Zebrafish Dataset Practical 3

All of these exercises will be done on the same dataset as yesterday's practicals. Make sure you've read the document that describes the dataset and make sure you've got the four required files (`Amp.counts.tsv`, `Amp.samples.tsv`, `Oxy.counts.tsv` and `Oxy.samples.tsv`) in your home directory. You will also need the significantly differentially expressed gene lists you made in Tuesday's practical and also the files that just contain their Ensembl IDs.

(If you haven't finished Tuesday's or yesterday's practical then please carry on with them, either before or after today's practical.)

You've been shown a number of tools today, including:

- PANTHER: http://pantherdb.org/webservices/go/overrep.jsp
- Princeton: https://go.princeton.edu/cgi-bin/GOTermFinder
- g:Profiler: https://biit.cs.ut.ee/gprofiler/gost
- Reactome: https://reactome.org/
- Cytoscape
- Ontologizer

We'd like you to practise using them with the zebrafish dataset. Please try using all of the tools you've been introduced to today.

As well as doing enrichments using all the significant genes, try just using all the genes whose expression goes up (i.e. a positive $\log_2$ fold change) and those whose expression goes down (i.e. a negative $\log_2$ fold change). How do the enrichments change?

Many of these tools produce a table of terms and adjusted p-values. How could you use R to present these data graphically? Try to make a plot showing the distribution of p-values for significant terms.