

Dicty analysis

```
library(ggplot2)
```

```
# load the data and visualization functions
```

```
source("global.R")
```

```
source("../utils/visualization_functions.R")
```

Summary statistics Dicty

```
all_data_dicty %>% summarise(Unique_bait = n_distinct(bait))
```

```
## # A tibble: 1 x 1
```

```
##   Unique_bait
```

```
##   <int>
```

```
## 1         12
```

```
all_data_dicty %>% group_by(bait, condition) %>% unite(bait_condition, c(bait, condition), sep = '_') %>%
```

```
## # A tibble: 1 x 1
```

```
##   Unique_bait_condition
```

```
##   <int>
```

```
## 1             15
```

```
# show how many instances of bait and condition there are
```

```
table(all_data_dicty$bait)
```

```
##
```

```
## Galpha2 Galpha4 Galpha8 Gbeta1 Gbeta2 Rac1 RapA RasB RasC RasG1
```

```
## 3206 1603 4100 876 876 1448 461 808 998 337
```

```
## Ric8 Roco4
```

```
## 1448 1448
```

```
table(all_data_dicty$condition)
```

```
##
```

```
## BACKGROUND GDP GppNHp Normal starv veg
```

```
## 1448 1436 1436 10083 1603 1603
```

```
# show spectral count per bait + condition
```

```
all_data_dicty %>% group_by(bait, condition) %>% summarise(.groups = 'drop', across(spectral_count, sum))
```

```
## # A tibble: 15 x 3
```

```
##   bait condition spectral_count
```

```
##   <fct> <fct> <dbl>
```

```
## 1 Galpha2 starv 20155
```

```
## 2 Galpha2 veg 30986
```

```
## 3 Galpha4 Normal 27070
```

```
## 4 Galpha8 GDP 13099
```

```
## 5 Galpha8 GppNHp 23136
```

```
## 6 Galpha8 Normal 23324
```

```
## 7 Gbeta1 Normal 16390
```

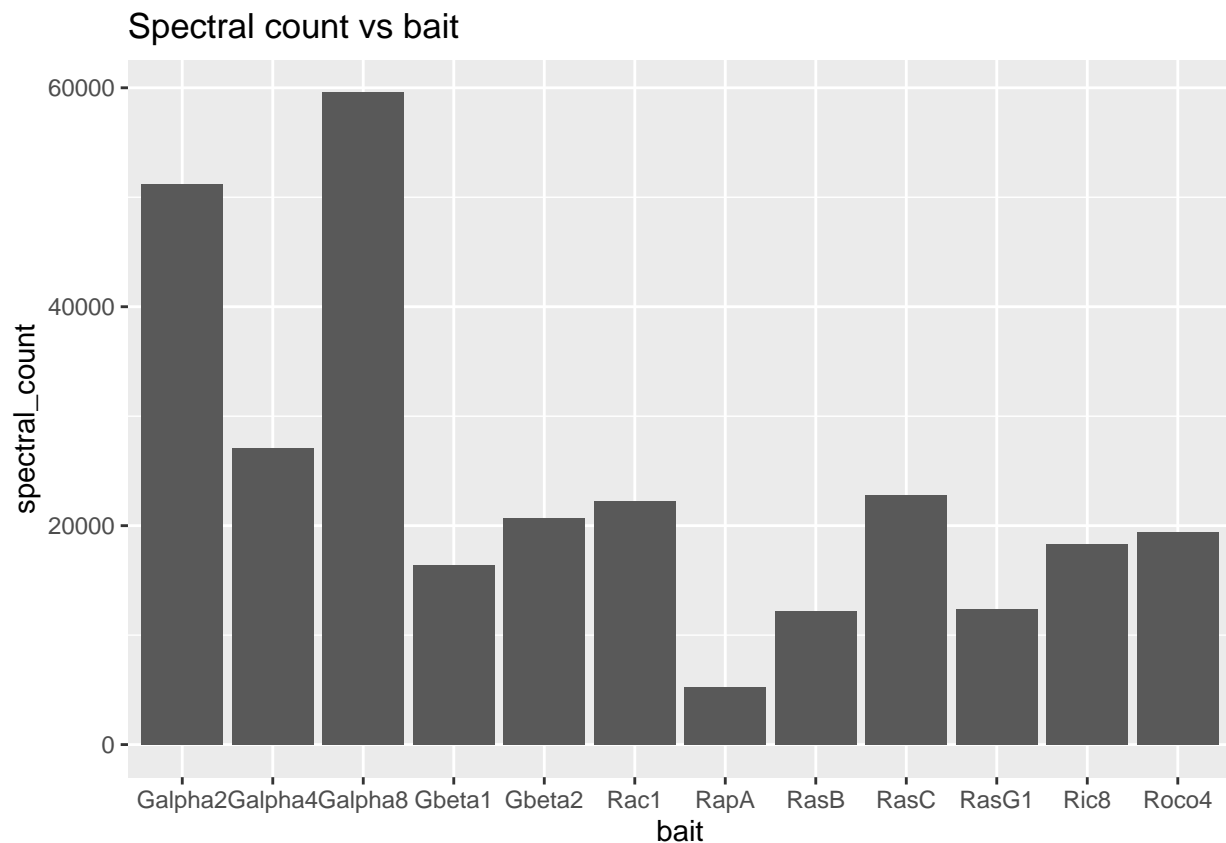
```
## 8 Gbeta2 Normal 20644
## 9 Rac1 Normal 22196
## 10 RapA Normal 5238
## 11 RasB Normal 12182
## 12 RasC Normal 22778
## 13 RasG1 Normal 12369
## 14 Ric8 Normal 18323
## 15 Roco4 BACKGROUND 19376
```

```
# show how many uniq uniprot id's are in the set (dicty has 6 chromosomes and 12500 proteins)
n_distinct(all_data_dicty$uniprot)
```

```
## [1] 2988
```

