

Analyze RNASeq Data from G9P2 RFI Lines Using QuasiSeq Package (paired end read)

Yet Nguyen

July 15, 2014

Results of Model 7

```
## Covariate Set
covset <- read.csv("covset.csv")
attach(covset)

## Load the result of model 11
load("Model7_resultdat2.RData")

## To get P.values of all Line Testings, use result$P.values[[3]][, 'Line'] To get Q.values of
## Line Testings, use result$Q.values[[3]][, 'Line']

scount <- read.table("paired end uniquely mapped reads count table.txt", header = T)

## List of Genes used to find DE Genes
scount <- scount[-c(which(scount[, 1] %in% "ENSSSCG00000007978"), which(scount[, 1] %in% "ENSSSCG00000014725")),
]
counts <- as.matrix(scount[rowSums(scount[, -1] > 0) > 3 & rowMeans(scount[, -1]) > 8, -1])

## List of Line DE Genes when FDR is controled at 0.05, 0.10, 0.15

degene05 <- which(result$Q.values[[3]][, "Line"] <= 0.05)
head(scount[rownames(counts[degene05, ]), 1]) # First 6 of the List of Line DE Genes when FDR is controled at 0.05

## [1] ENSSSCG00000000035 ENSSSCG00000000036 ENSSSCG00000000037 ENSSSCG00000000062
## [5] ENSSSCG00000000066 ENSSSCG00000000082
## 25322 Levels: ENSSSCG00000000001 ENSSSCG00000000002 ... ENSSSCG000000031070

degene10 <- which(result$Q.values[[3]][, "Line"] <= 0.1)
degene15 <- which(result$Q.values[[3]][, "Line"] <= 0.15)

## Total number of Line DE Genes when FDR is controlled at 0.05, 0.10, 0.15

degene <- c(length(degene05), length(degene10), length(degene15))

## List of Line DE Genes with log2(FC) >=1 when FDR is controlled at 0.05, 0.10, 0.15

lf105 <- degene05[abs(log2(apply(counts[degene05, Line == 1] + 1, 1, mean)/apply(counts[degene05,
Line == 2] + 1, 1, mean))) >= 1]

scount[rownames(counts[lf105, ]), 1] # List of Line DE Genes with log2(FC) >=1 when FDR is controled at 0.05

## [1] ENSSSCG000000000625 ENSSSCG000000000633 ENSSSCG000000000636 ENSSSCG000000000639
## [5] ENSSSCG000000000651 ENSSSCG00000001229 ENSSSCG00000001605 ENSSSCG00000001942
## [9] ENSSSCG00000001943 ENSSSCG00000003595 ENSSSCG00000003965 ENSSSCG00000005385
## [13] ENSSSCG00000006142 ENSSSCG00000006398 ENSSSCG00000008604 ENSSSCG00000009051
## [17] ENSSSCG00000009347 ENSSSCG00000009406 ENSSSCG00000010411 ENSSSCG00000012150
## [21] ENSSSCG00000012960 ENSSSCG00000014123 ENSSSCG00000014562 ENSSSCG00000014599
## [25] ENSSSCG00000014832 ENSSSCG00000014875 ENSSSCG00000015083 ENSSSCG00000015324
```

```

## [29] ENSSSCG000000015707 ENSSSCG000000015766 ENSSSCG000000015854 ENSSSCG000000016216
## [33] ENSSSCG000000016599 ENSSSCG000000016851 ENSSSCG000000016992 ENSSSCG000000017727
## [37] ENSSSCG000000020872 ENSSSCG000000020945 ENSSSCG000000021051 ENSSSCG000000021283
## [41] ENSSSCG000000021572 ENSSSCG000000022216 ENSSSCG000000023108 ENSSSCG000000023256
## [45] ENSSSCG000000023537 ENSSSCG000000023585 ENSSSCG000000024132 ENSSSCG000000024791
## [49] ENSSSCG000000024900 ENSSSCG000000025754 ENSSSCG000000026375 ENSSSCG000000026873
## [53] ENSSSCG000000027157 ENSSSCG000000027378 ENSSSCG000000027611 ENSSSCG000000028095
## [57] ENSSSCG000000028911 ENSSSCG000000029096 ENSSSCG000000030268 ENSSSCG000000030752
## [61] ENSSSCG000000030797
## 25322 Levels: ENSSSCG000000000001 ENSSSCG000000000002 ... ENSSSCG000000031070

lf110 <- degene10[abs(log2(apply(counts[degene10, Line == 1] + 1, 1, mean)/apply(counts[degene10,
  Line == 2] + 1, 1, mean))) >= 1]

scount[rownames(counts[lf110, ]), 1] # List of Line DE Genes with log2(FC) >=1 when FDR is controlled at 0.10

