## main

## August 16, 2021

## 1 Visualize GO analysis

```
[1]: import numpy as np
     import pandas as pd
[2]: def get_top_GO(tissue, fn, label):
         df = pd.read_csv(fn, sep='\t').sort_values('p_value').head(10)
         df['Log10'] = -np.log10(df['p_value'])
         df['Tissue'] = tissue
         df['Direction'] = label
         return df
[3]: tissue = 'caudate'
     config = {
         'All': '../../_m/DEGs_functional_enrichment.txt',
         'Up': '../../_m/upreg_DEGs_functional_enrichment.txt',
         'Down': '../../_m/downreg_DEGs_functional_enrichment.txt',
     }
     df = pd.DataFrame()
     for bias in ['All', 'Up', 'Down']:
         df = pd.concat([df, get_top_GO(tissue, config[bias], bias)], axis=0)
[4]: df.to_csv("%s_functional_analysis.txt" % tissue, sep='\t', index=False)
[5]: df.shape
[5]: (30, 17)
    1.1 Plot
[6]: %load_ext rpy2.ipython
[7]: \%\R -i df
     library(ggplot2)
     library(tidyverse)
     save_plot <- function(p, fn, w, h){</pre>
```

```
for(ext in c('.svg', '.png', '.pdf')){
             ggsave(file=paste0(fn,ext), plot=p, width=w, height=h)
         }
     }
     plot_GO <- function(){</pre>
         cbPalette <- c("#000000", "Red", "Blue")
         gg1 = df \% > \%
             ggplot(aes(x=Log10, y=term_name, color=Direction)) +
             geom_point(shape=18, alpha=0.8, size=4) + labs(y='', x='-Log10 (pu
     →adjust)') +
             theme_bw() +
             scale_colour_manual(name="Direction", values=cbPalette,
                                 labels=c("All", "Upregulated", "Downregulated")) +
             geom_vline(xintercept = -log10(0.05), linetype = "dotted") +
             theme(axis.text=element_text(size=14),
                   axis.title=element_text(size=18, face='bold'),
                   strip.text=element_text(size=18, face='bold'))
         return(gg1)
     }
    R[write to console]:
                           Attaching packages
                          tidyverse 1.3.1
    R[write to console]: tibble 3.1.2
                                              dplyr 1.0.7
            1.1.3
                         stringr 1.4.0
     tidyr
     readr 1.4.0
                         forcats 0.5.1
     purrr
             0.3.4
    R[write to console]:
                           Conflicts
    tidyverse_conflicts()
     dplyr::filter() masks stats::filter()
     dplyr::lag()
                     masks stats::lag()
[8]: %%R
     gg1 = plot_GO()
     print(gg1)
     save_plot(gg1, "caudate_G0_top10_stacked", 10, 6)
```

```
transcription factor IIIC multisubunit complex-
                     solute:sodium symporter activity-
        sodium:potassium:chloride symporter activity-
                   sodium:chloride symporter activity-
               SMG-1-UPF-ERF1-ERF3 complex (SURF)-
                           SMAD3-SKI-NCOR complex-
                           SMAD2-SKI-NCOR complex-
                                SHC3-GAB1 complex-
                           SETDB1-DNMT3B complex-
                           SETDB1-DNMT3A complex-
                 SETDB1-containing HMTase complex-
     Replication-coupled CAF-1-MBD1-ETDB1 complex-
                 potassium:sodium symporter activity-
                                                       Direction
                potassium: chloride symporter activity-
                                                          All
                            p400-associated complex-
                                                          Upregulate
                               NCOR-HDAC3 complex-
                                                          Downregu
                                      NCOR complex-
                      MBD1-MCAF1-SETDB1 complex-
                            LRRC8A-LRRC8D complex-
                    HDAC3-H1.3-SMRT-N-Cor complex-
actor: AP-2alpha; motif: NGCCYSNNGSN; match class: 1-
                     DNA ligase IV-condensin complex-
                                 Condensin I complex-
                    cation:chloride symporter activity-
                              CAP(C)-CAP(E) complex-
                     anion:sodium symporter activity-
            anion transmembrane transporter activity-
                             Angiogenin-PRI complex-
                     5S-DNA-TFIIIA-TFIIIC2 subcomplex-
                                                   2.0
```

## -Log10 (p adjust)

[]: