Statistical Analysis

The Role of the Glyoxylate Shunt in the Acclimation to Iron Limitation in Marine Heterotrophic Bacteria

16/02/2018

1. Bioreporter ICL-luc

Graphs and statistical analysis conducted on the results obtained from bioreporter measuring luminescence as a proxy for icl-expression. 3 experiments were conducted using the bioreporter.

Carbon-Sources Differences in luminescence under 4x different carbon sources.

Carbon-Concentration Differences in luminescence under 2x different glucose concentrations.

Fe-limitation Differences in luminescence under 5x different levels of Fe-limitation through the addition of DFOB.

1a. Luminescence for Different Carbon Concentrations

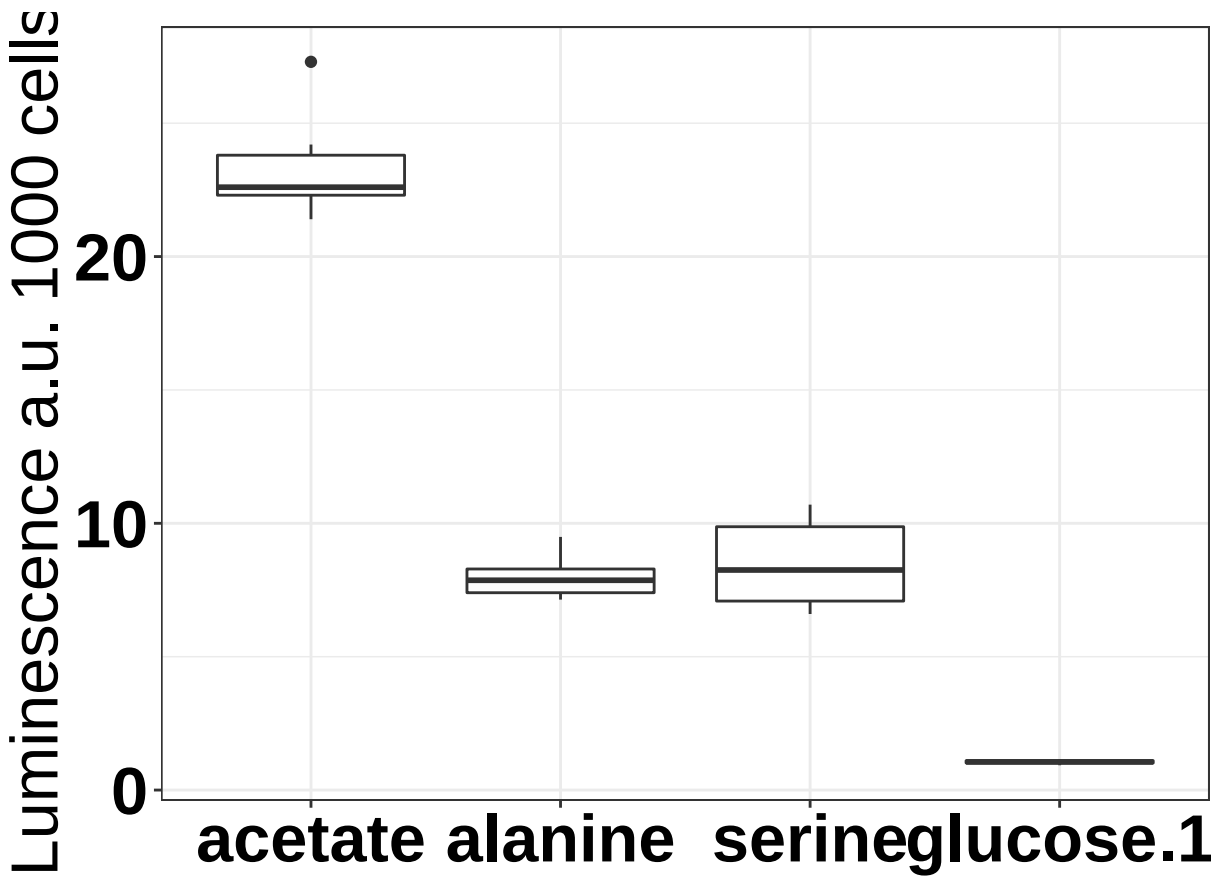
Luminescence is measured per 1000 cells.

```
library(ggplot2)
library(car)
library(knitr)

carb=read.table('PUBLICATION_carbonconc.txt', na.strings="NA", sep='\t', header=T, dec=',')
carb1=subset(carb, carbon!="glucose.6")
carb1 # inspect data
```

```
##       iron conc    carbon  counts lumi
## 1  replete  100 glucose.1 2174.42 1.11
## 2  replete  100 glucose.1 2190.52 1.01
## 3  replete  100 glucose.1 1802.27 1.15
## 4  replete  100 glucose.1 2331.95 0.93
## 5  replete  100 glucose.1 2110.26 1.08
## 6  replete  100 glucose.1 2180.65 1.03
## 13 replete  100   acetate 2567.66 24.2
## 14 replete  100   acetate 2634.09 22.6
## 15 replete  100   acetate 2154.61 27.3
## 16 replete  100   acetate 2574.61 22.2
## 17 replete  100   acetate 2615.06 21.4
## 18 replete  100   acetate 2521.43 22.6
## 19 replete  100   alanine 1827.21 9.49
## 20 replete  100   alanine 1742.34 8.34
## 21 replete  100   alanine 1586.88 8.13
## 22 replete  100   alanine 1755.45 7.14
## 23 replete  100   alanine 1625.58 7.33
## 24 replete  100   alanine 1621.62  7.6
## 25 replete  100    serine 2638.18 10.7
## 26 replete  100    serine 2293.12 10.2
## 27 replete  100    serine 1833.44 7.61
```

```
## 28 replete  100    serine 1966.49  6.6
## 29 replete  100    serine 2037.86 6.91
## 30 replete  100    serine 1828.77 8.89
```

```
carb1$conc=as.character(carb1$conc)
carb1$lumi=as.numeric(as.character(carb1$lumi))

carb1$carbon=factor(carb1$carbon,levels = c("acetate", "alanine", "serine", "glucose.1"))
c=ggplot(carb1, aes(x=carbon, y=lumi, fill=conc))+geom_boxplot()
c=c+theme_bw()
c=c+scale_y_continuous(expression(Luminescence~a.u.~"1000"~cells^-1))
c=c+scale_fill_manual(values=c("white"))
c=c+ guides(fill=FALSE)
c=c+ theme(axis.title.y = element_text(color="black",size=25, face="bold"), axis.text = element_text(co
c
```




```
ggsave("c.png", width = 10, height = 10)

##anova - untransformed data
cc=lm(lumi ~ carbon,
        data = carb1)
ccc <- Anova(cc, type = 3)
kable(ccc, digits = 10,results = 'asis')
```

	Sum Sq	Df	F value	Pr(>F)
(Intercept)	3280.68167	1	1596.0071	0
carbon	1591.48575	3	258.0787	0

