Analyze RNASeq Data from G9P2 RFI Lines Using QuasiSeq Package

June 27, 2014

Table of Contents

- 1. Data Summary
- 2. Model Selection
 - Backward Model Selection
 - AIC Model Selection
- 3. Results

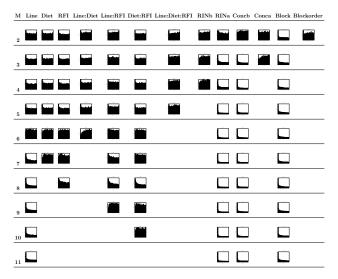
RNASeq Data Summary

- RNASeq data set is a 25320×31 table of count data corresponding to 25320 genes of 31 pigs from 2 Lines (high RFI Line and low RFI Line) and 2 Diets (high energy Diet 1 and low energy Diet 2).
- For Diet 1, the RNA data are from 7 low RFI line pigs and 8 high RFI line pigs. For Diet 2, the RNA data are from 8 low RFI line pigs and 8 high RFI line pigs.
- The number of genes analyzed is 15058. Those are genes with average counts greater than one and for which there are at least four samples with non-zero counts.

Metadata Summary

- The available metadata consists of infomation of 9 covariates for 31 samples of 31 pigs.
- Factors: Diet (2 levels), Line (2 levels), Block (4 levels), Blockorder (8 levels).
- Quantitative covariates: RFI (RFI values), RINb (RIN before globin depletion), RINa (RIN after globin depletion), Concb (RNA Concentration before globin depletion), Conca(RNA Concentration after globin depletion).

Backward Model Selection



AIC Model Selection

 AIC: Obtain Log Maximum Likelihood and calculate AIC for each gene,

$$AIC = -2Log.Max.Lik + 2p,$$

where p is the number of parameters in the model. Then, calculate mean of those values across all genes for each model.

AIC of above models

```
## M2 M3 M4 M5 M6 M7 M8 M9 M10 M11
## 339.9 322.2 321.0 318.7 316.6 314.5 312.8 310.9 309.0 307.0
```

 The best model is Model 11, which is consistent with the Backward Selection method.

Results of Model 11

Estimated number of DE Genes between two RFI Lines

[1] 3451

• When FDR is controlled at 0.05, 0.10, 0.15, the number of DE Genes between two RFI Lines (DEGs) and the number of DE Genes between two RFI Lines with log(fold change) at least 1 (log(FC) > 1) are shown in the table below

	FDR	DEGs	$\log(FC){>}1$
1	0.05	211	38
2	0.10	407	52
3	0.15	689	58