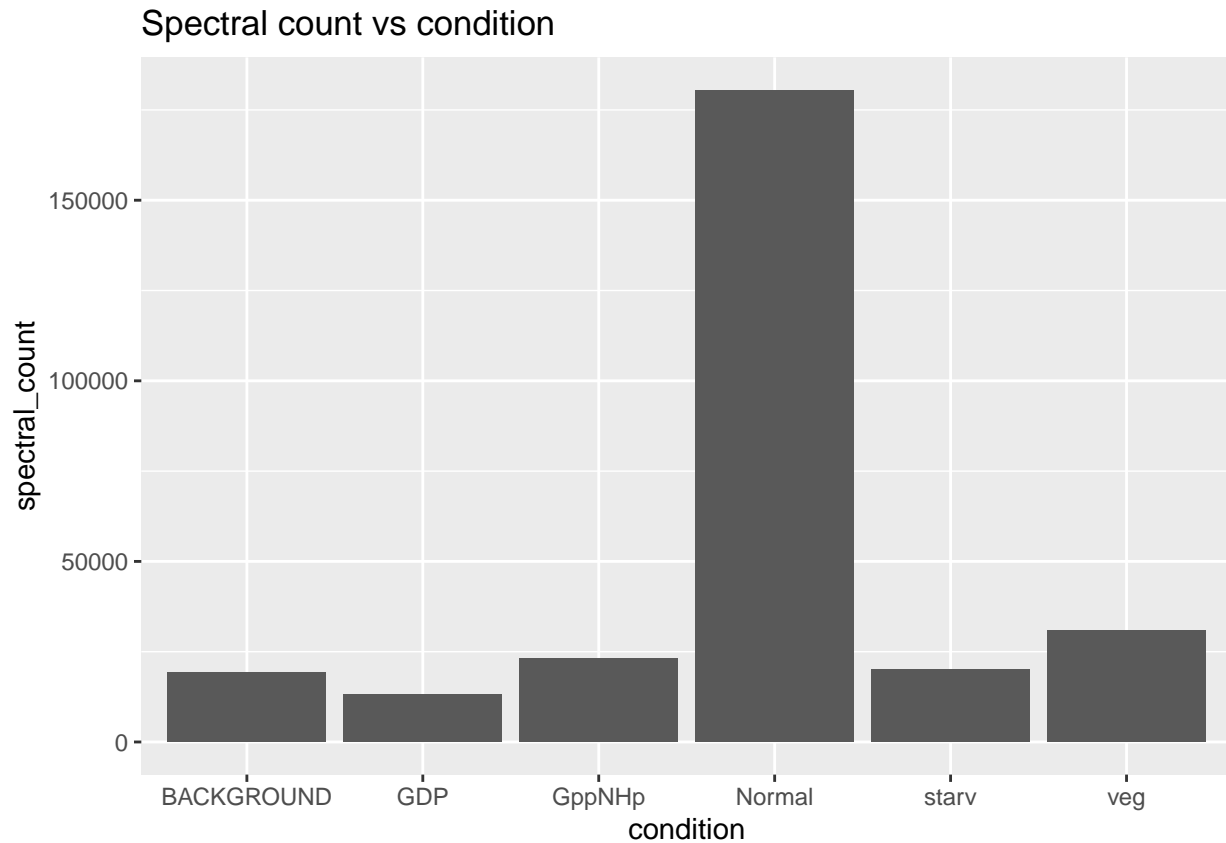
```
# barplot showing the spectral count for each bait
```

```
p <- ggplot(all_data_dicty, aes(x=bait, y=spectral_count))
p <- p + geom_col() + labs(title = 'Spectral count vs bait')
p
```