## [1] ENSSSCG000000000607 ENSSSCG000000000625 ENSSSCG000000000633 ENSSSCG000000000636
## [5] ENSSSCG000000000639 ENSSSCG000000000651 ENSSSCG000000000959 ENSSSCG000000000961
## [9] ENSSSCG000000001050 ENSSSCG000000001229 ENSSSCG000000001605 ENSSSCG000000001942
## [13] ENSSSCG000000001943 ENSSSCG000000001978 ENSSSCG000000002620 ENSSSCG000000003595
## [17] ENSSSCG000000003965 ENSSSCG000000005385 ENSSSCG000000006034 ENSSSCG000000006140
## [21] ENSSSCG000000006142 ENSSSCG000000006183 ENSSSCG000000006398 ENSSSCG000000007405
## [25] ENSSSCG000000007463 ENSSSCG000000008604 ENSSSCG000000008771 ENSSSCG000000009051
## [29] ENSSSCG000000009347 ENSSSCG000000009406 ENSSSCG000000009469 ENSSSCG000000010411
## [33] ENSSSCG000000010703 ENSSSCG000000012150 ENSSSCG000000012911 ENSSSCG000000012960
## [37] ENSSSCG000000013497 ENSSSCG000000013575 ENSSSCG000000014066 ENSSSCG000000014123
## [41] ENSSSCG000000014436 ENSSSCG000000014562 ENSSSCG000000014599 ENSSSCG000000014832
## [45] ENSSSCG000000014875 ENSSSCG000000014876 ENSSSCG000000015083 ENSSSCG000000015324
## [49] ENSSSCG000000015706 ENSSSCG000000015707 ENSSSCG000000015766 ENSSSCG000000015854
## [53] ENSSSCG000000016216 ENSSSCG000000016503 ENSSSCG000000016599 ENSSSCG000000016851
## [57] ENSSSCG000000016992 ENSSSCG000000017222 ENSSSCG000000017717 ENSSSCG000000017727
## [61] ENSSSCG000000020872 ENSSSCG000000020945 ENSSSCG000000021051 ENSSSCG000000021283
## [65] ENSSSCG000000021368 ENSSSCG000000021572 ENSSSCG000000022216 ENSSSCG000000023108
## [69] ENSSSCG000000023127 ENSSSCG000000023256 ENSSSCG000000023537 ENSSSCG000000023585
## [73] ENSSSCG000000024132 ENSSSCG000000024136 ENSSSCG000000024791 ENSSSCG000000024900
## [77] ENSSSCG000000025631 ENSSSCG000000025754 ENSSSCG000000026087 ENSSSCG000000026375
## [81] ENSSSCG000000026873 ENSSSCG000000027157 ENSSSCG000000027378 ENSSSCG000000027611
## [85] ENSSSCG000000028095 ENSSSCG000000028911 ENSSSCG000000029096 ENSSSCG000000029147
## [89] ENSSSCG000000030268 ENSSSCG000000030752 ENSSSCG000000030797 ENSSSCG000000030951
## 25322 Levels: ENSSSCG000000000001 ENSSSCG000000000002 ... ENSSSCG000000031070

lf115 <- degene15[abs(log2(apply(counts[degene15, Line == 1] + 1, 1, mean)/apply(counts[degene15,
  Line == 2] + 1, 1, mean))) >= 1]

scount[rownames(counts[lf115, ]), 1] # List of Line DE Genes with log2(FC) >=1 when FDR is controlled at 0.15

## [1] ENSSSCG000000000607 ENSSSCG000000000625 ENSSSCG000000000633 ENSSSCG000000000636
## [5] ENSSSCG000000000639 ENSSSCG000000000651 ENSSSCG000000000959 ENSSSCG000000000961
## [9] ENSSSCG000000001050 ENSSSCG000000001229 ENSSSCG000000001605 ENSSSCG000000001942
## [13] ENSSSCG000000001943 ENSSSCG000000001978 ENSSSCG000000002620 ENSSSCG000000003595
## [17] ENSSSCG000000003965 ENSSSCG000000004191 ENSSSCG000000005211 ENSSSCG000000005385
## [21] ENSSSCG000000006034 ENSSSCG000000006140 ENSSSCG000000006142 ENSSSCG000000006183
## [25] ENSSSCG000000006398 ENSSSCG000000007371 ENSSSCG000000007405 ENSSSCG000000007463
## [29] ENSSSCG000000008203 ENSSSCG000000008604 ENSSSCG000000008771 ENSSSCG000000009051
## [33] ENSSSCG000000009347 ENSSSCG000000009406 ENSSSCG000000009469 ENSSSCG000000009497
## [37] ENSSSCG000000009965 ENSSSCG000000010411 ENSSSCG000000010703 ENSSSCG000000012100
## [41] ENSSSCG000000012150 ENSSSCG000000012911 ENSSSCG000000012960 ENSSSCG000000013497
## [45] ENSSSCG000000013575 ENSSSCG000000014066 ENSSSCG000000014123 ENSSSCG000000014436
## [49] ENSSSCG000000014562 ENSSSCG000000014599 ENSSSCG000000014832 ENSSSCG000000014875
## [53] ENSSSCG000000014876 ENSSSCG000000015083 ENSSSCG000000015324 ENSSSCG000000015706
## [57] ENSSSCG000000015707 ENSSSCG000000015766 ENSSSCG000000015854 ENSSSCG000000016216
## [61] ENSSSCG000000016503 ENSSSCG000000016599 ENSSSCG000000016851 ENSSSCG000000016992
## [65] ENSSSCG000000017222 ENSSSCG000000017717 ENSSSCG000000017727 ENSSSCG000000020872
## [69] ENSSSCG000000020945 ENSSSCG000000021051 ENSSSCG000000021283 ENSSSCG000000021368
## [73] ENSSSCG000000021404 ENSSSCG000000021572 ENSSSCG000000022216 ENSSSCG000000022500
## [77] ENSSSCG000000023108 ENSSSCG000000023127 ENSSSCG000000023256 ENSSSCG000000023537
## [81] ENSSSCG000000023585 ENSSSCG000000024132 ENSSSCG000000024136 ENSSSCG000000024791
## [85] ENSSSCG000000024900 ENSSSCG000000025631 ENSSSCG000000025754 ENSSSCG000000026087
## [89] ENSSSCG000000026375 ENSSSCG000000026873 ENSSSCG000000027157 ENSSSCG000000027378

