Exercise 1

- We will use the tools on the DAVID web site to explore the annotations for the keratinocyte data.
- Upload the significant genes using the gene-specific p-value and a criterion of $p < 0.05$.
- Upload the full set of genes on the array as the background.
- Look at all three functional allocation tools and evaluate how useful they are in this case.
- Try this again with the posterior p-values and with the FDR corrected posterior p-values with $p < 0.10$. Does it make a difference in the interpretation?

Exercise 2

- In the bcmort data set, the four-level factor cohort can be considered the product of two two-level factors, say “period” (1981-1991 or 1991-2001) and “area” (Copenhagen/Fredriksberg and National). Generate those two factors.
- Fit a Poisson regression model to the data with age, period, and area as descriptors, as well as the three two-factor interaction terms. The interaction between period and area can be interpreted as the effect of screening (explain why).
- Try this regression using “family = quasipoisson” and see if it makes any difference.