Correlated protein-RNA associations reveal a requirement for HNRNPU in long-range Polycomb recruitment by the lncRNAs *Airn*, *Kcnq1ot1*, and *Xist*

McKenzie M. Murvin[1,2,3,4,5], Shuang Li[1,2,3], Samuel P. Boyson[1,2,3,5], Zhiyue Zhang[1,6], Rachel E. Cherney[1,2,3,5,7], J. Mauro Calabrese[1,2,3,8]

[1]Department of Pharmacology,

[2]RNA Discovery Center,

[3]Lineberger Comprehensive Cancer Center,

[4]Curriculum in Mechanistic, Interdisciplinary Studies of Biological Systems,

[5]Curriculum in Genetics and Molecular Biology, University of North Carolina, Chapel Hill, NC, 27599

[6]Present address: Computer Science Graduate Program, Princeton University, Princeton, NJ, 08544

[7]Present address: Korro Bio, Cambridge, MA 02141

[8]Corresponding author and lead contact; jmcalabr@med.unc.edu

Abstract

The lncRNAs *Airn*, *Kcnq1ot1*, and *Xist* recruit Polycomb Repressive Complexes (PRCs) and repress genes over multi-megabase genomic regions, but how they interact with proteins to carry out repression remains poorly understood. To investigate, we conducted RNA-immunoprecipitations from formaldehyde-crosslinked mouse trophoblast stem cells, using antibodies against 27 proteins known to associate with *Xist*. Relative to other chromatin-bound transcripts, *Airn*, *Kcnq1ot1*, and *Xist* exhibited enriched and highly correlated associations with several RNA-binding proteins. The degree of repression induced by *Airn*, *Kcnq1ot1*, and *Xist* unexpectedly correlated with the extent to which the lncRNAs partitioned protein associations into communities. HNRNPU, a chromatin-associated RNA-binding protein, exhibited enriched association with, and was required to induce PRC-directed modifications by, all three lncRNAs, without being necessary for proper localization of *Airn* or *Kcnq1ot1*. Our data nominate multivalent RNA-protein interaction domains, along with the chromatin-associated RNA-binding protein HNRNPU, as features important for long-range gene regulation induced by chromatin-associated RNAs.

Introduction

The majority of the mammalian genome is transcribed, and many of the RNA transcripts produced are classified as long noncoding RNAs (lncRNAs), which have little or no potential to encode for protein even after splicing. While most lncRNAs are functionally unannotated, a growing number are known to play important roles in health and disease via their ability to repress gene expression *in cis*, on the same chromosomal allele from which the lncRNAs are transcribed (Mattick et al. 2023). However, the mechanisms by which lncRNAs interact with specific proteins to carry out repression or other gene regulatory functions remain poorly understood.

The best-studied and most potent repressive lncRNA is *Xist*, an 18 kilobase (kb) spliced and polyadenylated transcript that silences gene expression over the 165 megabase (Mb) X-chromosome early during the development of eutherian females. On autosomes, at least two lncRNAs are also known to repress gene expression *in cis*: *Airn* and *Kcnq1ot1*. *Airn* is expressed from a locus on mouse chromosome 17 (chr17), where it can repress a genomic region that spans up to ~15 Mb (Andergassen et al. 2017; Andergassen et al. 2019; Schertzer et al. 2019a). *Kcnq1ot1* is expressed from a locus on mouse chr7 (human chr11), where it can repress a genomic region that spans up to ~3 Mb (Umlauf et al. 2004; Lewis et al. 2006; Pandey et al. 2008; Schertzer et al. 2019a). While *Xist* is 18kb long, robustly spliced and abundant, both *Airn* and *Kcnq1ot1* are upwards of 90 kb long, are predominantly unspliced, and are lower in abundance than *Xist* (Mitsuya et al. 1999; Smilinich et al. 1999; Mancini-Dinardo et al. 2006; Seidl et al. 2006; Schertzer et al. 2019a; Braceros et al. 2023; Trotman et al. 2025a).

Despite differing in length, splicing status, and abundance, accumulating data indicate that *Airn*, *Kcnq1ot1,* and *Xist* require several of the same protein cofactors for their repressive functions. These cofactors include at least three histone-modifying enzymes: the Polycomb Repressive Complex 1 (PRC1), which catalyzes the monoubiquitination of lysine 119 on Histone H2A; the Polycomb Repressive Complex 2 (PRC2), which catalyzes the trimethylation of lysine

27 on Histone H3; and G9a, which catalyzes the demethylation of lysine 9 on Histone H3 (Mager et al. 2003; Lewis et al. 2004; Nagano et al. 2008; Terranova et al. 2008; Wagschal et al. 2008; Sachani et al. 2018; Andergassen et al. 2021; Szanto et al. 2021; Masui et al. 2023). Additionally, *Airn*, *Kcnq1ot1,* and *Xist* require the RNA-binding protein (RBP) HNRNPK to recruit the PRCs to chromatin, and *Xist* and *Kcnq1ot1* each require an RBP called SPEN to induce gene silencing (Schertzer et al. 2019a; Quinodoz et al. 2021; Trotman et al. 2025a). These shared dependencies suggest that *Airn*, *Kcnq1ot1,* and *Xist* may engage with larger, shared subsets of RBPs to maintain repression. However, while protein interactions within *Xist* have been well-characterized (Chu et al. 2015; McHugh et al. 2015; Minajigi et al. 2015; Smola et al. 2016; Pintacuda et al. 2017; Bousard et al. 2019; Graindorge et al. 2019; Weidmann et al. 2021; Yu et al. 2021; Tsue et al. 2024), the full spectrum of RBPs that interact with *Airn* and *Kcnq1ot1*, their patterns of interaction within each lncRNA, and the functional relevance of the interactions are incompletely defined.

To establish maps of RNA-protein interactions and identify potential protein cofactors necessary for long-range gene repression by *Airn*, *Kcnq1ot1,* and *Xist*, we performed RNA immunoprecipitation coupled with sequencing (RIP-Seq) for a panel of RBPs in mouse trophoblast stem cells (TSCs), in which all three lncRNAs are endogenously expressed (Schertzer et al. 2019a; Braceros et al. 2023; Trotman et al. 2025a). Our comparative analyses revealed a high degree of similarity in the enrichment and patterns of protein associations between *Airn*, *Kcnq1ot1*, and *Xist*, supporting the view that these lncRNAs carry out repression via analogous mechanisms. Using CRISPR-mediated depletion, we found that the chromatin-associated RBP HNRNPU is required for the long-range induction of PRC-directed modifications by all three lncRNAs, without being necessary for the proper localization of *Airn* or *Kcnq1ot1*. The extent to which proteins partitioned into distinct clusters along the lengths of each of the three lncRNAs correlated with the genomic range over which they recruit PRCs and repress genes, nominating

multiprotein interaction domains as features important for long-range gene regulation by chromatin-associated RNAs.

Results

Protein association profiles of *Airn*, *Kcnq1ot1*, and *Xist* are highly similar and exhibit similarity to profiles of transcripts produced from both lncRNA and protein-coding genes

Our previous analysis of RIP-Seq experiments from formaldehyde-crosslinked TSCs showed that relative to other expressed, chromatin-associated transcripts, *Airn*, *Kcnq1ot1*, and *Xist* exhibited highly enriched associations with HNRNPK, RING1B, and RYBP, proteins that play important roles in repression mediated by all three lncRNAs (Schertzer et al. 2019a; Trotman et al. 2025a). We reasoned that other proteins involved in repression might also exhibit highly enriched association with *Airn*, *Kcnq1ot1*, and *Xist*. To investigate this hypothesis, we performed RIP-Seq in TSCs for a panel of 25 different RBPs and two components of PRC1 (RING1B and RYBP), all of which have been reported to associate with *Xist*, along with non-specific IgG controls (Figure 1A; Table S1; (Chu et al. 2015; McHugh et al. 2015; Minajigi et al. 2015; Smola et al. 2016; Pintacuda et al. 2017; Bousard et al. 2019; Graindorge et al. 2019; Weidmann et al. 2021; Yu et al. 2021; Tsue et al. 2024)). Read densities under peaks were highly correlated between replicates (Table S1). Sequences under peaks were enriched for expected sets of motifs in cases where motifs had been previously defined (Figures S1, S2). Likewise, we observed expected patterns of enrichment for select RBPs over their cognate binding regions in *Xist*, providing additional support that our RIP protocol can recover expected sets of RNA-protein interactions (Figure S3). Because we used formaldehyde as a crosslinking reagent, our RIPs are expected to recover both direct and indirect (protein-bridged) RNA-protein interactions; therefore, we use the term “association” rather than “binding” to describe RNA-protein interactions detected by this method.

To determine the extent to which the proteins selected for RIP-Seq associated with *Airn*, *Kcnq1ot1*, and *Xist* in comparison to other expressed transcripts, we summed the background-corrected signal under peak for each protein examined over the length of each expressed transcript. We then ranked the set of chromatin-enriched transcripts in TSCs by their overall levels of association with each protein (Figure 1B; (Trotman et al. 2025a)). Relative to their own expression levels (assessed by Total and Chromatin-associated RNA-Seq), *Airn* and *Kcnq1ot1* exhibited enriched associations with several RBPs that are known to be required for silencing by *Xist*, including HNRNPK, HNRNPU, RBM15, RING1B, RYBP, SPEN, and YTHDC1. *Airn* and *Kcnq1ot1* also exhibited enrichment for several proteins not enriched over *Xist*, including HNRNPL, PABPN1, SAFB, TIA1, and XRN2 (Figure 1B).

We next sought to quantify the similarity in protein-association rankings among *Airn*, *Kcnq1ot1,* and *Xist* relative to rankings in all other expressed, chromatin-enriched transcripts in TSCs. Of the set of 19295 chromatin-enriched transcripts detected by RNA-Seq in TSCs, 18820 had non-zero, background-corrected values for at least one protein examined by RIP-Seq and were included in the comparisons below. Examining this set of transcripts revealed that the protein association profiles of *Airn* and *Kcnq1ot1* were substantially more similar to each other than to nearly all other chromatin-enriched RNAs, including *Xist*. The association profile of *Xist* was more similar to that of *Airn* and *Kcnq1ot1* than it was similar to most other chromatin-associated RNAs, but only moderately so. More specifically, relative to *Airn*, out of 18820 chromatin-enriched transcripts covered by RIP-Seq data, the protein association profiles of *Kcnq1ot1* and *Xist* were the 92nd and 6888th most similar, respectively; relative to *Kcnq1ot1*, the profiles of *Airn* and *Xist* were the 270th and 9744th most similar, respectively; and relative to *Xist*, the protein association profiles of *Airn* and *Kcnq1ot1* were the 2352nd and 2601st most similar, respectively (Figure 1C; Tables S2, S3). Thus, when compared to *Xist*, *Airn* and *Kcnq1ot1* exhibited both shared and

different sets of enriched RBPs. In terms of their overall protein association levels, *Airn* and *Kcnq1ot1* are more similar to each other than to the majority of other chromatin-enriched RNAs.

Next, we examined the types of chromatin-associated transcripts that resembled *Airn*, *Kcnq1ot1,* and *Xist* in terms of their RIP-Seq association profiles. We used Pearson's correlation to identify the 1,000 transcripts whose background-corrected patterns of RIP-Seq signal across all datasets were the most similar to those of *Airn*, *Kcnq1ot1,* and *Xist.* In each set of top 1,000 similar transcripts, we evaluated the proportions of transcripts produced from lncRNA versus protein-coding genes and compared these proportions to those present in the complete set of chromatin-associated RNAs. We also compared the proportions of intron-excluding (presumably spliced) and intron-including (presumably nascent) transcripts. These analyses revealed that the majority of transcripts most similar to *Airn*, *Kcnq1ot1,* and *Xist* are produced from protein-coding genes (Figure 1D; over 70% in all three sets). Relative to expressed chromatin-associated transcripts, the sets of *Airn*- and *Xist*-similar transcripts exhibited a mild but significant enrichment for transcripts originating from lncRNA over protein-coding genes, whereas the set of *Kcnqot1*-similar transcripts exhibited a mild but significant enrichment for transcripts originating from protein-coding genes over lncRNAs (Figure 1D; p-adj.<0.01 for all comparisons; Fisher's exact). *Airn*- and *Kcnq1ot1*-similar transcripts exhibited significant enrichment for intron-including RNAs, whereas *Xist*-similar transcripts exhibited significant enrichment for intron-excluding RNAs (Figure 1D; p-adj.<0.01 for all comparisons; Fisher's exact). Thus, based on the set of proteins we examined by RIP-Seq, the majority of chromatin-associated transcripts most similar to *Airn* and *Kcnq1ot1*, in terms of their protein association profiles, are intron-containing RNAs produced from protein-coding genes. *Xist* exhibits a higher degree of similarity to intron-excluding transcripts, the majority of which are also produced from protein-coding genes.

Networks of protein association in *Airn*, *Kcnq1ot1*, *Xist* are highly similar and reveal a correlation between modularity and repressive potency

Having compared the aggregate levels of protein association between *Airn*, *Kcnq1ot1*, and *Xist*, we next sought to compare the patterns of protein association in each lncRNA. To do so, we compared background-corrected RIP-Seq read density in 25 nucleotide bins for all possible pairs of proteins in all chromatin-enriched transcripts in TSCs. We used these comparisons to create protein interaction networks, in which edges were defined by the Pearson's r value for each protein pair. Networks for *Airn*, *Kcnq1ot1*, and *Xist,* along with RIP-Seq wiggle density tracks for select proteins, are shown in Figure 2A. We then examined the similarity of the protein interaction networks in *Airn*, *Kcnq1ot1*, and *Xist* relative to all other chromatin-enriched RNAs. Of the 19295 chromatin-enriched RNA transcripts in TSCs, 19223 had non-zero binned values for at least one protein profiled by RIP-Seq and were included in the comparisons below. Consistent with *Airn* and *Kcnq1ot1* harboring a high degree of similarity in terms of their aggregate protein associations (Figure 1C), these analyses underscored a remarkable level of similarity between the networks of protein association in *Airn* and *Kcnq1ot1*. Relative to 19223 chromatin-enriched transcripts, the protein association network of *Kcnq1ot1* ranked 6th in similarity to *Airn*, and that of *Airn* ranked 1st in similarity to *Kcnq1ot1* (Figures 2A, B; Table S3). The protein association networks of *Airn* and *Kcnq1ot1* were also more similar to *Xist* than to most other chromatin-associated RNAs, ranking 292nd and 1464th, respectively (Figures 2A, B; Table S3).

Next, we examined the community structures within the protein interaction networks of *Airn*, *Kcnq1ot1*, and *Xist* relative to those of other chromatin-associated transcripts, using three orthogonal approaches. First, to quantify the average separation between communities, we examined two distributions derived from the protein interaction networks of *Airn*, *Kcnq1ot1*, and *Xist*: the distributions of Pearson's r values comparing all proteins within communities (intra-community r values) and the distributions comparing all proteins between communities (inter-

community r values). While the average intra-community r values were similar between all three lncRNAs, the average inter-community r values were lowest in *Xist*, followed by *Airn*, then *Kcnq1ot1* (Figure 2C, panel (i); p<2e-16, Wilcoxon rank test*).* Next, we calculated modularity, which is used to assess the extent of partitioning of communities in a network relative to a random-graph null model. Here, we found that the *Xist* network had the highest modularity, followed by *Airn*, then *Kcnq1ot1* (Figure 2C, panel (ii); Table S3). Finally, to assess the degree to which individual nodes within a community are separated from other communities, we calculated the average silhouette coefficients all nodes in each transcript's interaction network (Fowlkes and Mallows 1983). Average silhouette coefficients were higher in *Xist*, followed by *Airn*, followed by *Kcnq1ot1* (Figure 2C, panel (iii); Table S3). These data indicate that for the proteins we examined by RIP-Seq, the communities of association within *Xist* are more separated from each other than those in *Airn*, which are in turn more separated from each other than those in *Kcnq1ot1*. Relative to all other chromatin-associated transcripts, *Xist* and *Airn* exhibited above-average separation distances between communities, and *Xist* could be considered exceptional, with modularity and average silhouette width values that fell in the 99th percentiles. These trends are striking mirrors of the genomic ranges over which each lncRNA represses genes and recruits PRCs in TSCs; whereas *Xist*'s repressive range spans the entire X-chromosome, *Airn*'s spans 15 Mb, and *Kcnq1ot1*'s spans 3 Mb (Schertzer et al. 2019a).

Lastly, we investigated the extent to which intracommunity connections within *Airn*, *Kcnq1ot1*, and *Xist* represent prevalent versus rare events, seeking to determine whether the community structures in each lncRNA exhibited unusual properties relative to other chromatin-associated transcripts. To this end, using network structures from all chromatin-associated transcripts, we calculated *p* values describing the likelihood that each possible pair of proteins (n = 351 pairs from 27 total proteins examined by RIP-Seq) would be detected in the same community relative to randomized controls. By these analyses, protein pairs that were detected more frequently in

the transcriptome than by chance were classified as “prevalent,” and those that were detected less frequently than by chance were classified as “rare”. Of the 351 possible protein pairs, 145 and 191 were classified as prevalent and rare, respectively (p-adj <0.05; Table S4). These analyses showed that *Xist* had a higher number of rare intracommunity edges than many other chromatin-associated RNAs (87th percentile), whereas *Airn* and *Kcnq1ot1* had closer to average (58th) and lower (14th) percentile values, respectively (Figure 2D, panel (i); Table S4). About 30% (*Xist* and *Airn*) and 19% (*Kcnq1ot1*) of the edges within the three lncRNAs were classified as rare (Figure 2D, panel (ii); Table S4). Focusing on connections surrounding the communities in *Airn*, *Kcnq1ot1*, and *Xist* that contain the PRC1 components RING1B and RYBP, we observed rare edges connecting RING1B and RYBP to each of the RBPs HNRNPK, HNRNPL, and HNRNPU (Table S4). In contrast, edges between any combination of HNRNPK, HNRNPL, and HNRNPU were classified as prevalent, as were edges between RING1B and RYBP and the exoribonuclease XRN2, which was found in communities located adjacent to those containing RING1B and RYBP in *Airn* and *Kcnq1ot1* (Table S4). Together, these analyses demonstrate that relative to other chromatin-enriched transcripts, the RIP-Seq association networks of *Airn*, *Kcnq1ot1*, and *Xist* contain both prevalent and rare connected edges, suggesting that the lncRNAs engage with proteins in ways that are common in certain instances and uncommon in others.

Community structures demarcate protein interaction domains with distinct sequence signatures in *Airn*, *Kcnq1ot1*, and *Xist*

While the boundaries of protein-RNA interaction domains have been characterized in *Xist* (Navarro-Cobos and Brown 2025), the boundaries of where putative RNA-protein interaction domains might start and stop within the ~90 kb *Airn* and *Kcnq1ot1* are unknown. However, the communities of correlated protein associations suggest that like *Xist*, *Airn* and *Kcnq1ot1* also

organize protein interactions through underlying RNA domains characterized by distinct sequence signatures. To investigate, we normalized signal-under-peak for each protein examined by RIP-Seq, then created aggregate density profiles of RIP-Seq data merged by their community assignments in 50 nt bins across *Airn*, *Kcnq1ot1*, and *Xist*. These analyses revealed seven above-threshold communities in *Airn*, nine in *Kcnq1ot1*, and six in *Xist*, which together covered 87, 63, and 81% of the lncRNAs' gene bodies, respectively (98% of the spliced *Xist* transcript; Figure 3).

We next sought to determine the extent to which the community demarcations in each lncRNA were associated with distinct sequence signatures. Within each lncRNA, we divided larger community domains into 200 nt segments and used SEEKR to compare the *k*-mer content of segments within each community to segments in all other communities or in "null" segments of each lncRNA (i.e. segments that did not receive a community assignment). At the level of 4-mers, segments from four of the seven communities in *Airn*, from four of the nine communities in *Kcnq1ot1*, and from four of the five communities in *Xist* were significantly more similar to themselves than they were similar to segments from all other communities or the null within each respective lncRNA (Figure 3). Thus, communities of protein-RNA associations coincide with regions of distinct sequence within *Airn*, *Kcnq1ot1*, and *Xist*. Consistent with our observations regarding the extent of protein partitioning within each lncRNA (Figure 2), the sequence signatures underlying communities were more likely to be distinct in *Xist*, followed by *Airn*, then *Kcnq1ot1*.

HNRNPU is required for maintenance of PRC-directed modifications induced by *Airn*, *Kcnq1ot1*, and *Xist*

The ability of *Airn*, *Kcnq1ot1*, and *Xist* to recruit PRCs to chromatin is a central aspect of their function (Schertzer et al. 2019a; Trotman et al. 2025a). Therefore, we were intrigued by the community structure surrounding the PRC1 components RING1B and RYBP in the *Airn*,

Kcnq1ot1, and *Xist* networks (Figure 2A). In addition to the known PRC1-bridging factor HNRNPK, RING1B and RYBP formed connected edges with HNRNPU and XRN2, proteins that also exhibited highly enriched associations with *Airn* and *Kcnq1ot1* (Figures 1A, 2A; (Pintacuda et al. 2017; Colognori et al. 2019; Schertzer et al. 2019a; Trotman et al. 2025a)). HNRNPU is known to be required for the function of *Xist*, promoting its expression and stabilizing its association with chromatin (Hasegawa et al. 2010; Kolpa et al. 2016; Sakaguchi et al. 2016; Creamer et al. 2021; Kolpa et al. 2022; Sharp et al. 2025). However, HNRNPU has not been implicated in the functions of *Airn* or *Kcnq1ot1,* nor has it been directly implicated in PRC recruitment. And while XRN2 is a known *Xist*-associated protein that is important for the accumulation of H3K27me3 in autosomal regions (Bousard et al. 2019; Mattout et al. 2020; Zhou et al. 2022), it has not previously been implicated in the function of *Airn*, *Kcnq1ot1*, or *Xist*.

