

hw5

Joshua Burkhardt

December 8, 2015

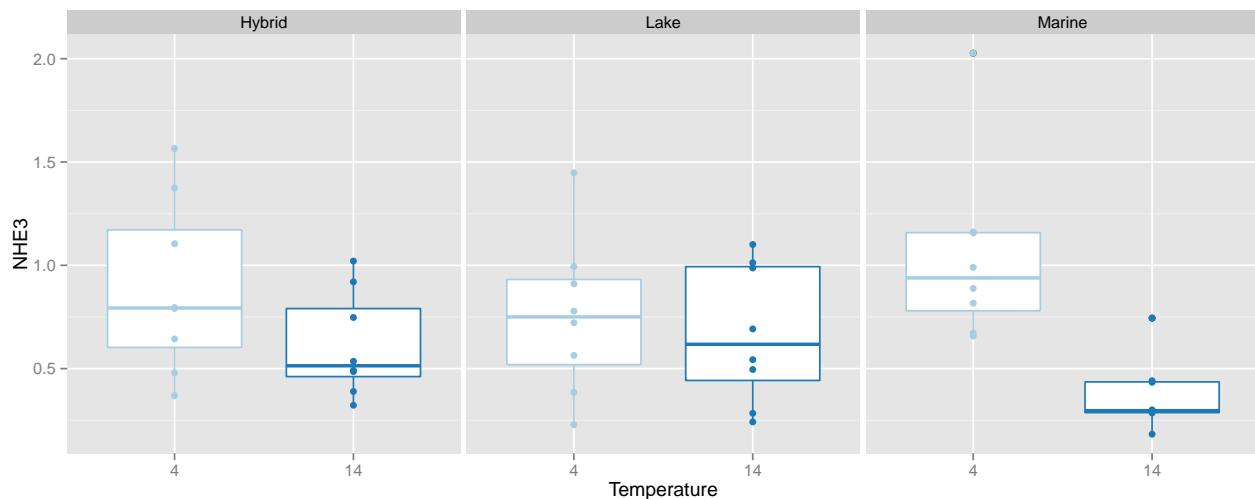
Homework 5: ANOVA

Analysis of the expression of an isoform of the electroneutral Na⁺/H⁺ exchanger (NHE3)

Re-create the following two plots (note: don't worry about colors, but you should be able to make plots that have the same x/y axes, faceting, and geoms using ggplot2):

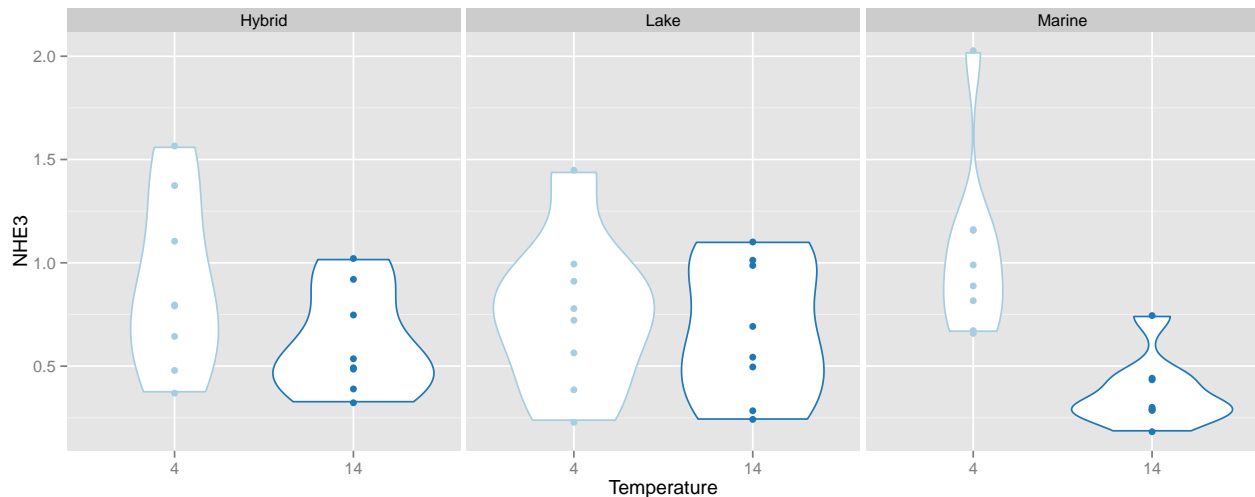
```
# for comma separated values
stickle <- read.table("~/Software/Probability/hw5/qPCR_data.csv",
                      header = TRUE, sep = ",")
stickle$Temperature <- as.factor(stickle$Temperature)
```

```
stickle %>%
  group_by(Temperature) %>%
  ggplot() +
  facet_wrap(~ Ecotype) +
  geom_boxplot(aes(x=Temperature, y=NHE3, color=Temperature)) +
  geom_point(aes(x=Temperature, y=NHE3, color=Temperature)) +
  scale_color_brewer(palette="Paired") +
  xlab("Temperature") +
  theme(legend.position="none")
```



```
stickle %>%
  group_by(Temperature) %>%
  ggplot() +
  facet_wrap(~ Ecotype) +
  geom_violin(aes(x=Temperature, y=NHE3, color=Temperature)) +
```

```
geom_point(aes(x=Temperature, y=NHE3,color=Temperature)) +
scale_color_brewer(palette="Paired") +
xlab("Temperature") +
theme(legend.position="none")
```



