Figure 4 ChIP-seq signals for various regulatory targets

Zheng Zuo

04/15/2022

Contents

```
require(dplyr)
require(GenomicRanges)

#FSD1, STOML3, MARK1, SPC24, WDR19, CERKL, TTC26, CEP162
load("./ZFP3.annotated.sites.RData")
Genes.Of.Interest <- c("FSD1", "STOML3", "MARK1", "SPC24", "WDR19", "CERKL", "TTC26",
                       "CEP162", "MAP1S", "CEP95", "PKN2", "C7orf43", "KAT2B")

require(EnsDb.Hsapiens.v86)
require(ChIPpeakAnno)
require(GenomicRanges)
ZFP3.peaks.regions <- ChIPpeakAnno::toGRanges(EnsDb.Hsapiens.v86, feature="gene") %>%
                      base::subset(gene_name %in% Genes.Of.Interest)


ZFP3.peaks.regions <- promoters(ZFP3.peaks.regions,
                                upstream = 500, downstream = 500)

ZFP3.peaks.regions
```

```
## GRanges object with 13 ranges and 1 metadata column:
##                    seqnames              ranges strand |   gene_name
##                       <Rle>           <IRanges>  <Rle> | <character>
##   ENSG00000065243     chr1   88683722-88684721      + |        PKN2
##   ENSG00000116141     chr1 220527683-220528682      + |       MARK1
##   ENSG00000133115    chr13   38990567-38991566      - |      STOML3
##   ENSG00000258890    chr17   64506088-64507087      + |       CEP95
##   ENSG00000105255    chr19     4304100-4305099      + |        FSD1
##               ...      ...                 ...    ... .         ...
##   ENSG00000114166     chr3   20039523-20040522      + |       KAT2B
##   ENSG00000157796     chr4   39181904-39182903      + |       WDR19
##   ENSG00000135315     chr6   84227136-84228135      - |      CEP162
##   ENSG00000146826     chr7 100158216-100159215      - |     C7orf43
##   ENSG00000105948     chr7 139133244-139134243      + |       TTC26
##   -------
##   seqinfo: 357 sequences from GRCh38 genome
```

```
hg38.genomeAxis.track <- Gviz::GenomeAxisTrack()

require(ensembldb)
require(EnsDb.Hsapiens.v86)
edb <- EnsDb.Hsapiens.v86
seqlevelsStyle(edb) <- "UCSC"

require(Gviz)
require(ggplotify)
hg38.geneRegion.track <- Gviz::GeneRegionTrack(ensembldb::getGeneRegionTrackForGviz(edb,
                                                                    chromosome = seqnames(ZFP3.peaks.regions)
                                                                    ),
                                               name = "Genes Model",
                                               transcriptAnnotation = "symbol",
                                               background.title = "brown",
                                               fill = "#777777")

ZFP3.ChIP.SKN.track <- Gviz::DataTrack(range = rtracklayer::import.bw("./ENCFF774MKP.bigWig",
                                                                      which = ZFP3.peaks.regions),
                       genome = "hg38", type = "l", col = "#5BB4E5",
                       groups = factor("SK-N-SH", levels = c("HEK293", "SK-N-SH")),
                       name = "ZFP3 ChIP-seq signal (SK-N-SH cells)", ylim = c(0, 100))


ZFP3.ChIP.H293.track <- Gviz::DataTrack(range = rtracklayer::import.bw("./ENCFF655DZE.bigWig",
                                                                      which = ZFP3.peaks.regions),
                       genome = "hg38", type = "l", col = "#E6A024",
                       groups = factor("HEK293", levels = c("HEK293", "SK-N-SH")),
                       name = "ZFP3 ChIP-seq signals (HEK293 cells)", ylim = c(0, 100))

ZFP3.ChIP.tracks <- Gviz::OverlayTrack(trackList=list(ZFP3.ChIP.H293.track,
                                                      ZFP3.ChIP.SKN.track
