

# Fig4 ECDF plots

Nikita Verheyden

2023-05-16

## Setup

dir

```
setwd("D:/Krueger_Lab/Publications/miR181_paper/Figure4")
```

## packages

```
source("D:/Krueger_Lab/Publications/miR181_paper/Supporting_scripts/themes/theme_paper.R")
library(ggplot2)
library(rtracklayer)
```

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

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

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

```
##
```

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

```
## 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:base':
```

```
##
```

```
##      expand.grid, I, unname
```

```
## Loading required package: IRanges
```

```
##
```

```
## Attaching package: 'IRanges'
```

```
## The following object is masked from 'package:grDevices':
##
## windows
## Loading required package: GenomeInfoDb
library(dplyr)

##
## Attaching package: 'dplyr'
## The following objects are masked from 'package:GenomicRanges':
##
## intersect, setdiff, union
## The following object is masked from 'package:GenomeInfoDb':
##
## intersect
## The following objects are masked from 'package:IRanges':
##
## collapse, desc, intersect, setdiff, slice, union
## The following objects are masked from 'package:S4Vectors':
##
## first, intersect, rename, setdiff, setequal, union
## The following objects are masked from 'package:BiocGenerics':
##
## combine, intersect, setdiff, union
## The following objects are masked from 'package:stats':
##
## filter, lag
## The following objects are masked from 'package:base':
##
## intersect, setdiff, setequal, union
```

## data

```
#Ribo profiling
RNA <- read.csv("D:/Krueger_Lab/Publications/miR181_paper/Figure3/RNA_masterframe.csv")
RPF <- read.csv("D:/Krueger_Lab/Publications/miR181_paper/Figure3/RPF_masterframe.csv")

#load the gtf file to compare genes
gff23 <- import.gff3("D:/Krueger_Lab/Ribo_Profiling/run15112022M23/ref_genome/gencode.vM23.annotation.gff3")

#targets
larget <- readRDS("D:/Krueger_Lab/Publications/miR181_paper/Figure3/mir181_bs_with_seeds.rds")
largetframe <- as.data.frame(larget)

#targets with introns and other
tject <- readRDS("D:/Krueger_Lab/Publications/miR181_paper/Figure1/mir181_binding_sites__venn_types/mir181_binding_sites__venn_types.rds")
names(tject) <- 1:length(tject$geneName)
tframe <- as.data.frame(tject)
head(tframe)

## seqnames start end width strand scoreSum scoreMean scoreMax
```

```

## 1 chr1 6245651 6245657 7 + 9.52553 4.762765 6.00678
## 2 chr1 6248341 6248347 7 + 92.68921 23.172303 48.76900
## 3 chr1 6248857 6248863 7 + 14.07133 7.035665 7.04425
## 4 chr1 6248918 6248924 7 + 38.91451 12.971503 20.65080
## 5 chr1 7170481 7170487 7 + 66.92218 13.384436 25.84490
## 6 chr1 9899605 9899611 7 + 25.15963 6.289907 8.61019
## geneType geneName geneID region BS_ID mir_IP
## 1 protein_coding Rb1cc1 ENSMUSG00000025907 cds 5 mmu-miR-181a-5p
## 2 protein_coding Rb1cc1 ENSMUSG00000025907 cds 8 mmu-miR-181a-5p
## 3 protein_coding Rb1cc1 ENSMUSG00000025907 cds 10 mmu-miR-181a-5p
## 4 protein_coding Rb1cc1 ENSMUSG00000025907 cds 11 mmu-miR-181a-5p
## 5 protein_coding Pcmt1 ENSMUSG00000051285 utr3 19 mmu-miR-181a-5p
## 6 protein_coding Sgk3 ENSMUSG00000025915 utr3 23 mmu-miR-181a-5p
## n_mir181 n_mir181a n_mir181b n_mir181c n_mir181d set WT KO
## 1 1 1 0 0 0 ago_bs_mir181_chi 1 1
## 2 5 5 0 0 0 ago_bs_mir181_chi 1 1
## 3 6 6 0 0 0 ago_bs_mir181_chi 1 0
## 4 6 6 0 0 0 ago_bs_mir181_chi 1 1
## 5 4 4 0 0 0 ago_bs_mir181_chi 1 1
## 6 1 1 0 0 0 ago_bs_mir181_chi NA NA
## geneID.2 geneName.1 region.1 counts.bs.1_KO counts.bs.2_KO
## 1 ENSMUSG00000025907 Rb1cc1 cds 4 3
## 2 ENSMUSG00000025907 Rb1cc1 cds 28 32
## 3 ENSMUSG00000025907 Rb1cc1 cds 13 11
## 4 ENSMUSG00000025907 Rb1cc1 cds 15 15
## 5 ENSMUSG00000051285 Pcmt1 utr3 12 22
## 6 <NA> <NA> <NA> NA NA
## counts.bs.3_KO counts.bs.4_WT counts.bs.5_WT counts.bs.6_WT
## 1 3 3 10 3
## 2 27 46 41 20
## 3 4 22 13 12
## 4 10 33 20 18
## 5 14 16 20 9
## 6 NA NA NA NA
## geneID.1 counts.bg.1_KO counts.bg.2_KO counts.bg.3_KO
## 1 ENSMUSG00000025907 1609 1973 1250
## 2 ENSMUSG00000025907 1609 1973 1250
## 3 ENSMUSG00000025907 1609 1973 1250
## 4 ENSMUSG00000025907 1609 1973 1250
## 5 ENSMUSG00000051285 1355 1706 1064
## 6 <NA> NA NA NA
## counts.bg.4_WT counts.bg.5_WT counts.bg.6_WT resBs.baseMean
## 1 2638 2231 1352 92.10645
## 2 2638 2231 1352 281.53271
## 3 2638 2231 1352 145.51107
## 4 2638 2231 1352 186.74162
## 5 1654 1348 755 151.36245
## 6 NA NA NA NA
## resBs.log2FoldChange resBs.lfcSE resBs.stat resBs.pvalue resBs.padj
## 1 -0.1093039 0.5923673 0.03419066 0.8533018 0.9652601
## 2 0.2749428 0.2351157 1.35874137 0.2437557 0.6729889
## 3 -0.1805519 0.3623758 0.25017050 0.6169550 0.8961239
## 4 -0.2606282 0.3062717 0.73169661 0.3923338 0.7868678
## 5 0.1466485 0.3122905 0.22052922 0.6386370 0.9013566

```

```
## 6          NA          NA          NA          NA          NA
##   resBg.baseMean resBg.log2FoldChange resBg.lfcSE resBg.stat resBg.pvalue
## 1          NA          NA          NA          NA          NA
## 2          NA          NA          NA          NA          NA
## 3          NA          NA          NA          NA          NA
## 4          NA          NA          NA          NA          NA
## 5          NA          NA          NA          NA          NA
## 6          NA          NA          NA          NA          NA
##   resBg.padj tpm.counts.bg.1_K0 tpm.counts.bg.2_K0 tpm.counts.bg.3_K0
## 1          NA          133.7259          117.9980          129.8669
## 2          NA          133.7259          117.9980          129.8669
## 3          NA          133.7259          117.9980          129.8669
## 4          NA          133.7259          117.9980          129.8669
## 5          NA          248.6210          225.2505          244.0445
## 6          NA          NA          NA          NA
##   tpm.counts.bg.4_WT tpm.counts.bg.5_WT tpm.counts.bg.6_WT
## 1          139.8635          146.2855          163.5360
## 2          139.8635          146.2855          163.5360
## 3          139.8635          146.2855          163.5360
## 4          139.8635          146.2855          163.5360
## 5          193.5994          195.1330          201.6149
## 6          NA          NA          NA
##   BS_ID.1 tpm_support_K0 tpm_support_WT tpm_supported down
## 1 ENSMUSG00000025907.bs5          3          3          TRUE FALSE
## 2 ENSMUSG00000025907.bs8          3          3          TRUE FALSE
## 3 ENSMUSG00000025907.bs10         3          3          TRUE FALSE
## 4 ENSMUSG00000025907.bs11         3          3          TRUE FALSE
## 5 ENSMUSG00000051285.bs4          3          3          TRUE FALSE
## 6          <NA>          NA          NA          NA    NA
```

#### #MMSat4

```
repeat_masker <- readRDS("D:/Krueger_Lab/Publications/miR181_paper/Figure2/MMSat4/repeat_masker.rds")
MMSAT4 <- repeat_masker[repeat_masker$repName == "MMSAT4"]
```

## colours

```
#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 <- "#008EAOFF"
farbe14 <- "#8A4198FF"
farbe15 <- "#5A9599FF"
farbe16 <- "#FF6348FF"
```

```

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

```

## inspect targetdata

We're keeping all of those targets for now but will analyze the in ecdf plots

```
table(targetframe$set)
```

```
##
##          ago_bs_mir181_chi ago_bs_mir181_chi&mir181_enriched
##                      5815                      1082
##          mir181_enriched
##                      3576

```

```
colnames(targetframe)
```

```
## [1] "seqnames"          "start"
## [3] "end"               "width"
## [5] "strand"            "scoreSum"
## [7] "scoreMean"         "scoreMax"
## [9] "geneType"          "geneName"
## [11] "geneID"            "region"
## [13] "mir_IP"            "n_mir181"
## [15] "n_mir181a"         "n_mir181b"
## [17] "n_mir181c"         "n_mir181d"
## [19] "set"               "mir181BS_ID"
## [21] "WT"                "KO"
## [23] "geneID.2"          "geneName.1"
## [25] "region.1"          "counts.bs.1_KO"
## [27] "counts.bs.2_KO"    "counts.bs.3_KO"
## [29] "counts.bs.4_WT"    "counts.bs.5_WT"
## [31] "counts.bs.6_WT"    "geneID.1"
## [33] "counts.bg.1_KO"    "counts.bg.2_KO"
## [35] "counts.bg.3_KO"    "counts.bg.4_WT"
## [37] "counts.bg.5_WT"    "counts.bg.6_WT"
## [39] "resBs.baseMean"    "resBs.log2FoldChange"
## [41] "resBs.lfcSE"        "resBs.stat"
## [43] "resBs.pvalue"       "resBs.padj"
## [45] "resBg.baseMean"    "resBg.log2FoldChange"
## [47] "resBg.lfcSE"        "resBg.stat"
## [49] "resBg.pvalue"       "resBg.padj"
## [51] "tpm.counts.bg.1_KO" "tpm.counts.bg.2_KO"
## [53] "tpm.counts.bg.3_KO" "tpm.counts.bg.4_WT"
## [55] "tpm.counts.bg.5_WT" "tpm.counts.bg.6_WT"
## [57] "BS_ID"             "tpm_support_KO"
## [59] "tpm_support_WT"    "tpm_supported"
## [61] "down"              "all_seeds_200down"
## [63] "first_seed_200down.start" "first_seed_200down.end"
## [65] "first_seed_200down.width" "first_seed_200down.type"
## [67] "first_seed_200down.wobble" "seed_repetitions.200down"
## [69] "seed_repetitions.200down.wobble" "all_seeds_200up"

```

```
## [71] "first_seed_200up.start"          "first_seed_200up.end"
## [73] "first_seed_200up.width"         "first_seed_200up.type"
## [75] "first_seed_200up.wobble"        "seed_repetitions.200up"
## [77] "seed_repetitions.200up.wobble"

#remove duplicates but keep NAs (and ask melina where the NAs come from)
targetframe <- targetframe[!(duplicated(targetframe$BS_ID) & !is.na(targetframe$BS_ID)), ]

## adjust type column
# Add "No seed" as a level to the factor column
targetframe$first_seed_200down.type <- factor(targetframe$first_seed_200down.type, levels = c(levels(targetframe$first_seed_200down.type), "No seed"))

# Replace NAs with "No seed"
targetframe$first_seed_200down.type[is.na(targetframe$first_seed_200down.type)] <- "No seed"
table(targetframe$first_seed_200down.type)

##
##      seed_8mer seed_7mer_m8 seed_7mer_a1      seed_6mer      No seed
##          245         1354          321         1054         7490

## adjust wobble column
targetframe$first_seed_200down.wobble[is.na(targetframe$first_seed_200down.wobble)] <- "No seed"
```

