

# ChIP-seq Analysis of TGF-beta Treated SMAD2 and SMAD3 Peaks

# Contents

|   |          |
|---|----------|
| <b>Methods and Code</b>   | <b>3</b> |
| Libraries . . . . .   | 3        |
| Annotating Peaks . . . . .  | 3        |
| GO Term Enrichment Analysis of Annotated Peaks . . . . .            | 6        |
| <b>Results</b>  | <b>7</b> |
| Peak Annotation Results . . . . .                                   | 7        |
| SMAD2 Peaks - Untreated . . . . .                                   | 7        |
| SMAD2 Peaks - Treated . . . . .                                     | 8        |
| SMAD3 Peaks - Untreated . . . . .                                   | 9        |
| SMAD3 Peaks - Treated . . . . .                                     | 10       |
| GO Term Analysis Results . . . . .                                  | 11       |
| GO Term Analysis for All Peaks . . . . .                            | 11       |
| GO Term Analysis for Peaks Restricted to Gene Regions . . . . .     | 17       |
| GO Term Analysis for Peaks Restricted to Promoter Regions . . . . . | 23       |

## Methods and Code

**Note:** the following code takes a long time to run. Hence, it shall just be displayed here for transparency but will not be run to compile the document.

Instead, the objects from the image saved in “annotation\_go\_rough\_code.RData” after running this code will be loaded to present and discuss the results.

### Libraries

The following libraries were loaded during this analysis:

1. tidyverse
2. formatR
3. GenomicRanges
4. org.Hs.eg.db
5. TxDb.Hsapiens.UCSC.hg38.knownGene
6. ChIPseeker
7. clusterProfiler
8. kableExtra

**Note:** Since the code takes a long time to run, the following file which contains the results of the analyses carried out by code will be loaded first, and the code will be displayed for transparency purposes, but will not be evaluated.

```
load("annotation_go_rough_code.RData")
```

### Annotating Peaks

Peaks were annotated using the **ChIPseeker** package. The annotated peaks for each of the conditions were further filtered into three subsets:

1. All peaks
2. Peaks with only promoter annotations.
3. Peaks lying within genic regions, i.e., peaks without the annotations “Distal Intergenic” and “Downstream”

The following code was used to read in the file names and create a character vector from them. Those names were then fed into a function to convert those files into Granges objects held in a list.

```
knitr::opts_chunk$set(echo = FALSE, tidy = TRUE)

## Takes a filename as string input, and if it has suffix '.narrowPeak',
## converts it to a GRanges object containing a q_value metadata column.
## Requires tidyverse and GRanges packages.

## checking for existence of file and correct extension
narrow_to_granges <- function(file) {
  stopifnot(file.exists(file), endsWith(x = file, suffix = ".narrowPeak"))

  bed <- as_tibble(read_tsv(file = file, col_names = c("chrom", "start", "end",
    "name", "score", "strand", "signal_value", "p_value", "q_value", "peak")))

  gr <- GRanges(seqnames = bed$chrom, ranges = IRanges(start = bed$start,
    end = bed$end), strand = if_else(condition = bed$strand == "+" | bed$strand ==
    "-", true = bed$strand, false = "*"), q_value = bed$q_value)

  return(gr)
}

### To read in multiple .narrowPeak files and create a GRanges list of GRanges
### objects with the input of respective files.

files <- dir(path = ".", pattern = "\\..narrowPeak$") ## specify files in character vector

grl <- lapply(files, narrow_to_granges) ## Create list of GRanges object
```

Here are the contents of the files vector.

```
files

## [1] "SMAD2_abInput_treated_peaks.narrowPeak"
## [2] "SMAD2_abInput_untreated_peaks.narrowPeak"
## [3] "SMAD3_LAP_treated_peaks.narrowPeak"
## [4] "SMAD3_LAP_untreated_peaks.narrowPeak"
```

The following code was then used to assign the Granges in the list to separate variables. The name of the variable to hold each GRanges object was derived from the respective file name from which its data was obtained. The names of the variables were stored as a character vector “v”. To names in v, the respective GRanges was assigned.

```
TxDb <- TxDb.Hsapiens.UCSC.hg38.knownGene

## vector of variable names for granges in list

v <- str_remove_all(string = files, pattern = "\\..narrowPeak$")

### Assign individual granges within grl to variable names in v.
```

```
for (i in seq_along(grl)) {
  assign(x = v[i], value = grl[[i]])
}
```

Here are the names of the variables held in vector v:

```
v
```

```
## [1] "SMAD2_abInput_treated_peaks" "SMAD2_abInput_untreated_peaks"
## [3] "SMAD3_LAP_treated_peaks"    "SMAD3_LAP_untreated_peaks"
```

A character vector of variable names was created to hold the annotated peak objects derived from the GRanges.

```
ap_v <- paste0("AP_", v) ## variable names for annotated peak objects

### Create annotated peak objects and assign them to names in ap_v

for (i in seq_along(grl)) {
  assign(x = ap_v[i], value = annotatePeak(peak = grl[[i]], TxDb = TxDb, annoDb = "org.Hs.eg.db"))
}

df_ap_v <- paste0("df_", ap_v) ## variable names for tibbles derived from annotated peaks

### Create tibbles from annotated peak objects listed in ap_v and assign them
### to names in df_ap_v

for (i in seq_along(ap_v)) {
  assign(x = df_ap_v[i], value = parse(text = ap_v[i]) %>% eval() %>% AnnotationDbi::as.data.frame() %>%
    as_tibble())
}

## variable names for tibbles derived from annotated peaks limited to
## promoter annotations

prom_df_ap_v <- paste0("prom_df_", ap_v)

### Create subset of tibbles from variable names listed in df_ap_v and assign
### them to variable names in prom_df_ap_v
for (i in seq_along(df_ap_v)) {
  assign(x = prom_df_ap_v[i], value = parse(text = df_ap_v[i]) %>% eval() %>%
    filter(str_detect(annotation, "Promoter")))
}

## variable names for tibbles derived from annotated peaks limited to gene
## region annotations

gen_df_ap_v <- paste0("gen_df_", ap_v)
```

```

### Filter tibbles with variable names listed in df_ap_v to discard rows with
### 'Distal Intergenic' or 'Downstream' annotations and assign them to names
### in gen_df_ap_v.

for (i in seq_along(df_ap_v)) {
  assign(x = gen_df_ap_v[i], value = parse(text = df_ap_v[i]) %>% eval() %>%
    filter(!str_detect(annotation, "Distal Intergenic|Downstream")))
}

```

## GO Term Enrichment Analysis of Annotated Peaks

GO term enrichment analysis of the annotated peaks was carried out using the **ClusterProfiler** package. Analysis was carried out for all four conditions and each subset of annotated peaks within each condition (all peaks, promoter-limited peaks and genic region-limited peaks).

**Note:** The following code takes *very* long to run.

```

ego_v <- paste0("ego_", v) ### Vector of variable names for go term enrichment objects

for (i in seq_along(ego_v)) {
  ego <- parse(text = df_ap_v[i]) %>% eval(expr = .) %>% dplyr::select(geneId) %>%
    unlist() %>% enrichGO(gene = ., OrgDb = org.Hs.eg.db, ont = "BP", pAdjustMethod = "BH",
      readable = TRUE)
  if (!nrow(as.data.frame(ego)) == 0)
    assign(x = ego_v[i], value = ego)
}

### Vector of variable names for go term enrichment objects limited to
### promoter peaks

prom_ego_v <- ego_v <- paste0("prom_ego_", v)

### Performing go term analysis using the clusterprofiler package and
### assigning the results to the variables in gen_ego_v

for (i in seq_along(prom_ego_v)) {
  ego <- parse(text = prom_df_ap_v[i]) %>% eval(expr = .) %>% dplyr::select(geneId) %>%
    unlist() %>% enrichGO(gene = ., OrgDb = org.Hs.eg.db, ont = "BP", pAdjustMethod = "BH",
      readable = TRUE)
  if (!nrow(as.data.frame(ego)) == 0)
    assign(x = prom_ego_v[i], value = ego)
}

### Vector of variable names for go term enrichment objects limited to genic
### region peaks

gen_ego_v <- paste0("gen_ego_", v)

### Performing go term analysis using the clusterprofiler package and

```

```
### assigning the results to the variables in gen_ego_v

for (i in seq_along(gen_ego_v)) {
  ego <- parse(text = gen_df_ap_v[i]) %>% eval(expr = .) %>% dplyr::select(geneId) %>%
    unlist() %>% enrichGO(gene = ., OrgDb = org.Hs.eg.db, ont = "BP", pAdjustMethod = "BH",
      readable = TRUE)
  if (!nrow(as.data.frame(ego)) == 0)
    assign(x = gen_ego_v[i], value = ego)
}
```

