Figure 3 Heatmap

true

Setup

Set directory

```
setwd("D:/Krueger_Lab/Publications/miR181_paper/Figure3")
```

Load packages

```
source("D:/Krueger_Lab/Publications/miR181_paper/Supporting_scripts/themes/theme_paper.R")
library(ComplexHeatmap)
```

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

```
## ========================================
## ComplexHeatmap version 2.15.2
## 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 either one:
## - Gu, Z. Complex Heatmap Visualization. iMeta 2022.
## - 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))
## ========================================
```

```
library(ggplot2)
library(circlize)
```

```
## ========================================
## circlize version 0.4.15
## CRAN page: https://cran.r-project.org/package=circlize
## Github page: https://github.com/jokergoo/circlize
## Documentation: https://jokergoo.github.io/circlize_book/book/
##
## If you use it in published research, please cite:
## Gu, Z. circlize implements and enhances circular visualization
##   in R. Bioinformatics 2014.
##
## This message can be suppressed by:
```

```
##   suppressPackageStartupMessages(library(circlize))
## ========================================
```

```
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
```

```
library(eulerr)
library(xlsx)
library(ggpubr)
```

Load data

```
#Ribo
RNA <- read.csv("D:/Krueger_Lab/Publications/miR181_paper_v07122022/Figure_3/RNA_masterframe.csv")
RPF <- read.csv("D:/Krueger_Lab/Publications/miR181_paper_v07122022/Figure_3/RPF_masterframe.csv")
#ms
ms <- as.data.frame(read.xlsx("D:/Krueger_Lab/R/ECDF plots/Kreuger_analysis_ms.xlsx",
                              sheetName = "Analysis"))
#adjust upper case lower case of gene names of MS data:
#function
capFirst <- function(s) {
  paste(toupper(substring(s, 1, 1)), substring(s, 2), sep = "")
}
ms$Gene.Symbol <- tolower(ms$Gene.Symbol)
ms$Gene.Symbol <- capFirst(ms$Gene.Symbol)
names(ms)[names(ms) == 'Gene.Symbol'] <- 'GeneName'
head(ms$GeneName)
```

```
## [1] "Ckb"     "Gnb4"    "Ccm2"    "Rnpep"   "Aldh1b1" "Macf1"
```

```
#diff clip
eclipGR <- readRDS("D:/Krueger_Lab/miReCLIP/Mirco/DifferentialBinding/BsDifferentialResult.rds")
eclip <- as.data.frame(eclipGR)
```

```
## Loading required package: GenomicRanges
```

```
## Loading required package: stats4
```

```
## Loading required package: BiocGenerics
```

```
##
## Attaching package: 'BiocGenerics'
```

```
## The following objects are masked from 'package:dplyr':
##
##     combine, intersect, setdiff, union
```

```
## The following objects are masked from 'package:stats':
##
##     IQR, mad, sd, var, xtabs
```

```
## The following objects are masked from 'package:base':
##
##     anyDuplicated, aperm, append, as.data.frame, basename, cbind,
##     colnames, dirname, do.call, duplicated, eval, evalq, Filter, Find,
##     get, grep, grepl, intersect, is.unsorted, lapply, Map, mapply,
##     match, mget, order, paste, pmax, pmax.int, pmin, pmin.int,
##     Position, rank, rbind, Reduce, rownames, sapply, setdiff, sort,
##     table, tapply, union, unique, unsplit, which.max, which.min

## Loading required package: S4Vectors

##
## Attaching package: 'S4Vectors'

## The following objects are masked from 'package:dplyr':
##
##     first, rename

## The following objects are masked from 'package:base':
##
##     expand.grid, I, unname

## Loading required package: IRanges

##
## Attaching package: 'IRanges'

## The following objects are masked from 'package:dplyr':
##
##     collapse, desc, slice

## The following object is masked from 'package:grDevices':
##
##     windows

## Loading required package: GenomeInfoDb
```

```
tframe <- eclip[eclip$res.log2FoldChange < 0 & eclip$res.padj <= 0.05,]
head(tframe)
```

```
##     seqnames    start      end width strand scoreSum scoreMean scoreMax WT KO
## 40      chr1 15829574 15829580     7      + 18.32669  18.32669 18.32669  1  0
## 114     chr1 43509974 43509980     7      + 41.39330  41.39330 41.39330  1  0
## 204     chr1 64576272 64576278     7      + 49.28560  49.28560 49.28560  1  0
## 212     chr1 64601677 64601683     7      + 22.84493  22.84493 22.84493  0  1
## 251     chr1 82896040 82896046     7      + 31.07340  31.07340 31.07340  1  0
## 254     chr1 85650006 85650012     7      + 64.77470  64.77470 64.77470  1  0
##                 geneID geneName region               gene_id counts.bs.1_KO
## 40  ENSMUSG00000025925    Terf1 intron ENSMUSG00000025925.14              2
## 114 ENSMUSG00000066877     Nck2 intron ENSMUSG00000066877.11              1
## 204 ENSMUSG00000025958    Creb1   utr3 ENSMUSG00000025958.14              5
## 212 ENSMUSG00000025958    Creb1   utr3 ENSMUSG00000025958.14              8
## 251 ENSMUSG00000026159    Agfg1   utr3 ENSMUSG00000026159.13              2
## 254 ENSMUSG00000026222    Sp100   utr5 ENSMUSG00000026222.16              8
##     counts.bs.2_KO counts.bs.3_KO counts.bs.4_WT counts.bs.5_WT counts.bs.6_WT
## 40               0              0             12              4              2
## 114              1              0             13              6              6
## 204              3              2             13             13              8
## 212              8              7             38             13             13
```

```
## 251              1              0             10              8             14
## 254              4              1             20             22             11
##     counts.bg.1_KO counts.bg.2_KO counts.bg.3_KO counts.bg.4_WT counts.bg.5_WT
## 40             398            580            396            623            466
## 114           4883           7795           4297           8365           6486
## 204           2524           2754           1831           3265           2963
## 212           2524           2754           1831           3265           2963
## 251            881           1074            671           1474           1202
## 254           1376           2007           1142           2498           1935
##     counts.bg.6_WT res.baseMean res.log2FoldChange res.lfcSE res.stat
## 40             254     34.60655          -3.220330 1.0850775 14.64585
## 114           3484    220.01721          -3.535134 1.0657093 20.82672
## 204           1788    134.36861          -1.583008 0.5372789 10.00932
## 212           1788    180.26615          -1.281178 0.3808920 12.29309
## 251            657     77.88992          -3.082714 0.8749484 21.32573
## 254            949    127.11799          -1.800166 0.4592001 18.56607
##       res.pvalue     res.padj                    BsID
## 40  1.297197e-04 0.0067674589  ENSMUSG00000025925.bs1
## 114 5.027663e-06 0.0004944622 ENSMUSG00000066877.bs21
## 204 1.557503e-03 0.0396302933  ENSMUSG00000025958.bs3
## 212 4.546375e-04 0.0170974938 ENSMUSG00000025958.bs11
## 251 3.874960e-06 0.0004061979  ENSMUSG00000026159.bs7
## 254 1.641154e-05 0.0012883428  ENSMUSG00000026222.bs1
```

```
#newdiffclip
mir181bs <- readRDS("D:/Krueger_Lab/Publications/miR181_paper/Figure1/mir181_binding_sites__venn_types/m
names(mir181bs) <- 1:length(mir181bs)
mir181df <- as.data.frame(mir181bs)

#Translational efficiency
TEframe <- read.csv("D:/Krueger_Lab/Publications/miR181_paper_v07122022/Supporting scipts/deltaTE/TE_m23
head(TEframe)
```

```
##   X  baseMean log2FoldChange     lfcSE        stat      pvalue      padj
## 1 1 2354.3220    -0.07560025 0.1336485 -0.56566464 0.571621774 0.9277760
## 2 2  579.2399    -0.06784576 0.1617420 -0.41946900 0.674873409 0.9512329
## 3 3   20.3333    -0.02453968 0.5874974 -0.04176985 0.966682171 0.9964083
## 4 4 1154.1664    -0.05314227 0.1359683 -0.39084299 0.695913297 0.9542063
## 5 5  635.5639     0.36967194 0.1364473  2.70926532 0.006743239 0.2886809
## 6 6  378.2069     0.10567650 0.1653495  0.63910971 0.522751566 0.9143531
##                    Gene gene_symbol
## 1  ENSMUSG00000000001.4       Gnai3
## 2 ENSMUSG00000000028.15       Cdc45
## 3 ENSMUSG00000000037.17       Scml2
## 4  ENSMUSG00000000056.7        Narf
## 5  ENSMUSG00000000078.7        Klf6
## 6 ENSMUSG00000000085.16       Scmh1
```

