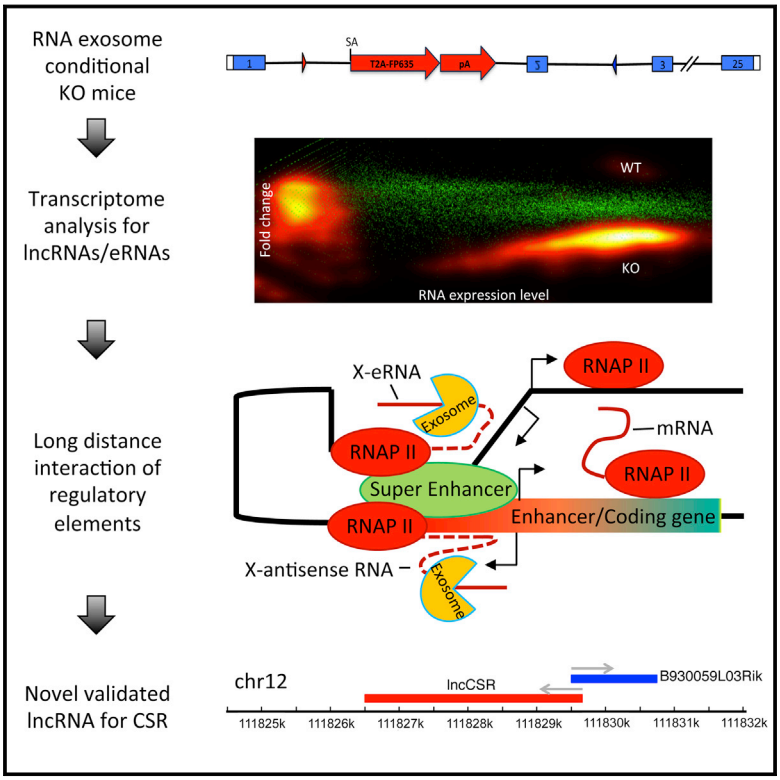
Article

Cell

RNA Exosome-Regulated Long Non-Coding RNA Transcription Controls Super-Enhancer Activity

Graphical Abstract



Highlights

- Identification of new lncRNAs and eRNAs from RNA exosome mutant transcriptomes
- Transcribed enhancer sequences are protected from genomic instability by RNA exosome
- Antisense RNAs substrates of the RNA exosome regulate super-enhancer activity
- eRNA-expressing element (lncRNA-CSR) promotes recombination in the IgH locus

Authors

Evangelos Pefanis, Jiguang Wang, ..., Raul Rabadan, Uttiya Basu

Correspondence

rr2579@cumc.columbia.edu (R.R.), ub2121@cumc.columbia.edu (U.B.)

In Brief

The RNA exosome regulates the expression of non-coding RNAs originating from enhancer regions, helping to coordinate their function with the activity of neighboring super-enhancers.

Pefanis et al., 2015, Cell *161*, 774–789
May 7, 2015 ©2015 Elsevier Inc.
http://dx.doi.org/10.1016/j.cell.2015.04.034

Cell

Article

RNA Exosome-Regulated Long Non-Coding RNA Transcription Controls Super-Enhancer Activity

Evangelos Pefanis,[1,4,6] Jiguang Wang,[2,6] Gerson Rothschild,[1,6] Junghyun Lim,[1,6] David Kazadi,[1] Jianbo Sun,[1] Alexander Federation,[5] Jaime Chao,[1] Oliver Elliott,[2] Zhi-Ping Liu,[3] Aris N. Economides,[4] James E. Bradner,[5] Raul Rabadan,[2,*] and Uttiya Basu[1,*]

[1]Department of Microbiology and Immunology

[2]Department of Biomedical Informatics and Department of Systems Biology

College of Physicians and Surgeons, Columbia University, New York, NY 10032, USA

[3]Department of Biomedical Engineering, School of Control Science and Engineering, Shandong University, Jinan, Shandong 250061, China

[4]Regeneron Pharmaceuticals and Regeneron Genetics Center, Tarrytown, NY 10591, USA

[5]Dana-Farber Cancer Institute and Harvard Medical School, Boston, MA 02115, USA

[6]Co-first author

*Correspondence: rr2579@cumc.columbia.edu (R.R.), ub2121@cumc.columbia.edu (U.B.)

http://dx.doi.org/10.1016/j.cell.2015.04.034

SUMMARY

We have ablated the cellular RNA degradation machinery in differentiated B cells and pluripotent embryonic stem cells (ESCs) by conditional mutagenesis of core (*Exosc3*) and nuclear RNase (*Exosc10*) components of RNA exosome and identified a vast number of long non-coding RNAs (lncRNAs) and enhancer RNAs (eRNAs) with emergent functionality. Unexpectedly, eRNA-expressing regions accumulate R-loop structures upon RNA exosome ablation, thus demonstrating the role of RNA exosome in resolving deleterious DNA/RNA hybrids arising from active enhancers. We have uncovered a distal divergent eRNA-expressing element (lncRNA-CSR) engaged in long-range DNA interactions and regulating *IgH* 3′ regulatory region super-enhancer function. CRISPR-Cas9-mediated ablation of lncRNA-CSR transcription decreases its chromosomal looping-mediated association with the *IgH* 3′ regulatory region super-enhancer and leads to decreased class switch recombination efficiency. We propose that the RNA exosome protects divergently transcribed lncRNA expressing enhancers by resolving deleterious transcription-coupled secondary DNA structures, while also regulating long-range super-enhancer chromosomal interactions important for cellular function.

INTRODUCTION

Recent advances in RNA biology have revealed a plethora of non-coding RNA transcripts whose identity and functions were previously unknown. It has been postulated that transcription control of coding genes is modulated by non-coding RNAs such as enhancer RNAs (eRNAs) (Kim et al., 2010) and long intergenic non-coding RNAs (lincRNAs) (Rinn and Chang, 2012). Of note, a significant number of non-coding RNAs are characterized as being expressed from regions proximal to the transcription start sites (TSSs) of coding genes. These transcripts include promoter-associated long RNAs (PALRs, >200 bp and bidirectional) (Kapranov et al., 2007), promoter-associated short RNAs (PASRs, 20–100 nt) (Kapranov et al., 2007), TSS-associated RNA (TSS-aRNA, small and divergently transcribed RNA) (Core et al., 2008; Seila et al., 2008), and transcription initiation RNAs (tiRNAs, 18 nt long and located 20 nt downstream of the coding TSS) (Taft et al., 2009). In addition, a large fraction of TSS-proximal transcriptional expenditure is dedicated to the production of unstable non-coding RNAs that are subject to RNA exosome-mediated degradation (PROMPTs, uaRNAs, xTSS-RNAs) (Flynn et al., 2011; Pefanis et al., 2014; Preker et al., 2008). Although the characteristics of these new RNA species may overlap, it is abundantly clear that these non-coding RNAs function in the regulation of transcription initiation and transcription elongation by various mechanisms, including control of RNA polII pausing and recruitment of chromatin modification factors (Flynn and Chang, 2012; Reyes-Turcu and Grewal, 2012; Shin et al., 2013).

Recently, some of these ncRNAs have been shown to be substrates of the RNA surveillance complex, RNA exosome (Andersson et al., 2014a, 2014b; Pefanis et al., 2014; Wan et al., 2012). The eukaryotic RNA exosome complex functions in both the nucleus and the cytoplasm. Nuclear exosome is involved in 3′-5′ processing of rRNAs, sn/snoRNAs, degradation of hypomodified tRNAs, and cryptic unstable transcripts (CUTs), whereas cytoplasmic exosome is responsible for the degradation of aberrant mRNA species subject to nonsense mediated decay, non-stop decay, or no-go decay (Schmid and Jensen, 2008; Chlebowski et al., 2013). The eukaryotic exosome complex is composed of a nine subunit core, consisting of six distinct proteins forming a "ring" and three distinct RNA-binding-domain-containing proteins forming a "cap" structure required for the stabilization of the core structure. Enzymatic activity of the exosome complex is provided through two additional subunits: Rrp44 (*Dis3*) and Rrp6 (*Exosc10*) (Houseley et al., 2006; Januszyk and Lima, 2011; Liu et al., 2006; Lorentzen et al., 2008). Rrp6 is a nuclear-specific 3′-5′ distributive exoribonuclease (Lykke-Andersen et al., 2009). Although in vitro Rrp6 and Dis3

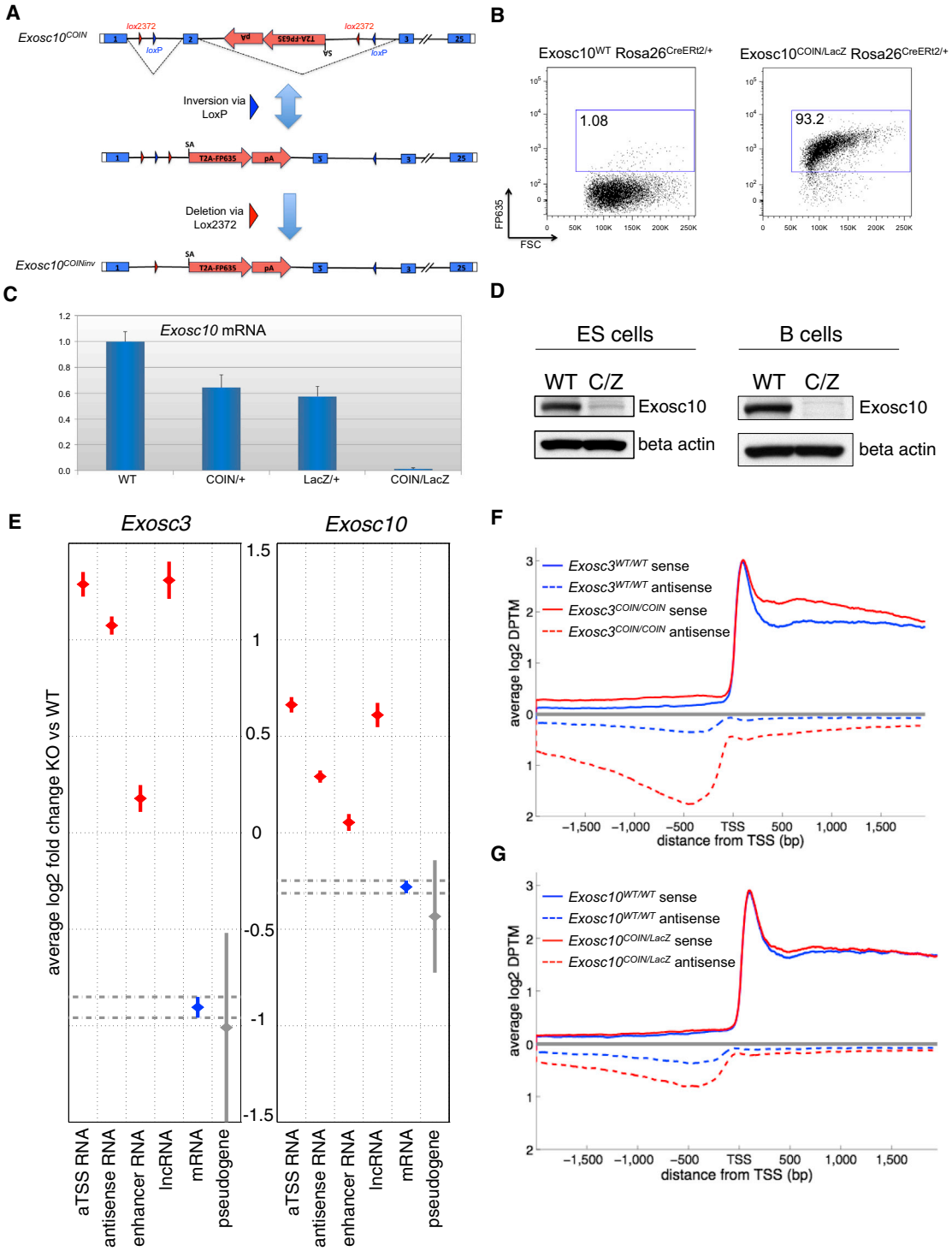


Figure 1. Generation of RNA Exosome Mutant ESCs and Transcriptome Analysis

(A) *Exosc10*^{COIN} allele and conversion to *Exosc10*^{COINinv}. Cre-mediated inversion of loxP pair (blue triangles) and subsequent deletion via lox2372 pair (red triangles). FP635-expressing terminal exon represented by red arrow. SA, splice acceptor.