On your own

This summary of results appears to focus on a two-way ANOVA predicting NHE3 gene expression levels from Ecotype and Temperature. Run this model on your own: interpret the regression coefficients from the general linear model, using `contr.sum` for all factor predictors, and conduct any post-hoc comparisons as you see appropriate to aid in interpreting the model.

```
nhe3_sum <- lm(NHE3 ~ Temperature * Ecotype,
               data = stickle,
               contrasts = list(Ecotype = contr.sum,
                               Temperature = contr.sum))
tidy(nhe3_sum)
```

	term	estimate	std.error	statistic	p.value
1	(Intercept)	0.72421309	0.05015096	14.4406639	0.00000000000000006881586
2	Temperature1	0.17256645	0.05015096	3.4409404	0.001323275632469792092646
3	Ecotype1	0.02805772	0.07092416	0.3956017	0.694401289411864852674228
4	Ecotype2	-0.01249763	0.07092416	-0.1762112	0.860974937159666442454409
5	Temperature1:Ecotype1	-0.03446151	0.07092416	-0.4858924	0.629568929580255787215037
6	Temperature1:Ecotype2	-0.13052787	0.07092416	-1.8403865	0.072784401825830249976867

```
attr(model.matrix(nhe3_sum), "contrasts") # contrast check
```

```
$Temperature
  [,1]
4     1
14    -1
```

```
$Ecotype
      [,1] [,2]
Hybrid    1    0
Lake      0    1
Marine   -1   -1
```

```
#interpreting regression coefficients
grandmean <- mean(stickle$NHE3)
```

```
# main effects
bTemp1 <- stickle %>%
  filter(Temperature == 4) %>%
  summarise(bTemp1 = mean(NHE3) - grandmean)
Ecotype1 <- stickle %>%
  filter(Ecotype == "Hybrid") %>%
  summarise(bEcotype1 = mean(NHE3) - grandmean)
Ecotype2 <- stickle %>%
  filter(Ecotype == "Lake") %>%
  summarise(bEcotype2 = mean(NHE3) - grandmean)

cbind(grandmean, bTemp1, Ecotype1, Ecotype2)
```

```
      grandmean    bTemp1  bEcotype1  bEcotype2
1 0.7242131 0.1725665 0.02805772 -0.01249763
```

```
Anova(nhe3_sum,type=2,white.adjust=TRUE)
```

Analysis of Deviance Table (Type II tests)

