

## combined pcr data

May 9, 2014

```
> input_file <- 'qRT-PCR combined data.csv'
> #import our data
> data <- read.csv(input_file)
> genes.of.interest <- c("Actb", "Cdkn1a", "Mef2c", "Myf5", "Myod1", "Myog" )
> #remove genes that we are not interested in
> data.of.interest <- droplevels(subset(data, gene %in% genes.of.interest))
> data.of.interest$Rapamycin <- grepl("RAPA", data.of.interest$treatment)
```

This uses the input file qRT-PCR combined data.csv. This is located in /Users/innocenceharvey/DrosophilaMuscleFunction/CellCulture. It was most recently run on Fri May 9 16:23:01 2014.

```
> #make the table
> shapiro.results <- data.frame(row.names=levels(data.of.interest$gene),
+                               pval.dms0 = rep(NA, length(levels(data.of.interest$gene))),
+                               pval.rapa = rep(NA, length(levels(data.of.interest$gene))))
> #first we subset to get data frame with just actin and dms0
> #subset(data.of.interest, gene=='Actb'&treatment=="DMSO")
> #get only the expression values
> #subset(data.of.interest, gene=='Actb'&treatment=="DMSO")$value
> #run a shapiro test on those values shapiro.test(subset(data.of.interest, gene=='Actb'&treatment=="DMSO")$value)
> #shapiro.test(subset(data.of.interest, gene=='Actb'&treatment=="DMSO")$value)$p.value
>
> #create a loop for DMSO
> for (test.gene in levels(data.of.interest$gene)) {
+   shapiro.results[test.gene, 'pval.dms0'] <- shapiro.test(subset(data.of.interest, gene==test.gene & treatment=="DMSO")$value)$p.value
+   shapiro.results[test.gene, 'pval.rapa'] <- shapiro.test(subset(data.of.interest, gene==test.gene & treatment=="RAPA")$value)$p.value
+ }
> shapiro.results$Normal <- apply(shapiro.results, 1, min) > 0.05
> shapiro.results
```

	pval.dms0	pval.rapa	Normal
Actb	0.8772361477	0.005041333	FALSE
Cdkn1a	0.5828581004	0.585257352	TRUE
Mef2c	0.7960543063	0.646668648	TRUE
Myf5	0.0465770009	0.573296479	FALSE

```

Myod1  0.0001314389 0.959376668 FALSE
Myog   0.0426887592 0.063660610 FALSE

> #test equal variance with Levene's test
> library(car)
> #make a table for results
> levene.results <- data.frame(row.names=levels(data.of.interest$gene),
+                               pval = rep(NA, length(levels(data.of.interest$gene))))
> for (gene.of.interest in levels(data.of.interest$gene)) {
+   levene.results[gene.of.interest, 'pval'] <- with(subset(data.of.interest, gene==gene.of.interest),
+   }
> levene.results$Equal.Variance <- apply(levene.results, 1, min) > 0.05
> levene.results

      pval Equal.Variance
Actb    0.0538042380      TRUE
Cdkn1a  0.7408066705      TRUE
Mef2c   0.0002193127     FALSE
Myf5    0.0689679076      TRUE
Myod1   0.2574946928      TRUE
Myog    0.0658846581      TRUE

> #do a t-test
> #first see if there is a normal distribution, if there is then could do a t-test. If not
>
> test.results <- data.frame(row.names=levels(data.of.interest$gene),
+                             pval = rep(NA, length(levels(data.of.interest$gene))),
+                             test = rep(NA, length(levels(data.of.interest$gene))))
> for (gene.of.interest in levels(data.of.interest$gene)) {
+   if (!(shapiro.results[gene.of.interest, 'Normal'])) {
+     test.results[gene.of.interest, 'pval'] <- wilcox.test(value~Rapamycin, data=subset(data.of.interest, gene==gene.of.interest))
+     test.results[gene.of.interest, 'test'] <- 'Wilcoxon Rank Sum Test'
+   }
+   else {
+     if (levene.results[gene.of.interest, 'Equal.Variance']){
+       test.results[gene.of.interest, 'pval'] <- t.test(value~Rapamycin, data=subset(data.of.interest, gene==gene.of.interest))
+       test.results[gene.of.interest, 'test'] <- 'Welch T-Test'
+     }
+     else {
+       test.results[gene.of.interest, 'pval'] <- t.test(value~Rapamycin, data=subset(data.of.interest, gene==gene.of.interest))
+       test.results[gene.of.interest, 'test'] <- 'Student T-Test'
+     }
+   }
+ }
> #run a one-way anova on all treatment groups from each experiment to see if they vary
> treatment.aov <- aov(value~gene * treatment*Exp, data=data.of.interest)
> summary (treatment.aov)