(B) Induction of fluorescent reporter FP635 in *Exosc10*$^{COIN/LacZ}$ *ROSA26*$^{CreERt2/+}$ B cells following 4-OHT treatment.

(C) qRT-PCR analysis of *Exosc10* mRNA expression in 4-OHT-treated, LPS+IL-4-stimulated B cells. Indicated *Exosc10* genotypes on *ROSA26*$^{CreERt2/+}$ background. Expression levels normalized to cyclophilin A (*Ppia*) and plotted relative to *Exosc10*$^{WT/WT}$. Splenic B cells were isolated and treated with 4-OHT for 24 hr, and the cells were then washed. Total cellular RNA was isolated after 72 hr of B cell culture. Three technical replicates; error bars represent SD.

(legend continued on next page)

bind the RNA exosome core (Exo9) independent of each other, Exo9 may interconnect the properties of the two RNase subunits in vivo (Schaeffer et al., 2009; Schaeffer and van Hoof, 2011; Wasmuth and Lima, 2012) so that different types of RNA substrates can be processed/degraded. Crystal structure analysis of an Rrp6-containing yeast RNA exosome complex suggests that Rrp6 may function in regulating the size of the central channel through which RNA traverses prior to degradation (Wasmuth et al., 2014). The true nature of Rrp6 function within the RNA exosome complex, via its distributive RNase activity and/or its contribution to central channel regulation, is incompletely understood. Moreover, mammalian RNA substrates of the RNA exosome complex with or without the Rrp6 component have not been systematically identified. The activity of the RNA exosome in co-transcriptionally degrading RNA plays a critical function in the nucleus, with recent observations in yeast and mammalian cells indicating a role for RNA degradation in early transcription termination (Colin et al., 2014; Hazelbaker et al., 2013; Lemay et al., 2014; Pefanis et al., 2014; Richard and Manley, 2009; Shah et al., 2014; Storb, 2014; Sun et al., 2013b). As such, the role of RNA exosome in chromatin-associated events is a major focus of ongoing research.

In this study, we reveal and analyze the transcriptomes of *Exosc3*- and *Exosc10*-ablated embryonic stem cells (ESCs) and B cells and identify a vast number of non-coding RNAs with emergent biological functionality. Strikingly, we find that the RNA exosome regulates the levels of divergently transcribed enhancer RNAs by promoting co-transcriptional silencing, thereby preventing the persistence of detrimental chromatin structures that can lead to genomic instability. Moreover, we provide evidence that RNA exosome substrate divergently transcribed loci may regulate interactions with super-enhancer loci. Thus, our study provides a mode of long-range chromatin regulation not previously described. As an example, we have identified the long non-coding RNA (lncRNA)-CSR-expressing locus and report its regulation of immunoglobulin heavy-chain DNA rearrangements by functionally interacting with the 3′ regulatory region super-enhancer sequence (3′RR).

RESULTS

RNA Exosome Mutant ESCs and Mouse Models

To ascertain the role of the RNA exosome complex in the degradation of non-coding RNAs, we have generated mouse conditional alleles of *Exosc10* (expressing the distributive nuclease subunit Rrp6) (Figures S1A and S1B) and *Exosc3* (expressing the RNA exosome core subunit Rrp40) (Pefanis et al., 2014). Using these two approaches, inducible RNA exosome deficiency was evaluated in either primary pluripotent embryonic stem cells or differentiated mature B cells. *Exosc10* and *Exosc3* allele schemes utilize Cre/lox conditional inversion (COIN) methodology to ablate normal gene expression upon exposure of the alleles to Cre recombinase activity (Economides et al., 2013; Pefanis et al., 2014). The salient feature of this approach, as utilized here, is the inversion of one or more endogenous coding exons resulting in the simultaneous "activation" of a fluorescent reporter terminal exon within the same locus (Figure 1A). $Exosc10^{COIN/WT}$ mice were crossed with mice heterozygous for a null allele of *Exosc10* ($Exosc10^{LacZ/WT}$) to derive ESCs and B cells of the genotype $Exosc10^{COIN/LacZ}$. Similarly, we have generated $Exosc3^{COIN/COIN}$ ESCs and B cells (Pefanis et al., 2014). Both $Exosc10^{COIN/LacZ}$ and $Exosc3^{COIN/COIN}$ cells also contain the inducible $ROSA26^{CreERt2}$ allele allowing for rapid ablation of RNA exosome activity upon tamoxifen treatment. When B cells from $Exosc10^{COIN/LacZ}$ mice were treated with 4-hydroxytamoxifen (4-OHT) ex vivo, inversion of the $Exosc10^{COIN}$ allele was observed in more than 90% of the cells (Figure 1B). qRT-PCR assays performed on total cellular RNA demonstrated nearly complete loss of *Exosc10* mRNA in 4-OHT-treated $Exosc10^{COIN/LacZ}$ B cells (Figure 1C). Western blotting of protein extracts from $Exosc10^{COIN/LacZ}$ B cells and ESCs demonstrated severe loss of Rrp6 protein following 4-OHT, indicating robust ablation of *Exosc10* expression (Figure 1D). The RNA exosome previously has been implicated in catalyzing class switch recombination (CSR) in B cells by supporting the activity of activation-induced cytidine deaminase (AID) (Basu et al., 2011). Consistent with these observations, *Exosc10*-deficient B cells display reduced CSR efficiency as compared to wild-type (WT) littermate control B cells (Figure S1C) despite comparable expression of AID (Figure S1D). Finally, RNA-seq analysis of $Exosc10^{COIN/LACZ}$ B cells and ESCs confirmed loss of *Exosc10* transcripts in both cell types (Figure S1E). Similarly, and consistent with previously published characterization of *Exosc3* ablation in $Exosc3^{COIN/COIN}$ B cells, RNA-seq analysis demonstrated a clear loss of *Exosc3* transcripts in both $Exosc3^{COIN/COIN}$ B cells and ESCs (Figure S1F).

Transcriptome of RNA Exosome Mutant ESCs and B Cells

We assembled the transcriptomes of littermate pairs of WT control and $Exosc10^{COIN/LacZ}$ or $Exosc3^{COIN/COIN}$ B cells and ESCs using next-generation RNA sequencing technology. The bioinformatics pipeline used for transcriptome reconstitution is outlined in Figure S2A and is described in further detail in the Extended Experimental Procedures. We find that, in the exotomes (exosome-deficient transcriptome) of $Exosc3^{COIN/COIN}$

(D) Western blot detection of Rrp6 (*Exosc10*) from $Exosc10^{WT/WT}$ (WT) and $Exosc10^{COIN/LacZ}$ (C/Z) protein extracts obtained from 4-OHT-treated ESCs and B cells. Indicated *Exosc10* genotypes on $ROSA26^{CreERt2/+}$ background.

(E) Genome-wide differential expression level analysis of RNA subsets in *Exosc3* (left) and *Exosc10* (right) ablated mouse ESCs relative to WT littermate-matched ESCs. The error bars represent confidence interval of mean value estimated by an improved version of the Tukey-Kramer method (see Extended Experimental Procedures).

(F and G) Genome-wide TSS proximal expression profile in *Exosc3* (F) and *Exosc10* (G) ablated mouse ESCs. Sense and antisense transcript levels 2 kb flanking the TSS of annotated coding transcripts are indicated. ESCs were treated with 4-OHT for 24 hr and then further cultured for an additional 48 hr before total RNA isolation.

See also Figure S1 and Tables S1, S2, S3, and S4.

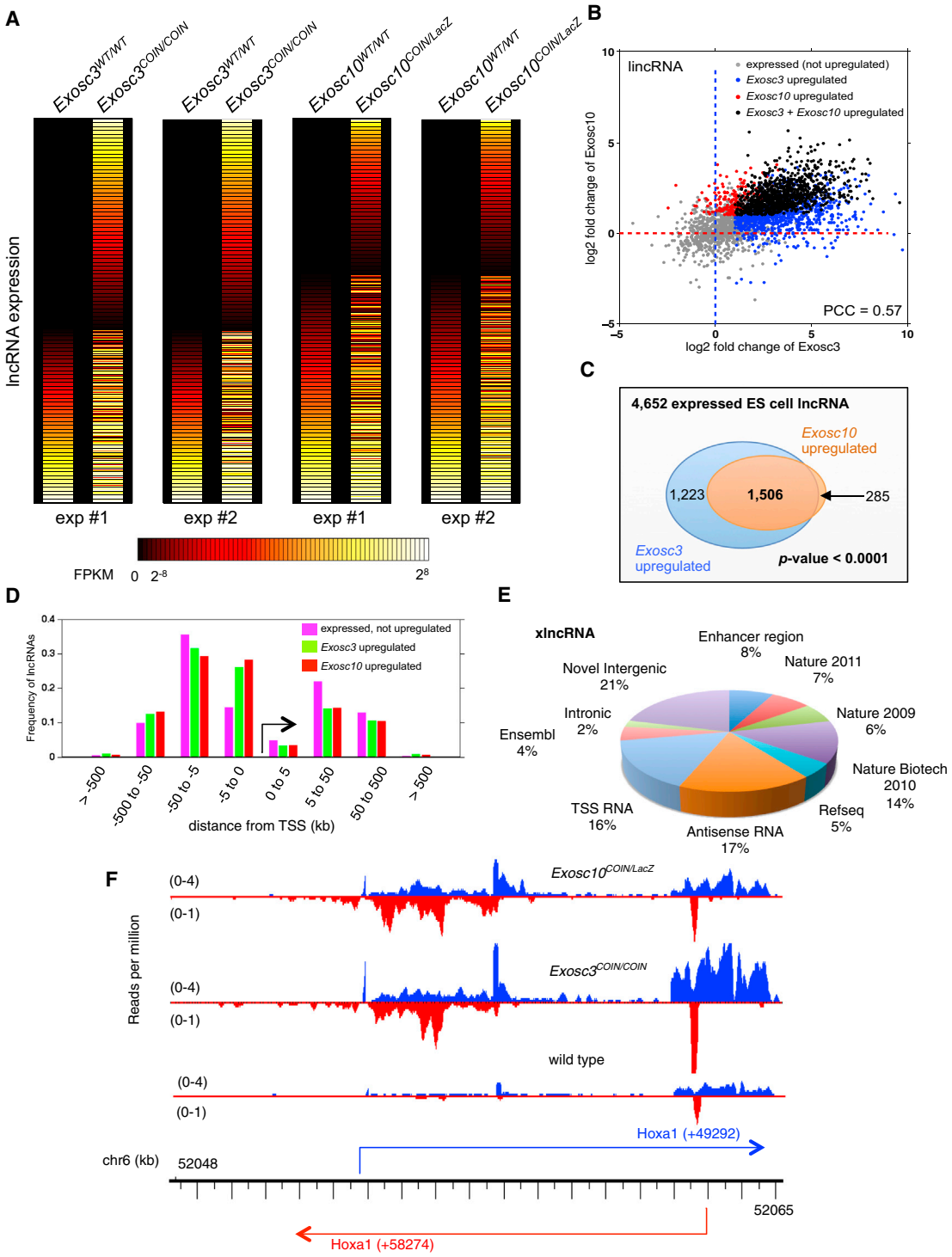


Figure 2. Identification and Characterization of RNA Exosome Targeted lncRNAs in ESCs

(A) Heatmap of lncRNAs expressed in *Exosc3*$^{WT/WT}$/*Exosc3*$^{COIN/COIN}$ and *Exosc10*$^{WT/WT}$/*Exosc10*$^{COIN/LacZ}$ genotype pairs. Horizontal lines represent different lncRNAs, which were ranked by their expression level in matched WT controls.