colour pattern

```
#colours
farbeneg <- "#b4b4b4"
farbe1 <- "#0073C2FF"
farbe2 <- "#EFC000FF"
farbe3 <- "#CD534CFF"
farbe4 <- "#7AA6DCFF"
```

```
farbe5 <- "#868686FF"
farbe6 <- "#003C67FF"
farbe7 <- "#8F7700FF"
farbe8 <- "#3B3B3BFF"
farbe9 <- "#A73030FF"
farbe10 <- "#4A6990FF"
farbe11 <- "#FF6F00FF"
farbe12 <- "#C71000FF"
farbe13 <- "#008EA0FF"
farbe14 <- "#8A4198FF"
farbe15 <- "#5A9599FF"
farbe16 <- "#FF6348FF"

RNApcol <- "#b56504"
RNAncol <- "#027d73"
RPFpcol <- "#c4c404"
RPFncol <- "#8d0391"
```

Venn diagram to figure out the dataset for the joint heatmaps

create column for heatmap

Make columns to indentify genes included in both datasets and that are significant

```
mID <- RNA[RNA$LFCandPADJSig %in% c("Significant up", "Significant down") & RNA$Gene %in%
              RPF[RPF$LFCandPADJSig %in% c("Significant up", "Significant down"),"Gene"],"Gene"]

RNA$mID <- "One dataset"
RNA$mID[RNA$Gene %in% mID] <- "Both datasets"
RPF$mID <- "One dataset"
RPF$mID[RPF$Gene %in% mID] <- "Both datasets"

head(RNA)
```

```
##   X                  Gene   baseMean log2FoldChange      lfcSE      stat
## 1 1  ENSMUSG00000104197.1 2009.89774       2.741830 0.10245760  26.76063
## 2 2 ENSMUSG00000004110.17 1028.70696      -2.019241 0.09914247 -20.36706
## 3 3 ENSMUSG00000027669.14  648.18701       1.905815 0.10065041  18.93499
## 4 4  ENSMUSG00000098206.1 4004.62856       2.148076 0.12889738  16.66501
## 5 5 ENSMUSG00000004552.16  516.71005      -2.252195 0.13648248 -16.50172
## 6 6  ENSMUSG00000069306.5   68.07176       2.763638 0.19108029  14.46323
##          pvalue          padj    WT_1411   WT_1601      WT_1710   KO_1411
## 1 9.288051e-158 1.183948e-153  487.52165  474.5874  405.8464624 3338.7077
## 2  3.277594e-92  2.088974e-88 1576.25174 1737.4574 1739.7225816  400.4560
## 3  5.872559e-80  2.495250e-76  248.88185  252.5740  252.2109307 1048.3637
## 4  2.354521e-62  7.503269e-59 1449.25030 1259.2746  985.7539896 5751.8330
## 5  3.566111e-61  9.091443e-58  872.62279 1026.4751  777.9463918  148.2821
## 6  2.068594e-47  4.394729e-44    2.04841    0.0000    0.8880667  123.7258
##     KO_1601   KO_1710    LFCandPADJSig   gene_symbol           mID
## 1 3428.5263 3924.1970   Significant up       Gm37632 Both datasets
## 2  357.0428  361.3113 Significant down       Cacna1e Both datasets
## 3 1043.6635 1043.4281   Significant up          Gnb4 Both datasets
## 4 6644.9630 7936.6965   Significant up A430106G13Rik Both datasets
## 5  112.9110  162.0230 Significant down          Ctse Both datasets
```

```
## 6  146.4791  135.2892   Significant up      Hist1h4m Both datasets
```

```
head(RPF)
```

```
##   X                  Gene   baseMean log2FoldChange      lfcSE       stat
## 1 1  ENSMUSG00000000001.4 3788.57802     0.05462417 0.05957247  0.9169365
## 2 2 ENSMUSG00000000028.15 1086.36203    -0.12956555 0.07692175 -1.6843811
## 3 3 ENSMUSG00000000037.17   19.45454    -0.06042604 0.15784031 -0.3828302
## 4 4  ENSMUSG00000000056.7 1454.64304     0.05456272 0.08781406  0.6213438
## 5 5  ENSMUSG00000000078.7  850.27809     0.65975317 0.07971920  8.2759636
## 6 6 ENSMUSG00000000085.16  337.25837    -0.06342545 0.12383411 -0.5121808
##         pvalue         padj    WT_1411    WT_1601    WT_1710    KO_1411
## 1 3.591759e-01 6.850504e-01 3687.40632 3612.48756 3862.10415 3905.91865
## 2 9.210804e-02 3.386014e-01 1091.79093 1147.15325 1182.54053 1081.59050
## 3 7.018457e-01 8.852986e-01   13.71269   22.17027   29.07887   28.37993
## 4 5.343734e-01 8.038393e-01 1561.28714 1359.98786 1348.39777 1344.36759
## 5 1.274210e-16 2.314117e-14  649.06709  668.27536  622.50312 1054.26168
## 6 6.085245e-01 8.452822e-01  325.18653  326.21972  396.33417  351.07019
##      KO_1601     KO_1710   LFCandPADJSig gene_symbol         mID
## 1 3984.05486 3679.496580 Not significant       Gnai3 One dataset
## 2 1030.01306  985.083926 Not significant       Cdc45 One dataset
## 3   17.83572    5.549769 Not significant       Scml2 One dataset
## 4 1559.88270 1553.935207 Not significant        Narf One dataset
## 5 1022.58151 1084.979761  Significant up        Klf6 One dataset
## 6  350.02608  274.713546 Not significant       Scmh1 One dataset
```

plot with venn

Make Venn diagrams for the significant genes

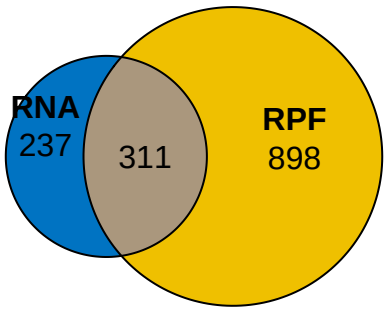
List

```
vlist <- list(RNA[RNA$LFCandPADJSig  %in% c("Significant up", "Significant down"),"Gene"],
              RPF[RPF$LFCandPADJSig  %in% c("Significant up", "Significant down"),"Gene"])
names(vlist) <- c("RNA", "RPF")


str(vlist)
```

```
## List of 2
##  $ RNA: chr [1:548] "ENSMUSG00000104197.1" "ENSMUSG00000004110.17" "ENSMUSG00000027669.14" "ENSMUSG0
##  $ RPF: chr [1:1209] "ENSMUSG00000000078.7" "ENSMUSG00000000168.9" "ENSMUSG00000000184.12" "ENSMUSG0
```

plot

```
plot(euler(vlist, shape="ellipse"), quantities = T, fills=c(farbe1, farbe2), main = "Overlap of differe
```




```
pdf("Venn_Overlap of differentially regulated genes.pdf", width = 3, height = 3)
plot(euler(vlist, shape="ellipse"), quantities = T, fills=c(farbe1, farbe2), main = "Overlap of differe
dev.off()
```

```
## pdf
##   2
```

