Figure 5B Chromatin signals correlation analysis

Zheng Zuo

04/15/2022

Contents

```r
require(dplyr)
require(GenomicRanges)

load("../ZFP3.annotated.sites.RData")

ZFP3.sites <- subset(ZFP3.annotated.sites,
                     predicted.Upstream.Energy<(-4))
```

```r
ZFP3.H293.bigWig <- "../ENCFF655DZE.bigWig"
ZFP3.SKN.bigWig  <- "../ENCFF774MKP.bigWig"

DNase.H293T.bigWig<- "../DNase-seq signals/ENCFF529BOG.H293T.bigWig"
DNase.SKN.bigWig  <- "../DNase-seq signals/ENCFF280RMA.SKN.bigWig"

H3K4me1.H293.bigWig <- "ENCFF003LZR.HEK293.H3K4me1.bigWig"
H3K4me3.H293.bigWig <- "ENCFF315TAU.HEK293.H3K4me3.bigWig"
H3K9me3.H293.bigWig <- "ENCFF758LNF.HEK293.H3K9me3.bigWig"
H3K27ac.H293.bigWig <- "ENCFF157TGK.HEK293.H3K27ac.bigWig"

H3K4me1.SKN.bigWig  <- "ENCFF056LYG.SKN.H3K4me1.bigWig"
H3K4me3.SKN.bigWig  <- "ENCFF943FNS.SKN.H3K4me3.bigWig"
H3K9me3.SKN.bigWig  <- "ENCFF655EVA.SKN.H3K9me3.bigWig"
H3K27ac.SKN.bigWig  <- "ENCFF287KJY.SKN.H3K27ac.bigWig"
require(genomation)

ZFP3.sites.w400    <- promoters(ZFP3.sites, upstream = 200, downstream = 228)


ZFP3.sites$score.ZFP3.HEK293   <- ScoreMatrixBin(target = ZFP3.H293.bigWig,
                                                 windows= ZFP3.sites.w400,
                                                 bin.num= 1, bin.op = "mean",
                                                 type  = "bigWig", strand.aware = TRUE)@.Data


ZFP3.sites$score.ZFP3.SKN      <- ScoreMatrixBin(target = ZFP3.SKN.bigWig,
                                                 windows= ZFP3.sites.w400,
                                                 bin.num= 1, bin.op = "mean",
                                                 type  = "bigWig", strand.aware = TRUE)@.Data
```

```r
ZFP3.sites$score.DNase.HEK293   <- ScoreMatrixBin(target = DNase.H293T.bigWig,
                                                  windows= ZFP3.sites.w400,
                                                  bin.num= 1, bin.op = "mean",
                                                  type   = "bigWig", strand.aware = TRUE)@.Data


ZFP3.sites$score.DNase.SKN      <- ScoreMatrixBin(target = DNase.SKN.bigWig,
                                                  windows= ZFP3.sites.w400,
                                                  bin.num= 1, bin.op = "mean",
                                                  type   = "bigWig", strand.aware = TRUE)@.Data


ZFP3.sites$score.H3K27ac.HEK293 <- ScoreMatrixBin(target = H3K27ac.H293.bigWig,
                                                  windows= ZFP3.sites.w400,
                                                  bin.num= 1, bin.op = "mean",
                                                  type   = "bigWig", strand.aware = TRUE)@.Data


ZFP3.sites$score.H3K27ac.SKN    <- ScoreMatrixBin(target = H3K27ac.SKN.bigWig,
                                                  windows= ZFP3.sites.w400,
                                                  bin.num= 1, bin.op = "mean",
                                                  type   = "bigWig", strand.aware = TRUE)@.Data


ZFP3.sites$score.H3K4me3.HEK293 <- ScoreMatrixBin(target = H3K4me3.H293.bigWig,
                                                  windows= ZFP3.sites.w400,
                                                  bin.num= 1, bin.op = "mean",
                                                  type   = "bigWig", strand.aware = TRUE)@.Data


ZFP3.sites$score.H3K4me3.SKN    <- ScoreMatrixBin(target = H3K4me3.SKN.bigWig,
                                                  windows= ZFP3.sites.w400,
                                                  bin.num= 1, bin.op = "mean",