To determine whether HNRNPU or XRN2 is involved in PRC recruitment by *Airn*, *Kcnq1ot1,* or *Xist*, we depleted each protein in TSCs using doxycycline-inducible CRISPR-Cas9 technology (Schertzer et al. 2019b). We additionally depleted YTHDC1, a protein that exhibited highly enriched association with *Airn*, *Kcnq1ot1* and *Xist* in our RIP-Seq experiments (Figure 1B). YTHDC1 has also been implicated in targeting PRC2 to specific genomic regions (Dou et al. 2023) and has been shown to promote silencing by *Xist* and to promote its abundance (Patil et al. 2016; Trotman et al. 2025b; Wei et al. 2025). As a positive control, we depleted the known PRC1-bridging factor HNRNPK (Schertzer et al. 2019a; Trotman et al. 2025a). As a negative control, we delivered a non-targeting (NTG) sgRNA along with the same doxycycline-inducible Cas9 construct used to deplete the proteins mentioned above (Braceros et al. 2023). Following four days of Cas9 induction by the addition of doxycycline ("(+) dox" condition), the levels of all four proteins were substantially reduced (Figures 4A,B, S7). To examine the extent to which depletion of our target proteins altered the levels of PRC1- and PRC2-directed modifications (H2AK119ub and H3K27me3, respectively) in the genomic regions targeted by *Airn*, *Kcnq1ot1* and *Xist,* we

performed allelic ChIP-Seq. As expected, in cells expressing Cas9 and the NTG sgRNA, H2AK119ub and H3K27me3 levels were unchanged or minimally changed (Figure S4). Also as expected, in cells depleted of HNRNPK, H2AK119ub and H3K27me3 levels were significantly reduced in the *Airn*, *Kcnq1ot1* and *Xist* target regions (Figures 4C, S5, S6). Depletion of XRN2 or YTHDC1 did not significantly reduce H2AK119ub or H3K27me3 levels in the *Airn*, *Kcnq1ot1* or *Xist* target regions (Figure S7). However, depletion of HNRNPU using separate sgRNAs significantly reduced the levels of H2AK119ub and H3K27me3 in the *Airn*, *Xist*, and, to a lesser extent, *Kcnq1ot1* target regions (Figure 4D and S6). Thus, HNRNPU is required to maintain PRC-directed histone modifications not only over the X-chromosome (Hasegawa et al. 2010), but also in the genomic regions targeted for repression by *Airn* and *Kcnq1ot1*.

HNRNPU and HNRNPK depletion derepress genes in the *Airn* and *Xist* but not *Kcnq1ot1* target domains

To determine whether HNRNPU is required to maintain gene repression in the regions targeted by *Airn*, *Kcnq1ot1*, and *Xist*, we performed RNA-Seq in four biological replicates, comparing HNRNPU-depleted to non-depleted TSCs using the CRISPR-Cas9 approach outlined above. In parallel, we performed RNA-Seq experiments comparing HNRNPK-depleted to non-depleted controls. We then examined the extent to which paternal expression bias was altered by HNRNPU and HNRNPK depletion, specifically looking at genes previously identified to be repressed by *Airn* and *Kcnq1ot1* in TSCs (Schertzer et al. 2019a), and at genes that are subject to, weakly, or strongly escape X-chromosome Inactivation in TSCs (defined in this study; Methods). Depletion of HNRNPU led to significant de-repression of *Airn* and *Xist* target genes of all classes but did not affect repression of *Kcnq1ot1* target genes (Figure 5; Table S5). Depletion of HNRNPK led to a more modest derepression of *Airn* target genes and a derepression of genes that strongly escape silencing by *Xist* (Figure 5; Table S5). Therefore, in TSCs, HNRNPU and

HNRNPK are required for the maintenance of gene repression mediated by *Airn* and *Xist*, but not by *Kcnq1ot1*.

HNRNPK but not HNRNPU depletion reduces associations between PRC1 and *Airn*, *Kcnq1ot1*, and *Xist*

We and others have previously found that HNRNPK bridges the association between PRC1 and *Airn*, *Kcnq1ot1*, and *Xist*, and through that bridging mechanism, HNRNPK is required for the accumulation of PRC1- and PRC2-directed modifications in each lncRNA's target region (Pintacuda et al. 2017; Schertzer et al. 2019a; Trotman et al. 2025a). Therefore, we sought to determine whether the requirement for HNRNPU in the maintenance of PRC-directed modifications in the *Airn*, *Kcnq1ot1*, and *Xist* target regions could also be explained by a reduced association between each lncRNA and PRC1. To investigate, we performed RNA immunoprecipitation followed by quantitative PCR and sequencing (RIP-qPCR and RIP-Seq) for the PRC1 component RING1B from formaldehyde-crosslinked TSCs under conditions in which HNRNPK and HNRNPU were depleted, as well as in non-depleted controls. By RIP-qPCR, after normalizing for levels of each lncRNA in the RNA input, we observed that HNRNPK depletion led to significant reductions in the level of RING1B associated with peaked regions in *Airn* and *Xist*, and also resulted in decreased association with *Kcnq1ot1*, as expected (Figure 6A; (Trotman et al. 2025a)). In contrast, although we observed a downward trend in *Airn*, HNRNPU depletion did not significantly reduce the input-normalized levels of association between RING1B and *Airn*, *Kcnq1ot1*, or *Xist* (Figure 6A). Western blots demonstrated that levels of HNRNPK and RING1B were unaltered by HNRNPU depletion (Figure 6B). Examination of RIP-Seq data demonstrated that HNRNPU depletion did not alter the pattern of RING1B association in *Airn*, *Kcnq1ot1*, or *Xist* (Figure 6C). Finally, while the patterns of association of HNRNPK and HNRNPU over *Airn* and *Kcnq1ot1* were highly correlated, they were less correlated over *Xist*, where HNRNPK, but not

HNRNPU, was enriched in clear peaks over the same regions that also exhibited peaked association with RING1B (Figure 6C). We conclude that HNRNPU affects the maintenance of PRC-directed modifications induced by *Airn*, *Kcnq1ot1*, and *Xist* via a mechanism that is at least partly distinct from that of HNRNPK.

HNRNPU depletion reduces the abundance of *Airn* without affecting its localization or half-life

HNRNPU has been shown to be important for promoting the association between chromatin and *Xist* as well as other classes of nuclear-retained RNAs (Hasegawa et al. 2010; Kolpa et al. 2016; Sakaguchi et al. 2016; Creamer et al. 2021; Kolpa et al. 2022; Sharp et al. 2025). We therefore sought to determine the effects of HNRNPU depletion on the localization, abundance, and half-lives of *Airn*, *Kcnq1ot1*, and *Xist*. We used single-molecule sensitivity RNA FISH to examine the localization of all three lncRNAs in HNRNPU-depleted cells and non-depleted controls. Looking first at *Xist*, the primary phenotype we observed was a reduced ability to detect *Xist* RNA FISH signal upon HNRNPU depletion, consistent with observations made in human cells (Kolpa et al. 2016). However, after setting imaging thresholds to enable visualization of the lower *Xist* RNA FISH signal, we did detect dispersal of *Xist* in the HNRNPU-depleted cells, again consistent with prior observations (Figures 7A, S8; (Hasegawa et al. 2010; Kolpa et al. 2016; Sakaguchi et al. 2016; Kolpa et al. 2022; Sharp et al. 2025)). In contrast, while the overall sizes of *Airn* and *Kcnq1ot1* RNA FISH foci were qualitatively smaller in HNRNPU-depleted cells, we detected similar numbers of foci for both *Airn* and *Kcnq1ot1* in both HNRNPU-depleted cells and non-depleted controls, and we did not detect obvious diffusion away from their sites of transcription (Figures 7B, S8). Examination of replicate RNA-Seq data revealed that HNRNPU and HNRNPK depletion significantly reduced the abundance of *Airn* and *Xist,* but not *Kcnq1ot1* (Figure 7C). The patterns of RNA-Seq read density and the half-lives of each lncRNA were

unaffected by the depletion of HNRNPK and HNRNPU (Figures 7D, 7E). Thus, both HNRNPU and HNRNPK are required to maintain wild-type levels of *Airn* and *Xist* without altering their stability, yet the proteins are not required to maintain levels or stability of *Kcnq1ot1*. Moreover, HNRNPU depletion does not lead to obvious dispersal of *Airn* or *Kcnq1ot1* from their sites of transcription.

Discussion

Our study reports a direct comparison of protein association patterns across *Airn*, *Kcnq1ot1*, and *Xist*, three evolutionarily unrelated lncRNAs that each induce long-range Polycomb recruitment and gene repression in a single cell type. Despite their different linear sequences, we identified shared enrichments for several of the same RBPs along the lengths of each lncRNA, supporting the view that they carry out repressive functions using mechanisms that are at least partly shared (Quinodoz et al. 2021; Trotman et al. 2025a).

At the whole-transcript level, relative to other chromatin-enriched RNAs, *Airn*, *Kcnq1ot1*, and *Xist* exhibited highly enriched associations with many proteins. Several proteins that were highly enriched over *Xist*, including CIZ1, HNRNPU, RBM15, PTBP1, SRSF1, and YTHDC1, have been shown to be required for *Xist* biogenesis or its repressive mechanism (Hasegawa et al. 2010; Moindrot et al. 2015; Patil et al. 2016; Ridings-Figueroa et al. 2017; Sunwoo et al. 2017; Pandya-Jones et al. 2020; Trotman et al. 2025b; Wei et al. 2025). Similarly, HNRNPK and HNRNPU, two proteins that we found to be highly enriched over *Airn* and *Kcnq1ot1*, have been shown to be required for repression by *Airn* and *Kcnq1ot1* ((Schertzer et al. 2019a; Trotman et al. 2025a) and this study). By extension, our data suggest that other enriched proteins, including HNRNPM, HNRNPL, PABPN1, and SAFB, or their paralogs, may be important for the biogenesis or silencing mechanism of *Airn*, *Kcnq1ot1*, or *Xist*. Future studies are needed to investigate this hypothesis. Enriched association by RIP-Seq does not indicate a role for a protein in mediating the biogenesis

or function of its associated lncRNA. For example, depletion of XRN2 and YTHDC1 did little to alter lncRNA abundance or PRC-recruitment by *Airn*, *Kcnq1ot1*, and *Xist*, even though both proteins exhibited enriched associations with our lncRNAs of interest. Conversely, a lack of enriched association does not preclude a role for a protein in mediating the biogenesis or function of its partner lncRNA. For example, in our RIP-Seq assays, the known *Xist* co-factor MATR3 did not exhibit peaks of association within *Xist* (Pandya-Jones et al. 2020). In working towards identifying proteins responsible for the regulatory functions of *Airn*, *Kcnq1ot1*, *Xist,* and other chromatin-enriched RNAs, it is important to recognize that many RBPs, including those listed above, harbor paralogs that could compensate for their depletion. For example, HNRNPUL1 has been shown to partially compensate for the role of HNRNPU in *Xist* localization in Neuro2A cells (Sakaguchi et al. 2016). Likewise, in our own preliminary studies, we observed that depletion of HNRNPL in TSCs led to the transcriptional upregulation of its paralog, HNRNPLL (data not shown).

In addition to their overall enriched associations, the patterns of protein association within *Airn*, *Kcnq1ot1,* and *Xist* were strikingly similar, supporting the notion that aspects of repressive mechanism or biogenesis are shared among all three lncRNAs, and between *Airn* and *Kcnq1ot1* especially. *Xist* is known to nucleate the formation of higher-order assemblies in which hundreds of molecules of proteins are thought to surround a single focus of lncRNA; gradients of effector proteins emanating outwards from each assembled focus are then thought to mediate repression by *Xist* on the megabase-scale (Pandya-Jones et al. 2020; Markaki et al. 2021; Jachowicz et al. 2022). Our data are consistent with the notion that *Airn* and *Kcnq1ot1* also nucleate the formation of higher-order protein assemblies to mediate repression, and that many of the proteins contained within the assemblies that surround *Xist* are also included in the assemblies that surround *Airn* and *Kcnq1ot1*. In terms of their biogenesis, *Airn* and *Kcnq1ot1* are both unspliced, which is an unusual property for transcripts of their length. Considering their remarkable similarities in protein

association patterns, it is also conceivable that *Airn* and *Kcnq1ot1* may evade or resist splicing using shared mechanisms.

We observed that the modularity of protein associations over *Airn*, *Kcnq1ot1*, and *Xist* correlated with the genomic range over which the lncRNAs recruit PRCs and repress genes. In TSCs, *Xist* exerts repressive activity over ~165 Mb, *Airn* over ~15 Mb, and *Kcnq1ot1* over ~3 Mb (Schertzer et al. 2019a; Braceros et al. 2023). By several metrics, *Xist* partitioned protein associations to a greater degree than *Airn*, which partitioned interactions to a greater degree than *Kcnq1ot1*. We speculate that the partitioning of protein associations by RIP is an indirect readout of the scale and structure of the protein assembly that surrounds each lncRNA. Under this model, the higher degree of partitioning in *Xist* and *Airn* versus *Kcnq1ot1* may be due to the presence of a greater number of RNA domains in the former lncRNAs that coordinate interactions with multiple proteins simultaneously. Because our RIP-Seq assays relied on formaldehyde as a crosslinking agent, the RNA-protein associations that we detected likely represent a combination of direct and indirect, or bridged, interactions. Future work is needed to distinguish between these possibilities and subsequently determine the mechanisms by which *Airn*, *Kcnq1ot1*, and *Xist* exhibit enriched associations with many of the RBPs that we profiled by RIP-Seq.

Based on (1) its enriched association with *Airn*, *Kcnq1ot1*, and *Xist*, (2) its pattern of association that was correlated with the PRC1 components RING1B and RYBP, and (3) its known requirement for the function of *Xist* (Hasegawa et al. 2010; Kolpa et al. 2016; Sakaguchi et al. 2016), we examined the role of HNRNPU in mediating repression by *Airn*, *Kcnq1ot1*, and *Xist*. Upon depletion of HNRNPU in TSCs, we observed striking reductions in levels of H2AK119ub and H3K27me3 over the inactive X and in genomic regions targeted by *Airn* and *Kcnq1ot1*, indicating that HNRNPU is required to maintain long-range PRC-directed modifications induced by all three lncRNAs. HNRNPU (and HNRNPK) were also required to maintain wild-type levels of gene repression by *Airn* and *Xist*. However, compared to the effects on PRC-directed modifications, the effects on gene repression were weak. This observation leads us to speculate

that the loss of repression upon HNRNPU/HNRNPK depletion is a secondary effect arising from loss of PRC-directed modifications. The reduction in PRC-directed modifications upon HNRNPU depletion occurred without an apparent loss of association between PRC1 and *Airn*, *Kcnq1ot1*, or *Xist*, suggesting that unlike HNRNPK, HNRNPU does not directly bridge the lncRNAs with PRC1 (Trotman et al. 2025a). Based on previous data suggesting HNRNPU depletion causes dispersal of *Xist* from the inactive X, we investigated whether HNRNPU was involved in tethering *Airn* or *Kcnq1ot1* to their sites of transcription (Hasegawa et al. 2010; Kolpa et al. 2016; Sakaguchi et al. 2016). However, HNRNPU depletion did not cause *Airn* or *Kcnq1ot1* to disperse from their sites of transcription. Likewise, HNRNPU depletion reduced the abundance of *Airn* and *Xist* without altering their stability. Interestingly, HNRNPK depletion also reduced abundance of *Airn* and *Xist* to a similar extent, suggesting a potential connection between the two RBPs.

Considered together, our findings suggest a revised model for the role of HNRNPU in long-range gene regulation mediated by chromatin-associated RNAs. HNRNPU is known to play an important architectural role in the nucleus, helping to keep transcribed regions in an accessible state by forming what has been proposed to be a dynamic, mesh-like network in concert with other chromatin-associated RNAs (Hasegawa et al. 2010; Sakaguchi et al. 2016; Nozawa et al. 2017; Creamer et al. 2021; Kolpa et al. 2022; Sharp et al. 2025). Thus, rather than serving purely as a tether of lncRNAs to chromatin, our data support the view that HNRNPU serves as an architectural factor that enables the induction of long-range PRC-directed modifications by lncRNAs via at least two possible mechanisms. On the one hand, via its ability to form mesh-like networks, HNRNPU may help to keep distal regions of chromatin accessible to the components of lncRNA-ribonucleoprotein assemblies that recruit PRCs and repress genes; this same mesh-like network may be responsible for keeping *Xist* tethered to chromatin around the inactive X-chromosome (Hasegawa et al. 2010; Sakaguchi et al. 2016; Nozawa et al. 2017; Creamer et al. 2021; Kolpa et al. 2022; Sharp et al. 2025). On the other hand, HNRNPU depletion reduced the abundance of *Airn* and *Xist* without altering their half-lives or causing *Airn* to disperse from

chromatin; and by RIP-Seq, HNRNPU was one of the most highly enriched proteins over *Airn*, *Kcnq1ot1*, and *Xist*, nominating it as a component of each lncRNA's ribonucleoprotein assembly. Thus, in addition to facilitating long-range Polycomb recruitment by maintaining an accessible chromatin architecture, HNRNPU may play an architectural role locally, by associating with nascent RNA to maintain appropriate expression levels of the *Airn* and *Xist* lncRNAs in what might otherwise be a naturally repressive environment.

Lastly, beyond *Airn*, *Kcnq1ot1*, and *Xist*, our study provides perspectives on the possible spectrum of chromatin-associated RNAs that might harbor an ability to regulate gene expression or chromatin state *in cis*. The transcripts most similar to *Airn*, *Kcnq1ot1*, and *Xist* in terms of protein association profiles were produced from both lncRNA and protein-coding genes. *Airn* and *Kcnq1ot1*, which are both unspliced, harbored protein association profiles that most frequently resembled other intron-containing, presumably nascent transcripts. In contrast, *Xist*, which is robustly spliced, harbored a protein association profile that most frequently resembled other intron-excluding, presumably spliced transcripts. Splicing often occurs co-transcriptionally and leads to the rapid nuclear export of processed RNA; however, recent studies suggest that post-transcriptional splicing or delayed export after splicing are common events (Mukherjee et al. 2017; Neugebauer 2019; Bedi et al. 2021; Reimer et al. 2021; Yeom et al. 2021; Ntini et al. 2023; Cote et al. 2024; Ietswaart et al. 2024; Choquet et al. 2025). Thus, it is conceivable that in addition to lncRNAs, RNAs produced from protein-coding genes might carry out lncRNA-like regulatory events prior to their clearance from chromatin. Indeed, the *TTN* pre-mRNA can form an architectural hub in the nucleus that serves to regulate gene expression at multiple levels (Kania et al. 2024). The extent to which other chromatin-associated RNAs might carry out *cis*-regulatory activities remains unclear. Our work implies that *Airn*, *Kcnq1ot1*, and *Xist* are exceptional in many respects. Accordingly, we would presume that in TSCs, any additional RNAs capable of *cis* repression would operate within genomic intervals smaller than those targeted by *Airn*, *Kcnq1ot1*, or *Xist*. Future studies are needed to investigate this hypothesis and the extent to which transcripts

produced from lncRNAs and protein-coding genes represent a source of regulatory RNAs in the mammalian transcriptome.

Acknowledgements

This work was supported by NIH grants R01GM136819, R01GM121806, and R35GM153293; NSF grant DBI-2228805; and the Yang Family Biomedical Scholar Fund (J.M.C.). M.M.M. was supported by T32GM119999 and F31HD111292. R.E.C. was supported by T32GM007092 and F31HD103334.

Author Contributions

M.M.M., S.L., and J.M.C. conceived the study; M.M.M., S.L., S.P.B., R.E.C., Z.Z., and J.M.C. performed the experiments and/or analyzed data; M.M.M. and J.M.C. acquired funding; and M.M.M. and J.M.C. wrote the paper with feedback from co-authors.

Declaration of Interests

The authors declare no competing interests.

Methods

TSC culture

The C/B TSC line used in this study was derived in (Calabrese et al. 2012). TSCs were cultured as in (Quinn et al. 2006). Briefly, TSCs were cultured on gelatin-coated, pre-plated irradiated mouse embryonic fibroblast (irMEF) feeder cells in TSC media (RPMI [Gibco 11875093], 20% qualified FBS [Gibco 26140079], 0.1 mM penicillin-streptomycin [Gibco 15140122], 1 mM sodium pyruvate [Gibco 11360070], 2 mM L-glutamine [Gibco 25030081], 100 μM β-mercaptoethanol [Sigma-Aldrich 63689]) supplemented with 25 ng/mL FGF4 (Gibco PHG0154) and 1 μg/mL Heparin (Sigma-Aldrich H3149) just before use, at 37°C in a humidified incubator at 5% CO_2. At passage, TSCs were trypsinized with 0.125% trypsin-EDTA in PBS solution (Gibco 25200-072) for ~4 min at room temperature and gently dislodged from the plate with a sterile, cotton-plugged Pasteur pipette. To deplete irMEFs from TSCs prior to all harvests, TSCs were pre-plated for 45 min at 37°C on a nongelatinized plate, which was then rinsed twice with a serological pipette, and the solution was transferred to a fresh culture plate. MEF-depleted TSCs were then cultured for three to four days in 70% irMEF-conditioned TSC media supplemented with growth factors prior to harvesting for molecular or genomic assays.