Results of analysis were saved in the following file:

```
save.image(file = "annotation_go_rough_code.RData")
```

## Results

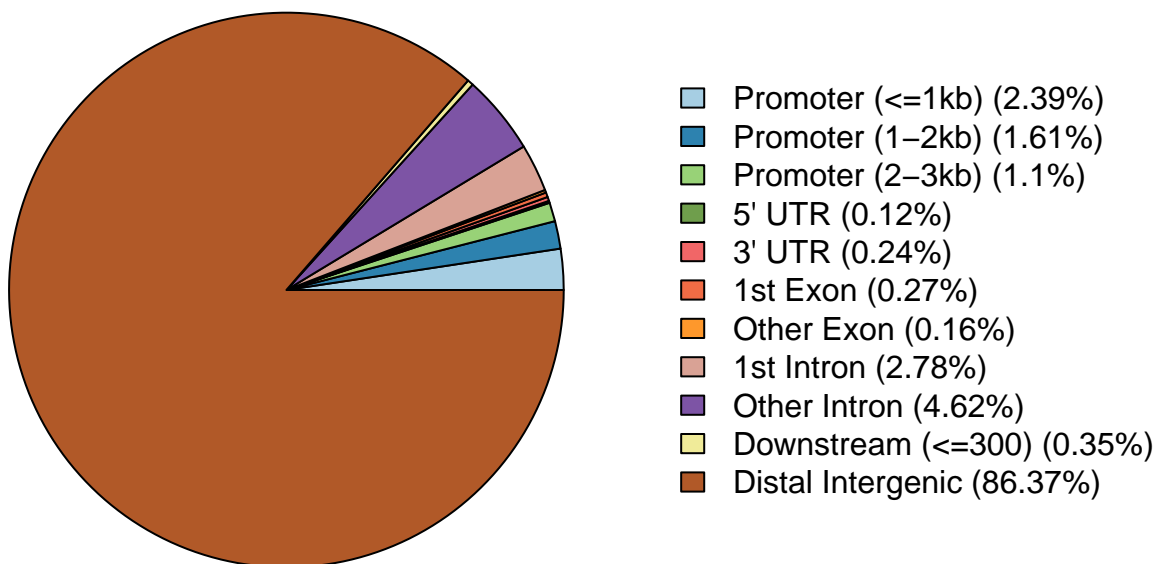
### Peak Annotation Results

#### SMAD2 Peaks - Untreated

Most of the untreated SMAD2 peaks were in the distal intergenic regions (about 86%).

```
plotAnnoPie(AP_SMAD2_abInput_untreated_peaks)
title(main = "SMAD2 Peaks: Untreated", line = -2, adj = 0)
```

#### SMAD2 Peaks: Untreated

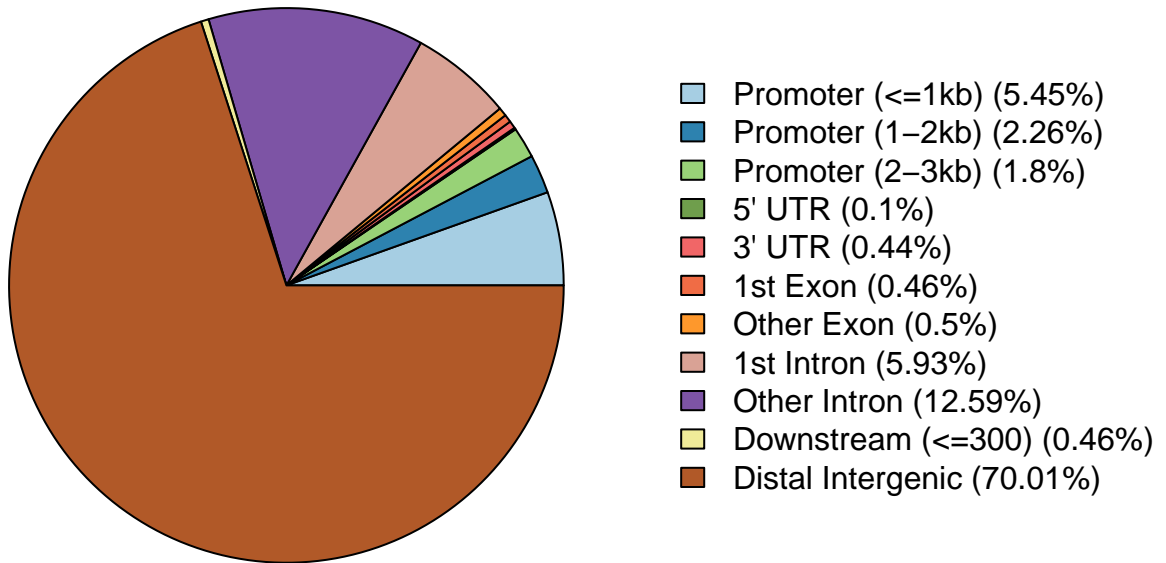


## SMAD2 Peaks - Treated

Similar to the untreated SMAD2 peaks, most of the treated SMAD2 peaks also lay within the intergenic regions ( about 71%), but in a lower proportion as compared to the untreated peaks (about 86%). On the other hand, the proportion of SMAD2 peaks in the genic regions rose upon treatment from 13.28% to 29.53%.

```
plotAnnoPie(AP_SMAD2_abInput_treated_peaks)
title(main = "SMAD2 Peaks: Treated", line = -2, adj = 0)
```

## SMAD2 Peaks: Treated



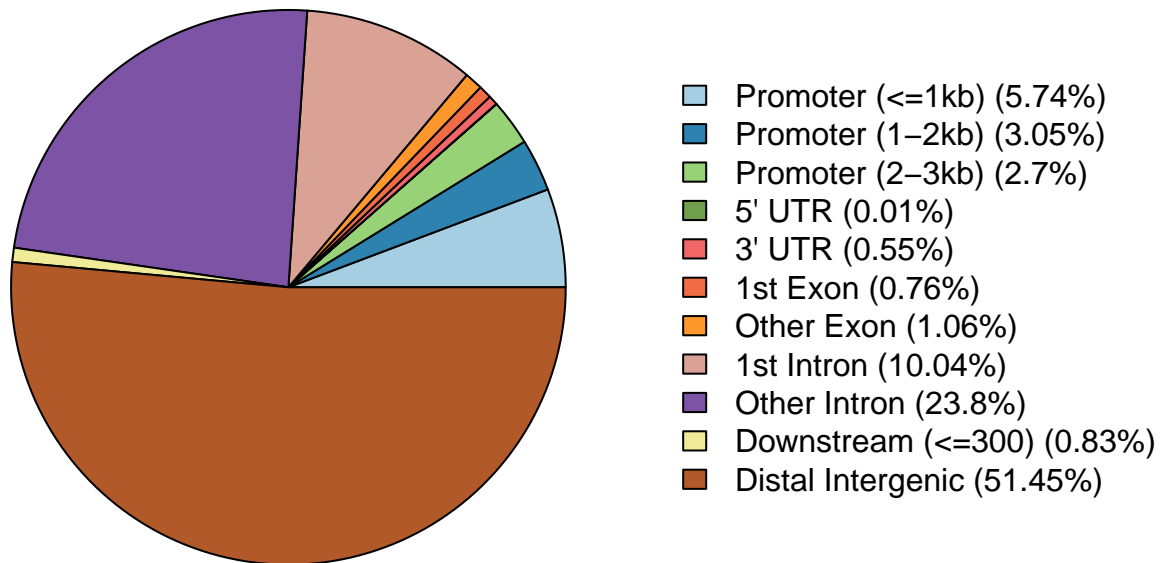


### SMAD3 Peaks - Untreated

Around half of the untreated SMAD3 peaks were in the intergenic regions, while a majority of the peaks in the genic regions (33.84 %) overlapped with intronic sequences. Compared to the SMAD2 peaks (both treated and untreated), the SMAD3 untreated peaks had a far lower proportion of peaks within the intergenic regions. On the other hand, the proportion of SMAD3 peaks in intronic regions was noticeably elevated as compared to the SMAD2 peaks.

```
plotAnnoPie(AP_SMAD3_LAP_untreated_peaks)
title(main = "SMAD3 Peaks: Untreated", line = -2, adj = 0)
```

