

# Secreted and Transmembrane

*Maggie Shaw, Temple University*

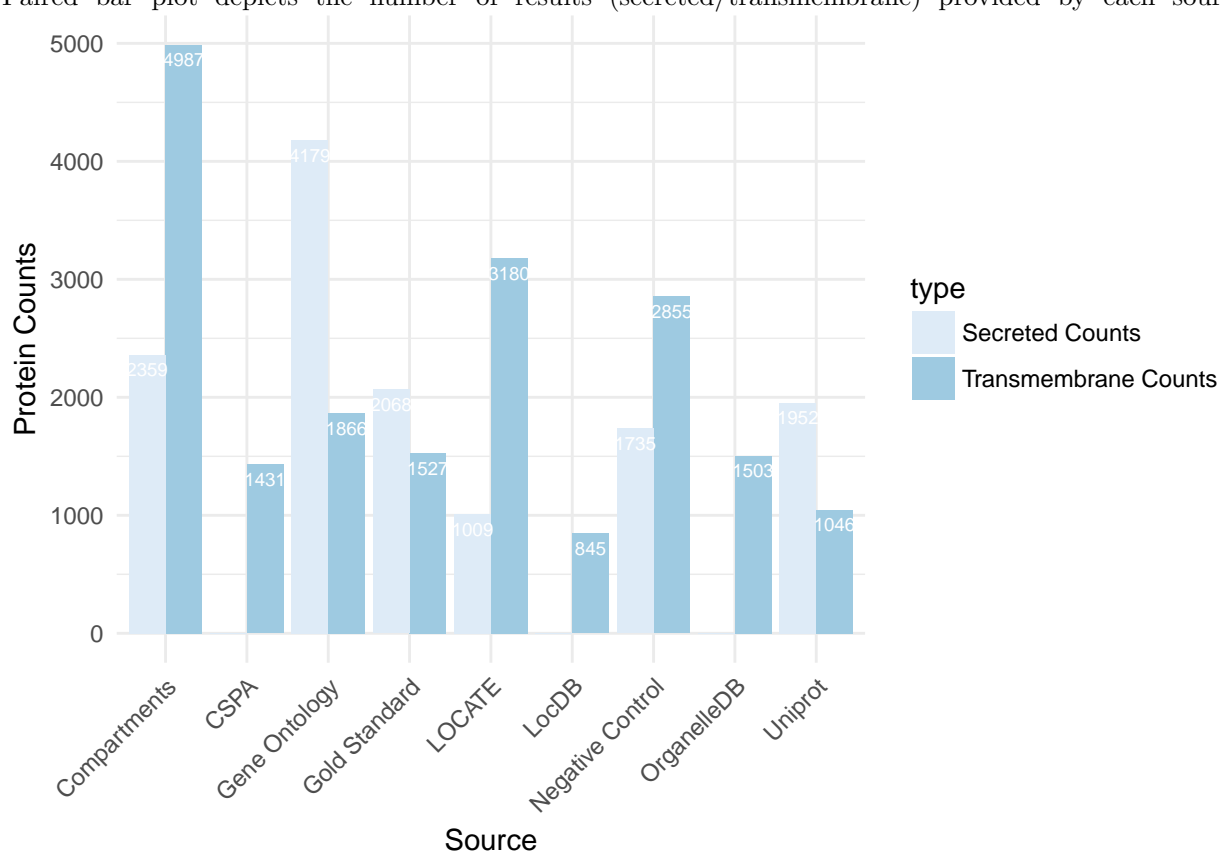
*5/3/2018*

## Objective:

Analyze the database results for each protein (indicated by the corresponding protein coding gene hgnc\_symbol) to determine a final comprehensive list of “transmembrane” and “secreted” proteins.

## Summary of each source

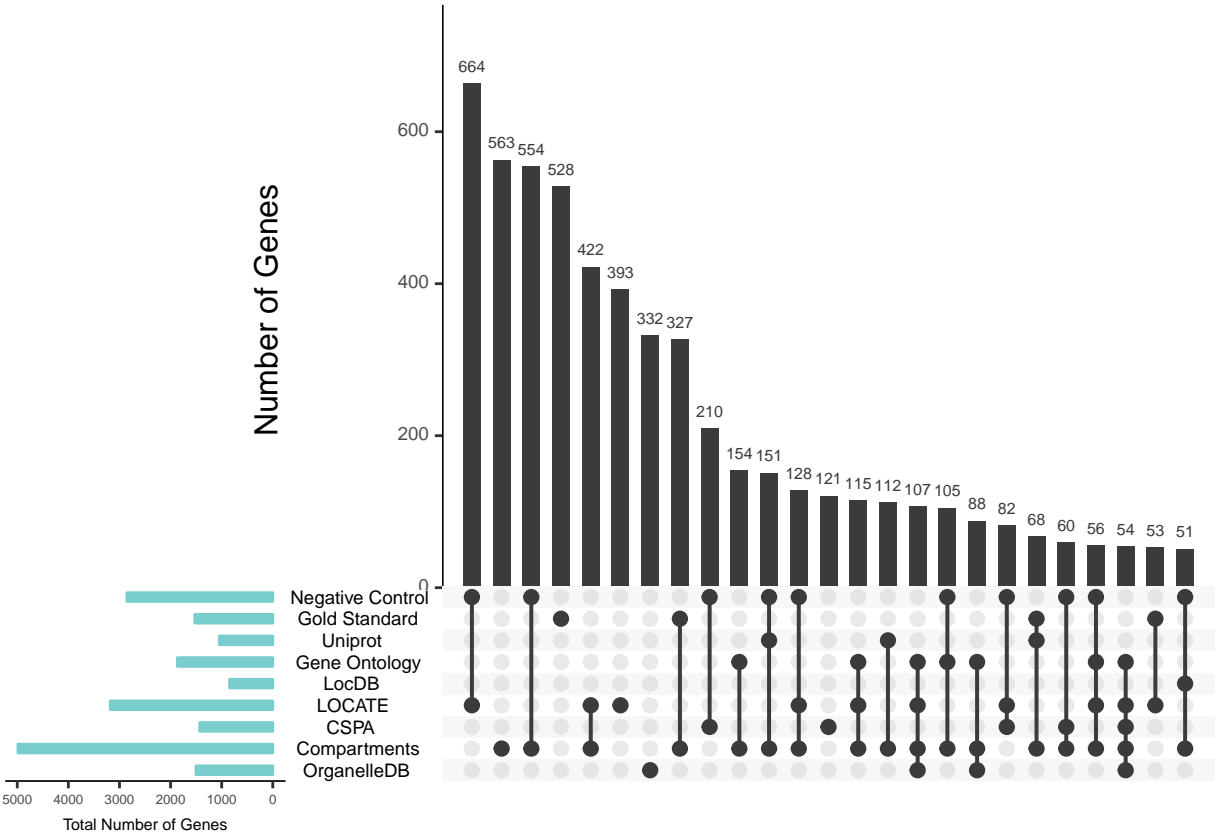
Paired bar plot depicts the number of results (secreted/transmembrane) provided by each source.



