

ribosome_heatmaps_adipogenesis

Sarah Hp

2022-07-08

Changes:

- Using DE calculated from only 4 time points (Pro -> D0 -> D3 -> D15)

From previous revisions:

- 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_adipogenesis/", fig.dir = "figs/")
```

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

```
##           Geneid Length gene_name
## 1 ENSG000000000003   4535   TSPAN6
## 2 ENSG000000000005   1610    TNMD
## 3 ENSG000000000419   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
```

```
combpval = read.delim("../03limma/nucleolus_adipogenesis_DE.tab", header=T)
head(combpval)
```

```
## Geneid Length gene_name
## 1 ENSG00000151726 6284 ACSL1
## 2 ENSG00000099194 5362 SCD
## 3 ENSG00000042445 4005 RETSAT
## 4 ENSG00000056998 3655 GYG2
## 5 ENSG00000076555 14505 ACACB
## 6 ENSG00000101938 3920 CHRDL1
## description logFC AveExpr t
## 1 acyl-CoA synthetase long chain family member 1 12.916437 8.810891 50.67161
## 2 stearoyl-CoA desaturase 16.490127 11.807874 43.63744
## 3 retinol saturase 8.195969 7.100134 36.32129
## 4 glycogenin 2 10.028582 4.980263 36.39842
## 5 acetyl-CoA carboxylase beta 14.871196 7.351638 33.29351
## 6 chordin like 1 14.207558 5.954537 32.38451
## P.Value adj.P.Val B
## 1 8.291877e-37 1.755722e-32 68.92770
```

```
## 2 2.897205e-34 3.067271e-30 64.45829
## 3 2.309287e-31 9.779370e-28 59.33478
## 4 2.133268e-31 9.779370e-28 58.59981
## 5 7.038353e-30 2.474282e-26 55.41491
## 6 1.960845e-29 4.613214e-26 54.48242
```

```
sig = merge(rpkm, combpval[c("gene_name", "adj.P.Val")])
head(sig); dim(sig)
```

```
##      gene_name      Geneid Length
## 1      A1BG ENSG00000121410   4006
## 2 A1BG-AS1 ENSG00000268895   2793
## 3      A2M ENSG00000175899   6384
## 4 A2M-AS1 ENSG00000245105   2816
## 5     A4GALT ENSG00000128274   3407
## 6      AAAS ENSG00000094914   4138
##                                     description day.2.D1G.bulk
## 1                                     alpha-1-B glycoprotein      0.01815337
## 2                                     A1BG antisense RNA 1      0.90975838
## 3                                     alpha-2-macroglobulin    0.21349869
## 4                                     A2M antisense RNA 1      0.29287176
## 5 alpha 1,4-galactosyltransferase (P blood group)      4.67272390
## 6                                     aladin WD repeat nucleoporin 6.47412375
##      day.2.D2A.bulk day0.D1G.bulk day0.D2A.bulk day1.D1G.bulk day1.D2A.bulk
## 1      0.1047458      0.05978501      0.06987478      0.04544531      0.0815610
## 2      1.3179450      0.90779680      1.23126472      1.87323332      2.0247272
## 3      0.8046160      7.33195064      14.48690784      6.17424097      12.9542154
## 4      0.2337510      0.49255301      0.29768876      0.40976779      0.4652485
## 5      2.0682521      8.79899844      2.55743385      4.76230682      1.3129724
## 6      7.4918064      6.71469625      8.02178923      8.66851010      7.7134610
##      day15.D1G.bulk day15.D1G.floating day15.D2A.bulk day15.D2A.floating
## 1      0.03912984      0.0185444      0.03159112      0.01945774
## 2      1.27934477      1.0516270      1.39386219      0.74776720
## 3      55.90851651      39.4762646      73.55229005      69.76889731
## 4      0.80400184      1.1234145      0.59104284      0.81515811
## 5      4.78652768      9.4369909      4.47173184      2.50902053
## 6      6.76318918      13.7434453      7.46425689      6.38822347
##      day3.D1G.bulk day3.D2A.bulk day9.D1G.bulk day9.D2A.bulk      adj.P.Val
## 1      0.04925457      0.05550053      0.02259233      0.05723545 4.206849e-01
## 2      1.72406789      1.63700087      1.06522748      1.13733587 1.941349e-01
## 3      7.65875953      10.30956590      9.76357919      14.88341115 2.466237e-18
## 4      0.46771549      0.33457401      0.48169174      0.48900749 2.002885e-06
## 5      6.63361018      2.40230131      4.70604627      0.93162324 1.130921e-01
## 6      9.03170037      9.52568667      6.86448447      8.35763234 8.865056e-01
```

```
## [1] 20590      19
```

```
#formatting
sig = sig[sig$adj.P.Val < 0.01,]
sig$adj.P.Val = NULL
#discard duplicate rownames
sig = sig[!duplicated(sig$gene_name),]
dim(sig)
```

```
## [1] 11109      18
```

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

remove floating adipocytes

```
sig = sig[!grepl("floating", colnames(sig))]
head(sig)
```

##		Geneid	Length		description	day.2.D1G.bulk
##	A2M	ENSG00000175899	6384		alpha-2-macroglobulin	0.2134987
##	A2M-AS1	ENSG00000245105	2816		A2M antisense RNA 1	0.2928718
##	AACS	ENSG00000081760	16094		acetoacetyl-CoA synthetase	1.1224629
##	AADAC	ENSG00000114771	1668		arylacetamide deacetylase	0.8549193
##	AADAT	ENSG00000109576	2981		aminoadipate aminotransferase	4.3500497
##	AAK1	ENSG00000115977	23608		AP2 associated kinase 1	4.8359569
##		day.2.D2A.bulk	day0.D1G.bulk	day0.D2A.bulk	day1.D1G.bulk	day1.D2A.bulk
##	A2M	0.804616	7.3319506	14.4869078	6.1742410	12.9542154
##	A2M-AS1	0.233751	0.4925530	0.2976888	0.4097678	0.4652485
##	AACS	1.085992	1.2582284	0.9649686	1.9610903	2.5120671
##	AADAC	8.040700	0.3219137	4.1762492	0.4514961	6.0855305
##	AADAT	4.043753	5.0225282	2.9179408	6.5711193	3.3549336
##	AAK1	3.816897	3.9496572	3.2858525	2.6913469	3.0596925
##		day15.D1G.bulk	day15.D2A.bulk	day3.D1G.bulk	day3.D2A.bulk	day9.D1G.bulk
##	A2M	55.9085165	73.5522900	7.6587595	10.309566	9.7635792
##	A2M-AS1	0.8040018	0.5910428	0.4677155	0.334574	0.4816917
##	AACS	3.5082122	7.0684261	5.1619360	7.084625	6.6211387
##	AADAC	0.3446196	0.6362802	0.6398332	6.148814	0.3752928
##	AADAT	1.9088162	1.4068448	5.2608906	2.226705	2.8830540
##	AAK1	2.7325238	3.0795301	2.2682384	2.929480	2.6410048
##		day9.D2A.bulk				
##	A2M	14.8834112				
##	A2M-AS1	0.4890075				
##	AACS	5.3953233				
##	AADAC	3.6222883				
##	AADAT	1.8764380				
##	AAK1	3.4250558				

