Figure 5A

RAC

24/08/2020

```
knitr::opts_chunk$set(warning=FALSE, message=FALSE, tidy.opts = list(width.cutoff = 60), tidy = TRUE)

library(ggplot2)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

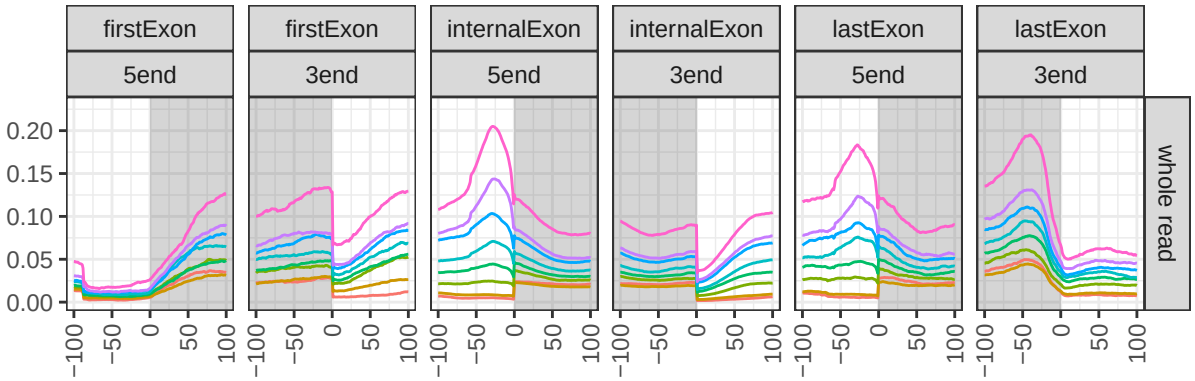
```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(tidyr)

COUNTSS="../../data/xiCLIP_read2_mono_multi_exonic.calculate_coverage_gene_structures_exonsover99nt_LIN

EXPRESSION_VECTOR_FILEPATH="../../data/log2_mean_cov_RNAseq_TTseq.RData"

ANNOTATION_BED_FILEPATH="../../data/hg38_HeLa_trimmed_loci_major_primary_isoform_annotated.exonNumber.s
```

#Figure 5 A Figure 5A was processed on the cluster for speed. It was processed in the same way that Figure 3B was

```
annotate_rect <- data.frame(XMIN = c(0, -Inf), XMAX = c(Inf,
    0), YMAX = c(Inf, Inf), YMIN = c(-Inf, -Inf), region = c("5end",
    "3end"))


DF <- read.table(COUNTSS, header = F) %>%
    setNames(c("Protein", "Rep", "Timepoint", "readType", "rem",
        "geneDescription", "DistToLandmark", "meanCoverage",
        "nAnno", "Timepoint_f", "nANNO2", "region")) %>%
    select(-rem, -nANNO2) %>%
    separate(geneDescription, into = c("TU", "ExonPosition"),
        sep = "-") %>%
    mutate_at(vars(DistToLandmark, meanCoverage, nAnno), .funs = as.numeric) %>%
    mutate(region = factor(region, levels = c("5end", "3end")),
        Timepoint = factor(Timepoint, levels = c("PBSDRB", "t00",
            "t05", "t10", "t15", "t20", "t40", "t60", "DMSO")),
        Timepoint_f = factor(Timepoint_f, levels = c("t00", "t05",
            "t10", "t15", "t20", "t40", "t60", "DMSO")), GeneStructure = factor(ExonPosition,
```

```
            levels = c("first", "internal", "last"))) %>%
    filter(!(Protein == "CBP20" & Rep == "3")) %>%
    mutate(readType = gsub("read2", "whole read", readType))

plot <- DF %>%
    filter(Protein == "RBM7") %>%
    filter(TU == "multiExonicGene" & Timepoint != "negative") %>%
    ggplot() + geom_rect(data = annotate_rect, aes(xmin = XMIN,
    xmax = XMAX, ymin = YMIN, ymax = YMAX), alpha = 0.25) + geom_line(aes(x = DistToLandmark,
    y = meanCoverage, col = Timepoint_f), stat = "summary") +
    facet_grid(readType ~ ExonPosition + region, scale = "free") +
    theme_bw()
# labs(subtitle = 'on cluster, read2 coverage ctrl_RNAseq
# log2 > 1')



plot + theme(axis.text.x = element_text(angle = 90, vjust = 0.5,
    hjust = 1), legend.position = "none") + xlab("") + ylab("")
```




```
# facet_grid(readType ~ GeneStructure + region, scale =
# 'free') + ggsave('fig5a.pdf', height = 2, width =6)
```