Reactome Pathway Results

Annika Jorgensen

2022-12-07

Title: “Reactome Pathway Results”
Author: Annika Jorgensen
Date: 12/07/2022 Purpose: The purpose of this file is to generate comparisons and find commonalities and differences between lists of pathways created by doing a Reactome pathway analysis. The comparisons are done for overall tumor tumor adjacent, male and female tumor tumor adjacent, HBV and HCV tumor tumor adjacent

```
setwd('~/R/Liver Cancer Project')
```

Tumor vs. Tumor-adjacent and sex stratified comparisons

A gene list was generated from a differential expression analysis (DEA) and input into Reactome. A CSV file containing all the pathway enrichment information was downloaded. The CSV files are then read in and stored into variables.

“tumor_tumor_adjacent” contains the pathway information from an overall tumor-tumor adjacent DEA.

“male_all” contains the pathway information from a male tumor-tumor adjacent DEA.

“female_all” contains the pathway information from a female tumor-tumor adjacent DEA.

```
tumor_tumor_adjacent <- read.csv("~/R/Liver Cancer Project/Pathway Analysis Results/Tumor Tumor
Adjacent.csv")

male_all <- read.csv("~/R/Liver Cancer Project/Pathway Analysis Results/Male_results.csv")

female_all <- read.csv("~/R/Liver Cancer Project/Pathway Analysis Results/Female_results.csv")
```

The pathway information is being filtered by a false discovery rate (FDR) of less than 0.05. The which() command is used to locate which rows in the “Entities.FDR” column (the column containing the FDRs) have values less than 0.05 in the file, and then outputs those rows into the “subsetted” variables.

```
tumor_tumor_adjacent_Subset <- tumor_tumor_adjacent[which(tumor_tumor_adjacent$Entities.FDR < 0.
05), ]

male_all_subset <-  male_all[which(male_all$Entities.FDR < 0.05), ]

female_all_subset <- female_all[which(female_all$Entities.FDR < 0.05), ]
```

Here, we are finding pathways that are common between multiple variables. The commands all have the same structure, where we locate which pathway names are in another variable and output those rows into a new variable. The contents of the shared variables are listed below:

“tumor_tumor_adjacent_male” are the pathways that are in common between the overall tumor-tumor adjacent and male tumor-tumor adjacent comparisons.

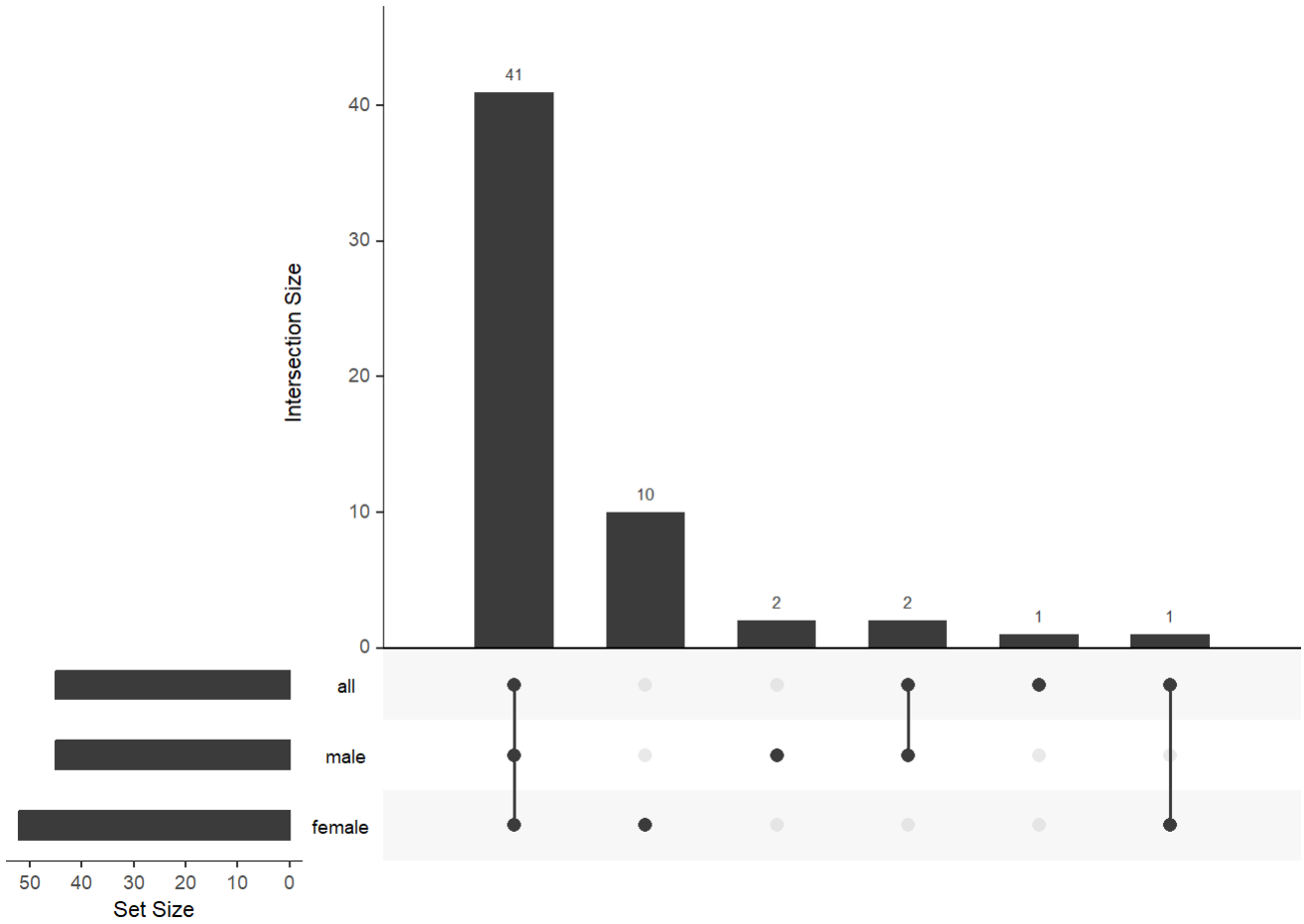
“tumor_tumor_adjacent_female” are the pathways that are in common between the overall tumor-tumor adjacent and female tumor-tumor adjacent comparisons.