(B) Distribution of lncRNAs stabilized in the *Exosc3*-exotome (blue), *Exosc10*-exotome (red), and both *Exosc3* and *Exosc10* exotomes (black).

(C) Venn diagram demonstrating the distribution of overlapping lncRNAs in *Exosc3* (blue) and *Exosc10* (orange) exotomes.

(D) Distribution of x-lncRNA TSS distances from closest neighboring coding gene TSS genome wide.

(legend continued on next page)

(Figure 1E, left) and *Exosc10*$^{COIN/LacZ}$ ESCs (Figure 1E, right), relative levels of lncRNAs, antisense RNAs, and eRNAs are significantly increased genome wide compared to WT control ESC transcriptomes. Comparing relative transcript accumulations of lncRNAs, antisense RNAs, and eRNAs indicates that these non-coding RNA subsets experience greater stabilization within the *Exosc3*$^{COIN/COIN}$ exotome in comparison to the *Exosc10*$^{COIN/LacZ}$ exotome genome wide. TSS antisense divergent RNAs are well-known substrates of the RNA exosome complex (Pefanis et al., 2014; Preker et al., 2008; Seila et al., 2008; Seila et al., 2009). Consistent with expectations, TSS-associated antisense RNAs are markedly stabilized within the *Exosc3*$^{COIN/COIN}$ ESC transcriptome (Figure 1F). A list of antisense RNA in the body of the genes and around the genic TSS from B cell exotome and ESC exotome are provided in Tables S1 and S2, respectively. Relative to *Exosc3*-deficient cells, TSS-associated antisense transcripts are moderately stabilized within the *Exosc10*$^{COIN/LacZ}$ ESC transcriptome (Figure 1G). Collectively, these results point toward a role for *Exosc10* in the degradation of a subset of RNA exosome-targeted lncRNAs (presumably fully represented via *Exosc3* ablation).

RNA Exosome Substrate Long Non-Coding RNA

Previously, it has been shown that enhancers express bidirectional, divergently transcribed, RNA exosome-sensitive, capped non-coding RNAs in human cell lines and primary mouse B cells (Andersson et al., 2014a, 2014b; Pefanis et al., 2014; Wan et al., 2012). Taking clues from these studies, we evaluated whether our RNA exosome mutant mouse models could be utilized for identifying eRNAs in pluripotent ESCs or lineage-committed matured B cells. Following the analysis pipeline described in the Extended Experimental Procedures, we observed that a subset of lncRNAs were strong substrates of RNA exosome. We describe such transcripts here as exosome substrate lncRNA (x-lncRNA). As shown via heatmap representation, both in *Exosc3*$^{WT/WT}$/*Exosc3*$^{COIN/COIN}$ and in *Exosc10*$^{WT/WT}$/*Exosc10*$^{COIN/LacZ}$ RNA-seq analysis pairs, multiple x-lncRNA loci are revealed in RNA exosome-deficient ESCs while weakly expressed in counterpart WT control cells (Figure 2A) (details of expression and genome coordinates of these transcripts supplied in Table S3). Next, we performed comparative expression analysis between *Exosc3* and *Exosc10* substrate x-lncRNAs and found that a significant number of, although not all, *Exosc3* x-lncRNAs also classify as *Exosc10* x-lncRNAs (Figure 2B). Specifically, of a total of 2,729 *Exosc3* x-lncRNAs in ESCs, 1,506 also fell within the cutoff for *Exosc10* x-lncRNAs (Figures 2C, S2B, and S2C; details in Table S3). Surprisingly, only 59% of *Exosc3* x-lncRNAs described here have been reported previously (Figures 2E and S2D). In fact, 236 of these identified x-lncRNAs are positioned close to enhancer sequences and thus may serve as RNA exosome target "x-eRNAs." Moreover, the accumulation of x-lncRNAs mostly maps within 5–50 kb from the TSS of known coding genes, making it possible that these lncRNAs regulate gene expression of distal genes via long-range chromatin interactions (Figure 2D). As indicated earlier, there are substantial numbers of lncRNAs that are quite unstably expressed in WT steady-state ESCs, but their identity cannot be confidently evaluated due to weak detection. However, RNA-seq analysis of *Exosc3*$^{COIN/COIN}$ and/or *Exosc10*$^{COIN/LacZ}$ cells provides a methodology for the detection and characterization of highly unstable lncRNA species. One such example is provided as the sense/antisense x-lncRNAs in the *Hoxa1* locus (Figure 2F). There are multiple species of antisense x-lncRNAs that are expressed in the *Hoxa1* locus (Figure 2F), whose detection is amplified in the *Exosc3*$^{COIN/COIN}$ or *Exosc10*$^{COIN/LacZ}$ exotomes.

RNA Exosome Substrate Enhancer RNA

Some enhancer RNAs (x-eRNAs) are predicted to form a subset of x-lncRNAs. Thus, we analyzed eRNA stability and identity in both *Exosc3* and *Exosc10* exotomes and found overlapping, as well as distinct, requirements for these two RNA exosome subunits (Figure 3A). All eRNAs that could be identified from ESCs are listed in Table S4. Of a total of 891 *Exosc3* x-eRNAs in ESCs, a subset of 423 displayed a significant enrichment with *Exosc10* loss (Figure 3B). In addition, 86% of the *Exosc3* x-eRNAs reported here are previously unrecognized. Of the 37 *Exosc3* x-eRNAs previously reported in VISTA, a subset of 18 was upregulated following *Exosc10* depletion (data not shown). In B cell exotomes, the degree of overlap between *Exosc3* and *Exosc10* x-eRNAs is reduced in comparison to ESC exotomes (Figure 3C). Of the 870 identified B cell *Exosc3* x-eRNAs, only 62 were *Exosc10* targets (Figure 3D). Representative *Exosc3* x-eRNAs within the *Cd83* locus were significantly upregulated in *Exosc3*$^{COIN/COIN}$ B cells and modestly increased in *Exosc10*$^{COIN/LacZ}$ B cells (Figure 3E).

x-lncRNA (or x-eRNA) expression is detectable in WT cells although significantly stabilized in *Exosc3*$^{COIN/COIN}$ cells (Figure 3F). Moreover, in both B cells (Figure 3G) and ESCs (Figure 3H), the degree of conservation for x-lncRNAs genome wide is greater than a random control set of sequences, albeit lower in amount than protein-coding DNA sequences in the mouse genome. To determine the conservation of lncRNAs that we have identified in this study, we compared x-lncRNAs with human genes (genome version hg19) using the LiftOver tool (https://genome.ucsc.edu/cgi-bin/hgLiftOver). The percentage of genes that are conserved between human and mouse is shown distributed with different cutoffs. In Figures 3G and 3H, equivalent numbers of coding genes/random genomic regions with similar length were generated as controls. For each group of genes, the percentage that is conserved between human and mouse (y axis) is calculated based on UCSC LiftOver tool

(E) The pie chart represents distribution of previously reported and newly identified lncRNAs from this study (Guttman et al., 2009, 2010, 2011). Each category represents RNAs unique to that category and non-overlapping with previous categories, with the initial category designated as "enhancer region lncRNAs" and proceeding clockwise.

(F) Expression of sense and antisense lncRNA expression profile at the *Hoxa1* locus identified from *Exosc3* and *Exosc10*-ablated ESCs. Based on RNA-seq read distribution, multiple lncRNAs are expressed in the sense and antisense directions.

See also Figure S2 and Table S3.

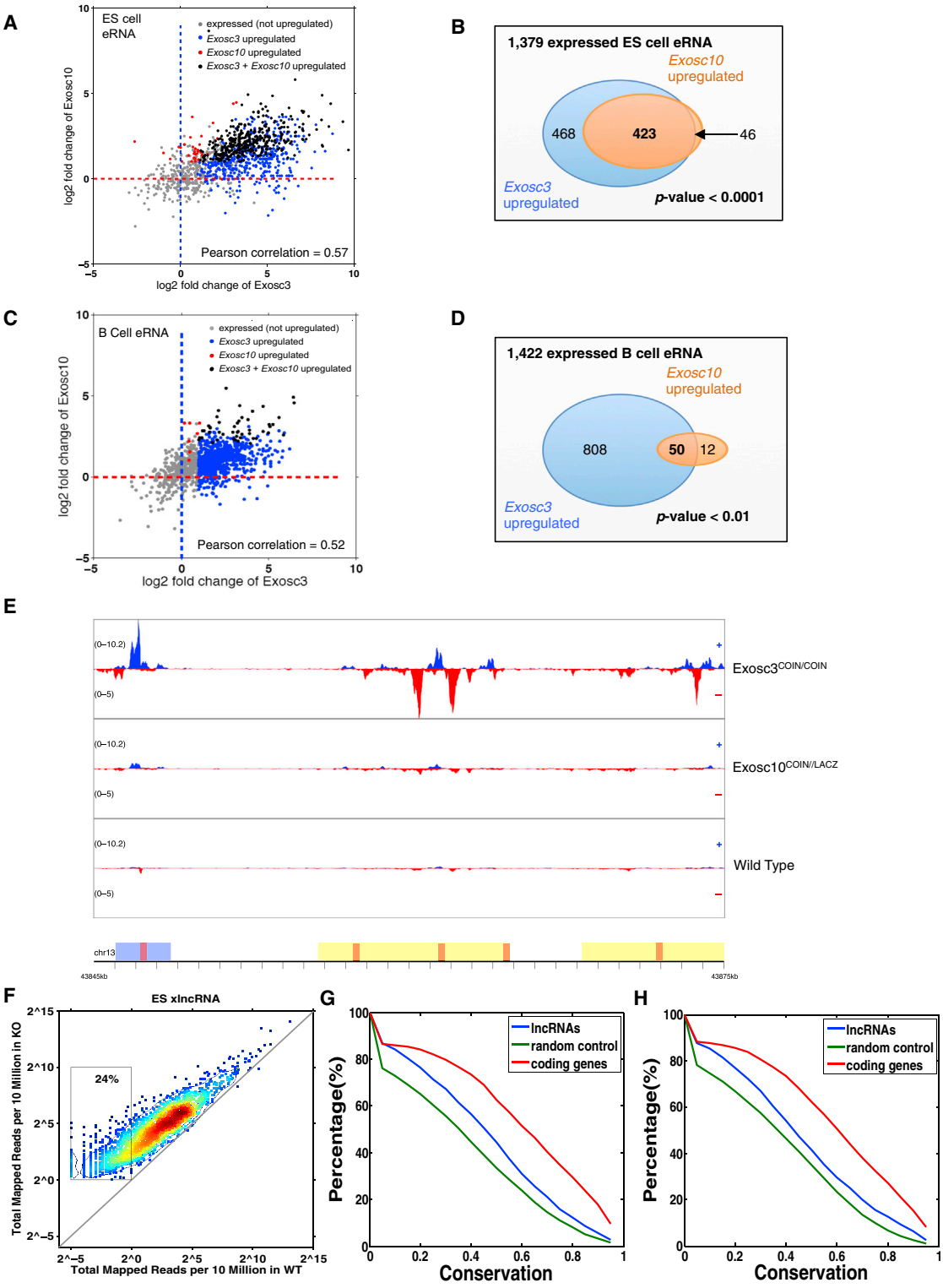


Figure 3. A Subset of Enhancer RNAs, x-eRNAs, Is an RNA Exosome Target in ESCs and B Cells

(A) Distribution of upregulated ESC x-eRNAs from *Exosc3* and *Exosc10* exotomes. Pearson correlation is indicated.

(B) Overlap of identified x-eRNAs stabilized from degradation in ESC *Exosc3* (blue circle) and *Exosc10* (orange circle) exotomes.

(C) Distribution of upregulated B cell x-eRNAs from *Exosc3* and *Exosc10* exotomes. Pearson correlation is indicated.