```
sig = sig[grepl("day(.2|0|3|15)", colnames(sig))]
#separate by donor
d1 = sig[grepl("D1G", colnames(sig), value=T)]
head(d1)
```

##		day.2.D1G.bulk	day0.D1G.bulk	day15.D1G.bulk	day3.D1G.bulk
##	A2M	0.2134987	7.3319506	55.9085165	7.6587595
##	A2M-AS1	0.2928718	0.4925530	0.8040018	0.4677155
##	AACS	1.1224629	1.2582284	3.5082122	5.1619360
##	AADAC	0.8549193	0.3219137	0.3446196	0.6398332
##	AADAT	4.3500497	5.0225282	1.9088162	5.2608906
##	AAK1	4.8359569	3.9496572	2.7325238	2.2682384

```
d2 = sig[grep("D2A", colnames(sig), value=T)]
head(d2)
```

```
##          day.2.D2A.bulk day0.D2A.bulk day15.D2A.bulk day3.D2A.bulk
## A2M          0.804616    14.4869078    73.5522900    10.309566
## A2M-AS1      0.233751    0.2976888     0.5910428     0.334574
## AACS         1.085992    0.9649686     7.0684261     7.084625
## AADAC        8.040700    4.1762492     0.6362802     6.148814
## AADAT        4.043753    2.9179408     1.4068448     2.226705
## AAK1         3.816897    3.2858525     3.0795301     2.929480
```

Get GO terms

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

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
```

```
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
```

```
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
```

Check the gene lists

```
summary(translation$external_gene_name %in% rownames(sig))
```

```
##      Mode   FALSE    TRUE
## logical    191    187
```

```
summary(ribogen$external_gene_name %in% rownames(sig))
```

```
##      Mode   FALSE    TRUE
## logical    333    275
```

Donor 1

heatmap formatting

```
colnames(d1) = gsub(".D1G.bulk","", colnames(d1))
d1 = d1[c("day.2", "day0", "day3", "day15")]
```

```
#create zscores of log transformed rpkm
```

```
zs = t(scale(t(log2(d1+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
```

```
##      day.2      day0      day3      day15
## Min.   :-1.4996364 Min.   :-1.49360 Min.   :-1.49784 Min.   :-1.499764
## 1st Qu.: -0.7525029 1st Qu.: -0.59169 1st Qu.: -0.64610 1st Qu.: -0.972699
## Median : 0.0000832 Median : -0.11097 Median : -0.10636 Median : -0.128307
## Mean   : 0.0670357 Mean   : -0.04112 Mean   : -0.02994 Mean   : 0.004022
## 3rd Qu.: 0.9457481 3rd Qu.: 0.48140 3rd Qu.: 0.58149 3rd Qu.: 1.097580
## Max.   : 1.5000000 Max.   : 1.49998 Max.   : 1.49928 Max.   : 1.500000
```

```
#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)
```

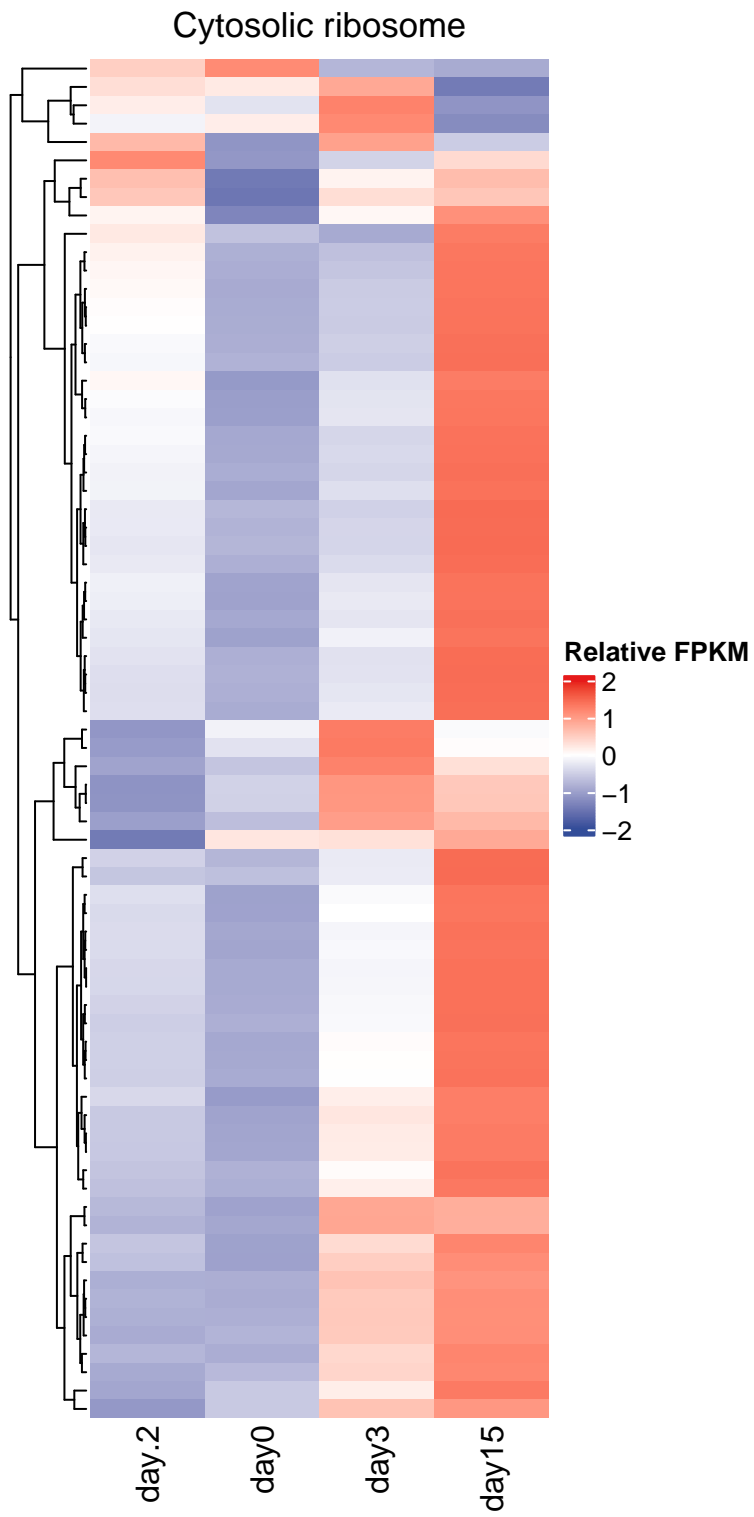
```
##      A2M      A2M-AS1      AACS      AADAC      AADAT      AAK1
## 0.9517668 0.9234575 0.8046453 -0.5930503 -0.6471557 -0.8605864
```

Cytosolic ribosome components with significantly different expression cross the timecourse.

```
summary(rownames(zs) %in% cyt_ribosome$external_gene_name) #72 genes
```

```
##      Mode  FALSE    TRUE  
## logical 11035     74
```

