ribosome_heatmaps_bulk_v_floating

Sarah Hp

2022-07-08

Showing bulk -> floating for donor 7 and 13. Rationale: to show the trend in differentiation -> day15 is because of the adipocytes in the population.

Maintain from previous heatmaps: printing lists of gene to annotate the terms/heatmaps

```
library(biomaRt)
library(ComplexHeatmap)
```

```
## Loading required package: grid
```

```
## ========================================
## ComplexHeatmap version 2.12.0
## Bioconductor page: http://bioconductor.org/packages/ComplexHeatmap/
## Github page: https://github.com/jokergoo/ComplexHeatmap
## Documentation: http://jokergoo.github.io/ComplexHeatmap-reference
##
## If you use it in published research, please cite:
## Gu, Z. Complex heatmaps reveal patterns and correlations in multidimensional
##   genomic data. Bioinformatics 2016.
##
## The new InteractiveComplexHeatmap package can directly export static
## complex heatmaps into an interactive Shiny app with zero effort. Have a try!
##
## This message can be suppressed by:
##   suppressPackageStartupMessages(library(ComplexHeatmap))
## ========================================
```

```
knitr::opts_chunk$set(echo = TRUE, dev = c("pdf"), fig.path = "ribosome_heatmaps_bulk_v_floating/", fig
```

```
rpkm = read.delim( "../03limma/adipogenesis_rpkm_tmm_means.tab", header=T) #rpkm table
head(rpkm); dim(rpkm)
```

```
##            Geneid Length gene_name
## 1 ENSG00000000003   4535    TSPAN6
## 2 ENSG00000000005   1610      TNMD
## 3 ENSG00000000419   1207      DPM1
## 4 ENSG00000000457   6883     SCYL3
## 5 ENSG00000000460   5967  C1orf112
## 6 ENSG00000000938   3474       FGR
##                                                 description day.2.D1G.bulk
```

```
## 1                                                tetraspanin 6     3.601999275
## 2                                                  tenomodulin     0.007936845
## 3 dolichyl-phosphate mannosyltransferase subunit 1, catalytic    62.130005013
## 4                                   SCY1 like pseudokinase 3     1.426204636
## 5                          chromosome 1 open reading frame 112     1.906747475
## 6              FGR proto-oncogene, Src family tyrosine kinase     0.005503088
##   day.2.D2A.bulk day0.D1G.bulk day0.D2A.bulk day1.D1G.bulk day1.D2A.bulk
## 1     6.76071437    2.69543757    6.28332533    2.21416364    6.40180026
## 2     0.00000000    0.00000000    0.02464032    0.03748234    0.25958703
## 3    41.18788108   46.57260499   38.24198388   62.18075805   44.14029328
## 4     1.33753816    1.58269525    1.83452213    1.67277113    1.78269269
## 5     1.65219243    0.36937923    0.56555920    0.35661924    0.73777550
## 6     0.01975517    0.06610962    0.03066300    0.02233808    0.04456323
##   day15.D1G.bulk day15.D1G.floating day15.D2A.bulk day15.D2A.floating
## 1      9.1422795         15.7180139      8.3310883         15.8591145
## 2      0.5985436          0.5635616      1.9493616          2.5943861
## 3     35.7545717         30.0118693     36.1959969         43.8454420
## 4      1.7310483          1.5638762      1.6389382          1.9744310
## 5      0.5774263          0.8213003      0.4421079          0.7067799
## 6      0.6295882          3.7540597      1.9196028          3.2901662
##   day3.D1G.bulk day3.D2A.bulk day9.D1G.bulk day9.D2A.bulk
## 1    2.76755264     7.6279970     2.3380573     8.8221868
## 2    0.06034418     0.1242015     0.3061671     2.7625195
## 3   56.03395361    46.3859709    49.9509805    45.7146885
## 4    1.98222151     1.9524608     2.0051649     2.0390762
## 5    0.39508980     0.5258706     0.3640591     0.6355628
## 6    0.16554146     0.1578544     0.3397660     0.3845655
```

```
## [1] 21174    18
```

```
#discard duplicate rownames
rpkm = rpkm[!duplicated(rpkm$gene_name),]
dim(rpkm)
```

```
## [1] 21131    18
```

```
rownames(rpkm) = rpkm$gene_name
rpkm$gene_name = NULL
```