Heatmap scaled together

Combine experiments into one df and scale

```
head(RNA)
```

```
##   X                  Gene   baseMean log2FoldChange      lfcSE      stat
## 1 1  ENSMUSG00000104197.1 2009.89774       2.741830 0.10245760  26.76063
## 2 2 ENSMUSG00000004110.17 1028.70696      -2.019241 0.09914247 -20.36706
## 3 3 ENSMUSG00000027669.14  648.18701       1.905815 0.10065041  18.93499
## 4 4  ENSMUSG00000098206.1 4004.62856       2.148076 0.12889738  16.66501
## 5 5 ENSMUSG00000004552.16  516.71005      -2.252195 0.13648248 -16.50172
## 6 6  ENSMUSG00000069306.5   68.07176       2.763638 0.19108029  14.46323
##          pvalue          padj    WT_1411   WT_1601      WT_1710   KO_1411
## 1 9.288051e-158 1.183948e-153  487.52165  474.5874  405.8464624 3338.7077
## 2  3.277594e-92  2.088974e-88 1576.25174 1737.4574 1739.7225816  400.4560
## 3  5.872559e-80  2.495250e-76  248.88185  252.5740  252.2109307 1048.3637
## 4  2.354521e-62  7.503269e-59 1449.25030 1259.2746  985.7539896 5751.8330
## 5  3.566111e-61  9.091443e-58  872.62279 1026.4751  777.9463918  148.2821
## 6  2.068594e-47  4.394729e-44    2.04841    0.0000    0.8880667  123.7258
##     KO_1601   KO_1710    LFCandPADJSig   gene_symbol           mID
## 1 3428.5263 3924.1970   Significant up       Gm37632 Both datasets
## 2  357.0428  361.3113 Significant down       Cacna1e Both datasets
## 3 1043.6635 1043.4281   Significant up          Gnb4 Both datasets
## 4 6644.9630 7936.6965   Significant up A430106G13Rik Both datasets
## 5  112.9110  162.0230 Significant down          Ctse Both datasets
## 6  146.4791  135.2892   Significant up      Hist1h4m Both datasets
```

```
RNAhframe <- RNA[RNA$Gene %in% mID, c(16, 9:14)]
RPFhframe <- RPF[RPF$Gene %in% mID, c(16, 9:14)]

colnames(RNAhframe) <- c("Gene", "WT1 RNA", "WT2 RNA", "WT3 RNA", "KO1 RNA", "KO2 RNA", "KO3 RNA")
colnames(RPFhframe) <- c("Gene", "WT1 RPF", "WT2 RPF", "WT3 RPF", "KO1 RPF", "KO2 RPF", "KO3 RPF")

head(RNAhframe)
```

```
##            Gene    WT1 RNA   WT2 RNA      WT3 RNA   KO1 RNA   KO2 RNA   KO3 RNA
## 1       Gm37632  487.52165  474.5874  405.8464624 3338.7077 3428.5263 3924.1970
## 2       Cacna1e 1576.25174 1737.4574 1739.7225816  400.4560  357.0428  361.3113
## 3          Gnb4  248.88185  252.5740  252.2109307 1048.3637 1043.6635 1043.4281
## 4 A430106G13Rik 1449.25030 1259.2746  985.7539896 5751.8330 6644.9630 7936.6965
## 5          Ctse  872.62279 1026.4751  777.9463918  148.2821  112.9110  162.0230
## 6      Hist1h4m    2.04841    0.0000    0.8880667  123.7258  146.4791  135.2892
```

```
head(RPFhframe)
```

```
##        Gene    WT1 RPF    WT2 RPF    WT3 RPF   KO1 RPF    KO2 RPF    KO3 RPF
## 16    Ccnd2 1744.12295 1558.88688 1980.59384 1293.9144 1432.06001 1104.40395
## 34     Ccm2 2743.19000 2397.55662 2641.86879 1770.0665 2025.84099 1823.09898
## 161  Gtf2h4 2547.29450 2770.01720 2432.93175 1552.4870 1418.68322 1603.88312
## 238  Rmnd5a 5937.59264 5952.40147 6067.78995 3991.0584 3981.08224 3857.08917
## 376  Cacnb3   67.25746   67.77769   74.31266   25.2266   25.26728   19.42419
## 452 Cacna1e  534.14173  627.73543  611.73317  145.0530  121.13429  122.09491
```

```
heatframe <- left_join(RNAhframe, RPFhframe, by="Gene")
rownames(heatframe) <- heatframe$Gene
heatframe <- heatframe[,-1]
heatframe <- heatframe[!(rownames(heatframe)=="") & !is.na(rownames(heatframe)),]
head(heatframe)
```

```
##                   WT1 RNA   WT2 RNA      WT3 RNA   KO1 RNA   KO2 RNA   KO3 RNA
## Gm37632         487.52165  474.5874  405.8464624 3338.7077 3428.5263 3924.1970
## Cacna1e        1576.25174 1737.4574 1739.7225816  400.4560  357.0428  361.3113
## Gnb4            248.88185  252.5740  252.2109307 1048.3637 1043.6635 1043.4281
## A430106G13Rik  1449.25030 1259.2746  985.7539896 5751.8330 6644.9630 7936.6965
## Ctse            872.62279 1026.4751  777.9463918  148.2821  112.9110  162.0230
## Hist1h4m          2.04841    0.0000    0.8880667  123.7258  146.4791  135.2892
##                   WT1 RPF    WT2 RPF    WT3 RPF   KO1 RPF   KO2 RPF   KO3 RPF
## Gm37632          37.87313   31.67182   38.77182  293.2592  330.7041  285.8131
## Cacna1e         534.14173  627.73543  611.73317  145.0530  121.1343  122.0949
## Gnb4            286.00743  282.51262  308.02058 1026.9329 1038.1878  943.4607
## A430106G13Rik   171.08207  156.45878  175.55019  491.9187  534.3286  471.7303
## Ctse           4012.59285 3587.78354 3646.70514  466.6921  526.1539  513.3536
## Hist1h4m         29.38433   20.90340   28.00187 3535.9285 4264.9675 3698.9208
```

scale

```
heat_scaled = t(scale(t(heatframe)))
colnames(heat_scaled) <- c(rep("WT", 3), rep("miR-181-KO", 3), rep("WT", 3), rep("miR-181-KO", 3))

head(as.data.frame(heat_scaled))
```

```
##                        WT         WT         WT miR-181-KO miR-181-KO miR-181-KO
## Gm37632       -0.3998356 -0.4084225 -0.4540584  1.4930180  1.5526471  1.8817144
## Cacna1e        1.4161464  1.6750556  1.6786937 -0.4722763 -0.5420014 -0.5351459
## Gnb4          -1.0122475 -1.0028800 -1.0038011  1.0161500  1.0042252  1.0036278
## A430106G13Rik -0.2526860 -0.3193752 -0.4153921  1.2576956  1.5712206  2.0246716
## Ctse          -0.3001186 -0.1971680 -0.3634715 -0.7848129 -0.8084815 -0.7756181
## Hist1h4m      -0.5802978 -0.5814903 -0.5809733 -0.5094641 -0.4962184 -0.5027326
##                        WT         WT         WT miR-181-KO miR-181-KO miR-181-KO
## Gm37632       -0.6983496 -0.7024666 -0.6977530 -0.5288031 -0.5039442 -0.5337465
```

```
## Cacna1e       -0.2575663 -0.1072472 -0.1329482 -0.8824742 -0.9208895 -0.9193467
## Gnb4          -0.9180547 -0.9269215 -0.8622042  0.9617771  0.9903324  0.7499964
## A430106G13Rik -0.7013751 -0.7065084 -0.6998066 -0.5887484 -0.5738608 -0.5958353
## Ctse           1.8009996  1.5167375  1.5561650 -0.5717481 -0.5319591 -0.5405244
## Hist1h4m      -0.5643844 -0.5693215 -0.5651892  1.4769271  1.9013324  1.5718120
```