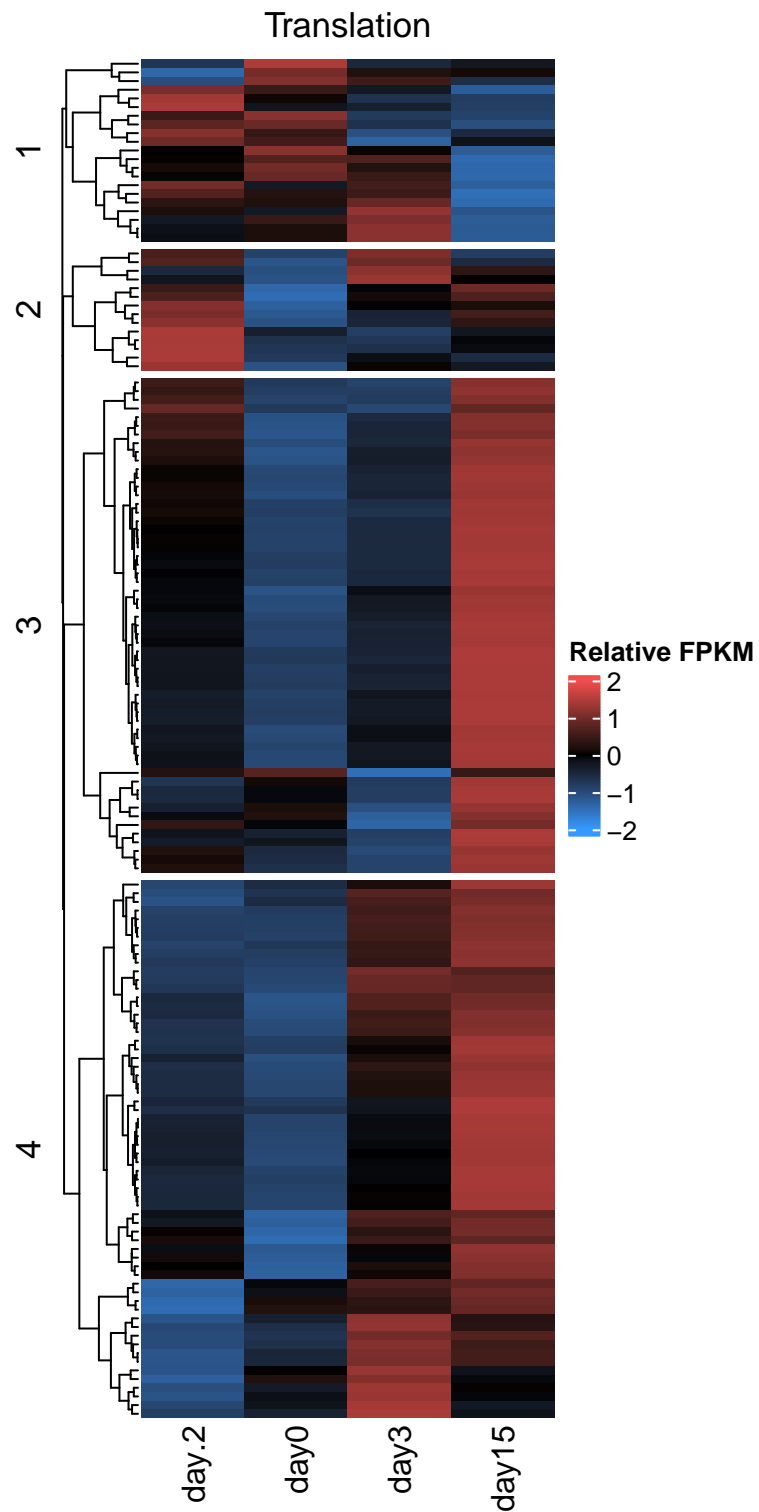
```
Heatmap(as.matrix(zs[rownames(zs) %in% cyt_ribosome$external_gene_name,]),  
        cluster_columns = F,  
        name="Relative FPKM", column_title = "Cytosolic ribosome",  
        col=circlize::colorRamp2(c(-2,0,2),c(rgb(0.2,0.3,0.6),rgb(1,1,1),rgb(0.9,0.1,0.1))),  
        row_names_gp = gpar(fontsize = 7),  
        show_row_names = FALSE,  
        row_dend_reorder = order[rownames(zs) %in% cyt_ribosome$external_gene_name])
```



Translation terms

And with splits


```
Heatmap(as.matrix(zs[rownames(zs) %in% translation$external_gene_name,]),
        cluster_columns = F,
        name="Relative FPKM",
        column_title = "Translation",
        col=circlize::colorRamp2(c(-2,0,2),c(rgb(0.2,0.6,1),rgb(0,0,0),rgb(0.95,0.3,0.3))),
        show_row_names = F, row_split=4)
```



```
Heatmap(as.matrix(zs[rownames(zs) %in% translation$external_gene_name,]),
        cluster_columns = F,
        name="Relative FPKM",
        column_title = "Translation",
        col=circlize::colorRamp2(c(-2,0,2),c(rgb(0.2,0.6,1),rgb(0,0,0),rgb(0.95,0.3,0.3))),
```

```
show_row_names = F, row_split=8)
```