Generation of Cas9/sgRNA-expressing TSCs

sgRNAs targeting HNRNPK, HNRNPU, XRN2, and YTHDC1 were designed using the Benchling CRISPR Guide RNA Design tool then cloned into BbsI-digested rtTA-BsmbI piggyBac vector (Addgene plasmid #126028; (Schertzer et al. 2019b)) using NEB Quick Ligase (NEB M2200S). HNRNPK sgRNAs are from (Schertzer et al. 2019a). The NTG sgRNA is from (Braceros et al. 2023). Plasmids were purified from High-Efficiency 5-alpha competent cells (NEB C2987H) using the PureLink HiPure Plasmid Midiprep Kit (Invitrogen K210004). DNA was quantified via Qubit 2.0 fluorometer (dsDNA broad-range kit, Thermo Fisher Q32853). Cloned sgRNA plasmids

are in the process of being deposited in Addgene. Individual or pooled sgRNA vectors were mixed with the TRE-Cas9-Cargo vector (Addgene plasmid #126029; (Schertzer et al. 2019b)) and the piggyBac transposase vector from (Kirk et al. 2018) at a 1:1:1 ratio in a total of 10 µl. The 10-µL (2.5 ug) plasmid mixtures were mixed with 1 million WT TSCs suspended in 100 µL Neon Buffer R (Invitrogen MPK10025). TSCs were then electroporated using a Neon Transfection System (Invitrogen) with two 30-ms pulse of 950 V (program 5) before seeding onto a well of a 6-well of gamma-irradiated drug-resistant (DR4) irMEF feeder cells (ATCC SCRC-1045) in TSC growth medium lacking penicillin-streptomycin. Media was changed the following morning with 70% conditioned media and growth factors. Starting 24 hours after electroporation, cells were selected with G418 (200 ug/ml; GIBCO) and Hygromycin B (150 ug/ml; GIBCO) for 9 days. TSCs were split every 2-3 days, and media was changed daily with 70% conditioned media, growth factors and antibiotics. Following 9 days of selection, TSCs were then split on to irMEFs with standard TSM and growth factors for expansion.

Western blots

TSCs were harvested by washing twice with 3 mL ice-cold PBS, scraping into 1 mL ice-cold PBS, and transferring to 1.7-mL tubes. Following centrifugation (1200 × g, 10 min, 4°C), supernatants were removed and cell pellets stored at -80°C. Cell pellets were thawed on ice and lysed by resuspension in 75-150 µL ice-cold RIPA buffer (10 mM Tris-HCl pH 7.5, 140 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 1% [v/v] IGEPAL CA-630, 0.1% [w/v] sodium deoxycholate, 0.1% [w/v] SDS) supplemented with fresh 1 mM PMSF and 1:100 protease inhibitor cocktail (Sigma P8340-5ML). Cells were thoroughly resuspended using 10 up-down strokes with a P1000 pipette, incubated with rocking for 30 minutes at 4 C, then sonicated on ice using a Sonics Vibracell VCX130 probe-tip sonicator at 30% amplitude (two 10-second pulses with 1-minute rest on ice). Cell debris was removed by centrifugation (16,100 × g, 15 min, 4°C), and cleared lysates were

transferred to pre-chilled tubes. Protein concentrations were determined using the Bio-Rad DC Protein Assay (Bio-Rad 5000112) and a standard curve of BSA.

Lysates were diluted 3:1 with 4× Laemmli sample buffer (Bio-Rad 1610747) containing 10% (v/v) β-mercaptoethanol. Equal amounts of protein (20 µg) were loaded onto pre-cast polyacrylamide SDS-PAGE gels (4-20% polyacrylamide, Bio-Rad) selected based on target protein molecular weights and ran in Tris-glycine-SDS running buffer (2.5 mM Tris, 19.2 mM glycine, 0.01% [w/v] SDS, pH 8.3) at 130 V at room temperature. Proteins were transferred to methanol-activated Immobilon-P PVDF membranes (Millipore Sigma IPVH00010), in Tris-glycine-methanol-SDS transfer buffer (2.5 mM Tris, 19.2 mM glycine, 20% [v/v] methanol, 0.001% [w/v] SDS, pH 8.3) at 4°C, either for 1 hour at 100 V or overnight at 20 V. Membranes were blocked for 30 minutes at room temperature with orbital shaking in 1x TBS with 1% Casein blocking buffer (Bio-Rad #1610782). Primary antibodies were applied at concentrations listed in Table S6 and incubated overnight at 4°C with gentle rocking. After washing (one quick TBS-T rinse followed by three 10-minute washes in TBS-T with orbital shaking), membranes were incubated with HRP-conjugated secondary antibodies (1:20,000 dilution, goat anti-mouse HRP (Thermo Fisher A16072); or 1:100,000 dilution, goat anti-rabbit HRP (Thermo Fisher G-21234); or 1:100,000 dilution, rabbit anti-goat HRP (Thermo Fisher 31402)) for 45 minutes at room temperature. Following a final round of TBS-T washes, protein bands were visualized using SuperSignal West Femto ECL substrate (Thermo Fisher 34094) and imaged on a Bio-Rad ChemiDoc system in chemiluminescence mode. Western blots assessing the extent of depletion of protein targets were performed for each (-) and (+) dox replicate in this study.

RIP-Seq

Prior to RIP, TSCs were grown to 75-85% confluency for one passage off irMEFs, trypsinized, and counted. TSCs were washed twice in cold 1xPBS and then rotated for 30 min in 10 mL of 0.3% formaldehyde (1 mL 16% methanol-free formaldehyde (Pierce, #28906) in 49 mL 1xPBS)

at 4°C. Formaldehyde was quenched with 1.2 mL of 2 M glycine for 5 min at room temperature with rotation. TSCs were washed 3x in cold 1xPBS, then resuspended in 1xPBS at 10 million cells per mL, aliquoted to 10 million cells per 1.7-mL tube and spun down. PBS was aspirated, and the pellets were snap frozen in a dry ice methanol bath and immediately transferred to -80°C. RIPs were performed as in (Trotman et al. 2025a). Protein A/G agarose beads (25 µL; Santa Cruz sc-2003) were washed three times in blocking buffer (0.5% BSA in 1xPBS). The washed beads were then resuspended in 300 µl of blocking buffer, and 10 µg of antibody was added (see Table S6 for exceptions). The beads and antibodies were then rotated overnight at 4°C. The next day, 10 million cells were resuspended in 500 µL RIPA Buffer (50 mM Tris-HCl, pH 8.0, 1% Triton X-100, 0.5% sodium deoxycholate, 0.1% SDS, 5 mM EDTA, 150 mM KCl) supplemented with 1:100 Protease Inhibitor Cocktail (PIC; Sigma Product #P8340), 2.5 µL SUPERase-In (Thermo Fisher Scientific AM2696) and 0.5 mM DTT and sonicated twice for 30 s on and 1 min off at 30% output using the Sonics Vibracell Sonicator (Model VCX130, Serial# 52223R). Cell lysates were then centrifuged at 15,000 × g for 15 min. Supernatants were transferred to a new tube and diluted 1:2 with 500ul in 500 µL fRIP Buffer (25 mM Tris-HCl pH 7.5, 5 mM EDTA, 0.5% NP-40, 150 mM KCl) supplemented as above with PIC, SUPERase-In, and DTT. The total lysate (25 µL) was saved for input. Antibody-bound beads were washed three times with 1 mL fRIP Buffer. To the washed beads, 500 µl of diluted cell lysate (5 million cell equivalents) was added. The beads and lysate were then rotated overnight at 4°C. The beads were then washed once with 1 mL fRIP Buffer and resuspended in 1 mL PolII ChIP Buffer (50 mM Tris-HCl pH 7.5, 140 mM NaCl, 1 mM EDTA, 1 mM EGTA, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% sodium dodecyl sulfate) before transferring to a new 1.7-mL tube. Samples were rotated at 4°C for five minutes, spun down at 1200xg, and the supernatant was aspirated. Samples were washed twice more with 1 mL PolII ChIP Buffer, once with 1 mL High Salt ChIP Buffer (50 mM Tris-HCl pH 7.5, 500 mM NaCl, 1 mM EDTA, 1 mM EGTA, 0.1% sodium deoxycholate, 0.1% sodium dodecylsulfate, 1% Triton X-100), and once in 1 mL LiCl Buffer (20 mM Tris pH 8.0, 1 mM EDTA, 250 mM LiCl, 0.5% NP-40, 0.5%

sodium deoxycholate); each wash included a five-minute rotation at 4°C. At the final wash, samples were transferred to a new 1.7-mL tube. After the final wash, inputs were thawed on ice and bead samples were resuspended in 100 µL of 1x reverse crosslinking buffer (1x PBS, 2% N-lauroylsarcosine, 10 mM EDTA) supplemented with 1 µL SUPERase-In, 20 µL Proteinase K, and 0.5 mM DTT). Samples were incubated for 1 h at 42°C, 1 h at 55°C, and 30 min at 65°C, and mixed by pipetting every 15 min. To the bead and buffer mixture, 1 mL TRIzol was added, vortexed, and 200 µL of $CHCl_3$ was added. Samples were vortexed and spun at 12,000 × *g* for 15 min at 4°C. The aqueous phase was extracted, and one volume of 100% ethanol was added. Samples were vortexed and applied to Zymo-Spin IC Columns (Zymo #R1013) and spun for 30 s at top speed on a benchtop microcentrifuge. Next, 400 µL of RNA Wash Buffer (Zymo #R1013) was added, and the samples were spun at top speed for 30 s. For each sample, 5 µL DNase I and 35 µL of DNA Digestion Buffer (Zymo #R1013) was added directly to the column matrix and incubated at room temperature for 20 min. Next, 400 µL of RNA Prep Buffer (Zymo #R1013) was added, and the columns were spun at top speed for 30 s. Next, 700 µL RNA Wash Buffer (Zymo #R1013) was added, and the columns were spun at top speed for 30 s. Next, 400 µL RNA Wash Buffer was added, and the columns were spun at top speed for 30 s. The flow-through was discarded, and the columns were spun again for 2 min to remove all traces of the wash buffer. Columns were transferred to a clean 1.7-mL tube, 15 µL of ddH2O was added to each column, and after a five-minute incubation, samples were spun at top speed to elute. 9 µL from each RIP-Seq replicate was mixed with 1 µL of a 1:250 dilution of ERCC RNA spike-in controls (Thermo Fisher 4456740) and subjected to strand-specific RNA-Seq using the KAPA RNA HyperPrep Kit with RiboErase (HMR) (Roche 08098140702) and KAPA Unique Dual-Indexed Adapters (Roche 08861919702) or TruSeq adapters. RNA fragmentation was performed at 85°C for 5 min. Libraries were quantified using a Qubit 2.0 fluorometer (dsDNA broad sensitivity kit, Thermo Fisher Q32850), pooled, and sequenced on Illumina NextSeq500 or Illumina NextSeq1000 platforms.

ChIP-Seq

ChIP-Seq experiments examining effects of protein depletion on chromatin modifications were performed after four days of addition of 1000 ng/mL of doxycycline (Sigma D9891) to sgRNA-expressing cells. (-) and (+) dox TSC cultures were crosslinked on the same day for each replicate performed. Prior to crosslinking for ChIP, TSCs were passaged once off irMEFs. Adhered TSCs were crosslinked with 0.6% formaldehyde (Fisher Scientific, cat #: BP531-500) in RPMI media with 10% FBS for 10 min at room temperature, then quenched with 125 mM glycine for 5 min at room temperature. Crosslinked TSCs were then washed twice with ice-cold PBS and scraped with ice-cold PBS containing 0.05% Tween (Fisher Scientific, cat #: EW-88065-31), PMSF, and PIC (Sigma Aldrich, cat #: P8340). TSCs were then spun at 3,000 × g for 5 min at 4 °C to remove PBS, followed by resuspension in ice-cold PBS with PIC and PMSF and divided into 5-million cell aliquots. All ChIPs were performed using 5 million cells, 5 µL of antibody, and 30 uL of Protein A/G agarose beads (Santa Cruz, cat #: sc-2003). Antibody-conjugated beads were prepared by incubating antibody with beads in 300 µL Blocking Buffer (PBS, 0.5% BSA [Invitrogen, cat #: AM2616]) overnight at 4 °C with rotation. Crosslinked TSCs were resuspended in 1 mL Lysis Buffer 1 (50 mM HEPES pH 7.5, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% NP-40, 0.25% Triton X-100, PIC) and incubated with rotation for 10 min at 4 °C. Cells were then resuspended in 1 mL Lysis Buffer 2 (10 mM Tris-HCl pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, PIC) for 10 min at room temperature. All buffer removal steps were performed with 5-min 1,200 × g spins at 4 °C. The extracted nuclei pellet was then resuspended and sonicated in 500 µL Lysis Buffer 3 (10 mM Tris-HCl pH 8.0, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% sodium-deoxycholate, 0.5% N- lauroylsarcosine, PIC) using a Vibra-Cell VX130 (Sonics) with the following parameters: 8–10 cycles of 30% intensity for 30 s with 1 min of rest on ice between cycles to obtain 100-500 bp fragments. Lysates were then spun for 30 min max speed at 4 °C. The detergent compatible protein assay kit (Biorad 5000113) was used along with BSA protein

standards to determine protein quantity of lysates, as well as sonicated chromatin from HEK293T cells crosslinked as above with 0.6% formaldehyde. For each ChIP in each sgRNA genotype, equal amounts of TSC lysate by protein content comparing the (-) and (+) dox conditions was retained. 5% of the standardized TSC lysate amount was then removed, and an equivalent quantity sonicated chromatin from HEK293T cells was added as a spike in control. Lysate volumes were brought to 500ul, Triton X-100 was then added to a final concentration of 1%, and 25 µL was removed to serve as input. Lysates were mixed with pre-conjugated antibody bead mixes and incubated overnight at 4 °C with rotation. The next day, the beads were washed five times with 1 mL RIPA Buffer (50 mM HEPES pH 7.5, 500 mM LiCl, 1 mM EDTA, 1% NP-40, 0.7% sodium deoxycholate, PIC) and once with 1 mL TE, each for 5 min at 4 °C with rotation and spun at 2,000 × g for 2 min for buffer removal. The washed beads were then resuspended in elution buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA, 1% SDS) and placed on a 65 °C heat block for 17 min with frequent vortexing. ChIP DNA was then reverse crosslinked in 0.5% SDS and 100mM NaCl overnight at 65 °C, followed by a 1 h RNaseA (3 µL; Thermo Scientific, cat #: EN0531) treatment at 37 °C and a 2.5 h Proteinase K (10 µL; Invitrogen, cat #: 25530015) treatment at 56 °C. DNA was then extracted with 1 volume of phenol:chloroform:isoamyl alcohol (Sigma-Aldrich, cat #: P3803) and precipitated with 2 volumes 100% ethanol, 1/10 volume 3M sodium-acetate pH 5.4, and 1/1000 volume linear acrylamide (Invitrogen, cat #: AM9520) overnight at -20 °C. DNA was then precipitated with a 30 min max speed spin at 4 °C, washed once with ice-cold 80% ethanol, and resuspended in TE buffer. ChIP-seq libraries were prepared with NEBNext End Repair Module (NEB, cat #: E6050S), A-tailing by Klenow Fragment (3'/50 exo-; NEB, cat #: M0212S), and TruSeq 6-bp index adaptor ligation by Quick ligase (NEB, cat #: M2200S), and NEBNext High-Fidelity 2X PCR Master Mix (NEB, cat #: M0541S). All DNA size-selection purification steps were performed using AMPure XP beads (Beckman Coulter, cat #: A63880). Single-end, 75- and 100-bp sequencing was performed using Illumina NextSeq 500 and 1000 systems, respectively.

Total RNA preparation

irMEF feeder-depleted TSCs were washed twice with ice-cold PBS and harvested with 1 mL TRIzol (Thermo Fisher 15596026). Total RNA was prepared via standard TRIzol-chloroform extraction, per manufacturer's instructions, except for the addition of 4 µL 5 µg/µL linear acrylamide (Thermo Fisher AM9520) to promote precipitation in isopropanol. RNase-free water was added to dried RNA pellets, allowed to incubate overnight at 4°C, and resuspended. RNA concentrations were measured via Qubit 2.0 fluorometer (RNA high sensitivity kit, Thermo Fisher Q32852), and integrity of RNA was confirmed by visualizing rRNA bands on an agarose gel.

Total RNA-Seq

9 µL of 100ng/ul solution of TSC RNA was mixed with 1 µL of a 1:250 dilution of ERCC RNA spike-in controls (Thermo Fisher 4456740), and subject to strand-specific RNA-Seq using the KAPA RNA HyperPrep Kit with RiboErase (HMR) (Roche 08098140702) and KAPA Unique Dual-Indexed Adapters (Roche 08861919702) or TruSeq adapters. RNA fragmentation was performed at 94°C for 6 min, diluted adapter stocks were 7 µM, and libraries were amplified using 11 cycles of PCR. Libraries were quantified using a Qubit 2.0 fluorometer (dsDNA broad sensitivity kit, Thermo Fisher Q32850), pooled, and sequenced on Illumina NextSeq500 or Illumina NextSeq1000 platforms.

RIP-qPCR

RIP-qPCR was performed as described above for RIP-Seq, with some exceptions. Crosslinking was performed as described above for RIP-Seq. Protein A/G agarose beads were prepared as described, except that all bead centrifugation steps in this RIP-qPCR protocol were 2000 × g for 1 min; 5 µL of 219 ng/µL anti-RING1B antibody (Cell Signaling 5694S) was used for each RIP. 10-million-cell pellets were thawed on ice, resuspended in 508 µL RIPA Buffer

containing PIC, SUPERase-In, and DTT, and sonicated and centrifuged as described. A total of 510 μL of cleared lysate was transferred to new tubes on ice, to which 510 μL fRIP Buffer containing PIC, SUPERase-In, and DTT was added and mixed well. 25 μL of this mixture was removed and stored at -20 °C for later processing as a 5% input sample. Antibody-bound beads were washed twice with 1 mL fRIP Buffer before adding 490 μL lysate mixture (~5 million cell equivalents) and 490 μL of a 1:1 mixture of RIPA Buffer and fRIP Buffer containing PIC, SUPERase-In, and DTT. Overnight bead-lysate incubation and wash steps were performed as previously described. The input-sample Reverse Crosslinking Buffer was prepared on ice by mixing, for each sample, 33 μL 3X reverse crosslinking buffer (3x PBS, 6% [w/v] N-lauroylsarcosine, 30 mM EDTA), 41 μL RNase-free water, 5 μL 100 mM DTT, 1 μL 20 U/μL SUPERase-In, and 10 μL 20 mg/mL proteinase K (Thermo Fisher 25530049). Bead-sample Reverse Crosslinking Buffer was prepared on ice by mixing, for each sample, 33 μL 3X reverse crosslinking buffer (3x PBS, 6% [w/v] N-lauroylsarcosine, 30 mM EDTA), 66 μL RNase-free water, 5 μL 100 mM DTT, 1 μL 20 U/μL SUPERase-In, and 10 μL 20 mg/mL proteinase K. The 25-μL 5% input samples were thawed on ice and mixed thoroughly by pipetting with 90 μL input-sample reverse crosslinking buffer. To washed beads, 115 μL bead-sample reverse crosslinking buffer was added forcefully to resuspend the beads, but the mixture was not otherwise mixed. Samples were heated as described to reverse crosslinks; bead-containing samples had their beads resuspended every 30 min by forcefully ejecting ~100 μL of the supernatant onto the settled beads to avoid losing beads to the inside of pipet tips. After heating, RNA was purified with TRIzol and the RNA Clean & Concentrator-5 kit (Zymo R1013) as described above, during which 580 μL of the TRIzol/chloroform aqueous phase was mixed with 580 μL of 100% ethanol.

Reverse transcription was performed using the Applied Biosystems High-Capacity cDNA Reverse Transcription Kit (Thermo Fisher 4368814) to assemble 20-μL random-primed reverse transcription reactions in 0.2-mL PCR strip-tubes on ice, each containing 5 μL of 5% input or 100% RING1B IP RNA sample and 0.5 μL 40 U/μL RNaseOUT RNase Inhibitor (Thermo Fisher

10777019). The reverse transcription was run with the following thermocycler parameters: 25°C for 10 min, 37°C for 120 min, 85°C for 5 min, hold at 4°C. A six-standard curve of 4-fold serial dilutions was constructed from each 5% (-) dox input sample: 8 µL of the previous standard was mixed with 24 µL of nuclease-free water. For each qPCR primer pair, a qPCR master mix was generated, containing for each 10-µL qPCR reaction: 5 µL iTaq Universal SYBR Green Supermix (Bio-Rad 1725124), 3 µL nuclease-free water, 0.5 µL 10 µM forward primer, and 0.5 µL 10 µM reverse primer (see Table S6 for oligonucleotide sequences). Technical-triplicate qPCR was performed with 9 µL qPCR master mix and 1 µL of standard-curve or 2-fold-diluted reverse-transcription product. A standard curve was run on every plate for each primer pair. Plates were firmly sealed with Microseal 'B' PCR Plate Sealing Film (Bio-Rad MSB1001), mixed several times by inversion and low intensity vortexing, and briefly pulsed down in a centrifuge. Reactions were run on a Bio-Rad C1000 Touch Thermal Cycler equipped with a CFX96 Real-Time System with the following program: 95°C for 10 min, 40 cycles of (95°C for 15 s, 60°C for 30 s, 72°C for 30 s, plateread). Using Bio-Rad CFX manager, first setting “Cq Determination Mode” to “Regression” and then using each primer pair’s standard curve (Cq vs. starting quantity) to convert Cq values to starting quantities for each technical replicate, the percent IP/input value for each RIP sample qPCR triplicate was calculated. qPCR technical triplicates were averaged for each biological replicate and plotted as a single dot in GraphPad Prism. Significance was determined via one-tailed t-test comparing the average of triplicate-averaged biological replicate values of (-) and (+) dox conditions.