Select day15s incl. floating adipocytes

```
rpkm = rpkm[grepl("day15", colnames(rpkm))]
head(rpkm)
```

```
##          day15.D1G.bulk day15.D1G.floating day15.D2A.bulk day15.D2A.floating
## TSPAN6        9.1422795         15.7180139      8.3310883         15.8591145
## TNMD          0.5985436          0.5635616      1.9493616          2.5943861
## DPM1         35.7545717         30.0118693     36.1959969         43.8454420
## SCYL3         1.7310483          1.5638762      1.6389382          1.9744310
## C1orf112      0.5774263          0.8213003      0.4421079          0.7067799
## FGR           0.6295882          3.7540597      1.9196028          3.2901662
```

Get GO terms

```
mart <- biomaRt::useMart(biomart = "ensembl",
  dataset = "hsapiens_gene_ensembl",
  host = "https://jan2019.archive.ensembl.org")
```

Ribosome may not be necessary

```
cyt_ribosome = biomaRt::getBM(c("external_gene_name", "ensembl_gene_id","go_linkage_type"),
               filters = "go",
               values = c("GO:0022625","GO:0022627"),
               mart = mart)
length(unique(cyt_ribosome$ensembl_gene_id)) #120 cytosolic ribosome genes
```

```
## [1] 120
```

```
ribogen = biomaRt::getBM(c("external_gene_name", "ensembl_gene_id","go_linkage_type"),
               filters = "go",
               values = c("GO:0042254"),
               mart = mart)
length(unique(ribogen$ensembl_gene_id)) #104
```

```
## [1] 104
```

```
norp_ribogen = ribogen[!grepl("^M?RP", ribogen$external_gene_name),]
length(unique(norp_ribogen$ensembl_gene_id)) #96
```

```
## [1] 96
```

```
translation = getBM(c("external_gene_name", "ensembl_gene_id"),
               filters = "go",
               values = "GO:0006412",
               mart = mart)
nrow(translation)
```

```
## [1] 378
```

```
norp_trans = translation[!grepl("^M?RP", translation$external_gene_name),]
nrow(norp_trans) #165 non RP translation genes
```

```
## [1] 213
```

Combined graph

heatmap formatting

```
colnames(rpkm) = gsub("day15.","", colnames(rpkm))
summary(rpkm)
```

```
##     D1G.bulk          D1G.floating         D2A.bulk          D2A.floating
##  Min.   :    0.000   Min.   :    0.00   Min.   :    0.000   Min.   :    0.000
##  1st Qu.:    0.166   1st Qu.:    0.13   1st Qu.:    0.179   1st Qu.:    0.145
##  Median :    1.180   Median :    1.14   Median :    1.228   Median :    1.153
##  Mean   :   15.932   Mean   :   29.40   Mean   :   16.354   Mean   :   21.881
##  3rd Qu.:    5.244   3rd Qu.:    5.83   3rd Qu.:    5.280   3rd Qu.:    5.683
##  Max.   :15166.018   Max.   :42799.40   Max.   :15948.458   Max.   :29147.912
```

```
nrow(rpkm) #removing genes with no expression between these time points (about 82 genes)
```

```
## [1] 21131
```

```
rpkm = rpkm[rowSums(rpkm) > 0,]
nrow(rpkm)
```

```
## [1] 21042
```

```
#create zscores of log transformed rpkms
zs = t(scale(t(log2(rpkm+1))))
 # before log scaling the median is always about -0.2, meaning theres a negative bias to the values
summary(zs) # the median is more variable between timepoints upon log scaling
```

```
##     D1G.bulk          D1G.floating          D2A.bulk          D2A.floating
##  Min.   :-1.49970   Min.   :-1.499941   Min.   :-1.49980   Min.   :-1.49861
##  1st Qu.:-0.79702   1st Qu.:-1.039466   1st Qu.:-0.54742   1st Qu.:-0.57729
##  Median :-0.14036   Median :-0.147745   Median :-0.09195   Median : 0.02641
##  Mean   :-0.02055   Mean   :-0.006961   Mean   :-0.03306   Mean   : 0.06058
##  3rd Qu.: 0.78628   3rd Qu.: 1.140106   3rd Qu.: 0.45003   3rd Qu.: 0.71552
##  Max.   : 1.50000   Max.   : 1.500000   Max.   : 1.50000   Max.   : 1.50000
```