```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
gene	5	18.380	3.676	39.433	< 2e-16 ***
treatment	3	2.688	0.896	9.610	1.18e-05 ***
Exp	4	1.722	0.431	4.619	0.001794 **
gene:treatment	13	15.703	1.208	12.957	< 2e-16 ***
gene:Exp	18	2.556	0.142	1.523	0.096578 .
treatment:Exp	4	2.038	0.510	5.466	0.000492 ***
gene:treatment:Exp	16	1.967	0.123	1.319	0.199912
Residuals	104	9.695	0.093		

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

> treatment.aov

Call:

aov(formula = value ~ gene \* treatment \* Exp, data = data.of.interest)

Terms:

	gene	treatment	Exp	gene:treatment	gene:Exp
Sum of Squares	18.380268	2.687616	1.722262	15.702784	2.555761
Deg. of Freedom	5	3	4	13	18

	treatment:Exp	gene:treatment:Exp	Residuals
Sum of Squares	2.038339	1.967020	9.695072
Deg. of Freedom	4	16	104

Residual standard error: 0.3053225

56 out of 120 effects not estimable

Estimated effects may be unbalanced

> #because there is an interaction among all of the variables we need to perform a two way a

>

> summary(data.of.interest\$Exp)

20-Feb dose response	feb28 plate 2	mar7 plate 2	Mar7 plate1
36	24	36	36

> summary(aov(value~gene\*treatment, data=subset(data.of.interest, Exp=='feb28 plate 2')))

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
gene	5	4.600	0.9200	12.622	4.47e-06 ***
treatment	1	0.195	0.1952	2.678	0.115
gene:treatment	5	4.600	0.9200	12.622	4.47e-06 ***
Residuals	24	1.749	0.0729		

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

> summary(aov(value~gene\*treatment, data=subset(data.of.interest, Exp=='mar7 plate 2')))

```

      Df Sum Sq Mean Sq F value    Pr(>F)
gene      5   3.270   0.6540    17.50 2.58e-07 ***
treatment 1   0.832   0.8320    22.26 8.50e-05 ***
gene:treatment 5   3.270   0.6540    17.50 2.58e-07 ***
Residuals 24   0.897   0.0374
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> summary(aov(value~gene*treatment, data=subset(data.of.interest, Exp=='Mar7 plate1'))

      Df Sum Sq Mean Sq F value    Pr(>F)
gene      5  2.4501   0.4900    18.89 1.27e-07 ***
treatment 1  0.5681   0.5681    21.91 9.36e-05 ***
gene:treatment 5  2.4501   0.4900    18.89 1.27e-07 ***
Residuals 24  0.6225   0.0259
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> summary(aov(value~gene*treatment, data=subset(data.of.interest, Exp=='dose response ')))

      Df Sum Sq Mean Sq F value    Pr(>F)
gene      3  2.5749   0.8583    53.585 1.22e-05 ***
treatment 3  1.8075   0.6025    37.616 4.60e-05 ***
gene:treatment 9  1.3712   0.1524     9.512 0.0021 **
Residuals  8  0.1281   0.0160
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> summary(aov(value~gene*treatment, data=subset(data.of.interest, Exp=='20-Feb'))

      Df Sum Sq Mean Sq F value    Pr(>F)
gene      5   7.559   1.5118     5.761 0.00125 **
treatment 1   0.023   0.0226     0.086 0.77158
gene:treatment 5   7.559   1.5118     5.761 0.00125 **
Residuals 24   6.298   0.2624
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> #test that the data are normally distributed
> #run shapiro
> shapiro.test(residuals(aov(value~gene*treatment, data=subset(data.of.interest, Exp=='feb20

Shapiro-Wilk normality test

data:  residuals(aov(value ~ gene * treatment, data = subset(data.of.interest, Exp == "f
W = 0.8009, p-value = 1.699e-05

> shapiro.test(residuals(aov(log(value)~gene*treatment, data=subset(data.of.interest, Exp==

