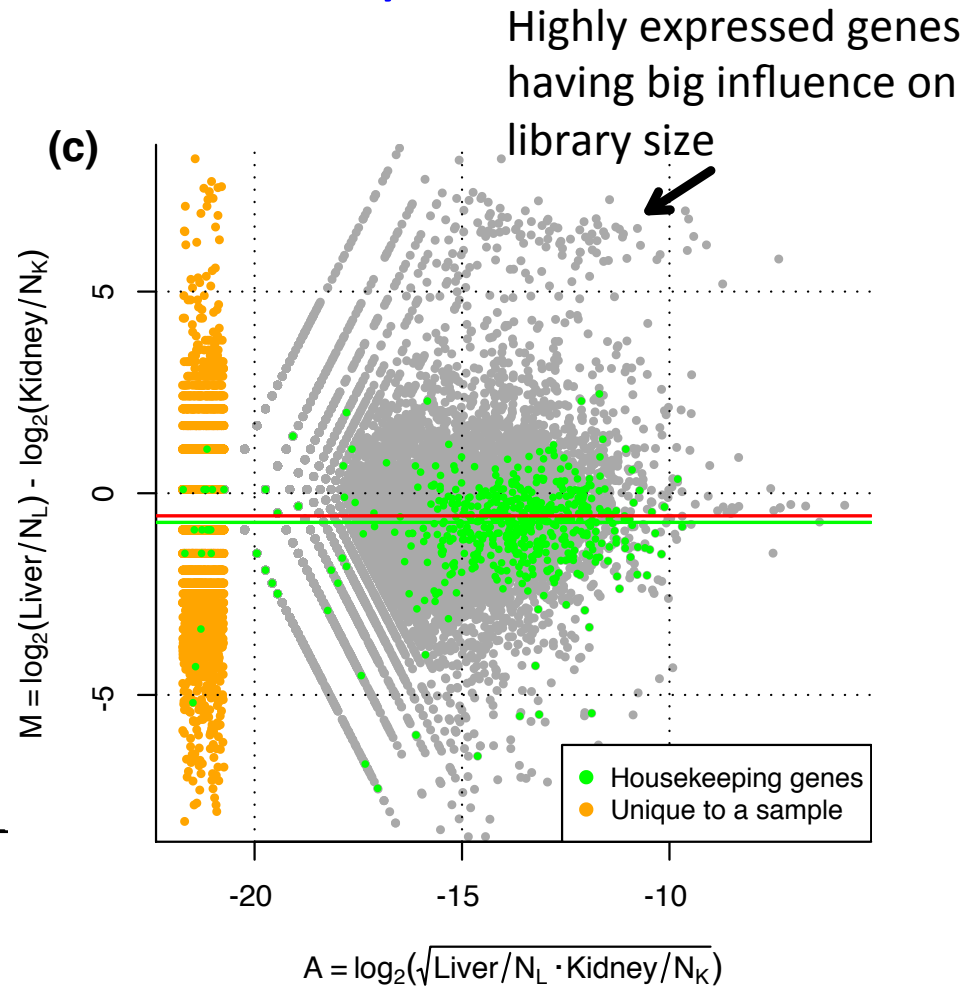
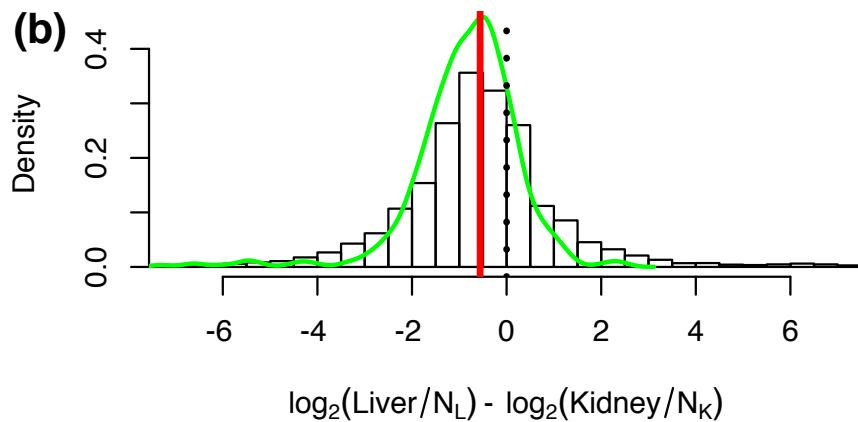
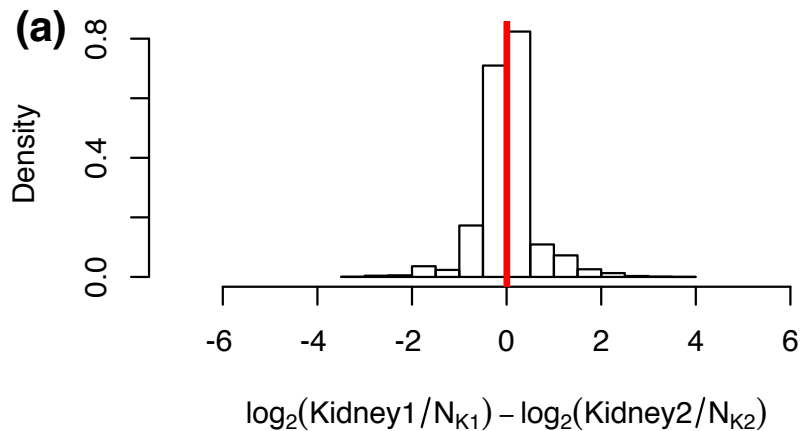


edgeR

An R/Bioconductor package
(Robinson et al, 2010)