```

```
## [93] ENSSSCG00000027611 ENSSSCG00000028095 ENSSSCG00000028911 ENSSSCG00000029096
## [97] ENSSSCG00000029147 ENSSSCG00000030268 ENSSSCG00000030752 ENSSSCG00000030797
## [101] ENSSSCG00000030951
## 25322 Levels: ENSSSCG00000000001 ENSSSCG00000000002 ... ENSSSCG00000031070

## Total number of Line DE Genes with log2(FC) >=1 when FDR is controlled at 0.05, 0.10,
## 0.15

lf1 <- c(length(lf105), length(lf110), length(lf115))
```

```
# Summary table
out <- data.frame(FDR = c(0.05, 0.1, 0.15), degene = degene, lf1 = lf1)
colnames(out) <- c("FDR", "DEGs", "log2(FC)>=1")
xtable(out)
```

	FDR	DEGs	log2(FC)>=1
1	0.05	771	61
2	0.10	1837	92
3	0.15	2936	101

1 List of Gene in Analysis and List of Pvalues, Qvalues, logFC

```
covset <- read.csv("covset.csv")
attach(covset)

## The following objects are masked from covset (position 3):
##
## baso, Block, Blockorder, Conca, Concb, Diet, eosi, iddam, idpig, idsire,
## lbaso, leosi, Line, llymp, lmono, lneut, llymp, mono, neut, RFI, RINa, RINb,
## X

## Load the result of model 11
load("Model7_resultdat2.RData")

## To get P.values of all Line Testings, use result$P.values[[3]][, 'Line'] To get Q.values of
## Line Testings, use result$Q.values[[3]][, 'Line']

scount <- read.table("paired end uniquely mapped reads count table.txt", header = T)

## List of Genes used to find DE Genes
scount <- scount[-c(which(scount[, 1] %in% "ENSSSCG00000007978"), which(scount[, 1] %in% "ENSSSCG00000014725")),
]

kept_gene <- scount[rowSums(scount[, -1] > 0) > 3 & rowMeans(scount[, -1]) > 8, ]

load("Model7_resultdat2.RData")
Line_pvalue <- result$P.values[[3]][, "Line"]
Line_qvalue <- result$Q.values[[3]][, "Line"]
log2fc <- log2(apply(kept_gene[, -1][, Line == 1] + 1, 1, mean)/apply(kept_gene[, -1][, Line ==
2] + 1, 1, mean))

kept_gene_ex <- cbind(kept_gene, line_pvalue = Line_pvalue, line_qvalue = Line_qvalue, log2fc = log2fc)
de_gene_ex <- subset(kept_gene_ex, line_qvalue <= 0.05)

de_log2_gene_ex <- subset(de_gene_ex, abs(log2fc) >= 1)
dim(de_log2_gene_ex)

## [1] 61 35
```