```

```

Shapiro-Wilk normality test

data: residuals(aov(log(value) ~ gene * treatment, data = subset(data.of.interest, Exp
W = 0.9764, p-value = 0.6229

> #run levene's test for equal variance
> library(car)
> leveneTest(log(value)~gene*treatment, data=subset(data.of.interest, Exp=='feb28 plate 2'))

Levene's Test for Homogeneity of Variance (center = median)
      Df F value Pr(>F)
group 11  0.5941  0.815
      24

> #so the assumptions for this model are OK
> summary(aov(log(value)~gene*treatment, data=subset(data.of.interest, Exp=='feb28 plate 2'))

              Df Sum Sq Mean Sq F value    Pr(>F)
gene             5   7.425   1.4851    36.32 1.96e-10 ***
treatment        1   2.445   2.4449    59.80 5.73e-08 ***
gene:treatment    5   7.404   1.4807    36.22 2.02e-10 ***
Residuals       24   0.981   0.0409
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> #pos hoc testing for this model
> feb28.t.tests <- data.frame(row.names=levels(data.of.interest$gene))
> for (test.gene in levels(data.of.interest$gene)) {
+ feb28.t.tests[test.gene,'pval'] <- t.test(log(value)~treatment, data=subset(data.of.interest, Exp=='feb28 plate 2'))
+ }
> feb28.t.tests

              pval
Actb    0.0220569009
Cdkn1a  0.0072971453
Mef2c   0.0009335880
Myf5    0.0223318648
Myod1   0.1805853201
Myog    0.0007129232

```

```

> shapiro.test(residuals(aov(value~gene*treatment, data=subset(data.of.interest, Exp=='mar7

Shapiro-Wilk normality test

data: residuals(aov(value ~ gene * treatment, data = subset(data.of.interest, Exp == "mar7
W = 0.9278, p-value = 0.0215

> shapiro.test(residuals(aov(log(value)~gene*treatment, data=subset(data.of.interest, Exp==

```

```

Shapiro-Wilk normality test

data: residuals(aov(log(value) ~ gene * treatment, data = subset(data.of.interest, Exp
W = 0.9761, p-value = 0.6135

> leveneTest(log(value)~gene*treatment, data=subset(data.of.interest, Exp=='mar7 plate 2'))

Levene's Test for Homogeneity of Variance (center = median)
      Df F value Pr(>F)
group 11  0.7292  0.701
      24

> #assumptions for this model are OK
> summary(aov(log(value)~gene*treatment, data=subset(data.of.interest, Exp=='mar7 plate 2'))

      Df Sum Sq Mean Sq F value    Pr(>F)
gene      5  9.163    1.833   41.05 5.40e-11 ***
treatment  1  5.497    5.497  123.13 6.21e-11 ***
gene:treatment  5  8.883    1.777   39.79 7.50e-11 ***
Residuals  24  1.071    0.045
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> #post hoc testing for this model
> mar7.2.t.tests <- data.frame(row.names=levels(data.of.interest$gene))
> for (test.gene in levels(data.of.interest$gene)) {
+ mar7.2.t.tests[test.gene, 'pval'] <- t.test(log(value)~treatment, data=subset(data.of.inter
+ }
> mar7.2.t.tests

      pval
Actb 0.0077264194
Cdkn1a 0.0119573375
Mef2c 0.0003566990
Myf5 0.0006424166
Myod1 0.2006887579
Myog 0.0194950170

> #
> shapiro.test(residuals(aov(value~gene*treatment, data=subset(data.of.interest, Exp=='Mar7

Shapiro-Wilk normality test

data: residuals(aov(value ~ gene * treatment, data = subset(data.of.interest, Exp == "M
W = 0.9762, p-value = 0.6176

> leveneTest(value~gene*treatment, data=subset(data.of.interest, Exp=='Mar7 plate1'))

```

Levene's Test for Homogeneity of Variance (center = median)

```
Df F value Pr(>F)
group 11 0.7787 0.6578
      24
```

```
> #assumptions for this model are OK (no log transform needed)
> summary(aov(value~gene*treatment, data=subset(data.of.interest, Exp=='Mar7 plate1')))
```

```
          Df Sum Sq Mean Sq F value    Pr(>F)
gene          5  2.4501   0.4900    18.89 1.27e-07 ***
treatment      1  0.5681   0.5681    21.91 9.36e-05 ***
gene:treatment  5  2.4501   0.4900    18.89 1.27e-07 ***
Residuals     24  0.6225   0.0259
```

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

```
> #post hoc testing for this model
> mar7.1.t.tests <- data.frame(row.names=levels(data.of.interest$gene))
> for (test.gene in levels(data.of.interest$gene)) {
+ mar7.1.t.tests[test.gene,'pval'] <- t.test(value~treatment, data=subset(data.of.interest,
+ })
> mar7.1.t.tests
```

```
          pval
Actb    0.003494267
Cdkn1a  0.024727887
Mef2c   0.001769465
Myf5    0.048016215
Myod1   0.418954729
Myog    0.004332754
```

```
> shapiro.test(residuals(aov(value~gene*treatment, data=subset(data.of.interest, Exp=='20-Feb'))
```