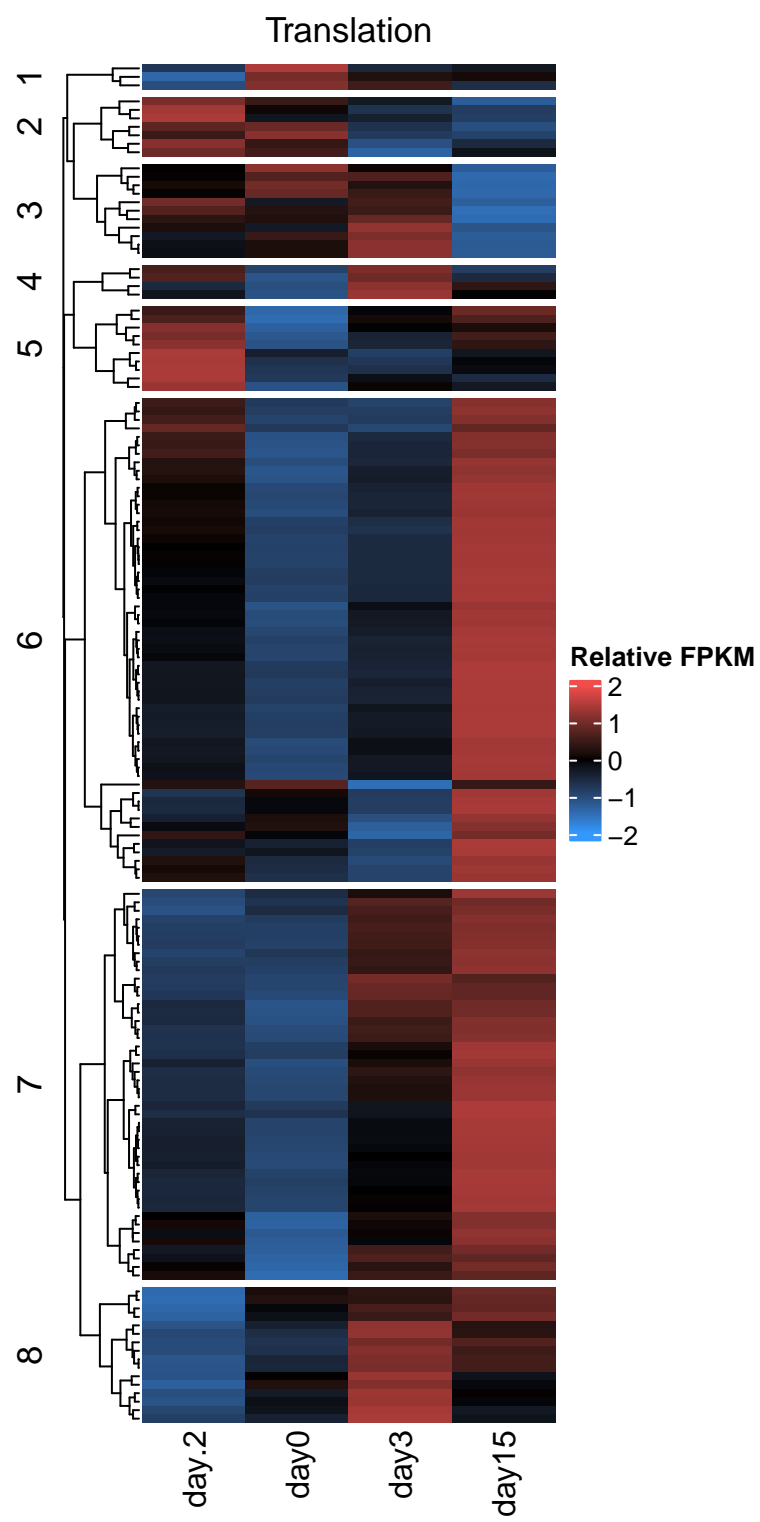
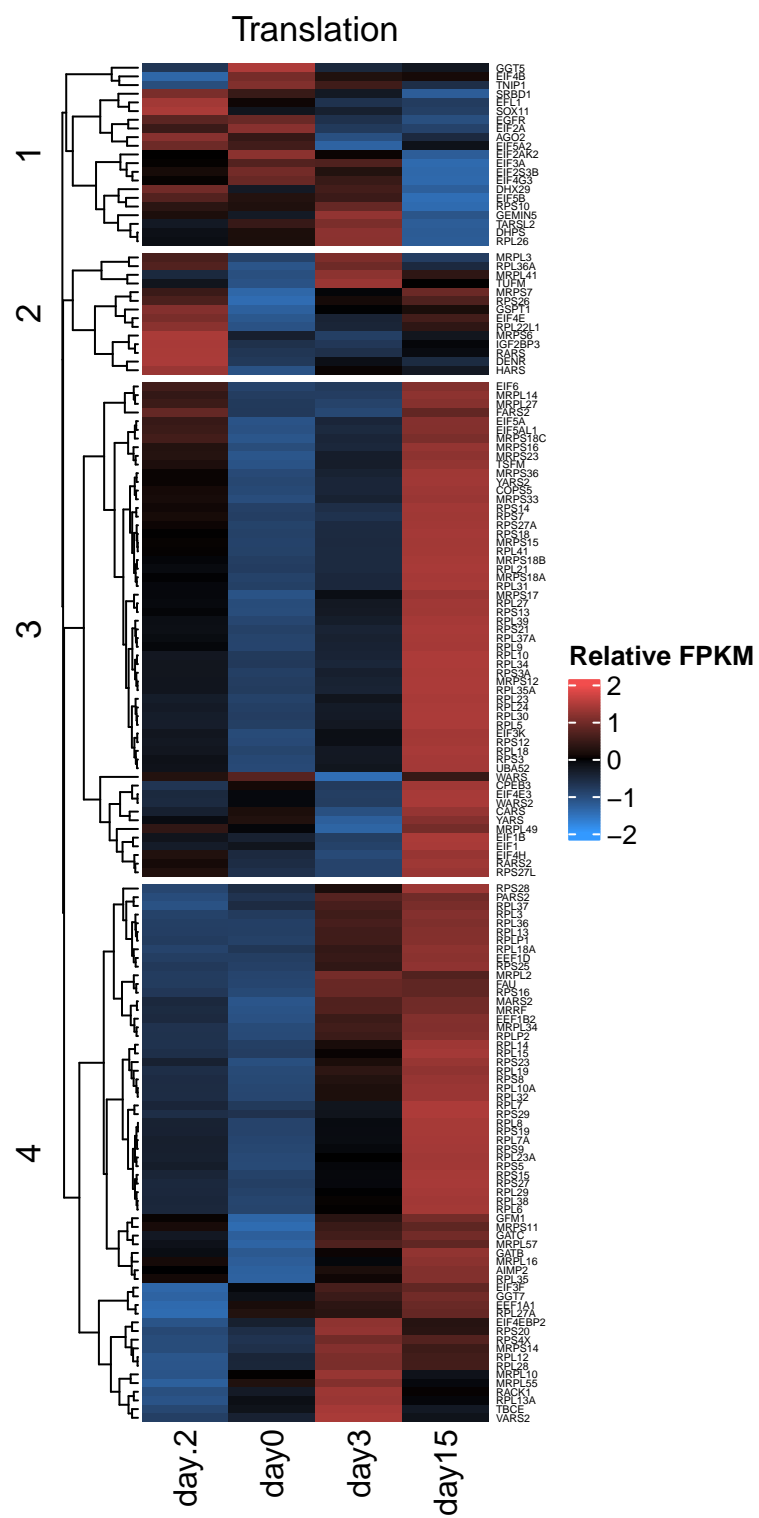


Figure3a Lower panel

```
Heatmap(as.matrix(zs[rownames(zs) %in% translation$external_gene_name,]),
        cluster_columns = F,
        name="Relative FPKM",
        column_title = "Translation",
        col=circlize::colorRamp2(c(-2,0,2),c(rgb(0.2,0.6,1),rgb(0,0,0),rgb(0.95,0.3,0.3))),
        show_row_names = T, row_split=4, row_names_gp = gpar(fontsize = 4),
        row_dend_reorder = order[rownames(zs) %in% translation$external_gene_name])
```



Ribosome biogenesis

```
Heatmap(as.matrix(zs[rownames(zs) %in% ribogen$external_gene_name,]),
        cluster_columns = F,
        name="Relative FPKM",
        column_title = "Ribosome biogenesis",
        height=0.25*length(unique(translation$external_gene_name)),
        width=0.5*ncol(d1),
        col=circlize::colorRamp2(c(-2,0,2),c(rgb(0.2,0.6,1),rgb(0,0,0),rgb(0.95,0.3,0.3))),
        row_names_gp = gpar(fontsize = 8), row_split=6,
        row_dend_reorder = order[rownames(zs) %in% ribogen$external_gene_name])
```