“male_female” are the pathways that are in common between the male and female tumor-tumor adjacent comparisons.

```
 tumor_tumor_adjacent_male <- tumor_tumor_adjacent_Subset[which(tumor_tumor_adjacent_Subset$Path
way.name %in% male_all_subset$Pathway.name), ]

tumor_tumor_adjacent_female <- tumor_tumor_adjacent_Subset[which(tumor_tumor_adjacent_Subset$Pat
hway.name %in% female_all_subset$Pathway.name), ]

male_female <- male_all_subset[which(female_all_subset$Pathway.name %in% male_all_subset$Pathwa
y.name), ]
```

This code generates an upset plot for the overall tumor-tumor adjacent, male, and female tumor-tumor adjacent comparisons. The upset plot functions like a multiple Venn diagram, illustrating how many pathways are shared between multiple sets. In this specific plot, 41 pathways are shared by all, 10 are unique to female, 2 are unique to male, 2 are shared by tumor-tumor adjacent and male only, 1 is unique to tumor-tumor adjacent, and 1 is shared by female and tumor-tumor adjacent only.

```
input <- list(tumor_tumor_adjacent_Subset$Pathway.name, male_all_subset$Pathway.name, female_all
_subset$Pathway.name)

names(input) <- c("all", "male", "female")

upset(fromList(input), order.by= "freq", nsets=3)
```




```
pdf("~/R/Liver Cancer Project/Pathway Analysis Results/upsetplot_sex_stratified.pdf", width = 1
2, height = 12)
 upset(fromList(input), order.by= "freq", nsets=3)
 dev.off()
```

```
## png
##   2
```

This chunk generates a variable called “male_shared” which stores the pathway information for all 41 pathways that are enriched in all of the comparisons. The chunk also writes out a CSV file containing the names of these pathways. This is achieved by selecting pathways from the “tumor_tumor_adjacent_male” variable (which contains pathways shared by the overall tumor-tumor adjacent comparison and the male tumor-tumor adjacent comparison) that are also present in the “female_all_subset” variable (which contains all filtered pathway enrichment information from the female tumor-tumor adjacent comparison). The selected pathways are then stored in the “male_shared” variable.

```
male_shared <- tumor_tumor_adjacent_male[which(tumor_tumor_adjacent_male$Pathway.name %in% femal
e_all_subset$Pathway.name), ]

write.csv(male_shared$Pathway.name, "~/R/Liver Cancer Project/Pathway Analysis Results/Shared Pa
thways.csv")
```

