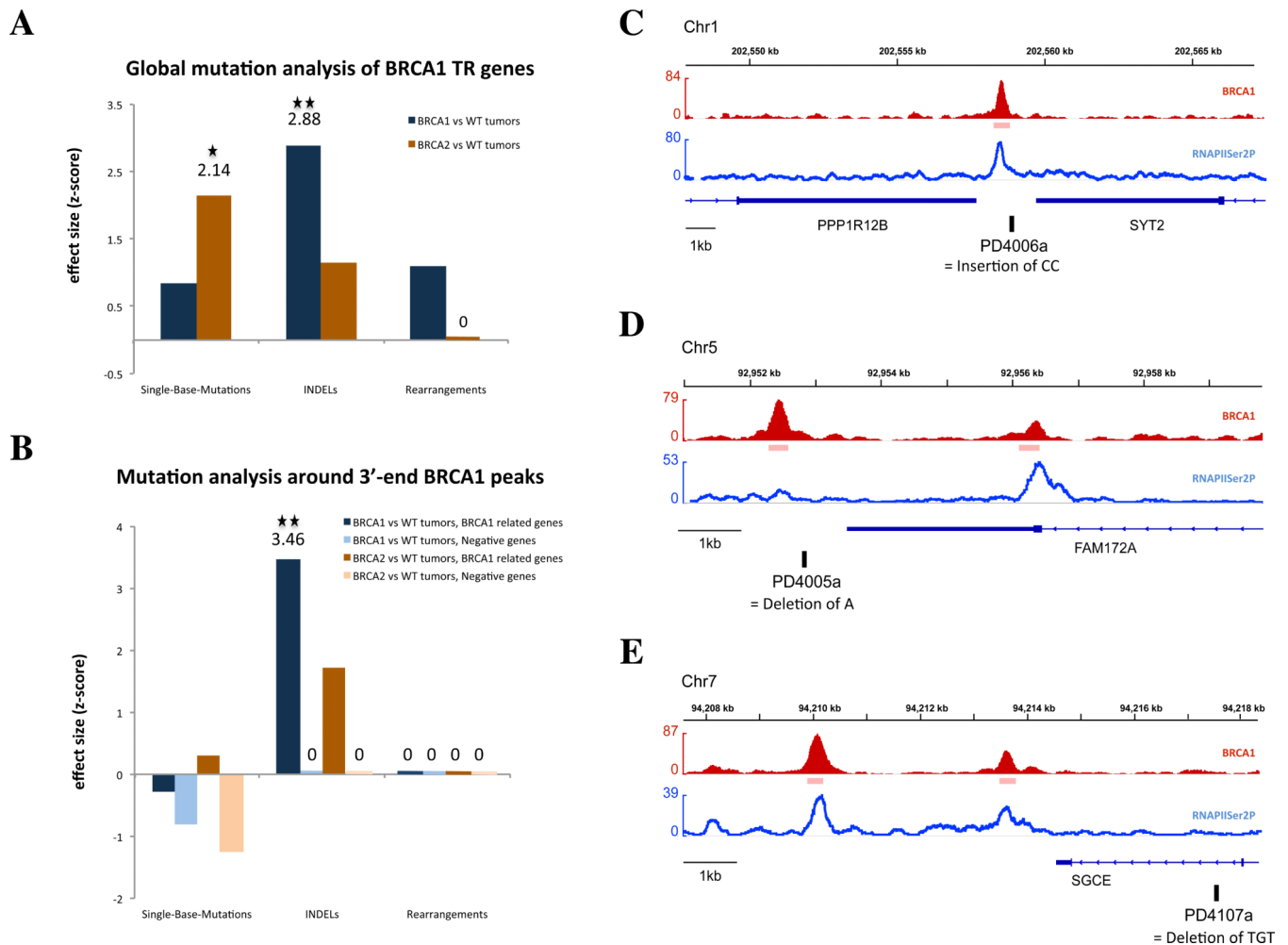
(B) Total number of observed and expected overlaps between TR and BRCA1 peaks ( $p_{\text{poiss}}=2.2E-177$ ). “Filtered”: BRCA1 ChIP-seq peaks divested of those overlapping promoter regions and transcripts ( $p_{\text{poiss}}=1.2E-42$ ). “Expressed filtered”: BRCA1 peaks present in TR region of expressed genes ( $p_{\text{poiss}}=5E-24$ ).

(C) Gene expression comparison between genes with (red) and without (blue) BRCA1 bound to relevant TR. Boxplots reflect the median [50<sup>th</sup> percentile] of mRNA expression. \*\*\* $p=1.1E-7$ , Mann-Whitney test. Outliers have been omitted from the plot.

(D) Overlap between BRCA1 TR and RNAPIISer2P peaks, overlap enrichment over random (\*\* $p < 1.0E-7$ ), see Methods.

(E) Boxplots showing DRIP-seq signals (RPKM) of DRIP samples compared with DRIP +RH controls (treated with RNaseH1) in BRCA1 TR candidate regions, \*\* $p = 5.101E-5$ , paired Wilcoxon test.

See also Figures S5 and Table S1.



**Figure 6. BRCA1 deficient breast cancers reveal genomic abnormalities at and near BRCA1-associated termination sites**

(A and B) Global mutational analysis carried out in the 184 BRCA1 TR genes using the complete whole-genome catalog of somatic mutations from 21 breast cancers (Nik-Zainal et al., 2012). Effect size comparison (one tailed CMH z-score) between the different tumor subgroups when testing the region from TSS -1250bp to TTS + 5kb (A) or  $\pm 4$ kb from TTS (B). WT tumors=non BRCA1/BRCA2 and negative genes=CoTC genes. Statistical significant: \* $p < 0.05$  and \*\* $p < 0.01$ .

(C-E) ChIP-seq profiles of BRCA1 (red) and RNAPIISer2P (blue) in BRCA1 TR genes and location of indels (black boxes). Chr, chromosome.

See also Figure S6 and Table S2.