

# GSEA

James Adler

9/8/2021

## 1. Packages and files

Files:

- /projects/bgmp/shared/Bi623/Assignment\_GSEA/CompCoag\_pouch\_multivar.tsv
- /projects/bgmp/shared/Bi623/Assignment\_GSEA/pouch\_RNAseq.tsv
- /projects/bgmp/shared/Bi623/Assignment\_GSEA/pouch\_TMM\_values.csv
- /projects/bgmp/shared/Bi623/Assignment\_GSEA/stickleback\_CPM.tsv

```
# load in kegg_pouch
kegg_pouch = read.delim("pouch_RNAseq.tsv", sep="\t", stringsAsFactors=FALSE)

head(kegg_pouch)
```

## 2. Perform GSEA and KEGG pathways using gage

```
# load data
data(kegg.sets.ko)
data(sigmet.idx.ko)
kegg.sets.ko = kegg.sets.ko[sigmet.idx.ko]
head(kegg.sets.ko, 3)

# check class
class(kegg.sets.ko)
```

What class of object is kegg.sets.ko, and what kind of information does it contain?

**Class:** List

**Contents:** Many different pathways and their associated IDs.

```
# important variables
pouch_foldchanges = kegg_pouch$logFC

# assign kegg ids if present
names(pouch_foldchanges) = kegg_pouch$ko_ID

head(pouch_foldchanges)
```

```

# test for enrichment genes with extreme values
pouch_test = gage(pouch_foldchanges, gsets=kegg.sets.ko,
                  same.dir=FALSE)

# look at top entries
lapply(pouch_test, head)
head(pouch_test$greater, 30)

# subset for FDR controlled at 0.1
str(pouch_test$greater)
pouch_test.greater.01 = pouch_test$greater[which(pouch_test$greater[, 'q.val'] < 0.1),] # subset function
head(pouch_test.greater.01)

```

Which KEGG pathways are enriched for genes with exceptional pregnancy-specific gene expression?

- Cytokine-cytokine receptor interaction
- Neuroactive ligand-receptor interaction
- Neuroactive ligand-receptor interaction
- Complement and coagulation cascades

Which two of these are related to the immune system?