(D) Overlap of identified x-eRNAs stabilized from degradation in B cell *Exosc3* (blue circle) and *Exosc10* (orange circle) exotomes.

(legend continued on next page)

with given cutoff (x axis) (details in Extended Experimental Procedures). Taking these observations into account, it is likely that many x-lncRNAs (and their subset x-eRNAs) are biologically functional. The dependency of the RNA exosome complex on Rrp6 (*Exosc10*) to degrade various subsets of ncRNAs may vary based on the type of ncRNA and/or the cell type. For example, xTSS-RNAs (one type of antisense RNA) in B cells (Figure S3A) or in ESCs (Figure S3B) have markedly increased representation in *Exosc3* exotomes in comparison to *Exosc10* exotomes. In contrast, antisense RNA levels arising from gene bodies were similar between *Exosc3* and *Exosc10* B cell (Figure S3C) and ESC exotomes (Figure S3D). Finally, to ascertain whether any major pathway was affected in the cells following RNA exosome activity depletion at the time points of RNA extraction, we performed gene set enrichment analysis (GSEA) in $Exosc3^{WT/WT}$ and $Exosc3^{COIN/COIN}$ ESCs. As would be expected, there were some perturbations in gene expression profiles in $Exosc3^{COIN/COIN}$ ESCs, specifically gene sets related to organic acid transport and carboxylic acid transport (for details for GSEA of upregulated and downregulated pathways in $Exosc3^{COIN/COIN}$ cells, see Tables S5 and S6, respectively.)

RNA Exosome-Mediated Degradation of eRNAs Protects Cells from Genomic Instability by Preventing Formation of DNA/RNA Hybrids and by Promoting Heterochromatin Marks at Divergent Enhancers

Regions of the B cell genome beyond the *Ig* loci are susceptible to hypermutation due to AID activity and may then undergo chromosomal translocations involving *Ig* genes. Genomic loci susceptible to AID-induced chromosomal translocation break points may also accumulate x-eRNA reads in $Exosc3^{COIN/COIN}$ B cells in comparison to $Exosc3^{WT/WT}$ B cells. We observed that some *IgH* translocation partners identified through translocation capture techniques show x-eRNA expressing divergently transcribed enhancers as recurrent translocation hotspots. These include the *Birc3* enhancer (Figure S4C), as well as the *Ncoa3* enhancer (Figure S4D). These enhancer regions display overlapping sense and antisense RNA exosome substrate transcripts. Genomic overlaps between translocation breakpoints and x-eRNA-expressing regions provide evidence that RNA exosome-regulated enhancers in the B cell genome could be sensitive to DNA double-strand breaks resulting from AID, a physiologically expressed DNA mutator. Indeed, recently it has been ascertained that Rrp6 (*Exosc10*) plays a role in DNA double-strand break repair by affecting recruitment of ssDNA binding protein RPA (Manfrini et al., 2015; Marin-Vicente et al., 2015). In fact, multiple studies indicate that AID-induced chromosomal translocation sites in the B cell genome harbor RPA for DNA double-strand break repair (Qian et al., 2014; Yamane et al., 2013).

Antisense RNAs that form co-transcriptional RNA/DNA hybrid structures called R-loops can initiate premature transcription termination and be a source of genomic instability (Bhatia et al., 2014; Pefanis et al., 2014; Skourti-Stathaki et al., 2014). In addition, such antisense RNAs can be substrates of the Dicer/Argonaute complex (Skourti-Stathaki et al., 2014) and RNA exosome (Pefanis et al., 2014). To investigate AID-independent DNA break formation in ESCs, we looked at whether x-eRNA-expressing regions are susceptible to genomic instability in RNA exosome-deficient cells due to formation of persistent R-loop structures. ESCs were irradiated with ionizing radiation (20 Gy) and allowed to recover over a period of 30 min. We evaluated three x-eRNA expressing loci neighboring *Klf6*, *Bcl6*, and *Cd38*. x-eRNA arising from these enhancer loci display divergent transcription and are sensitive to *Exosc3* function (Figures S4E–S4G). We evaluated the accumulation of DNA double-strand-break-associated γ-H2AX foci at divergent x-eRNA-expressing regions in $Exosc3^{COIN/COIN}$ and $Exosc10^{COIN/LacZ}$ cells. γ-H2AX accumulation at x-eRNA-expressing sequences was significantly enhanced in both *Exosc3* and *Exosc10* ablated ESC lines, implying a greater propensity for these sequences to undergo DNA double-strand breaks in the absence of functional RNA exosome complex (Figure 4A). Using immunoprecipitation assays with anti-DNA/RNA hybrid S9.6 antibody, we found that, in $Exosc3^{COIN/COIN}$ and $Exosc10^{COIN/LacZ}$ cells, x-eRNA-expressing regions are significantly enriched for RNase-H-sensitive DNA/RNA hybrid structures (Figure 4B). In contrast, an enhancer region in the ESC genome that does not demonstrate divergent transcription was not enriched for γ-H2AX foci or R-loops (Figures S4I and S4J, respectively). These observations point toward the possibility that RNA exosome mutant ESCs are more prone to genomic instability insults at divergently transcribed enhancer sequences. Telomeric fluorescence in situ hybridization (FISH) assays performed on IR-treated $Exosc3^{COIN/COIN}$ cells revealed a significantly greater frequency of chromosomal alteration in comparison to control $Exosc3^{WT/WT}$ cells (Figures S4A and S4B). Taken together, RNA exosome-mediated degradation of RNA in DNA/RNA hybrids at divergently transcribed enhancer sequences might serve as a mechanism for the maintenance of genomic integrity in mammalian cells.

The established roles of H3K9me2 and HP1γ chromatin marks in the cellular processes of chromatin condensation and transcriptional repression have recently been identified to appear at sites of transcription termination of antisense non-coding

(E) x-eRNA stabilization at annotated enhancers within the *Cd83* locus in $Exosc3^{WT/WT}$, $Exosc10^{COIN/LacZ}$, and $Exosc3^{COIN/COIN}$ B cells. Sense and antisense RNA indicated in blue and red, respectively.

(F) Expression levels of x-lncRNAs in ES cells identified from $Exosc3^{WT/WT}$ and $Exosc3^{COIN/COIN}$ transcriptomes.

(G and H) Sequence conservation plot of coding genes (red), identified x-lncRNAs (blue), and random control (green) from B cells (G) and ESCs (H). To measure how conserved the lncRNAs we have identified are, we compared the lncRNAs with human genes (genome version hg19) by LiftOver tool (https://genome.ucsc.edu/cgi-bin/hgLiftOver). The percentage of genes that are conserved between human and mouse is shown according to different cutoffs. The same number of coding genes/random genomic regions with similar length are generated as controls. For each group of genes, the percentage of genes conserved between human and mouse (y axis) is calculated based on UCSC Liftover tool with given cutoff (x axis).

See also Figure S3 and Table S4.

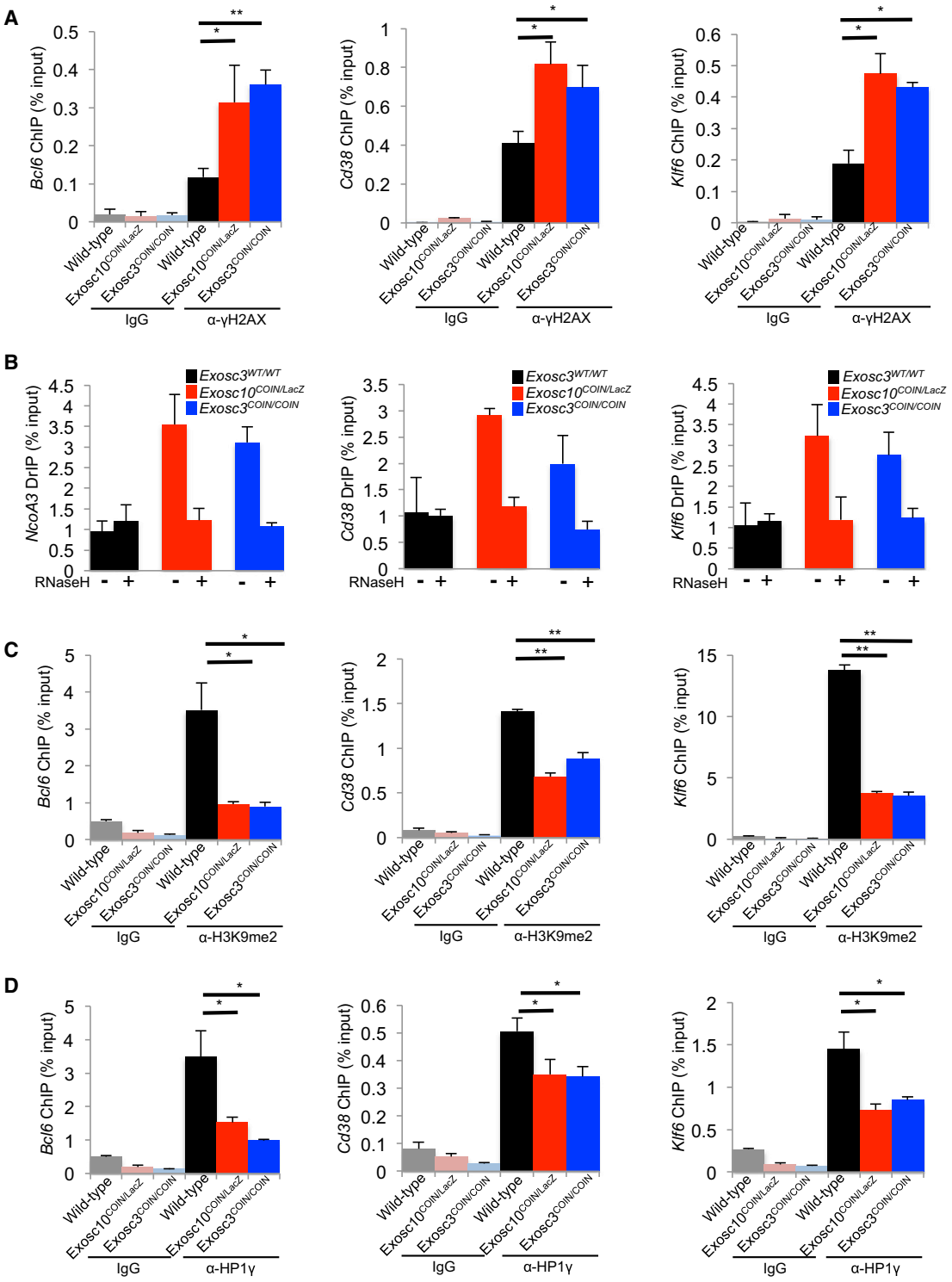


Figure 4. Genomic Instability in RNA Exosome-Deficient ESCs, along with Accumulation of DNA/RNA Hybrids and Loss of Chromatin-Silencing Markers H3K9me2 and HP1γ at x-eRNA-Expressing Sequences

(A) γH2AX immunoprecipitation for DNA double-strand breaks at enhancer sequences resident in the *Bcl6* (left), *Cd38* (middle), and *Klf6* (right) loci in WT, *Exosc10*$^{COIN/LacZ}$, and *Exosc3*$^{COIN/COIN}$ ES cells following ionizing radiation treatment.

(B) DNA/RNA hybrid immunoprecipitation at *Ncoa3* (left), *Cd38* (middle), and *Klf6* (right) enhancers in WT, *Exosc10*$^{COIN/LacZ}$, and *Exosc3*$^{COIN/COIN}$ ES cells.