Normalization using TMM (Trimmed Mean of M-values)

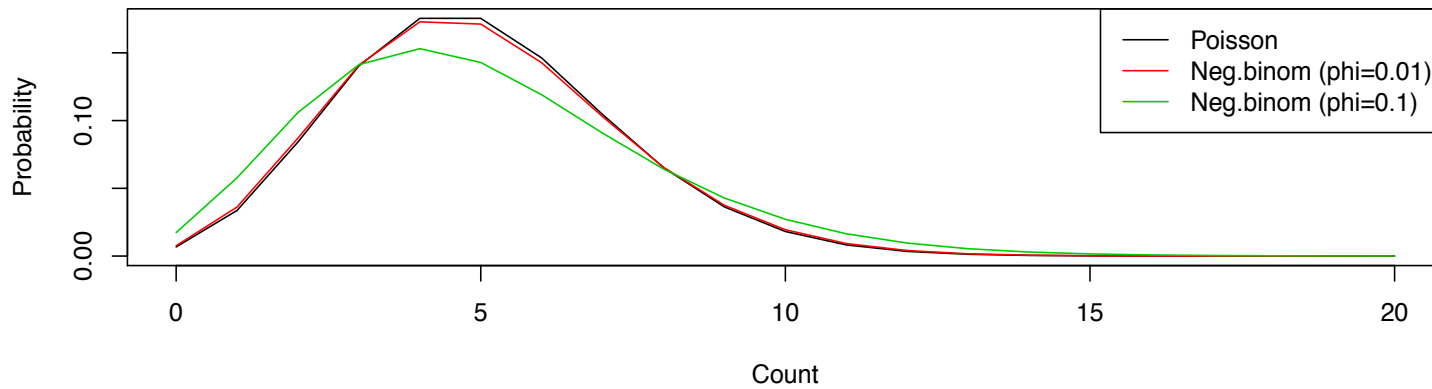


Count distribution

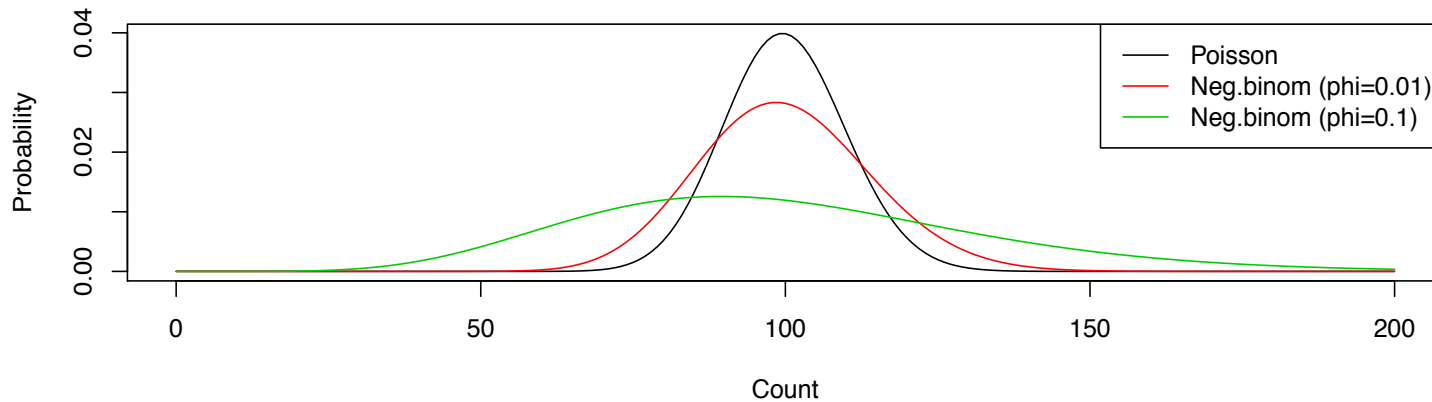
- What is the distribution of counts for a particular RNA
 - Counts from technical replicates are approximately Poisson distributed.
 - Biological replicates exhibit more variance, for which the negative binomial distribution gives a better fit. (Ballard et al, 2010)

Poisson vs negative binomial distribution

Mean=5



Mean=100



The edgeR procedure

- Counts are normalized using TMM (Trimmed mean of M-values)
- A negative binomial distribution is assumed and the extra dispersion parameter is estimated. The parameter can be common to all genes, gene-specific, or a combination