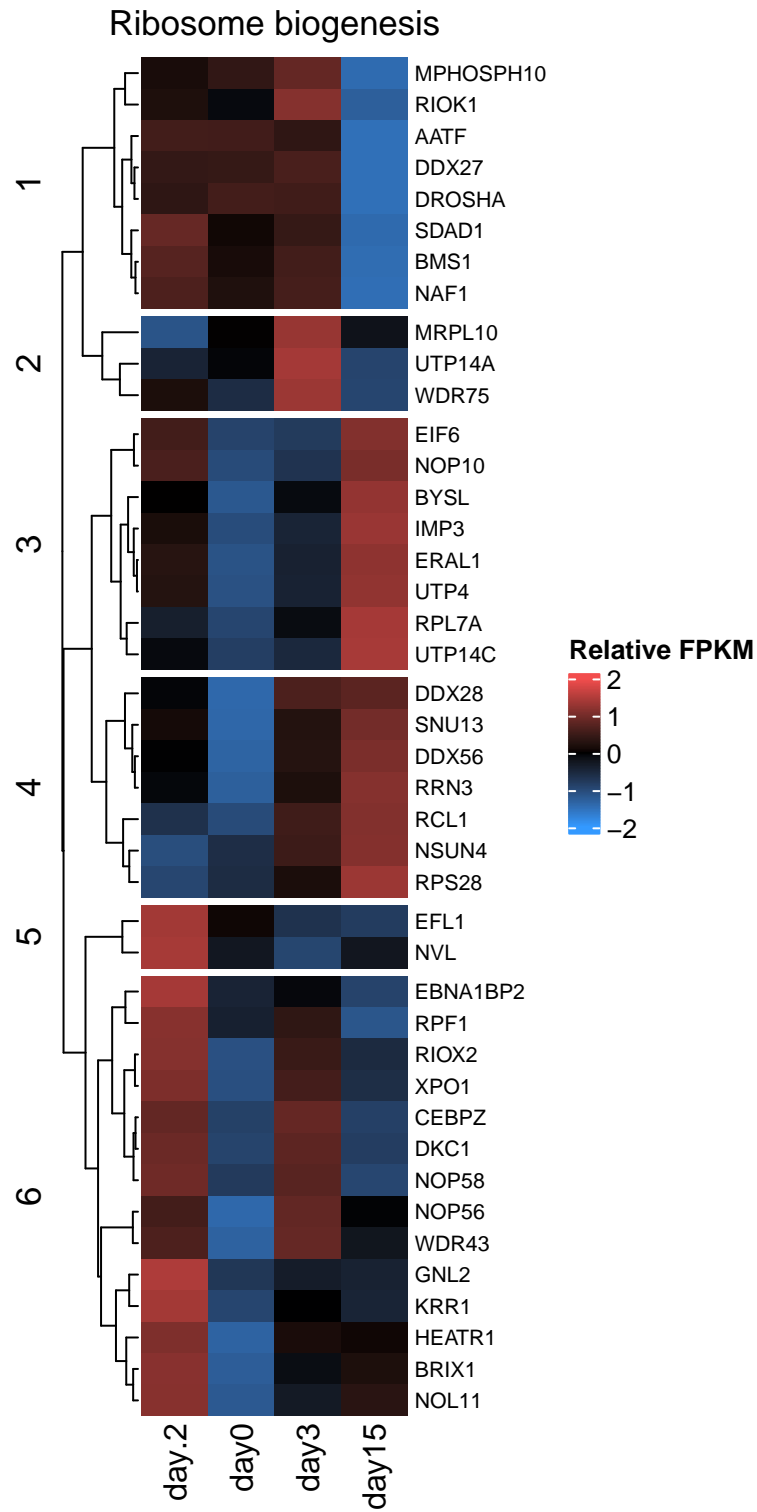
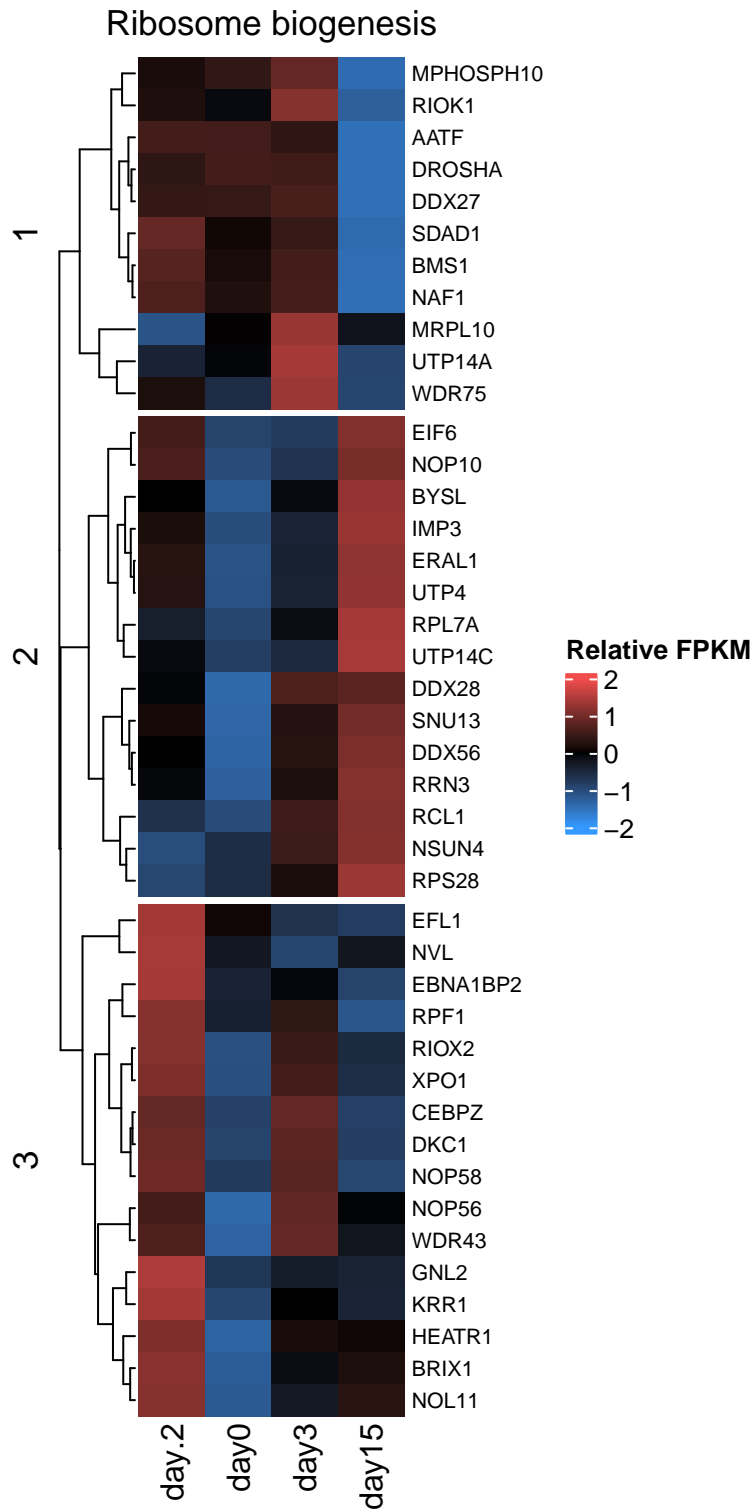


Figure3a Upper panel

```
Heatmap(as.matrix(zs[rownames(zs) %in% ribogen$external_gene_name,]),
        cluster_columns = F,
        name="Relative FPKM",
        column_title = "Ribosome biogenesis",
        height=0.25*length(unique(translation$external_gene_name)),
        width=0.5*ncol(d1),
        col=circlize::colorRamp2(c(-2,0,2),c(rgb(0.2,0.6,1),rgb(0,0,0),rgb(0.95,0.3,0.3))),
        row_names_gp = gpar(fontsize = 8), row_split=3,
        row_dend_reorder = order[rownames(zs) %in% ribogen$external_gene_name])
```

Donor 2

heatmap formatting

```
colnames(d2) = gsub(".D2A.bulk","", colnames(d2))
d2 = d2[c("day.2", "day0", "day3", "day15")]

#create zscores
zs13 = t(scale(t(log2(d2+1))))
head(zs13)
```

```
##           day.2      day0      day3      day15
## A2M      -1.2620822  0.14616894 -0.05976171  1.1756750
## A2M-AS1  -0.8671337 -0.41005001 -0.15649903  1.4336827
## AACS     -0.8282284 -0.90301406  0.86687604  0.8643664
## AADAC     0.8252467  0.08943798  0.51544871 -1.4301334
## AADAT     1.1513684  0.34304062 -0.27813901 -1.2162700
## AAK1      1.3737926  0.05460247 -0.92580611 -0.5025889
```

