

Unibind_to_pdf

2024-01-29

```
library(ggplot2)
library(here); i_am("R/Figure3/Fig3D_Unibind_to_pdf.Rmd")

## here() starts at /projects/imb-pkbphil/sp/rnaseq/six_donor_trans/splicing_paper

## here() starts at /projects/imb-pkbphil/sp/rnaseq/six_donor_trans/splicing_paper

LEAF = here("31_leafcutter/promoter_binding/")

results = read.delim(file.path(LEAF, "UniBind_twoSets_jxq3t3oo_20240129/allEnrichments.tsv"))
head(results)
```

```
##   userSet dbSet collection pValueLog oddsRatio support rnkPV rnkOR rnkSup
## 1      1  3283      PPARG  5.617537 13.602278      24    1    80   1324
## 2      1  3290      PPARG  3.688366  3.654232      32    2    96    825
## 3      1  3284      PPARG  3.452267  2.106618      82    3   182    13
## 4      1  3291      PPARG  3.049748  3.255681      29    4   102   1013
## 5      1  3285      PPARG  2.897468  1.975921      75    5   281    28
## 6      1  3297      PPARG  2.784890  2.360692      43    6   165   429
##   maxRnk meanRnk  b  c  d
## 1   1324     468  2 180 205
## 2    825     308 10 172 197
## 3    182      66 50 122 157
## 4   1013     373 10 175 197
## 5    281     105 47 129 160
## 6    429     200 21 161 186
##                                     description
## 1 PPARG Simpson Golabi Behmel Syndrome adipocytes PPARG
## 2      PPARG subcutaneous white adipose tissue PPARG
## 3      PPARG hMADS-3 (adipose-derived stem cells) PPARG
## 4      PPARG subcutaneous white adipose tissue PPARG
## 5      PPARG hMADS-3 (adipose-derived stem cells) PPARG
## 6      PPARG hASC (preadipocytes) PPARG
##                                     cellType tissue antibody
## 1 Simpson Golabi Behmel Syndrome adipocytes  NULL  PPARG
## 2      subcutaneous white adipose tissue  NULL  PPARG
## 3      hMADS-3 (adipose-derived stem cells)  NULL  PPARG
## 4      subcutaneous white adipose tissue  NULL  PPARG
## 5      hMADS-3 (adipose-derived stem cells)  NULL  PPARG
## 6      hASC (preadipocytes)  NULL  PPARG
##                                     treatment dataSource
## 1      NULL      GTRD
## 2      GTRD
```