## ECDF plots

each code chunk is a split of the main target table that is then used for a specific ecdf plot

### targetfiles

```
#old target with introns
#RNA
RNA$oldtarg <- "Non-target"
RNA$oldtarg[RNA$gene_symbol %in% tframe$geneName] <- "Target"
#RPF
RPF$oldtarg <- "Non-target"
RPF$oldtarg[RPF$gene_symbol %in% tframe$geneName] <- "Target"

#giant frame
#RNA
RNA$target <- "Non-target"
RNA$target[RNA$gene_symbol %in% targetframe$geneName] <- "Target"
#RPF
RPF$target <- "Non-target"
RPF$target[RPF$gene_symbol %in% targetframe$geneName] <- "Target"

#ECDF
#old targets with introns
#RNA
targetoldECDFRNA <- ggplot(RNA, aes(as.numeric(log2FoldChange), colour=factor(oldtarg, levels = c("Non-target", "Target")))) +
  stat_ecdf(geom="step", linewidth=2) +
  scale_colour_manual(values = c("black", "red")) +
  coord_cartesian(xlim = c(-0.75, 0.75)) +
  theme_paper() +
```

```

scale_y_continuous("Cumulative density") + scale_x_continuous("log2FoldChange") +
ggtitle("RNA targetset with introns")

## 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.

targetoldECDFRNA

#RPF
targetoldECDFRPF <- ggplot(RPF, aes(as.numeric(log2FoldChange), colour=factor(olddtarg, levels = c("Non-
stat_ecdf(geom="step", linewidth=2) +
scale_colour_manual(values = c("black", farbe1)) +
coord_cartesian(xlim = c(-0.75, 0.75)) +
theme_paper() +
scale_y_continuous("Cumulative density") + scale_x_continuous("log2FoldChange") +
ggtitle("RPF targetset with introns")

targetoldECDFRPF

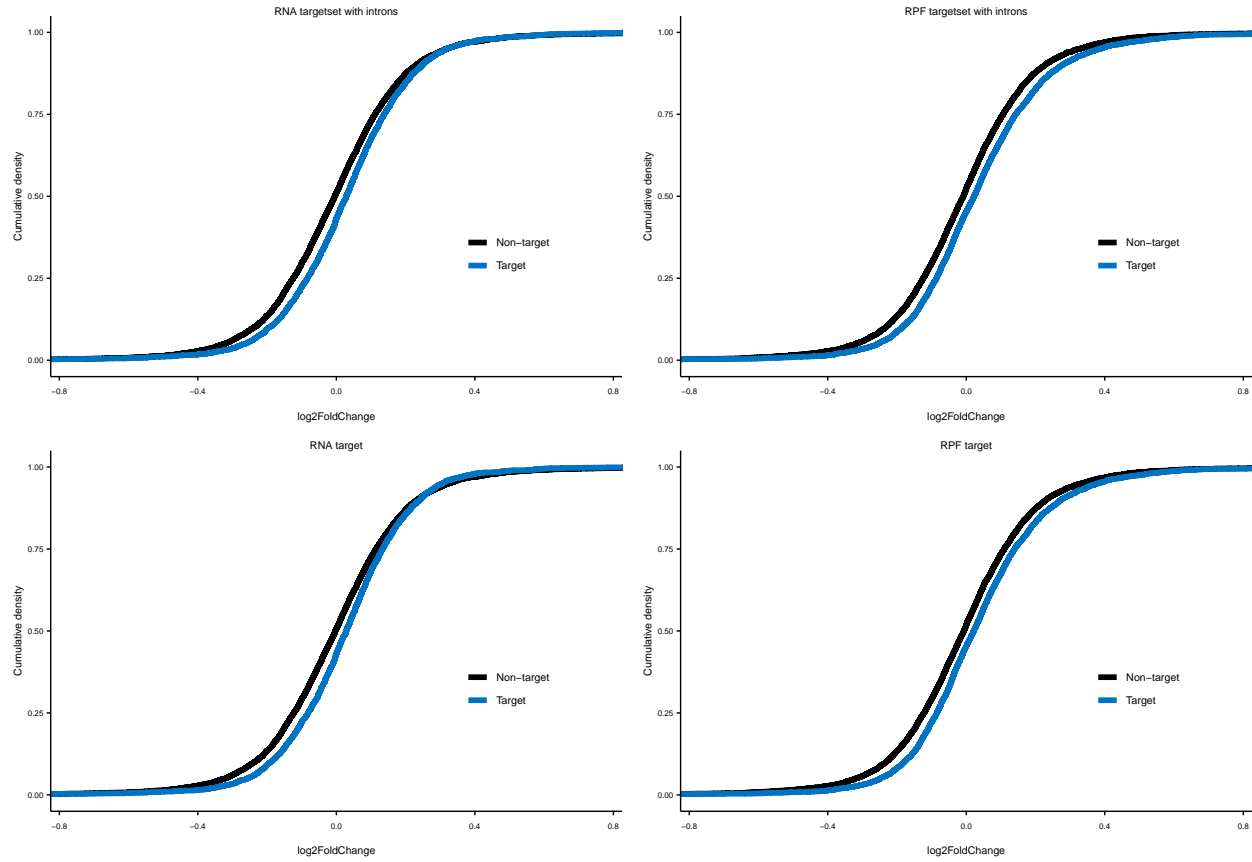
#targets
#RNA
targetECDFRNA <- ggplot(RNA, aes(as.numeric(log2FoldChange), colour=factor(target, levels = c("Non-targ
stat_ecdf(geom="step", linewidth=2) +
scale_colour_manual(values = c("black", farbe1)) +
coord_cartesian(xlim = c(-0.75, 0.75)) +
theme_paper() +
scale_y_continuous("Cumulative density") + scale_x_continuous("log2FoldChange") +
ggtitle("RNA target")

targetECDFRNA

#RPF
targetECDFRPF <- ggplot(RPF, aes(as.numeric(log2FoldChange), colour=factor(target, levels = c("Non-targ
stat_ecdf(geom="step", linewidth=2) +
scale_colour_manual(values = c("black", farbe1)) +
coord_cartesian(xlim = c(-0.75, 0.75)) +
theme_paper() +
scale_y_continuous("Cumulative density") + scale_x_continuous("log2FoldChange") +
ggtitle("RPF target")

targetECDFRPF

```



## datasets

### #RNA

```
RNA$targetset <- "Non-target"
RNA$targetset[RNA$gene_symbol %in% targetframe[targetframe$set == "ago_bs_mir181_chi", "geneName"]] <- "ago_bs_mir181_chi"
RNA$targetset[RNA$gene_symbol %in% targetframe[targetframe$set == "mir181_enriched", "geneName"]] <- "mir181_enriched"
RNA$targetset[RNA$gene_symbol %in% targetframe[targetframe$set == "ago_bs_mir181_chi&mir181_enriched", "geneName"]] <- "ago_bs_mir181_chi&mir181_enriched"

table(RNA$targetset)
```

```
##
## ago_bs_mir181_chi      both  mir181_enriched  Non-target
##                   783      666      1522      10330
```

### #RPF

```
RPF$targetset <- "Non-target"
RPF$targetset[RPF$gene_symbol %in% targetframe[targetframe$set == "ago_bs_mir181_chi", "geneName"]] <- "ago_bs_mir181_chi"
RPF$targetset[RPF$gene_symbol %in% targetframe[targetframe$set == "mir181_enriched", "geneName"]] <- "mir181_enriched"
RPF$targetset[RPF$gene_symbol %in% targetframe[targetframe$set == "ago_bs_mir181_chi&mir181_enriched", "geneName"]] <- "ago_bs_mir181_chi&mir181_enriched"

table(RPF$targetset)
```

```
##
## ago_bs_mir181_chi      both  mir181_enriched  Non-target
##                   782      666      1509      8412
```



```

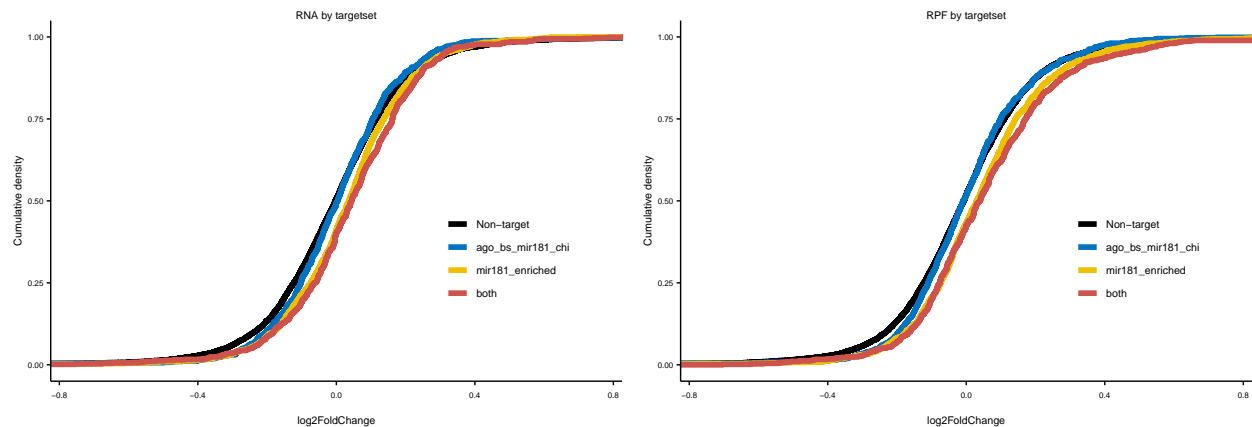
# ecdf plots
#RNA
setECDFRNA <- ggplot(RNA, aes(as.numeric(log2FoldChange), colour=factor(targetset, levels = c("Non-target", "ago_bs_mir181_chi", "mir181_enriched", "both")))) +
  stat_ecdf(geom="step", linewidth=2) +
  scale_colour_manual(values = c("black", farbe1, farbe2, farbe3)) +
  coord_cartesian(xlim = c(-0.75, 0.75)) +
  theme_paper() +
  scale_y_continuous("Cumulative density") + scale_x_continuous("log2FoldChange") +
  ggtitle("RNA by targetset")

setECDFRNA

#RPF
setECDFRPF <- ggplot(RPF, aes(as.numeric(log2FoldChange), colour=factor(targetset, levels = c("Non-target", "ago_bs_mir181_chi", "mir181_enriched", "both")))) +
  stat_ecdf(geom="step", linewidth=2) +
  scale_colour_manual(values = c("black", farbe1, farbe2, farbe3)) +
  coord_cartesian(xlim = c(-0.75, 0.75)) +
  theme_paper() +
  scale_y_continuous("Cumulative density") + scale_x_continuous("log2FoldChange") +
  ggtitle("RPF by targetset")

setECDFRPF

```



## region (single targets)

```

#get number of binding sites per gene to be able to sort for singles
bsnum <- as.data.frame(table(targetframe$geneName))
colnames(bsnum) <- c("geneName", "BS_number")

#RNA
RNA$region_single <- "Non-target"
RNA$region_single[RNA$gene_symbol %in% targetframe[targetframe$region == "utr5", "geneName"]] <- "5'UTR"
RNA$region_single[RNA$gene_symbol %in% targetframe[targetframe$region == "cds", "geneName"]] <- "CDS"
RNA$region_single[RNA$gene_symbol %in% targetframe[targetframe$region == "utr3", "geneName"]] <- "3'UTR"
RNA$region_single[RNA$gene_symbol %in% bsnum[bsnum$BS_number > 1, "geneName"]] <- "multiple"

table(RNA$region_single)