Single-molecule RNA-FISH

Stellaris RNA FISH probes were designed against *Kcnq1ot1* utilizing the Stellaris RNA FISH Probe Designer (LGC, Biosearch Technologies, Petaluma, CA) available online at www.biosearchtech.com/stellarisdesigner (version 4.2; parameters: masking level 5; max number of 48 probes; oligo length 20 nt; min. spacing length 2 nt; Quasar 670). *Kcnq1ot1* probe

sequences are listed in Table S6. Probes to *Airn* were designed in (Trotman et al., 2025). Probes to *Xist* were designed by Stellaris (product # SMF-3011-1, Quasar 570). TSCs were hybridized with the Stellaris RNA FISH probes following the manufacturer's instructions for adherent cells, available online at www.biosearchtech.com/stellarisprotocols. Briefly, following four days of culture in 70% conditioned media with growth factors on coverslips, and with the addition of 1000 ng/mL of doxycycline (Sigma D9891), or no doxycycline, sgRNA-expressing cells were fixed with 3.7% formaldehyde and permeabilized overnight in 70% ethanol. Cells were then incubated in Wash Buffer A with formamide for five minutes at room temperature, before hybridization with 125 nM of *Airn/Kcnq1ot1* or *Xist* probes in hybridization buffer at 37°C for 4 hours, then washed for 30 minutes at 37°C with Wash Buffer A. Cells were then incubated with Wash Buffer A consisting of 5 ng/mL DAPI for 30 minutes at 37°C. Cells were then washed for 5 minutes at room temperature once with Wash Buffer B, then fixed with ProLong™ Glass Antifade Mountant (Thermo Fisher P36982) for 48 hours at room temperature in the dark.

Two biological replicates per (-) dox and (+) dox condition were imaged. Images were taken with a Leica DMi8 inverted brightfield microscope, using Leica Application Suite X software version 3.7.5.24914 and the 63x/1.6 objective (pixel size .104 uM). Fields of cells were chosen by scanning for areas with flat easily identifiable individual nuclei. Acquisition setting for each channel (DAPI, 570, 670) were optimized to avoid saturated pixels (low florescence intensity, and exposure time with log histogram signal not piling up on the right) using the (-) dox sample and kept consistent for the (+) dox sample of that replicate. For each condition, 3-5 Z-stack images with a Z-slice of size 0.2 µm were acquired for each channel sequentially, starting with the highest wavelength channel and moving to the lowest wavelength channel to avoid photobleaching. Z-stacks were set by identifying the limits of the Z position where cells went out of focus. Images were deconvolved with Huygens Essential version 20.04.0p3 64b (Scientific Volume Imaging, The Netherlands, http://svi.nl) using the standard deconvolution profile under batch express. For representative images, Z-series are displayed as maximum z-projections, and (-) dox and (+) dox

images were simultaneously adjusted in FIJI (version 2.16.0/1.54p) to have the same threshold, brightness and contrast.

Half-life Analysis

RNA half-lives were measured using a timecourse of flavopiridol (Alvocidib) (Selleck chemicals S1230). MEF-depleted Cas9/sgRNA-expressing TSCs were induced for 3 days with 1000 ng/mL doxycycline prior to start of the flavopiridol time course, for a total of 4 days. TSCs were then treated with 1 uM flavopiridol (24 µl of 50 ug/ml flavopiridol in DMSO added to 3 mL of media) for 30 minutes, 1 hour, 3 hours, 8 hours, (or without flavopiridol, “0 h”) and lysed with TRIzol. To measure levels of *Airn*, *Kcnq1ot1*, *Xist*, and *Gapdh* at each time point, RNA was extracted and subjected to RT-qPCR as above, with the following exceptions: cDNA was diluted 1:20 in nuclease-free water and qPCR was carried out with 4 µL of diluted cDNA and 6 µL of master mix, composed of 5 µL iTaq Universal SYBR Green Supermix (Bio-Rad 1725124) and 1 µL of a 5 µM primer mix per well (see Table S6 for oligonucleotide sequences). RNA levels were normalized to *Gapdh* at each time point and calculated as the percentage of RNA relative to the 0 h time point. For (-) dox HNRNPU depletion cells, only one biological replicate was used, as the 0 h sample was compromised. To estimate half-life, *Gapdh*-normalized data for each replicate were averaged and fit to a non-linear one-phase decay model in GraphPad Prism using the equation $Y=(Y0)*\exp(-K*X)$. Error bars represent the standard deviation between biological replicates.

Calibrated Allelic ChIP-Seq analysis

ChIP-Seq and Input reads were aligned to a custom genome consisting of concatenated mm10 and hg38 genomes, and independently, also aligned to a concatenated genomes consisting of a version of mm10 that had been modified to incorporate CAST single-nucleotide polymorphisms (SNPs) downloaded from the Sanger Mouse Genomes Project on 7/30/2020

(Keane et al. 2011) as well as the hg38 genome. Alignments were performed using Bowtie2 with default parameters (Langmead and Salzberg 2012). Aligned reads that had a MAPQ greater than or equal to 30 were extracted with SAMtools (Li et al. 2009). Reads aligned to either human or mouse in both genomes were counted. Then mouse reads were retained and human reads removed from both B6 and CAST alignments. Allele-specific read retention (i.e., reads that overlap at least one B6 or CAST SNP) was performed as in (Calabrese et al. 2012; Calabrese et al. 2015) using a custom Perl script (intersect_reads_snps18.pl). For tiling density plots, B6- or CAST-specific reads were summed in 10 kb using a custom perl script (ase_analyzer8_hDbed.pl). To normalize to human spike-in reads, binned counts of the ChIP were then divided by a "ori ratio" as in (Hu et al. 2015). Ori ratios were calculated in two steps: first, multiplying the spike-in read ratio (human/mouse) from input samples by the target read ratio (mouse/human) from IP samples; second, normalizing to the smallest value between (-) dox and (+) dox conditions. Alternatively, binned counts were divided by the total number of mm10 aligned reads in the dataset and multiplied by a million (i.e., RPM). Normalized counts were then divided by the number of B6/CAST SNPs detected in the bin genomic coordinates (i.e., SNP-norm RPM). Finally, bins were averaged every 9 bins in 1bin increments. For tiling plots with multiple replicates, bins were then averaged among samples. For box plots showing difference in B6 and CAST alleles, bins were averaged every 9 bins in 1bin increments for each allele and CAST reads were subtracted from B6 reads in each bin. The B6-CAST value for bins within the lncRNA target regions identified in (Schertzer et al. 2019b) were then plotted as box plots for individual replicates. Statistical significance of differences for each replicate between the (-) and (+) dox conditions was determined using one-sided Student's t-tests. All plots were generated using ggplot2 (Wickham 2016) in RStudio.

Allelic RNA-seq analysis

Allelic RNA-seq analysis was performed essentially as described (Trotman et al. 2025a). RNA-seq reads were aligned to mm10, and independently, aligned to a version of mm10 (Perez et al. 2025) that had been modified to incorporate CAST single-nucleotide polymorphisms (SNPs) downloaded from the Sanger Mouse Genomes Project on 7/30/2020 (Keane et al. 2011). Alignment was performed with STAR (v2.7.9a; (Dobin et al. 2013)), using multi-sample two-pass mapping and the option "--outFilterMultimapNmax 1" to consider only uniquely mapping reads. Using a custom Perl script (intersect_reads_snps16.pl) aligned reads were parsed to identify reads clearly originating from either the B6 or CAST allele (i.e. reads that overlap at least one B6 or CAST SNP). Reads marked as either B6 or CAST were then assigned to genes using a custom Perl script (ase_analyzer10.pl) and the GTF file gencode.vM25.basic.annotation.gtf (Mudge et al. 2025). The ratio of B6/(B6+CAST) reads was calculated for each gene from each sample. The *Airn*- and *Kcnq1ot1*-target genes analyzed in this study were reported in Figure 1 of (Schertzer et al. 2019a). Lists of different X-linked genes analyzed in this study (inactivated, weak escapees, strong escapees) were defined as follows: we required that each gene was represented by an average of at least 5 (B6+CAST) reads in each HNRNPK and HNRNPU (-) and (+) dox replicate (sixteen total replicates). From this list, B6/(B6+CAST) ratio were calculated for genes in the (-) dox replicates from HNRNPK and HNRNPU and average for each gene. Genes whose average, B6/(B6+CAST) ratios were < 0.01 were defined as "inactivated", those whose ratios were ≥ 0.01 but < 0.10 were defined as "weak escapees", and those whose ratios were ≥ 0.1 were defined as "strong escapees". Statistical significance of differences for each gene class between the (-) and (+) dox conditions were determined using average, within-genotype B6/(B6+CAST) ratios and two-tailed Student's t-tests.

Quantitation of Airn, Kcnq1ot1, and Xist abundance by RNA-Seq

For each sample, RNA-seq reads from each replicate were aligned with STAR (v2.7.9a; (Dobin et al. 2013)) to the mouse GRCm38 (mm10; (Perez et al. 2025)) genome with reference to GTF file gencode.vM25.basic.annotation.gtf (Mudge et al. 2025), using multi-sample two-pass alignment that incorporated novel splice junctions discovered among all samples in a given experiment. STAR was run with the option "--outFilterMultimapNmax 1" to consider only uniquely mapped reads. For each sample, reads mapping to *Airn*, *Kcnq1ot1*, and *Xist* were counted using featureCounts (Subread v2.0.4; (Liao et al. 2014)) and RPM normalized using the total number of reads uniquely mapped by STAR to the mm10 genome.

RNA-FISH quantification

RNA FISH signal was quantified in FIJI (version 2.16.0/1.54p) as follows. First, channels were split and maximum z-projections were generated. Then, a threshold was set on the (– dox) projection and applied identically to the paired (+ dox) image to create binary masks of RNA FISH foci (presence/absence as shown in Fig. S8A). For each image, with only the DAPI channel visible, identifiable individual nuclei were manually outlined using the FIJI selection tool. These outlines were then saved as regions of interest (ROIs) as an overlay (example shown in Figure S8A, right). Then, the overlay was applied to the corresponding binary RNA FISH image and as each nucleus was counted and labeled, the number of discrete foci (0, 1, or ≥ 2) for each nucleus was recorded. At least 3 fields per condition ((-) and (+) dox) were analyzed among two biological replicates, fixed and imaged on separate days, for a total of > 200 nuclei per condition. The fraction of nuclei in each foci category (0, 1, ≥ 2) was plotted in GraphPad Prism and statistical significance assessed by chi square test.

RIP-Seq alignments and peak calling

Reads from individual RIP-Seq replicates were aligned to mm10 using STAR (Dobin et al. 2013) and filtered for MAPQ of ≥30 using SAMtools (Li et al. 2009). Code used for peak definition was the same code used in (Trotman et al. 2025a) and is outlined on https://github.com/CalabreseLab/Airn_Xist_manuscript/tree/main. For each protein profiled by RIP, after alignment and MAPQ of ≥30 filtering, sequencing data from all replicates were concatenated and split into two files using SAMtools (Li et al. 2009), corresponding to alignments that mapped to the positive and negative strands of the genome, respectively. Using a custom perl script, the strand information within the positive and negative strand alignment files was randomized so as to better match the criteria of the MACS peak caller, which uses the average distance between positive and negative strand alignments to estimate the fragment length (Zhang et al. 2008). Putative peaks were called on strand-randomized positive and negative strand alignment files, respectively, using default MACS parameters and not providing a background file (Zhang et al. 2008). Peak bed files were converted to SAF format and reads under each putative peak were counted for each RIP replicate as well as for concatenated the set of MAPQ of ≥30 filtered reads from all IgG replicates using featureCounts (Liao et al. 2014). Counts per peak per dataset were converted into RPM by dividing by the total number of sequenced reads per dataset and multiplying by 1 million. We defined as peaks those regions whose RPM-normalized values in the RIPs were >2x those in IgG control in at least two RIP replicates.

Identification of motifs enriched under RIP-Seq peaks

For each protein analyzed by RIP-Seq, peaks were ranked by the amount of averaged RPM-normalized RIP signal per replicate subtracted by the signal from RPM-normalized IgG control. The top 1000 peaks most enriched in each RIP-Seq were selected for motif analyses. The sequences underneath each peak were extracted from mm10 genome using bedtools getfasta (Quinlan and Hall 2010). For each protein analyzed, a set of 1000 control sequences of lengths

identical to those in the corresponding peak file was generated, but whose sequence content was randomly generated in python using the mononucleotide frequency of the mm10 primary assembly. Peak sequences were then searched for motifs relative to control sequences using streme and the options: --maxw 8 --minw 4 --nmotifs 10 --rna (Bailey 2021).

Wiggle track generation

Wiggle tracks were generated using protocols and code from (Li et al. 2024). Sequencing reads were aligned to the mm10 genome using STAR (v2.7.10a; (Langmead et al. 2009)). Replicates were downsampled and merged using SAMtools (v1.18; (Li et al. 2009)). Merged BAM files were filtered for MAPQ ≥ 30 and then split by strand with SAMtools and a custom script. Bam files were converted to BED12 format using Bedtools (v2.29; (Quinlan and Hall 2010)). Wiggle tracks were generated using a custom script which included normalizing wiggles to total aligned reads. WIG files were converted to BigWig format and then plotted over respective genome coordinates using PlotGardner in RStudio (Kramer et al. 2022).

RIP-Seq signal per transcript

RIP-Seq signal per transcript was calculated as in (Trotman et al. 2025a). To rank each expressed chromatin-associated transcript by its level of association with each protein profiled by RIP in TSCs, we first created a version of the GENCODE vM25 basic transcriptome that included one representative intron-containing (i.e., nascent) transcript for each gene that began at the gene's first annotated transcription start and the last annotated transcription end (Frankish et al. 2023). We used kallisto (Bray et al. 2016) with the options "--single -s 300 -l 100 --rf-stranded" to align RNA-Seq data from total cellular RNA, as well as TSC cytoplasmic and chromatin-associated fractions. We defined as expressed those transcripts whose median TPM expression

value from 8 independent replicates of total RNA-Seq was >0.0625 (Table S2). We defined as chromatin-associated any expressed transcript for which [chromatin_TPM]/([chromatin_TPM] +[cytoplasmic_TPM]) was >0.75. We excluded any transcript whose length was less than 500 nt.

To calculate RIP signal over each expressed transcript, we selected only the subset of reads from each RIP-Seq and IgG dataset that aligned under the peaks defined for each protein target of RIP. To do this for a given protein target, RIP-Seq and IgG reads were aligned to mm10 using STAR and filtered for MAPQ ≥30 using SAMtools (Li et al. 2009; Dobin et al. 2013). Samtools view was used to split alignments by strand. For each stranded alignment file, we used samtools view to select the subset of RIP-Seq and IgG reads that aligned under each peak. Still using samtools view, we converted bam alignments back into fastq format. Fastq data were aligned with kallisto to the same GENCODE vM25 basic transcriptome used above, which also contained one representative nascent transcript for each gene, using the options [-l 200 -s 50 --rf-stranded] (Bray et al. 2016). For each transcript, TPM counts from the IgG datasets were subtracted from TPM counts in RIP-Seq datasets, and these IgG-corrected values were used to generate transcript rankings shown in Figure 1.

RIP-Seq network analyses

Bedtools multicov was used to calculate replicate-averaged, RPM-normalized read densities in 25 nt bins across the set of 19295 chromatin-associated transcripts in TSCs for each RIP-Seq protein target along with IgG controls (Quinlan and Hall 2010). Signal in IgG was subtracted from signal in RIP; negative values were set to 0. Pairwise Pearson's correlation values for all 27 proteins were calculated between all transcript rows in the resulting matrix. Pearson's correlations were assigned the value "NA" if they involved signal values of "zero" across all bins in a given transcript. For transcripts that had non-zero signal values in at least one RIP, NA values were replaced with "0" values. Seventy-two transcripts had NA values for all positions across all RIPs

and were removed from subsequent network comparisons. To assign RIP protein targets to communities within each transcript's association network, we required that all RIPs had non-zero values in at least one 25 nt bin; this requirement removed 5464 transcripts and enabled our subsequent calculation of *p* values ascribing likelihood of the prevalence and rarity of edges between protein pairs across the chromatin-associated TSC transcriptome. Communities were assigned using the Leiden algorithm, using only non-negative edge weights/Pearson's r values in the adjacency matrix. To make the plot in Figure 2Ci, we calculated the distribution of edge weights (Pearson's r values) between all nodes within communities and across different communities. We calculated the modularity of each transcript's network using the function modularity from the R package igraph (Csardi et al. 2025) and calculated silhouette width using the function silhouette in the R package cluster (Maechler et al. 2025).

To determine whether protein pairs present within the same communities represented prevalent or rare events among the chromatin-associated TSC transcriptome we analyzed networks from the 13831 chromatin-associated transcripts that had at least one non-zero bin value in each RIP. Within these networks, the 27 proteins profiled by RIP-Seq could be present in 351 possible pairs (27*26/2). If the network structure is fixed, and node IDs are shuffled to simulate randomization, for each network, the probability that any two proteins fall in the same community (the intracommunity probability) can be estimated as the sum of intracommunity edges divided by the total number of edges in that specific network. Because each network has a different intracommunity probability, (defined below as "prob"), we used an application of the Poisson Binomial Distribution and the cumulative distribution function in the poibin R package to determine if an intracommunity protein pairs that appears x number of times among the 13831 networks were rare [ppoibin (x, prob)] or prevalent [1 - ppoibin(x - 1, prob)] (Hong and Team. 2024). Due to numerical instability / rounding errors in computing the Poisson binomial cumulative distribution function, *p* values corresponding to values of less than 1.0e-12 were set to < 1.0e-12, before correcting for multiple testing using the Benjamini-Hochberg method.

RIP-Seq community demarcations

To assign sequences in *Airn*, *Kcnq1ot1*, and *Xist* to individual communities, we extracted the reads under peaks from each RIP-Seq replicate merged by protein target and calculated the RPM based on the total number of reads under peak for each protein target. From this RPM value, we then subtracted the RPM value from merged IgG control under the same set of peaks for each corresponding protein target. IgG-normalized reads for each protein target for each community shown in Figure 2 were summed in 50 nt bins along the lengths of the *Airn*, *Kcnq1ot1*, and *Xist* genes. Any 50 nt bin that had an IgG-corrected RPM value of >0.125 was retained as belonging to the community in question and not the null.

To perform *k*-mer comparisons between community sequences, we took all consecutive binned community assignments and used SEEKR at *k*-mer length $k = 4$, log.post, and a background set of GENCODE vM25 canonical, non-redundant lncRNAs >500 nt long, to calculate the Pearson's r value of all possible pairwise comparisons for segments from within each community to itself, and from within each community to segments from all other communities and the null (Li et al. 2024). Prior to comparisons using SEEKR community assignments >350 nt in length were divided into segments of ~200 nt. We used Wilcoxon-ranked sum tests to calculate *p* values comparing the distribution of r values from within each community to all other communities and the null. We corrected the *p* values using the Benjamini-Hochberg method.

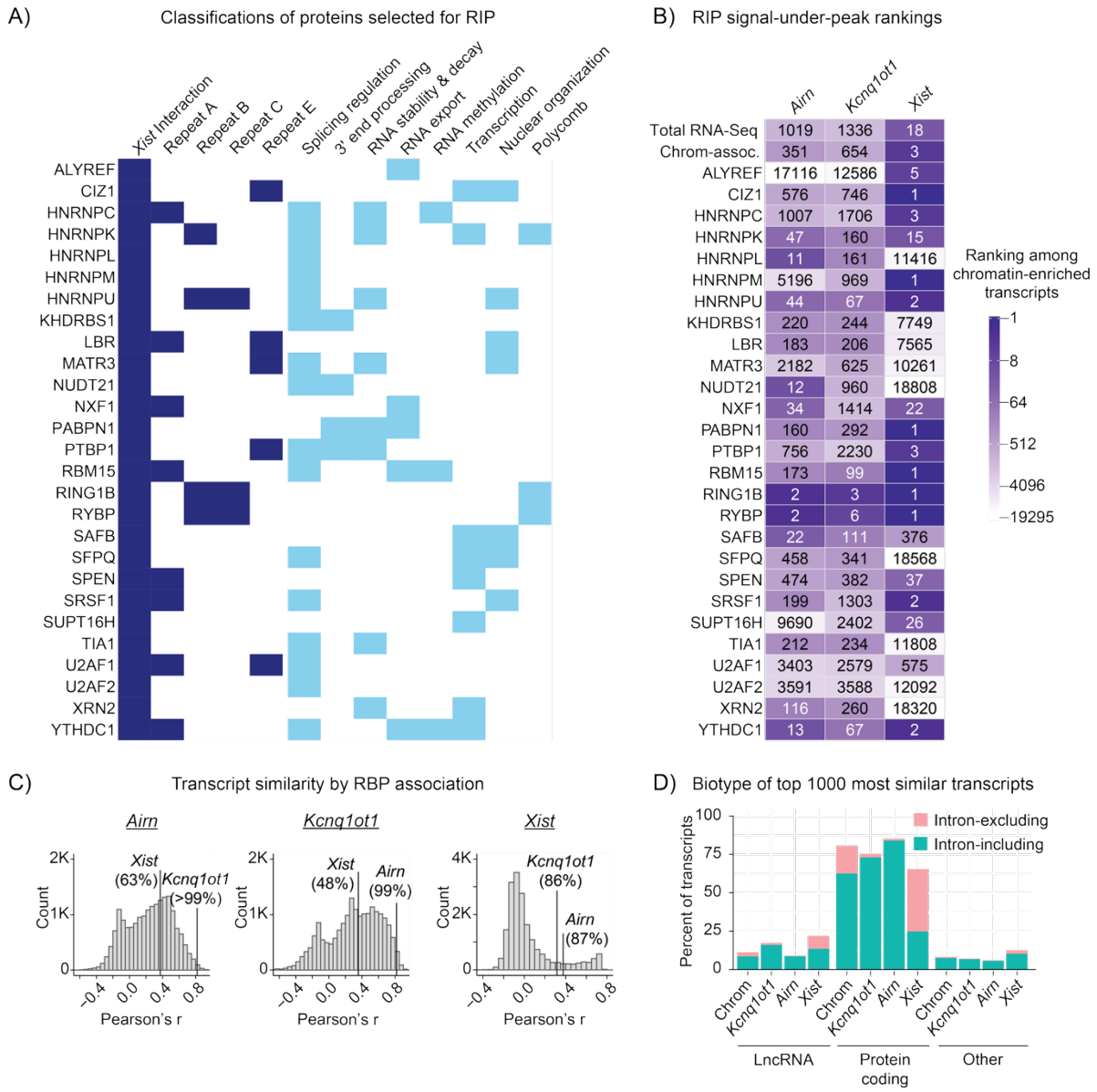

Figure 1. Protein association profiles of *Airn*, *Kcnq1ot1*, and *Xist* are highly similar and exhibit similarity to profiles of transcripts produced from both lncRNA and protein-coding genes. (A) 27 proteins selected for RIP-Seq profiling in TSCs, along with their links to *Xist* biology and select nuclear processes. "*Xist* Interaction," whether the protein is *Xist*-associated; "Repeat A/B/C/E," whether the protein is enriched over the specified *Xist* Repeat; (Chu et al. 2015; McHugh et al. 2015; Minajigi et al. 2015; Smola et al. 2016; Pintacuda et al. 2017; Bousard et al.