cluster without heatmap (this is not done yet)

```
# heat_ks <- kmeans(heat_scaled, centers = 5)
# heat_k_names <- as.data.frame(heat_ks$cluster)
# merge(heat_scaled, heat_k_names, by=0)
```

make heatmaps

make annotations for the heatmap

```
ha1 <- HeatmapAnnotation(Genotype=colnames(heat_scaled),col = list(Genotype= c("WT"=farbe1, "miR-181-KO
```

make annotation heatmaps as seperate heatmaps

ms

```
mshmat <- data.frame(row.names = rownames(heatframe), rownames(heatframe))
colnames(mshmat) <- "GeneName"
mshmat <- left_join(mshmat, ms[,c("GeneName", "Log2.FC.")], by="GeneName")
mshma <- as.data.frame(mshmat[,"Log2.FC."])
rownames(mshma) <- mshmat$GeneName
colnames(mshma) <- "Log2FC in MS"
head(mshma)
```

```
##               Log2FC in MS
## Gm37632                 NA
## Cacna1e                 NA
## Gnb4              1.572372
## A430106G13Rik           NA
## Ctse                    NA
## Hist1h4m                NA
```

```
hmms <- Heatmap(mshma, show_row_dend = F, show_column_dend = F, show_row_names = F, colorRamp2(c(-0.5, 
                name = "Mass spec")
```

```
## Warning: The input is a data frame-like object, convert it to a matrix.
```

Targets

```
targethmt <- data.frame(row.names = rownames(heatframe), rownames(heatframe))
colnames(targethmt) <- "GeneName"
targethmt$Target <- "Non-target"
targethmt[targethmt$GeneName %in% mir181df$geneName, "Target"] <- "Target"

targethm <- as.data.frame(targethmt[,"Target"])
rownames(targethm) <- targethmt$GeneName
colnames(targethm) <- "miR-181 target"
head(targethm)
```

```
##               miR-181 target
## Gm37632           Non-target
```

```
## Cacna1e               Target
## Gnb4              Non-target
## A430106G13Rik     Non-target
## Ctse                  Target
## Hist1h4m          Non-target
```

```
hmtarget <- Heatmap(targethm, show_row_dend = F, show_column_dend = F, show_row_names = F , col = c(far
                    name = "miR-181 target")
```

```
## Warning: The input is a data frame-like object, convert it to a matrix.
```

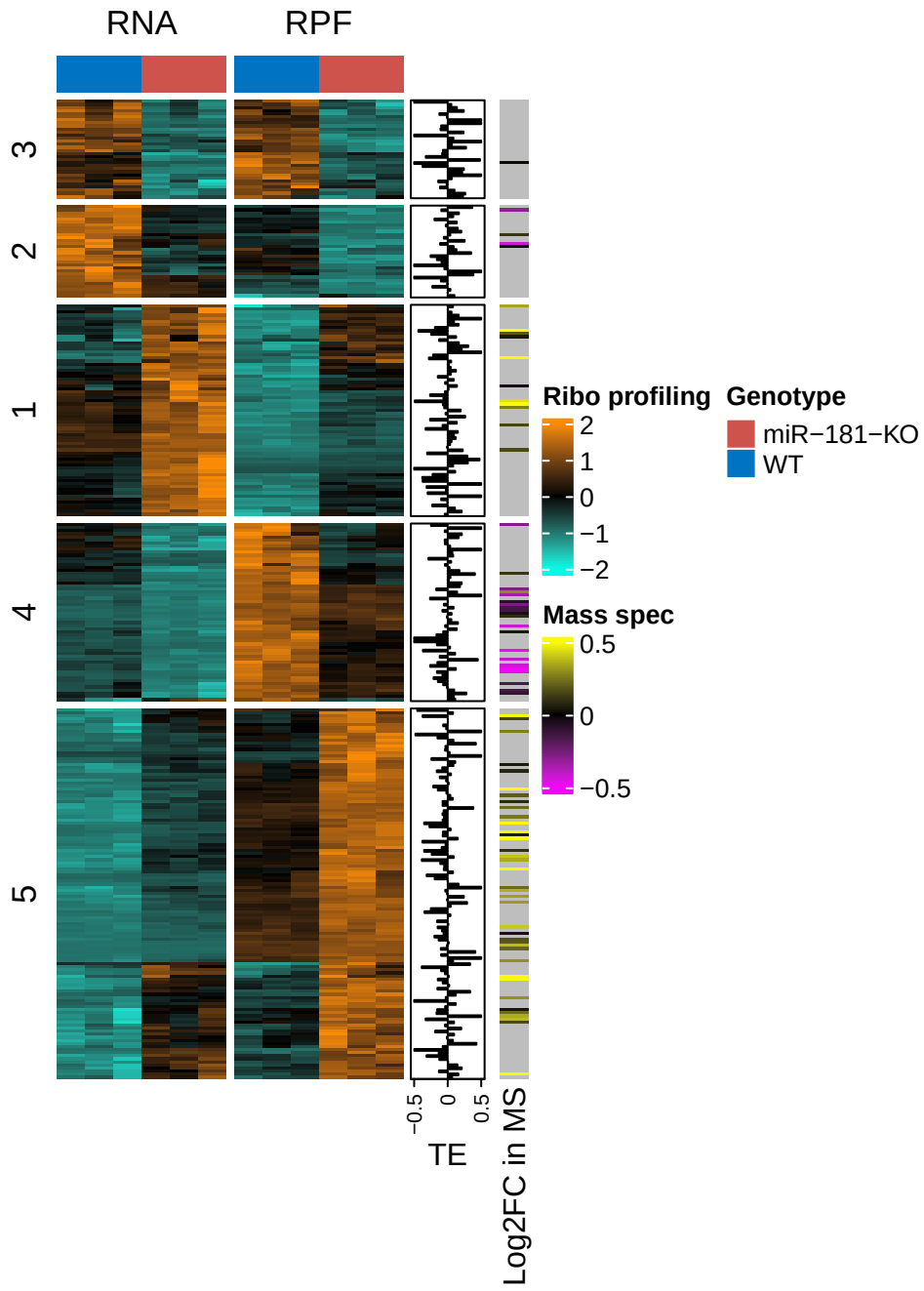
```
## Warning: Note: not all columns in the data frame are numeric. The data frame
## will be converted into a character matrix.
```

TE annotation

```
TEsel = TEframe[unique(TEframe$gene_symbol) %in% rownames(heatframe), c("log2FoldChange")]
names(TEsel) = TEframe[unique(TEframe$gene_symbol) %in% rownames(heatframe), c("gene_symbol")]

TEsel = TEsel[order(factor(names(TEsel), levels = rownames(heat_scaled)))]

#set maximum
TElim = 0.5

TEsel[TEsel > TElim] = TElim
TEsel[TEsel < -TElim] = -TElim

haTE = rowAnnotation(TE = anno_barplot(TEsel, ylim = c(-TElim,TElim)))
```

plot

```
set.seed(666)


HRNA <- Heatmap(heat_scaled, show_row_dend = F, show_column_dend = F, show_row_names = F,show_column_na
                column_order=1:12 , top_annotation =c(ha1), column_split = c(rep("RNA",6), rep("RPF",6)
                name = "Ribo profiling", right_annotation = haTE)


HRNA + hmms
```