This code creates a variable that contains the pathway information for the 10 pathways that are unique to the female tumor-tumor adjacent comparison. This is done by finding the pathways in the “female_all_subset” variable that are **not** in the “male_all_subset” variable (which contains all the filtered pathway enrichment information from the male tumor-tumor adjacent comparison) and storing it in a variable called “female_unique”. The “female_unique” variable is then overwritten by finding the pathways that are in “female_unique” but **not** in the “tumor_tumor_adjacent_Subset” variable (which contains all the filtered pathway enrichment information from the overall tumor-tumor adjacent comparison).

```
female_unique <- female_all_subset[-which(female_all_subset$Pathway.name %in%male_all_subset$Pat
hway.name), ]

female_unique <- female_unique[-which(female_unique$Pathway.name %in% tumor_tumor_adjacent_Subse
t$Pathway.name), ]
```

The code creates a variable that contains the pathway information for the two pathways that are unique to the male tumor-tumor adjacent comparison. The process is similar to the one used in the previous code chunk.

```
male_unique <- male_all_subset[-which(male_all_subset$Pathway.name %in% female_all_subset$Pathwa
y.name), ]

male_unique <- male_unique[-which(male_unique$Pathway.name %in% tumor_tumor_adjacent_Subset$Path
way.name), ]
```

The code creates a variable that contains the pathway information for the one pathway that is unique to the overall tumor-tumor adjacent comparison. The process is similar to the one used to generate the “female_unique” variable.

```
tumor_tumor_adjacent_unique <- tumor_tumor_adjacent_Subset[-which(tumor_tumor_adjacent_Subset$Pa
thway.name %in% female_all_subset$Pathway.name), ]

tumor_tumor_adjacent_unique <- tumor_tumor_adjacent_unique[-which(tumor_tumor_adjacent_unique$Pa
thway.name %in% male_all_subset$Pathway.name), ]
```

This code generates the two pathways that are shared by the male tumor-tumor adjacent and the overall tumor-tumor adjacent comparisons. The process used is a combination of the one used to output pathways that are unique to a specific comparison and the one used to output the pathway enrichment information shared by all comparisons.

```
shared_male_tumor <- tumor_tumor_adjacent_Subset[which(tumor_tumor_adjacent_Subset$Pathway.name
%in% male_all_subset$Pathway.name), ]

shared_male_tumor <- shared_male_tumor[-which(shared_male_tumor$Pathway.name %in% female_all_sub
set$Pathway.name), ]
```

This chunk generates the one pathway that is shared by overall tumor-tumor adjacent comparisons and female tumor-tumor adjacent comparisons.

```
shared_female_tumor <- tumor_tumor_adjacent_Subset[which(tumor_tumor_adjacent_Subset$Pathway.nam
e %in% female_all_subset$Pathway.name), ]

shared_female_tumor <- shared_female_tumor[-which(shared_female_tumor$Pathway.name %in% male_all
_subset$Pathway.name), ]
```

Etiology Specific Pathway Analysis

Here we are reading a csv that contains in the pathway enrichment information from the HBV and HCV tumor-tumor adjacent comparisons and storing into the “HBV_all” and “HCV_all” variables. “HBV_all” contains the information from the HBV comparisons and “HCV_all” contains the information about the HCV comparisons.

```
HBV_all <- read.csv("~/R/Liver Cancer Project/Pathway Analysis Results/HBV_results.csv")
HCV_all <- read.csv("~/R/Liver Cancer Project/Pathway Analysis Results/HCV_results.csv")
```

This code is filtering the pathways in “HBV_all” and “HCV_all” by their FDR values. “HBV_all_subset” and “HCV_all_subset” contain pathways with false discovery rates less than 0.05

```
HBV_all_subset <-  HBV_all[which(HBV_all$Entities.FDR < 0.05), ]
HCV_all_subset <- HCV_all[which(HCV_all$Entities.FDR < 0.05), ]
```