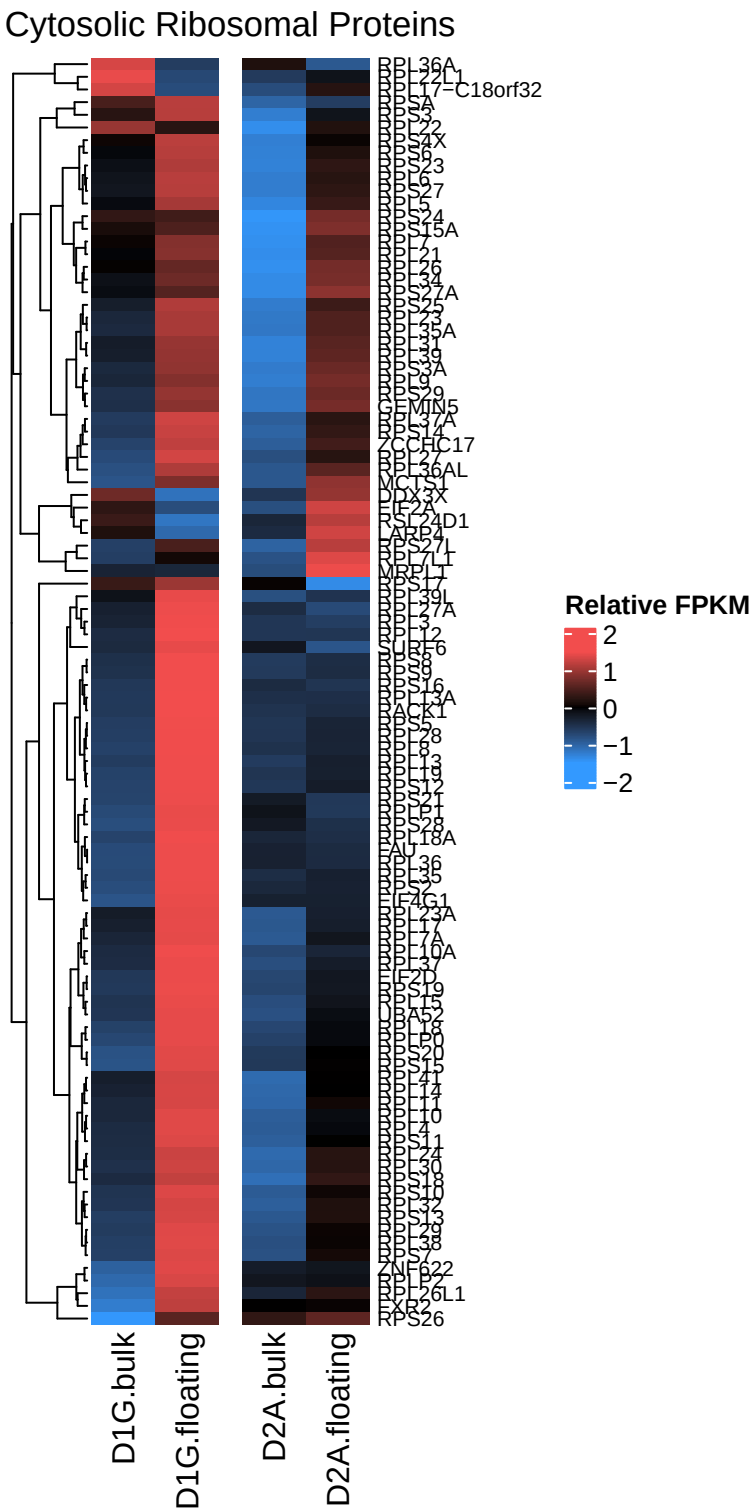
```
#create vector for ordering
#order = as.vector(zs[,"day15"] - zs[,"day.2"])
order = apply(zs, 1, cor, y=1:ncol(zs)) #uses pearson correlation
head(order)
```

```
##      TSPAN6        TNMD        DPM1       SCYL3    C1orf112         FGR
## 0.383526243 0.931291453 0.666902097 0.571797813 0.003836649 0.643553999
```