Heatmap scald seperately

scale

```
RNA2hframe <- RNAhframe
rownames(RNA2hframe) <- RNA2hframe$Gene
RNA2hframe <- RNA2hframe[,-1]
head(RNA2hframe)
```

```
##                 WT1 RNA   WT2 RNA     WT3 RNA   KO1 RNA   KO2 RNA   KO3 RNA
```

```
## Gm37632        487.52165  474.5874  405.8464624 3338.7077 3428.5263 3924.1970
## Cacna1e       1576.25174 1737.4574 1739.7225816  400.4560  357.0428  361.3113
## Gnb4           248.88185  252.5740  252.2109307 1048.3637 1043.6635 1043.4281
## A430106G13Rik 1449.25030 1259.2746  985.7539896 5751.8330 6644.9630 7936.6965
## Ctse           872.62279 1026.4751  777.9463918  148.2821  112.9110  162.0230
## Hist1h4m         2.04841    0.0000    0.8880667  123.7258  146.4791  135.2892
```

```
RPF2hframe <- RPFhframe
rownames(RPF2hframe) <- RPF2hframe$Gene
RPF2hframe <- RPF2hframe[,-1]
head(RPF2hframe)
```

```
##            WT1 RPF    WT2 RPF    WT3 RPF   KO1 RPF    KO2 RPF    KO3 RPF
## Ccnd2   1744.12295 1558.88688 1980.59384 1293.9144 1432.06001 1104.40395
## Ccm2    2743.19000 2397.55662 2641.86879 1770.0665 2025.84099 1823.09898
## Gtf2h4  2547.29450 2770.01720 2432.93175 1552.4870 1418.68322 1603.88312
## Rmnd5a  5937.59264 5952.40147 6067.78995 3991.0584 3981.08224 3857.08917
## Cacnb3    67.25746   67.77769   74.31266   25.2266   25.26728   19.42419
## Cacna1e  534.14173  627.73543  611.73317  145.0530  121.13429  122.09491
```

```
RNA_scaled = as.data.frame(t(scale(t(RNA2hframe))))
colnames(RNA_scaled) <- c(rep("WT", 3), rep("miR-181-KO", 3))
head(RNA_scaled)
```

```
##                       WT         WT         WT miR-181-KO miR-181-KO miR-181-KO
## Gm37632       -0.8881526 -0.8956985 -0.9358019  0.7752263  0.8276264  1.1168002
## Cacna1e        0.7594651  0.9830633  0.9862052 -0.8714075 -0.9316233 -0.9257027
## Gnb4          -0.9182414 -0.9097510 -0.9105859  0.9202455  0.9094371  0.9088957
## A430106G13Rik -0.8190839 -0.8799774 -0.9676499  0.5600372  0.8463151  1.2603587
## Ctse           0.8486754  1.2155368  0.6229191 -0.8785180 -0.9628606 -0.8457527
## Hist1h4m      -0.8940090 -0.9217461 -0.9097210  0.7536002  1.0616980  0.9101779
```

```
RPF_scaled = as.data.frame(t(scale(t(RPF2hframe))))
colnames(RPF_scaled) <- c(rep("WT", 3), rep("miR-181-KO", 3))
head(RPF_scaled)
```

```
##                WT        WT        WT miR-181-KO miR-181-KO miR-181-KO
## Ccnd2   0.7151707 0.1267205 1.4663815 -0.7150330 -0.2761778 -1.3170618
## Ccm2    1.2148270 0.3908552 0.9732826 -1.1050482 -0.4952952 -0.9786214
## Gtf2h4  0.8317081 1.2073894 0.6388049 -0.8462999 -1.0719957 -0.7596067
## Rmnd5a  0.8681965 0.8814091 0.9843593 -0.8685120 -0.8774128 -0.9880401
## Cacnb3  0.8070374 0.8273071 1.0819261 -0.8305928 -0.8290080 -1.0566697
## Cacna1e 0.6815571 1.0485294 0.9857860 -0.8440244 -0.9378073 -0.9340408
```

now combine the tables

```
RNA_scaled$gene.symbol <- rownames(RNA_scaled)
colnames(RNA_scaled) <- c(rep("WT_RNA", 3), rep("miR-181-KO_RNA", 3), "gene.symbol")
RPF_scaled$gene.symbol <- rownames(RPF_scaled)
colnames(RPF_scaled) <- c(rep("WT_RPF", 3), rep("miR-181-KO_RPF", 3), "gene.symbol")


SepScaleframe <- merge(RNA_scaled,
                       RPF_scaled, by="gene.symbol")
```

```
## Warning in merge.data.frame(RNA_scaled, RPF_scaled, by = "gene.symbol"): column
```

```
## names 'WT_RNA', 'WT_RNA', 'miR-181-KO_RNA', 'miR-181-KO_RNA' are duplicated in
## the result
```

```
SepScaleframe = SepScaleframe[!(SepScaleframe$gene.symbol == ""),]

rownames(SepScaleframe) <- SepScaleframe$gene.symbol
SepScaleframe <- SepScaleframe[,-1]
colnames(SepScaleframe) <- c(rep("WT", 3), rep("miR-181-KO", 3), rep("WT", 3), rep("miR-181-KO", 3))
head(SepScaleframe)
```

```
##                         WT         WT         WT miR-181-KO miR-181-KO
## 2010016I18Rik  -0.2852270 -1.1619350 -0.9454200  0.1187576  1.3311953
## 2510009E07Rik  -0.8990716 -0.8810933 -0.9549938  0.8714430  0.8718551
## 4930523C07Rik  -0.5093680 -0.8683728 -1.2011361  0.4487555  0.8731558
## 5830411N06Rik   0.7512700  0.1727972  1.4868718 -0.7829998 -0.4556643
## 9930111J21Rik2  1.0142836  0.8153492  0.9035324 -0.9110550 -0.9110550
## A430035B10Rik   1.5087331  0.7002925  0.2304264 -1.1538509 -0.7487766
##                 miR-181-KO         WT         WT         WT miR-181-KO
## 2010016I18Rik   0.9426291 -1.1351577 -0.6698709 -0.8295780  1.2851310
## 2510009E07Rik   0.9918607 -0.8886151 -0.9452000 -0.9026435  0.8877953
## 4930523C07Rik   1.2569655 -0.9653176 -0.9695535 -0.7732223  1.0190113
## 5830411N06Rik  -1.1722748  1.4113550 -0.0066162  0.9717884 -1.0438427
## 9930111J21Rik2 -0.9110550  0.9061925  0.9021519  0.9301430 -0.9128291
## A430035B10Rik  -0.5368244  1.4080276  0.6025263  0.5038469 -1.2628822
##                 miR-181-KO miR-181-KO
## 2010016I18Rik   0.5960406  0.7534351
## 2510009E07Rik   0.9755001  0.8731633
## 4930523C07Rik   1.0253338  0.6637484
## 5830411N06Rik  -0.4589426 -0.8737418
## 9930111J21Rik2 -0.9128291 -0.9128291
## A430035B10Rik  -0.7319144 -0.5196042
```

make heatmaps

make annotations for the heatmap

```
ha3 <- HeatmapAnnotation(Genotype=colnames(SepScaleframe),col = list(Genotype= c("WT"=farbe1, "miR-181-
ha4 <- HeatmapAnnotation(Experiment=c(rep("RNA",6), rep("RPF",6)),col = list(Experiment= c("RNA"=farbe2
```