Response: NHE3

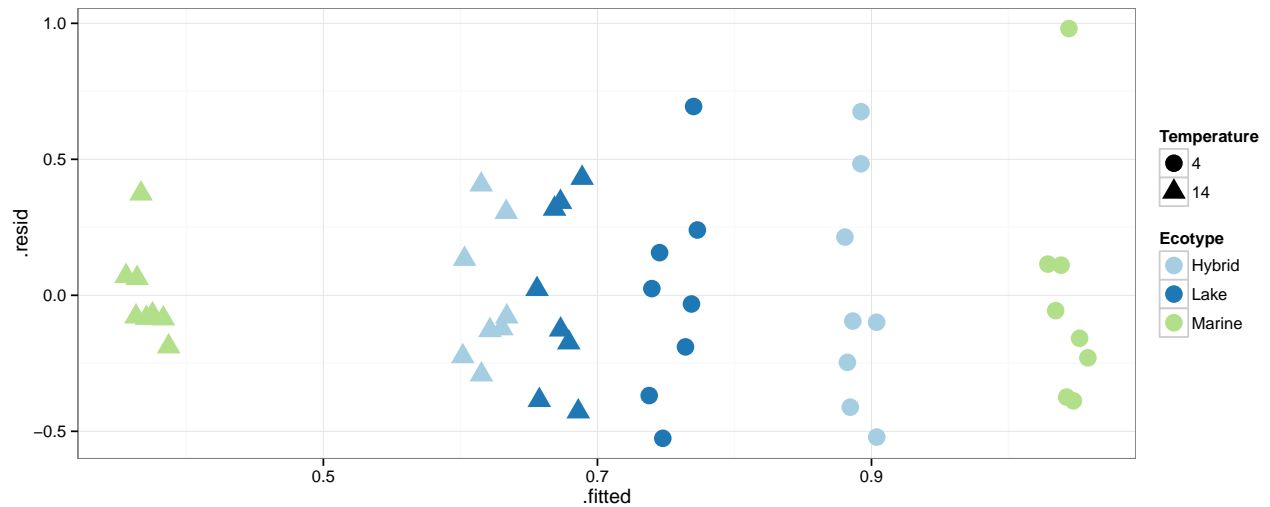
	Df	F	Pr(>F)
Temperature	1	11.2473	0.001698 **
Ecotype	2	1.6883	0.197134
Temperature:Ecotype	2	2.6904	0.079538 .
Residuals	42		

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

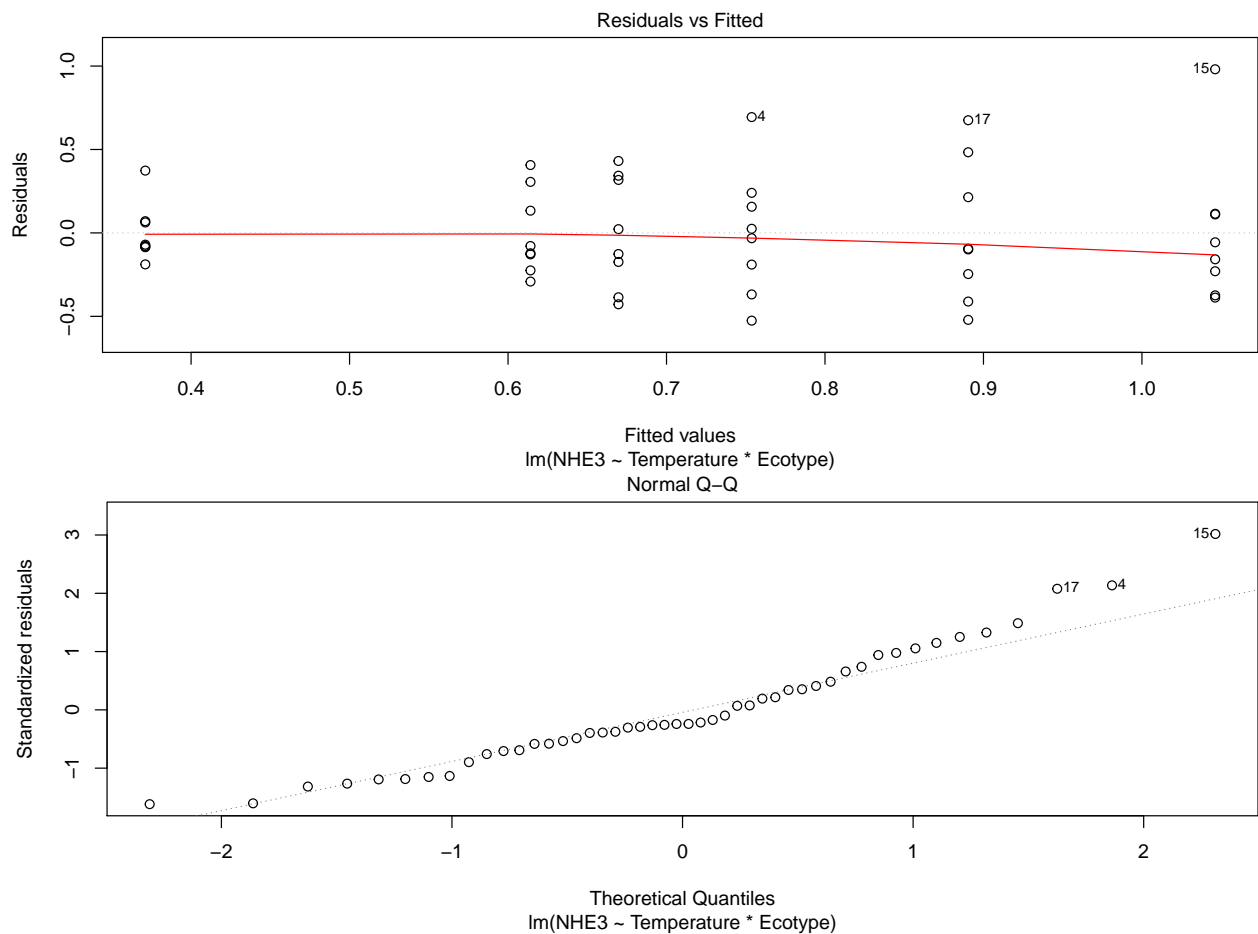
It appears the mean NHE3 differed depending on Temperature but not Ecotype.

