Heatmap plotting

Chia-Yu Chen

4/16/2021

Abstract

This is for plotting the heatmap of module-metabolite and species-metabolite together. To plot this heatmap, one needs to prepare correlation tables (association and p values) in long-format beforehand.

Contents

Data preparation

```
rm(list = ls())
setwd("~/Documents/Lab/Ellen_gremFree/analysis/")
library(tidyverse)
library(readxl)
library(gplots)
```

Read in module-metabolites correation p values and association table

```
module_corr_q <- read.table("Functiona_module_LowConc_metabolite_corr_q_noMF.txt",
    header = T, sep = "\t", row.names = 1, stringsAsFactors = F)  # this is p value table
module_assoc <- read.table("Functional_module_LowConc_metabolite_assoc_noMF.txt",
    header = T, sep = "\t", row.names = 1, stringsAsFactors = F)  # this is association table
```

Take a look at module_corr_q

```
module_corr_q[1:4, 1:4]
```

```
##                           M00004     M00006     M00008     M00015
## Octadecanoylcarnitine  0.1787842 0.08648855 0.05699321 0.14983414
## Ceramide.d18.1.16.0.   0.6802192 0.42483661 0.35231918 0.09123641
## Ceramide.d18.1.24.1.   0.5154956 0.37975066 0.25350368 0.06021169
## Cholesteryl.ester.14.0 0.7386646 0.43545930 0.25665975 0.06676265
```

Take a look at module_assoc

```
module_assoc[1:4, 1:4]
```

```
##                           M00004     M00006     M00008     M00015
## Octadecanoylcarnitine  -0.4323529 -0.5705882 -0.6441176 -0.4676471
## Ceramide.d18.1.16.0.   -0.1235294 -0.2441176 -0.2882353 -0.5617647
```

```
## Ceramide.d18.1.24.1.   -0.1970588 -0.2705882 -0.3617647 -0.6323529
## Cholesteryl.ester.14.0 -0.1000000 -0.2382353 -0.3588235 -0.6147059
```

Read in species-metabolites correlation p values and association table

```
species_corr_q <- read.table("Species_LowConc_metabolite_corr_q.txt",
    header = T, sep = "\t", row.names = 1, stringsAsFactors = F)
species_assoc <- read.table("Species_LowConc_metabolite_assoc.txt",
    header = T, sep = "\t", row.names = 1, stringsAsFactors = F)
```

Add description for module-metabolites first before merging

```
module.defs.column1 <- read.table("../../Salt/TS4/Kegg/module.defs.column1.defs",
    sep = "\t", stringsAsFactors = F, quote = "", header = F)
names <- rbind(module.defs.column1)
colnames(names) <- c("Module", "Description")
names$Description <- gsub("\"", "", names$Description, fixed = TRUE)

module_corr_q_t <- as.data.frame(t(module_corr_q))
module_corr_q_t <- module_corr_q_t %>% rownames_to_column("Module")
module_corr_q_t_des <- inner_join(names, module_corr_q_t, "Module")

a <- str_split_fixed(module_corr_q_t_des$Description, ",", n = 2)
module_corr_q_t_des$Description <- a[, 1]
module_corr_q_t_des$Module <- paste(module_corr_q_t_des$Module,
    module_corr_q_t_des$Description, sep = " ")
module_corr_q_t_des$Description <- NULL
module_corr_q_t_des <- module_corr_q_t_des %>% column_to_rownames("Module")
module_corr_q_des <- t(module_corr_q_t_des)  #Take this for merging

module_assoc_des <- module_assoc
colnames(module_assoc_des) <- colnames(module_corr_q_des)  #Take this for merging
```

Combine the two tables

```
merge_corr_q <- cbind(species_corr_q, module_corr_q_des)
merge_assoc <- cbind(species_assoc, module_assoc_des)

merge_corr_q_long <- merge_corr_q %>%
  rownames_to_column("Metabolite") %>%
  pivot_longer(cols = c(-1))
merge_corr_q_long_sig <- merge_corr_q_long %>%
  filter(value < 0.05)
colnames(merge_corr_q_long_sig) <- c("Metabolite", "Module_or_species", "Spearman_p_fdr")

merge_assoc <- as.data.frame(merge_assoc)
merge_assoc_long <- merge_assoc %>%
  rownames_to_column("Metabolite") %>%
  pivot_longer(cols = c(-1))
colnames(merge_assoc_long) <- c("Metabolite", "Module_or_species", "Spearman_rho")

finalSig <- merge_corr_q_long_sig %>%
  inner_join(merge_assoc_long, by=c("Metabolite","Module_or_species"))
```