2019; Graindorge et al. 2019; Weidmann et al. 2021; Yu et al. 2021; Tsue et al. 2024). **(B)** Ranking of signal-under-peaks within *Airn*, *Kcnq1ot1*, and *Xist* relative to signal-under-peak within all TSC chromatin-enriched transcripts. “Total RNA-Seq,” ranking from total RNA-Seq. “Chrom-assoc,” ranking from chromatin-associated RNA-Seq data (Trotman et al. 2025a). **(C)** Histograms of Pearson’s r values comparing signal-under-peak for all 27 proteins in *Airn*, *Kcnq1ot1*, and *Xist* relative to all other TSC chromatin-enriched transcripts. Percentile of similarity ranking relative to 18820 chromatin-enriched transcripts is labeled. **(D)** GENCODE biotype of the corresponding gene and intron status of 1000 transcripts most similar to *Airn*, *Kcnq1ot1*, and *Xist* compared to all chromatin-enriched transcripts (Mudge et al. 2025). See also Figures S1-3, Tables S1-3.

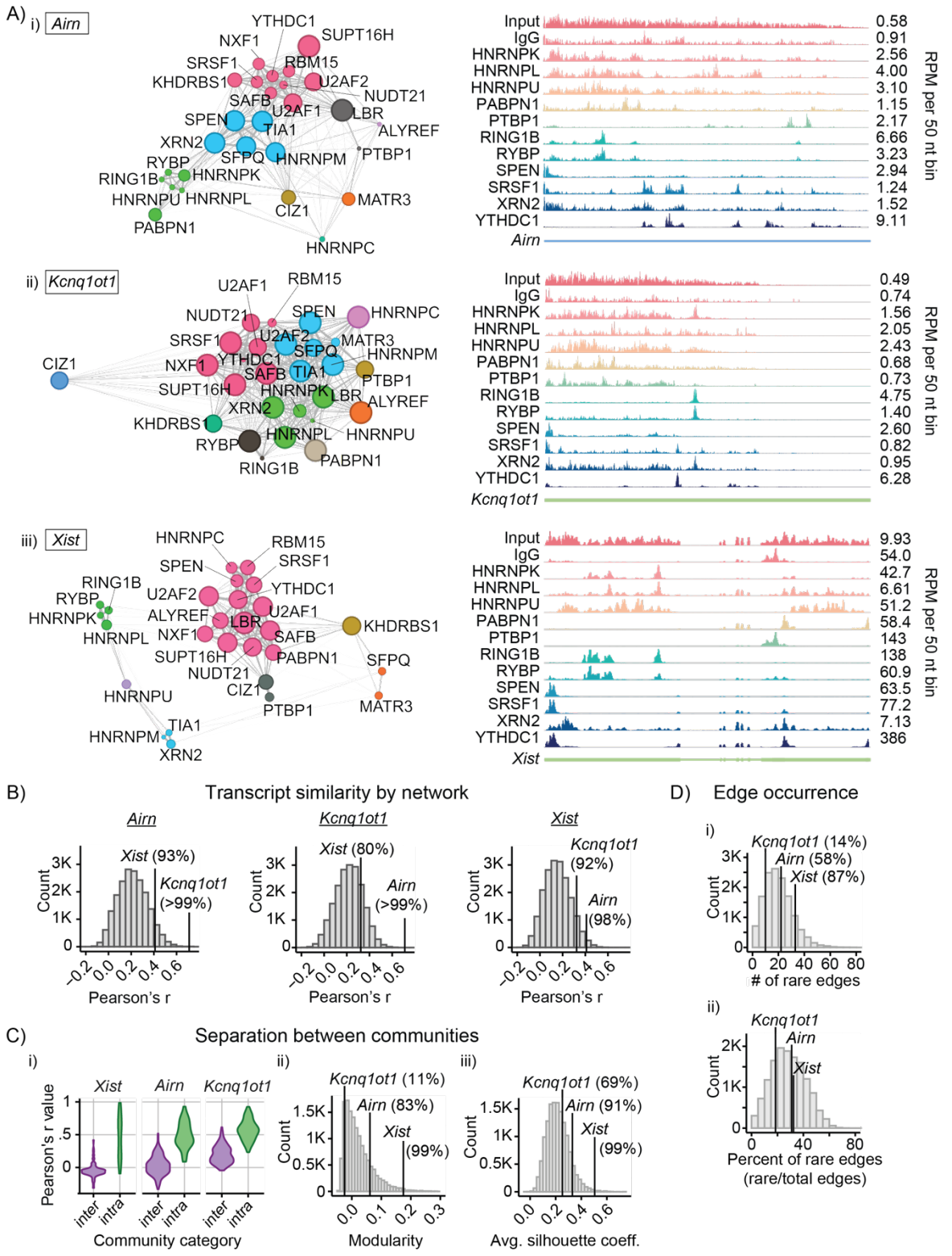

Figure 2. Networks of protein association in *Airn*, *Kcnq1ot1*, and *Xist* are highly similar and reveal a correlation between modularity and repressive potency. (A) Network graphs of

protein association in (i) *Airn*, (ii) *Kcnq1ot1*, and (iii) *Xist*. Drawn edges represent nodes connected by Pearson's r values of >0.0. Right-hand panels: UCSC wiggle density profiles of RIP-Seq data from selected proteins. **(B)** Histograms of Pearson's r values comparing RIP-Seq association networks in *Airn*, *Kcnq1ot1*, and *Xist* relative to other TSC chromatin-enriched transcripts. **(C)** Metrics quantifying the degree of separation between RIP communities in *Airn*, *Kcnq1ot1*, and *Xist*. Distributions of inter- versus intra-community Pearson's r values (i), modularity (ii), and average silhouette distance (iii) quantifying connectivity of RIP-Seq association networks in *Airn*, *Kcnq1ot1*, and *Xist*. **(D)** Number (i) and proportion (ii) of rare edges in *Airn*, *Kcnq1ot1*, and *Xist*. For all histograms: Percentile of similarity ranking relative to other chromatin-enriched transcripts in category is labeled. See also Tables S3, S4.

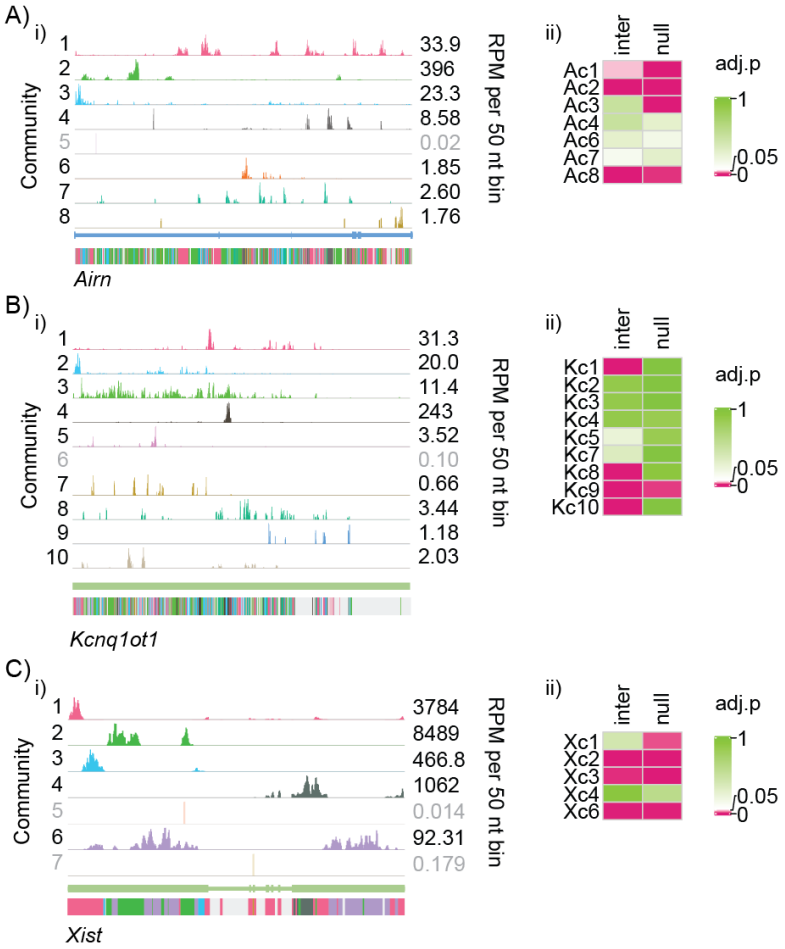

Figure 3. Community structures demarcate protein interaction domains with distinct sequence signatures in *Airn*, *Kcnq1ot1*, and *Xist*. (A-C) Panels i) Wiggle density tracks of IgG-corrected RIP-Seq data under peak merged by communities outlined in Figure 2, in *Airn* (A), *Kcnq1ot1* (B), and *Xist* (C). Communities that did not harbor above-threshold read density or are marked by only a single enriched region are greyed out. Panels ii) Benjamini-Hochberg corrected *p* values of Wilcoxon-rank sum tests comparing the distributions of Pearson's r values derived from all pairwise comparisons of segments from within a single community to themselves, versus the distributions of Pearson's r values derived from all pairwise comparisons of segments from

within a single community to segments from all other communities or the null for *Airn* (A), *Kcnq1ot1* (B), and *Xist* (C).

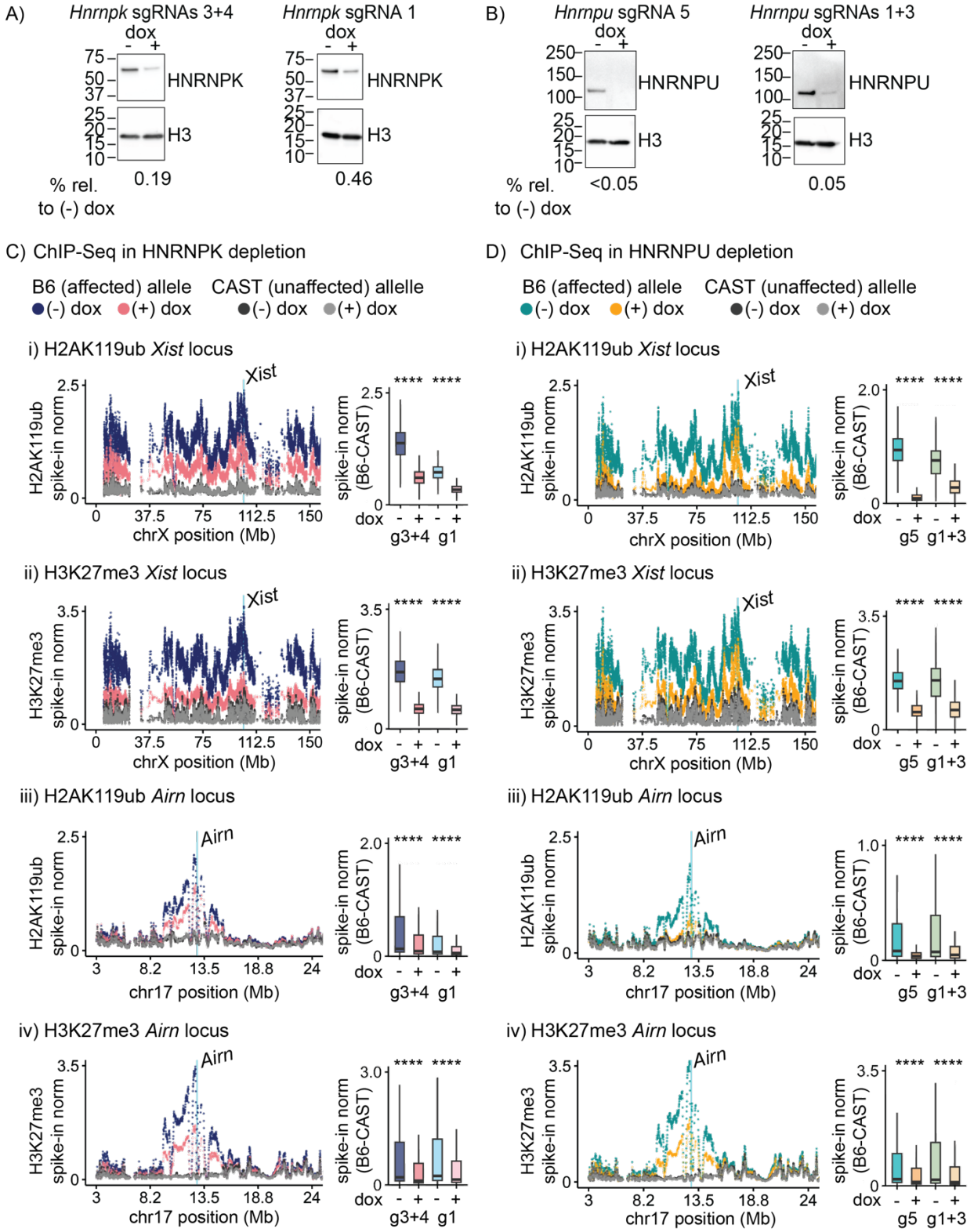

Figure 4. HNRNPU is required for maintenance of PRC-directed modifications induced by

***Airn*, *Kcnq1ot1*, and *Xist*. (A, B)** Western blots showing depletion of HNRNPK (A) and HNRNPU (B) relative to H3 control in polyclonal TSC cultures after four days of Cas9 induction. Each protein was depleted using two separate groups of sgRNAs. **(C, D)** H2AK119ub and H3K27me3 levels with and without doxycycline treatment to induce Cas9 in TSCs expressing sgRNAs targeting HNRNPK (C) and HNRNPU (D). H2AK119ub and H3K27me3 levels are shown in panels (i, iii) and (ii, iv), respectively, over the *Xist* (i, ii) and *Airn* (iii, iv) target domains, on B6 and CAST alleles with and without doxycycline treatment to induce Cas9. Tiling density plots and box-and-whisker plots of spike-in normalized H2AK119ub and H3K27me3 ChIP-Seq signal per 10 kb bin are displayed on the left and right, respectively, in both (C) and (D). Tiling density plots represent the data averaged between the two different sgRNA-expressing populations, and box-and-whisker plots show [B6-CAST] values for each sgRNA experiment. ****: $p \leq 0.0001$; Student's t-test comparing data in (-) vs. (+) doxycycline conditions. See also Figures S4-7.

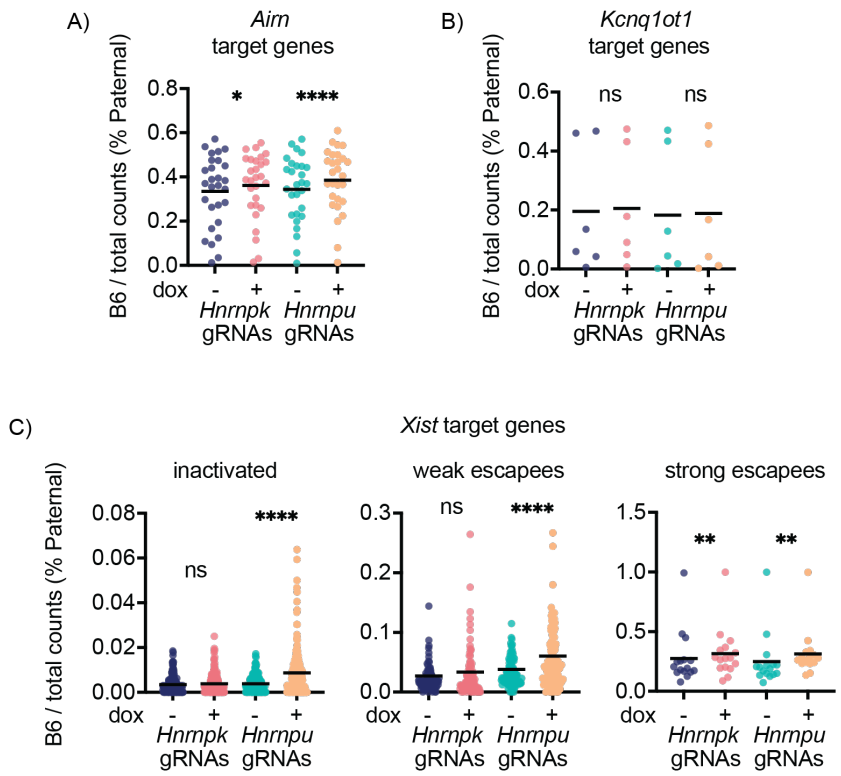

Figure 5. HNRNPU and HNRNPK depletion derepress genes in *Airn* and *Xist* but not *Kcnq1ot1* target domains. (A, B, C) RNA-Seq showing percent paternal expression bias for *Airn* (A), *Kcnq1ot1* (B), and *Xist* (C) target genes for both HNRNPK and HNRNPU- depleted cells and non-depleted (-dox) controls. *Airn* and *Kcnq1ot1* target genes are from (Schertzer et al. 2019a). Each individual dot displays the average % paternal expression derived from four independent RNA-Seq experiments. *Xist* target genes were classified as inactivated, weak, or strongly escaping based on (-) dox RNA-Seq samples collected in this study. *, **, ****: $p \leq 0.05$, 0.01, and 0.0001, respectively; Student's t-test. See also Table S5.

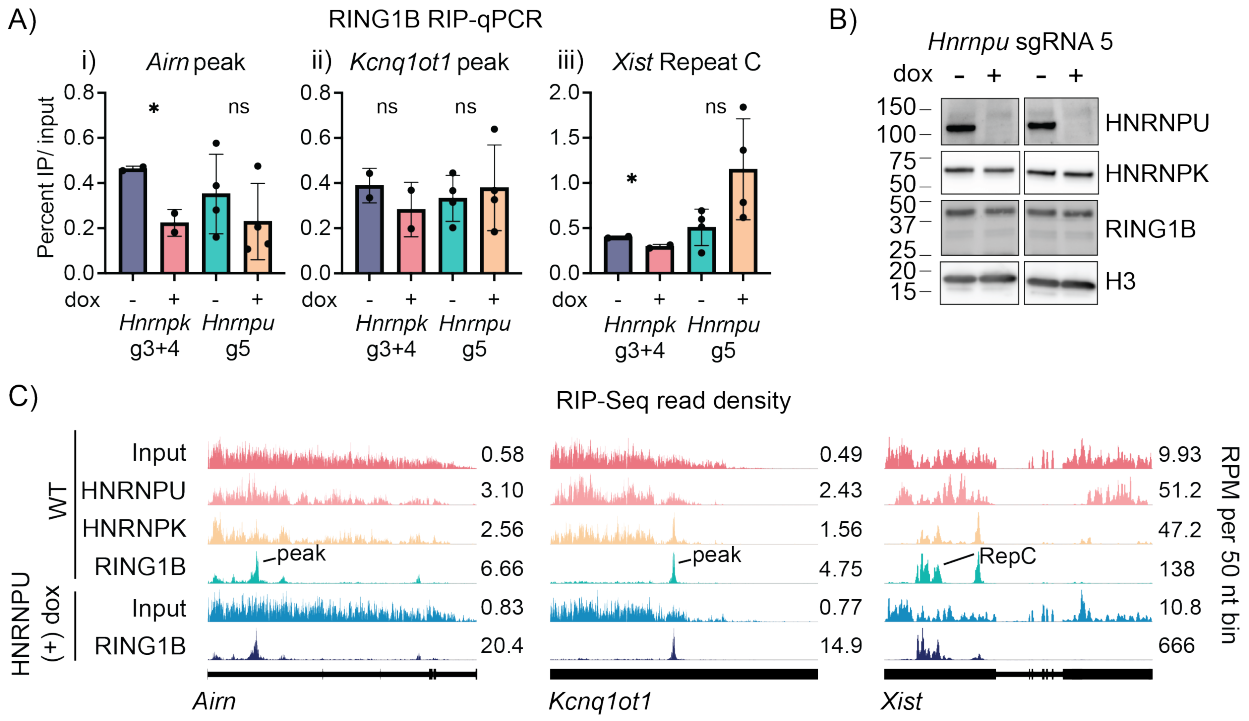

Figure 6. HNRNPK but not HNRNPU depletion reduces associations between PRC1 and *Airn*, *Kcnq1ot1*, and *Xist*. (A) Input-normalized RING1B RIP-qPCR in *Airn* (i), *Kcnq1ot1* (ii), and *Xist* (iii) in HNRNPK and HNRNPU-depleted cells and paired non-depleted controls. Locations of qPCR primer pairs are shown in (C). Each dot represents the average of qPCR technical triplicates from a biological replicate RIP experiment. *, $p \leq 0.05$; Student's t-test. **(B)** Western blots showing levels of HNRNPU, HNRNPK, RING1B, and H3 with and without Cas9 induction by doxycycline in HNRNPU sgRNA-expressing TSCs. **(C)** Wiggle density profiles in 50 nt bins of Input, HNRNPK, and HNRNPK RIP-Seq data from wild-type TSCs ("WT"; same data analyzed in Figures 1 and 2) and from HNRNPU depleted TSCs ("HNRNPU (+) dox.": same data analyzed in Figure 5A (+) dox condition).