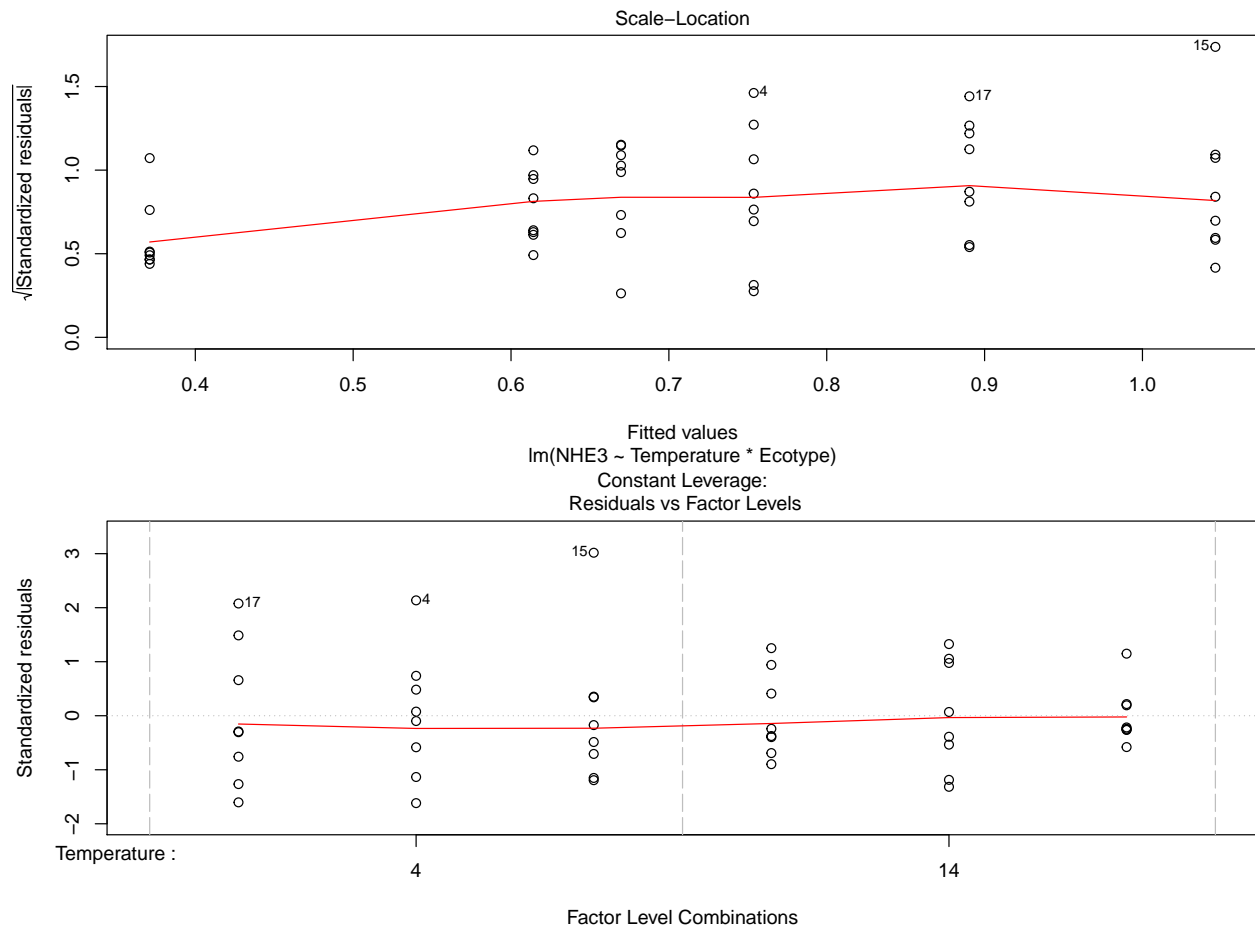
```
nhe3_twowayvars <- augment(nhe3_sum)
nhe3_resid <- ggplot(data = nhe3_twowayvars, aes(x = .fitted,
  y = .resid,
  colour = Ecotype)) +
```

```
geom_point(aes(shape = Temperature), size = 5, position = position_jitter(width = .02)) +
scale_colour_brewer(type = "qual", palette = 3) +
theme_bw()
suppressWarnings(print(nhe3_resid))
```



```
plot(nhe3_sum)
```





```
testInteractions(nhe3_sum, pairwise = "Temperature", adjustment = "BH")
```