                                                     ))
```

```
ZFP3.highlight.sites <- subsetByOverlaps(ZFP3.annotated.sites, subset(ZFP3.peaks.regions, gene_name ==
                                         ignore.strand = TRUE) %>%
                        subset(predicted.Full.Energy < -7)


highlightBox <-          Gviz::HighlightTrack(trackList = list(ZFP3.ChIP.tracks),
                                              start     = start(ZFP3.highlight.sites),
                                              end       = end(ZFP3.highlight.sites),
                                              chromosome= seqnames(ZFP3.highlight.sites), col = NA

                                    )

as.ggplot(function(){
  ZFP3.peaks.regions %>%
  subset(gene_name == "STOML3") %>%
  Gviz::plotTracks(trackList = list(hg38.geneRegion.track,
                                    highlightBox),
                                    collapseTranscripts = "longest",
                                    chromosome = as.character(seqnames(.)),
```

```
                                    from = start(.),
                                    to =   end(.),
                                    window = -1,
                                    windowSize = 4, ylim = c(0, 300),
                                    reverseStrand = TRUE, legend = FALSE)
}) -> plot.STOML3

#ggsave("STOML3 track.svg", height = 1.3, width = 8)
```

```
ZFP3.highlight.sites <- subsetByOverlaps(ZFP3.annotated.sites, subset(ZFP3.peaks.regions, gene_name ==
                                    ignore.strand = TRUE) %>%
                        subset(predicted.Full.Energy < -7)


highlightBox <-          Gviz::HighlightTrack(trackList = list(ZFP3.ChIP.tracks),
                                              start     = start(ZFP3.highlight.sites),
                                              end       = end(ZFP3.highlight.sites),
                                              chromosome= seqnames(ZFP3.highlight.sites), col = NA

                                     )

as.ggplot(function(){
  ZFP3.peaks.regions %>%
  subset(gene_name == "FSD1") %>%
  Gviz::plotTracks(trackList = list(hg38.geneRegion.track,
                                    highlightBox
                                    ),
                                    collapseTranscripts = "longest",
                                    chromosome = as.character(seqnames(.)),
                                    from = start(.),
                                    to =   end(.),
                                    window = -1,
                                    windowSize = 4, ylim = c(0, 2100),
                                    reverseStrand = FALSE, legend = FALSE)
}) -> plot.FSD1


#ggsave("FSD1 track.svg", height = 1.6, width = 8)
```

```
ZFP3.highlight.sites <- subsetByOverlaps(ZFP3.annotated.sites, subset(ZFP3.peaks.regions, gene_name ==
                                    ignore.strand = TRUE) %>%
                        subset(predicted.Full.Energy < -7)

highlightBox <-          Gviz::HighlightTrack(trackList = list(ZFP3.ChIP.tracks),
                                              start     = start(ZFP3.highlight.sites),
                                              end       = end(ZFP3.highlight.sites),
                                              chromosome= seqnames(ZFP3.highlight.sites), col = NA

                                     )

as.ggplot(function(){
  ZFP3.peaks.regions %>%
  subset(gene_name == "MARK1") %>%
```

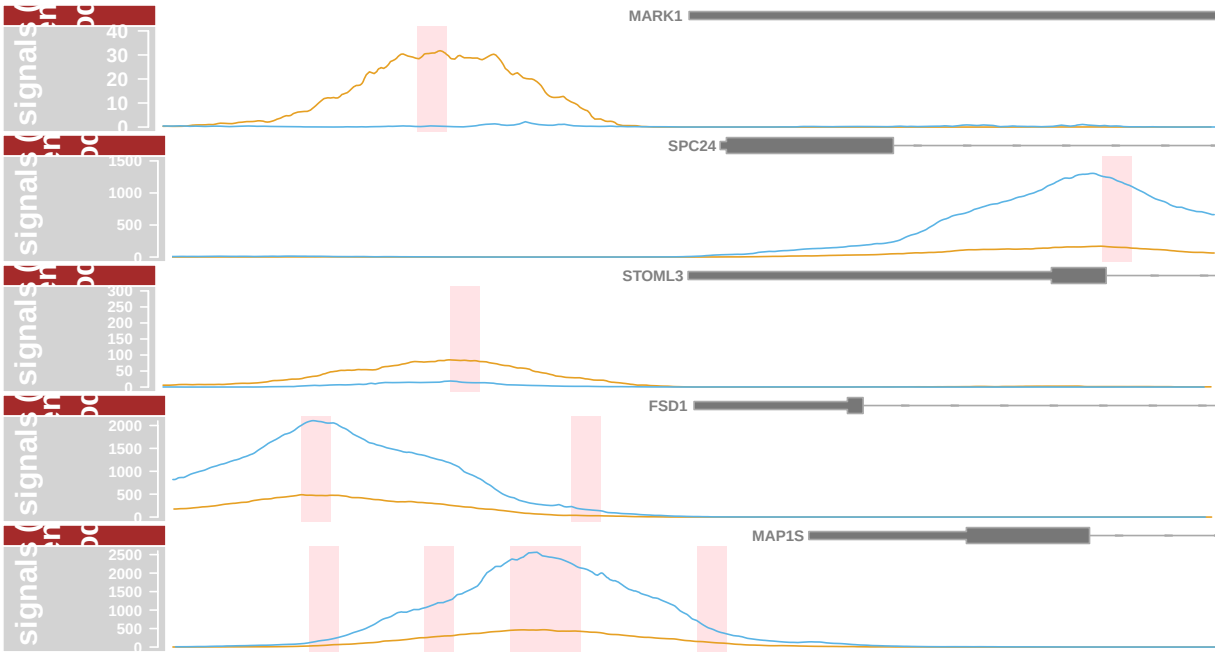
```