	Sum Sq	Df	F value	Pr(>F)
Residuals	41.11112	20	NA	NA

```
shapiro.test(residuals(cc))  #ok
```

```
##
##  Shapiro-Wilk normality test
##
## data:  residuals(cc)
## W = 0.92028, p-value = 0.05926
```

```
leveneTest(cc) #ok
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##       Df F value  Pr(>F)
## group  3  2.4944 0.08934 .
##       20
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
cccc <- aov(lumi ~ carbon, data = carb1)
TukeyHSD(cccc)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = lumi ~ carbon, data = carb1)
##
## $carbon
##                          diff        lwr        upr     p adj
## alanine-acetate   -15.378333 -17.695178 -13.061488 0.0000000
## serine-acetate    -14.898333 -17.215178 -12.581488 0.0000000
## glucose.1-acetate -22.331667 -24.648512 -20.014822 0.0000000
## serine-alanine      0.480000  -1.836845   2.796845 0.9369716
## glucose.1-alanine  -6.953333  -9.270178  -4.636488 0.0000003
## glucose.1-serine   -7.433333  -9.750178  -5.116488 0.0000001
```

Summary for different carbon sources

Normality ($p = 0.05926$) and homoskedasticity ($p = 0.08934$) were met. Significant differences ($p<0.0005$) in the luminescence between different carbon sources. A tukey test reveals that there was no significant difference between alanine and serine ($p = 0.9369716$).

1b. Different Glucose Concentrations

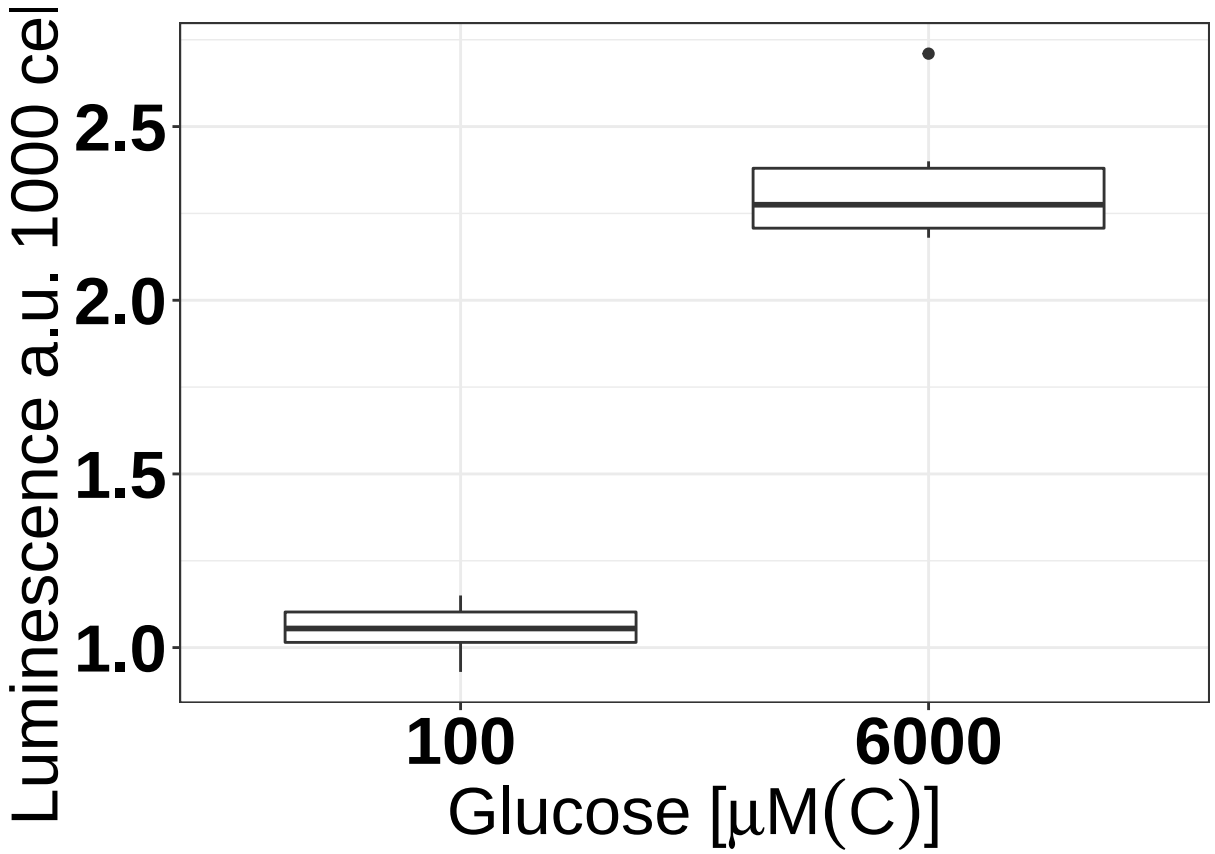
Luminescence is measured per 1000 cells.

```
glu=read.table('glucose.txt', na.strings="NA", sep='\t', header=T, dec=',')
glu=subset(glu, carbon=="glucose")
glu # inspect data
```

```
##       iron conc  carbon  counts lumi
## 1  replete  100 glucose 2174.42 1.11
```

```
## 2  replete  100 glucose 2190.52 1.01
## 3  replete  100 glucose 1802.27 1.15
## 4  replete  100 glucose 2331.95 0.93
## 5  replete  100 glucose 2110.26 1.08
## 6  replete  100 glucose 2180.65 1.03
## 7  replete 6000 glucose 2605.52 2.23
## 8  replete 6000 glucose 2760.13  2.2
## 9  replete 6000 glucose 2445.97  2.4
## 10 replete 6000 glucose 2694.87 2.18
## 11 replete 6000 glucose 2616.04 2.32
## 12 replete 6000 glucose 2088.18 2.71
```

```
glu$conc=as.character(glu$conc)
glu$lumi=as.numeric(as.character(glu$lumi))

gluc=ggplot(glu, aes(x=conc, y=lumi, fill=carbon))+geom_boxplot()
gluc=gluc+theme_bw()
gluc=gluc+scale_x_discrete(expression("Glucose"~"["*mu*M*(C)*"]"))+scale_y_continuous(expression(Lumines
gluc=gluc+scale_fill_manual(values=c("white"))
gluc=gluc+ guides(fill=FALSE)
gluc=gluc+ theme(axis.title = element_text(color="black",size=25, face="bold"), axis.text = element_tex
gluc
```




```
ggsave("gluc.png", width = 5, height = 10)

##anova - untransformed data
gluco=lm(lumi ~ conc,
        data = glu)
```

```
glucose <- Anova(gluco, type = 3)
kable(glucose, digits = 10,results = 'asis')
```

	Sum Sq	Df	F value	Pr(>F)
(Intercept)	6.6360167	1	289.6770	1.03e-08
conc	4.9794083	1	217.3623	4.13e-08
Residuals	0.2290833	10	NA	NA

```
shapiro.test(residuals(gluco))  #ok
```

```
##
##  Shapiro-Wilk normality test
##
## data:  residuals(gluco)
## W = 0.86549, p-value = 0.0573
```

```
leveneTest(gluco) #ok
```

```
## Warning in leveneTest.default(y = y, group = group, ...): group coerced to
## factor.
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##       Df F value Pr(>F)
## group  1  1.4108 0.2624
##       10
```

Summary for different glucose concentrations

Normality (p = 0.0573) and homoskedasticity (p = 0.2624) were met. Significant differences (p<0.0005) in the luminescence between different glucose concentratrions.