F Test:

P-value adjustment method: BH

	Value	Df	Sum of Sq	F	Pr(>F)
4-14	0.34513	1	1.4294	11.84	0.001323 **
Residuals		42	5.0705		

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

```
testInteractions(nhe3_sum, pairwise = "Temperature", adjustment = "holm")
```

F Test:

P-value adjustment method: holm

	Value	Df	Sum of Sq	F	Pr(>F)
4-14	0.34513	1	1.4294	11.84	0.001323 **
Residuals		42	5.0705		

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

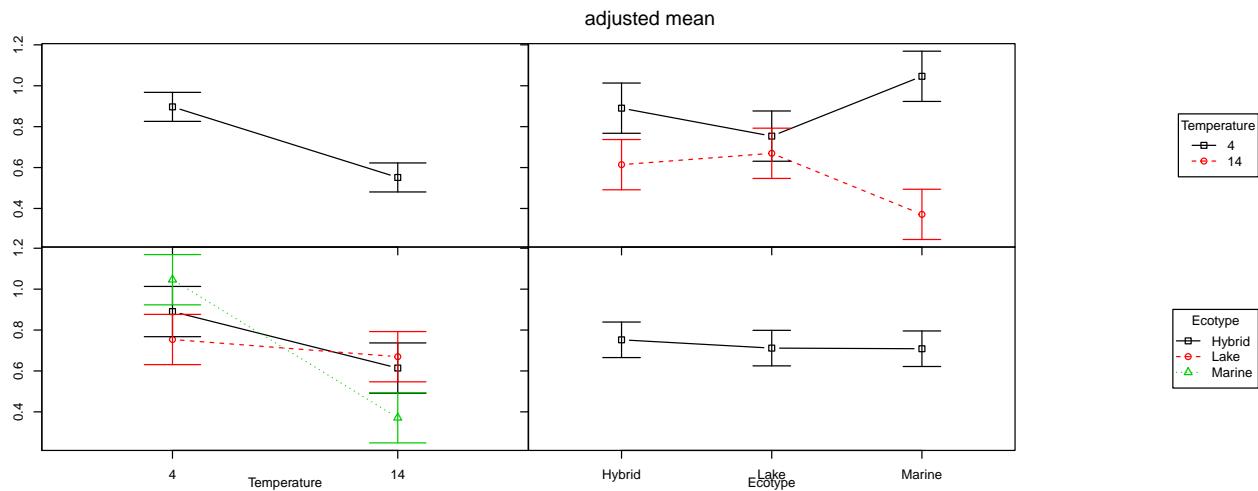
```
plot(interactionMeans(nhe3_sum))
```

```
testInteractions(nhe3_sum, fixed = "Ecotype", pairwise = "Temperature", adjustment = "bonferroni")
```

P-value adjustment method: bonferroni

	Value	Df	Sum of Sq	F	Pr(>F)
4-14 : Hybrid	0.27621	1	0.3052	2.5278	0.358074
4-14 : Lake	0.08408	1	0.0283	0.2342	1.000000
4-14 : Marine	0.67511	1	1.8231	15.1012	0.001067 **
Residuals		42	5.0705		

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



The authors' statement, "substantial change with temperature in the marine ecotype that was not observed in the freshwater ecotype under the conditions tested here," seems to be supported by our interaction test. Indeed, the freshwater ecotype does not exhibit a significant interaction between temperature and NHE3 expression, while the marine ecotype does.