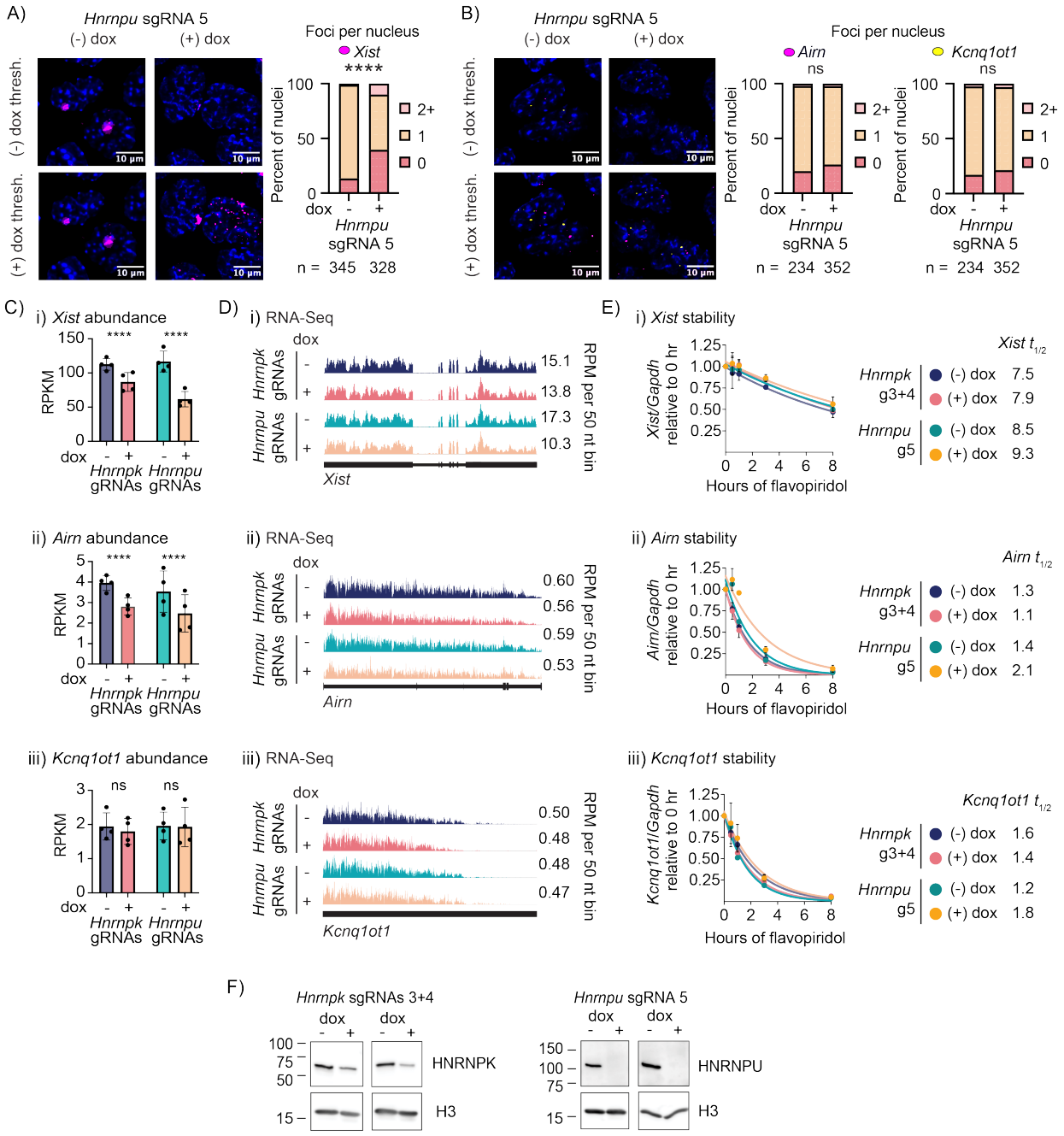

Figure 7. HNRNPU depletion reduces the abundance of *Airn* without affecting its localization or half-life. (A, B) Single-molecule sensitivity RNA FISH detecting *Xist* (A) and *Airn* and *Kcnq1ot1* (magenta and yellow, respectively;B) in HNRNPU-depleted and non-depleted TSCs. Representative FISH images with signal thresholds set against TSCs lacking Cas9

expression ("(-) dox thresh.") and after four days of Cas9 induction ("(+) dox thresh."). Stacked bar plots show counts of lncRNA foci per nucleus in the (-) dox and (+) dox conditions. ****, p ≤0.0001; Chi-squared. **(C,D,E)** Expression levels (C), Wiggle density tracks displaying RNA-Seq read density per 50 nt bin (D), and expression levels after treatment with flavopiridol (E) of *Xist* (i), *Airn* (ii), and *Kcnq1ot1* (iii) in the (-) dox and (+) dox conditions in TSCs expressing HNRNPK- and HNRNPU-targeting sgRNAs. RPKM, reads per kilobase per million aligned reads. Dots, values from individual biological replicates. ****, p ≤0.0001; Student's t-test. (**F)** Representative western blots demonstrating successful depletion of HNRNPK and HNRNPU in (+) dox conditions. See also Figure S8.

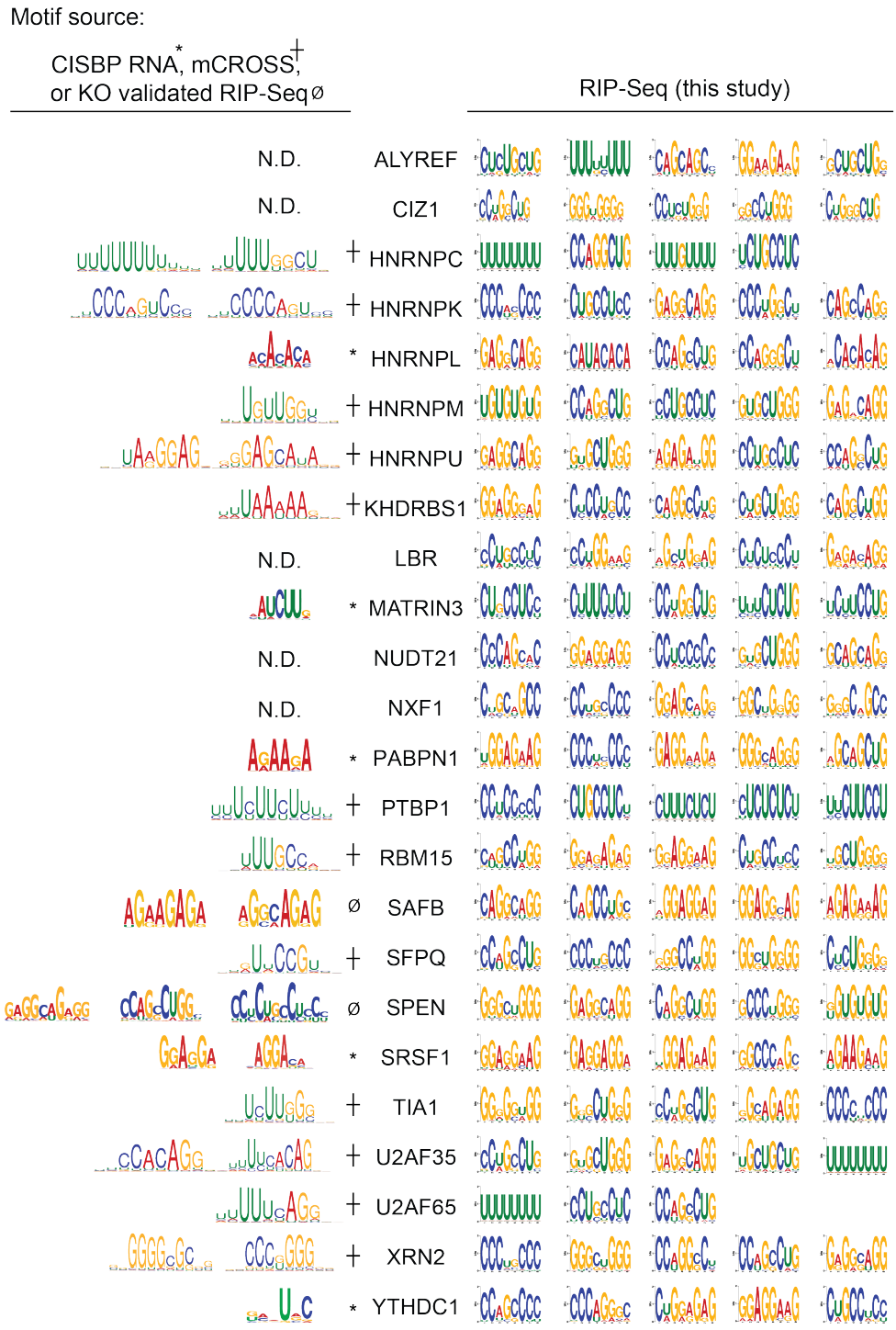

Figure S1. Motifs for RNA-binding proteins derived from RIP-Seq in this study versus motifs derived from prior studies (in vitro binding assays, CLIP, or knockout validated RIP-Seq). Left-hand panels, motifs reported in CISBP-RNA (Ray et al. 2013), mCROSS (Feng et al. 2019), or in the case of SAFB and SPEN, motifs derived from studies which used the same

antibodies in this study and derived motifs by comparing RIP-Seq from wild-type cells versus cells in which the denoted protein was knocked out (Cherney et al. 2023; Trotman et al. 2025b). Right-hand panels, up to the top 5 motifs derived from RIP-Seq experiments in this study. Of note, prior CLIP-Seq performed to map the RNA targets bound by ALYREF, NXF1, and NUDT21 (a.k.a. CPSF5/CFIm25) failed to identify strong consensus motifs (Martin et al. 2012; Viphakone et al. 2019).

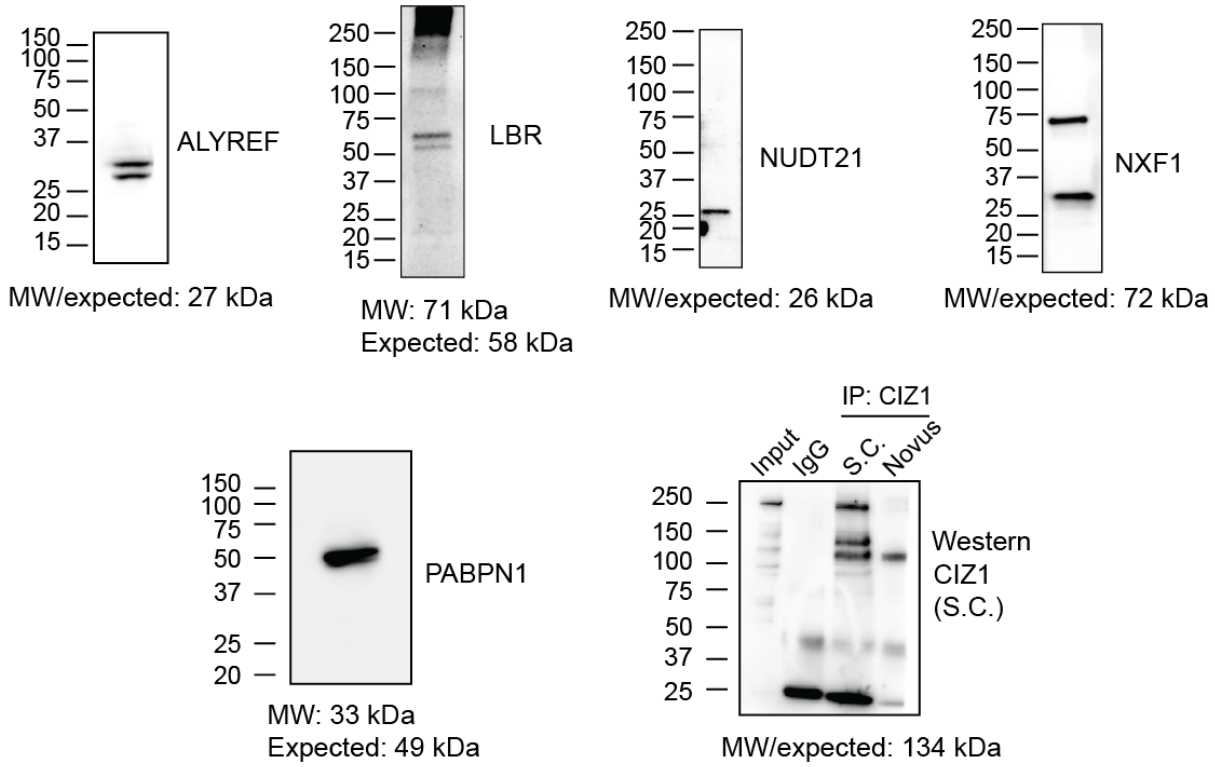

Figure S2. Western blots examining specificity of RIP antibodies raised against RIP-target proteins that lack previously reported consensus RNA interaction motifs. CIZ1 was only detectable by IP-Western; we used the CIZ1 antibody from Novus for RIP-Seq in this study. S.C., CIZ1 antibody from Santa Cruz Biotechnology (sc-393021). IP western was performing using RIP-seq washes as described.

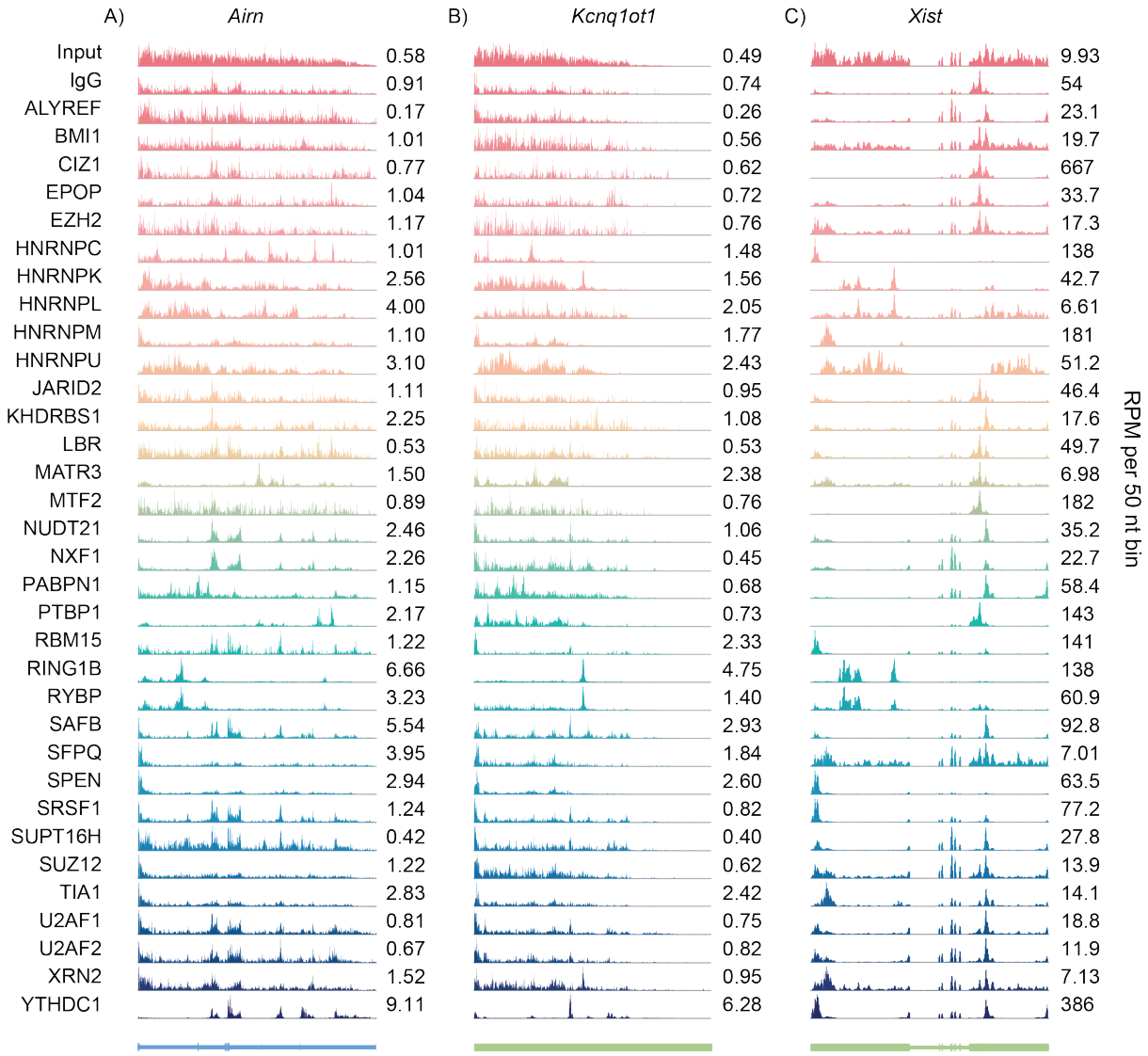

Figure S3. Wiggle density tracks of RIP-Seq data over *Airn* (A), *Kcnq1ot1* (B), and *Xist* (C) genes. y-axes, RPM per 50 nt bin.

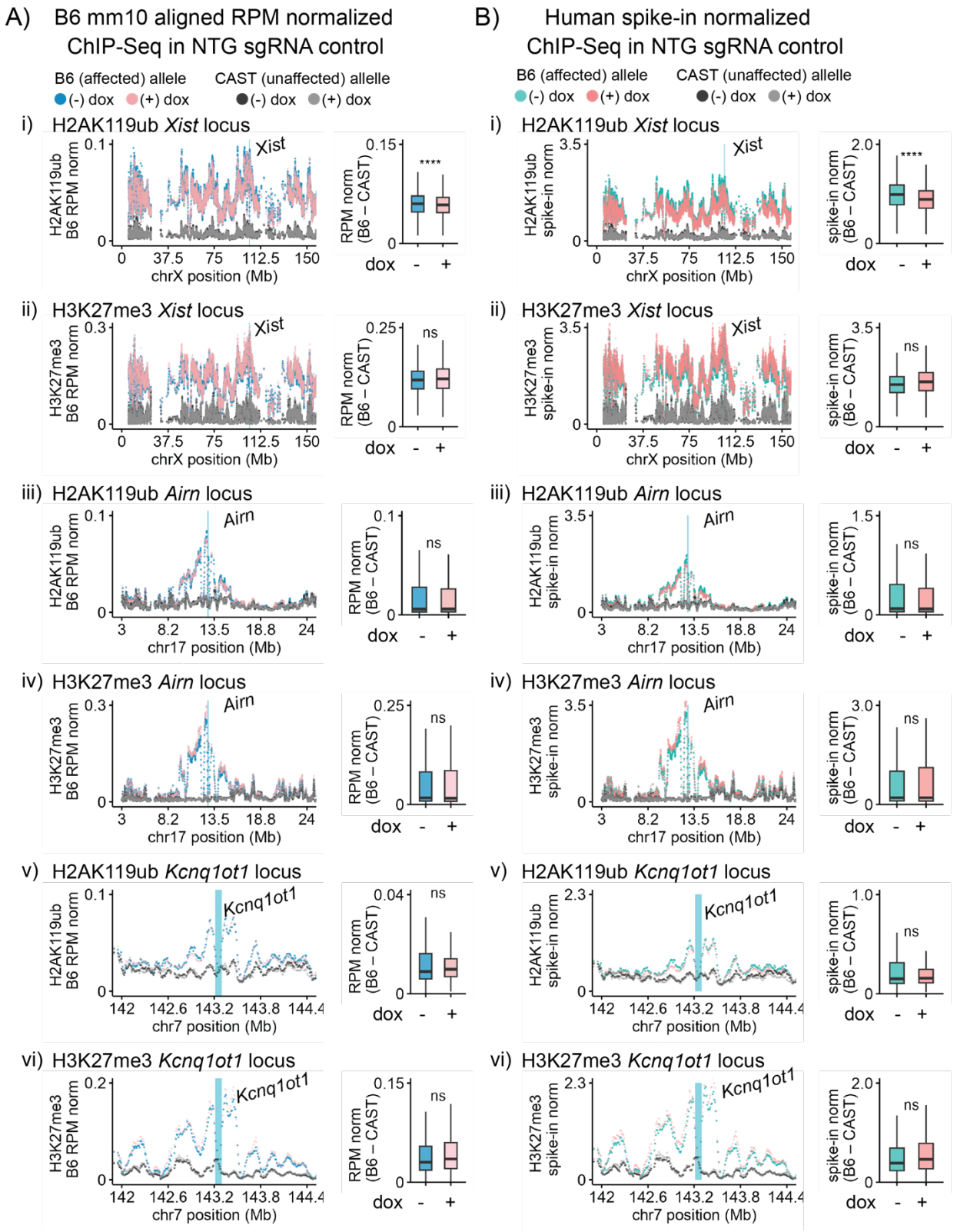
A) B6 mm10 aligned RPM normalized ChIP-Seq in NTG sgRNA control
B6 (affected) allele
(-) dox
(+) dox
CAST (unaffected) allele
(-) dox
(+) dox
i) H2AK119ub Xist locus
H2AK119ub B6 RPM norm
Xist
chrX position (Mb)
RPM norm (B6 – CAST)
****
dox - +
ii) H3K27me3 Xist locus
H3K27me3 B6 RPM norm
ns
iii) H2AK119ub Airn locus
Airn
chr17 position (Mb)
iv) H3K27me3 Airn locus
v) H2AK119ub Kcnq1ot1 locus
Kcnq1ot1
chr7 position (Mb)
vi) H3K27me3 Kcnq1ot1 locus
B) Human spike-in normalized ChIP-Seq in NTG sgRNA control
H2AK119ub spike-in norm
H3K27me3 spike-in norm
spike-in norm (B6 – CAST)

Figure S4. H2AK119ub and H3K27me3 ChIP-Seq over *Xist*, *Airn*, and *Kcnq1ot1* target domains in TSCs expressing non-targeting control sgRNA. RPM-normalized (A) and human Spike-In normalized (B) H2AK119ub and H3K27me3 ChIP-Seq data from TSCs expressing non-targeting sgRNAs shown over *Xist*, *Airn*, and *Kcnq1ot1* loci. H2AK119ub and H3K27me3 levels are shown in panels (i, iii, v) and (ii, iv, vi), respectively, over the *Xist* (i, ii), *Airn* (iii, iv), and *Kcnq1ot1* (v, vi) target domains, on B6 and CAST alleles with and without Cas9 induction by doxycycline treatment. Tiling density plots and box-and-whisker plots of spike-in normalized H2AK119ub and H3K27me3 ChIP-Seq signal per 10 kb bin are displayed on the left and right, respectively, in both (A) and (B). *, **, ****: $p \leq 0.05$, 0.01, and 0.0001, respectively; Student's t-test.