1c. Different levels of Fe-limitation

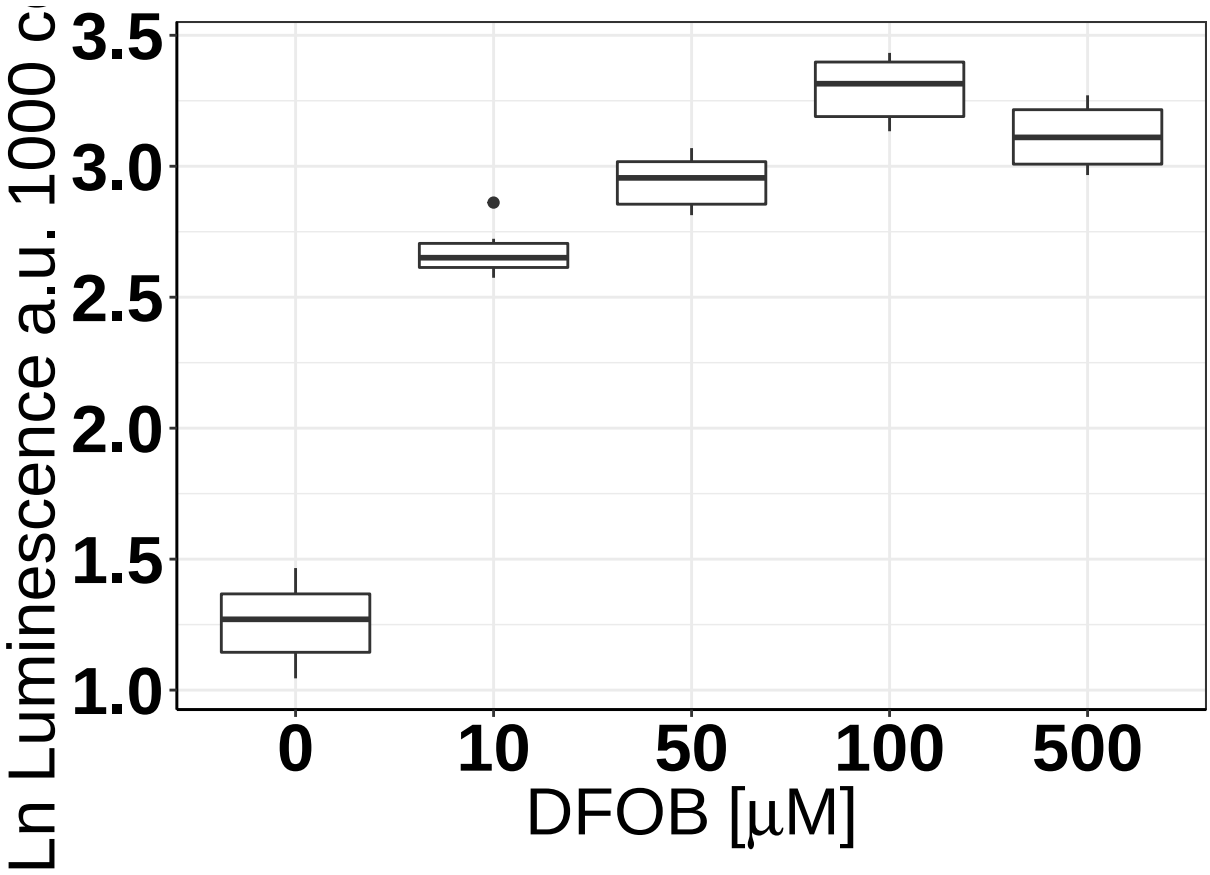
Luminescence is measured per 1000 cells. Data is log transformed in order to account for a skewed dataset. This is likely because of an increased variance under extreme Fe-limitation using much lower cell counts as opposed to Fe-repete conditions.

```
exp=read.table('icl_expression.txt', na.strings="NA", sep='\t', header=T, dec=',')
exp1=subset(exp, iron=="replete")
exp1 # inspect data
```

```
##       iron dfob luminescence
## 30 replete    0  2.842889088
## 31 replete    0  4.041059807
## 32 replete    0  4.331404024
## 33 replete   10  13.48810378
## 34 replete   10  14.13670492
## 35 replete   10  15.22671627
## 36 replete   50  16.87912485
## 37 replete   50  19.47637717
## 38 replete   50  18.95493243
## 39 replete  100  30.96405413
```

```
## 40 replete  100  30.29318557
## 41 replete  100  23.62707085
## 42 replete  500  23.55657247
## 43 replete  500  19.89361131
## 44 replete  500  21.34877925
## 45 replete    0  3.020356847
## 46 replete    0  3.594393491
## 47 replete    0  3.528756866
## 48 replete   10  13.12134066
## 49 replete   10  17.48387984
## 50 replete   10  14.19237542
## 51 replete   50  20.77272413
## 52 replete   50  21.52469875
## 53 replete   50  16.66442481
## 54 replete  100  28.75253252
## 55 replete  100  26.34827204
## 56 replete  100  22.95847964
## 57 replete  500  25.39999756
## 58 replete  500  26.33389248
## 59 replete  500  19.42245774
```

```
exp1$dfob=as.factor(exp1$dfob)
exp1$luminescence=as.numeric(as.character(exp1$luminescence))
exp1$loglumi=log(exp1$luminescence)
exp1$loglumi=as.numeric(as.character(exp1$loglumi))

lum=ggplot(exp1, aes(x=factor(dfob), y=as.numeric(as.character(loglumi)), fill=dfob))+geom_boxplot(fill=
lum=lum+theme_bw()
lum=lum+scale_x_discrete(expression("DFOB"~"["*mu*M*"]"))+scale_y_continuous(expression(Ln~Luminescence=
#lum=lum+scale_fill_manual(values="grey")
lum=lum+theme(axis.line = element_line(colour = "black",
                                       size = 0.5, linetype = "solid"))
lum=lum+theme(axis.title = element_text(color="black",size=25, face="bold"), axis.text = element_text(co
lum
```




```
ggsave("lum.png", width = 10, height = 10)

# ANOVA one way
lumi=lm(loglumi ~ dfob,
            data = exp1)

ll <- Anova(lumi, type = 3)
kable(ll, digits = 10 ,results = 'asis')
```