```

```
##
```

```
##      3'UTR      5'UTR      CDS      multiple Non-target
##      660       86      451      1774      10330
```

```
#RPF
```

```
RPF$region_single <- "Non-target"
RPF$region_single[RPF$gene_symbol %in% targetframe[targetframe$region == "utr5", "geneName"]] <- "5'UTR"
RPF$region_single[RPF$gene_symbol %in% targetframe[targetframe$region == "cds", "geneName"]] <- "CDS"
RPF$region_single[RPF$gene_symbol %in% targetframe[targetframe$region == "utr3", "geneName"]] <- "3'UTR"
RPF$region_single[RPF$gene_symbol %in% bsnum[bsnum$BS_number > 1, "geneName"]] <- "multiple"
```

```
table(RPF$region_single)
```

```
##
##      3'UTR      5'UTR      CDS      multiple Non-target
##      657       84      450      1766      8412
```

```
# ECDF plots
```

```
#RNA
```

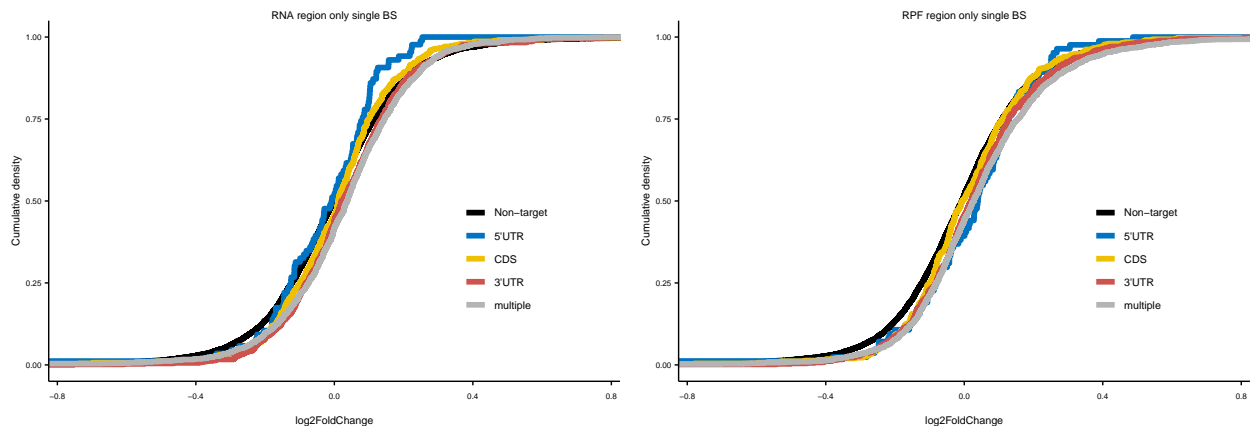
```
regsingECDFRNA <- ggplot(RNA, aes(as.numeric(log2FoldChange), colour=factor(region_single, levels = c("Non-target", "5'UTR", "CDS", "3'UTR", "multiple")),
  stat_ecdf(geom="step", linewidth=2) +
  scale_colour_manual(values = c("black", farbe1, farbe2, farbe3, farbeneg)) +
  coord_cartesian(xlim = c(-0.75, 0.75)) +
  theme_paper() +
  scale_y_continuous("Cumulative density") + scale_x_continuous("log2FoldChange") +
  ggtitle("RNA region only single BS")
```

```
regsingECDFRNA
```

```
#RPF
```

```
regsingECDFRPF <- ggplot(RPF, aes(as.numeric(log2FoldChange), colour=factor(region_single, levels = c("Non-target", "5'UTR", "CDS", "3'UTR", "multiple")),
  stat_ecdf(geom="step", linewidth=2) +
  scale_colour_manual(values = c("black", farbe1, farbe2, farbe3, farbeneg)) +
  coord_cartesian(xlim = c(-0.75, 0.75)) +
  theme_paper() +
  scale_y_continuous("Cumulative density") + scale_x_continuous("log2FoldChange") +
  ggtitle("RPF region only single BS")
```

```
regsingECDFRPF
```



## number of target sites

```
colnames(bsnum) <- c("gene_symbol", "BS_number")

#RNA
RNAnum <- left_join(RNA, bsnum, by="gene_symbol")
RNAnum$BS_number[is.na(RNAnum$BS_number)] <- "Non-target"
RNAnum$BS_num_plot <- ifelse(RNAnum$BS_number == "Non-target", "Non-target",
                             ifelse(RNAnum$BS_number == 1, "One bs",
                                     ifelse(RNAnum$BS_number == 2, "Two bs", "More")))

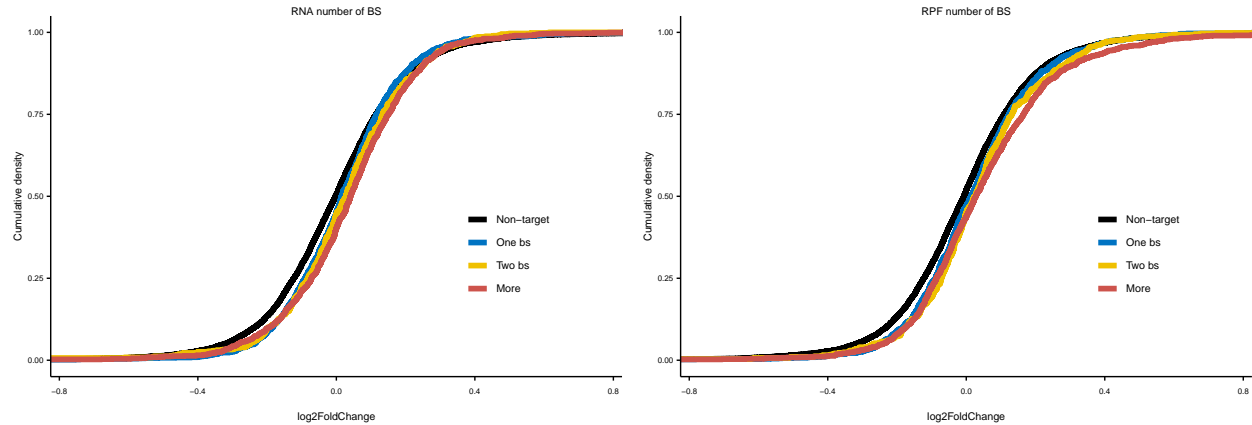
#RPF
RPFnum <- left_join(RPF, bsnum, by="gene_symbol")
RPFnum$BS_number[is.na(RPFnum$BS_number)] <- "Non-target"
RPFnum$BS_num_plot <- ifelse(RPFnum$BS_number == "Non-target", "Non-target",
                             ifelse(RPFnum$BS_number == 1, "One bs",
                                     ifelse(RPFnum$BS_number == 2, "Two bs", "More")))

#ecdf plots
#RNA
numECDFRNA <- ggplot(RNAnum, aes(as.numeric(log2FoldChange), colour=factor(BS_num_plot, levels = c("Non-
  stat_ecdf(geom="step", linewidth=2) +
  scale_colour_manual(values = c("black", farbe1, farbe2, farbe3)) +
  coord_cartesian(xlim = c(-0.75, 0.75)) +
  theme_paper() +
  scale_y_continuous("Cumulative density") + scale_x_continuous("log2FoldChange") +
  ggtitle("RNA number of BS")

numECDFRNA

#RPF
numECDFRPF <- ggplot(RPFnum, aes(as.numeric(log2FoldChange), colour=factor(BS_num_plot, levels = c("Non-
  stat_ecdf(geom="step", linewidth=2) +
  scale_colour_manual(values = c("black", farbe1, farbe2, farbe3)) +
  coord_cartesian(xlim = c(-0.75, 0.75)) +
  theme_paper() +
  scale_y_continuous("Cumulative density") + scale_x_continuous("log2FoldChange") +
  ggtitle("RPF number of BS")

numECDFRPF
```



## MMSat4

```
mmsat4frame <- as.data.frame(subsetByOverlaps(gff23, MMSAT4))
```

*#RNA*

```
RNA$tvsmmsat4 <- "Non-target"
```

```
RNA$tvsmmsat4[RNA$gene_symbol %in% mmsat4frame$gene_name] <- "MMSat4"
```

```
RNA$tvsmmsat4[RNA$gene_symbol %in% tframe$geneName] <- "miR-181 target"
```

```
RNA$tvsmmsat4[RNA$gene_symbol %in% tframe$geneName & RNA$gene_symbol %in% mmsat4frame$gene_name] <- "both"
```

```
table(RNA$tvsmmsat4)
```

```
##
```

```
##          both miR-181 target          MMSat4          Non-target
```

```
##          103          3441          141          9616
```

*#RPF*

```
RPF$tvsmmsat4 <- "Non-target"
```

```
RPF$tvsmmsat4[RPF$gene_symbol %in% mmsat4frame$gene_name] <- "MMSat4"
```

```
RPF$tvsmmsat4[RPF$gene_symbol %in% tframe$geneName] <- "miR-181 target"
```

```
RPF$tvsmmsat4[RPF$gene_symbol %in% tframe$geneName & RPF$gene_symbol %in% mmsat4frame$gene_name] <- "both"
```

```
table(RPF$tvsmmsat4)
```

```
##
```

```
##          both miR-181 target          MMSat4          Non-target
```

```
##          100          3405          131          7733
```

*#RNA*

```
tolECDFRNA <- ggplot(RNA, aes(as.numeric(log2FoldChange), colour=factor(tvsmmsat4, levels = c("Non-target", "MMSat4", "miR-181 target", "both")))) +
  stat_ecdf(geom="step", linewidth=2) +
  scale_colour_manual(values = c("black", "blue", "yellow", "red")) +
  coord_cartesian(xlim = c(-0.75, 0.75)) +
  theme_paper() +
  scale_y_continuous("Cumulative density") + scale_x_continuous("log2FoldChange") +
  ggtitle("RNA Mmsat4 vs target")
```

```
tolECDFRNA
```

*#RPF*

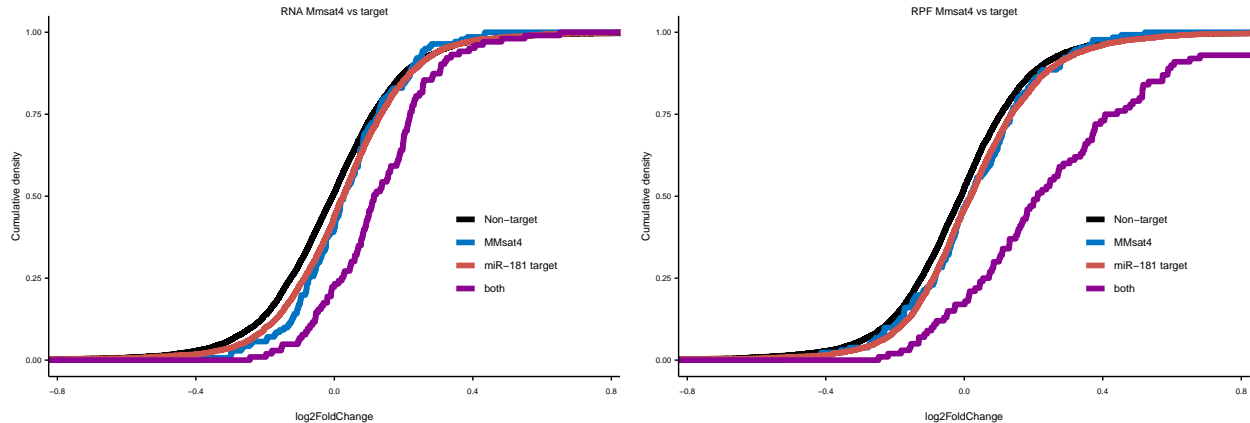
```
tolECDFRPF <- ggplot(RPF, aes(as.numeric(log2FoldChange), colour=factor(tvsmmsat4, levels = c("Non-target", "MMSat4", "miR-181 target", "both")))) +
  stat_ecdf(geom="step", linewidth=2) +
```

```

scale_colour_manual(values = c("black", farbe1, farbe3, RPFncol)) +
coord_cartesian(xlim = c(-0.75, 0.75)) +
theme_paper() +
scale_y_continuous("Cumulative density") + scale_x_continuous("log2FoldChange") +
ggtitle("RPF Mmsat4 vs target")