                                                  type   = "bigWig", strand.aware = TRUE)@.Data


ZFP3.sites$score.H3K4me1.HEK293 <- ScoreMatrixBin(target = H3K4me1.H293.bigWig,
                                                  windows= ZFP3.sites.w400,
                                                  bin.num= 1, bin.op = "mean",
                                                  type   = "bigWig", strand.aware = TRUE)@.Data


ZFP3.sites$score.H3K4me1.SKN    <- ScoreMatrixBin(target = H3K4me1.SKN.bigWig,
                                                  windows= ZFP3.sites.w400,
                                                  bin.num= 1, bin.op = "mean",
                                                  type   = "bigWig", strand.aware = TRUE)@.Data


ZFP3.sites$score.H3K9me3.HEK293 <- ScoreMatrixBin(target = H3K9me3.H293.bigWig,
                                                  windows= ZFP3.sites.w400,
                                                  bin.num= 1, bin.op = "mean",
                                                  type   = "bigWig", strand.aware = TRUE)@.Data
```

```r
ZFP3.sites$score.H3K9me3.SKN      <- ScoreMatrixBin(target = H3K9me3.SKN.bigWig,
                                                    windows= ZFP3.sites.w400,
                                                    bin.num= 1, bin.op = "mean",
                                                    type   = "bigWig", strand.aware = TRUE)@.Data

ggLower <- function(data, mapping, ...) {
  ggplot(data, mapping) +
    geom_point(size = 0.4, alpha = 0.4) +
    ggpubr::stat_cor(aes(label = ..r.label..), size = 2, na.rm = TRUE) +
    theme_bw() +
    theme(panel.grid.minor.x = element_blank(),
        panel.grid.minor.y = element_blank()) %>%
    return()
    #geom_abline(intercept = 0, slope = 1, col = 'darkred')
}

ggDiag <- function(data, mapping, ...) {
  ggplot(data, mapping) +
    geom_histogram(aes(y=..density..), bins = 20, fill = 'steelblue', color='black', alpha=.4) +
    geom_density(aes(y=..density..)) +
    theme_minimal() +
    theme(panel.grid.minor.x = element_blank(),
          panel.grid.minor.y = element_blank()) %>%
  return()
}




require(GGally)
ZFP3.sites %>% as_tibble() %>%
  mutate(logFC.ZFP3    = log10(score.ZFP3.SKN/score.ZFP3.HEK293),
         logFC.H3K4me3 = log10(score.H3K4me3.SKN/score.H3K4me3.HEK293),
         logFC.H3K4me1 = log10(score.H3K4me1.SKN/score.H3K4me1.HEK293),
         logFC.H3K27ac = log10(score.H3K27ac.SKN/score.H3K27ac.HEK293),
         logFC.H3K9me3 = log10(score.H3K9me3.SKN/score.H3K9me3.HEK293),
         logFC.DNase   = log10(score.DNase.SKN/score.DNase.HEK293)) %>%
  mutate(Feature = if_else(abs(distancetoFeature)<=500, "TSS proximal", "TSS distal")) %>%
 # na.omit(cols = c("logFC.ZFP3", "logFC.DNase","logFC.H3K4me3", "logFC.H3K4me1","logFC.H3K27ac", "logF
  dplyr::filter(is.finite(logFC.ZFP3),
                is.finite(logFC.DNase), is.finite(logFC.H3K4me3),
                is.finite(logFC.H3K4me1), is.finite(logFC.H3K9me3)) %>%
  dplyr::filter(Property == "merged") %>%
  GGally::ggpairs(
                  lower = list(continuous = ggLower),
                  diag  = list(continuous = ggDiag),
                  upper = list(continuous = wrap(ggally_cor, display_grid = FALSE)),
                  columns = c("logFC.DNase","logFC.H3K27ac",
                              "logFC.H3K4me1","logFC.H3K4me3",
                              "logFC.H3K9me3", "logFC.ZFP3"),
                  columnLabels = c("DNase-seq", "H3K27ac","H3K4me1", "H3K4me3", "H3K9me3", "ZFP3"))
```
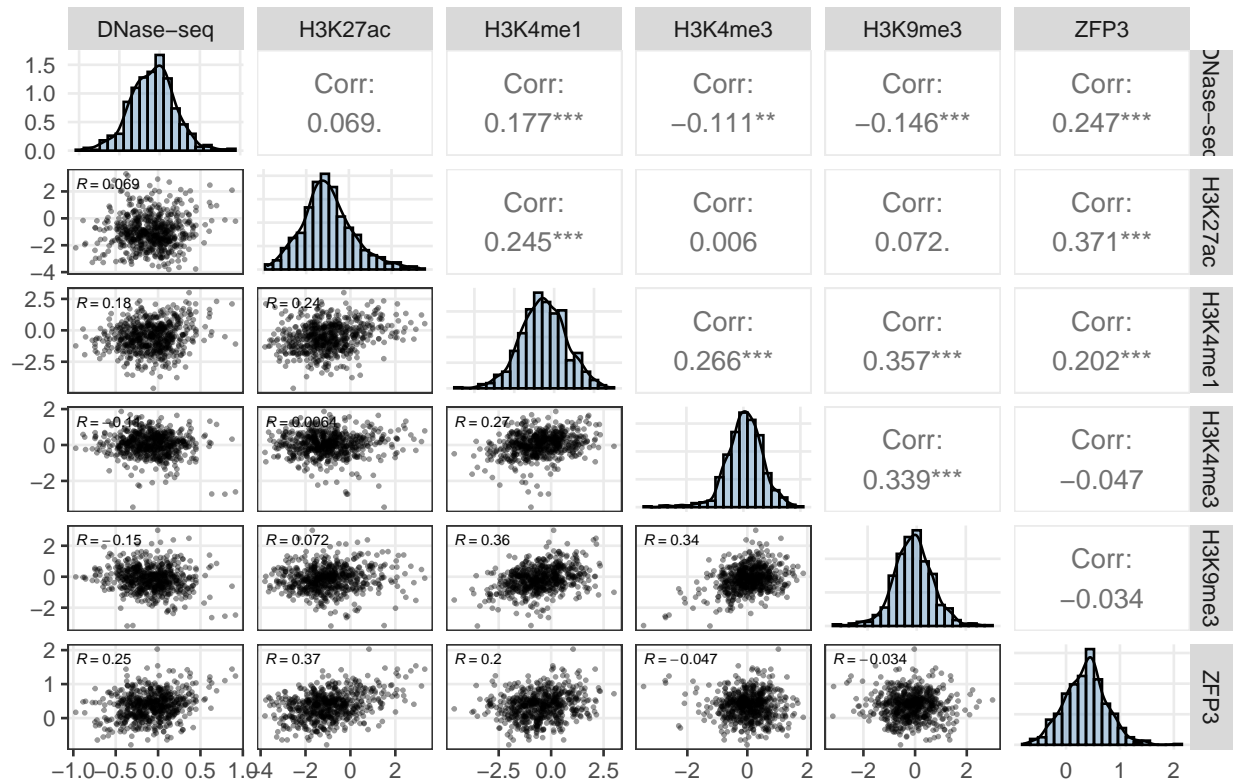
```
#ggthemes::scale_color_colorblind()

#ggsave("Signals correlations map.svg", width = 7, height = 4.5)
```