	Sum Sq	Df	F value	Pr(>F)
(Intercept)	9.5077553	1	590.7227	0
dfob	15.9142706	4	247.1909	0
Residuals	0.4023781	25	NA	NA

```
shapiro.test(residuals(lumi))  #ok
```

```
##
##  Shapiro-Wilk normality test
##
## data:  residuals(lumi)
## W = 0.96025, p-value = 0.3143
```

```
leveneTest(lumi) #ok
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##       Df F value Pr(>F)
```

```
## group  4  0.5933 0.6707
##       25
```

```
# Summary of the analysis
lll <- aov(loglumi ~ dfob, data = exp1)
TukeyHSD(lll)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = loglumi ~ dfob, data = exp1)
##
## $dfob
##               diff         lwr        upr     p adj
## 10-0     1.4181572  1.20304179 1.63327252 0.0000000
## 50-0     1.6834245  1.46830908 1.89853982 0.0000000
## 100-0    2.0361739  1.82105852 2.25128926 0.0000000
## 500-0    1.8549864  1.63987105 2.07010179 0.0000000
## 50-10    0.2652673  0.05015193 0.48038267 0.0103848
## 100-10   0.6180167  0.40290137 0.83313211 0.0000001
## 500-10   0.4368293  0.22171389 0.65194463 0.0000291
## 100-50   0.3527494  0.13763407 0.56786481 0.0005298
## 500-50   0.1715620 -0.04355340 0.38667733 0.1650511
## 500-100 -0.1811875 -0.39630284 0.03392789 0.1290819
```

Summary for different Fe-limitation.

Data was normally distributed ($p = 0.3143$) and met requirements for homoskedasticity ($p = 0.6707$). Data is highly significant between the luminescence and different levels of iron limitation ($p > 0.000001$). A Tukey Test reveals that these differences mainly lie between Fe-replete conditions and strong iron limitation and that at strong iron limitation luminescence does not differ significantly amongst each other. At 500 umol DFOB the luminescence seems to decrease, but this is likely due the fact that growth rate is so severely impacted by Fe-limitation that the results of luminescence data become unreliable.

2. ICL knockout and WT

Comparative experiments for both growth and respiration rates between an ICL knockout and WT were conducted when subjected to different levels of iron limitation (0, 10 and 100 DFOB)

```
data1=read.table('icl_alldates1.txt', na.strings="NA", sep='\t', header=T, dec=',')
data1 # inspect data
```

```
##        date    iron iron_v strain conditions substrate respiration growth
## 1  20141015      Fe      0     WT      wt_fe   glucose        3.71  10.28
## 2  20141013      Fe      0     WT      wt_fe   glucose        3.26   9.85
## 3  20141013      Fe      0     WT      wt_fe   glucose        2.92  10.45
## 4  20141013      Fe      0     WT      wt_fe   glucose        3.17  10.97
## 5  20141013      Fe      0     WT      wt_fe   glucose        2.89  10.45
## 6  20141020      Fe      0     KO      ko_fe   glucose        4.79  10.37
## 7  20141022      Fe      0     KO      ko_fe   glucose        4.62  10.11
## 8  20141022      Fe      0     KO      ko_fe   glucose        4.51   9.94
## 9  20141022      Fe      0     KO      ko_fe   glucose        4.84  10.54
## 10 20141105      Fe      0     KO      ko_fe   glucose        3.67   7.69
```

```
## 11 20141105      Fe      0     KO      ko_fe   glucose        4.12   8.47
## 12 20141105      Fe      0     KO      ko_fe   glucose        4.05   8.99
## 13 20151011  10DFOB     10     WT  wt_10dfob   glucose        3.08   6.19
## 14 20151011  10DFOB     10     WT  wt_10dfob   glucose        1.94   5.99
## 15 20151011  10DFOB     10     WT  wt_10dfob   glucose        1.86    6.4
## 16 20151011  10DFOB     10     KO  ko_10dfob   glucose         1.4   4.04
## 17 20151011  10DFOB     10     KO  ko_10dfob   glucose        2.17   6.73
## 18 20151011  10DFOB     10     KO  ko_10dfob   glucose        2.31   7.68
## 19 20141020 100DFOB    100     KO ko_100dfob   glucose        0.78   1.21
## 20 20141020 100DFOB    100     KO ko_100dfob   glucose        0.97   1.93
## 21 20141020 100DFOB    100     KO ko_100dfob   glucose         0.9   1.17
## 22 20141025 100DFOB    100     WT wt_100dfob   glucose        1.41   3.02
## 23 20141025 100DFOB    100     WT wt_100dfob   glucose        1.63   3.44
## 24 20141025 100DFOB    100     WT wt_100dfob   glucose        1.66   3.08
## 25 20141119      Fe      0     WT      wt_fe   glucose        3.88   9.85
## 26 20141119      Fe      0     WT      wt_fe   glucose        6.62   6.46
## 27 20141119      Fe      0     WT      wt_fe   glucose        5.25    8.9
```

2a. Growth Rate

Data was log transformed in order to account for a skewed dataset. Unequal sample sizes (different sizes for each condition) were accounted for by accounting for it within the parameters (contrasts) of an ANOVA after assumptions have been met.

```
data1=read.table('icl_alldates1.txt', na.strings="NA", sep='\t', header=T, dec=',')

data1$strain=as.factor(data1$strain)
data1$iron=as.factor(data1$iron)
data1$growth=as.numeric(as.character(data1$growth))
data1$respiration=as.numeric(as.character(data1$respiration))

data1$loggrowth=log(data1$growth)
data1 # inspect data
```

```
##        date    iron iron_v strain conditions substrate respiration growth
## 1  20141015      Fe      0     WT      wt_fe   glucose        3.71  10.28
## 2  20141013      Fe      0     WT      wt_fe   glucose        3.26   9.85
## 3  20141013      Fe      0     WT      wt_fe   glucose        2.92  10.45
## 4  20141013      Fe      0     WT      wt_fe   glucose        3.17  10.97
## 5  20141013      Fe      0     WT      wt_fe   glucose        2.89  10.45
## 6  20141020      Fe      0     KO      ko_fe   glucose        4.79  10.37
## 7  20141022      Fe      0     KO      ko_fe   glucose        4.62  10.11
## 8  20141022      Fe      0     KO      ko_fe   glucose        4.51   9.94
## 9  20141022      Fe      0     KO      ko_fe   glucose        4.84  10.54
## 10 20141105      Fe      0     KO      ko_fe   glucose        3.67   7.69
## 11 20141105      Fe      0     KO      ko_fe   glucose        4.12   8.47
## 12 20141105      Fe      0     KO      ko_fe   glucose        4.05   8.99
## 13 20151011  10DFOB     10     WT  wt_10dfob   glucose        3.08   6.19
## 14 20151011  10DFOB     10     WT  wt_10dfob   glucose        1.94   5.99
## 15 20151011  10DFOB     10     WT  wt_10dfob   glucose        1.86   6.40
## 16 20151011  10DFOB     10     KO  ko_10dfob   glucose        1.40   4.04
## 17 20151011  10DFOB     10     KO  ko_10dfob   glucose        2.17   6.73
## 18 20151011  10DFOB     10     KO  ko_10dfob   glucose        2.31   7.68
## 19 20141020 100DFOB    100     KO ko_100dfob   glucose        0.78   1.21
```