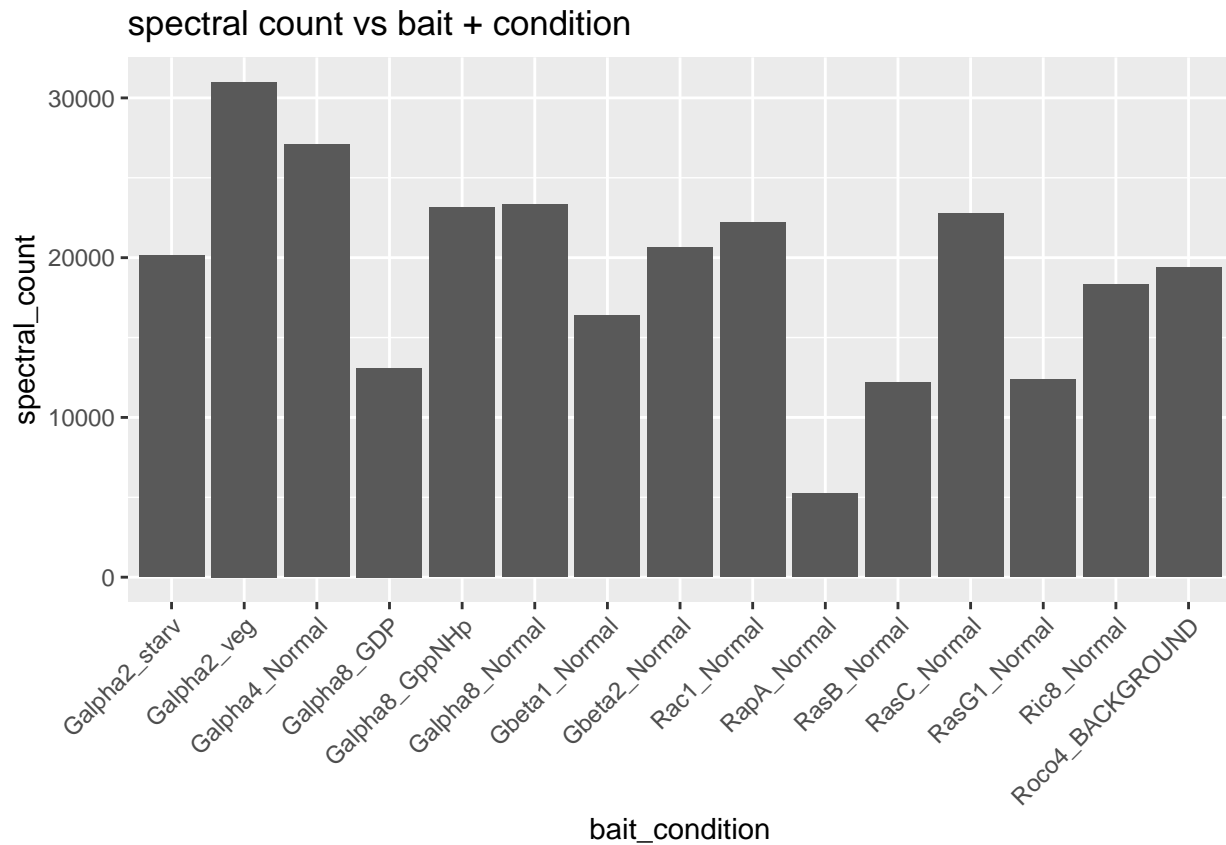
```
# barplot showing the spectral count for each condition
```

```
p <- ggplot(all_data_dicty, aes(x=condition, y=spectral_count))
p <- p + geom_col() + labs(title = 'Spectral count vs condition')
p
```