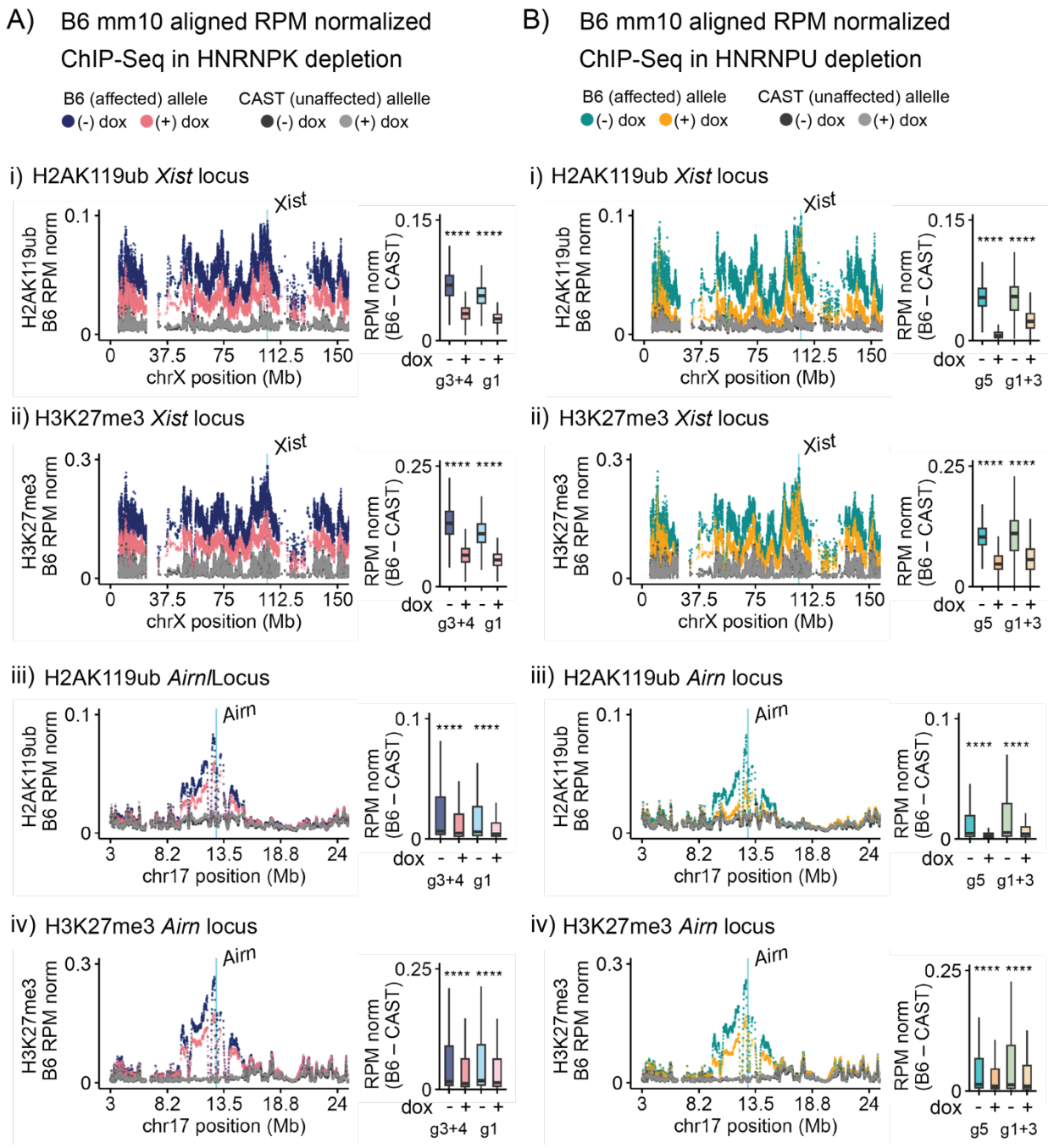

Figure S5. RPM normalized H2AK119ub and H3K27me3 ChIP-Seq over *Xist* and *Airn* target domains in HNRNPK- and HNRNPU-targeting sgRNA expressing TSCs. RPM-normalized H2AK119ub and H3K27me3 ChIP-Seq data in TSCs expressing HNRNPK-targeting (A) and HNRNPU-targeting (B) sgRNAs shown over *Xist* and *Airn* loci. H2AK119ub and H3K27me3 levels are shown in panels (i, iii) and (ii, iv), respectively, over the *Xist* (i, ii) and *Airn* (iii, iv) target

domains, on B6 and CAST alleles before and after Cas9 induction with dox. Tiling density plots and box-and-whisker plots of spike-in normalized H2AK119ub and H3K27me3 ChIP-Seq signal per 10 kb bin are displayed on the left and right, respectively, in both (A) and (B). Tiling density plots represent the data averaged between the two different sgRNA-expressing populations, and box-and-whisker plots show [B6-CAST] values for each sgRNA experiment. *, **, ****: $p \leq 0.05$, 0.01, and 0.0001, respectively; Student's t-test.

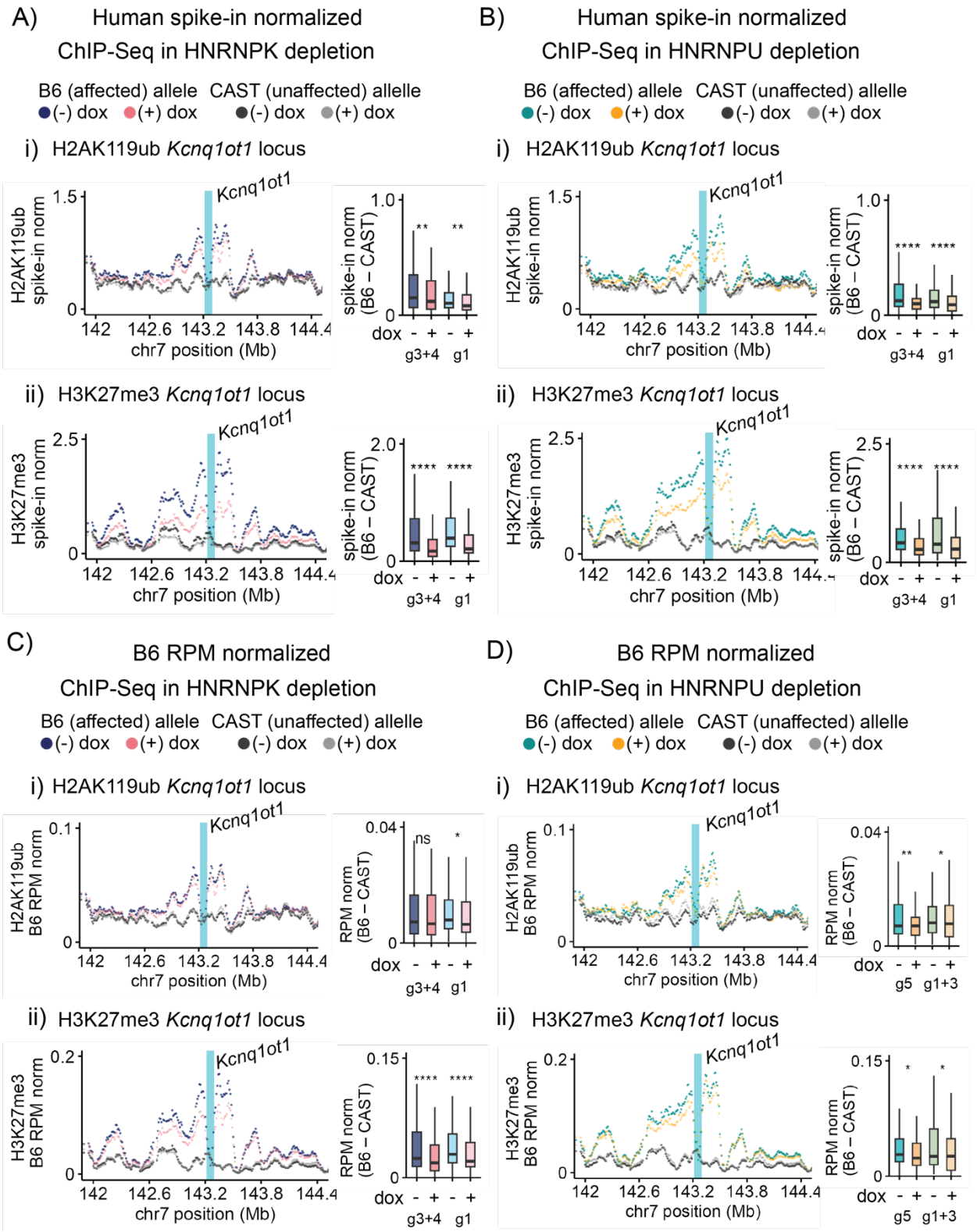
A)
Human spike-in normalized
ChIP-Seq in HNRNPK depletion
B6 (affected) allele
(-) dox
(+) dox
CAST (unaffected) allelle
(-) dox
(+) dox
i) H2AK119ub Kcnq1ot1 locus
H2AK119ub spike-in norm
Kcnq1ot1
chr7 position (Mb)
spike-in norm (B6 – CAST)
dox
g3+4
g1
ii) H3K27me3 Kcnq1ot1 locus
H3K27me3 spike-in norm
B)
Human spike-in normalized
ChIP-Seq in HNRNPU depletion
C)
B6 RPM normalized
ChIP-Seq in HNRNPK depletion
H2AK119ub B6 RPM norm
H3K27me3 B6 RPM norm
RPM norm (B6 – CAST)
D)
B6 RPM normalized
ChIP-Seq in HNRNPU depletion
g5
g1+3

Figure S6. RPM and human spike-in normalized H2AK119ub and H3K27me3 ChIP-Seq over *Kcnq1ot1* target domain in TSCs expressing HNRNPK- and HNRNPU-targeting sgRNAs. H2AK119ub and H3K27me3 ChIP-Seq data in TSCs expressing HNRNPK-targeting (A,C) and HNRNPU-targeting (B,D) sgRNAs shown over *Kcnq1ot1* target domain. Data are shown using both human Spike-In (A,B) and RPM (C,D) normalization schemes. H2AK119ub and H3K27me3 levels are shown in panels (i) and (ii), respectively, on B6 and CAST alleles before and after four days of Cas9 induction with dox. Tiling density plots and box-and-whisker plots of spike-in normalized H2AK119ub and H3K27me3 ChIP-Seq signal per 10 kb bin are displayed on the left and right, respectively, in (A-D). Tiling density plots represent the data averaged between the two different sgRNA-expressing populations, and box-and-whisker plots show [B6-CAST] values for each sgRNA experiment. *, **, ****: $p \leq 0.05$, 0.01, and 0.0001, respectively; Student's t-test.

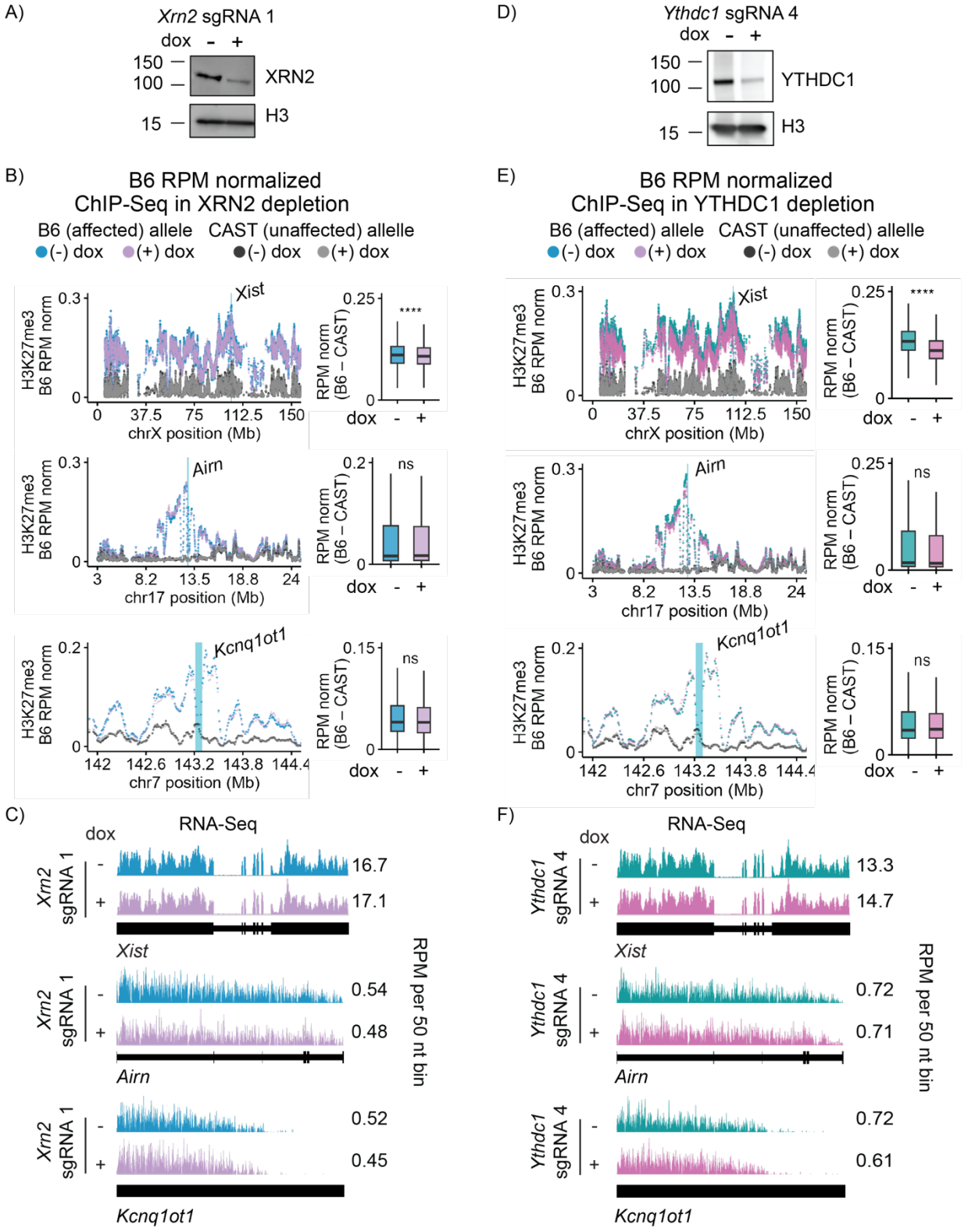
A)
Xrn2 sgRNA 1
dox - +
150
100
XRN2
15
H3
D)
Ythdc1 sgRNA 4
dox - +
150
100
YTHDC1
15
H3
B)
B6 RPM normalized
ChIP-Seq in XRN2 depletion
B6 (affected) allele
CAST (unaffected) allele
(-) dox
(+) dox
(-) dox
(+) dox
H3K27me3
B6 RPM norm
Xist
chrX position (Mb)
RPM norm
(B6 – CAST)
****
dox - +
Airn
chr17 position (Mb)
ns
Kcnq1ot1
chr7 position (Mb)
E)
B6 RPM normalized
ChIP-Seq in YTHDC1 depletion
C)
RNA-Seq
dox
Xrn2
sgRNA 1
16.7
17.1
Xist
0.54
0.48
Airn
0.52
0.45
Kcnq1ot1
RPM per 50 nt bin
F)
Ythdc1
sgRNA 4
13.3
14.7
0.72
0.71
0.72
0.61

Figure S7. RPM normalized H3K27me3 ChIP-Seq over *Xist*, *Airn*, and *Kcnq1ot1* target domains in TSCs expressing XRN2- and YTHDC1-targeting sgRNAs. (A, D) Western blot of XRN2 and histone H3 in (-) and (+) dox conditions, in TSCs expressing *Xrn2*-targeting (A) and *Ythdc1*-targeting (D) sgRNAs. **(B, E)** RPM normalized H3K27me3 ChIP-Seq data in TSCs expressing *Xrn2*-targeting (B) and *Ythdc1*-targeting (E) sgRNAs, shown over *Xist*, *Airn*, and *Kcnq1ot1* target domains. H3K27me3 levels are shown on B6 and CAST alleles in (-) and (+) dox conditions (four days of Cas9 induction). Tiling density plots and box-and-whisker plots of RPM normalized H3K27me3 ChIP-Seq signal per 10 kb bin are displayed on the left and right, respectively, in both (B) and (E). *, **, ****: $p \leq 0.05$, 0.01, and 0.0001, respectively; Student's t-test. Data shown are from a single experiment. **(C, F)** Wiggle density plots showing RNA-Seq data over *Xist*, *Airn*, and *Kcnq1ot1* in (-) and (+) dox conditions in TSCs expressing *Xrn2*-targeting (C) and *Ythdc1*-targeting (F) sgRNAs.

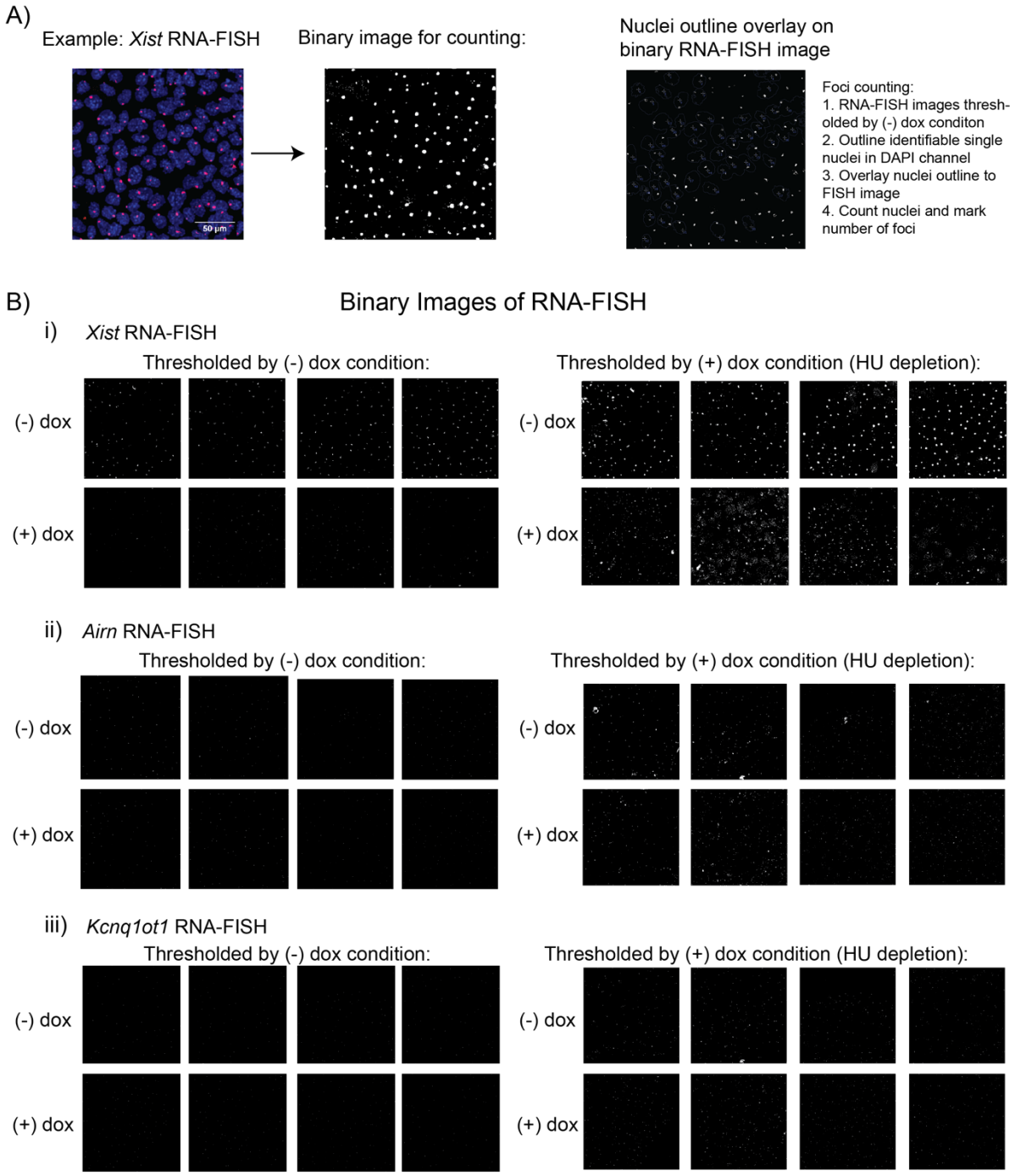

Figure S8. Overview of RNA FISH quantitation strategy and representative images. (A) Overview of foci counting strategy. **(B)** Representative *Xist* (i), *Airn* (ii), and *Kcnq1ot1* (iii) RNA

FISH images thresholded by signal in the (-) dox condition (left-hand panels) and (+) dox conditions (right-hand panels).

References

Andergassen D, Dotter CP, Wenzel D, Sigl V, Bammer PC, Muckenhuber M, Mayer D, Kulinski TM, Theussl HC, Penninger JM et al. 2017. Mapping the mouse Allelome reveals tissue-specific regulation of allelic expression. *eLife* **6**.

Andergassen D, Muckenhuber M, Bammer PC, Kulinski TM, Theussl HC, Shimizu T, Penninger JM, Pauler FM, Hudson QJ. 2019. The Airn lncRNA does not require any DNA elements within its locus to silence distant imprinted genes. *PLoS Genet* **15**: e1008268.