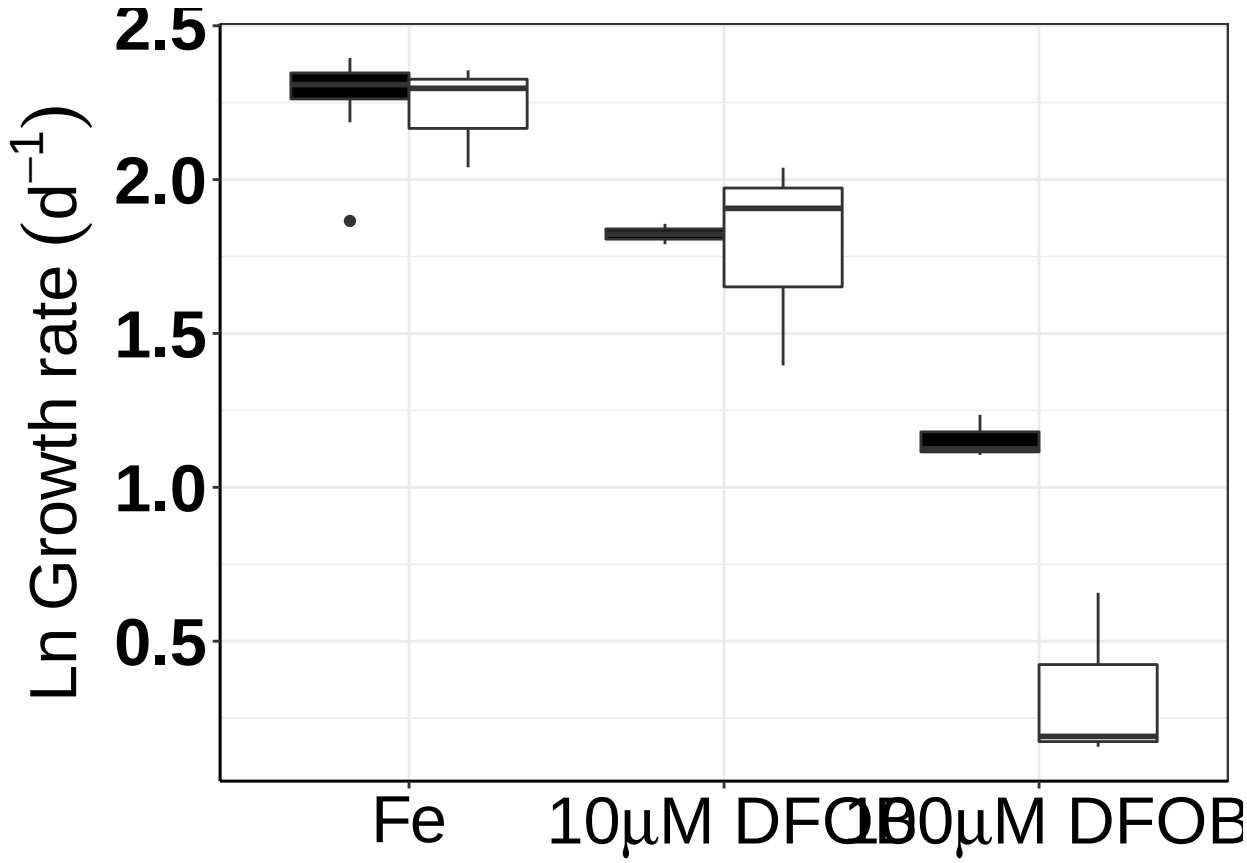
```
## 20 20141020 100DFOB    100     KO ko_100dfob   glucose        0.97   1.93
## 21 20141020 100DFOB    100     KO ko_100dfob   glucose        0.90   1.17
## 22 20141025 100DFOB    100     WT wt_100dfob   glucose        1.41   3.02
## 23 20141025 100DFOB    100     WT wt_100dfob   glucose        1.63   3.44
## 24 20141025 100DFOB    100     WT wt_100dfob   glucose        1.66   3.08
## 25 20141119      Fe      0     WT      wt_fe   glucose        3.88   9.85
## 26 20141119      Fe      0     WT      wt_fe   glucose        6.62   6.46
## 27 20141119      Fe      0     WT      wt_fe   glucose        5.25   8.90
##    loggrowth
## 1  2.3302003
## 2  2.2874715
## 3  2.3466020
## 4  2.3951643
## 5  2.3466020
## 6  2.3389170
## 7  2.3135250
## 8  2.2965670
## 9  2.3551775
## 10 2.0399208
## 11 2.1365305
## 12 2.1961128
## 13 1.8229351
## 14 1.7900914
## 15 1.8562980
## 16 1.3962447
## 17 1.9065751
## 18 2.0386195
## 19 0.1906204
## 20 0.6575200
## 21 0.1570037
## 22 1.1052568
## 23 1.2354715
## 24 1.1249296
## 25 2.2874715
## 26 1.8656293
## 27 2.1860513
```

```
### graph ###

levels(data1$strain)
```

```
## [1] "KO" "WT"
```

```
data1$strain <- factor(data1$strain, levels = rev(levels(data1$strain)))

graph=ggplot(data1, aes(x=iron, y=loggrowth, fill=strain))+geom_boxplot()
graph=graph+theme_bw()
graph=graph+scale_x_discrete(("DFOB"~"["*mu*M*"]"), limits=c("Fe", "10DFOB", "100DFOB"), labels=c("Fe",

graph=graph+scale_fill_manual(values=c("black","white"), name="Bacteria", labels=c("WT", "KO_ICL"))
graph=graph+theme(axis.line = element_line(colour = "black",
size = 0.5, linetype = "solid"), axis.title.y = element_text(color="black",size=25, face="bold"), axis.
axis.title.x = element_blank())
graph=graph+ guides(fill=FALSE) #WT = black, KO = white
```

```
graph
```




```
ggsave("growth.png", width = 10, height = 10)

##anova - untransformed data
orig=lm(growth ~ iron * strain,
            data = data1,
            contrasts = list(iron = "contr.sum", strain = "contr.poly"))
gg <- Anova(orig, type = 3)

##anova - log transformed data
model=lm(loggrowth ~ iron * strain,
            data = data1,
            contrasts = list(iron = "contr.sum", strain = "contr.poly"))

gg <- Anova(model, type = 3)
kable(gg, digits = 10,results = 'asis')
```

