Riborex Manual

Wenzheng Li, Weili Wang, Philip J. Uren, Luiz OF Penalva, Andrew D. Smith

Introduction

Riborex is a R package for identifying differentially translated genes from Ribo-seq data. Riborex integrates both RNA- and Ribo-seq read count data into a single generalized linear model (GLM) and generates a modified design matrix reflecting the integration. At its core, Riborex applies existing RNA-seq analysis tools such as edgeR, DESeq2 and Voom to this modified design matrix and identifies differential translation across conditions.

Detailed example

First, we need to load Riborex library.

```
library(riborex)
```

The input for Riborex are two read count tables summarized from RNA-seq and Ribo-seq data respectively. The read count table should be organized as a data frame with rows correspond to genes and columns correspond to samples as shown below.

```
data(riborexdata)
RNACntTable <- riborexdata$rna
RiboCntTable <- riborexdata$ribo
```

We can check the first five lines of the table:

```
head(RNACntTable,5)
```

```
##                    BN_336 BN_337 BN_338 BN_339
## ENSRNOG00000000017      7     11      4      4
## ENSRNOG00000000024   2467   2478   3258   2316
## ENSRNOG00000000033    206    282    330    244
## ENSRNOG00000000034    758    672   1335    767
## ENSRNOG00000000036    237    163    211    189
```

```
head(RiboCntTable,5)
```

```
##                    BN_341 BN_342 BN_343 BN_344
## ENSRNOG00000000017     15      5     10      2
## ENSRNOG00000000024   5206   5921   2864   1985
## ENSRNOG00000000033     30     30     23     13
## ENSRNOG00000000034    943    775    842    311
## ENSRNOG00000000036     80     49     30      7
```

Then we need to prepare two vectors to indicate the treatments of samples in RNA- and Ribo-seq data. Both RNA-seq and Ribo-seq can have different number of samples in control and treated condtions, and RNA-seq and Ribo-seq data can have different number of samples.

```r
rnaCond <- c("control", "control", "treated", "treated")
riboCond <- c("control", "control", "treated", "treated")
```

After the two read count table and two condition vectors are ready, we can use riborex (), and we can choose which engine to use. By default, DESeq2 is used as the engine if you don't specify the engine option. Use help(riborex) in R to see more details about this function.

```r
res.deseq2 <- riborex(RNACntTable, RiboCntTable, rnaCond, riboCond)
```

The format of the result is the same when DESeq2 is used in RNA-seq analysis.

```r
res.deseq2
```

```
## log2 fold change (MAP): EXTRA2 treated vs control
## Wald test p-value: EXTRA2 treated vs control
## DataFrame with 13916 rows and 6 columns
##                       baseMean log2FoldChange      lfcSE       stat
##                      <numeric>      <numeric>  <numeric>  <numeric>
## ENSRNOG00000000017    7.694934     0.60242172 0.6633528  0.9081467
## ENSRNOG00000000024 3544.238071    -0.09262781 0.1967739 -0.4707321
## ENSRNOG00000000033  111.439451     0.12090513 0.4390872  0.2753556
## ENSRNOG00000000034  783.974916     0.07400898 0.3243979  0.2281426
## ENSRNOG00000000036   99.151141    -0.65373460 0.4894467 -1.3356604
## ...                        ...            ...        ...        ...
## ENSRNOG00000061895  105.24110     0.42802240 0.4224309  1.0132366
## ENSRNOG00000061899   40.04057    -0.55910038 0.5828614 -0.9592339
## ENSRNOG00000061910 4237.53481     0.29957699 0.2366771  1.2657623
## ENSRNOG00000061928 1651.61680    -0.07583546 0.2370871 -0.3198633
## ENSRNOG00000061989  107.36281    -0.06620802 0.4009571 -0.1651250
##                        pvalue      padj
##                     <numeric> <numeric>
## ENSRNOG00000000017 0.3638007 0.7494602
## ENSRNOG00000000024 0.6378320 0.9022741
## ENSRNOG00000000033 0.7830430 0.9509349
## ENSRNOG00000000034 0.8195354 0.9607316
## ENSRNOG00000000036 0.1816603 0.5432690
## ...                      ...       ...
## ENSRNOG00000061895 0.3109472 0.7034611
## ENSRNOG00000061899 0.3374409 0.7261818
## ENSRNOG00000061910 0.2055982 0.5775201
## ENSRNOG00000061928 0.7490719 0.9411896
## ENSRNOG00000061989 0.8688456 0.9744379
```

Also, you can use summary () for your results.

```r
summary(res.deseq2)
```

```
##
## out of 13916 with nonzero total read count
## adjusted p-value < 0.1
## LFC > 0 (up)       : 1139, 8.2%
```

```
## LFC < 0 (down)    : 1274, 9.2%
## outliers [1]      : 0, 0%
## low counts [2]    : 270, 1.9%
## (mean count < 3)
## [1] see 'cooksCutoff' argument of ?results
## [2] see 'independentFiltering' argument of ?results
```

And results can be saved by:

```
write.table(res.deseq2, "riborex_res_deseq2.txt", quote=FALSE)
```

If you want to use edgeR as your engine, you can use riborex () as:

```
res.edgeR <- riborex(RNACntTable, RiboCntTable, rnaCond, riboCond, "edgeR")
```

The format of the result is the same when edgeR is used in RNA-seq analysis.

```
head(res.edgeR$table)
```

```
##                      logFC     logCPM          LR     PValue       FDR
## ENSRNOG00000000017  1.3629124 -2.456968 1.80670421 0.1789041 0.4626704
## ENSRNOG00000000024 -0.3012715  6.212404 2.13615823 0.1438618 0.4037879
## ENSRNOG00000000033  0.0717875  1.313235 0.02631400 0.8711358 0.9636269
## ENSRNOG00000000034 -0.1343023  4.029136 0.12544284 0.7232047 0.9111520
## ENSRNOG00000000036 -0.8253990  1.132478 2.15594580 0.1420190 0.4007169
## ENSRNOG00000000040 -0.1905669 -1.555003 0.07536038 0.7836862 0.9338705
```

For edgeR engine, you can also choose to estimate dispersion of RNA-seq and Ribo-seq data separately by specifying engine as "edgeRD".

```
res.edgeRD <- riborex(RNACntTable, RiboCntTable, rnaCond, riboCond, "edgeRD")
```

If you want to use Voom as the engine, you can run riborex () as:

```
res.voom <- riborex(RNACntTable, RiboCntTable, rnaCond, riboCond, "Voom")
```

The format of the result is the same when Voom is used in RNA-seq analysis.

```
head(res.voom)
```

```
##                      logFC     AveExpr          t    P.Value adj.P.Val
## ENSRNOG00000000017  1.39674189 -2.8437969  1.2847048 0.2268559 0.5243467
## ENSRNOG00000000024 -0.30047073  6.0075571 -1.8737990 0.0893910 0.2991861
## ENSRNOG00000000033  0.07661457  0.7070877  0.1687028 0.8692764 0.9540355
## ENSRNOG00000000034 -0.13777833  3.9701893 -0.3933633 0.7020190 0.8871606
## ENSRNOG00000000036 -0.89165405  0.7106061 -1.3890222 0.1939374 0.4826701
## ENSRNOG00000000040 -0.20967655 -1.7042630 -0.2991046 0.7707701 0.9157225
##                           B
## ENSRNOG00000000017 -4.928058
## ENSRNOG00000000024 -5.682138
## ENSRNOG00000000033 -6.320171
## ENSRNOG00000000034 -7.130133
## ENSRNOG00000000036 -5.425369
## ENSRNOG00000000040 -5.854482
```