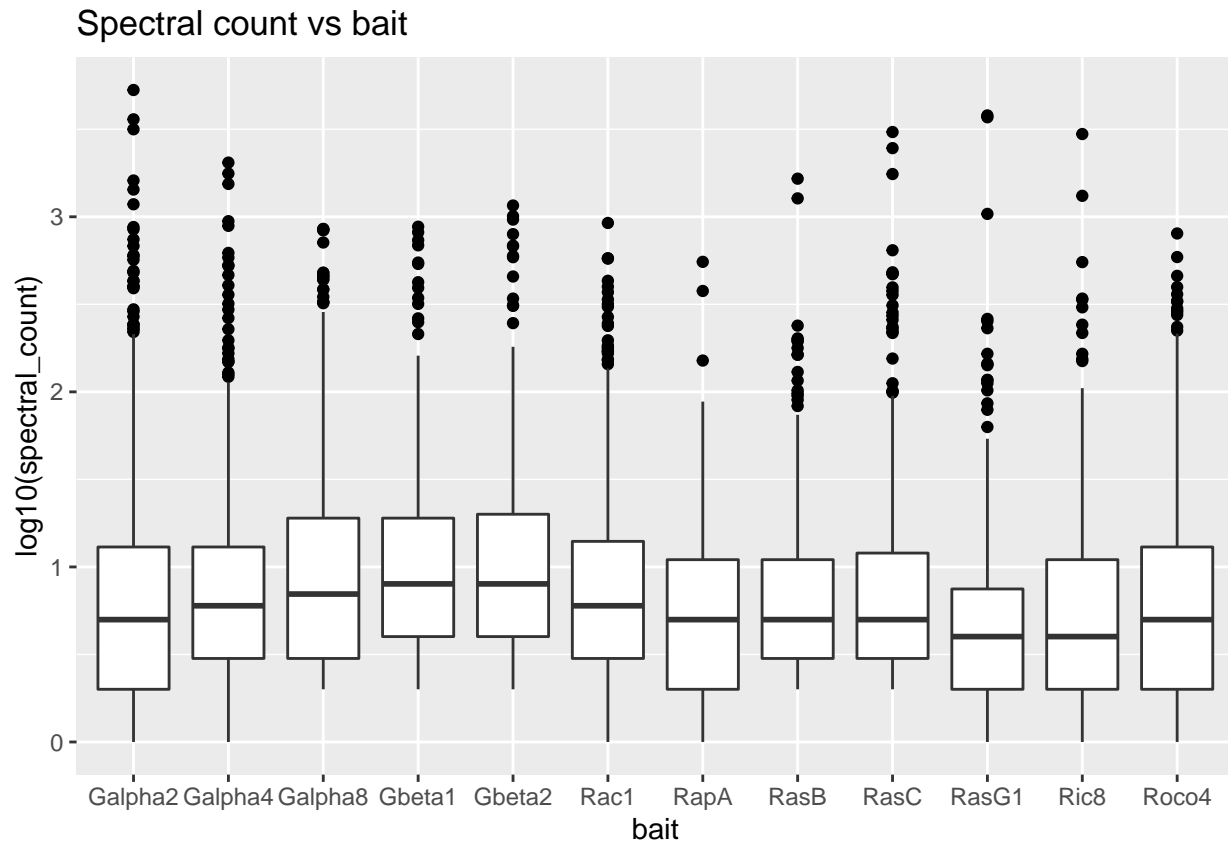
```
# combined spectral count for bait + condition, sum spectral count per group
my_data <- all_data_dicty %>% group_by(bait, condition) %>%
  # .groups = drop is the default, will get warnings if you not specify explicitly
  summarise(.groups = 'drop', across(spectral_count, sum)) %>%
  unite(bait_condition, c(bait, condition), sep = "_")

p <- ggplot(my_data, aes(x = bait_condition, y = spectral_count))
p <- p + geom_col() + labs(title = 'spectral count vs bait + condition')
p + theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



```
# boxplot count vs bait
p <- ggplot(all_data_dicty, aes(x=bait, y =log10(spectral_count)))
p <- p + geom_boxplot(outlier.colour = "black") + labs(title = 'Spectral count vs bait')
p
```

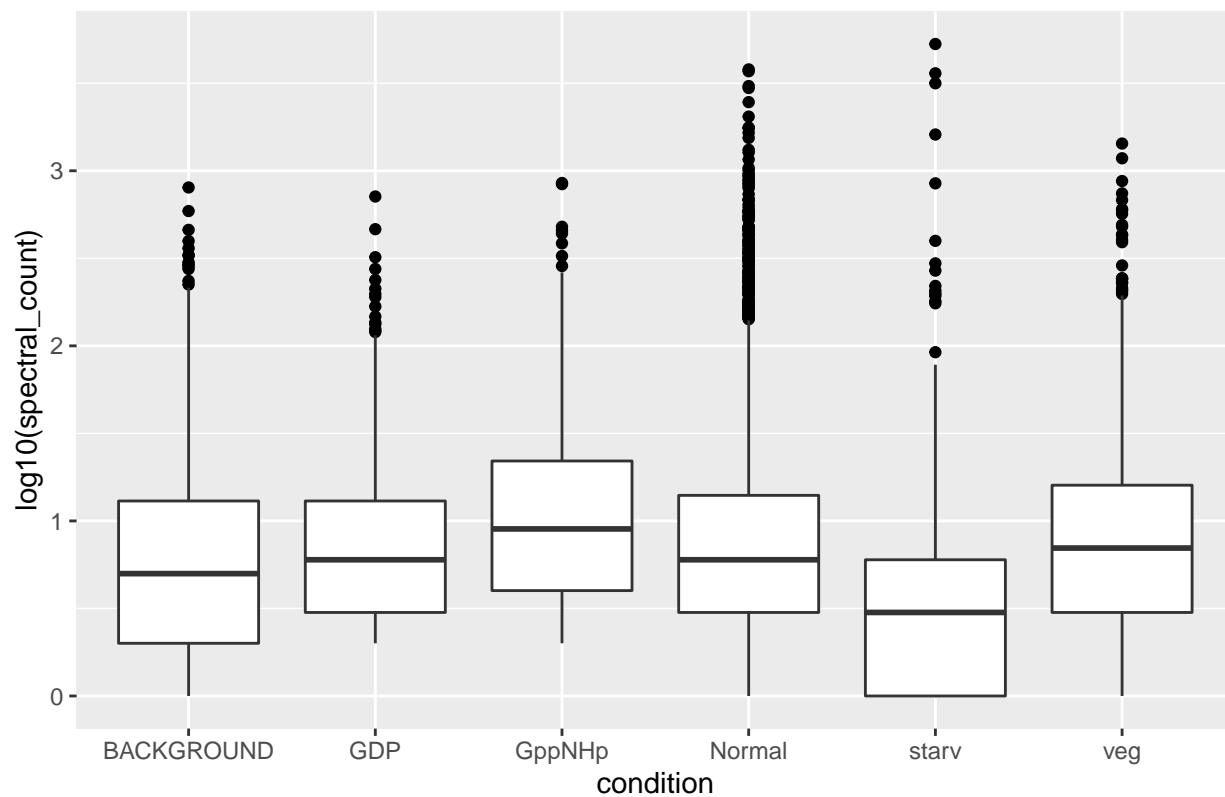
```
## Warning: Removed 3896 rows containing non-finite values (stat_boxplot).
```