	Sum Sq	Df	F value	Pr(>F)
(Intercept)	57.4258289	1	1764.01211	0.0000000000
iron	9.6632427	2	148.41821	0.0000000000
strain	0.4824459	1	14.81982	0.0009301739
iron:strain	0.7402989	2	11.37029	0.0004508332
Residuals	0.6836361	21	NA	NA

```r
#check normality
shapiro.test(residuals(model))  #ok
```

```
##
##  Shapiro-Wilk normality test
##
## data:  residuals(model)
## W = 0.93822, p-value = 0.1101
```

```r
leveneTest(model) #ok
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##       Df F value Pr(>F)
## group  5  0.7097 0.6228
##       21
```

```r
rrr <- aov(loggrowth ~ iron*strain, data = data1)
TukeyHSD(rrr)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = loggrowth ~ iron * strain, data = data1)
##
## $iron
##                       diff       lwr       upr     p adj
## 10DFOB-100DFOB 1.0566603 0.7940928 1.3192278 0.0000000
## Fe-100DFOB     1.5029958 1.2833161 1.7226756 0.0000000
## Fe-10DFOB      0.4463355 0.2266558 0.6660153 0.0001285
##
## $strain
##             diff        lwr         upr     p adj
## KO-WT -0.2009208 -0.3454421 -0.05639959 0.0087373
##
## $`iron:strain`
##                                 diff         lwr         upr     p adj
## 10DFOB:WT-100DFOB:WT     0.66788886  0.20700445  1.12877328 0.0021867
## Fe:WT-100DFOB:WT         1.10042970  0.71828453  1.48257487 0.0000002
## 100DFOB:KO-100DFOB:WT   -0.82017126 -1.28105568 -0.35928685 0.0002042
## 10DFOB:KO-100DFOB:WT     0.62526049  0.16437608  1.08614491 0.0042514
## Fe:KO-100DFOB:WT         1.08431652  0.69479810  1.47383495 0.0000003
## Fe:WT-10DFOB:WT          0.43254084  0.05039567  0.81468601 0.0206944
## 100DFOB:KO-10DFOB:WT    -1.48806013 -1.94894454 -1.02717571 0.0000000
## 10DFOB:KO-10DFOB:WT     -0.04262837 -0.50351278  0.41825605 0.9996826
## Fe:KO-10DFOB:WT          0.41642766  0.02690923  0.80594608 0.0316879
## 100DFOB:KO-Fe:WT        -1.92060096 -2.30274613 -1.53845579 0.0000000
## 10DFOB:KO-Fe:WT         -0.47516921 -0.85731437 -0.09302404 0.0095158
## Fe:KO-Fe:WT             -0.01611318 -0.30825200  0.27602564 0.9999752
## 10DFOB:KO-100DFOB:KO     1.44543176  0.98454734  1.90631617 0.0000000
## Fe:KO-100DFOB:KO         1.90448779  1.51496936  2.29400621 0.0000000
## Fe:KO-10DFOB:KO          0.45905603  0.06953760  0.84857445 0.0149940
```

Summary Growth Rate Data Analysis

Data reveals a skewed data set where differences between the KO and WT are shown particularly in 100 DFOB conditions, while differences do not seem to be present for 10 DFOB and Fe conditions. In this particular case, it is appropriate to conduct a log transformation in order to discern a pattern. This is highlighted by the fact that a log transformation gives a normal distribution of the data ($p = 0.08762$) using a shapiro wilk test. Homoskedasticity was also met ($p = 0.5947$).

The first ANOVA table shows the data untransformed. While the second ANOVA table shows data after transformation. This is followed by tests to check whether data meet the assumptions of the ANOVA (**note that assumptions of normality were not met for untransformed data)

A tukey test reveals a significant differences betweeen WT and KO grown in strong iron limitation of 100uM DFOB ($p = 0.0000433$).

2b. Respiration Rate

Data was trimmed and log transformed in order to account for skewed data set. Unequal sample sizes (different sizes for each condition) were accounted for by accounting for it within the parameters (contrasts) of an ANOVA after assumptions have been met.

```
data2=read.table('icl_alldates.txt', na.strings="NA", sep='\t', header=T, dec=',')

data2$respiration=as.numeric(as.character(data2$respiration))
data2$logr=log(data2$respiration)
data2$logr=as.numeric(as.character(data2$logr))
data2 # new dataset with log.respiration
```

```
##        date    iron iron_v strain conditions substrate respiration growth
## 1  20141015      Fe      0     WT      wt_fe   glucose        3.71  10.28
## 2  20141013      Fe      0     WT      wt_fe   glucose        3.26   9.85
## 3  20141013      Fe      0     WT      wt_fe   glucose        2.92  10.45
## 4  20141013      Fe      0     WT      wt_fe   glucose        3.17  10.97
## 5  20141013      Fe      0     WT      wt_fe   glucose        2.89  10.45
## 6  20141020      Fe      0     KO      ko_fe   glucose        4.79  10.37
## 7  20141022      Fe      0     KO      ko_fe   glucose        4.62  10.11
## 8  20141022      Fe      0     KO      ko_fe   glucose        4.51   9.94
## 9  20141022      Fe      0     KO      ko_fe   glucose        4.84  10.54
## 10 20141105      Fe      0     KO      ko_fe   glucose        3.67   7.69
## 11 20141105      Fe      0     KO      ko_fe   glucose        4.12   8.47
## 12 20141105      Fe      0     KO      ko_fe   glucose        4.05   8.99
## 13 20151011  10DFOB     10     WT  wt_10dfob   glucose        3.08   6.19
## 14 20151011  10DFOB     10     WT  wt_10dfob   glucose        1.94   5.99
## 15 20151011  10DFOB     10     WT  wt_10dfob   glucose        1.86    6.4
## 16 20151011  10DFOB     10     KO  ko_10dfob   glucose        1.40   4.04
## 17 20151011  10DFOB     10     KO  ko_10dfob   glucose        2.17   6.73
## 18 20151011  10DFOB     10     KO  ko_10dfob   glucose        2.31   7.68
## 19 20141020 100DFOB    100     KO ko_100dfob   glucose        0.78   1.21
## 20 20141020 100DFOB    100     KO ko_100dfob   glucose        0.97   1.93
## 21 20141020 100DFOB    100     KO ko_100dfob   glucose        0.90   1.17
## 22 20141025 100DFOB    100     WT wt_100dfob   glucose        1.41   3.02
## 23 20141025 100DFOB    100     WT wt_100dfob   glucose        1.63   3.44
## 24 20141025 100DFOB    100     WT wt_100dfob   glucose        1.66   3.08
## 25 20141025 100DFOB    100     WT wt_100dfob   glucose        1.02   3.33
## 26 20141119      Fe      0     WT      wt_fe   glucose        3.88   9.85
```

```
## 27 20141119      Fe      0     WT      wt_fe   glucose        6.62   6.46
## 28 20141119      Fe      0     WT      wt_fe   glucose        5.25    8.9
##           logr
## 1   1.31103188
## 2   1.18172720
## 3   1.07158362
## 4   1.15373159
## 5   1.06125650
## 6   1.56653041
## 7   1.53039471
## 8   1.50629715
## 9   1.57691472
## 10  1.30019166
## 11  1.41585316
## 12  1.39871688
## 13  1.12492960
## 14  0.66268797
## 15  0.62057649
## 16  0.33647224
## 17  0.77472717
## 18  0.83724752
## 19 -0.24846136
## 20 -0.03045921
## 21 -0.10536052
## 22  0.34358970
## 23  0.48858001
## 24  0.50681760
## 25  0.01980263
## 26  1.35583515
## 27  1.89009537
## 28  1.65822808
```

```
wtiron=subset(data2, conditions=="wt_fe") #selects from column only what includes ... (one factor)
koiron=subset(data2, conditions=="ko_fe")
wtmed=subset(data2, conditions=="wt_10dfob")
komed=subset(data2, conditions=="ko_10dfob")
wtext=subset(data2, conditions=="wt_100dfob")
koext=subset(data2, conditions=="ko_100dfob")

#comparing mean and median for logr data (skewed dataset?)
mean(data2$logr) #0.94
```

```
## [1] 0.9396264
```

```
median(data2$logr) #1.10
```

```
## [1] 1.098257
```

```
#compare subsets to identify potential outliers
mean(koiron$logr) #1.47
```

```
## [1] 1.4707
```

```
median(koiron$logr) #1.51
```

```
## [1] 1.506297
```

```
mean(wtiron$logr) #1.33
```

```
## [1] 1.335436
```

```
median(wtiron$logr) #1.25
```

```
## [1] 1.24638
```

```
mean(wtmed$logr) #0.80
```

```
## [1] 0.8027314
```

```
median(wtmed$logr) #0.66
```

```
## [1] 0.662688
```

```
mean(komed$logr) #0.65
```

```
## [1] 0.6494823
```

```
median(komed$logr) #0.77
```

```
## [1] 0.7747272
```

```
mean(koext$logr) #-0.13
```

```
## [1] -0.1280937
```

```
median(koext$logr) #-0.10
```

```
## [1] -0.1053605
```

```
mean(wtext$logr) #0.34
```

```
## [1] 0.3396975
```

```
median(wtext$logr) #0.42
```

```
## [1] 0.4160849
```

```
 #compare subsets
mean(koiron$logr, trim=0.25) #1.51 (remove outlier)
```

```
## [1] 1.483558
```

```
mean(wtiron$logr, trim=0.25) #1.25 (remove outlier)
```

```
## [1] 1.250581
```

```
mean(wtext$logr, trim=0.25) #0.42 (remove outlier)
```

```
## [1] 0.4160849
```

```
trim=read.table('icl_alldates_outliers.txt', na.strings="NA", sep='\t', header=T, dec=',')

trim$strain
```

```
##  [1] WT WT WT WT WT KO KO KO KO KO KO WT WT WT KO KO KO KO KO KO WT WT WT
## [24] WT WT
## Levels: KO WT
```