Supp Figure 6b

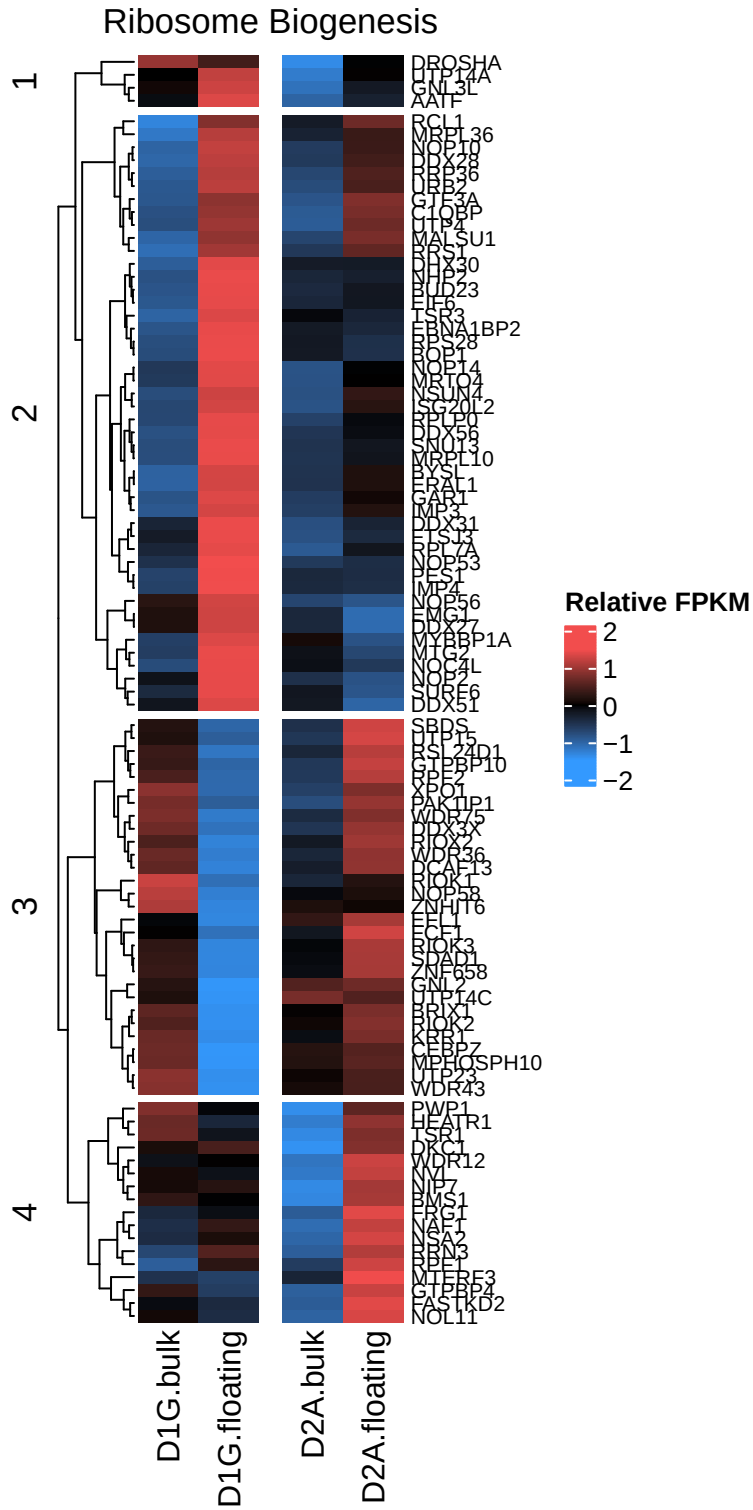
```
Heatmap(as.matrix(zs[rownames(zs) %in% cyt_ribosome$external_gene_name,]),
        cluster_columns = F,
        name="Relative FPKM",
        column_title = "Cytosolic Ribosomal Proteins",
```

```
        col=circlize::colorRamp2(c(-1.5,0,1.5),c(rgb(0.2,0.6,1),rgb(0,0,0),rgb(0.95,0.3,0.3))),
        show_row_names = T, row_names_gp = gpar(fontsize=8),
        row_dend_reorder = order[rownames(zs) %in% cyt_ribosome$external_gene_name],
        column_split = c("D07","D07","D13","D13"), column_gap = unit(3, "mm"))
```



