**Fig1 B**

Based on pipeline set up in <https://github.com/CalabreseLab/Airn_Xist_manuscript/tree/main/RIPanalysis> we got all34pr\_kallisto\_igg\_rpm\_TPMs\_updated\_04\_10\_24.csv, where kallisto is used to count the aligned reads per transcripts and IGG reads are subtracted from each RBP to get normalized rpm: RBP\_less\_igg.

By using addtpms.R, we merged table TSC-exp\_ESC-exp\_Chrom-assoc\_4\_9\_2024.csv

and all34pr\_kallisto\_igg\_rpm\_TPMs\_updated\_04\_10\_24.csv

into TSC-exp\_ESC-exp\_Chrom-assoc\_igg\_rpm\_4\_10\_2024.csv.

This table was filtered for length >=500, median exp >0.0625, chrom-fraction >0.75 and saved as TSC-exp\_ESC-exp\_Chrom-assoc\_igg\_rpm\_4\_10\_2024\_filtered\_07222024.csv

input: TSC-exp\_ESC-exp\_Chrom-assoc\_igg\_rpm\_4\_10\_2024.csv

code: Fig1B.R

plot:Fig1B

**Fig1C**

input: TSC-exp\_ESC-exp\_Chrom-assoc\_igg\_rpm\_4\_10\_2024\_filtered\_07222024.csv

code: XAK\_RIP\_r.R

plot:Fig1C

**Fig1D**

Based on the RIP pearson’s correlation r value calculated for XAK in Fig1C, and the biotype assignments for each transcript from vM25 gtf from GENCODE, get the top 1000 ranked transcripts that are most similar to X or A or K. Categorize the top 1000 transcripts into biotypes: protein coding, lncRNA and others, and also intron including (unspliced isoform), intron excluding (spliced isoform). Plot the percentage of each category (chrom divide by 19295, XAK top1000, divide by 1000). Perform Fisher’s exact test to query: whether count of transcripts from lncRNA vs protein coding differ in the top 1000 for AKX vs in the total chromatin-associated population; and within each biotype, whether the proportions of intron-excluding vs including differ in the top 1000 for AKX vs the total chromatin-associated population.

input: allgenes\_XAKRIPr\_XAKnetworkr\_common\_sil\_modu\_genetype\_06302025.csv

code: XAK\_RIPr\_biotype\_Fisher.R

plot:Fig1D