  Gviz::plotTracks(trackList = list(hg38.geneRegion.track,
                                    highlightBox),
                                    collapseTranscripts = "longest",
                                    chromosome = as.character(seqnames(.)),
                                    from = start(.),
                                    to =   end(.),
                                    window = -1,
                                    windowSize = 4, ylim = c(0, 40),
                                    reverseStrand = FALSE, legend = FALSE)
  }) -> plot.MARK1


#ggsave("MARK1 track.svg", height = 1.6, width = 8)
```

```
ZFP3.highlight.sites <- subsetByOverlaps(ZFP3.annotated.sites, subset(ZFP3.peaks.regions, gene_name ==
                                         ignore.strand = TRUE) %>%
                         subset(predicted.Full.Energy < -7)

highlightBox <-          Gviz::HighlightTrack(trackList = list(ZFP3.ChIP.tracks),
                                              start     = start(ZFP3.highlight.sites),
                                              end       = end(ZFP3.highlight.sites),
                                              chromosome= seqnames(ZFP3.highlight.sites), col = NA

                                     )

as.ggplot(function(){
  ZFP3.peaks.regions %>%
  subset(gene_name == "SPC24") %>%
  Gviz::plotTracks(trackList = list(hg38.geneRegion.track,
                                    highlightBox),
                                    collapseTranscripts = "longest",
                                    chromosome = as.character(seqnames(.)),
                                    from = start(.),
                                    to =   end(.),
                                    window = -1,
                                    windowSize = 4, ylim = c(0, 1500),
                                    reverseStrand = TRUE, legend = FALSE)
  }) -> plot.SPC24


#ggsave("SPC24 track.svg", height = 1.6, width = 8)
```

```
ZFP3.highlight.sites <- subsetByOverlaps(ZFP3.annotated.sites, subset(ZFP3.peaks.regions, gene_name ==
                                         ignore.strand = TRUE) %>%
                         subset(predicted.Full.Energy < -7)

highlightBox <-          Gviz::HighlightTrack(trackList = list(ZFP3.ChIP.tracks),
                                              start     = start(ZFP3.highlight.sites),
                                              end       = end(ZFP3.highlight.sites),
                                              chromosome= seqnames(ZFP3.highlight.sites), col = NA

                                     )
```

```
as.ggplot(function(){
  ZFP3.peaks.regions %>%
  subset(gene_name == "MAP1S") %>%
  Gviz::plotTracks(trackList = list(hg38.geneRegion.track,
                                    highlightBox),
                                    collapseTranscripts = "longest",
                                    chromosome = as.character(seqnames(.)),
                                    from = start(.),
                                    to =   end(.),
                                    window = -1,
                                    windowSize = 4, ylim = c(0, 2600),
                                    reverseStrand = FALSE, legend = FALSE)}) -> plot.MAP1S

#ggsave("MAP1S track.svg", height = 1.6, width = 8)
```

```
cowplot::plot_grid(plot.MARK1,
                   plot.SPC24,
                   plot.STOML3,
                   plot.FSD1,
                   plot.MAP1S,
                   nrow = 6)
```




```
#ggsave("Genome tracks.svg", height = 5, width = 8)
```