### SMAD3 Peaks: Untreated

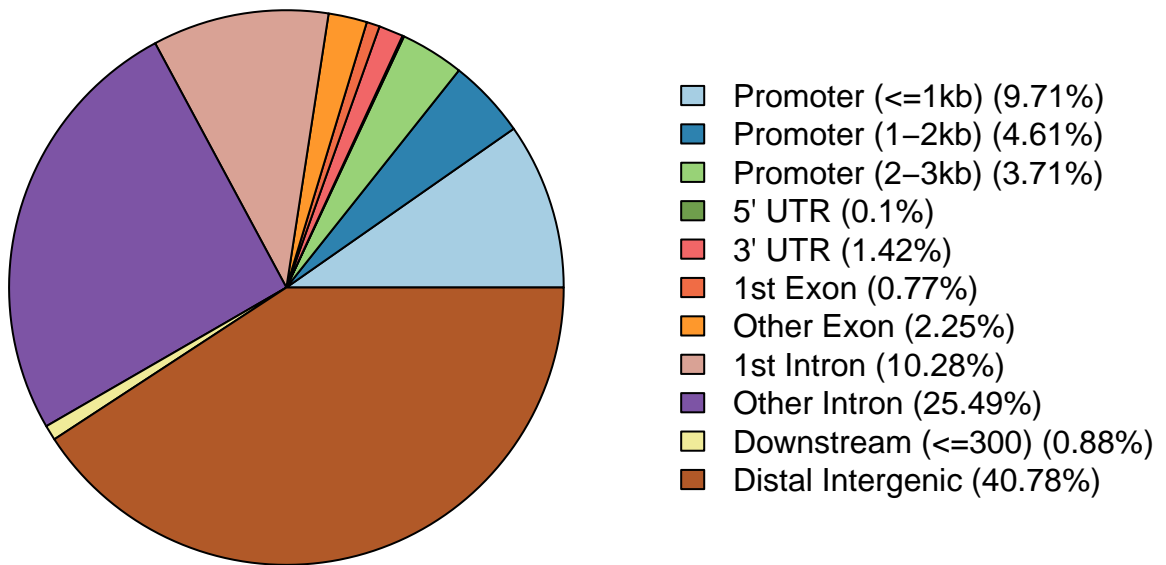


## SMAD3 Peaks - Treated

TGF- $\beta$  ligand treatment reduced the proportion of SMAD3 peaks in the intergenic regions (52.28% without treatment vs. 41.66% with treatment). For the most part, those peaks seem to have been displaced to the promoter regions, which accounted for a total of 11.49% before treatment but rose to 18.03% after treatment.

```
plotAnnoPie(AP_SMAD3_LAP_treated_peaks)
title(main = "SMAD3 Peaks: Treated", line = -2, adj = 0)
```

## SMAD3 Peaks: Treated



## GO Term Analysis Results

GO term analysis was limited to biological process (BP) terms; GO Terms relating to Cellular Component and Molecular Function were not analysed.

### GO Term Analysis for All Peaks

No enriched BP GO terms were found within annotations for untreated SMAD2 peaks. This may have been due to the majority of these peaks being located in the intergenic regions. For treated SMAD2 peaks, GO terms related to processes in neuronal development were over-represented.

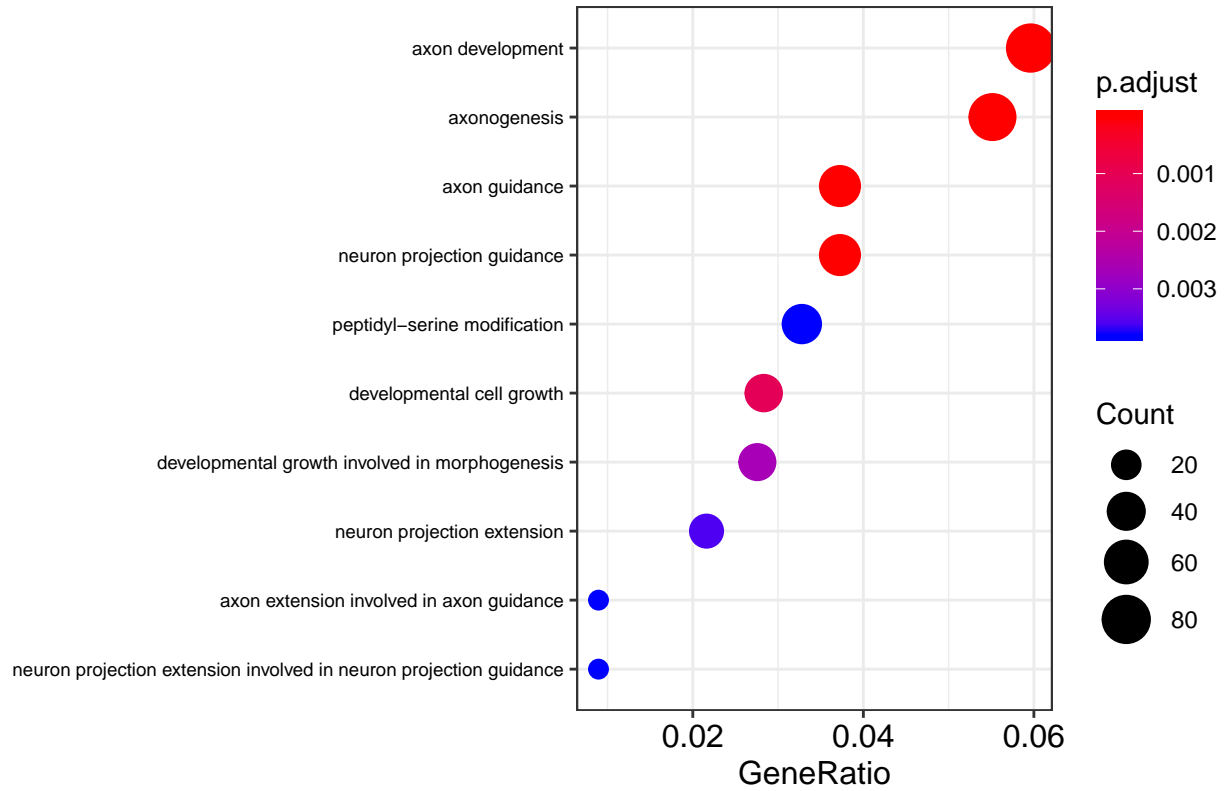
**Note:** Only the first 20 enriched GO terms for each peak subset and condition are shown. Terms were arranged in ascending order by adjusted p value, then in descending order by gene ratio and absolute count.

```
ego_SMAD2_abInput_treated_peaks@result %>%  
  as_tibble() %>%  
  arrange(p.adjust, desc(Count), desc(GeneRatio)) %>% filter(p.adjust < 0.05) %>%  
  dplyr::select(Description, p.adjust, GeneRatio, BgRatio) %>%  
  head(20) %>%  
  rename(Description = "Enriched GO Terms in Treated SMAD2 Peaks") %>%  
  kable(x = ., "latex", longtable = T) %>% kable_styling(latex_options = c("repeat_header"))
```

| Enriched GO Terms in Treated SMAD2 Peaks                           | p.adjust  | GeneRatio | BgRatio   |
|--|-----------|-----------|-----------|
| axon development   | 0.0000000 | 80/1342   | 493/18493 |
| axonogenesis   | 0.0000001 | 74/1342   | 449/18493 |
| axon guidance  | 0.0000002 | 50/1342   | 259/18493 |
| neuron projection guidance   | 0.0000002 | 50/1342   | 260/18493 |
| developmental cell growth  | 0.0010681 | 38/1342   | 226/18493 |
| developmental growth involved in morphogenesis                     | 0.0026364 | 37/1342   | 227/18493 |
| neuron projection extension  | 0.0035945 | 29/1342   | 161/18493 |
| peptidyl-serine modification                                       | 0.0037980 | 44/1342   | 301/18493 |
| axon extension involved in axon guidance                           | 0.0037980 | 12/1342   | 37/18493  |
| neuron projection extension involved in neuron projection guidance | 0.0037980 | 12/1342   | 37/18493  |
| regulation of developmental growth                                 | 0.0043011 | 47/1342   | 334/18493 |
| cell junction assembly   | 0.0043011 | 37/1342   | 238/18493 |
| regulation of axonogenesis   | 0.0066075 | 30/1342   | 180/18493 |
| axon extension   | 0.0108430 | 22/1342   | 116/18493 |
| cell growth  | 0.0121295 | 60/1342   | 485/18493 |
| peptidyl-serine phosphorylation                                    | 0.0121295 | 40/1342   | 282/18493 |
| regulation of cell morphogenesis involved in differentiation       | 0.0125971 | 41/1342   | 293/18493 |
| cell junction organization   | 0.0161102 | 40/1342   | 287/18493 |
| smooth muscle contraction  | 0.0165590 | 20/1342   | 105/18493 |
| mesenchymal cell development                                       | 0.0165590 | 17/1342   | 82/18493  |

```
dotplot(ego_SMAD2_abInput_treated_peaks) + ggtitle(label = "Enriched BP GO Terms in Treated SMAD2 Peaks")
  theme(plot.title = element_text(hjust = 0.8, size = 15), axis.text.y = element_text(size = 7))
```

