Cell_cycle_down_gnes

William_Salvidge

20/03/2017

```
library(knitr)
opts_chunk$set(tidy.opts=list(width.cutoff=60),tidy=TRUE)
```

```
# cell_cycle_peaks is in a email from Nicole on gmail date
# 11-01-2017

library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
cell_cycle_peaks = read.table("/Users/Will/Downloads/cell_cycle_peaks",
    header = T)

set1_down_357_veg = read.table("/Users/Will/Desktop/Set1/SET1_RNA-seq_1/Analysis/Developmental_timecours
    header = F)

colnames(cell_cycle_peaks)[1] <- "gene_ID"

colnames(set1_down_357_veg) <- "gene_ID"

cell_cycle_cluster_1 = filter(cell_cycle_peaks, cluster == "X1")
cell_cycle_cluster_2 = filter(cell_cycle_peaks, cluster == "X2")
cell_cycle_cluster_3 = filter(cell_cycle_peaks, cluster == "X3")
cell_cycle_cluster_4 = filter(cell_cycle_peaks, cluster == "X4")
cell_cycle_cluster_5 = filter(cell_cycle_peaks, cluster == "X5")
cell_cycle_cluster_6 = filter(cell_cycle_peaks, cluster == "X6")
cell_cycle_cluster_7 = filter(cell_cycle_peaks, cluster == "X7")
cell_cycle_cluster_8 = filter(cell_cycle_peaks, cluster == "X8")
cell_cycle_cluster_9 = filter(cell_cycle_peaks, cluster == "X9")

cell_cycle_cluster_1 = filter(cell_cycle_cluster_1, tp == "X1")
cell_cycle_cluster_2 = filter(cell_cycle_cluster_2, tp == "X2")
cell_cycle_cluster_3 = filter(cell_cycle_cluster_3, tp == "X3")
cell_cycle_cluster_4 = filter(cell_cycle_cluster_4, tp == "X4")
cell_cycle_cluster_5 = filter(cell_cycle_cluster_5, tp == "X5")
cell_cycle_cluster_6 = filter(cell_cycle_cluster_6, tp == "X6")
```

```
cell_cycle_cluster_7 = filter(cell_cycle_cluster_7, tp == "X7")
cell_cycle_cluster_8 = filter(cell_cycle_cluster_8, tp == "X8")
cell_cycle_cluster_9 = filter(cell_cycle_cluster_9, tp == "X9")

# cluster1 = 916 etc... 916 + 315 + 118 + 304 + 122 + 271 +
# 378 + 235 + 213 = 2872

cell_cycle_peaks_tp1 = filter(cell_cycle_peaks, tp == "X1")

set1_cell_cycle_peaks = filter(cell_cycle_peaks_tp1, cell_cycle_peaks_tp1$gene_ID %in%
    set1_down_357_veg$gene_ID)

set1_cell_cycle_cluster_1 = filter(set1_cell_cycle_peaks, cluster ==
    "X1")
set1_cell_cycle_cluster_2 = filter(set1_cell_cycle_peaks, cluster ==
    "X2")
set1_cell_cycle_cluster_3 = filter(set1_cell_cycle_peaks, cluster ==
    "X3")
set1_cell_cycle_cluster_4 = filter(set1_cell_cycle_peaks, cluster ==
    "X4")
set1_cell_cycle_cluster_5 = filter(set1_cell_cycle_peaks, cluster ==
    "X5")
set1_cell_cycle_cluster_5 = filter(set1_cell_cycle_peaks, cluster ==
    "X5")
set1_cell_cycle_cluster_6 = filter(set1_cell_cycle_peaks, cluster ==
    "X6")
set1_cell_cycle_cluster_7 = filter(set1_cell_cycle_peaks, cluster ==
    "X7")
set1_cell_cycle_cluster_8 = filter(set1_cell_cycle_peaks, cluster ==
    "X8")
set1_cell_cycle_cluster_9 = filter(set1_cell_cycle_peaks, cluster ==
    "X9")

# Test if overall there are more set1 genes with cell cycle
# regulation than we might expect. Need to ask Nicole what X
# is (156, 2872, (X - 2872), 725)

# None of these hypergeometric tests gives an
# over-representation of set1 genes in any of the clusters.
# Except for cluster 1 which Nicole says she has doubts about
# set1 cluster 1 0.0198839

phyper(50, 916, (2872 - 916), 125, lower.tail = F)
```

```
## [1] 0.0198839
```

```
# set1 cluster 2 0.1348028

phyper(17, 315, (2872 - 315), 125, lower.tail = F)
```

```
## [1] 0.1348028
```

```r
# set1 cluster 3 0.5913242

phyper(4, 118, (2872 - 118), 125, lower.tail = F)
```

```
## [1] 0.5913242
```

```r
# set1 cluster 4 0.9981732

phyper(4, 304, (2872 - 304), 125, lower.tail = F)
```

```
## [1] 0.9981732
```

```r
# set1 cluster 5 0.4404463

phyper(5, 122, (2872 - 122), 125, lower.tail = F)
```

```
## [1] 0.4404463
```

```r
# set1 cluster 6 0.2874349

phyper(13, 271, (2872 - 271), 125, lower.tail = F)
```

```
## [1] 0.2874349
```

```r
# set1 cluster 7 0.8588959

phyper(12, 378, (2872 - 378), 125, lower.tail = F)
```

```
## [1] 0.8588959
```

```r
# set1 cluster 8 0.5799583

phyper(9, 235, (2872 - 235), 125, lower.tail = F)
```

```
## [1] 0.5799583
```

```r
# set1 cluster 9 0.2127731

phyper(11, 213, (2872 - 213), 125, lower.tail = F)
```

```
## [1] 0.2127731
```