```

toIECDFRPF



## Mmsat4 in 3'UTR and CDS

Here we took all targets (genes) that also contain a MMsat4 element and split them by 3'UTR or UTR. The location of the MMsat4 within the gene is not considered in this analysis.

```
rframe <- tframe[tframe$geneName %in% mmsat4frame$gene_name,]
```

*#RNA*

```

RNA$regMmsat4 <- "Non-target"
RNA$regMmsat4[RNA$gene_symbol %in% rframe[rframe$region == "cds", "geneName"]] <- "cds"
RNA$regMmsat4[RNA$gene_symbol %in% rframe[rframe$region == "utr3", "geneName"]] <- "utr3"
RNA$regMmsat4[RNA$gene_symbol %in% rframe[rframe$region == "utr3", "geneName"] &
  RNA$gene_symbol %in% rframe[rframe$region == "cds", "geneName"]] <- "both"

```

```
table(RNA$regMmsat4)
```

```
##
##      both      cds Non-target      utr3
##      159       51      13069       22
```

*#RPF*

```

RPF$regMmsat4 <- "Non-target"
RPF$regMmsat4[RPF$gene_symbol %in% rframe[rframe$region == "cds", "geneName"]] <- "cds"
RPF$regMmsat4[RPF$gene_symbol %in% rframe[rframe$region == "utr3", "geneName"]] <- "utr3"
RPF$regMmsat4[RPF$gene_symbol %in% rframe[rframe$region == "utr3", "geneName"] &
  RPF$gene_symbol %in% rframe[rframe$region == "cds", "geneName"]] <- "both"

```

```
table(RPF$regMmsat4)
```

```
##
##      both      cds Non-target      utr3
##      15       49      11283       22
```

```

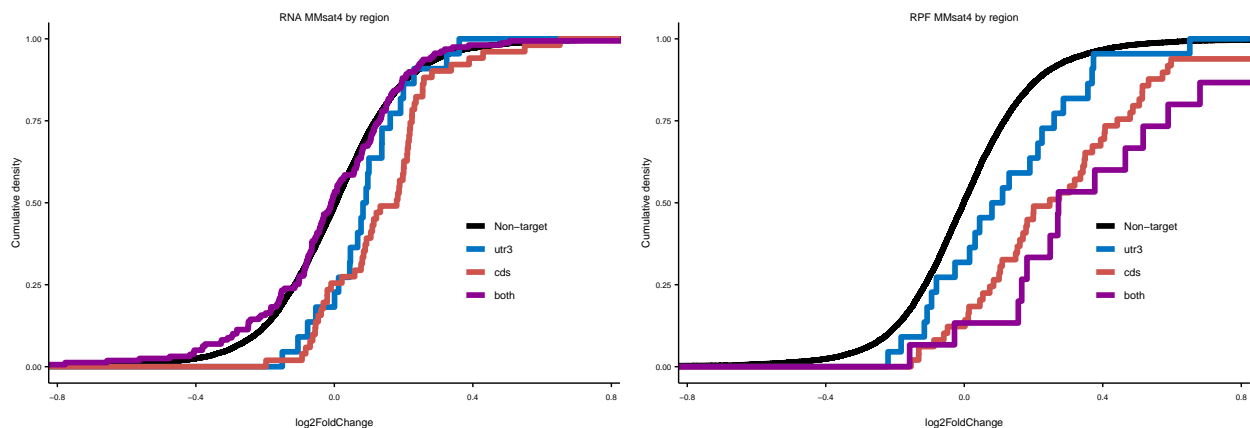
#ecdf plots
#RNA
posECDFRNA <- ggplot(RNA, aes(as.numeric(log2FoldChange), colour=factor(regMMsat4, levels = c("Non-target", "utr3", "cds", "both")))) +
  stat_ecdf(geom="step", linewidth=2) +
  scale_colour_manual(values = c("black", farbe1, farbe3, RPFncol)) +
  coord_cartesian(xlim = c(-0.75, 0.75)) +
  theme_paper() +
  scale_y_continuous("Cumulative density") + scale_x_continuous("log2FoldChange") +
  ggtitle("RNA MMsat4 by region")

posECDFRNA

#RPF
posECDFRPF <- ggplot(RPF, aes(as.numeric(log2FoldChange), colour=factor(regMMsat4, levels = c("Non-target", "utr3", "cds", "both")))) +
  stat_ecdf(geom="step", linewidth=2) +
  scale_colour_manual(values = c("black", farbe1, farbe3, RPFncol)) +
  coord_cartesian(xlim = c(-0.75, 0.75)) +
  theme_paper() +
  scale_y_continuous("Cumulative density") + scale_x_continuous("log2FoldChange") +
  ggtitle("RPF MMsat4 by region")

posECDFRPF

```



by type of MRE (wobble and non wobble combined)

```

#RNA
RNA$MREtype <- "Non-target"
RNA$MREtype[RNA$gene_symbol %in% targetframe[targetframe$first_seed_200down.type == "No seed", "geneName"]]
RNA$MREtype[RNA$gene_symbol %in% targetframe[targetframe$first_seed_200down.type == "seed_6mer", "geneName"]]
RNA$MREtype[RNA$gene_symbol %in% targetframe[targetframe$first_seed_200down.type == "seed_7mer_a1", "geneName"]]
RNA$MREtype[RNA$gene_symbol %in% targetframe[targetframe$first_seed_200down.type == "seed_7mer_m8", "geneName"]]
RNA$MREtype[RNA$gene_symbol %in% targetframe[targetframe$first_seed_200down.type == "seed_8mer", "geneName"]]

#RPF
RPF$MREtype <- "Non-target"
RPF$MREtype[RPF$gene_symbol %in% targetframe[targetframe$first_seed_200down.type == "No seed", "geneName"]]
RPF$MREtype[RPF$gene_symbol %in% targetframe[targetframe$first_seed_200down.type == "seed_6mer", "geneName"]]
RPF$MREtype[RPF$gene_symbol %in% targetframe[targetframe$first_seed_200down.type == "seed_7mer_a1", "geneName"]]
RPF$MREtype[RPF$gene_symbol %in% targetframe[targetframe$first_seed_200down.type == "seed_7mer_m8", "geneName"]]

```

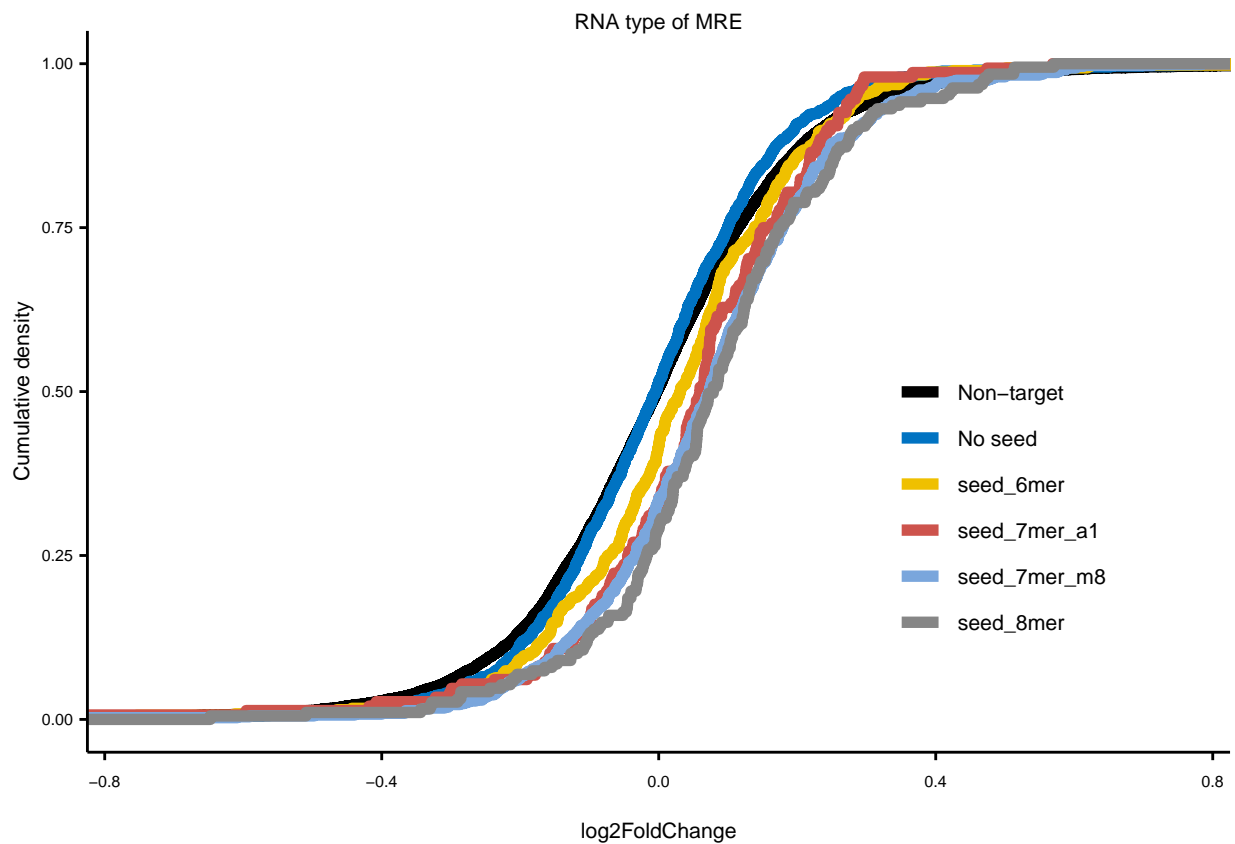
```
RPF$MREtype[RPF$gene_symbol %in% largetframe[largetframe$first_seed_200down.type == "seed_8mer", "geneName"]]
```

```
# ecdf plots
```

```
#RNA
```

```
typeECDFRNA <- ggplot(RNA, aes(as.numeric(log2FoldChange),  
                                colour=factor(MREtype, levels = c("Non-target", "No seed", "seed_6mer", "seed_7mer_a1", "seed_7mer_m8", "seed_8mer"))),  
  stat_ecdf(geom="step", linewidth=2) +  
  scale_colour_manual(values = c("black", farbe1, farbe2, farbe3, farbe4, farbe5)) +  
  coord_cartesian(xlim = c(-0.75, 0.75)) +  
  theme_paper() +  
  scale_y_continuous("Cumulative density") + scale_x_continuous("log2FoldChange") +  
  ggtitle("RNA type of MRE")
```

