

## Gene ontology and pathway analysis

Finally, let's try some simple gene ontology (GO) and pathway analysis. There are quite a few BioConductor packages for this (as you will find if you spend a few minutes browsing <http://www.bioconductor.org/packages/2.12/bioc/> (<http://www.bioconductor.org/packages/2.12/bioc/>)) but we will take a brief look at one of those, topGO.

Start by installing the two packages in the usual BioConductor way.

### topGO

topGO is a Gene Ontology enrichment analysis package. It needs us to provide a list of all the genes we have analyzed, and a measure of their significance in the DE analysis. We will use the adjusted p value, padj, for this.

```
library(topGO)
all <- res$padj
names(all) <- rownames(res)
all <- na.omit(all)
```

Then we can create a topGOdata object which we can then manipulate in different ways. You can choose between looking at CC (cellular component), MF (molecular function) or BP (biological process).

```
G0data <- new("topGOdata",ontology = "CC", allGenes = all, geneSel=function(p) p < 0.01, description
="ProstateDEGenes", annot=annFUN.org, mapping="org.Hs.eg.db", ID="Ensembl")
```

Then we can perform a test to assess significant overrepresentation of genes in GO categories:

```
resultFisher <- runTest(G0data, algorithm = "classic", statistic = "fisher")
```

The results can be displayed in a table:


```
GenTable(G0data, classicFisher = resultFisher, topNodes = 10)
```

If we want to get fancy, we can also plot the results as a nice graph, but we need to install another package for that:

```
biocLite('Rgraphviz')
library(Rgraphviz)
showSigOfNodes(G0data, score(resultFisher),firstSigNodes=5,useInfo='all')
```

## QUESTION 3 (1/1 point)

What theme do the most enriched GO terms relate to?

- ☐ Nucleic acids (e g DNA and RNA)
- ☒ Physical components (e g membranes, organelles and extracellular matrix) 
- ☐ Processes such as cell division or differentiation

[Final Check](#)[Save](#)[Hide Answer](#)

You have used 1 of 2 submissions

## References on RNA sequencing

Wang et al. RNA-Seq: a revolutionary tool for transcriptomics. Nature Rev Genet 10:5763, 2009.

Mortazavi et al. Mapping and quantifying mammalian transcriptomes by RNASeq. Nature 5:621628,2008.

Auer and Doerge. Statistical Design and Analysis of RNA Sequencing Data. Genetics 185:405–416, 2010.

[Show Discussion](#)[New Post](#)

EdX offers interactive online classes and MOOCs from the world's best universities. Online courses from MITx, HarvardX, BerkeleyX, UTx and many other universities. Topics include biology, business, chemistry, computer science, economics, finance, electronics, engineering, food and nutrition, history, humanities, law, literature, math, medicine, music, philosophy, physics, science, statistics and more. EdX is a non-profit online initiative created by founding partners Harvard and MIT.

© 2014 edX, some rights reserved.

Terms of Service and Honor Code (<https://www.edx.org/edx-terms-service>)

Privacy Policy (Revised 4/16/2014) (<https://www.edx.org/edx-privacy-policy>)

### About & Company Info

About

(<https://www.edx.org/about-us>)

News

(<https://www.edx.org/news>)

Contact

(<https://www.edx.org/contact>)

FAQ

(<https://www.edx.org/student-faq>)

edX Blog

(<https://www.edx.org/edx-blog>)

Donate to edX

(<https://www.edx.org/donate>)

Jobs at edX

(<https://www.edx.org/jobs>)


### Follow Us

 Twitter

(<https://twitter.com/edXOnline>)

 Facebook

(<http://www.facebook.com/EdxOnline>)

 Meetup

(<http://www.meetup.com/edX-Global-Community>)

 LinkedIn

(<http://www.linkedin.com/company/edx>)

 Google+

(<https://plus.google.com/+edXOnline>)