Shapiro-Wilk normality test

```
data: residuals(aov(value ~ gene * treatment, data = subset(data.of.interest, Exp == "20-Feb"))
W = 0.7907, p-value = 1.071e-05
```

```
> shapiro.test(residuals(aov(log(value)~gene*treatment, data=subset(data.of.interest, Exp=='20-Feb'))
```

Shapiro-Wilk normality test

```
data: residuals(aov(log(value) ~ gene * treatment, data = subset(data.of.interest, Exp == "20-Feb"))
W = 0.9749, p-value = 0.5724
```

```
> leveneTest(log(value)~gene*treatment, data=subset(data.of.interest, Exp=='20-Feb'))
```

Levene's Test for Homogeneity of Variance (center = median)

```

      Df F value Pr(>F)
group 11  0.6446 0.7737
      24

```

> #assumptions for this model are OK

> summary(aov(log(value)~gene\*treatment, data=subset(data.of.interest, Exp=='20-Feb')))

```

              Df Sum Sq Mean Sq F value    Pr(>F)
gene              5   7.013   1.4027    16.95 3.43e-07 ***
treatment         1   0.840   0.8403    10.16 0.00396 **
gene:treatment    5   6.745   1.3490    16.30 4.88e-07 ***
Residuals        24   1.986   0.0827

```

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

> #pos hoc testing for this model

> feb20.t.tests <- data.frame(row.names=levels(data.of.interest\$gene))

> for (test.gene in levels(data.of.interest\$gene)) {

+ feb20.t.tests[test.gene,'pval'] <- t.test(log(value)~treatment, data=subset(data.of.interest,

+ })

> feb20.t.tests

```

              pval
Actb    0.047362045
Cdkn1a  0.098430095
Mef2c   0.004045253
Myf5    0.035140962
Myod1   0.044237560
Myog    0.031918406

```

> shapiro.test(residuals(aov(value~gene\*treatment, data=subset(data.of.interest, Exp=='dose

Shapiro-Wilk normality test

data: residuals(aov(value ~ gene \* treatment, data = subset(data.of.interest, Exp == "dose

W = 0.7866, p-value = 0.0001792

> shapiro.test(residuals(aov(log(value)~gene\*treatment, data=subset(data.of.interest, Exp=='dose

Shapiro-Wilk normality test

data: residuals(aov(log(value) ~ gene \* treatment, data = subset(data.of.interest, Exp == "dose

W = 0.9458, p-value = 0.2191

> leveneTest(log(value)~gene\*treatment, data=subset(data.of.interest, Exp=='dose response '))



Levene's Test for Homogeneity of Variance (center = median)

```

      Df    F value    Pr(>F)
group 15 3.1203e+30 < 2.2e-16 ***
      8

```

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

```
> #does not meet levenes test for equal variance so must run Welch's...?
```

```
> #dose response need to do anova instead of t-test due to 3 different groups
```

```
> dose.response.aov <- data.frame(row.names=levels(data.of.interest$gene))
```

```
> kruskal.test(value~treatment, data=subset(data.of.interest, Exp=='dose response '&gene=='treatment'))
```

```
[1] 0.3116156
```

```
> for (test.gene in c("Actb","Mef2c","Myod1","Myog" )) {
```

```
+ dose.response.aov[test.gene,'pval-low'] <- t.test(mu=1, subset(data.of.interest, Exp=='dose response '&gene==test.gene))
```

```
+ dose.response.aov[test.gene,'pval-high'] <- t.test(mu=1, subset(data.of.interest, Exp=='dose response '&gene==test.gene))
```

```
+ }
```

```
> dose.response.aov
```

	pval-low	pval-high
Actb	0.133498028	0.781207337
Cdkn1a	NA	NA
Mef2c	0.006423015	0.004777345
Myf5	NA	NA
Myod1	0.030032472	0.104138664
Myog	0.013727306	0.034779776

```
> library(plyr)
```

```
> complete.experiments <- c('20-Feb','feb28 plate 2','mar7 plate 2', 'Mar7 plate1')
```

```
> summary.data <- ddply(subset(data.of.interest, Exp %in% complete.experiments), .(gene, Rapamycin),
```

```
+   mean = mean(value),
```

```
+   sd = sd(value),
```

```
+   se = sd(value)/sqrt(length(value)),
```

```
+   n = length(value))
```

```
>
```