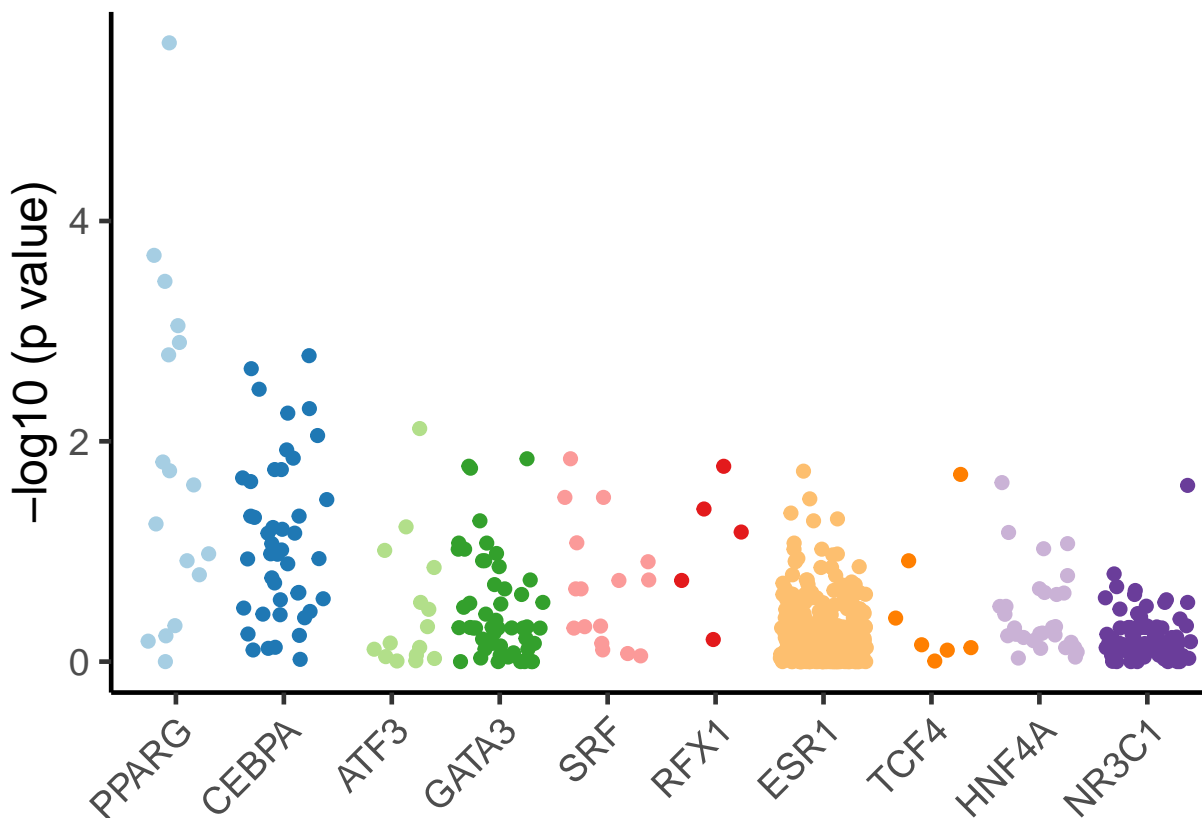
```
## 3 differentiaiton inducators      GTRD
## 4                                GTRD
## 5 differentiaiton inducators      GTRD
## 6                                GEO
##
##                                     filename
## 1 EXP000907_Simpson-Golabi-Behmel-Syndrome-adipocytes_PPARG_MACS_PPARG_MA0065.2.damo.pwm.bed
## 2     EXP034425_subcutaneous-white-adipose-tissue_PPARG_MACS_PPARG_MA0065.2.damo.pwm.bed
## 3     EXP033780_hMADS-3--adipose-derived-stem-cells-_PPARG_MACS_PPARG_MA0065.2.damo.pwm.bed
## 4     EXP034426_subcutaneous-white-adipose-tissue_PPARG_MACS_PPARG_MA0065.2.damo.pwm.bed
## 5     EXP033781_hMADS-3--adipose-derived-stem-cells-_PPARG_MACS_PPARG_MA0065.2.damo.pwm.bed
## 6                                GSE21366_pparg_asc_MACS_PPARG_MA0066.1.damo.pwm.bed
##   size
## 1  5143
## 2  9105
## 3 60933
## 4  6490
## 5 63954
## 6 16549
```

```
top10 = results[results$collection %in% head(unique(results$collection), n=10),]
str(top10)
```

```
## 'data.frame':   545 obs. of  22 variables:
## $ userSet      : int  1 1 1 1 1 1 1 1 1 1 ...
## $ dbSet        : int  3283 3290 3284 3291 3285 3297 348 337 336 310 ...
## $ collection   : chr  "PPARG" "PPARG" "PPARG" "PPARG" ...
## $ pValueLog    : num  5.62 3.69 3.45 3.05 2.9 ...
## $ oddsRatio    : num  13.6 3.65 2.11 3.26 1.98 ...
## $ support      : int  24 32 82 29 75 43 9 30 29 40 ...
## $ rnkPV        : int  1 2 3 4 5 6 7 8 9 10 ...
## $ rnkOR        : int  80 96 182 102 281 165 1 133 137 178 ...
## $ rnkSup       : int  1324 825 13 1013 28 429 2564 937 1013 502 ...
## $ maxRnk       : int  1324 825 182 1013 281 429 2564 937 1013 502 ...
## $ meanRnk      : int  468 308 66 373 105 200 857 359 386 230 ...
## $ b            : int  2 10 50 10 47 21 0 12 12 21 ...
## $ c            : int  180 172 122 175 129 161 195 174 175 164 ...
## $ d            : int  205 197 157 197 160 186 207 195 195 186 ...
## $ description: chr  "PPARG Simpson Golabi Behmel Syndrome adipocytes PPARG" "PPARG subcutaneous whi
## $ cellType     : chr  "Simpson Golabi Behmel Syndrome adipocytes" "subcutaneous white adipose tissue"
## $ tissue       : chr  "NULL" "NULL" "NULL" "NULL" ...
## $ antibody     : chr  "PPARG" "PPARG" "PPARG" "PPARG" ...
## $ treatment    : chr  "NULL" "" "differentiaiton inducators" "" ...
## $ dataSource   : chr  "GTRD" "GTRD" "GTRD" "GTRD" ...
## $ filename     : chr  "EXP000907_Simpson-Golabi-Behmel-Syndrome-adipocytes_PPARG_MACS_PPARG_MA0065.2.
## $ size         : int  5143 9105 60933 6490 63954 16549 8140 42883 42489 24225 ...
```

```
top10$collection = factor(top10$collection, levels=head(unique(results$collection), n=10))
top10$is_adipose = grepl("adipo", top10$cellType)
```

```
ggplot(top10) + geom_jitter(aes(x=collection, y=pValueLog, color=collection), show.legend = F, size=2) +
  theme_classic(base_size=18) + theme(axis.text.x = element_text(angle=45, hjust=1), axis.title.x = element_text()) +
  scale_color_brewer(palette = "Paired") + ylab("-log10 (p value)")
```



```
ggsave(file.path(LEAF, "Beige_unibind_results.pdf"), width=4, height=4)

ggplot(top10) + geom_jitter(aes(x=collection, y=pValueLog, color=is_adipose)) +
  theme_classic(base_size=18) + theme(axis.text.x = element_text(angle=45, hjust=1))
```