## Enriched BP GO Terms in Treated SMAD2 Peaks

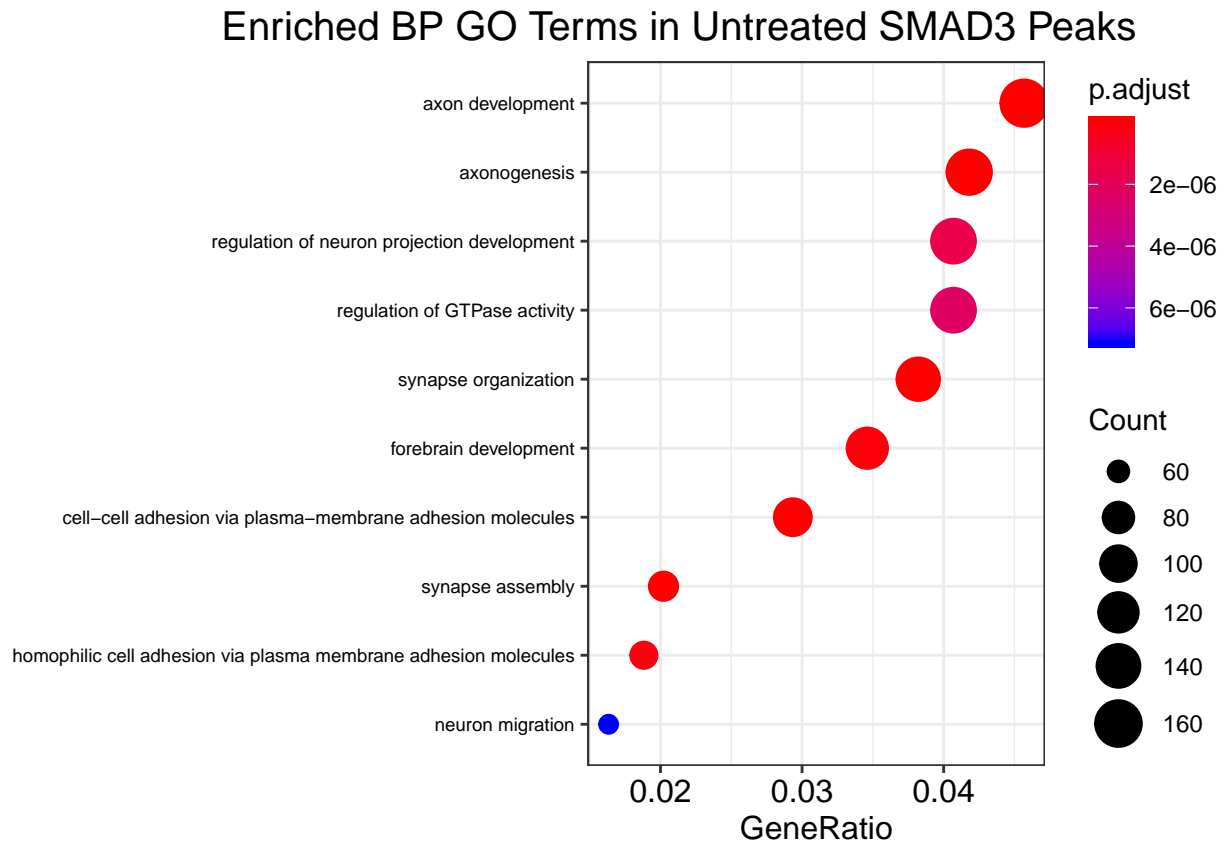


Untreated SMAD3 peaks also showed an enrichment in neuronal development-related terms. In addition, peaks in regions related to GTPase activity and cell-cell adhesion were also enriched.

```
ego_SMAD3_LAP_untreated_peaks@result %>%
  as_tibble() %>%
  arrange(p.adjust, desc(Count) , desc(GeneRatio)) %>%
  filter(p.adjust < 0.05) %>%
  dplyr::select(Description, p.adjust, GeneRatio, BgRatio) %>%
  head(20) %>%
  rename(Description = "Enriched GO Terms in Untreated SMAD3 Peaks") %>%
kable( x = ., "latex", longtable = T) %>% kable_styling(latex_options = c("repeat_header"))
```

| Enriched GO Terms in Untreated SMAD3 Peaks                      | p.adjust | GeneRatio | BgRatio   |
|---|----------|-----------|-----------|
| synapse organization  | 0.00e+00 | 138/3612  | 384/18493 |
| cell-cell adhesion via plasma-membrane adhesion molecules       | 0.00e+00 | 106/3612  | 270/18493 |
| axon development  | 0.00e+00 | 165/3612  | 493/18493 |
| synapse assembly  | 0.00e+00 | 73/3612   | 165/18493 |
| axonogenesis  | 0.00e+00 | 151/3612  | 449/18493 |
| forebrain development   | 1.00e-07 | 125/3612  | 374/18493 |
| homophilic cell adhesion via plasma membrane adhesion molecules | 2.00e-07 | 68/3612   | 167/18493 |
| regulation of neuron projection development                     | 1.60e-06 | 147/3612  | 478/18493 |
| regulation of GTPase activity                                   | 2.30e-06 | 147/3612  | 481/18493 |
| neuron migration  | 7.10e-06 | 59/3612   | 149/18493 |
| positive regulation of GTPase activity                          | 1.50e-05 | 125/3612  | 405/18493 |
| modulation of chemical synaptic transmission                    | 1.56e-05 | 128/3612  | 418/18493 |
| regulation of trans-synaptic signaling                          | 1.69e-05 | 128/3612  | 419/18493 |
| neuron projection guidance                                      | 2.69e-05 | 87/3612   | 260/18493 |
| regulation of synapse organization                              | 2.69e-05 | 75/3612   | 214/18493 |
| postsynapse organization  | 2.69e-05 | 60/3612   | 159/18493 |
| positive regulation of neurogenesis                             | 2.84e-05 | 135/3612  | 453/18493 |
| regulation of synapse structure or activity                     | 3.65e-05 | 76/3612   | 220/18493 |
| axon guidance   | 3.95e-05 | 86/3612   | 259/18493 |
| regulation of small GTPase mediated signal transduction         | 6.12e-05 | 101/3612  | 321/18493 |

```
dotplot(ego_SMAD3_LAP_untreated_peaks) + ggtitle(label = "Enriched BP GO Terms in Untreated SMAD3 Peaks")
```