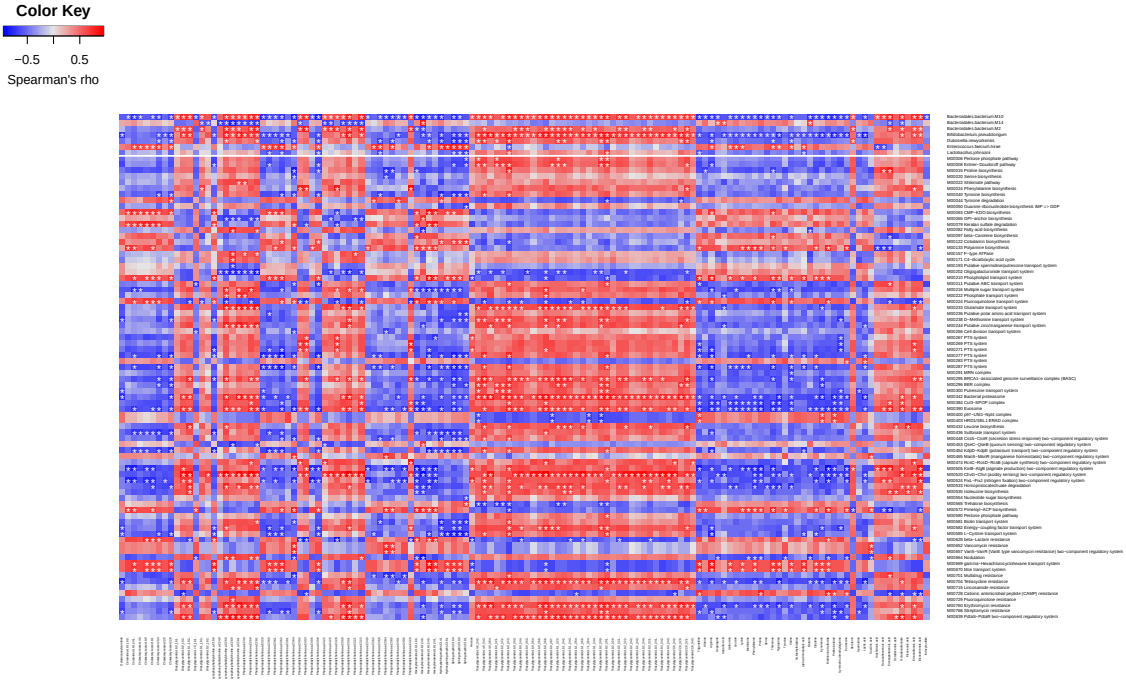
Prepare for heatmap plotting

```
# Select only the modules_species that are in finalSig from
# corr/assoc table
merge_corr_sig <- merge_corr_q[, colnames(merge_corr_q) %in%
    unique(finalSig$Module_or_species)]
merge_assoc_sig <- merge_assoc[, colnames(merge_assoc) %in% unique(finalSig$Module_or_species)]

# Then select the metabolites that are in finalSig from
# corr/assoc table
merge_corr_sig <- merge_corr_sig[rownames(merge_corr_sig) %in%
    unique(finalSig$Metabolite), ]
merge_assoc_sig <- merge_assoc_sig[rownames(merge_assoc_sig) %in%
    unique(finalSig$Metabolite), ]
assoc_rowname <- rownames(merge_assoc_sig)
merge_assoc_sig <- apply(merge_assoc_sig, 2, as.numeric)
rownames(merge_assoc_sig) <- assoc_rowname
```

```
## Transpose the tables
merge_assoc_plotting_t <- t(merge_assoc_sig)

## Create a matrix specifying the label of each cell in heatmap
## Only significant correlations are labelled with asterisk
merge_corr_sig_t <- t(merge_corr_sig)
asterisk <- ifelse(merge_corr_sig_t < 0.05, yes = "*", no = "")
```

Plotting

```
my_palette <- colorRampPalette(c("blue", "#E4E4E4", "red"))(n = 55)
par(mar=c(1,1,1,1))
heatmap.2(merge_assoc_plotting_t,
          dendrogram = "none",
          scale = "none",
          col = c(my_palette), # Color gradient
          trace = "none",
          density.info = "none",
          key.xlab = "Spearman's rho",
          margins = c(12, 15),#Set margin to proper value so that the text won't be cut
          cexRow = 0.3, # Size of row names
          cexCol = 0.2, # Size of col names
          lhei = c(1, 6), # adjust vertical plot layout
          lwid = c(1, 6), # adjust horizontal plot layout
          key = T,
          Rowv = F, #Suppress rows being reordered
          Colv = F, #Suppress cols being reordered
          keysize = 0.1,
          cellnote = asterisk,
          notecex= 0.8,
          notecol="white",
          hclustfun = function(x) hclust(x, method = "single"),
          rowsep=c(7), #Seperation line after 7th value
          sepwidth=c(0.1, 0.1)) #Separation width
```


Color Key
−0.5
0.5
Spearman's rho