```
# boxplot count vs condition
p <- ggplot(all_data_dicty, aes(x=condition, y =log10(spectral_count)))
p <- p + geom_boxplot(outlier.colour = "black") + labs(title = 'Spectral count vs condition')
p
```

```
## Warning: Removed 3896 rows containing non-finite values (stat_boxplot).
```

Spectral count vs condition



```
# boxplot combined spectral count for bait + condition
my_data <- all_data_dicty %>% group_by(bait, condition) %>%
  unite(bait_condition, c(bait, condition), sep = "_")

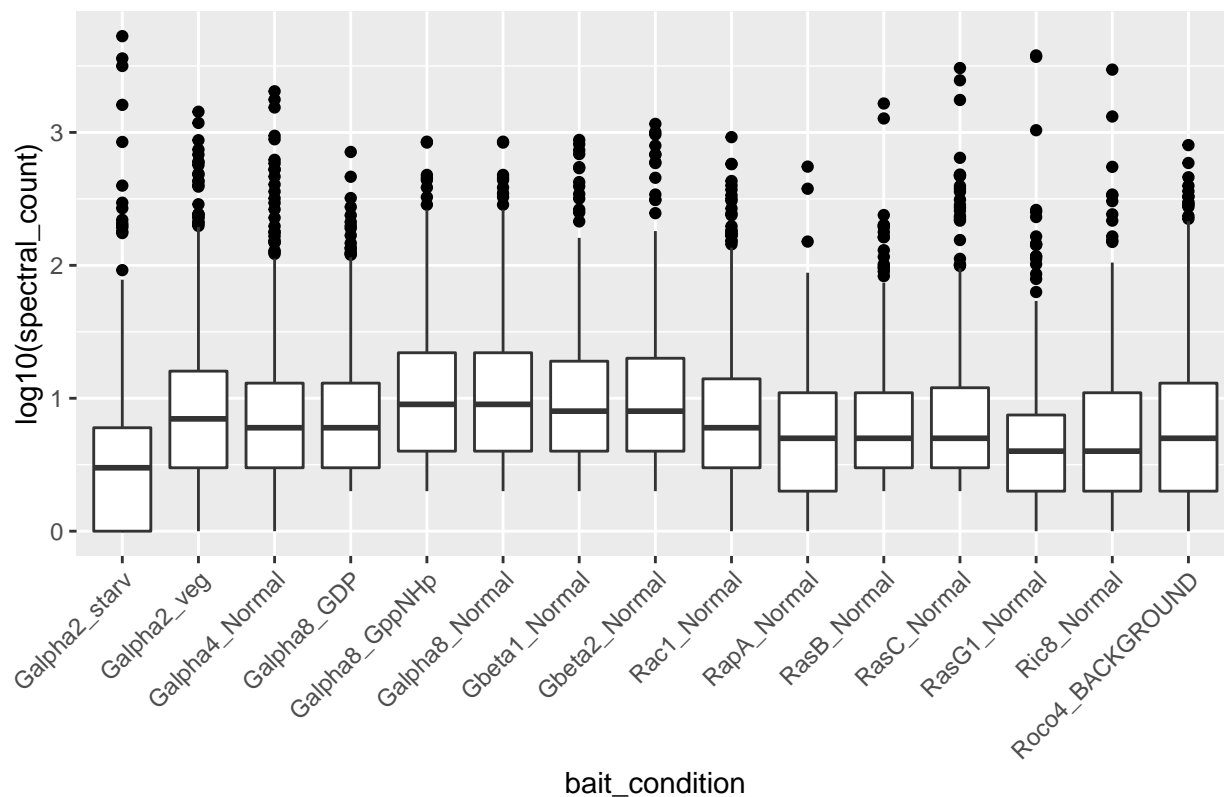
my_data$bait_condition <- as.factor(my_data$bait_condition)
p <- ggplot(my_data, aes(x = bait_condition, y = log10(spectral_count)))