Andergassen D, Smith ZD, Kretzmer H, Rinn JL, Meissner A. 2021. Diverse epigenetic mechanisms maintain parental imprints within the embryonic and extraembryonic lineages. *Dev Cell* **56**: 2995-3005 e2994.

Bailey TL. 2021. STREME: Accurate and versatile sequence motif discovery. *Bioinformatics*.

Bedi K, Magnuson BR, Narayanan I, Paulsen M, Wilson TE, Ljungman M. 2021. Co-transcriptional splicing efficiencies differ within genes and between cell types. *RNA* **27**: 829-840.

Bousard A, Raposo AC, Zylicz JJ, Picard C, Pires VB, Qi Y, Gil C, Syx L, Chang HY, Heard E et al. 2019. The role of Xist-mediated Polycomb recruitment in the initiation of X-chromosome inactivation. *EMBO Rep* **20**: e48019.

Braceros AK, Schertzer MD, Omer A, Trotman JB, Davis ES, Dowen JM, Phanstiel DH, Aiden EL, Calabrese JM. 2023. Proximity-dependent recruitment of Polycomb repressive complexes by the lncRNA Airn. *Cell Rep* **42**: 112803.

Bray NL, Pimentel H, Melsted P, Pachter L. 2016. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* **34**: 525-527.

Calabrese JM, Starmer J, Schertzer MD, Yee D, Magnuson T. 2015. A survey of imprinted gene expression in mouse trophoblast stem cells. *G3 (Bethesda)* **5**: 751-759.

Calabrese JM, Sun W, Song L, Mugford JW, Williams L, Yee D, Starmer J, Mieczkowski P, Crawford GE, Magnuson T. 2012. Site-specific silencing of regulatory elements as a mechanism of X inactivation. *Cell* **151**: 951-963.

Cherney RE, Eberhard QE, Giri G, Mills CA, Porrello A, Zhang Z, White D, Trotman JB, Herring LE, Dominguez D et al. 2023. SAFB associates with nascent RNAs and can promote gene expression in mouse embryonic stem cells. *RNA* **29**: 1535-1556.

Choquet K, Patop IL, Churchman LS. 2025. The regulation and function of post-transcriptional RNA splicing. *Nat Rev Genet* **26**: 378-394.

Chu C, Zhang QC, da Rocha ST, Flynn RA, Bharadwaj M, Calabrese JM, Magnuson T, Heard E, Chang HY. 2015. Systematic discovery of Xist RNA binding proteins. *Cell* **161**: 404-416.

Colognori D, Sunwoo H, Kriz AJ, Wang CY, Lee JT. 2019. Xist Deletional Analysis Reveals an Interdependency between Xist RNA and Polycomb Complexes for Spreading along the Inactive X. *Mol Cell* **74**: 101-117 e110.

Cote A, O'Farrell A, Dardani I, Dunagin M, Cote C, Wan Y, Bayatpour S, Drexler HL, Alexander KA, Chen F et al. 2024. Post-transcriptional splicing can occur in a slow-moving zone around the gene. *eLife* **12**.

Creamer KM, Kolpa HJ, Lawrence JB. 2021. Nascent RNA scaffolds contribute to chromosome territory architecture and counter chromatin compaction. *Mol Cell* **81**: 3509-3525 e3505.

Csardi G, Nepusz T, Muller K, Horvat S, Traag V, Zanini F, Noom D. 2025. igraph for R: R interface of the igraph library for graph theory and network analysis.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15-21.

Dou X, Xiao Y, Shen C, Wang K, Wu T, Liu C, Li Y, Yu X, Liu J, Dai Q et al. 2023. RBFOX2 recognizes N(6)-methyladenosine to suppress transcription and block myeloid leukaemia differentiation. *Nat Cell Biol* **25**: 1359-1368.

Feng H, Bao S, Rahman MA, Weyn-Vanhentenryck SM, Khan A, Wong J, Shah A, Flynn ED, Krainer AR, Zhang C. 2019. Modeling RNA-Binding Protein Specificity In Vivo by Precisely Registering Protein-RNA Crosslink Sites. *Mol Cell* **74**: 1189-1204 e1186.

Fowlkes EB, Mallows CL. 1983. A Method for Comparing Two Hierarchical Clusterings. *Journal of the American Statistical Association* **78**: 553–569.

Frankish A, Carbonell-Sala S, Diekhans M, Jungreis I, Loveland JE, Mudge JM, Sisu C, Wright JC, Arnan C, Barnes I et al. 2023. GENCODE: reference annotation for the human and mouse genomes in 2023. *Nucleic Acids Res* **51**: D942-D949.

Graindorge A, Pinheiro I, Nawrocka A, Mallory AC, Tsvetkov P, Gil N, Carolis C, Buchholz F, Ulitsky I, Heard E et al. 2019. In-cell identification and measurement of RNA-protein interactions. *Nature communications* **10**: 5317.

Hasegawa Y, Brockdorff N, Kawano S, Tsutui K, Tsutui K, Nakagawa S. 2010. The matrix protein hnRNP U is required for chromosomal localization of Xist RNA. *Dev Cell* **19**: 469-476.

Hong Y, Team. RC. 2024. poibin: The Poisson Binomial Distribution.

Hu B, Petela N, Kurze A, Chan KL, Chapard C, Nasmyth K. 2015. Biological chromodynamics: a general method for measuring protein occupancy across the genome by calibrating ChIP-seq. *Nucleic Acids Res* **43**: e132.

Ietswaart R, Smalec BM, Xu A, Choquet K, McShane E, Jowhar ZM, Guegler CK, Baxter-Koenigs AR, West ER, Fu BXH et al. 2024. Genome-wide quantification of RNA flow across subcellular compartments reveals determinants of the mammalian transcript life cycle. *Mol Cell* **84**: 2765-2784 e2716.

Jachowicz JW, Strehle M, Banerjee AK, Blanco MR, Thai J, Guttman M. 2022. Xist spatially amplifies SHARP/SPEN recruitment to balance chromosome-wide silencing and specificity to the X chromosome. *Nat Struct Mol Biol* **29**: 239-249.

Kania EE, Fenix A, Marciniak DM, Lin Q, Bianchi S, Hristov B, Li S, Camplisson CK, Fields R, Beliveau BJ et al. 2024. Nascent transcript O-MAP reveals the molecular architecture of a single-locus subnuclear compartment built by RBM20 and the TTN RNA. *bioRxiv*.

Keane TM, Goodstadt L, Danecek P, White MA, Wong K, Yalcin B, Heger A, Agam A, Slater G, Goodson M et al. 2011. Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature* **477**: 289-294.

Kirk JM, Kim SO, Inoue K, Smola MJ, Lee DM, Schertzer MD, Wooten JS, Baker AR, Sprague D, Collins DW et al. 2018. Functional classification of long non-coding RNAs by k-mer content. *Nat Genet* **50**: 1474-1482.

Kolpa HJ, Creamer KM, Hall LL, Lawrence JB. 2022. SAF-A mutants disrupt chromatin structure through dominant negative effects on RNAs associated with chromatin. *Mamm Genome* **33**: 366-381.

Kolpa Heather J, Fackelmayer Frank O, Lawrence Jeanne B. 2016. SAF-A Requirement in Anchoring XIST RNA to Chromatin Varies in Transformed and Primary Cells. *Developmental Cell* **39**: 9-10.

Kramer NE, Davis ES, Wenger CD, Deoudes EM, Parker SM, Love MI, Phanstiel DH. 2022. Plotgardener: cultivating precise multi-panel figures in R. *Bioinformatics* **38**: 2042-2045.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**: 357-359.

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**: R25.

Lewis A, Green K, Dawson C, Redrup L, Huynh KD, Lee JT, Hemberger M, Reik W. 2006. Epigenetic dynamics of the Kcnq1 imprinted domain in the early embryo. *Development* **133**: 4203-4210.

Lewis A, Mitsuya K, Umlauf D, Smith P, Dean W, Walter J, Higgins M, Feil R, Reik W. 2004. Imprinting on distal chromosome 7 in the placenta involves repressive histone methylation independent of DNA methylation. *Nat Genet* **36**: 1291-1295.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078-2079.

Li S, Eberhard QE, Ni L, Calabrese JM. 2024. Improved functions for nonlinear sequence comparison using SEEKR. *RNA* **30**: 1408-1421.

Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**: 923-930.

Maechler M, Rousseeuw P, Struyf A, Hubert M, Hornik K. 2025. cluster: Cluster Analysis Basics and Extensions.

Mager J, Montgomery ND, de Villena FP, Magnuson T. 2003. Genome imprinting regulated by the mouse Polycomb group protein Eed. *Nat Genet* **33**: 502-507.

Mancini-Dinardo D, Steele SJ, Levorse JM, Ingram RS, Tilghman SM. 2006. Elongation of the Kcnq1ot1 transcript is required for genomic imprinting of neighboring genes. *Genes Dev* **20**: 1268-1282.

Markaki Y, Gan Chong J, Wang Y, Jacobson EC, Luong C, Tan SYX, Jachowicz JW, Strehle M, Maestrini D, Banerjee AK et al. 2021. Xist nucleates local protein gradients to propagate silencing across the X chromosome. *Cell* **184**: 6174-6192 e6132.

Martin G, Gruber AR, Keller W, Zavolan M. 2012. Genome-wide analysis of pre-mRNA 3' end processing reveals a decisive role of human cleavage factor I in the regulation of 3' UTR length. *Cell Rep* **1**: 753-763.

Masui O, Corbel C, Nagao K, Endo TA, Kezuka F, Diabangouaya P, Nakayama M, Kumon M, Koseki Y, Obuse C et al. 2023. Polycomb repressive complexes 1 and 2 are each essential for maintenance of X inactivation in extra-embryonic lineages. *Nat Cell Biol* **25**: 134-144.

Mattick JS, Amaral PP, Carninci P, Carpenter S, Chang HY, Chen LL, Chen R, Dean C, Dinger ME, Fitzgerald KA et al. 2023. Long non-coding RNAs: definitions, functions, challenges and recommendations. *Nat Rev Mol Cell Biol* **24**: 430-447.

Mattout A, Gaidatzis D, Padeken J, Schmid CD, Aeschimann F, Kalck V, Gasser SM. 2020. LSM2-8 and XRN-2 contribute to the silencing of H3K27me3-marked genes through targeted RNA decay. *Nat Cell Biol* **22**: 579-590.

McHugh CA, Chen CK, Chow A, Surka CF, Tran C, McDonel P, Pandya-Jones A, Blanco M, Burghard C, Moradian A et al. 2015. The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature* **521**: 232-236.

Minajigi A, Froberg JE, Wei C, Sunwoo H, Kesner B, Colognori D, Lessing D, Payer B, Boukhali M, Haas W et al. 2015. A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. *Science*.

Mitsuya K, Meguro M, Lee MP, Katoh M, Schulz TC, Kugoh H, Yoshida MA, Niikawa N, Feinberg AP, Oshimura M. 1999. LIT1, an imprinted antisense RNA in the human KvLQT1 locus identified by screening for differentially expressed transcripts using monochromosomal hybrids (vol 8, 1209, 1999). *Human Molecular Genetics* **8**: 1585-1585.

Moindrot B, Cerase A, Coker H, Masui O, Grijzenhout A, Pintacuda G, Schermelleh L, Nesterova TB, Brockdorff N. 2015. A Pooled shRNA Screen Identifies Rbm15, Spen, and Wtap as Factors Required for Xist RNA-Mediated Silencing. *Cell Rep* **12**: 562-572.

Mudge JM, Carbonell-Sala S, Diekhans M, Martinez JG, Hunt T, Jungreis I, Loveland JE, Arnan C, Barnes I, Bennett R et al. 2025. GENCODE 2025: reference gene annotation for human and mouse. *Nucleic Acids Res* **53**: D966-D975.

Mukherjee N, Calviello L, Hirsekorn A, de Pretis S, Pelizzola M, Ohler U. 2017. Integrative classification of human coding and noncoding genes through RNA metabolism profiles. *Nat Struct Mol Biol* **24**: 86-96.

Nagano T, Mitchell JA, Sanz LA, Pauler FM, Ferguson-Smith AC, Feil R, Fraser P. 2008. The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. *Science* **322**: 1717-1720.

Navarro-Cobos MJ, Brown CJ. 2025. Human XIST: Origin and Divergence of a cis-Acting Silencing RNA. *Noncoding RNA* **11**.

Neugebauer KM. 2019. Nascent RNA and the Coordination of Splicing with Transcription. *Cold Spring Harb Perspect Biol* **11**.

Nozawa RS, Boteva L, Soares DC, Naughton C, Dun AR, Buckle A, Ramsahoye B, Bruton PC, Saleeb RS, Arnedo M et al. 2017. SAF-A Regulates Interphase Chromosome Structure through Oligomerization with Chromatin-Associated RNAs. *Cell* **169**: 1214-1227 e1218.

Ntini E, Budach S, Vang Orom UA, Marsico A. 2023. Genome-wide measurement of RNA dissociation from chromatin classifies transcripts by their dynamics and reveals rapid dissociation of enhancer lncRNAs. *Cell Syst* **14**: 906-922 e906.

Pandey RR, Mondal T, Mohammad F, Enroth S, Redrup L, Komorowski J, Nagano T, Mancini-Dinardo D, Kanduri C. 2008. Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Mol Cell* **32**: 232-246.

Pandya-Jones A, Markaki Y, Serizay J, Chitiashvili T, Mancia Leon WR, Damianov A, Chronis C, Papp B, Chen CK, McKee R et al. 2020. A protein assembly mediates Xist localization and gene silencing. *Nature* **587**: 145-151.

Patil DP, Chen CK, Pickering BF, Chow A, Jackson C, Guttman M, Jaffrey SR. 2016. m(6)A RNA methylation promotes XIST-mediated transcriptional repression. *Nature* **537**: 369-+.

Perez G, Barber GP, Benet-Pages A, Casper J, Clawson H, Diekhans M, Fischer C, Gonzalez JN, Hinrichs AS, Lee CM et al. 2025. The UCSC Genome Browser database: 2025 update. *Nucleic Acids Res* **53**: D1243-D1249.

Pintacuda G, Wei G, Roustan C, Kirmizitas BA, Solcan N, Cerase A, Castello A, Mohammed S, Moindrot B, Nesterova TB et al. 2017. hnRNPK Recruits PCGF3/5-PRC1 to the Xist RNA B-Repeat to Establish Polycomb-Mediated Chromosomal Silencing. *Mol Cell* **68**: 955-969 e910.

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**: 841-842.

Quinn J, Kunath T, Rossant J. 2006. Mouse trophoblast stem cells. *Methods Mol Med* **121**: 125-148.

Quinodoz SA, Jachowicz JW, Bhat P, Ollikainen N, Banerjee AK, Goronzy IN, Blanco MR, Chovanec P, Chow A, Markaki Y et al. 2021. RNA promotes the formation of spatial compartments in the nucleus. *Cell* **184**: 5775-5790 e5730.

Ray D, Kazan H, Cook KB, Weirauch MT, Najafabadi HS, Li X, Gueroussov S, Albu M, Zheng H, Yang A et al. 2013. A compendium of RNA-binding motifs for decoding gene regulation. *Nature* **499**: 172-177.

Reimer KA, Mimoso CA, Adelman K, Neugebauer KM. 2021. Co-transcriptional splicing regulates 3' end cleavage during mammalian erythropoiesis. *Mol Cell* **81**: 998-1012 e1017.

Ridings-Figueroa R, Stewart ER, Nesterova TB, Coker H, Pintacuda G, Godwin J, Wilson R, Haslam A, Lilley F, Ruigrok R et al. 2017. The nuclear matrix protein CIZ1 facilitates localization of Xist RNA to the inactive X-chromosome territory. *Genes Dev* **31**: 876-888.

Sachani SS, Landschoot LS, Zhang L, White CR, MacDonald WA, Golding MC, Mann MRW. 2018. Nucleoporin 107, 62 and 153 mediate Kcnq1ot1 imprinted domain regulation in extraembryonic endoderm stem cells. *Nature communications* **9**: 2795.

Sakaguchi T, Hasegawa Y, Brockdorff N, Tsutsui K, Tsutsui KM, Sado T, Nakagawa S. 2016. Control of Chromosomal Localization of Xist by hnRNP U Family Molecules. *Dev Cell* **39**: 11-12.

Schertzer MD, Braceros KCA, Starmer J, Cherney RE, Lee DM, Salazar G, Justice M, Bischoff SR, Cowley DO, Ariel P et al. 2019a. lncRNA-Induced Spread of Polycomb Controlled by Genome Architecture, RNA Abundance, and CpG Island DNA. *Mol Cell* **75**: 523-537 e510.

Schertzer MD, Thulson E, Braceros KCA, Lee DM, Hinkle ER, Murphy RM, Kim SO, Vitucci ECM, Calabrese JM. 2019b. A piggyBac-based toolkit for inducible genome editing in mammalian cells. *RNA* **25**: 1047-1058.

Seidl CIM, Stricker SH, Barlow DP. 2006. The imprinted Air ncRNA is an atypical RNAPII transcript that evades splicing and escapes nuclear export. *Embo Journal* **25**: 3565-3575.

Sharp JA, Sparago E, Thomas R, Alimenti K, Wang W, Blower MD. 2025. Role of the SAF-A/HNRNPU SAP domain in X chromosome inactivation, nuclear dynamics, transcription, splicing, and cell proliferation. *PLoS Genet* **21**: e1011719.

Smilinich NJ, Day CD, Fitzpatrick GV, Caldwell GM, Lossie AC, Cooper PR, Smallwood AC, Joyce JA, Schofield PN, Reik W et al. 1999. A maternally methylated CpG island in KvLQT1 is associated with an antisense paternal transcript and loss of imprinting in Beckwith-Wiedemann syndrome. *P Natl Acad Sci USA* **96**: 8064-8069.

Smola MJ, Christy TW, Inoue K, Nicholson CO, Friedersdorf M, Keene JD, Lee DM, Calabrese JM, Weeks KM. 2016. SHAPE reveals transcript-wide interactions, complex structural domains, and protein interactions across the Xist lncRNA in living cells. *Proc Natl Acad Sci U S A* **113**: 10322-10327.

Sunwoo H, Colognori D, Froberg JE, Jeon Y, Lee JT. 2017. Repeat E anchors Xist RNA to the inactive X chromosomal compartment through CDKN1A-interacting protein (CIZ1). *Proc Natl Acad Sci U S A* **114**: 10654-10659.

Szanto A, Aguilar R, Kesner B, Blum R, Wang D, Cifuentes-Rojas C, Del Rosario BC, Kis-Toth K, Lee JT. 2021. A disproportionate impact of G9a methyltransferase deficiency on the X chromosome. *Genes Dev* **35**: 1035-1054.

Terranova R, Yokobayashi S, Stadler MB, Otte AP, van Lohuizen M, Orkin SH, Peters AH. 2008. Polycomb group proteins Ezh2 and Rnf2 direct genomic contraction and imprinted repression in early mouse embryos. *Dev Cell* **15**: 668-679.

Trotman JB, Abrash EW, Murvin MM, Braceros AK, Li S, Boyson SP, Salcido RT, Cherney RE, Bischoff SR, Kaufmann K et al. 2025a. Isogenic comparison of Airn and Xist reveals core principles of Polycomb recruitment by lncRNAs. *Mol Cell* **85**: 1117-1133 e1114.

Trotman JB, Porrello A, Schactler SA, DeLeon LE, Eberhard QE, Boyson SP, Zhang Z, Lee DM, Kirik SE, Sultana N et al. 2025b. Xist Repeat A coordinates an assembly of SR proteins to recruit SPEN and induce gene silencing. *bioRxiv*.

Tsue AF, Kania EE, Lei DQ, Fields R, McGann CD, Marciniak DM, Hershberg EA, Deng X, Kihiu M, Ong SE et al. 2024. Multiomic characterization of RNA microenvironments by oligonucleotide-mediated proximity-interactome mapping. *Nat Methods* **21**: 2058-2071.

Umlauf D, Goto Y, Cao R, Cerqueira F, Wagschal A, Zhang Y, Feil R. 2004. Imprinting along the Kcnq1 domain on mouse chromosome 7 involves repressive histone methylation and recruitment of Polycomb group complexes. *Nat Genet* **36**: 1296-1300.

Viphakone N, Sudbery I, Griffith L, Heath CG, Sims D, Wilson SA. 2019. Co-transcriptional Loading of RNA Export Factors Shapes the Human Transcriptome. *Mol Cell* **75**: 310-323 e318.

Wagschal A, Sutherland HG, Woodfine K, Henckel A, Chebli K, Schulz R, Oakey RJ, Bickmore WA, Feil R. 2008. G9a histone methyltransferase contributes to imprinting in the mouse placenta. *Mol Cell Biol* **28**: 1104-1113.

Wei G, Coker H, Rodermund L, Almeida M, Roach HL, Nesterova T, Brockdorff N. 2025. N6-methyladenosine and the NEXT complex direct Xist RNA turnover and X inactivation dynamics. *bioRxiv*.

Weidmann CA, Mustoe AM, Jariwala PB, Calabrese JM, Weeks KM. 2021. Analysis of RNA-protein networks with RNP-MaP defines functional hubs on RNA. *Nat Biotechnol* **39**: 347-356.

Wickham H. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag, New York.

Yeom KH, Pan Z, Lin CH, Lim HY, Xiao W, Xing Y, Black DL. 2021. Tracking pre-mRNA maturation across subcellular compartments identifies developmental gene regulation through intron retention and nuclear anchoring. *Genome Res* **31**: 1106-1119.

Yu B, Qi Y, Li R, Shi Q, Satpathy AT, Chang HY. 2021. B cell-specific XIST complex enforces X-inactivation and restrains atypical B cells. *Cell* **184**: 1790-1803 e1717.

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W et al. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**: R137.

Zhou H, Stein CB, Shafiq TA, Shipkovenska G, Kalocsay M, Paulo JA, Zhang J, Luo Z, Gygi SP, Adelman K et al. 2022. Rixosomal RNA degradation contributes to silencing of Polycomb target genes. *Nature* **604**: 167-174.