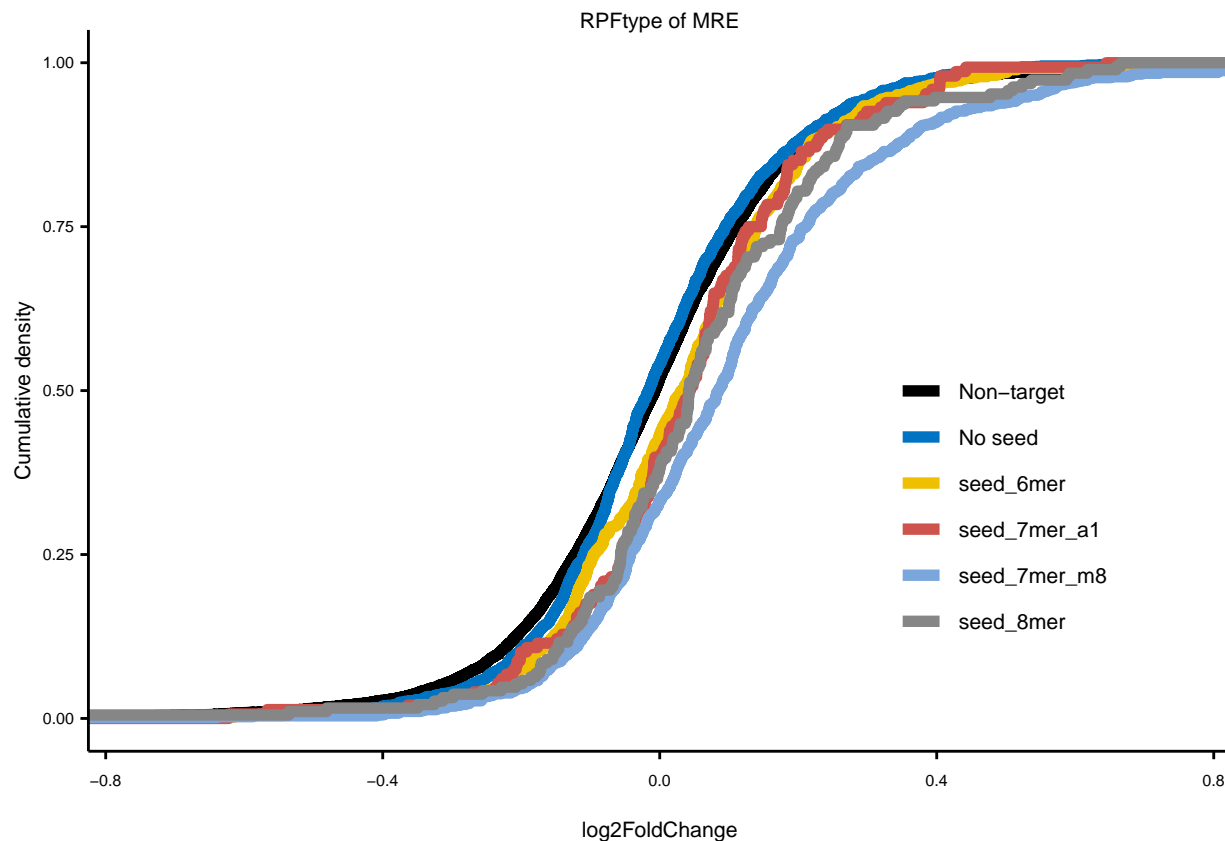
```
typeECDFRNA
```



```
#RPF
```

```
typeECDFRPF <- ggplot(RPF, aes(as.numeric(log2FoldChange),  
                                colour=factor(MREtype, levels = c("Non-target", "No seed", "seed_6mer", "seed_7mer_a1", "seed_7mer_m8", "seed_8mer"))),  
  stat_ecdf(geom="step", linewidth=2) +  
  scale_colour_manual(values = c("black", farbe1, farbe2, farbe3, farbe4, farbe5)) +  
  coord_cartesian(xlim = c(-0.75, 0.75)) +  
  theme_paper() +  
  scale_y_continuous("Cumulative density") + scale_x_continuous("log2FoldChange") +  
  ggtitle("RPFtype of MRE")
```

```
typeECDFRPF
```



divide by wobble and non wobble in seed

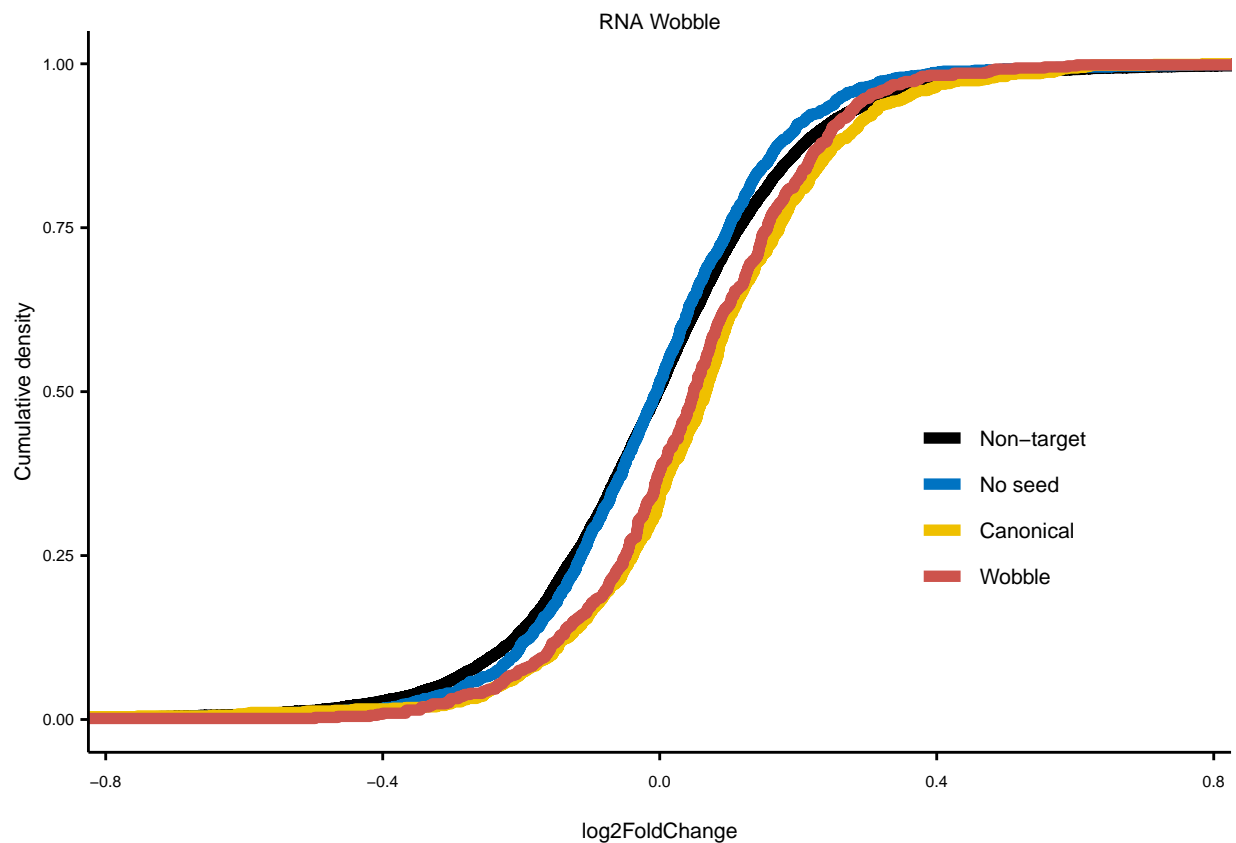
```
#RNA
RNA$wobble <- "Non-target"
RNA$wobble[RNA$gene_symbol %in% targetframe[targetframe$first_seed_200down.wobble == "No seed", "geneName"]]
RNA$wobble[RNA$gene_symbol %in% targetframe[targetframe$first_seed_200down.wobble == FALSE, "geneName"]]
RNA$wobble[RNA$gene_symbol %in% targetframe[targetframe$first_seed_200down.wobble == TRUE, "geneName"]]

#RPF
RPF$wobble <- "Non-target"
RPF$wobble[RPF$gene_symbol %in% targetframe[targetframe$first_seed_200down.wobble == "No seed", "geneName"]]
RPF$wobble[RPF$gene_symbol %in% targetframe[targetframe$first_seed_200down.wobble == FALSE, "geneName"]]
RPF$wobble[RPF$gene_symbol %in% targetframe[targetframe$first_seed_200down.wobble == TRUE, "geneName"]]

#ECDF plots
#RNA
wobbleECDFRNA <- ggplot(RNA, aes(as.numeric(log2FoldChange),
                                colour=factor(wobble, levels = c("Non-target", "No seed", "Canonical", "Wobble"))),
  stat_ecdf(geom="step", linewidth=2) +
  scale_colour_manual(values = c("black", farbe1, farbe2, farbe3)) +
  coord_cartesian(xlim = c(-0.75, 0.75)) +
  theme_paper() +
  scale_y_continuous("Cumulative density") + scale_x_continuous("log2FoldChange") +
  ggtitle("RNA Wobble")
```



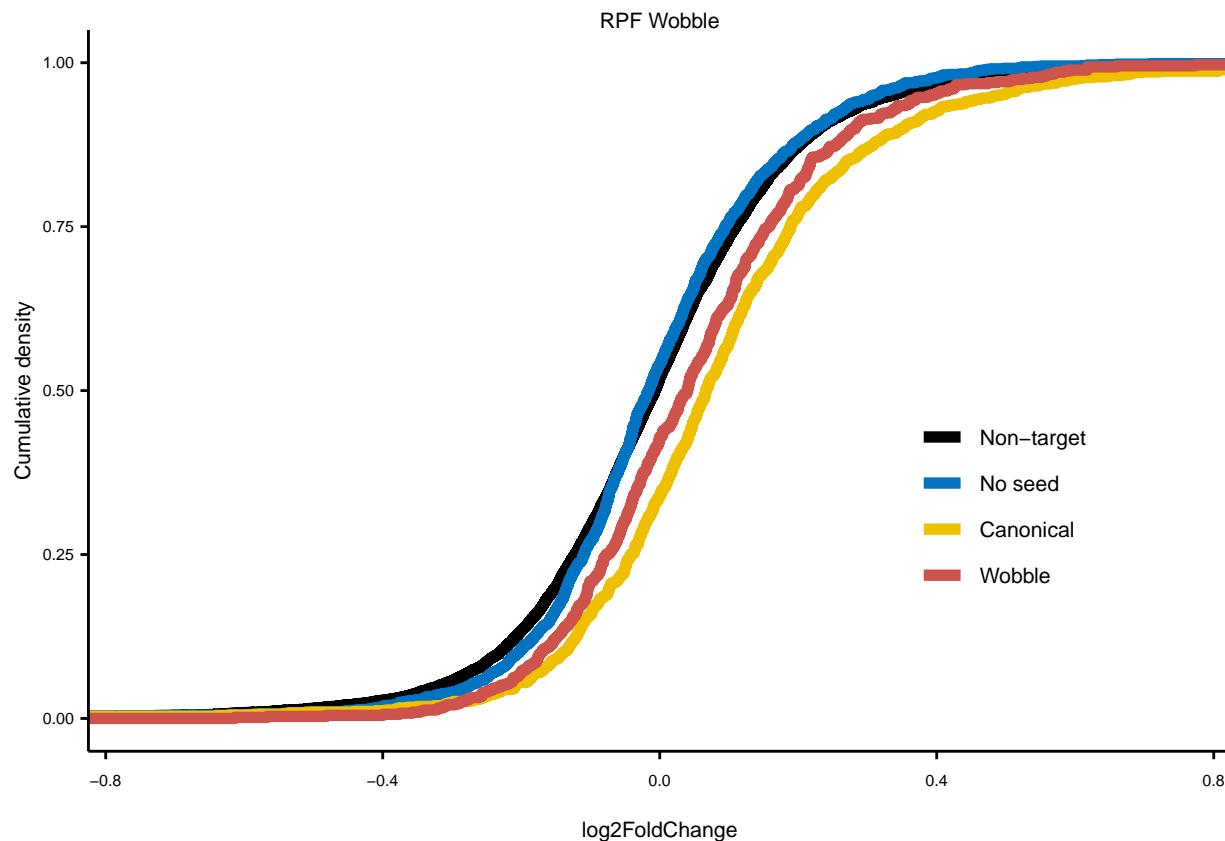
wobbleECDFRNA



*#RPF*

```
wobbleECDFRPF <- ggplot(RPF, aes(as.numeric(log2FoldChange),  
                                colour=factor(wobble, levels = c("Non-target", "No seed", "Canonical", "Wobble")),  
  stat_ecdf(geom="step", linewidth=2) +  
  scale_colour_manual(values = c("black", farbe1, farbe2, farbe3)) +  
  coord_cartesian(xlim = c(-0.75, 0.75)) +  
  theme_paper() +  
  scale_y_continuous("Cumulative density") + scale_x_continuous("log2FoldChange") +  
  ggtitle("RPF Wobble")
```