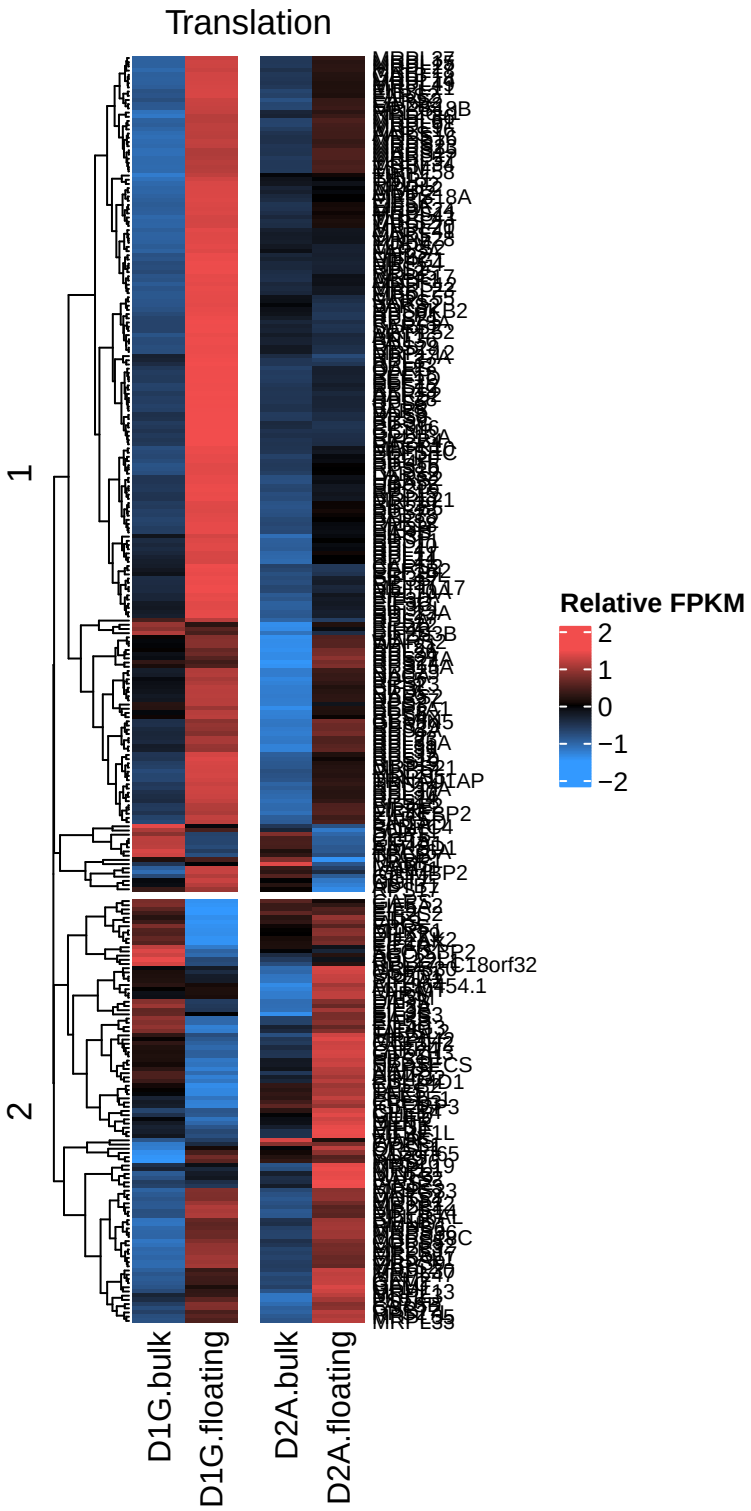
Other heatmaps

```
Heatmap(as.matrix(zs[rownames(zs) %in% ribogen$external_gene_name,]),
        cluster_columns = F,
        name="Relative FPKM",
        column_title = "Ribosome Biogenesis",
        col=circlize::colorRamp2(c(-1.5,0,1.5),c(rgb(0.2,0.6,1),rgb(0,0,0),rgb(0.95,0.3,0.3))),
        show_row_names = T, row_names_gp = gpar(fontsize=8),
        row_dend_reorder = order[rownames(zs) %in% ribogen$external_gene_name],
        row_split=4, column_split = c("D07","D07","D13","D13"), column_gap = unit(3, "mm"),
        )
```

