Distribution of m6A sites (MetaTX)

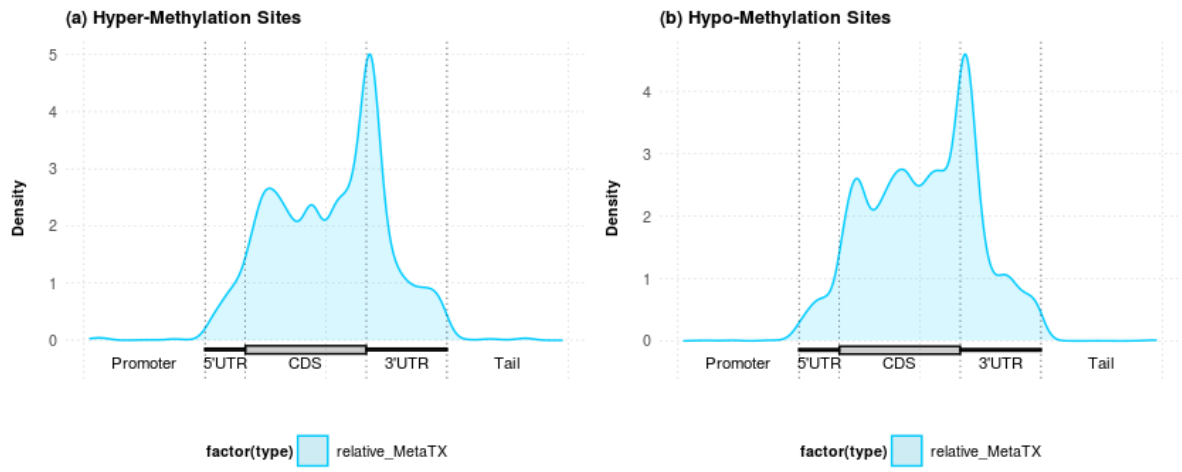
We are going to use MetaTX to visualize the distribution of methylation sites.

Visualization of the Distribution of Peaks

The following code produces separate figures of the distribution of hyper-methylation sites and hypo-methylation sites.

```
# Load libraries
library(MetaTX)
library(rtracklayer)
library(readr)
library(bedr)
library(genomation)
library(GenomicRanges)

# Import BED file from exomePeak2
file <- "exomePeak2_output_diff_1strand/DiffMod.bed"
gr_obj =  import(file)

# Separate by hyper and hypo methylation sites
data1 <- data[strand(gr_obj) == "+",]
data2 <- data[strand(gr_obj) == "-",]
df1 <- data.frame(seqnames=seqnames(data1),
                  starts=start(data1),
                  ends=end(data1),
                  names=elementMetadata(data1)$name,
                  scores=elementMetadata(data1)$score,
                  strands=strand(data1))
df2 <- data.frame(seqnames=seqnames(data2),
                  starts=start(data2),
                  ends=end(data2),
                  names=elementMetadata(data2)$name,
                  scores=elementMetadata(data2)$score,
                  strands=strand(data2))
write.table(df1, file="Mod_metaTX_diff_pos.bed", quote=F, sep="\t", row.names=F,
col.names=F)
write.table(df2, file="Mod_metaTX_diff_neg.bed", quote=F, sep="\t", row.names=F,
col.names=F)

# Import separated bed files
file_pos <- "Mod_metaTX_diff_pos.bed"
file_neg <- "Mod_metaTX_diff_neg.bed"
gr_obj_pos <- import(file_pos)
gr_obj_neg <- import(file_neg)
gr_obj_pos <- resize(gr_obj_pos, width = 1, fix = "center")
gr_obj_neg <- resize(gr_obj_neg, width = 1, fix = "center")

# Download information about mRNA components
txdb <- TxDb.Hsapiens.UCSC.hg19.knownGene
cds_by_tx0_1 <- cdsBy(txdb, "tx")
fiveUTR_tx0_1 <- fiveUTRsByTranscript(txdb,use.names=FALSE)
threeUTR_tx0_1 <- threeUTRsByTranscript(txdb,use.names=FALSE)
```

```
# Map peaks to the RNA model
remap_results_m6A_1 <- remapCoord(features = gr_obj_pos, txdb = txdb, num_bin =
10, includeNeighborDNA = TRUE, cds_by_tx0 = cds_by_tx0_1, fiveUTR_tx0 =
fiveUTR_tx0_1,
                                  threeUTR_tx0 = threeUTR_tx0_1)
# Plot 1
p1 <-  metaTXplot(remap_results_m6A_1,
                  num_bin              = 10,
                  includeNeighborDNA   = TRUE,
                  relativeProportion   = c(1, 3, 2, 3),
                  title  = '(a) Hyper-Methylation Sites',
                  legend = 'relative',
                  type = 'relative'
)

# Map peaks to the RNA model
remap_results_m6A_2 <- remapCoord(features = gr_obj_neg, txdb = txdb, num_bin =
10, includeNeighborDNA = TRUE, cds_by_tx0 = cds_by_tx0_1, fiveUTR_tx0 =
fiveUTR_tx0_1,
                                  threeUTR_tx0 = threeUTR_tx0_1)
# Plot 1
p2 <-  metaTXplot(remap_results_m6A_2,
                  num_bin              = 10,
                  includeNeighborDNA   = TRUE,
                  relativeProportion   = c(1, 3, 2, 3),
                  title  = '(b) Hypo-Methylation Sites',
                  legend = 'relative',
                  type = 'relative'
)

# Plot all
ggdraw() +
  draw_plot(p1, 0, 0, .5, 1) +
  draw_plot(p2, .5, 0, .5, 1)
```



Report Isoform Probabilities

MetaTX also provides a function for returning the probabilities of a particular feature being located on different isoforms.

```
isoform_probs <- isoformProb(remap_results_m6A_1, num_bin = 10,
includeNeighborDNA = TRUE, lambda = 2)
write.csv(isoform_probs, "isoform_probs.csv")
```

Here are the first few rows of the outputs.

index_trans (double)	index_methyl (double)	seqnames (character)	methyl_pos (double)	strand (character)	trans_ID (double)	isoform_prob (double)
1	1	chr19	58867266	-	70456	0.00000000
2	1	chr19	58867266	-	70457	0.51380127
3	1	chr19	58867266	-	70458	0.48619873
4	2	chr20	43269048	-	72132	0.00000000
5	3	chr18	25532067	-	65378	0.43630632
6	3	chr18	25532067	-	65379	0.56369368
7	4	chr3	101396111	+	14200	1.00000000
8	5	chrX	119387453	+	76492	0.45666710
9	5	chrX	119387453	+	76493	0.47472665