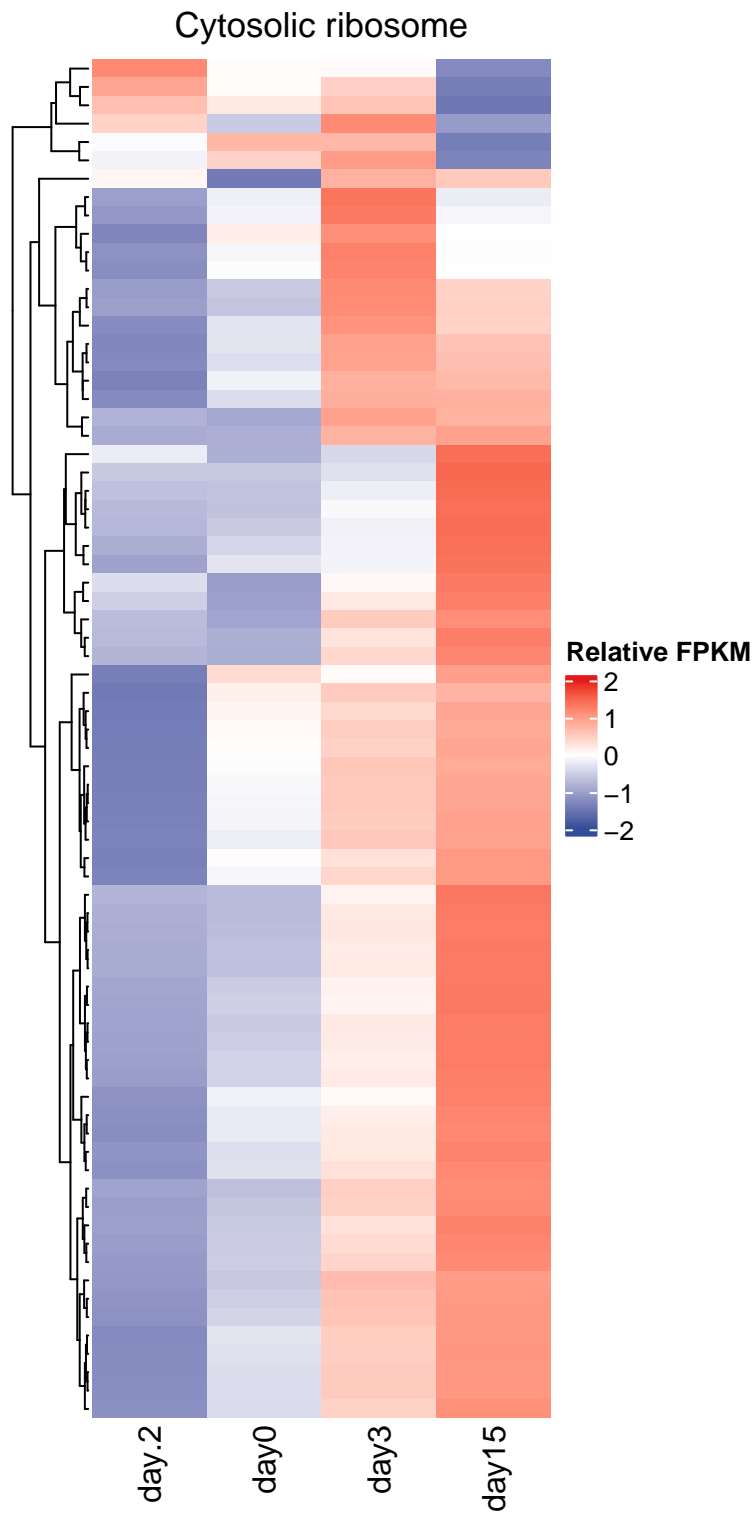
```
# create order vector
#order13 = as.vector(zs13[, "day15"] - zs13[, "day.2"])
order13 = apply(zs13, 1, cor, y=1:ncol(zs13))
```

Cytosolic ribosome components with significantly different expression cross the timecourse. Day 9 and floating adipocytes share high expression of many of these genes.

```
summary(rownames(zs13) %in% cyt_ribosome$external_gene_name) #72 genes
```

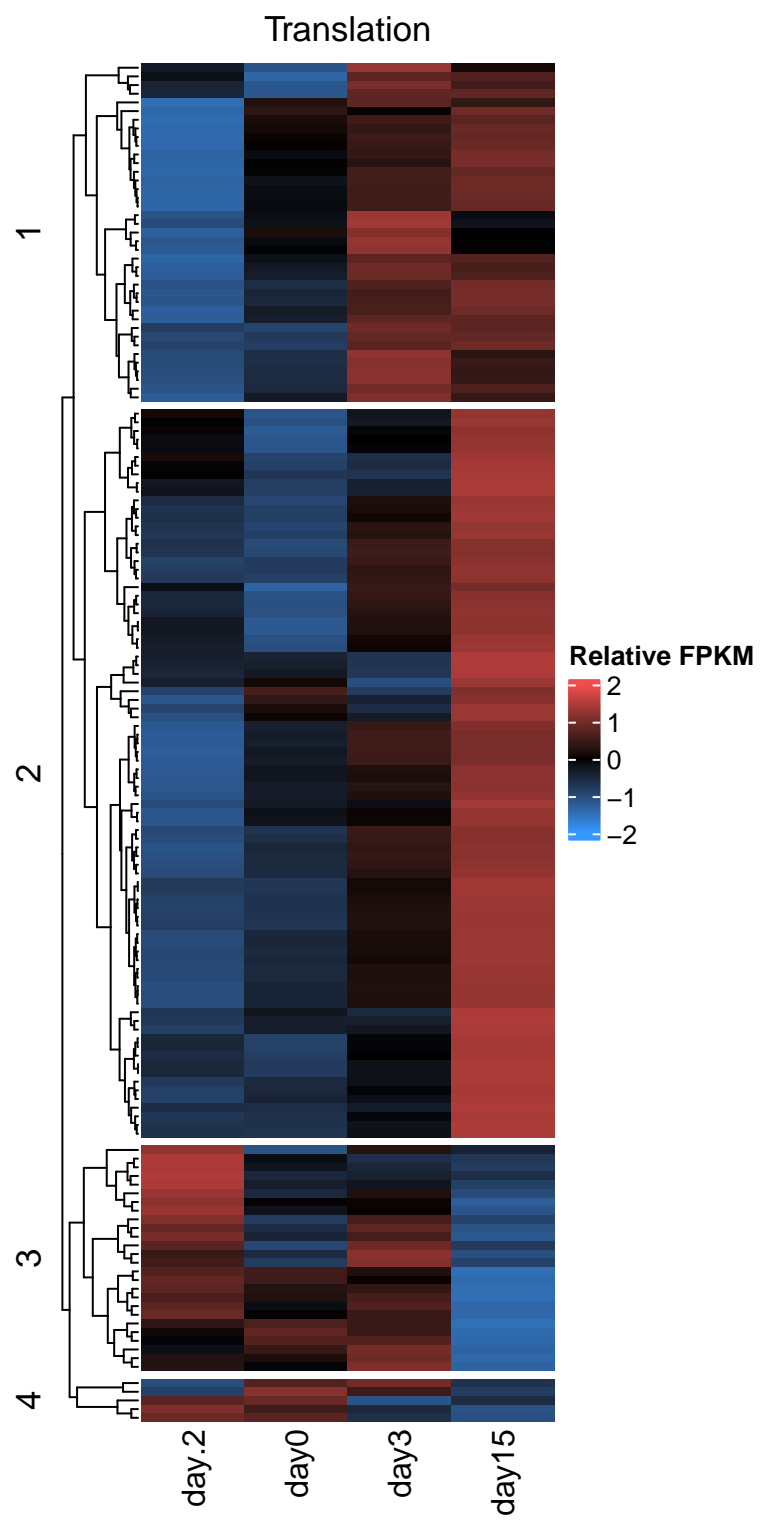
```
##      Mode   FALSE    TRUE
## logical  11035     74
```

```
Heatmap(as.matrix(zs13[rownames(zs13) %in% cyt_ribosome$external_gene_name,]),
        cluster_columns = F,
        name="Relative FPKM", column_title = "Cytosolic ribosome",
        height=0.5*length(unique(cyt_ribosome$external_gene_name)),
        width=0.5*ncol(d1),
        col=circlize::colorRamp2(c(-2,0,2),c(rgb(0.2,0.3,0.6),rgb(1,1,1),rgb(0.9,0.1,0.1))),
        row_dend_reorder = order13[rownames(zs13) %in% cyt_ribosome$external_gene_name],
        show_row_names = F)
```