wobbleECDFRPF



## mRNA target expression levels (10% bins or quartiles)

```
#RNA
tRNA <- RNA[RNA$target == "Target",]
tRNA <- tRNA[order(tRNA$baseMean, decreasing = T),]
RNA$expressionlvl <- "Non-target"
RNA$expressionlvl[RNA$gene_symbol %in% tRNA[1:(0.25*length(tRNA$gene_symbol)), "gene_symbol"]] <- "High"
RNA$expressionlvl[RNA$gene_symbol %in% tRNA[(0.25*length(tRNA$gene_symbol)):(0.5*length(tRNA$gene_symbol)), "gene_symbol"]] <- "Mid-high"
RNA$expressionlvl[RNA$gene_symbol %in% tRNA[(0.5*length(tRNA$gene_symbol)):(0.75*length(tRNA$gene_symbol)), "gene_symbol"]] <- "Mid-low"
RNA$expressionlvl[RNA$gene_symbol %in% tRNA[(0.75*length(tRNA$gene_symbol)):length(tRNA$gene_symbol), "gene_symbol"]] <- "Low"

#RPF
tRPF <- RPF[RPF$target == "Target",]
tRPF <- tRPF[order(tRPF$baseMean, decreasing = T),]
RPF$expressionlvl <- "Non-target"
RPF$expressionlvl[RPF$gene_symbol %in% tRPF[1:(0.25*length(tRPF$gene_symbol)), "gene_symbol"]] <- "High"
RPF$expressionlvl[RPF$gene_symbol %in% tRPF[(0.25*length(tRPF$gene_symbol)):(0.5*length(tRPF$gene_symbol)), "gene_symbol"]] <- "Mid-high"
RPF$expressionlvl[RPF$gene_symbol %in% tRPF[(0.5*length(tRPF$gene_symbol)):(0.75*length(tRPF$gene_symbol)), "gene_symbol"]] <- "Mid-low"
RPF$expressionlvl[RPF$gene_symbol %in% tRPF[(0.75*length(tRPF$gene_symbol)):length(tRPF$gene_symbol), "gene_symbol"]] <- "Low"

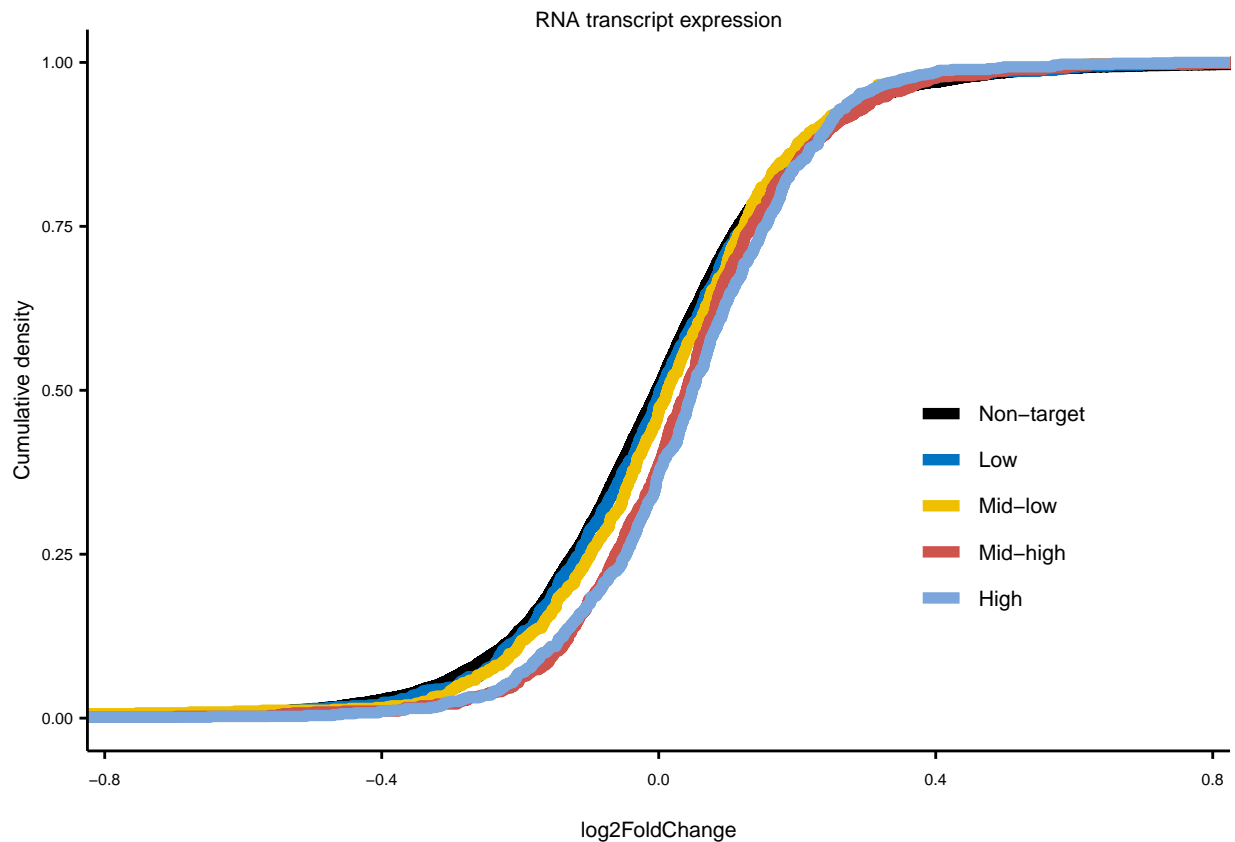
#ECDF plots
#RNA
expECDFRNA <- ggplot(RNA, aes(as.numeric(log2FoldChange),
                              colour=factor(expressionlvl, levels = c("Non-target", "Low", "Mid-low", "Mid-high", "High"))))
```

```

stat_ecdf(geom="step", linewidth=2) +
scale_colour_manual(values = c("black", farbe1, farbe2, farbe3, farbe4)) +
coord_cartesian(xlim = c(-0.75, 0.75)) +
theme_paper() +
scale_y_continuous("Cumulative density") + scale_x_continuous("log2FoldChange") +
ggtitle("RNA transcript expression")

```

expECDFRNA



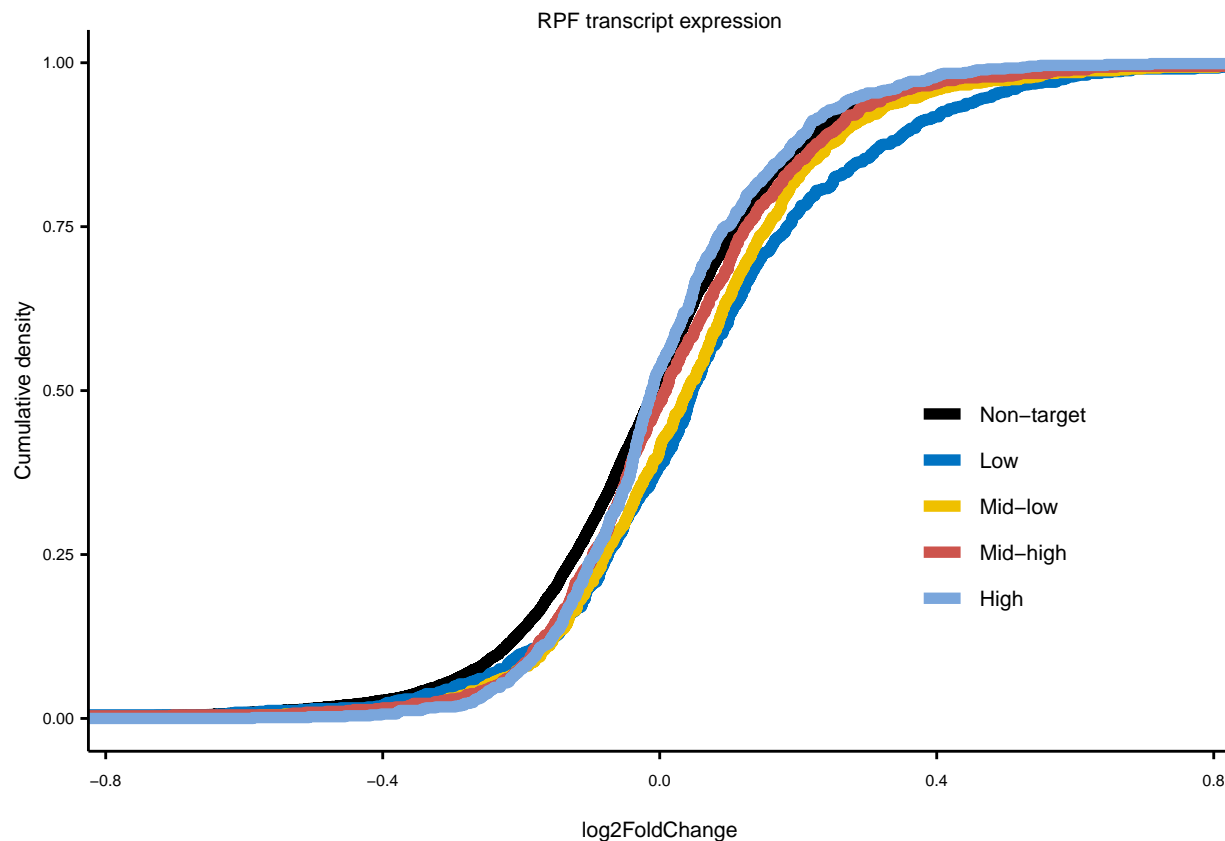
*#RPF*

```

expECDFRPF <- ggplot(RPF, aes(as.numeric(log2FoldChange),
                                colour=factor(expressionlvl, levels = c("Non-target", "Low", "Mid-low", "Mid-high", "High")))) +
  stat_ecdf(geom="step", linewidth=2) +
  scale_colour_manual(values = c("black", farbe1, farbe2, farbe3, farbe4)) +
  coord_cartesian(xlim = c(-0.75, 0.75)) +
  theme_paper() +
  scale_y_continuous("Cumulative density") + scale_x_continuous("log2FoldChange") +
  ggtitle("RPF transcript expression")