p <- p + geom_boxplot(outlier.colour = "black") + labs(title = 'spectral count vs bait + condition')

p + theme(axis.text.x = element_text(angle = 45, hjust = 1))

## Warning: Removed 3896 rows containing non-finite values (stat_boxplot).
```

spectral count vs bait + condition



uniq proteins per bait_condition

```
bait_condition <- group_by(all_data_dicty, bait, condition)
bait_condition %>% summarise(n_distinct(uniprot))
```

```
## `summarise()` regrouping output by 'bait' (override with `.groups` argument)
```

```
## # A tibble: 15 x 3
## # Groups:   bait [12]
##   bait    condition `n_distinct(uniprot)`
##   <fct>   <fct>          <int>
## 1 Galpha2 starv          1603
## 2 Galpha2 veg            1603
## 3 Galpha4 Normal         1603
## 4 Galpha8 GDP            1436
## 5 Galpha8 GppNHp         1436
## 6 Galpha8 Normal         1228
## 7 Gbeta1 Normal           876
## 8 Gbeta2 Normal           876
## 9 Rac1 Normal           1448
## 10 RapA Normal            461
## 11 RasB Normal            808
## 12 RasC Normal            998
## 13 RasG1 Normal           337
## 14 Ric8 Normal           1448
## 15 Roco4 BACKGROUND      1448
```

```
# uniq per condition
```

```
group_by(all_data_dicty, condition) %>% summarise(n_distinct(uniprot))
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

```
## # A tibble: 6 x 2
##   condition `n_distinct(uniprot)`
##   <fct>          <int>
## 1 BACKGROUND      1448
## 2 GDP              1436
## 3 GppNHp           1436
## 4 Normal           2933
## 5 starv            1603
## 6 veg              1603
```

```
# uniq per bait
```

```
group_by(all_data_dicty, bait) %>% summarise(n_distinct(uniprot))
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

```
## # A tibble: 12 x 2
##   bait      `n_distinct(uniprot)`
##   <fct>          <int>
## 1 Galpha2        1603
## 2 Galpha4        1603
## 3 Galpha8        1446
## 4 Gbeta1         876
## 5 Gbeta2         876
## 6 Rac1           1448
## 7 RapA           461
## 8 RasB           808
## 9 RasC           998
## 10 RasG1         337
## 11 Ric8          1448
## 12 Roco4         1448
```