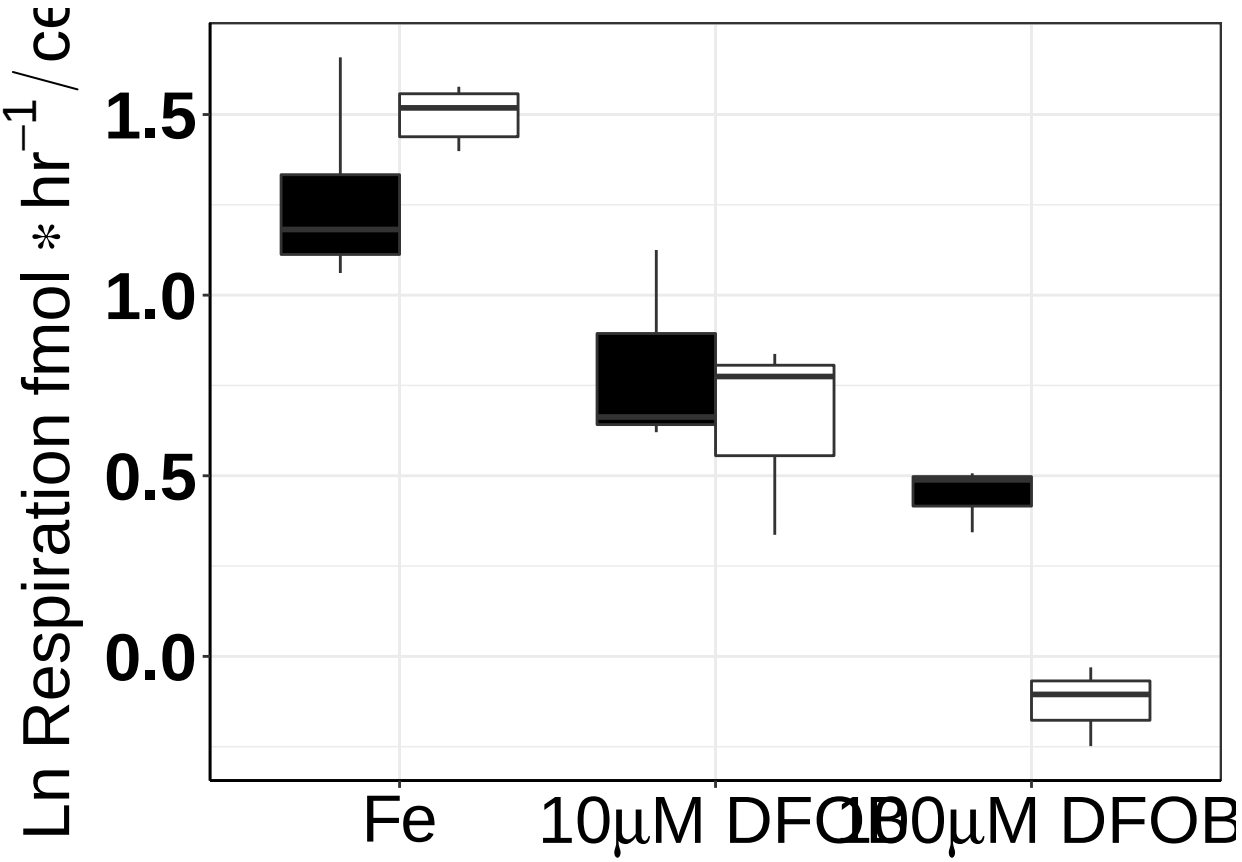
```
trim$strain=factor(trim$strain, levels = c('WT','KO'), ordered = TRUE)
trim$respiration=as.numeric(as.character(trim$respiration))
```

```
trim$logresp=log(trim$respiration)
trim# inspect data
```

```
##    X20141015    iron iron_v strain conditions substrate respiration growth
## 1   20141013      Fe      0     WT      wt_fe   glucose        3.71  10.28
## 2   20141013      Fe      0     WT      wt_fe   glucose        3.26   9.85
## 3   20141013      Fe      0     WT      wt_fe   glucose        2.92  10.45
## 4   20141013      Fe      0     WT      wt_fe   glucose        3.17  10.97
## 5   20141020      Fe      0     WT      wt_fe   glucose        2.89  10.45
## 6   20141022      Fe      0     KO      ko_fe   glucose        4.79  10.37
## 7   20141022      Fe      0     KO      ko_fe   glucose        4.62  10.11
## 8   20141022      Fe      0     KO      ko_fe   glucose        4.51   9.94
## 9   20141105      Fe      0     KO      ko_fe   glucose        4.84  10.54
## 10  20141105      Fe      0     KO      ko_fe   glucose        4.12   8.47
## 11  20151011      Fe      0     KO      ko_fe   glucose        4.05   8.99
## 12  20151011  10DFOB     10     WT  wt_10dfob   glucose        3.08   6.19
## 13  20151011  10DFOB     10     WT  wt_10dfob   glucose        1.94   5.99
## 14  20151011  10DFOB     10     WT  wt_10dfob   glucose        1.86    6.4
## 15  20151011  10DFOB     10     KO  ko_10dfob   glucose        1.40   4.04
## 16  20151011  10DFOB     10     KO  ko_10dfob   glucose        2.17   6.73
## 17  20141020  10DFOB     10     KO  ko_10dfob   glucose        2.31   7.68
## 18  20141020 100DFOB    100     KO ko_100dfob   glucose        0.78   1.21
## 19  20141020 100DFOB    100     KO ko_100dfob   glucose        0.97   1.93
## 20  20141025 100DFOB    100     KO ko_100dfob   glucose        0.90   1.17
## 21  20141025 100DFOB    100     WT wt_100dfob   glucose        1.41   3.02
## 22  20141025 100DFOB    100     WT wt_100dfob   glucose        1.63   3.44
## 23  20141119 100DFOB    100     WT wt_100dfob   glucose        1.66   3.08
## 24  20141119      Fe      0     WT      wt_fe   glucose        3.88   9.85
## 25  20141119      Fe      0     WT      wt_fe   glucose        5.25    8.9
##        logresp
## 1   1.31103188
## 2   1.18172720
## 3   1.07158362
## 4   1.15373159
## 5   1.06125650
## 6   1.56653041
## 7   1.53039471
## 8   1.50629715
## 9   1.57691472
## 10  1.41585316
## 11  1.39871688
## 12  1.12492960
## 13  0.66268797
## 14  0.62057649
## 15  0.33647224
## 16  0.77472717
## 17  0.83724752
## 18 -0.24846136
## 19 -0.03045921
## 20 -0.10536052
## 21  0.34358970
## 22  0.48858001
## 23  0.50681760
## 24  1.35583515
```

```
## 25  1.65822808
```

```
graph2=ggplot(trim, aes(x=iron, y=logresp, fill=strain))+geom_boxplot()
graph2=graph2+scale_fill_manual(values=c("black","white"), name="Bacteria")
graph2=graph2+theme_bw()
graph2=graph2+scale_x_discrete(("DFOB"~"["*mu*M*"]"), limits=c("Fe", "10DFOB", "100DFOB"), labels=c("Fe
#"Ln Resp (fmol*hr -1)")
graph2=graph2+theme(axis.line = element_line(colour = "black", size = 0.5, linetype = "solid"),
                    axis.title.y = element_text(color="black",size=25, face="bold"), axis.text = elemen
                    axis.title.x = element_blank())
graph2=graph2+ guides(fill=FALSE) #WT = black, KO = white
graph2
```




```
ggsave("resp_outlier.png", width = 10, height = 10)

trim$logresp=as.numeric(as.character(trim$logresp))
r=lm(logresp ~ iron * strain,
     data = trim,
     contrasts = list(iron = "contr.sum", strain = "contr.poly"))