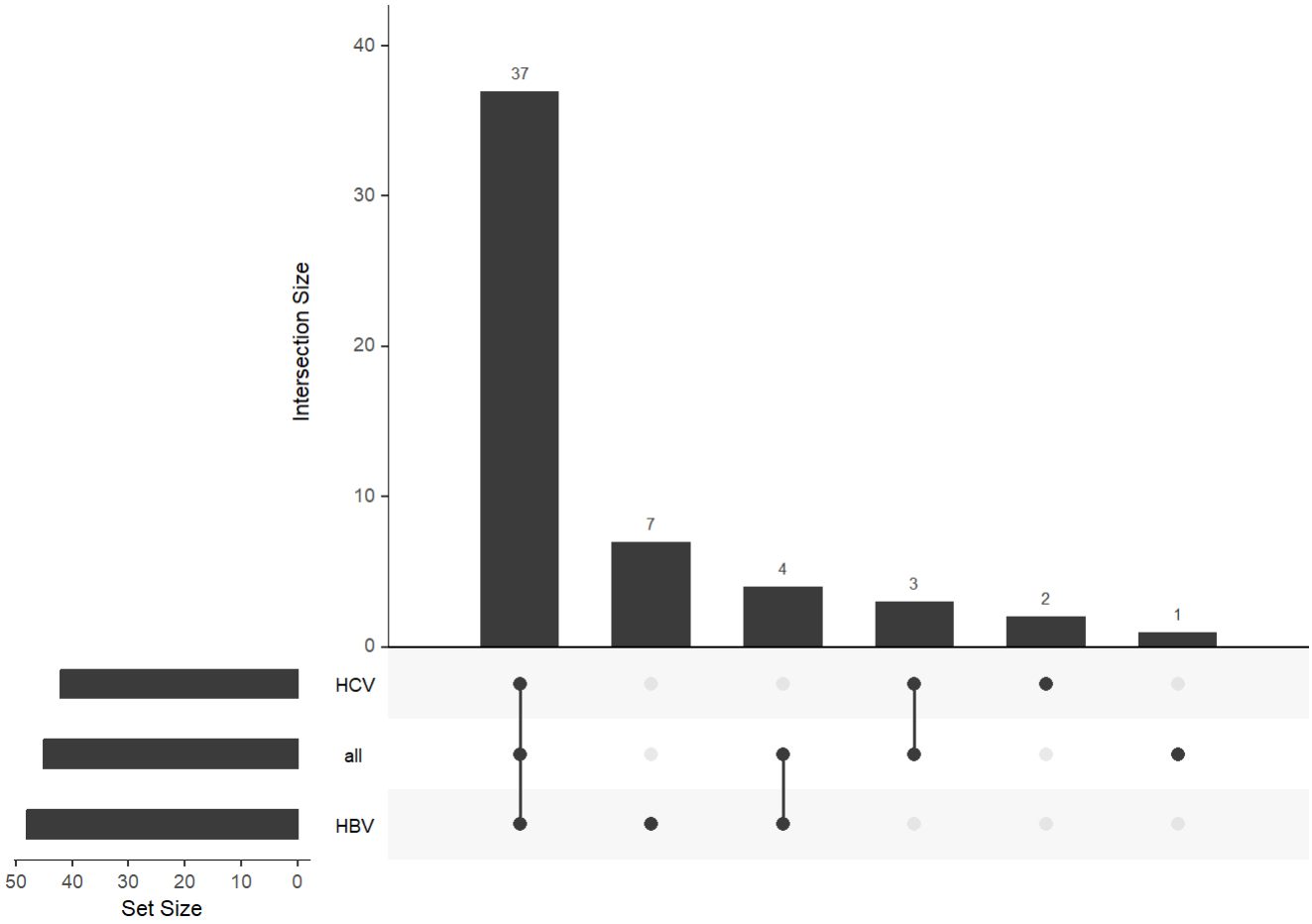
This code creates variables that have paths that are shared. The contents of the variables are listed below

1. The variable “tumor_tumor_adjacent_HBV” contains the pathway enrichment information of pathways shared between the overall tumor-tumor adjacent comparison and the HBV tumor-tumor adjacent comparison.
2. The variable “tumor_tumor_adjacent_HCV” contains the pathway enrichment information of pathways shared between the overall tumor-tumor adjacent comparison and the HCV tumor-tumor adjacent comparison.
3. The variable “HBV_HCV” contains the pathway enrichment information of pathways shared between the HBV and HCV tumor-tumor adjacent comparisons.

```
tumor_tumor_adjacent_HBV <- tumor_tumor_adjacent_Subset[which(tumor_tumor_adjacent_Subset$Pathwa
y.name %in% HBV_all_subset$Pathway.name), ]

tumor_tumor_adjacent_HCV <- tumor_tumor_adjacent_Subset[which(tumor_tumor_adjacent_Subset$Pathwa
y.name %in% HCV_all_subset$Pathway.name), ]

HBV_HCV <- HBV_all_subset[which(HBV_all_subset$Pathway.name %in% HCV_all_subset$Pathway.name), ]
```