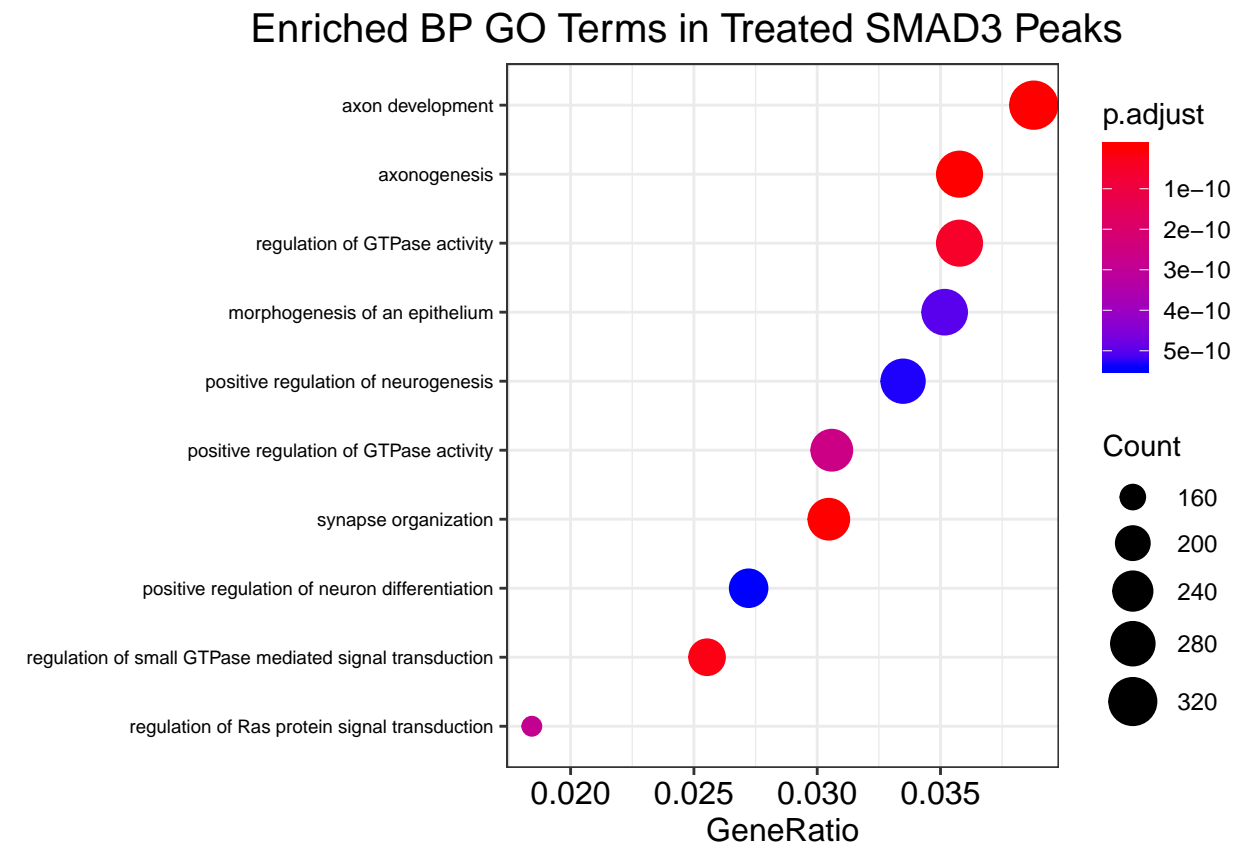


Similar terms related to neuronal development also appeared in treated SMAD3 peaks. However, additional enriched terms such as “regulation of small GTPase-mediated signal transduction”, as well as “regulation of Ras protein signal transduction” were also found to be enriched.

```
ego_SMAD3_LAP_treated_peaks@result %>%
  as_tibble() %>%
  arrange(p.adjust, desc(Count) , desc(GeneRatio)) %>%
  filter(p.adjust < 0.05) %>%
  dplyr::select(Description, p.adjust, GeneRatio, BgRatio) %>%
  head(20) %>%
  rename(Description = "Enriched GO Terms in Treated SMAD3 Peaks") %>%
  kable(x = ., "latex", longtable = T) %>%
  kable_styling(latex_options = c("repeat_header"))
```

| Enriched GO Terms in Treated SMAD3 Peaks                | p.adjust | GeneRatio | BgRatio   |
|---|----------|-----------|-----------|
| axon development  | 0        | 322/8303  | 493/18493 |
| axonogenesis  | 0        | 297/8303  | 449/18493 |
| synapse organization                                    | 0        | 253/8303  | 384/18493 |
| regulation of small GTPase mediated signal transduction | 0        | 212/8303  | 321/18493 |
| regulation of GTPase activity                           | 0        | 297/8303  | 481/18493 |
| positive regulation of GTPase activity                  | 0        | 254/8303  | 405/18493 |
| regulation of Ras protein signal transduction           | 0        | 153/8303  | 222/18493 |
| morphogenesis of an epithelium                          | 0        | 292/8303  | 479/18493 |
| positive regulation of neurogenesis                     | 0        | 278/8303  | 453/18493 |
| positive regulation of neuron differentiation           | 0        | 226/8303  | 356/18493 |
| synapse assembly  | 0        | 119/8303  | 165/18493 |
| neuron projection guidance                              | 0        | 172/8303  | 260/18493 |
| epithelial tube morphogenesis                           | 0        | 201/8303  | 313/18493 |
| axon guidance   | 0        | 171/8303  | 259/18493 |
| developmental growth involved in morphogenesis          | 0        | 153/8303  | 227/18493 |
| regulation of trans-synaptic signaling                  | 0        | 257/8303  | 419/18493 |
| modulation of chemical synaptic transmission            | 0        | 256/8303  | 418/18493 |
| regulation of neuron projection development             | 0        | 284/8303  | 478/18493 |
| Ras protein signal transduction                         | 0        | 259/8303  | 430/18493 |
| positive regulation of cell projection organization     | 0        | 227/8303  | 371/18493 |

```
dotplot(ego_SMAD3_LAP_treated_peaks) + ggtitle(label = "Enriched BP GO Terms in Treated SMAD3 Peaks") +
```





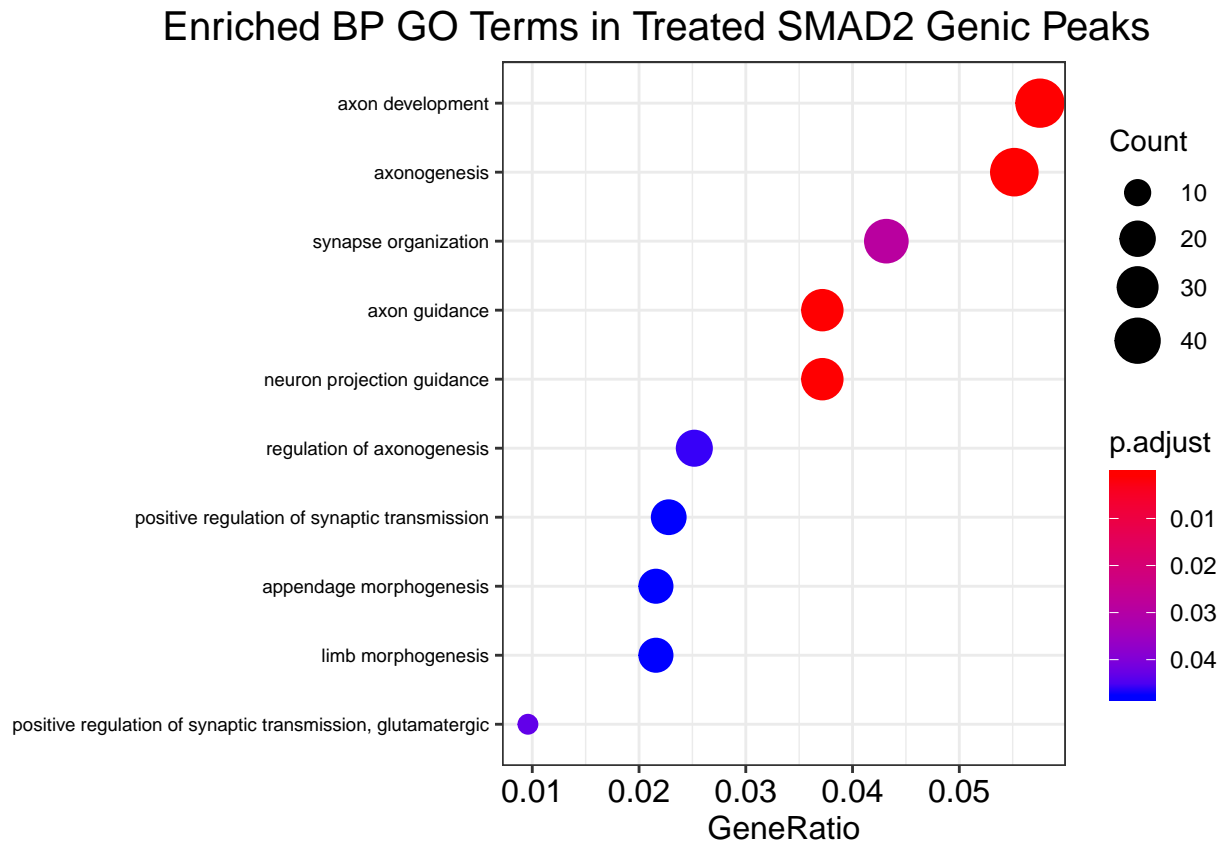
## GO Term Analysis for Peaks Restricted to Gene Regions

The enriched GO Terms for SMAD2 peaks limited to gene regions did not differ much from their counterparts where all the peaks were considered. The pattern remains the same, with terms associated with neuronal development being predominantly featured, although terms for limb and appendage morphogenesis also turn up.

```
gen_ego_SMAD2_abInput_treated_peaks@result %>%
  as_tibble() %>%
  arrange(p.adjust, desc(Count) , desc(GeneRatio)) %>%
  filter(p.adjust < 0.05) %>%
  dplyr::select(Description, p.adjust, GeneRatio, BgRatio) %>%
  head(20) %>%
  rename(Description = "Enriched GO Terms in Treated SMAD2 Peaks Restricted to Gene Regions") %>%
  kable(x = ., "latex", longtable = T) %>%
  kable_styling(latex_options = c("repeat_header"))
```

| Enriched GO Terms in Treated SMAD2 Peaks Restricted to Gene Regions | p.adjust  | GeneRatio | BgRatio   |
|---|-----------|-----------|-----------|
| axonogenesis  | 0.0009522 | 46/834    | 449/18493 |
| axon development  | 0.0010088 | 48/834    | 493/18493 |
| axon guidance   | 0.0010088 | 31/834    | 259/18493 |
| neuron projection guidance  | 0.0010088 | 31/834    | 260/18493 |
| synapse organization  | 0.0287847 | 36/834    | 384/18493 |
| positive regulation of synaptic transmission, glutamatergic         | 0.0433635 | 8/834     | 31/18493  |
| regulation of axonogenesis  | 0.0461952 | 21/834    | 180/18493 |
| peptidyl-serine modification  | 0.0474595 | 29/834    | 301/18493 |
| positive regulation of synaptic transmission                        | 0.0474595 | 19/834    | 159/18493 |
| appendage morphogenesis   | 0.0474595 | 18/834    | 146/18493 |
| limb morphogenesis  | 0.0474595 | 18/834    | 146/18493 |

```
dotplot(gen_ego_SMAD2_abInput_treated_peaks) + ggtitle(label = "Enriched BP GO Terms in Treated SMAD2 G
```