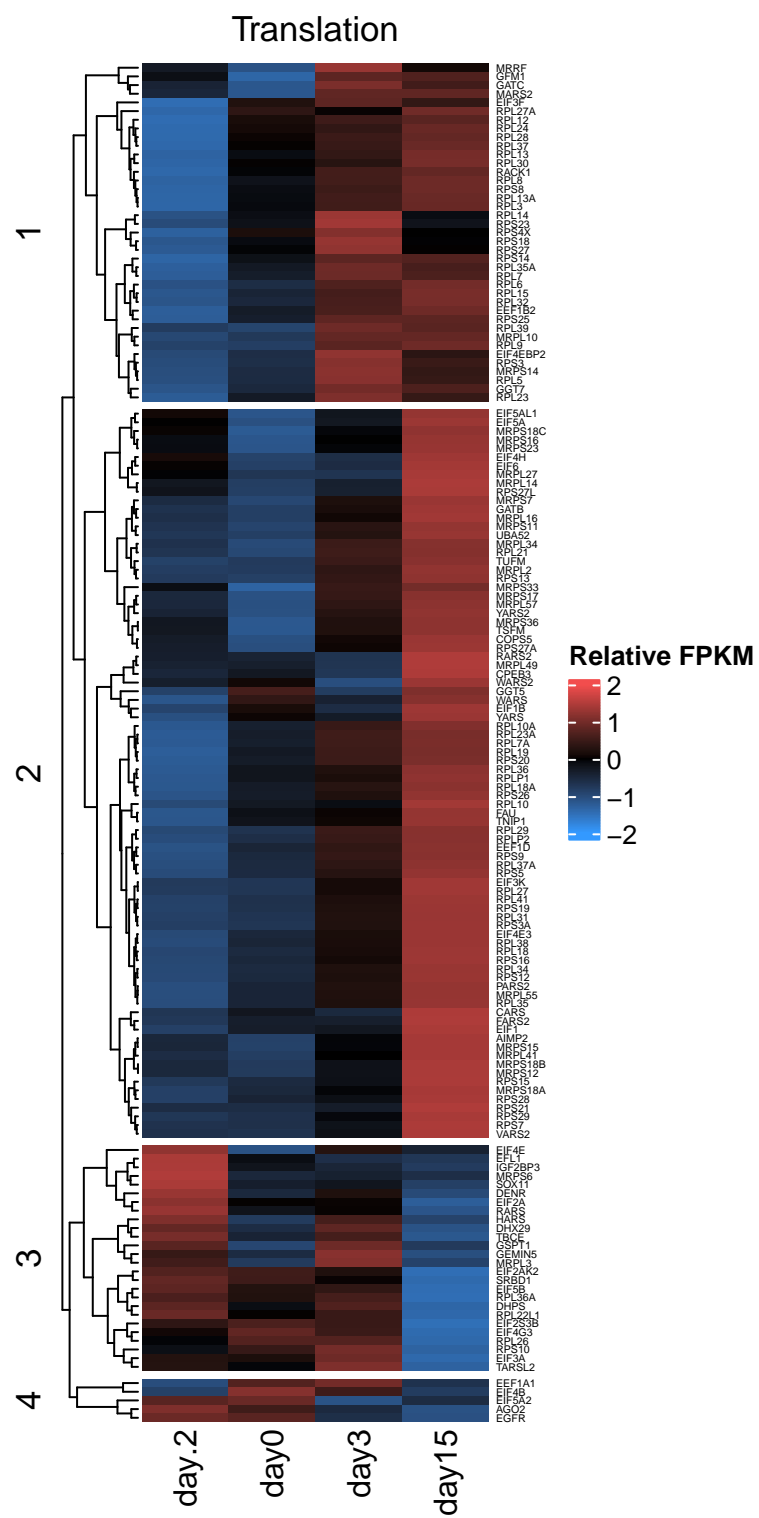
D2 Translation Heatmaps

```
Heatmap(as.matrix(zs13[rownames(zs13) %in% translation$external_gene_name,]),
        cluster_columns = F,
        name="Relative FPKM",
        column_title = "Translation",
        col=circlize::colorRamp2(c(-2,0,2),c(rgb(0.2,0.6,1),rgb(0,0,0),rgb(0.95,0.3,0.3))),
        show_row_names = F, row_split = 4,
        row_dend_reorder = order13[rownames(zs13) %in% translation$external_gene_name])
```



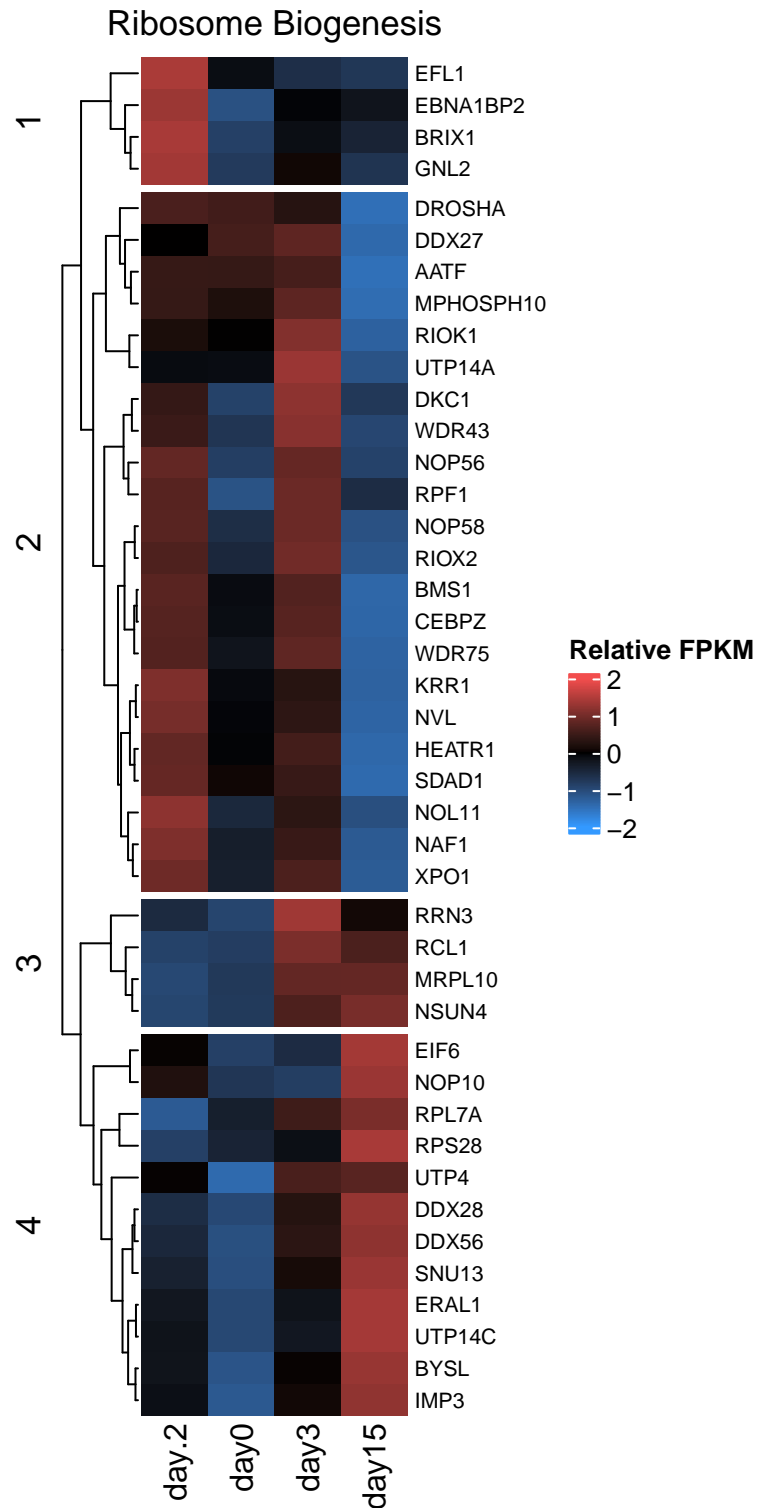
Supp Figure 6a lower panel

```
Heatmap(as.matrix(zs13[rownames(zs13) %in% translation$external_gene_name,]),
        cluster_columns = F,
        name="Relative FPKM",
        column_title = "Translation",
        col=circlize::colorRamp2(c(-2,0,2),c(rgb(0.2,0.6,1),rgb(0,0,0),rgb(0.95,0.3,0.3))) ,
        show_row_names = T, row_split = 4, row_names_gp = gpar(fontsize=4),
        row_dend_reorder = order13[rownames(zs13) %in% translation$external_gene_name])
```