This code creates an upset plot using the pathway names from the overall tumor-tumor adjacent, HBV, and HCV datasets. The upset plot reveals that 41 pathways are shared by all comparisons, 7 are unique to HBV, 4 are shared by overall tumor-tumor adjacent and HBV only, 3 are shared by overall tumor-tumor adjacent and HCV only, 2 are unique to HCV, and one is unique to overall tumor-tumor adjacent. The plot is outputted to a PDF file with formatting modifications made to line weight and text size.

```
input <- list(tumor_tumor_adjacent_Subset$Pathway.name, HBV_all_subset$Pathway.name, HCV_all_sub
set$Pathway.name)

names(input) <- c("all", "HBV", "HCV")

upset(fromList(input), order.by= "freq", nsets=3)
```




```
pdf("~/R/Liver Cancer Project/Pathway Analysis Results/upsetplot_etiology_stratified.pdf", width
= 12, height = 12)
 upset(fromList(input), order.by= "freq", nsets=3, line.size=1.5, text.scale= c(5,5,4.5,2.0,5,
5))
 dev.off()
```

```
## png
##   2
```

This code creates a variable that stores the 37 pathways shared by all the comparisons.

```
shared_etiology <- tumor_tumor_adjacent_HBV[which(tumor_tumor_adjacent_HBV$Pathway.name %in% HCV
_all_subset$Pathway.name), ]

write.csv(shared_etiology, "~/R/Liver Cancer Project/Pathway Analysis Results/all_shared_etiolog
y_pathways.csv")
```

This code creates a variable that stores the 7 pathways that are unique to the HBV tumor-tumor adjacent comparisons.

```
HBV_unique <- HBV_all_subset[-which(HBV_all_subset$Pathway.name %in% HCV_all_subset$Pathway.nam
e), ]

HBV_unique <- HBV_unique[-which(HBV_unique$Pathway.name %in% tumor_tumor_adjacent_Subset$Pathwa
y.name), ]
```

This code creates a variable that stores the 2 pathways that are unique to HCV tumor-tumor adjacent comparisons.

```
HCV_unique <- HCV_all_subset[-which(HCV_all_subset$Pathway.name %in% HBV_all_subset$Pathway.nam
e), ]

HCV_unique<- HCV_unique[-which(HCV_unique$Pathway.name %in% tumor_tumor_adjacent_Subset$Pathway.
name), ]
```

This code creates a variable that stores the 4 pathways that are shared by overall tumor-tumor adjacent and HBV only.

```
HBV_and_tumor_tumor_adjacent <- tumor_tumor_adjacent_Subset[which(tumor_tumor_adjacent_Subset$Pa
thway.name %in% HBV_all_subset$Pathway.name), ]

HBV_and_tumor_tumor_adjacent <- HBV_and_tumor_tumor_adjacent[-which(HBV_and_tumor_tumor_adjacent
$Pathway.name %in% HCV_all_subset$Pathway.name), ]
```

This code creates a variable that stores the 3 pathways that are shared by tumor-tumor adjacent and HCV only.

```
HCV_and_tumor_tumor_adjacent <- tumor_tumor_adjacent_Subset[which(tumor_tumor_adjacent_Subset$Pa
thway.name %in% HCV_all_subset$Pathway.name), ]

HCV_and_tumor_tumor_adjacent <-
HCV_and_tumor_tumor_adjacent[-which(HCV_and_tumor_tumor_adjacent$Pathway.name %in% HBV_all_subse
t$Pathway.name), ]
```

This code creates a variable that stores the one pathway that is unique to the overall tumor-tumor adjacent comparison.

```
unique_to_all <- tumor_tumor_adjacent_Subset[-which(tumor_tumor_adjacent_Subset$Pathway.name %in% HBV_all_subset$Pathway.name), ]

unique_to_all <- unique_to_all[-which(unique_to_all$Pathway.name %in% HCV_all_subset$Pathway.name), ]
```