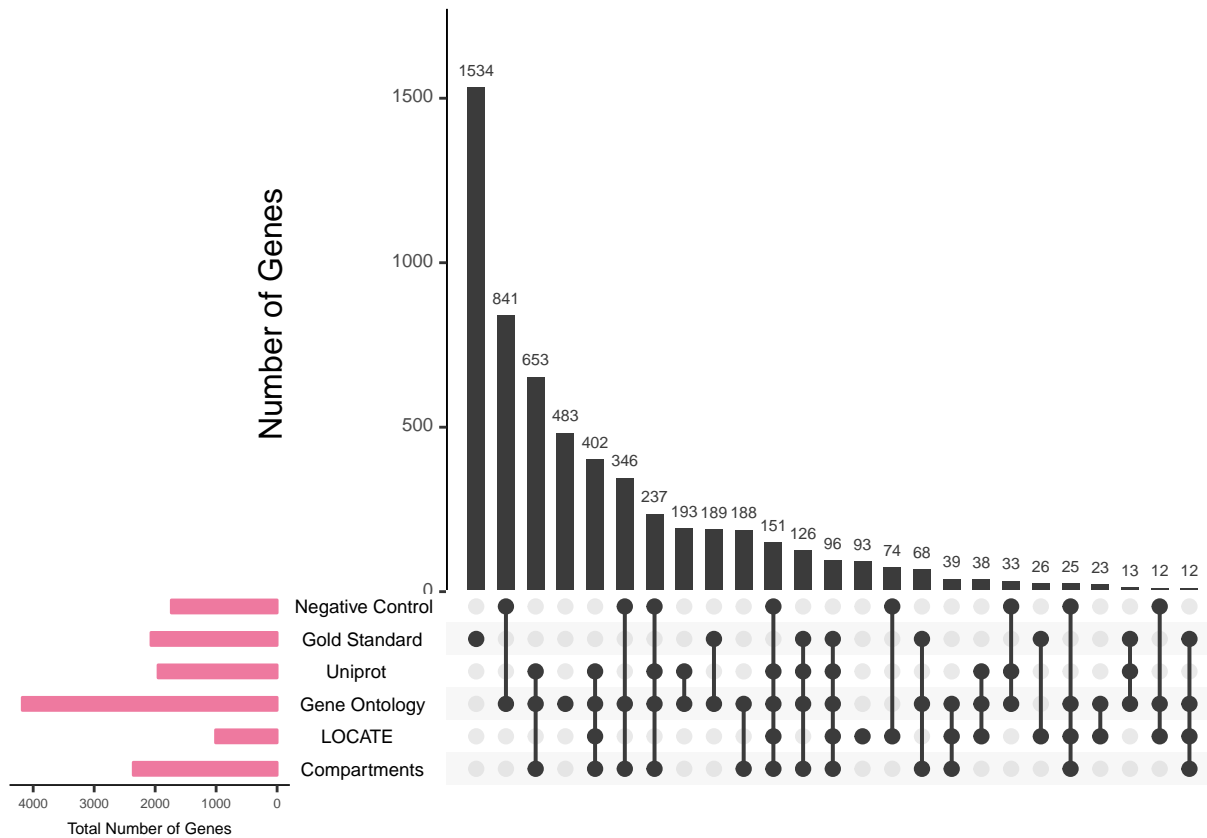
## Upset Plots as a Comparison Figure

Upset plots are used to compare all the overlapping proteins between sources.

“Transmembrane”



“Secreted”



## Sum Scores to Produce Final Lists

Sum the scores from each source (including “gold standard”, excluding “negative control”), and only include those which occur in at least half of the sources, to produce the final lists. Remove the other columns except the number of sources the protein occurred in and the hgnc\_symbol.

## Warning: package 'bindrcpp' was built under R version 3.4.4

```
## HGNC Symbol Number of Databases Confirmed
## 1 ABCB11 5
## 2 ABCB4 6
## 3 ABCC1 6
## 4 ABCC2 6
## 5 ABCC3 5
## 6 ABCC6 4

## HGNC Symbol Number of Databases Confirmed
## 1 FUCA2 4
## 2 SEMA3F 3
## 3 HSPB6 3
## 4 CCL26 3
## 5 PRSS21 3
## 6 CX3CL1 3
```