D2 Ribosome Biogenesis Heatmaps

```
Heatmap(as.matrix(zs13[rownames(zs13) %in% ribogen$external_gene_name,]),
        cluster_columns = F,
        name="Relative FPKM",
        column_title = "Ribosome Biogenesis",
        col=circlize::colorRamp2(c(-2,0,2),c(rgb(0.2,0.6,1),rgb(0,0,0),rgb(0.95,0.3,0.3))),
        show_row_names = T, row_names_gp = gpar(fontsize=8),
        row_dend_reorder = order13[rownames(zs13) %in% ribogen$external_gene_name],
        row_split=4)
```

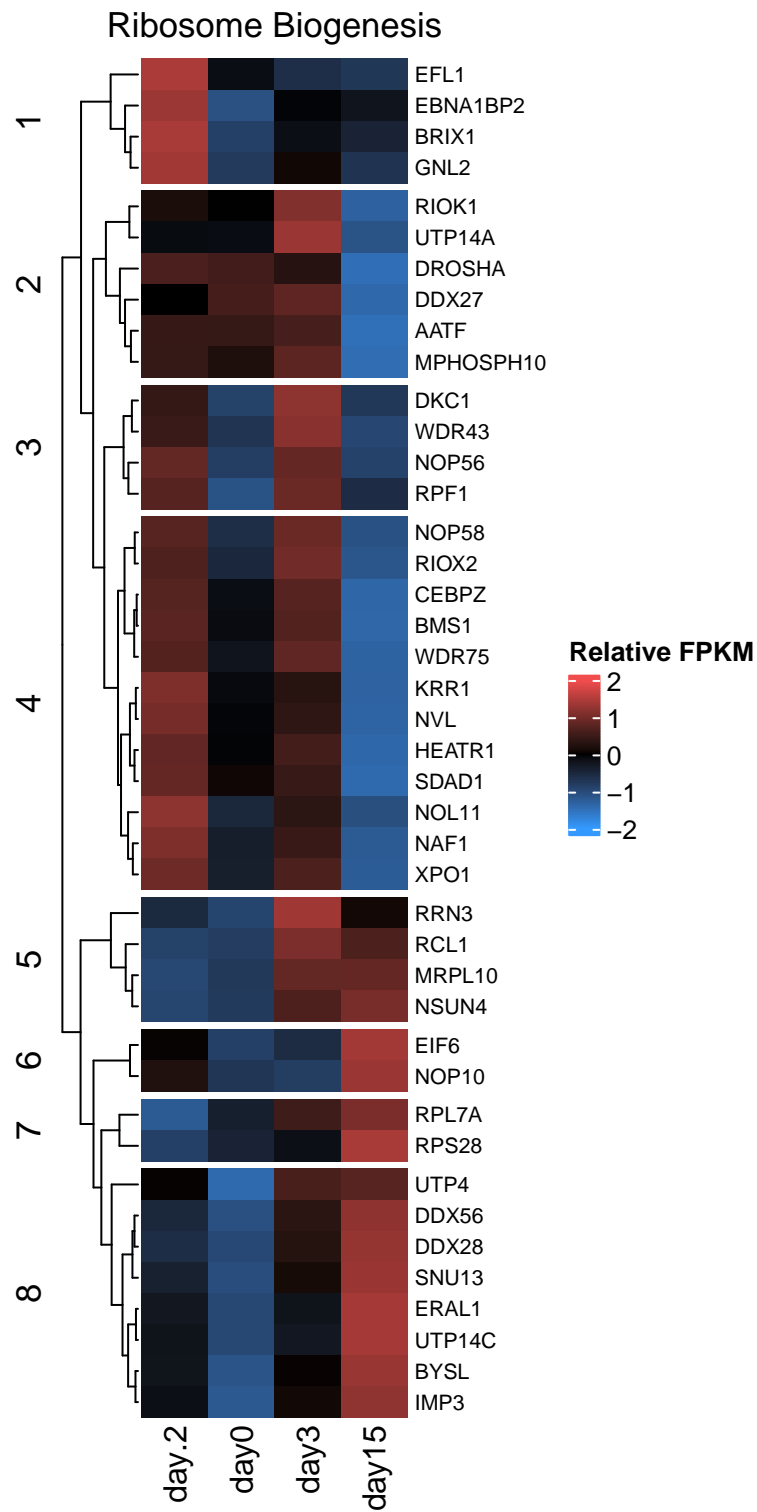



```
Heatmap(as.matrix(zs13[rownames(zs13) %in% ribogen$external_gene_name,]),
  cluster_columns = F,
  name="Relative FPKM",
  column_title = "Ribosome Biogenesis",
  col=circlize::colorRamp2(c(-2,0,2),c(rgb(0.2,0.6,1),rgb(0,0,0),rgb(0.95,0.3,0.3))),
```

```

show_row_names = T, row_names_gp = gpar(fontsize=8),
row_dend_reorder = order13[rownames(zs13) %in% ribogen$external_gene_name],
row_split=8)

```



Supp Figure 6a upper panel

```
Heatmap(as.matrix(zs13[rownames(zs13) %in% ribogen$external_gene_name,]),
        cluster_columns = F,
        name="Relative FPKM",
        column_title = "Ribosome Biogenesis",
        col=circlize::colorRamp2(c(-2,0,2),c(rgb(0.2,0.6,1),rgb(0,0,0),rgb(0.95,0.3,0.3))),
        show_row_names = T, row_names_gp = gpar(fontsize=8),
        row_dend_reorder = order13[rownames(zs13) %in% ribogen$external_gene_name],
        row_split=2)
```