(legend continued on next page)

RNAs (Skourti-Stathaki et al., 2014). Analysis of H3K9me2 (Figure 4C) and HP1γ (Figure 4D) occupancy revealed decreased levels of these repressive chromatin marks at x-eRNA-expressing loci in *Exosc3*$^{COIN/COIN}$ and *Exosc10*$^{COIN/LacZ}$ cells. Thus, RNA exosome-mediated regulation of x-eRNA levels in cells could occur via two distinct mechanisms, namely via post-transcriptional RNA degradation or possibly through repression of RNA synthesis by promoting early transcription termination. In summary, we provide evidence that x-eRNA-expressing DNA sequences generate potentially deleterious DNA/RNA hybrids that might contribute to genomic instability.

x-eRNAs Have Biological Function at Super-Enhancer Sequences

Because enhancers are well-known modulators of gene expression, we evaluated x-eRNAs that arose from our analyses for functionality in controlling gene expression. We observed two peaks of sense and antisense transcription at regions upstream of the *Tgfbr2* gene (Figure S5A). Using CRISPR-Cas9-mediated deletion of these lncRNA-expressing potential enhancer sequences in B cell line CH12F3, we observed a substantial decrease in the expression of *Tgfbr2* mRNA by individually knocking out either of the two *Tgfbr2* x-eRNA elements (Figure S5B).

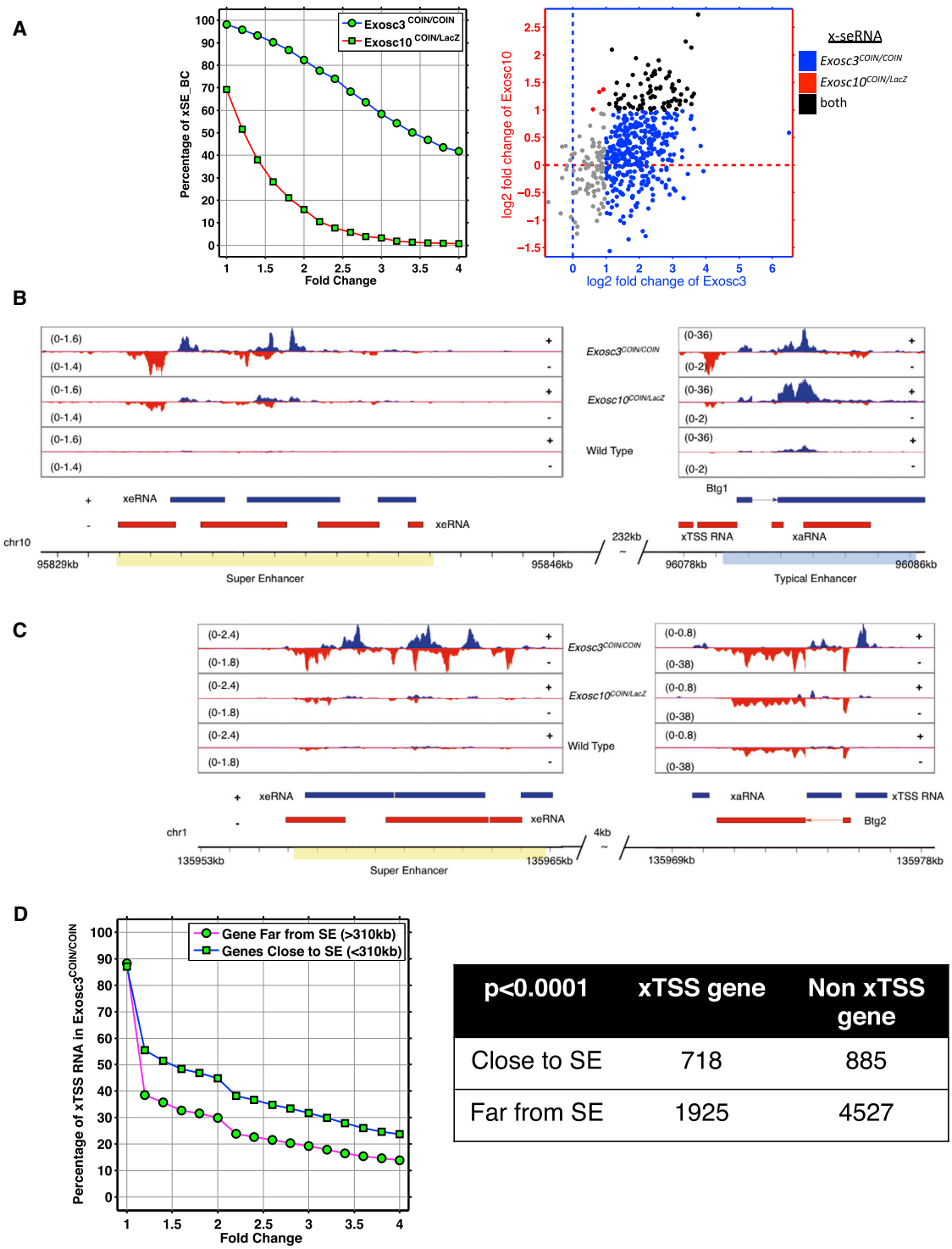
We considered whether super-enhancer sequences, which are characterized by high density of individual enhancers and high regional enrichment for active chromatin marks, can generate RNA exosome substrate super-enhancer RNAs (x-seRNAs). As super-enhancer coordinates and functions can be identified in B cells using previously published bioinformatic pipelines (Lovén et al., 2013; Meng et al., 2014), we evaluated the expression of x-seRNAs in these cells. Our analysis revealed a significant enrichment of x-seRNAs in both *Exosc3* and *Exosc10* exotomes (Figure 5A). Relative to *Exosc3*$^{COIN/COIN}$ cells, *Exosc10*-deficient cells retained significantly greater x-seRNA degradation activity, potentially due to RNA exosome complexes in these cells possessing the ability to utilize either the *Exosc10*-encoded Rrp6 or *Dis3*-encoded Rrp44 nuclease subunit in the degradation of x-seRNAs. We hypothesized that synthesis of antisense RNAs (either xTSS-RNA or those in the body of a gene) may functionally engage with super-enhancer elements to form higher-order chromosomal structures that may enable their local expression control. We sought such examples, i.e., super-enhancer sequences neighboring RNA exosome-sensitive antisense RNA (x-asRNA)-expressing genes and illustrate two examples here. First, a super-enhancer (Chr 10SE)-enhancer (overlapping the *Btg1* gene) pair separated by a distance of 232 kb from each other was found to express both x-seRNAs and xTSS-RNAs, respectively (Figure 5B). Accordingly, both the Chr 10SE x-seRNA and *Btg1* xTSS-RNA are contained within the *Exosc3* and *Exosc10* exotomes. As a second example, we identified a Chr1 SE that closely paired with an x-asRNA arising within the *Btg2* locus. In this case, the separation of the SE and *Btg2* was a mere 4 kb, with both the x-seRNA and the x-asRNA being part of the *Exosc3* and *Exosc10* exotomes (Figure 5C). A statistical analysis of the proximity between xTSS-RNA-expressing genes and x-seRNA-expressing super-enhancer sequences illustrates a remarkable correlation that genes less than 310 kb from a SE are statistically far more likely to express antisense xTSS-RNAs ($p < 0.0001$; Figure 5D). The 310 kb distance between xTSS-RNA and x-seRNA-expressing sequences was set based on a genome-wide statistical analysis of distance between these elements in B cells. Beyond a distance of 310 kb from a super-enhancer, there is a consistent decrease in correlation of x-TSS-RNA expression (Figure S5C; details in Extended Experimental Procedures). These observations at individual loci such as *Btg1* and *Btg2*, along with genome-wide analyses, support a model whereby super-enhancer and counterpart gene interactions are controlled by expression and/or processing of RNA exosome substrate non-coding RNAs.

Molecular Evidence that Antisense RNA/Super-Enhancer RNA Expression Regulates Long-Range *IgH* Locus Recombination

A pair of divergently transcribed x-lncRNAs was found to be expressed at a 2.6 Mb distal region downstream of the 3′RR of the *IgH* locus. Both members of this x-lncRNA pair—named here as B930059L03Rik and lncRNA-CSR—were significantly more stable in *Exosc3*$^{COIN/COIN}$ and *Exosc10*$^{COIN/LacZ}$ B cells but also detectably expressed in WT control B cells (Figure 6C). A detailed map of this lncRNA-locus is shown in Figure S6A; no transcription factor binding sites were computationally predicted to overlap this region (Figure S6A). We proceeded to delete the lncRNA-CSR locus in CH12F3 cells using CRISPR-Cas9 and demonstrated complete loss of expression of lncRNA-CSR (Figure 6A). We found that lncRNA-CSR homozygous deleted CH12F3 cells expressed similar levels of the *IgH* locus recombination catalyst enzyme AID (Figure S5D). When lncRNA-CSR-deficient CH12F3 cells were assayed for CSR efficiency, they showed substantial defect for isotype switching to IgA (Figures 6B and S5E). Chromosome conformation capture (3C) (using lncRNA-CSR 3C primer Figure S6A and HS4 region primer Figure S6B) was performed to assess the interaction frequency of the lncRNA-CSR locus with regions of the *IgH* locus 3′RR super-enhancer (for details see Extended Experimental Procedures). Remarkably, we observed that the HS4 region of the *IgH* locus 3′RR interacts with the lncRNA-CSR locus. Deletion of the lncRNA-CSR sequence substantially decreased the interaction frequency between the deleted locus and the 3′RR HS4 region, whereas the canonical 3′RR and Eμ interaction remained similar (Figure 6D). As can be seen from RNA-seq data, the antisense super-enhancer RNA peak corresponding to 3′RR HS4 (strongly visible in the *Exosc3*$^{COIN/COIN}$ track) also corresponds

(C) Immunoprecipitation for heterochromatin marker H3K9me2 at enhancer sequences resident in the *Bcl6* (left), *Cd38* (middle), and *Klf6* (right) loci in WT, *Exosc10*$^{COIN/LacZ}$, and *Exosc3*$^{COIN/COIN}$ ES cells.

(D) Immunoprecipitation for heterochromatin marker HP1γ at enhancer sequences resident in the *Bcl6* (left), *Cd38* (middle), and *Klf6* (right) loci in WT, *Exosc10*$^{COIN/LacZ}$, and *Exosc3*$^{COIN/COIN}$ ES cells. Each plot is a representation of three independent experiments performed (*$p < 0.05$; **$p < 0.01$ by t test).
See also Figure S4.



p<0.0001	xTSS gene	Non xTSS gene
Close to SE	718	885
Far from SE	1925	4527

Figure 5. Super-Enhancer Sequences and Neighboring Conventional Enhancers or Coding Genes Express RNA Exosome Substrate Antisense RNAs

(A) Left: expression of super-enhancer RNAs at 529 annotated super-enhancers within *Exosc3* (*Exosc3*$^{COIN/COIN}$) and *Exosc10* (*Exosc10*$^{COIN/LacZ}$) exotomes. x axis indicates the cutoff of fold change, and y axis indicates the fraction of super-enhancers with higher expression (given x axis) compared with WT. Right: expression of *Exosc3* (blue), *Exosc10* (red), and overlapping (black) x-seRNAs in B cells.

(B) A super-enhancer resident in chromosome 10 (SEChr10) and neighboring conventional enhancer element resident at the *Btg1* locus express sense (blue) and antisense (red) x-eRNAs.

(legend continued on next page)

to the region of interaction with lncRNA-CSR based on DNA sequencing results from 3C assays (Figure 6C, bottom). The 3′RR HS4 region expresses multiple distinct x-seRNAs, as can be seen from the non-overlapping RNA-seq reads from the *Exosc3*$^{COIN/COIN}$ transcriptome (Figure S6C). It is likely that the lncRNA-CSR element functions as a distal enhancer-like sequence and promotes the CSR-stimulating activity of the 3′RR super-enhancer via the interaction of the antisense lncRNA-CSR and the HS4 x-seRNA-expressing DNA regions. Thus, we provide functional evidence that RNA exosome substrate antisense RNA-expressing elements can interact with super-enhancer RNA-expressing regions to catalyze genomic rearrangement and organization.