TE annotation

```
TEsel2 = TEframe[unique(TEframe$gene_symbol) %in% rownames(heatframe), c("log2FoldChange")]
names(TEsel2) = TEframe[unique(TEframe$gene_symbol) %in% rownames(heatframe), c("gene_symbol")]

TEsel2 = TEsel2[order(factor(names(TEsel2), levels = rownames(SepScaleframe)))]

TEsel2[TEsel2 > TElim] = TElim
TEsel2[TEsel2 < -TElim] = -TElim

haTE2 = rowAnnotation(TE = anno_barplot(TEsel2, ylim = c(-TElim,TElim)))
```

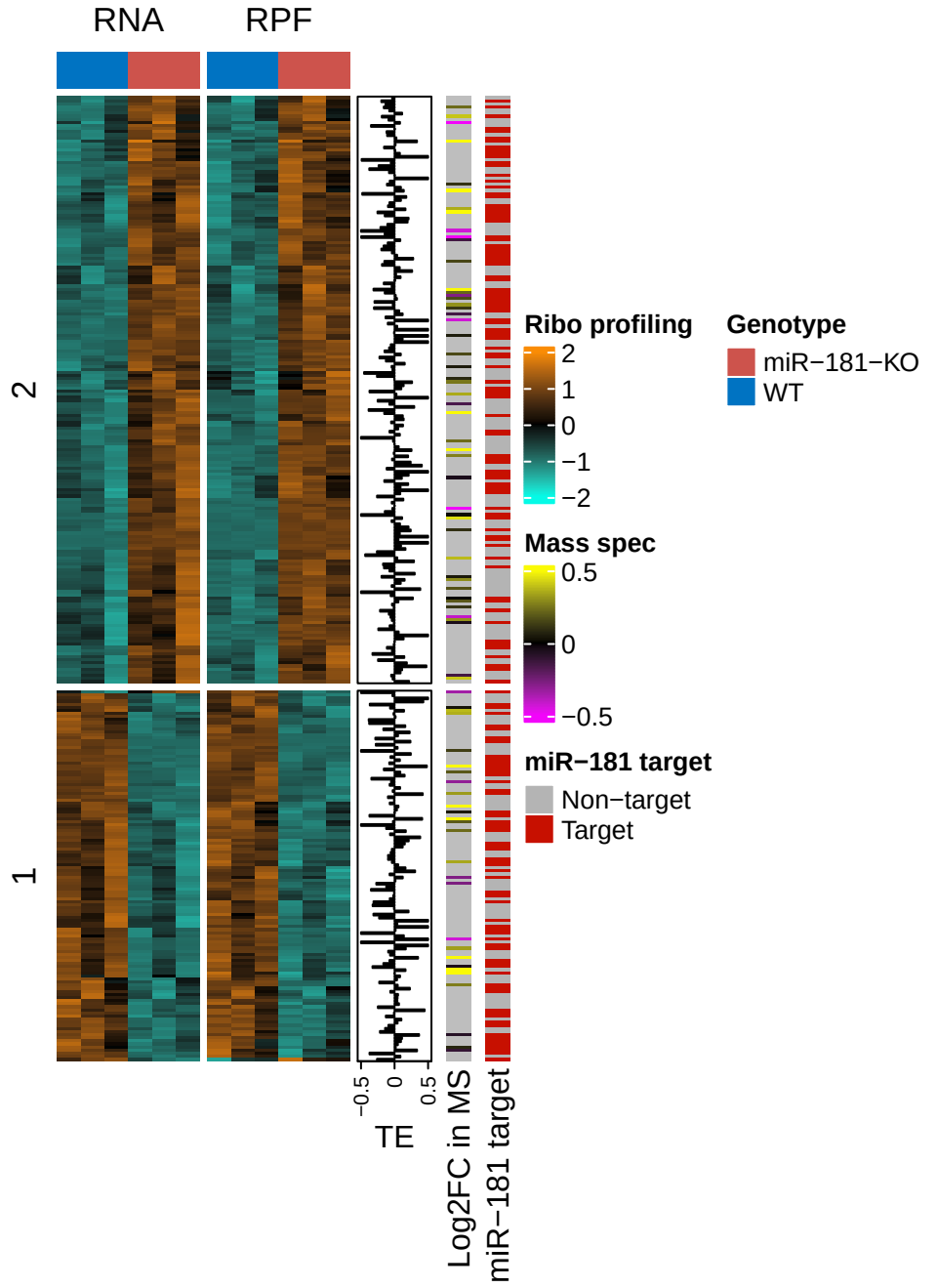
plot

```
set.seed(666)
HRNAsep <- Heatmap(SepScaleframe, show_row_dend = F, show_column_dend = F, show_row_names = F,show_colu
                name = "Ribo profiling", right_annotation = haTE2)
```

```
## Warning: The input is a data frame-like object, convert it to a matrix.
HRNAsep + hmms + hmtarget
```

```
## Warning: Row names of heatmap 2 are not consistent with the main heatmap (1). It
## may lead to wrong conclusion of your data. Please double check.
```

```
## Warning: Row names of heatmap 3 are not consistent with the main heatmap (1). It
## may lead to wrong conclusion of your data. Please double check.
```



Heatmap (x-mean)/mean

different scale that kathi proposed

```
kheatframe <- (heatframe - rowMeans(heatframe))/rowMeans(heatframe)
head(kheatframe)
```

```
##                   WT1 RNA    WT2 RNA    WT3 RNA    KO1 RNA    KO2 RNA    KO3 RNA
## Gm37632       -0.5526462 -0.5645148 -0.6275920  2.0636249  2.1460432  2.6008746
## Cacna1e        1.2695844  1.5016982  1.5049598 -0.4233987 -0.4859078 -0.4797618
## Gnb4          -0.6158363 -0.6101373 -0.6106977  0.6182106  0.6109557  0.6105923
## A430106G13Rik -0.3318564 -0.4194403 -0.5455407  1.6517508  2.0635079  2.6590320
## Ctse          -0.3394868 -0.2230316 -0.4111501 -0.8877611 -0.9145345 -0.8773601
## Hist1h4m      -0.9979493 -1.0000000 -0.9991109 -0.8761352 -0.8533564 -0.8645588
##                   WT1 RPF     WT2 RPF    WT3 RPF    KO1 RPF    KO2 RPF
## Gm37632       -0.9652473 -0.97093769 -0.9644227 -0.7309030 -0.6965433
## Cacna1e       -0.2309098 -0.09614785 -0.1191889 -0.7911438 -0.8255834
## Gnb4          -0.5585308 -0.56392526 -0.5245522  0.5851309  0.6025035
## A430106G13Rik -0.9211265 -0.92786827 -0.9190666 -0.7732122 -0.7536601
## Ctse           2.0372466  1.71569622  1.7602957 -0.6467474 -0.6017391
## Hist1h4m      -0.9705827 -0.97907312 -0.9719667  2.5399000  3.2697578
##                   KO3 RPF
## Gm37632       -0.7377356
## Cacna1e       -0.8242002
## Gnb4           0.4562867
## A430106G13Rik -0.7825195
## Ctse          -0.6114279
## Hist1h4m       2.7030753
```

make heatmaps

make annotations for the heatmap

```
ha5 <- HeatmapAnnotation(Genotype=colnames(kheatframe),col = list(Genotype= c("WT"=farbe1, "miR-181-KO"=
ha6 <- HeatmapAnnotation(Experiment=c(rep("RNA",6), rep("RPF",6)),col = list(Experiment= c("RNA"=farbe2
```

TE annotation

```
TEsel3 = TEframe[unique(TEframe$gene_symbol) %in% rownames(heatframe), c("log2FoldChange")]
names(TEsel3) = TEframe[unique(TEframe$gene_symbol) %in% rownames(heatframe), c("gene_symbol")]

TEsel3 = TEsel3[order(factor(names(TEsel3), levels = rownames(kheatframe)))]


TEsel3[TEsel3 > TElim] = TElim
TEsel3[TEsel3 < -TElim] = -TElim

haTE3 = rowAnnotation(TE = anno_barplot(TEsel3, ylim = c(-TElim,TElim)))
```