rr <- Anova(r, type = 3)
rr
```

```
## Anova Table (Type III tests)
##
## Response: logresp
##              Sum Sq Df  F value    Pr(>F)
## (Intercept) 12.4676  1 371.6482 6.206e-14 ***
```

```
## iron         6.4046  2  95.4573 1.227e-10 ***
## strain       0.1430  1   4.2637  0.052855 .
## iron:strain  0.7057  2  10.5184  0.000841 ***
## Residuals    0.6374 19
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#check normality
shapiro.test(residuals(r))  #ok
```

```
##
##  Shapiro-Wilk normality test
##
## data:  residuals(r)
## W = 0.96141, p-value = 0.4433
```

```
leveneTest(r) #ok
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##       Df F value Pr(>F)
## group  5  0.5322 0.7493
##       19
```

```
rrr <- aov(logresp ~ iron*strain, data = trim)
TukeyHSD(rrr)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = logresp ~ iron * strain, data = trim)
##
## $iron
##                      diff       lwr       upr     p adj
## 10DFOB-100DFOB 0.5669891 0.2983457 0.8356325 0.0001014
## Fe-100DFOB     1.2091978 0.9795478 1.4388477 0.0000000
## Fe-10DFOB      0.6422086 0.4125587 0.8718586 0.0000027
##
## $strain
##             diff        lwr      upr     p adj
## KO-WT -0.0491496 -0.2026142 0.104315 0.5107161
##
## $`iron:strain`
##                                diff         lwr        upr     p adj
## 10DFOB:WT-100DFOB:WT     0.3564022 -0.11612065  0.8289251 0.2111553
## Fe:WT-100DFOB:WT         0.8098700  0.41051530  1.2092248 0.0000493
## 100DFOB:KO-100DFOB:WT   -0.5744228 -1.04694569 -0.1018999 0.0120298
## 10DFOB:KO-100DFOB:WT     0.2031532 -0.26936969  0.6756761 0.7499119
## Fe:KO-100DFOB:WT         1.0527887  0.64357190  1.4620056 0.0000018
## Fe:WT-10DFOB:WT          0.4534678  0.05411306  0.8528225 0.0206399
## 100DFOB:KO-10DFOB:WT    -0.9308250 -1.40334794 -0.4583022 0.0000719
## 10DFOB:KO-10DFOB:WT     -0.1532490 -0.62577194  0.3192738 0.9038090
## Fe:KO-10DFOB:WT          0.6963865  0.28716966  1.1056033 0.0004284
## 100DFOB:KO-Fe:WT        -1.3842928 -1.78364757 -0.9849381 0.0000000
## 10DFOB:KO-Fe:WT         -0.6067168 -1.00607157 -0.2073621 0.0014928
## Fe:KO-Fe:WT              0.2429187 -0.07905138  0.5648888 0.2109102
## 10DFOB:KO-100DFOB:KO     0.7775760  0.30505311  1.2500989 0.0006277
```

```
## Fe:KO-100DFOB:KO       1.6272115  1.21799470  2.0364284 0.0000000
## Fe:KO-10DFOB:KO        0.8496355  0.44041870  1.2588524 0.0000362
```

Summary Respiration Rate Data Analysis

Data again was log transformed due to skewedness of data. Assumptions of the ANOVA were met through a shapiro test (p = 0.4433) on the log transformed residuals to test for normality while a levene test (p = 0.7493) was conducted to confirm homoskedastiicty.

Data is significant for interaction effects between strain and iron conditions and a tukey test reveals that these differences are statistically significant between WT and KO grown in 100DFOB conditions (p = 0.0120298). No significant differences were found between Fe-replete and intermediate levels of Fe-limitation.

```
library(cowplot)
```

```
##
## Attaching package: 'cowplot'
```

```
## The following object is masked from 'package:ggplot2':
##
##     ggsave
```

```
#cowplot --> to place graphs next to eachother
plot1=plot_grid(c, gluc, lum, graph, graph2, labels = c("", "", "", "", ""))
save_plot("graphs2.png", plot1,
          ncol = 3, # we're saving a grid plot of 2 columns
          nrow = 2, # and 1 rows
          # each individual subplot should have an aspect ratio of 1.0 (spacing)
          base_aspect_ratio = 1.0
)
```