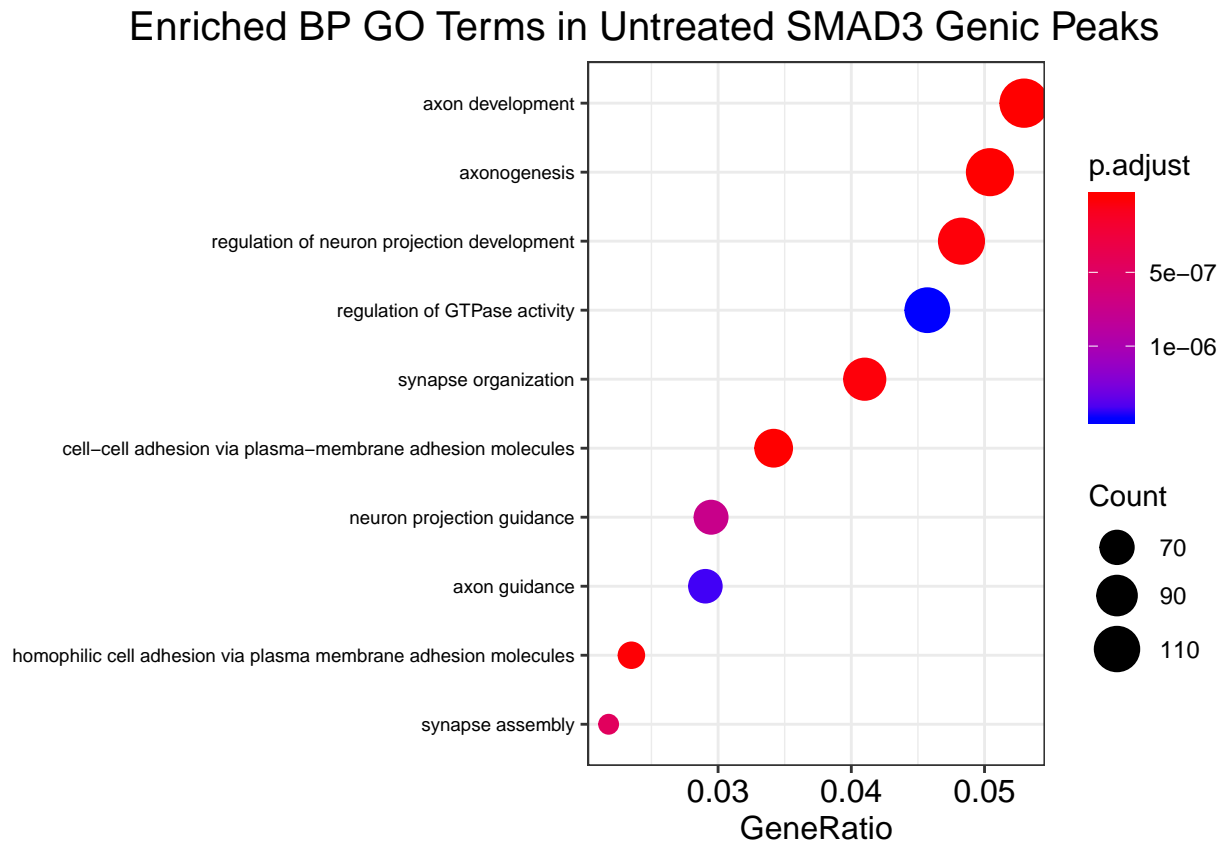


Untreated SMAD3 peaks in gene regions showed a similar neuronal development theme for the most significantly enriched genes.

```
gen_ego_SMAD3_LAP_untreated_peaks@result %>%
  as_tibble() %>% arrange(p.adjust, desc(Count) , desc(GeneRatio)) %>%
  filter(p.adjust < 0.05) %>%
  dplyr::select(Description, p.adjust, GeneRatio, BgRatio) %>%
  head(20) %>%
  rename(Description = "Enriched GO Terms in Untreated SMAD3 Peaks Restricted to Gene Regions") %>%
  kable(x = ., "latex", longtable = T) %>%
  kable_styling(latex_options = c("repeat_header"))
```

| Enriched GO Terms in Untreated SMAD3 Peaks Restricted to Gene Regions | p.adjust | GeneRatio | BgRatio   |
|---|----------|-----------|-----------|
| axonogenesis  | 0.00e+00 | 118/2341  | 449/18493 |
| axon development  | 0.00e+00 | 124/2341  | 493/18493 |
| cell-cell adhesion via plasma-membrane adhesion molecules             | 0.00e+00 | 80/2341   | 270/18493 |
| homophilic cell adhesion via plasma membrane adhesion molecules       | 0.00e+00 | 55/2341   | 167/18493 |
| regulation of neuron projection development                           | 0.00e+00 | 113/2341  | 478/18493 |
| synapse organization  | 0.00e+00 | 96/2341   | 384/18493 |
| synapse assembly  | 5.00e-07 | 51/2341   | 165/18493 |
| neuron projection guidance  | 7.00e-07 | 69/2341   | 260/18493 |
| axon guidance   | 1.40e-06 | 68/2341   | 259/18493 |
| regulation of GTPase activity   | 1.50e-06 | 107/2341  | 481/18493 |
| modulation of chemical synaptic transmission                          | 1.60e-06 | 96/2341   | 418/18493 |
| regulation of trans-synaptic signaling                                | 1.70e-06 | 96/2341   | 419/18493 |
| neuron migration  | 1.80e-06 | 46/2341   | 149/18493 |
| positive regulation of GTPase activity                                | 4.80e-06 | 92/2341   | 405/18493 |
| positive regulation of neurogenesis                                   | 5.00e-06 | 100/2341  | 453/18493 |
| forebrain development   | 1.39e-05 | 85/2341   | 374/18493 |
| sensory system development  | 1.39e-05 | 82/2341   | 357/18493 |
| eye development   | 1.39e-05 | 81/2341   | 351/18493 |
| visual system development   | 1.46e-05 | 81/2341   | 352/18493 |
| positive regulation of neuron differentiation                         | 2.32e-05 | 81/2341   | 356/18493 |

```
dotplot(gen_ego_SMAD3_LAP_untreated_peaks ) + ggtitle(label = "Enriched BP GO Terms in Untreated SMAD3 (
```