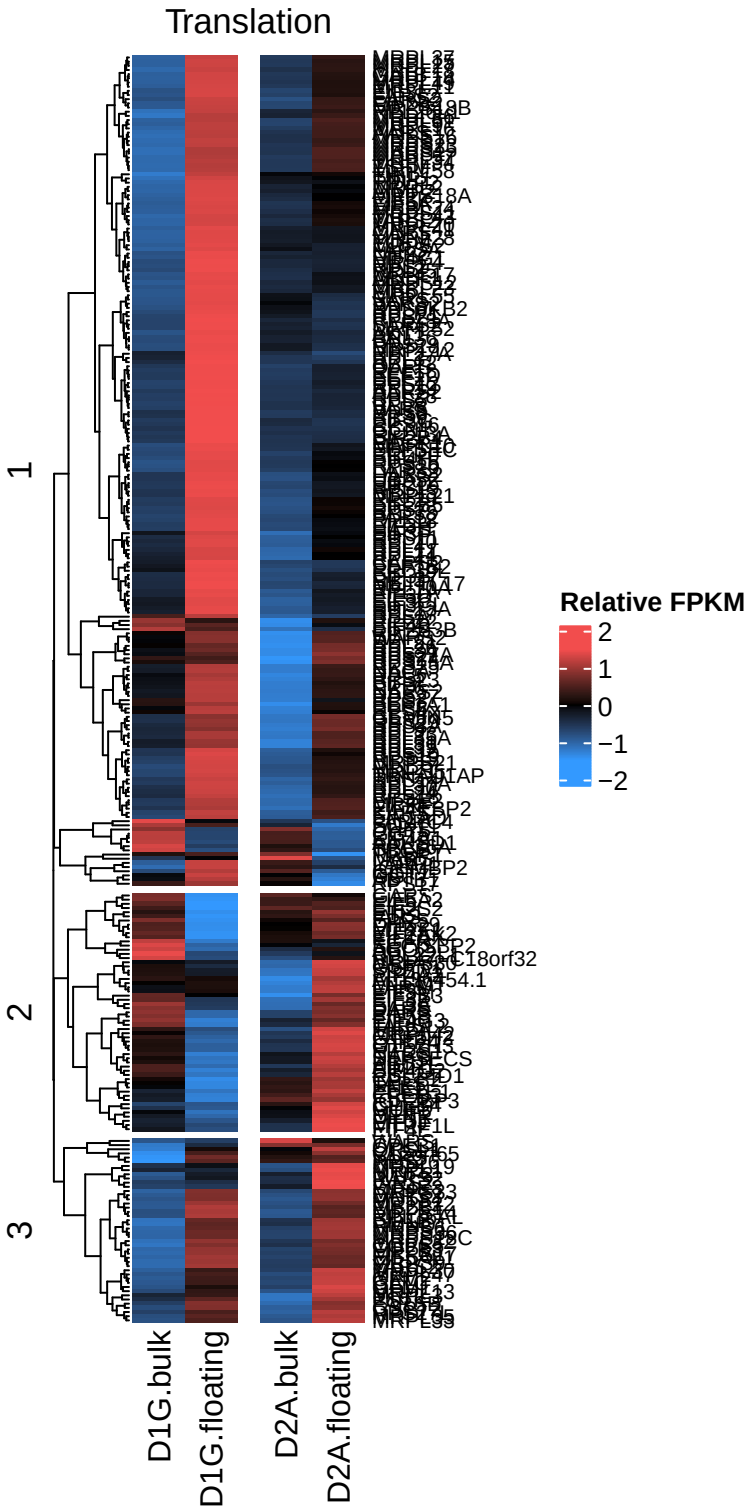



```
Heatmap(as.matrix(zs[rownames(zs) %in% translation$external_gene_name,]),
        cluster_columns = F,
        name="Relative FPKM",
        column_title = "Translation",
        col=circlize::colorRamp2(c(-1.5,0,1.5),c(rgb(0.2,0.6,1),rgb(0,0,0),rgb(0.95,0.3,0.3))),
```

```
        show_row_names = T, row_names_gp = gpar(fontsize=8),
        row_dend_reorder = order[rownames(zs) %in% translation$external_gene_name],
        row_split=2, column_split = c("D07","D07","D13","D13"), column_gap = unit(3, "mm"))
```



```
Heatmap(as.matrix(zs[rownames(zs) %in% translation$external_gene_name,]),
        cluster_columns = F,
        name="Relative FPKM",
        column_title = "Translation",
        col=circlize::colorRamp2(c(-1.5,0,1.5),c(rgb(0.2,0.6,1),rgb(0,0,0),rgb(0.95,0.3,0.3))),
        show_row_names = T, row_names_gp = gpar(fontsize=8),
        row_dend_reorder = order[rownames(zs) %in% translation$external_gene_name],
        row_split=3, column_split = c("D07","D07","D13","D13"), column_gap = unit(3, "mm"))
```


Translation
1
2
3
D1G.bulk
D1G.floating
D2A.bulk
D2A.floating
Relative FPKM
2
1
0
−1
−2