```

expECDFRPF



## Region with multiple sites ordered by “importance”

they were sorted in the following order: 5utr cds 3UTR

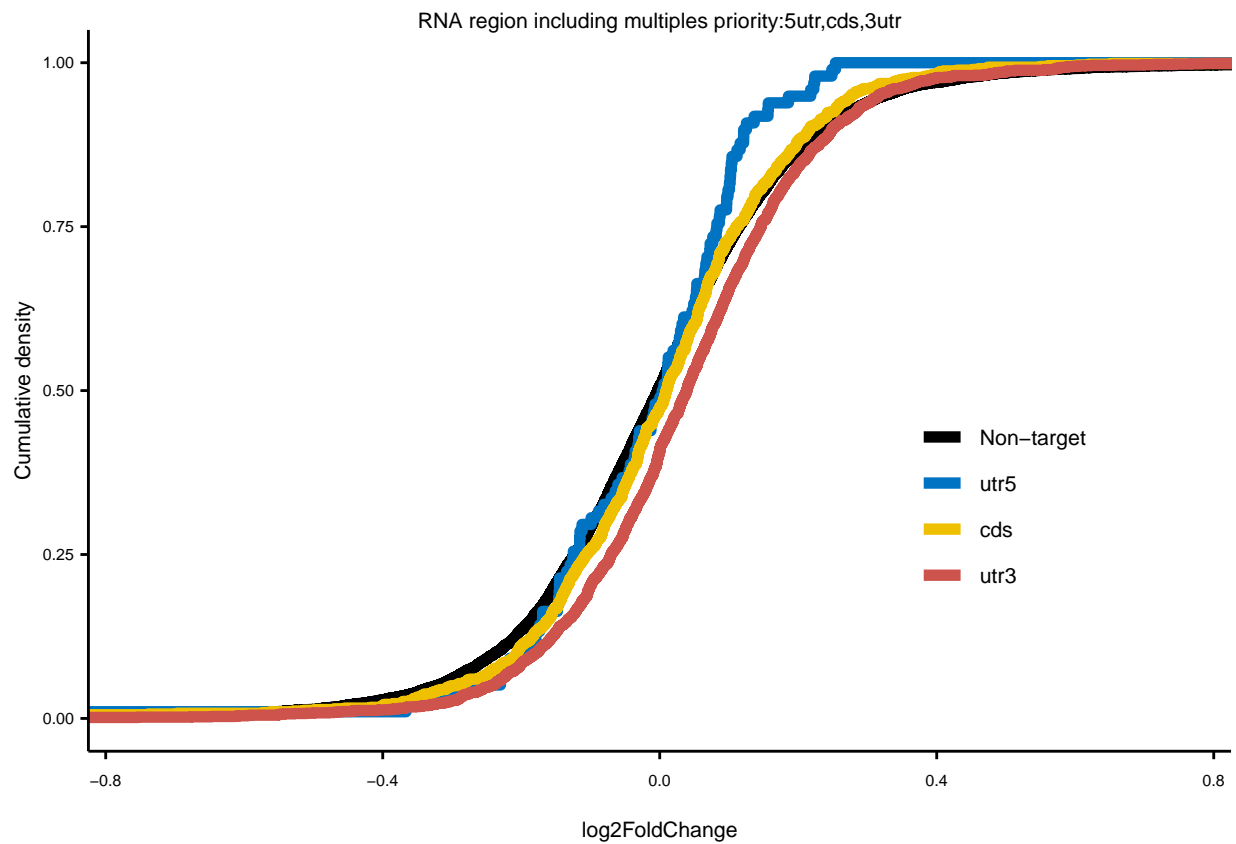
```
RNA$region_multiple <- "Non-target"
RNA$region_multiple[RNA$gene_symbol %in% targetframe[targetframe$region == "utr5", "geneName"]] <- "utr5"
RNA$region_multiple[RNA$gene_symbol %in% targetframe[targetframe$region == "cds", "geneName"]] <- "cds"
RNA$region_multiple[RNA$gene_symbol %in% targetframe[targetframe$region == "utr3", "geneName"]] <- "utr3"

RPF$region_multiple <- "Non-target"
RPF$region_multiple[RPF$gene_symbol %in% targetframe[targetframe$region == "utr5", "geneName"]] <- "utr5"
RPF$region_multiple[RPF$gene_symbol %in% targetframe[targetframe$region == "cds", "geneName"]] <- "cds"
RPF$region_multiple[RPF$gene_symbol %in% targetframe[targetframe$region == "utr3", "geneName"]] <- "utr3"

# ECDF plots

#RNA
regionmultiECDFRNA <- ggplot(RNA, aes(as.numeric(log2FoldChange),
                                     colour=factor(region_multiple, levels = c("Non-target", "utr5", "cds", "utr3")))) +
  stat_ecdf(geom="step", linewidth=2) +
  scale_colour_manual(values = c("black", farbe1, farbe2, farbe3)) +
  coord_cartesian(xlim = c(-0.75, 0.75)) +
  theme_paper() +
  scale_y_continuous("Cumulative density") + scale_x_continuous("log2FoldChange") +
  ggtitle("RNA region including multiples priority:5utr,cds,3utr")
```

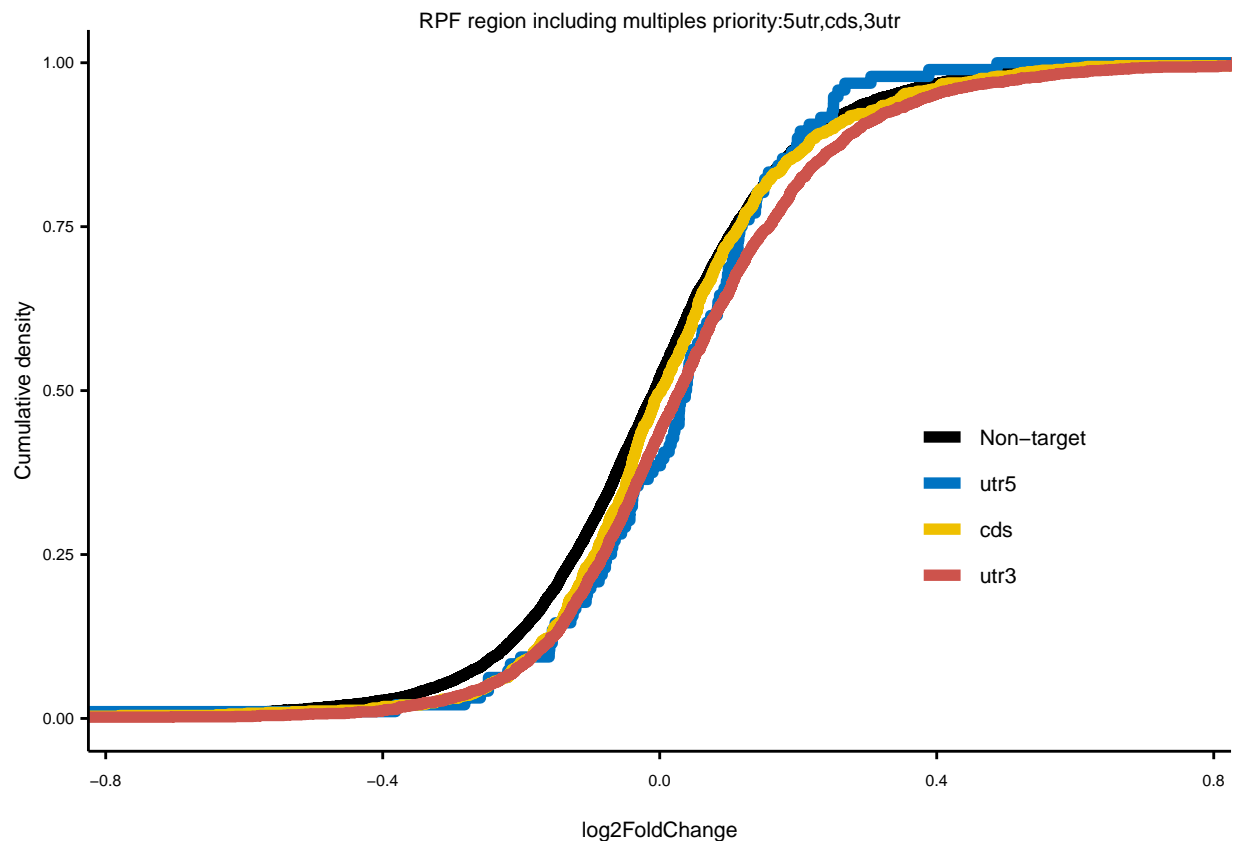
```
regionmultiECDFRNA
```



```
#RPF
```

```
regionmultiECDFRPF <- ggplot(RPF, aes(as.numeric(log2FoldChange),  
                                     colour=factor(region_multiple, levels = c("Non-target", "utr5", "cds", "utr3"))),  
  stat_ecdf(geom="step", linewidth=2) +  
  scale_colour_manual(values = c("black", farbe1, farbe2, farbe3)) +  
  coord_cartesian(xlim = c(-0.75, 0.75)) +  
  theme_paper() +  
  scale_y_continuous("Cumulative density") + scale_x_continuous("log2FoldChange") +  
  ggtitle("RPF region including multiples priority:5utr,cds,3utr"))
```

```
regionmultiECDFRPF
```



```
head(targetframe)
```

```
##      seqnames      start      end width strand scoreSum scoreMean scoreMax
## 1      chr1 6245651 6245657      7      + 9.52553 4.762765 6.00678
## 2      chr1 6248341 6248347      7      + 92.68921 23.172303 48.76900
## 3      chr1 6248857 6248863      7      + 14.07133 7.035665 7.04425
## 4      chr1 6248918 6248924      7      + 38.91451 12.971503 20.65080
## 5      chr1 7170481 7170487      7      + 66.92218 13.384436 25.84490
## 6      chr1 9899605 9899611      7      + 25.15963 6.289907 8.61019
##      geneType geneName      geneID region      mir_IP n_mir181
## 1 protein_coding Rb1cc1 ENSMUSG00000025907 cds mmu-miR-181a-5p 1
## 2 protein_coding Rb1cc1 ENSMUSG00000025907 cds mmu-miR-181a-5p 5
## 3 protein_coding Rb1cc1 ENSMUSG00000025907 cds mmu-miR-181a-5p 6
## 4 protein_coding Rb1cc1 ENSMUSG00000025907 cds mmu-miR-181a-5p 6
## 5 protein_coding Pcmt1d1 ENSMUSG00000051285 utr3 mmu-miR-181a-5p 4
## 6 protein_coding Sgk3 ENSMUSG00000025915 utr3 mmu-miR-181a-5p 1
##      n_mir181a n_mir181b n_mir181c n_mir181d      set mir181BS_ID WT KO
## 1      1      0      0      0      0 ago_bs_mir181_chi 1 1 1
## 2      5      0      0      0      0 ago_bs_mir181_chi 2 1 1
## 3      6      0      0      0      0 ago_bs_mir181_chi 3 1 0
## 4      6      0      0      0      0 ago_bs_mir181_chi 4 1 1
## 5      4      0      0      0      0 ago_bs_mir181_chi 5 1 1
## 6      1      0      0      0      0 ago_bs_mir181_chi 6 NA NA
##      geneID.2 geneName.1 region.1 counts.bs.1_KO counts.bs.2_KO
## 1 ENSMUSG00000025907 Rb1cc1 cds 4 3
## 2 ENSMUSG00000025907 Rb1cc1 cds 28 32
```