```
# TODO, check why we have the same numbers for some conditions or baits
#filter(all_data_dicty, bait == "Gbeta2")
#filter(all_data_dicty, bait == "Gbeta1")
#Shows the same number of rows retrieved, counts look different for now....
```

Summary statistics Neutro

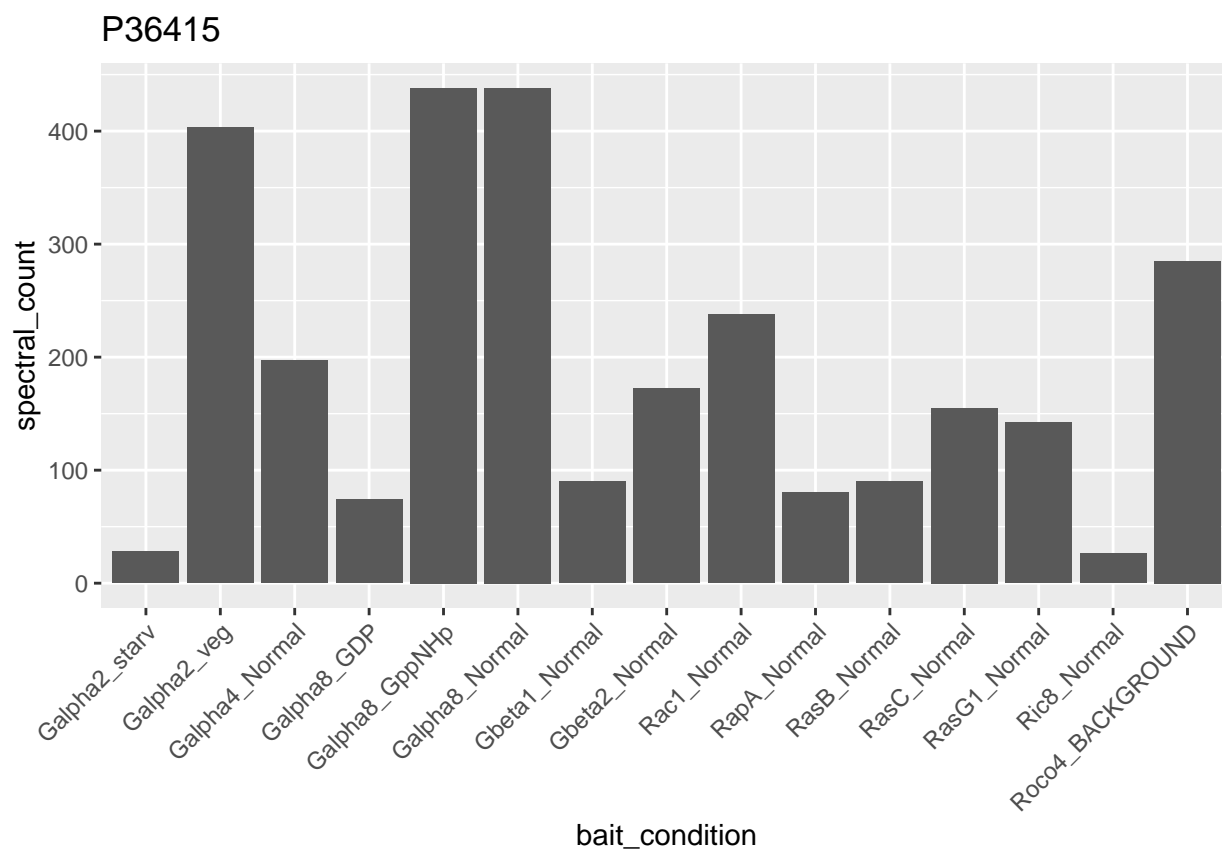
```
table(all_data_neutro$bait)
```

```
##
##      GST  Galphai Galphai2    Rac
##      942    3622     942    942
```

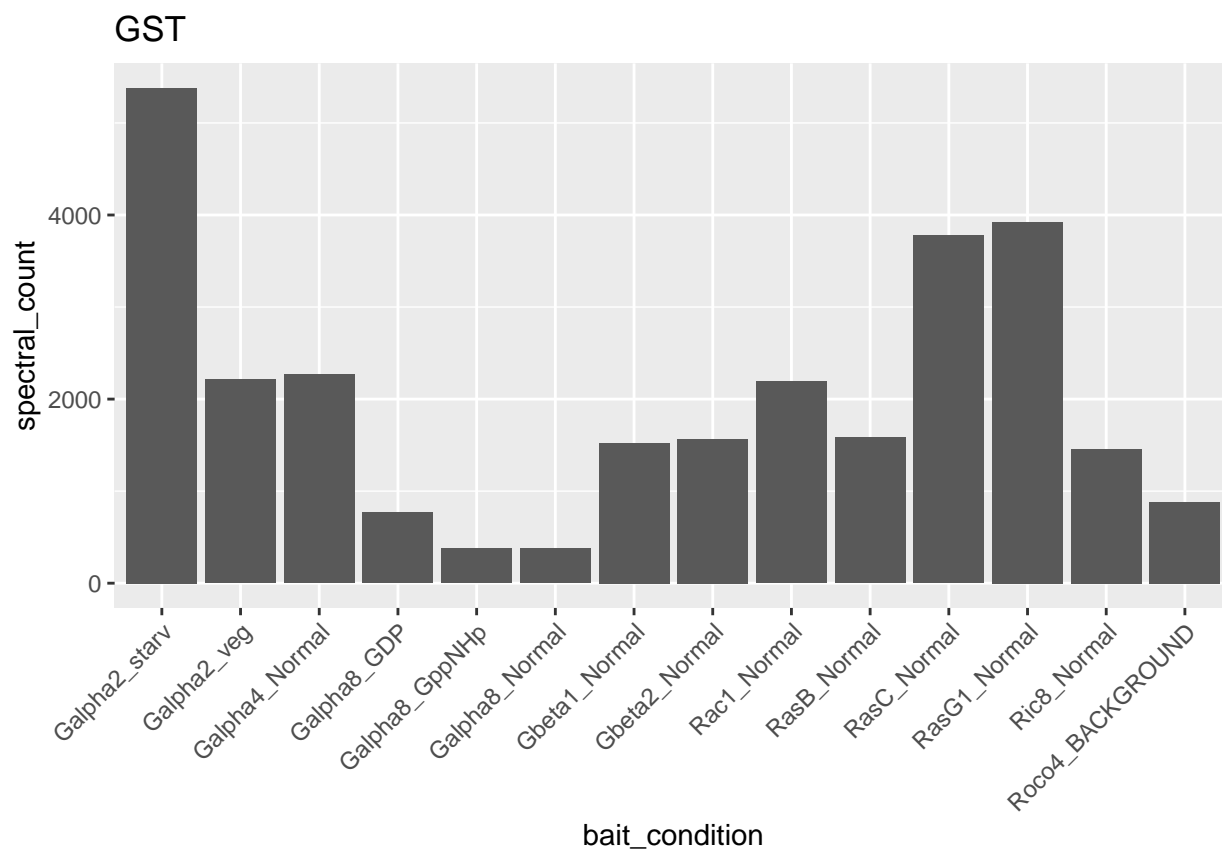
```
table(all_data_neutro$condition)
```