We wanted to investigate the molecular mechanism of lncRNA-CSR transcription on the activity of 3′RR function in promoting CSR. The 3′RR is known to regulate transcription of switch region germline transcripts (GLTs) (Birshtein, 2014; Pinaud et al., 2011). IgSμ transcript levels were comparable between parental (WT) and ΔlncRNA-CSR CH12F3 clones (Figure 7A). On the other hand, we observed a significant suppression of IgA germline transcripts (IgSα) in the ΔlncRNA-CSR CH12F3 clones (Figure 7B). These observations point toward a role for lncRNA-CSR/HS4 interaction in regulating the transcription of downstream switch sequence transcripts at the Sα locus. Whether this transcription regulation is similarly enforced at other switch regions can only be determined by generating mouse models deleted of the lncRNA-CSR locus. There is accumulation of long-range DNA rearrangements between the *IgH* (Klein et al., 2011) and lncRNA-CSR loci in B cells that overexpress AID (Figure S7A). Deletion of the lncRNA-CSR locus (Figure S6A) is presumed to disrupt its divergent transcription. We find, at least in these cells in which the transcription divergence is lost, H3K9me2 levels are decreased, raising the possibility that some level of heterochromatinization of these divergent sequences is important for their molecular activity to promote 3′RR interaction (Figure 7D). These observations are consistent with enhancer heterochromatinization regulation in ESCs by RNA exosome, as shown in Figures 4C and 4D. Finally, we evaluated the effect on 3′RR HS4-lncRNA-CSR interaction in B cells deficient in RNA exosome activity (*Exosc3*$^{COIN/COIN}$). We find that, in the absence of *Exosc3*, B cells have increased HS4-lncRNA-CSR interaction frequency relative to WT B cells (Figure 7C). However, increased interaction is not sufficient to promote CSR because the RNA exosome also regulates AID's DNA deamination activity in B cells (Basu et al., 2011; Pefanis et al., 2014; Sun et al., 2013a).

DISCUSSION

We envision that the identification of vast numbers of RNA exosome-targeted ncRNAs will enable the elucidation of their physiological roles in various developmental and gene expression regulatory pathways. Although many lncRNAs and their functions have been described (Bonasio and Shiekhattar, 2014; Rinn and Chang, 2012; Sauvageau et al., 2013), our study identifies a subclass targeted by RNA exosome (x-lncRNA), many of which have not been reported previously. To explore, visualize, and analyze the landscape of these x-lncRNAs, we have generated a public browser showing strand-specific transcripts in the absence and presence of the RNA exosome complex subunits (see Extended Experimental Procedures). Such a tool may shed greater light on co-transcriptional processing dynamics at individual loci of interest and allow for generation of new hypotheses.

Recent findings have revealed the existence of vast numbers of intergenic and intragenic enhancer elements throughout the mammalian genome (Bonasio and Shiekhattar, 2014; Lam et al., 2014). How their activity is regulated is an exciting and open question. Enhancers generate eRNA transcripts whose biological role and regulation beyond chromatin remodeling are not well appreciated. In this study, we unravel the role of RNA exosome-mediated degradation of eRNAs expressed from divergently transcribed loci. We demonstrate that enhancer RNAs generate complexes with single-strand DNA that are protected from being converted to sites of genomic instability by the rapid action of the RNA exosome complex. The formation of R-looped DNA secondary structures can arise from failure to undergo proper transcriptional termination (Skourti-Stathaki et al., 2014). Early transcription termination serves as a mechanism for co-transcriptional RNA exosome recruitment (Lemay et al., 2014; Pefanis et al., 2014). Thus, in the absence of RNA exosome, x-eRNAs may accumulate not solely due to lack of RNA degradation but also due to failure of transiently forming R-loop structure-induced termination at enhancer loci (Skourti-Stathaki et al., 2014). Divergent transcription can create enhanced negative DNA supercoiling that, in turn, promotes the generation of ssDNA structures surrounding enhancer TSSs (Rhee and Pugh, 2012), thereby promoting DNA double-strand breaks and genomic instability (Pefanis et al., 2014). Such breaks could be caused by the activity of an endogenous DNA mutator such as cytidine deaminase AID or due to collisions of replication forks with stalled RNA polymerase complexes at these enhancer sequences (Kim and Jinks-Robertson, 2012). Sense/antisense x-eRNA pairs that form within the R-loop bubble may result in dsRNA that can be processed by RNAi factors, eventually leading to local accumulation of chromatin condensation marks such as H3K9me2 and HP1γ (Skourti-Stathaki et al., 2014). Lack of RNA exosome activity may skew the ratio or abundance of sense and antisense eRNA transcripts, leading to impairment of RNAi pathway recruitment and heterochromatinization. Thus, RNA exosome may play an important role in promoting transcription termination-coupled silencing of divergent enhancer sequences genome wide.

Super-enhancers are large, densely packed enhancer elements that are occupied by master regulators of transcription and mediator proteins (Hnisz et al., 2013; Whyte et al., 2013). These elements are responsible for controlling transcription of diverse sets of tissue-specific gene expression programs. B cell

(C) A super-enhancer resident in chromosome 1 (SEChr1) and neighboring a conventional enhancer element resident at the *Btg2* locus express sense (blue) and antisense (red) x-eRNAs.
(D) Genome-wide correlation between proximity of super-enhancer location and antisense x-TSS-RNA expression at neighboring genes in B cells.
See also Figure S5.

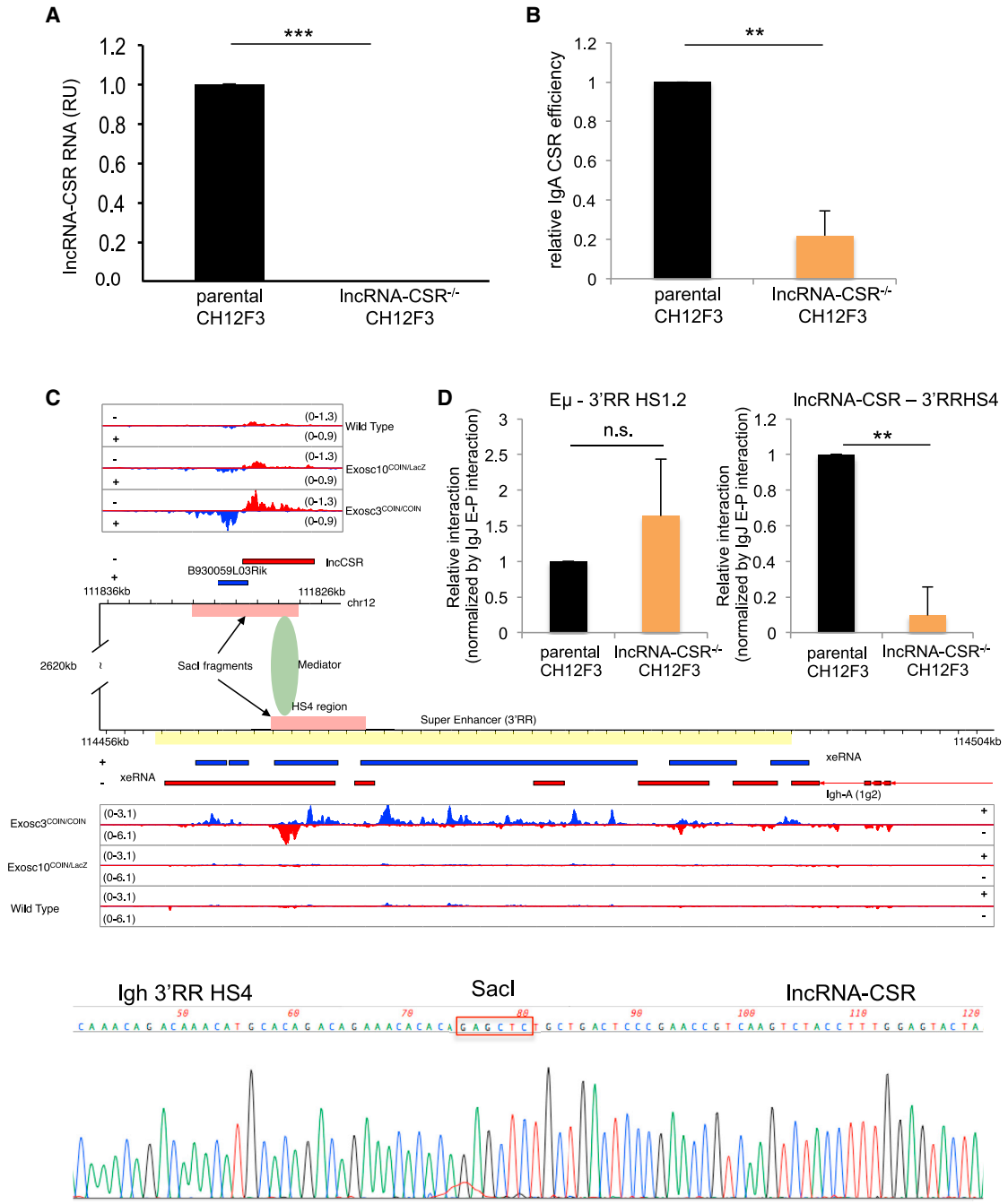


Figure 6. Identification of Divergently Expressed lncRNA-CSR at an Enhancer Region Controlling *IgH* Recombination in B Cells

(A) lncRNA-CSR expression in parental and *lncRNA-CSR*$^{-/-}$ CH12F3 cells following CRISPR/Cas9 mediated deletion.

(B) IgA class switch recombination efficiency of lncRNA-CSR-deleted CH12F3 cells obtained from 18 independent lines of *lncRNA-CSR*$^{-/-}$ CH12F3 cells.

(C) Top: expression profile of the lncRNA-CSR divergently transcribed enhancer locus that is stabilized in *Exosc3*$^{COIN/COIN}$ and *Exosc10*$^{COIN/LacZ}$ B cells. Middle: sense (blue) and antisense (red) tracks for 3′ regulatory region super-enhancer transcription in *Exosc3*$^{COIN/COIN}$ and *Exosc10*$^{COIN/LacZ}$ B cells. Bottom: DNA sequencing of the 3C-derived joint PCR product of the super-enhancer *IgH* 3′RR HS4 sequence with the lncRNA-CSR enhancer sequence. The SacI site is contributed from the lncRNA-CSR locus and the HS4 locus and demonstrates the joining of the two pieces of DNA in the 3C assay.

(D) 3C assay determination of relative interaction frequency of Eμ with 3′RR HS1.2 and lncRNA-CSR locus with 3′RR HS4. **$p < 0.01$ and ***$p < 0.001$ by t test. See also Figures S5 and S6.