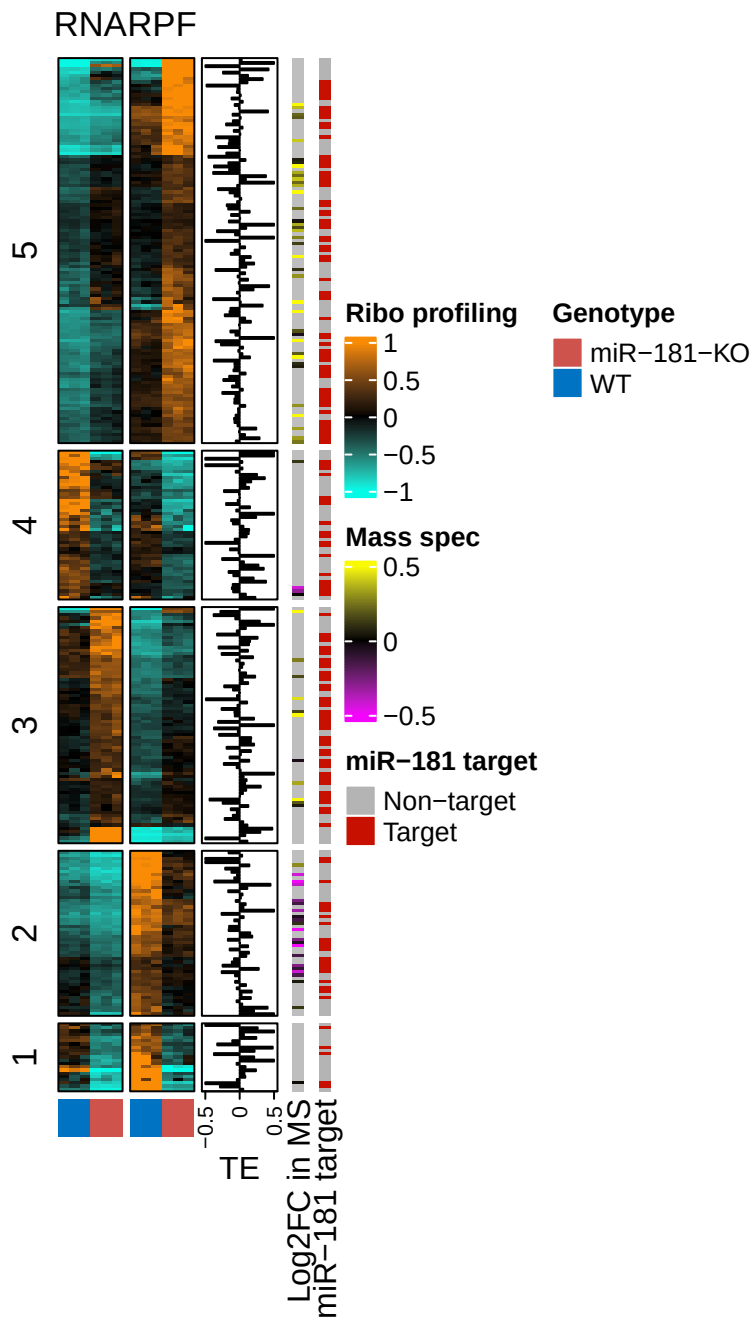
plot

```
kmcent <- 5

set.seed(6)
HRNAk <- Heatmap(kheatframe, show_row_dend = F, show_column_dend = F, show_row_names = F, show_column_na
                 name = "Ribo profiling", right_annotation = haTE3,
```

```
                border = TRUE
                )
```

```
## Warning: The input is a data frame-like object, convert it to a matrix.
```

```
HRNAk + hmms + hmtarget
```



cluster seperately

cluster

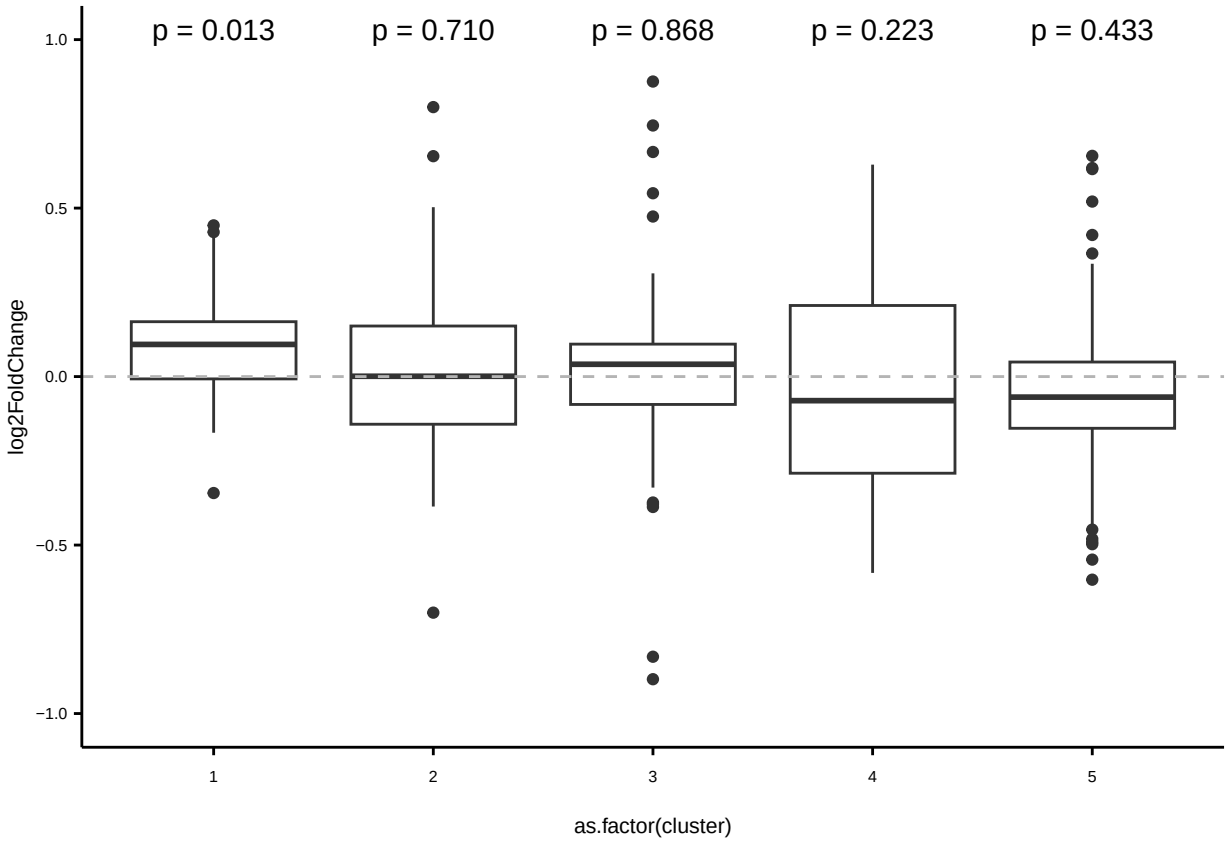
```
set.seed(6)
heat_ks <- kmeans(kheatframe, centers = kmcent)
heat_k_names <- as.data.frame(heat_ks$cluster)
heat_k_names$gene_symbol <- rownames(heat_k_names)
colnames(heat_k_names) <- c("cluster", "gene_symbol")

#merge with TE frame to get those clustered
cTEframe <- merge(TEframe, heat_k_names, by="gene_symbol")
```