```
##
##   AluF  AluFl    GDP GppNHp
##   670   670   2282  2826
```

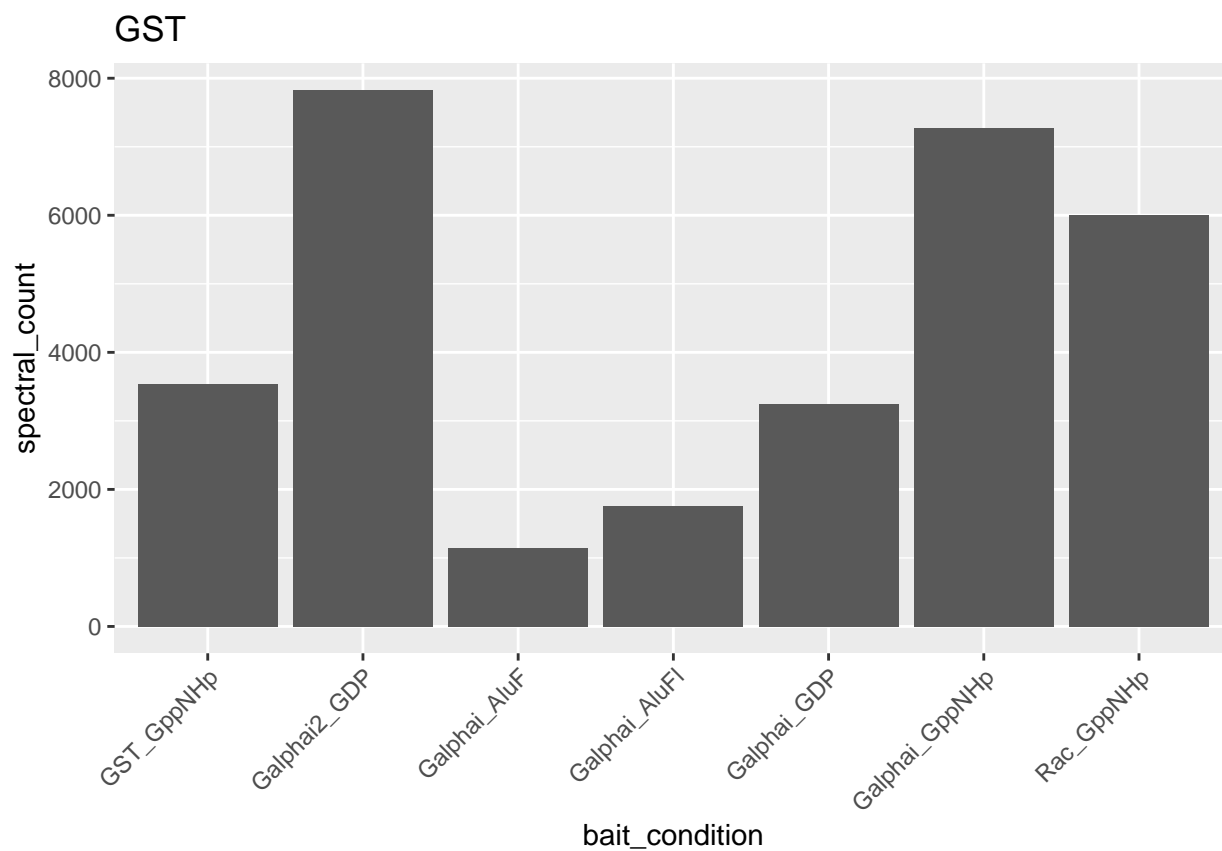
```
plot_spectral_bait_condition_using_uniprot('P36415', 'dicty')
```

```
plot_spectral_bait_condition_using_uniprot('GST', 'dicty')
```



```
plot_spectral_bait_condition_using_uniprot('GST', 'neutro')
```



```
plot_spectral_bait_condition_using_uniprot('Q54YT4', 'dicty')
```