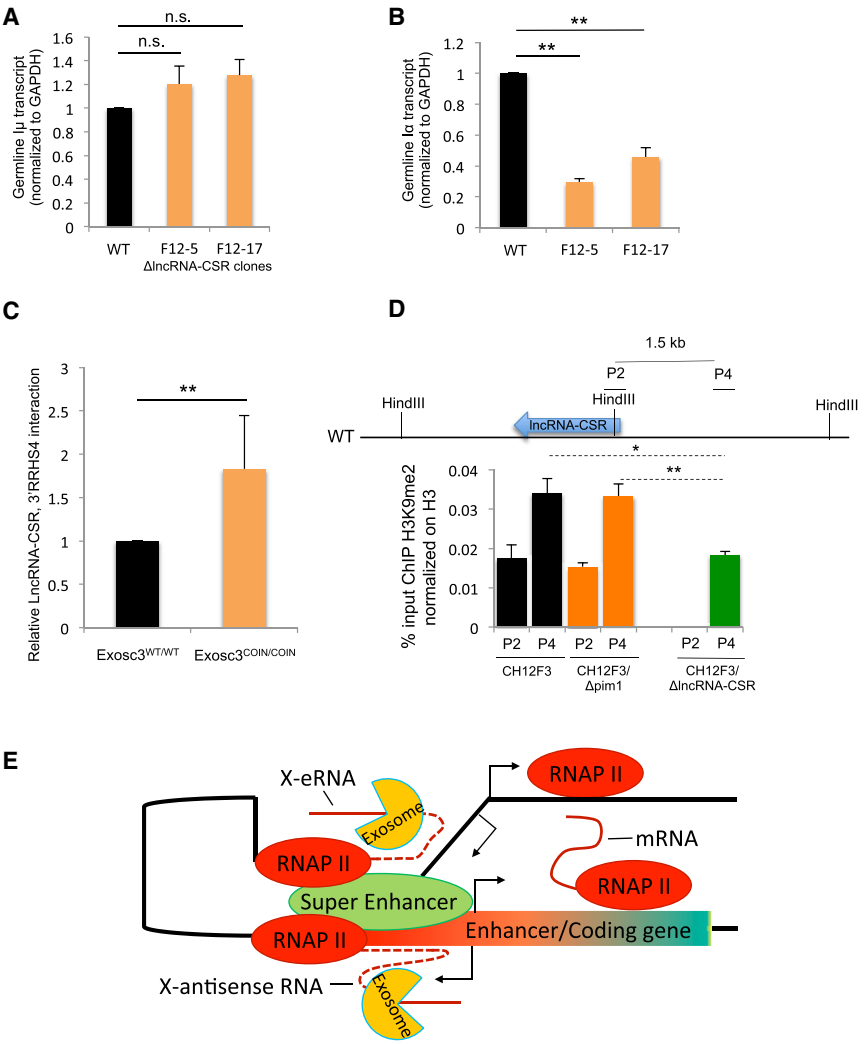


Figure 7. Mechanism of lncRNA-CSR-Mediated Suppression of 3′RR Super-Enhancer Function

(A) Germline transcripts at Iμ in parental cells and two separate clones of lncRNA-CSR knockouts. Three independent sets of RNA were isolated for each cell line and assayed by qRT-PCR.

(B) Germline transcripts at Iα in parental cells and two separate clones of lncRNA-cSR knockouts. Three independent sets of RNA were isolated for each cell line and assayed by qRT-PCR. **$p < 0.01$ by t test.

(C) Chromosomal conformation capture performed on *Exosc3*$^{WT/WT}$ and *Exosc3*$^{COIN/COIN}$ B cells that were stimulated for CSR with LPS+IL4 for 24 hr. The frequency of 3′RRHS4 interaction with the lncRNA-CSR locus was measured by normalizing to an interaction downstream of the *Calr* gene locus. The experiment is a representation of three independently performed assays. **$p < 0.01$ by t test.

(D) The accumulation of H3K9me2 marks (normalized to the presence of H3) in parental (CH12F3 cells), a random CRISPR/Cas9 mutated ΔPim1 (xTSS-RNA mutated) cell, and ΔlncRNA-CSR. Experiment is a representation of three independently performed assays. The ChIP assay primer pairs for various regions surrounding the lncRNA-CSR locus are shown in the top panel; *$p < 0.05$ and **$p < 0.01$ by t test.

(E) A model of RNA exosome substrate x-seRNA-expressing super-enhancer interaction with the divergently transcribing promoter of another enhancer or protein coding gene. We postulate that the activity of RNA exosome to process the x-seRNA and x-eRNAs has a role in titrating the proper level of interaction between regulatory elements that ultimately control gene expression.

See also Figure S7.

super-enhancers have been found to overlap large regions of the human genome susceptible to mutations in diffuse large B cell lymphomas (Chapuy et al., 2013) (Meng et al., 2014; Qian et al., 2014). We evaluated super-enhancers for the presence of RNA exosome-regulated transcripts and correspondingly identified x-seRNAs. Genes or canonical enhancers in proximity to super-enhancers express high levels of RNA exosome-regulated antisense RNAs around their TSSs (xTSS-RNAs) or within gene bodies (x-asRNAs). We hypothesize that super-enhancers may interact with genes under their regulation via mechanisms that depend upon transcription of RNA exosome-regulated transcripts. A test of this hypothesis was undertaken, and we observed that the divergently transcribed lncRNA-CSR enhancer element interacts with the HS4 region of the 3′ regulatory region super-enhancer of the *IgH* locus to control class switch recombination. The dependence of a super-enhancer function on an interacting lncRNA-expressing divergent enhancer provides a newly identified mechanism of gene expression regulation (see Figure 7E for a proposed model). Whether the interaction is dependent upon direct RNA-protein complexes that are co-transcriptionally generated at the cognate pairs of enhancer/promoter and super-enhancer loci is a question of immediate interest. Furthermore, the observation that 3′RR x-seRNAs and lncRNA-CSR are substrates of RNA exosome provides the possibility that RNA exosome regulates long-distance genomic interactions either through its RNA degradation activities and/or through its ability to terminate transcription of ncRNAs at enhancers and super-enhancers.

EXPERIMENTAL PROCEDURES

Details of ChIP experiments, DNA/RNA hybrid immunoprecipitation, and 3C can be found in the Extended Experimental Procedures.

Exosc10^{COIN} Allele Design and Construction

A mouse *Exosc10* locus containing bacterial artificial chromosome (clone bMQ169f23) was modified using bacterial homologous recombination. Briefly, a *lox*2372-*lox*P array was inserted in the first intron of *Exosc10*. In a subsequent recombination event, an inverted *lox*2372-*lox*P array, inverted FP635 expressing terminal exon (COIN module) in antisense orientation to *Exosc10* transcription, and an FRT-flanked neo^r selection cassette were inserted within

a non-conserved region of *Exosc10* exon 2. The *Exosc10* COIN module contains a 3′ splice acceptor sequence immediately followed by an in-frame T2A-FP635-pA cassette. *Exosc10*^{COINneo} BAC recombinants were screened by PCR across all four modified junctions and confirmed using restriction digestion and pulse field electrophoresis. A 20 kb fragment containing the entire *Exosc10*^{COINneo} modification was then subcloned into a plasmid containing a diphtheria toxin A (DTA) cassette. *Exosc10*^{COINneo} homology arms in the DTA vector were 6.7 and 8.2 kb. Linearized *Exosc10*^{COINneo} targeting vector was electroporated into ROSA26$^{CreERt2/+}$, 129S6/SvEv x C57BL/6 hybrid ESCs. Correctly targeted ESC clones were identified using external Southern blotting probes for both the upstream and downstream homology arms on HindIII or NsiI-digested genomic DNA, respectively. *Exosc10*$^{COIN/+}$ chimeric mice were created via blastocyst injection of targeted ESCs. Mice with the greatest ESC-derived coat color contribution were crossed with Tg(ACTB:FLPe) mice to delete the neo^{r} selection cassette and germline transmit the *Exosc10*^{COIN} allele. The FLPe transgene was eliminated during backcrossing. All mouse experiments were conducted in accordance with approved Columbia University Institutional Animal Care and Use Committee protocols.

RNA-Seq Analysis

rRNA-depleted total RNA was prepared using the Ribo-Zero rRNA removal kit (Epicentre). Libraries were prepared with Illumina TruSeq and TruSeq Stranded total RNA sample prep kits and then sequenced with 50–60 million of 2 × 100 bp paired raw passing filters reads on an Illumina HiSeq 2000 V3 instrument at the Columbia Genome Center. The details of generation of exotomes from *Exosc3*-deficient or *Exosc10*-deficient B cells and ESCs and their subsequent analysis are described in the Extended Experimental Procedures.

Transcriptome Reconstitution

Details of transcriptome reconstitution of the *Exosc3* and *Exosc10* exotomes from B cells and ESCs are described in detail in the Extended Experimental Procedures, and the data are provided in Tables S1, S2, S3, and S4 and in the "Exotome browser," which can be accessed from (http://rabadan.c2b2.columbia.edu/cgi-bin/hgGateway).

ACCESSION NUMBERS

The accession number for the RNA-sequencing data reported in this paper is Sequence Read Archive (SRA): SRP042355.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, seven figures, and six tables and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2015.04.034.

AUTHOR CONTRIBUTIONS

E.P., J.W., and U.B. planned the studies, and E.P., J.W., G.R., R.R., and U.B. interpreted the data. Experiments were performed as follows: E.P. and J.C., undertook mouse model generation and RNA-seq studies; G.R. and J.L. performed 3C, ChIP, DRIP, and CRISPR/Cas9; J.W. conducted bioinformatic studies and designed all the pipelines for determining various RNA exosome substrate non-coding RNAs; A.N.E. advised on mouse model construct; R.R. oversaw bioinformatics and guided J.W.; D.K. and J.S. prepared and analyzed metaphases of RNA exosome mutant cells; J.W. and O.E. prepared exotome browser; A.F. and J.E.B. identified the super-enhancer sequences in B cells; and E.P., J.W., and U.B. wrote the manuscript, which was further refined by all of the other authors.

ACKNOWLEDGMENTS

We thank Christopher Lima (MSKCC, New York), Frederic Chedin (University of California, Davis), Saul Silverstein, Stephen Goff, Sankar Ghosh, Lorraine Symington, and members of the Basu lab for critical input and reagents. We thank Olivier Courrone of Columbia University genome center for RNA sequencing and Victor Lin of Herbert Irving Cancer Center Transgenic facility for targeting and generation of *Exosc3*^{COIN} and *Exosc10*^{COIN} allele ESCs and mouse models. This work was supported by grants from NIH (1DP2OD008651-01) and NIAID (1R01AI099195-01A1) to U.B.; NIH (1R01CA185486-01; 1R01CA179044-01A1; 1U54CA121852-05) to R.R.; and F31AI098411-01A1 to D.K. U.B. is a scholar of the Irma Hirschl Charitable Trust and the Leukemia and Lymphoma Society.

Received: November 30, 2014
Revised: March 11, 2015
Accepted: April 20, 2015
Published: May 7, 2015

REFERENCES

Andersson, R., Gebhard, C., Miguel-Escalada, I., Hoof, I., Bornholdt, J., Boyd, M., Chen, Y., Zhao, X., Schmidl, C., Suzuki, T., et al.; FANTOM Consortium (2014a). An atlas of active enhancers across human cell types and tissues. Nature *507*, 455–461.

Andersson, R., Refsing Andersen, P., Valen, E., Core, L.J., Bornholdt, J., Boyd, M., Heick Jensen, T., and Sandelin, A. (2014b). Nuclear stability and transcriptional directionality separate functionally distinct RNA species. Nat. Commun. *5*, 5336.

Basu, U., Meng, F.L., Keim, C., Grinstein, V., Pefanis, E., Eccleston, J., Zhang, T., Myers, D., Wasserman, C.R., Wesemann, D.R., et al. (2011). The RNA exosome targets the AID cytidine deaminase to both strands of transcribed duplex DNA substrates. Cell *144*, 353–363.

Bhatia, V., Barroso, S.I., García-Rubio, M.L., Tumini, E., Herrera-Moyano, E., and Aguilera, A. (2014). BRCA2 prevents R-loop accumulation and associates with TREX-2 mRNA export factor PCID2. Nature *511*, 362–365.

Birshtein, B.K. (2014). Epigenetic Regulation of Individual Modules of the immunoglobulin heavy chain locus 3′ Regulatory Region. Front. Immunol. *5*, 163.

Bonasio, R., and Shiekhattar, R. (2014). Regulation of transcription by long noncoding RNAs. Annu. Rev. Genet. *48*, 433–455.

Chapuy, B., McKeown, M.R., Lin, C.Y., Monti, S., Roemer, M.G., Qi, J., Rahl, P.B., Sun, H.H., Yeda, K.T., Doench, J.G., et al. (2013). Discovery and characterization of super-enhancer-associated dependencies in diffuse large B cell lymphoma. Cancer Cell *24*, 777–790.

Chlebowski, A., Lubas, M., Jensen, T.H., and Dziembowski, A. (2013). RNA decay machines: the exosome. Biochim. Biophys. Acta *1829*, 552–560.

Colin, J., Candelli, T., Porrua, O., Boulay, J., Zhu, C., Lacroute, F., Steinmetz, L.M., and Libri, D. (2014). Roadblock termination by reb1p restricts cryptic and readthrough transcription. Mol. Cell *56*, 667–680.