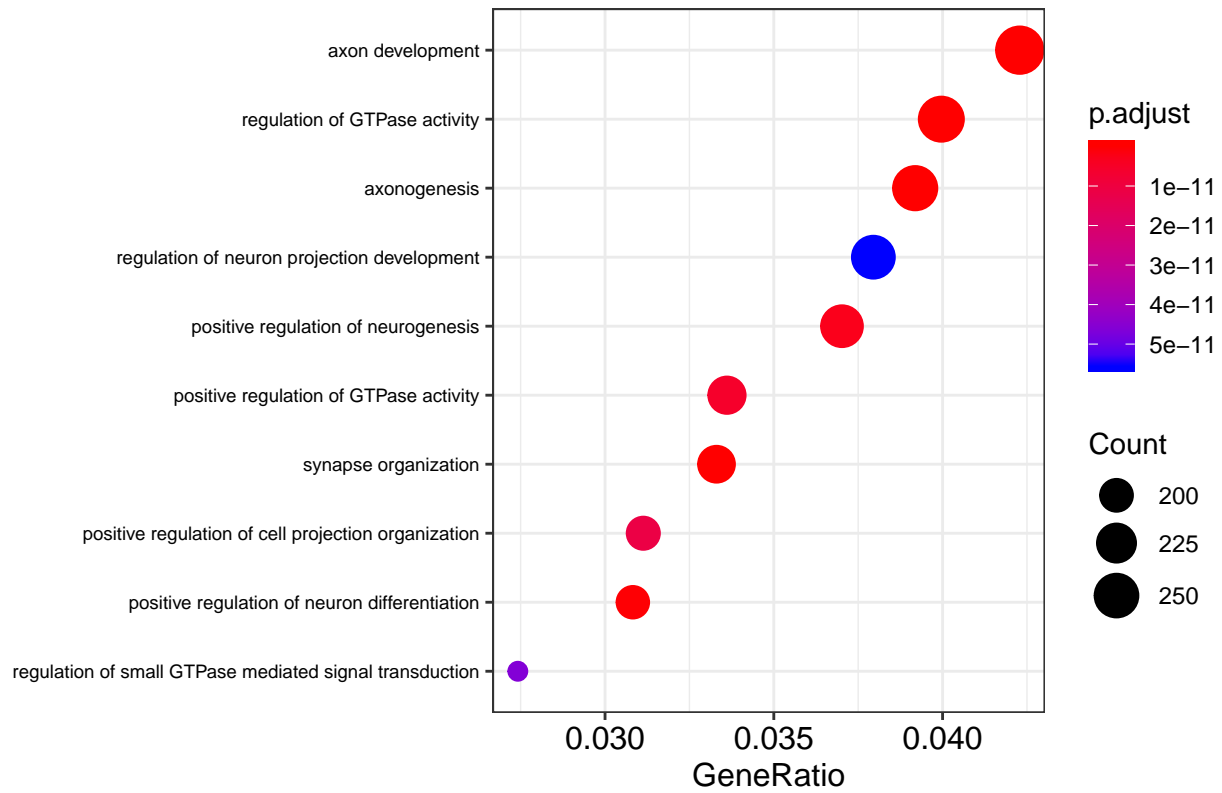
Treated SMAD3 peaks in gene regions did not show too many changes either, as compared to the the set of all treated SMAD3 peaks.

```
kable(gen_ego_SMAD3_LAP_treated_peaks@result %>% as_tibble() %>% arrange(p.adjust, desc(Count) , desc(G
```

| Enriched GO Terms in Treated SMAD3 Peaks Restricted to Gene Regions | p.adjust | GeneRatio | BgRatio   |
|---|----------|-----------|-----------|
| axon development  | 0        | 273/6456  | 493/18493 |
| axonogenesis  | 0        | 253/6456  | 449/18493 |
| regulation of GTPase activity                                       | 0        | 258/6456  | 481/18493 |
| synapse organization  | 0        | 215/6456  | 384/18493 |
| positive regulation of neuron differentiation                       | 0        | 199/6456  | 356/18493 |
| positive regulation of neurogenesis                                 | 0        | 239/6456  | 453/18493 |
| positive regulation of GTPase activity                              | 0        | 217/6456  | 405/18493 |
| positive regulation of cell projection organization                 | 0        | 201/6456  | 371/18493 |
| regulation of small GTPase mediated signal transduction             | 0        | 177/6456  | 321/18493 |
| regulation of neuron projection development                         | 0        | 245/6456  | 478/18493 |
| regulation of Ras protein signal transduction                       | 0        | 130/6456  | 222/18493 |
| regulation of axonogenesis  | 0        | 110/6456  | 180/18493 |
| neuron projection guidance  | 0        | 147/6456  | 260/18493 |
| axon guidance   | 0        | 146/6456  | 259/18493 |
| modulation of chemical synaptic transmission                        | 0        | 213/6456  | 418/18493 |
| regulation of trans-synaptic signaling                              | 0        | 213/6456  | 419/18493 |
| positive regulation of neuron projection development                | 0        | 147/6456  | 269/18493 |
| developmental growth involved in morphogenesis                      | 0        | 128/6456  | 227/18493 |
| regulation of cell morphogenesis involved in differentiation        | 0        | 157/6456  | 293/18493 |
| synapse assembly  | 0        | 99/6456   | 165/18493 |

```
dotplot(gen_ego_SMAD3_LAP_treated_peaks )+ ggtitle(label = "Enriched BP GO Terms in Treated SMAD3 Genic
```

## Enriched BP GO Terms in Treated SMAD3 Genic Peaks



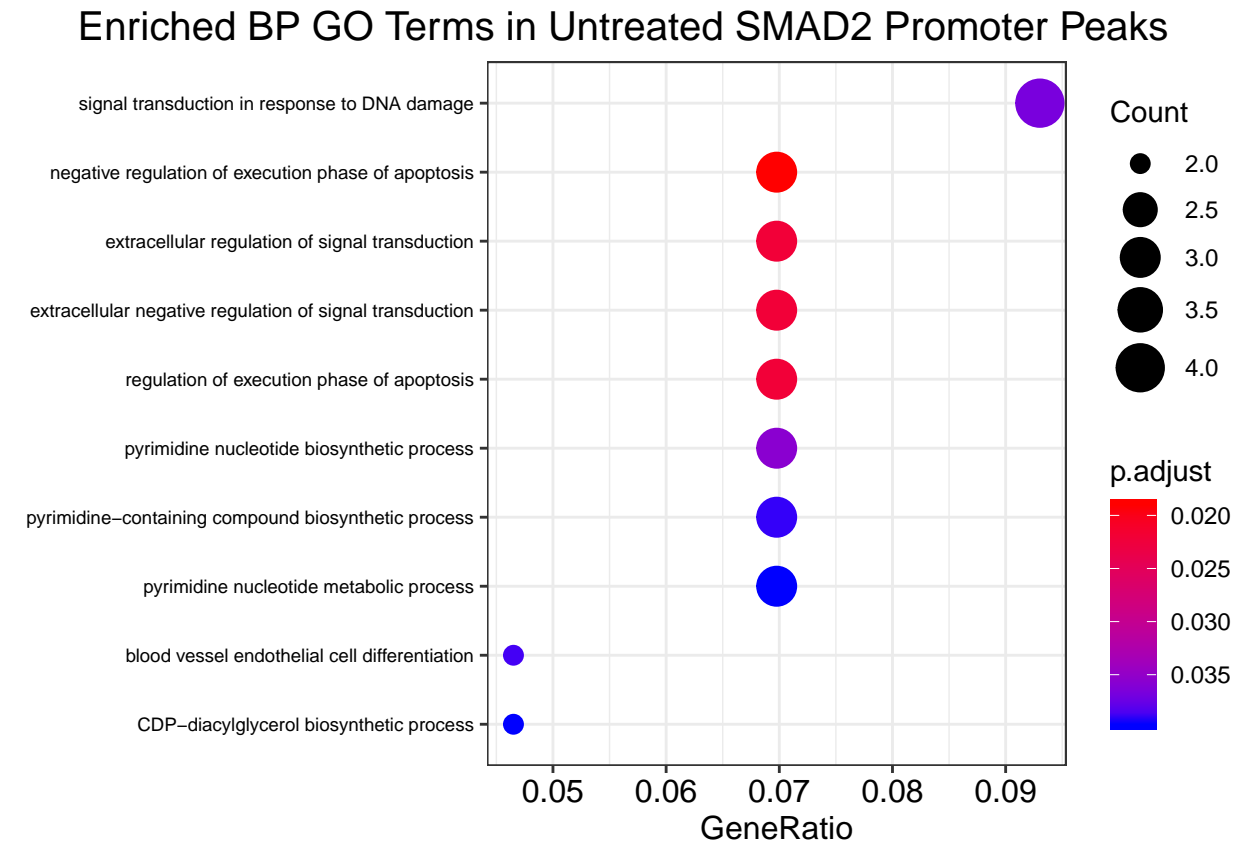
## GO Term Analysis for Peaks Restricted to Promoter Regions

When peaks were limited to promoter regions, the enriched GO term sets showed noticeable differences from sets derived from all peaks as well as sets derived from peaks limited to gene regions. No enriched GO terms were found for treated promoter-proximal SMAD2 peaks.

```
prom_ego_SMAD2_abInput_untreated_peaks@result %>%
  as_tibble() %>%
  arrange(p.adjust, desc(Count) , desc(GeneRatio)) %>%
  filter(p.adjust < 0.05) %>%
  dplyr::select(Description, p.adjust, GeneRatio, BgRatio) %>%
  head(20) %>%
  rename(Description = "Enriched GO Terms in Untreated SMAD2 Peaks Restricted to Promoters") %>%
  kable(x = ., "latex", longtable = T) %>%
  kable_styling(latex_options = c("repeat_header"))
```

| Enriched GO Terms in Untreated SMAD2 Peaks Restricted to Promoters | p.adjust  | GeneRatio | BgRatio   |
|--|-----------|-----------|-----------|
| negative regulation of execution phase of apoptosis                | 0.0190118 | 3/43      | 23/18493  |
| extracellular regulation of signal transduction                    | 0.0220964 | 3/43      | 36/18493  |
| extracellular negative regulation of signal transduction           | 0.0220964 | 3/43      | 36/18493  |
| regulation of execution phase of apoptosis                         | 0.0220964 | 3/43      | 38/18493  |
| pyrimidine nucleotide biosynthetic process                         | 0.0356605 | 3/43      | 48/18493  |
| signal transduction in response to DNA damage                      | 0.0367235 | 4/43      | 130/18493 |
| blood vessel endothelial cell differentiation                      | 0.0387759 | 2/43      | 11/18493  |
| pyrimidine-containing compound biosynthetic process                | 0.0391230 | 3/43      | 58/18493  |
| pyrimidine nucleotide metabolic process                            | 0.0396258 | 3/43      | 64/18493  |
| CDP-diacylglycerol biosynthetic process                            | 0.0396258 | 2/43      | 13/18493  |
| CDP-diacylglycerol metabolic process                               | 0.0396258 | 2/43      | 14/18493  |
| morphogenesis of an endothelium                                    | 0.0396258 | 2/43      | 15/18493  |
| endothelial tube morphogenesis                                     | 0.0396258 | 2/43      | 15/18493  |

```
dotplot(prom_ego_SMAD2_abInput_untreated_peaks) + ggtitle(label = "Enriched BP GO Terms in Untreated SMAD2 Promoter Peaks")
```