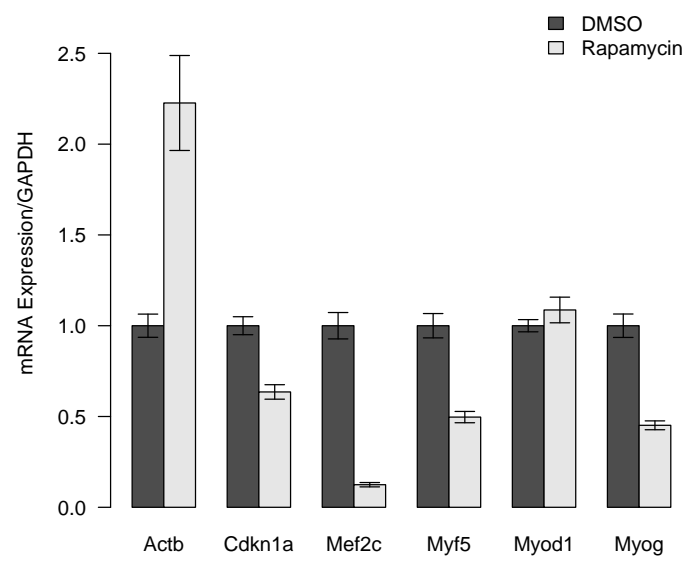


Figure 1: Combined Data from All Experiments

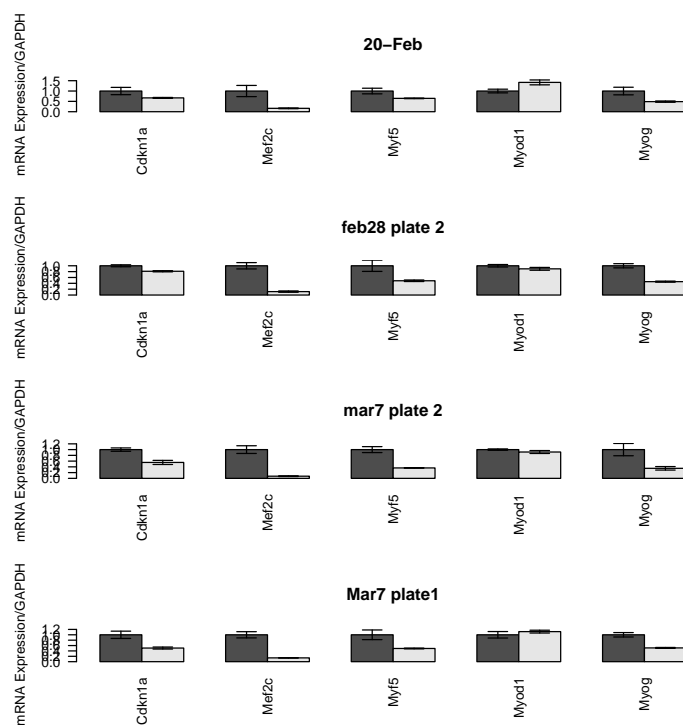


Figure 2: Data for Each Experiment

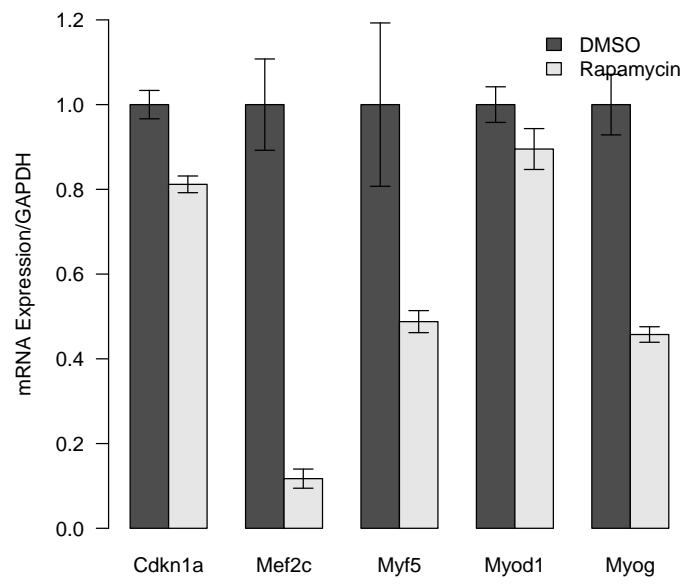


Figure 3: Feb 28 Data