plot TE by cluster

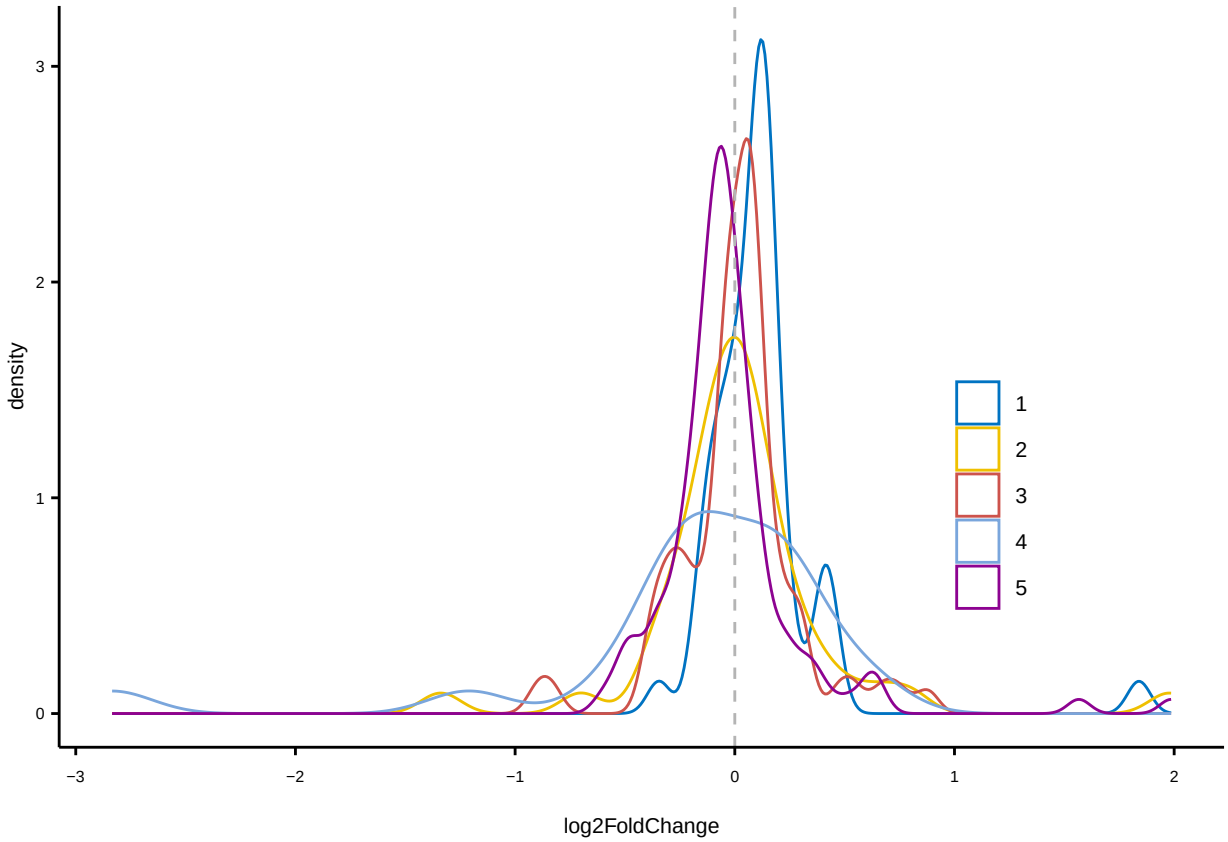
```
tecplot <- ggplot(cTEframe, aes(x=as.factor(cluster), y=log2FoldChange)) +
  geom_boxplot() +
  coord_cartesian(ylim = c(-1,1)) +
  geom_hline(yintercept = 0, colour=farbeneg,linetype = "dashed") +
  stat_compare_means(label = "p.format", method = "t.test", ref.group = ".all.", label.y=1)   +
  theme_paper()
```

```
## Warning: The `size` argument of `element_line()` is deprecated as of ggplot2 3.4.0.
## i Please use the `linewidth` argument instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```

```
tecplot
```




```
techist <- ggplot(cTEframe, aes(colour=as.factor(cluster), x=log2FoldChange)) +
  geom_density() +
  scale_colour_manual(values = c(farbe1, farbe2, farbe3, farbe4, RPFncol)) +
  geom_vline(xintercept = 0, colour=farbeneg,linetype = "dashed") +
  theme_paper()


techist
```




```
#export
pdf("D:/Krueger_Lab/Publications/miR181_paper/Figure3/TEbox_clustered.pdf", height = 2, width = 2)
tecplot
dev.off()
```

```
## pdf
##   2
```

```
pdf("D:/Krueger_Lab/Publications/miR181_paper/Figure3/TEhist_clustered.pdf", height = 2, width = 2)
techist
dev.off()
```

```
## pdf
##   2
```

#session info

```
sessionInfo()
```

```
## R version 4.2.3 (2023-03-15 ucrt)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows 10 x64 (build 19045)
##
## Matrix products: default
##
## locale:
## [1] LC_COLLATE=German_Germany.utf8  LC_CTYPE=German_Germany.utf8
## [3] LC_MONETARY=German_Germany.utf8 LC_NUMERIC=C
## [5] LC_TIME=German_Germany.utf8
```

```
##
## attached base packages:
## [1] stats4    grid      stats     graphics  grDevices utils     datasets
## [8] methods   base
##
## other attached packages:
##  [1] GenomicRanges_1.50.2  GenomeInfoDb_1.34.9   IRanges_2.32.0
##  [4] S4Vectors_0.36.2      BiocGenerics_0.44.0   ggpubr_0.6.0
##  [7] xlsx_0.6.5            eulerr_7.0.0          dplyr_1.1.2
## [10] circlize_0.4.15       ggplot2_3.4.2         ComplexHeatmap_2.15.2
##
## loaded via a namespace (and not attached):
##  [1] Rcpp_1.0.10            tidyr_1.3.0            png_0.1-8
##  [4] xlsxjars_0.6.1         digest_0.6.31          foreach_1.5.2
##  [7] utf8_1.2.3             R6_2.5.1               backports_1.4.1
## [10] evaluate_0.21          highr_0.10             pillar_1.9.0
## [13] zlibbioc_1.44.0        GlobalOptions_0.1.2    rlang_1.1.0
## [16] rstudioapi_0.14        car_3.1-2              magick_2.7.4
## [19] GetoptLong_1.0.5       rmarkdown_2.21         labeling_0.4.2
## [22] polyclip_1.10-4        RCurl_1.98-1.12        munsell_0.5.0
## [25] broom_1.0.4            polylabelr_0.2.0       compiler_4.2.3
## [28] xfun_0.39              pkgconfig_2.0.3        shape_1.4.6
## [31] htmltools_0.5.4        tidyselect_1.2.0       tibble_3.2.1
## [34] GenomeInfoDbData_1.2.9 codetools_0.2-19       matrixStats_0.63.0
## [37] fansi_1.0.4            crayon_1.5.2           withr_2.5.0
## [40] bitops_1.0-7           gtable_0.3.3           lifecycle_1.0.3
## [43] magrittr_2.0.3         scales_1.2.1           cli_3.6.0
## [46] carData_3.0-5          farver_2.1.1           XVector_0.38.0
## [49] ggsignif_0.6.4         doParallel_1.0.17      generics_0.1.3
## [52] vctrs_0.6.2            rjson_0.2.21           RColorBrewer_1.1-3
## [55] Cairo_1.6-0            iterators_1.0.14       tools_4.2.3
## [58] glue_1.6.2             purrr_1.0.1            abind_1.4-5
## [61] parallel_4.2.3         fastmap_1.1.1          yaml_2.3.7
## [64] clue_0.3-64            colorspace_2.1-0       cluster_2.1.4
## [67] rstatix_0.7.2          rJava_1.0-6            knitr_1.42
```