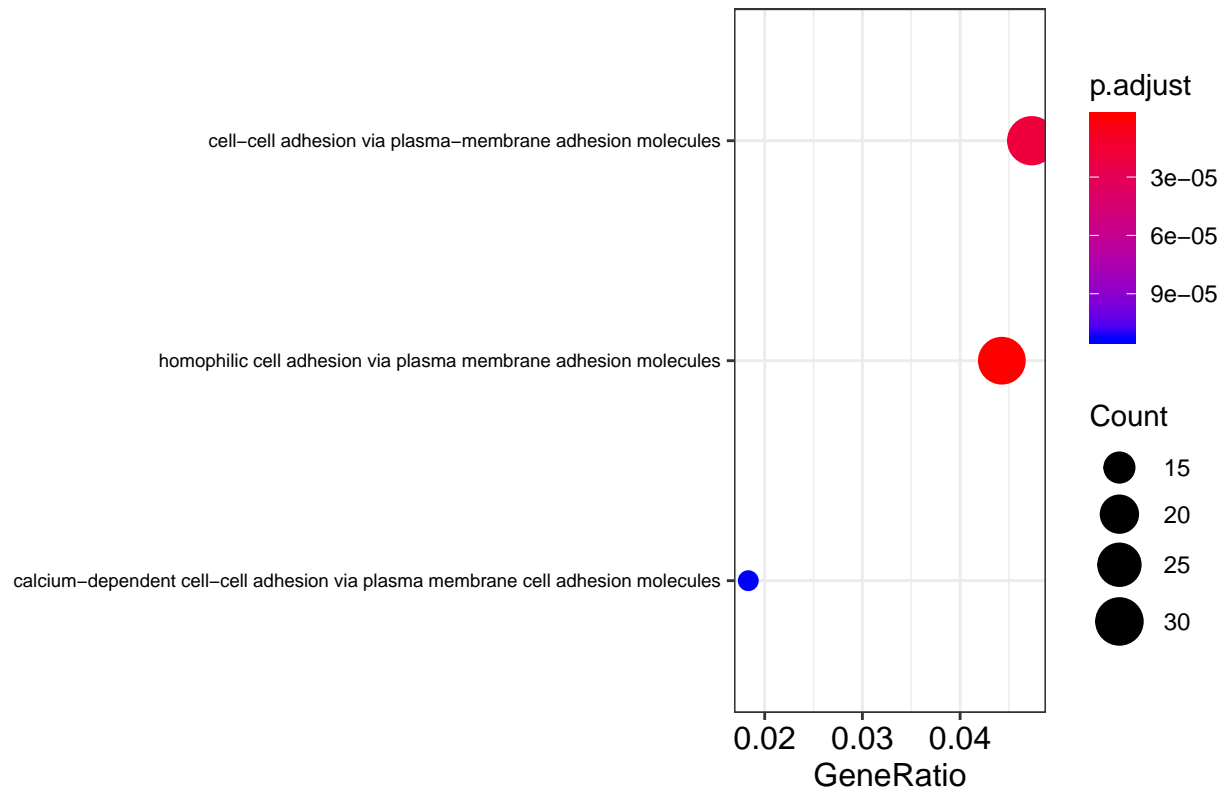
Untreated SMAD3 peaks restricted to promoters showed the least number of enriched terms, all of which were related to cell-cell adhesion.

```
prom_ego_SMAD3_LAP_untreated_peaks@result %>%
  as_tibble() %>%
  arrange(p.adjust, desc(Count) , desc(GeneRatio)) %>%
  filter(p.adjust < 0.05) %>%
  dplyr::select(Description, p.adjust, GeneRatio, BgRatio) %>%
  head(20) %>%
  rename(Description = "Enriched GO Terms in Untreated SMAD3 Peaks Restricted to Promoters") %>%
  kable(x = . , "latex", longtable = T) %>% kable_styling(latex_options = c("repeat_header"))
```

| Enriched GO Terms in Untreated SMAD3 Peaks Restricted to Promoters               | p.adjust  | GeneRatio | BgRatio   |
|--|-----------|-----------|-----------|
| homophilic cell adhesion via plasma membrane adhesion molecules                  | 0.0000000 | 29/655    | 167/18493 |
| cell-cell adhesion via plasma-membrane adhesion molecules                        | 0.0000194 | 31/655    | 270/18493 |
| calcium-dependent cell-cell adhesion via plasma membrane cell adhesion molecules | 0.0001125 | 12/655    | 48/18493  |

```
dotplot(prom_ego_SMAD3_LAP_untreated_peaks) + ggtitle(label = "Enriched BP GO Terms in Untreated SMAD3 L
```

## Enriched BP GO Terms in Untreated SMAD3 Promoter Peaks



Treatment with ligand seemed to show many more enriched GO terms in promoter-restricted SMAD3 peaks, as compared to untreated promoter-restricted SMAD3 peaks.

```
prom_ego_SMAD3_LAP_treated_peaks@result %>%
  as_tibble() %>%
  arrange(p.adjust, desc(Count) , desc(GeneRatio)) %>%
  filter(p.adjust < 0.05) %>%
  dplyr::select(Description, p.adjust, GeneRatio, BgRatio) %>%
  head(20) %>%
  rename(Description = "Enriched GO Terms in Treated SMAD3 Peaks Restricted to Promoters") %>%
  kable(x = . , "latex", longtable = T) %>%
  kable_styling(latex_options = c("repeat_header"))
```

| Enriched GO Terms in Treated SMAD3 Peaks Restricted to Promoters | p.adjust  | GeneRatio | BgRatio   |
|--|-----------|-----------|-----------|
| regulation of small GTPase mediated signal transduction          | 0.0000001 | 99/2990   | 321/18493 |
| regulation of Ras protein signal transduction                    | 0.0000001 | 76/2990   | 222/18493 |
| regulation of GTPase activity                                    | 0.0000017 | 129/2990  | 481/18493 |
| positive regulation of GTPase activity                           | 0.0000017 | 113/2990  | 405/18493 |
| regulation of Rho protein signal transduction                    | 0.0000017 | 51/2990   | 136/18493 |
| Rho protein signal transduction                                  | 0.0002438 | 61/2990   | 199/18493 |
| activation of GTPase activity                                    | 0.0007166 | 34/2990   | 91/18493  |
| Ras protein signal transduction                                  | 0.0013685 | 107/2990  | 430/18493 |
| epithelial tube morphogenesis                                    | 0.0045268 | 81/2990   | 313/18493 |
| extracellular matrix organization                                | 0.0230640 | 85/2990   | 348/18493 |
| pinocytosis  | 0.0230640 | 11/2990   | 19/18493  |
| morphogenesis of an epithelium                                   | 0.0276288 | 110/2990  | 479/18493 |
| extracellular structure organization                             | 0.0276288 | 95/2990   | 402/18493 |
| regulation of phosphoprotein phosphatase activity                | 0.0349439 | 35/2990   | 114/18493 |
| endosomal transport  | 0.0392397 | 55/2990   | 208/18493 |
| neural tube closure  | 0.0392397 | 29/2990   | 89/18493  |
| tube closure   | 0.0401784 | 29/2990   | 90/18493  |
| histone H3-K36 methylation                                       | 0.0401784 | 8/2990    | 12/18493  |
| distal tubule development  | 0.0401784 | 8/2990    | 12/18493  |
| tube formation   | 0.0492196 | 41/2990   | 145/18493 |

```
dotplot(prom_ego_SMAD3_LAP_treated_peaks) + ggtitle(label = "Enriched BP GO Term in Treated SMAD3 Promoter Peaks")
```

