RIP data analysis pipeline:

code: run\_3pr\_classify.sh which calls SL\_3pr\_classify.sh

code: run\_vet\_peaks\_bedfiles.sh which calls SL\_vet\_peaks\_bedfiles.sh

merge the columns:

ls \*\_classify/\*\_vet\_peaks.txt | wc -l # this will be the number of columns

paste \*\_classify/\*\_vet\_peaks.txt > vet\_peaks\_34pr\_5stats.txt

cat vet\_peaks\_34pr\_5stats.txt | tr '\t' ',' > vet\_peaks\_34pr\_5stats\_231128.csv

code: run\_rank\_rip\_signal\_2reps.sh which calls SL\_rank\_rip\_signal\_2reps.sh for all protein except:

For nxf1 and u2af65 run:

sbatch SL\_rank\_rip\_signal.sh nxf1

sbatch SL\_rank\_rip\_signal.sh u2af65

Copy over all {protein}\_kallisto\_igg\_rpm.csv to folder RIP\_data\_process\kallisto\_igg\_rpm\_all

And then merge all {protein}\_rpm and {protein}\_rpm\_over\_igg to all34pr\_kallisto\_igg\_rpm\_TPMs\_updated\_04\_10\_24.csv in R, using code: kallisto\_igg\_rpm\_merge.R (make sure the gene names are all right) 114783 rows

The output file: all34pr\_kallisto\_igg\_rpm\_TPMs\_updated\_04\_10\_24.csv

merge two files to add chr and cyto info and chromatin enrichment scores, with addtpms.R:

TSC-exp\_ESC-exp\_Chrom-assoc\_4\_9\_2024.csv

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

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