##	3	ENSMUSG00000025907	Rb1cc1	cds	13	11
##	4	ENSMUSG00000025907	Rb1cc1	cds	15	15
##	5	ENSMUSG00000051285	Pcmdt1	utr3	12	22
##	6	<NA>	<NA>	<NA>	NA	NA
##		counts.bs.3_KO	counts.bs.4_WT	counts.bs.5_WT	counts.bs.6_WT	
##	1	3	3	10	3	
##	2	27	46	41	20	
##	3	4	22	13	12	
##	4	10	33	20	18	
##	5	14	16	20	9	
##	6	NA	NA	NA	NA	
##		geneID.1	counts.bg.1_KO	counts.bg.2_KO	counts.bg.3_KO	
##	1	ENSMUSG00000025907	1609	1973	1250	
##	2	ENSMUSG00000025907	1609	1973	1250	
##	3	ENSMUSG00000025907	1609	1973	1250	
##	4	ENSMUSG00000025907	1609	1973	1250	
##	5	ENSMUSG00000051285	1355	1706	1064	
##	6	<NA>	NA	NA	NA	
##		counts.bg.4_WT	counts.bg.5_WT	counts.bg.6_WT	resBs.baseMean	
##	1	2638	2231	1352	92.10645	
##	2	2638	2231	1352	281.53271	
##	3	2638	2231	1352	145.51107	
##	4	2638	2231	1352	186.74162	
##	5	1654	1348	755	151.36245	
##	6	NA	NA	NA	NA	
##		resBs.log2FoldChange	resBs.lfcSE	resBs.stat	resBs.pvalue	resBs.padj
##	1	-0.1093039	0.5923673	0.03419066	0.8533018	0.9652601
##	2	0.2749428	0.2351157	1.35874137	0.2437557	0.6729889
##	3	-0.1805519	0.3623758	0.25017050	0.6169550	0.8961239
##	4	-0.2606282	0.3062717	0.73169661	0.3923338	0.7868678
##	5	0.1466485	0.3122905	0.22052922	0.6386370	0.9013566
##	6	NA	NA	NA	NA	NA
##		resBg.baseMean	resBg.log2FoldChange	resBg.lfcSE	resBg.stat	resBg.pvalue
##	1	NA	NA	NA	NA	NA
##	2	NA	NA	NA	NA	NA
##	3	NA	NA	NA	NA	NA
##	4	NA	NA	NA	NA	NA
##	5	NA	NA	NA	NA	NA
##	6	NA	NA	NA	NA	NA
##		resBg.padj	tpm.counts.bg.1_KO	tpm.counts.bg.2_KO	tpm.counts.bg.3_KO	
##	1	NA	133.7259	117.9980	129.8669	
##	2	NA	133.7259	117.9980	129.8669	
##	3	NA	133.7259	117.9980	129.8669	
##	4	NA	133.7259	117.9980	129.8669	
##	5	NA	248.6210	225.2505	244.0445	
##	6	NA	NA	NA	NA	
##		tpm.counts.bg.4_WT	tpm.counts.bg.5_WT	tpm.counts.bg.6_WT		
##	1	139.8635	146.2855	163.5360		
##	2	139.8635	146.2855	163.5360		
##	3	139.8635	146.2855	163.5360		
##	4	139.8635	146.2855	163.5360		
##	5	193.5994	195.1330	201.6149		
##	6	NA	NA	NA		
##		BS_ID	tpm_support_KO	tpm_support_WT	tpm_supported	down

```

## 1 ENSMUSG00000025907.bs5      3      3      TRUE FALSE
## 2 ENSMUSG00000025907.bs8      3      3      TRUE FALSE
## 3 ENSMUSG00000025907.bs10     3      3      TRUE FALSE
## 4 ENSMUSG00000025907.bs11     3      3      TRUE FALSE
## 5 ENSMUSG00000051285.bs4      3      3      TRUE FALSE
## 6      <NA>      NA      NA      NA      NA
## all_seeds_200down first_seed_200down.start first_seed_200down.end
## 1      NA, NA, ....      NA      NA
## 2      NA, NA, ....      NA      NA
## 3      121:122,....      121      127
## 4      60:61, c....      60      66
## 5      c(17, 17....      17      24
## 6      168:167,....      167      173
## first_seed_200down.width first_seed_200down.type first_seed_200down.wobble
## 1      NA      No seed      No seed
## 2      NA      No seed      No seed
## 3      7      seed_7mer_m8      FALSE
## 4      7      seed_7mer_m8      FALSE
## 5      8      seed_8mer      FALSE
## 6      7      seed_7mer_m8      TRUE
## seed_repetitions.200down seed_repetitions.200down.wobble all_seeds_200up
## 1      NA      NA      NA, NA, ....
## 2      NA      NA      NA, NA, ....
## 3      1      0      NA, NA, ....
## 4      1      0      NA, NA, ....
## 5      1      0      NA, NA, ....
## 6      0      1      NA, NA, ....
## first_seed_200up.start first_seed_200up.end first_seed_200up.width
## 1      NA      NA      NA
## 2      NA      NA      NA
## 3      NA      NA      NA
## 4      NA      NA      NA
## 5      NA      NA      NA
## 6      NA      NA      NA
## first_seed_200up.type first_seed_200up.wobble seed_repetitions.200up
## 1      <NA>      NA      NA
## 2      <NA>      NA      NA
## 3      <NA>      NA      NA
## 4      <NA>      NA      NA
## 5      <NA>      NA      NA
## 6      <NA>      NA      NA
## seed_repetitions.200up.wobble
## 1      NA
## 2      NA
## 3      NA
## 4      NA
## 5      NA
## 6      NA

```

## Export

“D:/Krueger\_Lab/Publications/miR181\_paper/Figure4”



```
#target
```

```
pdf("D:/Krueger_Lab/Publications/miR181_paper/Figure4/targetoldintronECDFRNA.pdf", width=2, height = 2)  
targetoldECDFRNA  
dev.off()
```

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

```
pdf("D:/Krueger_Lab/Publications/miR181_paper/Figure4/targetoldintronECDFRPF.pdf", width=2, height = 2)  
targetoldECDFRPF  
dev.off()
```

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

```
pdf("D:/Krueger_Lab/Publications/miR181_paper/Figure4/targetECDF_RNA.pdf", width=2, height = 2)  
targetECDFRNA  
dev.off()
```

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

```
pdf("D:/Krueger_Lab/Publications/miR181_paper/Figure4/targetECDF_RPF.pdf", width=2, height = 2)  
targetECDFRPF  
dev.off()
```

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

```
#set
```

```
pdf("D:/Krueger_Lab/Publications/miR181_paper/Figure4/setECDF_RNA.pdf", width=2, height = 2)  
setECDFRNA  
dev.off()
```

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

```
pdf("D:/Krueger_Lab/Publications/miR181_paper/Figure4/setECDF_RPF.pdf", width=2, height = 2)  
setECDFRPF  
dev.off()
```

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

```
#region single targets
```

```
pdf("D:/Krueger_Lab/Publications/miR181_paper/Figure4/regsingECDF_RNA.pdf", width=2, height = 2)  
regsingECDFRNA  
dev.off()
```

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

```
pdf("D:/Krueger_Lab/Publications/miR181_paper/Figure4/regsingECDF_RPF.pdf", width=2, height = 2)  
regsingECDFRPF  
dev.off()
```

```
## pdf
```

```

## 2
#number

pdf("D:/Krueger_Lab/Publications/miR181_paper/Figure4/numECDF_RNA.pdf", width=2, height = 2)
numECDFRNA
dev.off()

## pdf
## 2

pdf("D:/Krueger_Lab/Publications/miR181_paper/Figure4/numECDF_RPF.pdf", width=2, height = 2)
numECDFRPF
dev.off()

## pdf
## 2
# MMsat4 vs target

pdf("D:/Krueger_Lab/Publications/miR181_paper/Figure4/MMsat4vsTargetECDF_RNA.pdf", width=2, height = 2)
tolECDFRNA
dev.off()

## pdf
## 2

pdf("D:/Krueger_Lab/Publications/miR181_paper/Figure4/MMsat4vsTargetECDF_RPF.pdf", width=2, height = 2)
tolECDFRPF
dev.off()

## pdf
## 2
# region with MMsat4 cds and 3'utr

pdf("D:/Krueger_Lab/Publications/miR181_paper/Figure4/MMsat4byRegionECDF_RNA.pdf", width=2, height = 2)
posECDFRNA
dev.off()

## pdf
## 2

pdf("D:/Krueger_Lab/Publications/miR181_paper/Figure4/MMsat4byRegionECDF_RPF.pdf", width=2, height = 2)
posECDFRPF
dev.off()

## pdf
## 2
# MRE type

pdf("D:/Krueger_Lab/Publications/miR181_paper/Figure4/MRtypeECDF_RNA.pdf", width=2, height = 2)
typeECDFRNA
dev.off()

## pdf
## 2

pdf("D:/Krueger_Lab/Publications/miR181_paper/Figure4/MRtypeECDF_RPF.pdf", width=2, height = 2)
typeECDFRPF

```

```

dev.off()

## pdf
## 2
# Wobble

pdf("D:/Krueger_Lab/Publications/miR181_paper/Figure4/wobbleECDF_RNA.pdf", width=2, height = 2)
wobbleECDFRNA
dev.off()

## pdf
## 2

pdf("D:/Krueger_Lab/Publications/miR181_paper/Figure4/wobbleECDF_RPF.pdf", width=2, height = 2)
wobbleECDFRPF
dev.off()

## pdf
## 2
# transcript expression

pdf("D:/Krueger_Lab/Publications/miR181_paper/Figure4/expresssionECDF_RNA.pdf", width=2, height = 2)
expECDFRNA
dev.off()

## pdf
## 2

pdf("D:/Krueger_Lab/Publications/miR181_paper/Figure4/expresssionECDF_RPF.pdf", width=2, height = 2)
expECDFRPF
dev.off()

## pdf
## 2
# transcript expression

pdf("D:/Krueger_Lab/Publications/miR181_paper/Figure4/regionmultiECDF_RNA.pdf", width=2, height = 2)
regionmultiECDFRNA
dev.off()

## pdf
## 2

pdf("D:/Krueger_Lab/Publications/miR181_paper/Figure4/regionmultiECDF_RPF.pdf", width=2, height = 2)
regionmultiECDFRPF
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      stats      graphics  grDevices  utils      datasets  methods
## [8] base
##
## other attached packages:
## [1] dplyr_1.1.2      rtracklayer_1.58.0  GenomicRanges_1.50.2
## [4] GenomeInfoDb_1.34.9  IRanges_2.32.0      S4Vectors_0.36.2
## [7] BiocGenerics_0.44.0  ggplot2_3.4.2
##
## loaded via a namespace (and not attached):
## [1] lattice_0.20-45      tidyr_1.3.0
## [3] Rsamtools_2.14.0     Biostrings_2.66.0
## [5] digest_0.6.31        utf8_1.2.3
## [7] R6_2.5.1             backports_1.4.1
## [9] evaluate_0.21        highr_0.10
## [11] pillar_1.9.0         zlibbioc_1.44.0
## [13] rlang_1.1.0          rstudioapi_0.14
## [15] car_3.1-2            Matrix_1.5-3
## [17] rmarkdown_2.21       labeling_0.4.2
## [19] BiocParallel_1.32.6  RCurl_1.98-1.12
## [21] munsell_0.5.0        DelayedArray_0.23.2
## [23] broom_1.0.4          compiler_4.2.3
## [25] xfun_0.39            pkgconfig_2.0.3
## [27] htmltools_0.5.4      tidyselect_1.2.0
## [29] SummarizedExperiment_1.28.0 tibble_3.2.1
## [31] GenomeInfoDbData_1.2.9  codetools_0.2-19
## [33] matrixStats_0.63.0    XML_3.99-0.14
## [35] fansi_1.0.4          crayon_1.5.2
## [37] withr_2.5.0          ggpubr_0.6.0
## [39] GenomicAlignments_1.34.1 bitops_1.0-7
## [41] grid_4.2.3           gtable_0.3.3
## [43] lifecycle_1.0.3      magrittr_2.0.3
## [45] scales_1.2.1         cli_3.6.0
## [47] carData_3.0-5        farver_2.1.1
## [49] XVector_0.38.0       ggsignif_0.6.4
## [51] generics_0.1.3       vctrs_0.6.2
## [53] rjson_0.2.21         restfulr_0.0.15
## [55] tools_4.2.3          Biobase_2.58.0
## [57] glue_1.6.2           purrr_1.0.1
## [59] MatrixGenerics_1.10.0 abind_1.4-5
## [61] parallel_4.2.3       fastmap_1.1.1
## [63] yaml_2.3.7           colorspace_2.1-0
## [65] rstatix_0.7.2        knitr_1.42
## [67] BiocIO_1.8.0

```