Core, L.J., Waterfall, J.J., and Lis, J.T. (2008). Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science *322*, 1845–1848.

Economides, A.N., Frendewey, D., Yang, P., Dominguez, M.G., Dore, A.T., Lobov, I.B., Persaud, T., Rojas, J., McClain, J., Lengyel, P., et al. (2013). Conditionals by inversion provide a universal method for the generation of conditional alleles. Proc. Natl. Acad. Sci. USA *110*, E3179–E3188.

Flynn, R.A., and Chang, H.Y. (2012). Active chromatin and noncoding RNAs: an intimate relationship. Curr. Opin. Genet. Dev. *22*, 172–178.

Flynn, R.A., Almada, A.E., Zamudio, J.R., and Sharp, P.A. (2011). Antisense RNA polymerase II divergent transcripts are P-TEFb dependent and substrates for the RNA exosome. Proc. Natl. Acad. Sci. USA *108*, 10460–10465.

Guttman, M., Amit, I., Garber, M., French, C., Lin, M.F., Feldser, D., Huarte, M., Zuk, O., Carey, B.W., Cassady, J.P., et al. (2009). Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature *458*, 223–227.

Guttman, M., Garber, M., Levin, J.Z., Donaghey, J., Robinson, J., Adiconis, X., Fan, L., Koziol, M.J., Gnirke, A., Nusbaum, C., et al. (2010). Ab initio

reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. Nat. Biotechnol. *28*, 503–510.

Guttman, M., Donaghey, J., Carey, B.W., Garber, M., Grenier, J.K., Munson, G., Young, G., Lucas, A.B., Ach, R., Bruhn, L., et al. (2011). lincRNAs act in the circuitry controlling pluripotency and differentiation. Nature *477*, 295–300.

Hazelbaker, D.Z., Marquardt, S., Wlotzka, W., and Buratowski, S. (2013). Kinetic competition between RNA Polymerase II and Sen1-dependent transcription termination. Mol. Cell *49*, 55–66.

Hnisz, D., Abraham, B.J., Lee, T.I., Lau, A., Saint-André, V., Sigova, A.A., Hoke, H.A., and Young, R.A. (2013). Super-enhancers in the control of cell identity and disease. Cell *155*, 934–947.

Houseley, J., LaCava, J., and Tollervey, D. (2006). RNA-quality control by the exosome. Nat. Rev. Mol. Cell Biol. *7*, 529–539.

Januszyk, K., and Lima, C.D. (2011). Structural components and architectures of RNA exosomes. Adv. Exp. Med. Biol. *702*, 9–28.

Kapranov, P., Cheng, J., Dike, S., Nix, D.A., Duttagupta, R., Willingham, A.T., Stadler, P.F., Hertel, J., Hackermüller, J., Hofacker, I.L., et al. (2007). RNA maps reveal new RNA classes and a possible function for pervasive transcription. Science *316*, 1484–1488.

Kim, N., and Jinks-Robertson, S. (2012). Transcription as a source of genome instability. Nat. Rev. Genet. *13*, 204–214.

Kim, T.K., Hemberg, M., Gray, J.M., Costa, A.M., Bear, D.M., Wu, J., Harmin, D.A., Laptewicz, M., Barbara-Haley, K., Kuersten, S., et al. (2010). Widespread transcription at neuronal activity-regulated enhancers. Nature *465*, 182–187.

Klein, I.A., Resch, W., Jankovic, M., Oliveira, T., Yamane, A., Nakahashi, H., Di Virgilio, M., Bothmer, A., Nussenzweig, A., Robbiani, D.F., et al. (2011). Translocation-capture sequencing reveals the extent and nature of chromosomal rearrangements in B lymphocytes. Cell *147*, 95–106.

Lam, M.T., Li, W., Rosenfeld, M.G., and Glass, C.K. (2014). Enhancer RNAs and regulated transcriptional programs. Trends Biochem. Sci. *39*, 170–182.

Lemay, J.F., Larochelle, M., Marguerat, S., Atkinson, S., Bähler, J., and Bachand, F. (2014). The RNA exosome promotes transcription termination of backtracked RNA polymerase II. Nat. Struct. Mol. Biol. *21*, 919–926.

Liu, Q., Greimann, J.C., and Lima, C.D. (2006). Reconstitution, activities, and structure of the eukaryotic RNA exosome. Cell *127*, 1223–1237.

Lorentzen, E., Basquin, J., Tomecki, R., Dziembowski, A., and Conti, E. (2008). Structure of the active subunit of the yeast exosome core, Rrp44: diverse modes of substrate recruitment in the RNase II nuclease family. Mol. Cell *29*, 717–728.

Lovén, J., Hoke, H.A., Lin, C.Y., Lau, A., Orlando, D.A., Vakoc, C.R., Bradner, J.E., Lee, T.I., and Young, R.A. (2013). Selective inhibition of tumor oncogenes by disruption of super-enhancers. Cell *153*, 320–334.

Lykke-Andersen, S., Brodersen, D.E., and Jensen, T.H. (2009). Origins and activities of the eukaryotic exosome. J. Cell Sci. *122*, 1487–1494.

Manfrini, N., Trovesi, C., Wery, M., Martina, M., Cesena, D., Descrimes, M., Morillon, A., d'Adda di Fagagna, F., and Longhese, M.P. (2015). RNA-processing proteins regulate Mec1/ATR activation by promoting generation of RPA-coated ssDNA. EMBO Rep. *16*, 221–231.

Marin-Vicente, C., Domingo-Prim, J., Eberle, A.B., and Visa, N. (2015). RRP6/EXOSC10 is required for the repair of DNA double-strand breaks by homologous recombination. J. Cell Sci. *128*, 1097–1107.

Meng, F.L., Du, Z., Federation, A., Hu, J., Wang, Q., Kieffer-Kwon, K.R., Meyers, R.M., Amor, C., Wasserman, C.R., Neuberg, D., et al. (2014). Convergent transcription at intragenic super-enhancers targets AID-initiated genomic instability. Cell *159*, 1538–1548.

Pefanis, E., Wang, J., Rothschild, G., Lim, J., Chao, J., Rabadan, R., Economides, A.N., and Basu, U. (2014). Noncoding RNA transcription targets AID to divergently transcribed loci in B cells. Nature *514*, 389–393.

Pinaud, E., Marquet, M., Fiancette, R., Péron, S., Vincent-Fabert, C., Denizot, Y., and Cogné, M. (2011). The IgH locus 3′ regulatory region: pulling the strings from behind. Adv. Immunol. *110*, 27–70.

Preker, P., Nielsen, J., Kammler, S., Lykke-Andersen, S., Christensen, M.S., Mapendano, C.K., Schierup, M.H., and Jensen, T.H. (2008). RNA exosome depletion reveals transcription upstream of active human promoters. Science *322*, 1851–1854.

Qian, J., Wang, Q., Dose, M., Pruett, N., Kieffer-Kwon, K.R., Resch, W., Liang, G., Tang, Z., Mathé, E., Benner, C., et al. (2014). B cell super-enhancers and regulatory clusters recruit AID tumorigenic activity. Cell *159*, 1524–1537.

Reyes-Turcu, F.E., and Grewal, S.I. (2012). Different means, same end-heterochromatin formation by RNAi and RNAi-independent RNA processing factors in fission yeast. Curr. Opin. Genet. Dev. *22*, 156–163.

Rhee, H.S., and Pugh, B.F. (2012). Genome-wide structure and organization of eukaryotic pre-initiation complexes. Nature *483*, 295–301.

Richard, P., and Manley, J.L. (2009). Transcription termination by nuclear RNA polymerases. Genes Dev. *23*, 1247–1269.

Rinn, J.L., and Chang, H.Y. (2012). Genome regulation by long noncoding RNAs. Annu. Rev. Biochem. *81*, 145–166.

Sauvageau, M., Goff, L.A., Lodato, S., Bonev, B., Groff, A.F., Gerhardinger, C., Sanchez-Gomez, D.B., Hacisuleyman, E., Li, E., Spence, M., et al. (2013). Multiple knockout mouse models reveal lincRNAs are required for life and brain development. eLife *2*, e01749.

Schaeffer, D., and van Hoof, A. (2011). Different nuclease requirements for exosome-mediated degradation of normal and nonstop mRNAs. Proc. Natl. Acad. Sci. USA *108*, 2366–2371.

Schaeffer, D., Tsanova, B., Barbas, A., Reis, F.P., Dastidar, E.G., Sanchez-Rotunno, M., Arraiano, C.M., and van Hoof, A. (2009). The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities. Nat. Struct. Mol. Biol. *16*, 56–62.

Schmid, M., and Jensen, T.H. (2008). The exosome: a multipurpose RNA-decay machine. Trends Biochem. Sci. *33*, 501–510.

Seila, A.C., Calabrese, J.M., Levine, S.S., Yeo, G.W., Rahl, P.B., Flynn, R.A., Young, R.A., and Sharp, P.A. (2008). Divergent transcription from active promoters. Science *322*, 1849–1851.

Seila, A.C., Core, L.J., Lis, J.T., and Sharp, P.A. (2009). Divergent transcription: a new feature of active promoters. Cell Cycle *8*, 2557–2564.

Shah, S., Wittmann, S., Kilchert, C., and Vasiljeva, L. (2014). lncRNA recruits RNAi and the exosome to dynamically regulate pho1 expression in response to phosphate levels in fission yeast. Genes Dev. *28*, 231–244.

Shin, J.H., Wang, H.L., Lee, J., Dinwiddie, B.L., Belostotsky, D.A., and Chekanova, J.A. (2013). The role of the Arabidopsis Exosome in siRNA-independent silencing of heterochromatic loci. PLoS Genet. *9*, e1003411.

Skourti-Stathaki, K., Kamieniarz-Gdula, K., and Proudfoot, N.J. (2014). R-loops induce repressive chromatin marks over mammalian gene terminators. Nature *516*, 436–439.

Storb, U. (2014). Why does somatic hypermutation by AID require transcription of its target genes? Adv. Immunol. *122*, 253–277.

Sun, J., Keim, C.D., Wang, J., Kazadi, D., Oliver, P.M., Rabadan, R., and Basu, U. (2013a). E3-ubiquitin ligase Nedd4 determines the fate of AID-associated RNA polymerase II in B cells. Genes Dev. *27*, 1821–1833.

Sun, J., Rothschild, G., Pefanis, E., and Basu, U. (2013b). Transcriptional stalling in B-lymphocytes: a mechanism for antibody diversification and maintenance of genomic integrity. Transcription *4*, 127–135.

Taft, R.J., Glazov, E.A., Cloonan, N., Simons, C., Stephen, S., Faulkner, G.J., Lassmann, T., Forrest, A.R., Grimmond, S.M., Schroder, K., et al. (2009). Tiny RNAs associated with transcription start sites in animals. Nat. Genet. *41*, 572–578.

Wan, Y., Qu, K., Ouyang, Z., Kertesz, M., Li, J., Tibshirani, R., Makino, D.L., Nutter, R.C., Segal, E., and Chang, H.Y. (2012). Genome-wide measurement of RNA folding energies. Mol. Cell *48*, 169–181.

Wasmuth, E.V., and Lima, C.D. (2012). Exo- and endoribonucleolytic activities of yeast cytoplasmic and nuclear RNA exosomes are dependent on the non-catalytic core and central channel. Mol. Cell *48*, 133–144.

Wasmuth, E.V., Januszyk, K., and Lima, C.D. (2014). Structure of an Rrp6-RNA exosome complex bound to poly(A) RNA. Nature *511*, 435–439.

Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell *153*, 307–319.

Yamane, A., Robbiani, D.F., Resch, W., Bothmer, A., Nakahashi, H., Oliveira, T., Rommel, P.C., Brown, E.J., Nussenzweig, A., Nussenzweig, M.C., and Casellas, R. (2013). RPA accumulation during class switch recombination represents 5′-3′ DNA-end resection during the S-G2/M phase of the cell cycle. Cell Rep. *3*, 138–147.