Figure-S4A.R

sokole

2024-07-28

```r
# This Script Generates Figure S4A
# Script By: Eishani Kumar Sokolowski

# Empty the environment & suppress warnings
rm(list = ls())
options(warn=-1)

# Loading libraries
library(tidyverse)
```

```
## ── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──
## ✔ dplyr     1.1.4     ✔ readr     2.1.5
## ✔ forcats   1.0.0     ✔ stringr   1.5.1
## ✔ ggplot2   3.5.1     ✔ tibble    3.2.1
## ✔ lubridate 1.9.3     ✔ tidyr     1.3.1
## ✔ purrr     1.0.2
## ── Conflicts ──────────────────────────────────── tidyverse_conflicts() ──
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()    masks stats::lag()
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts
to become errors
```

```r
library(ggplot2)
library(ggrepel)
library(ggpubr)
library(data.table)
```

```
##
## Attaching package: 'data.table'
##
## The following objects are masked from 'package:lubridate':
##
##     hour, isoweek, mday, minute, month, quarter, second, wday, week,
##     yday, year
##
## The following objects are masked from 'package:dplyr':
##
##     between, first, last
##
## The following object is masked from 'package:purrr':
##
##     transpose
```

```
library(dplyr)
library(ggpubr)
library(rstatix)
```

```
##
## Attaching package: 'rstatix'
##
## The following object is masked from 'package:stats':
##
##     filter
```

```
################################################################################
# Preparing the file
################################################################################

# Loading the file
df <- read.csv("./Final_Closing_distal_and_proximal_DAPs_with_SNPs.csv")

# Removing traits with 0 overlapping SNPs
traits.to.remove <- c("Disposition_index", "HOMA-IR")

# Making dataframe
df <- df[!df$Trait %in% traits.to.remove,]

# Changing Type Name
df$Type <- gsub("ERS.specific","C_ERS.specific",df$Type)
df$Type <- gsub("INF.specific","B_INF.specific",df$Type)
df$Type <- gsub("Shared","A_Shared",df$Type)

# Factor levels
df$Trait <- factor(df$Trait, levels = c("HOMA-B",
                                          "Insulinogenic_Index",
                                          "Proinsulin",
                                          "2hrGlu",
                                          "HbA1c",
                                          "FI",
                                          "FG",
                                          "Type 1 Diabetes",
                                          "Type 2 Diabetes"))


# Plotting the barplot
p <- ggplot(data=df, aes(x=Number.of.Overlapping.Variants, y=Trait, fill=Type)) +
  geom_bar(stat="identity", color="black", position = "dodge")+
  theme_bw() + scale_fill_manual(values=c('tan4','darkorange','forestgreen')) + xlim(0,4
0)

# View
ggarrange(p, nrow=1, ncol=1)
```