

Analysis script for Berkhout et al 2014

bwb

11/05/2021

Read me

Analyses for Berkhout et al 2014 (doi: 10.1016/j.ijpara.2014.07.002). The analyses have been left unchanged from the original publication, but some minor updates have been included. These are: - three separate scripts are combined to a single .Rmd file - paths to data files are made relative - obsolete code is removed - some annotation is added - the order of some of the analyses is modified to match the order in the paper - statistics reported in the paper are marked ('reported in paper')

doi for data: 10.5281/zenodo.4749477 last modified: 17/05/2021 contact: papers[at]boriswberkhout.com

Loading packages

```
require("lme4", quietly = T)
library("data.table", quietly = T)
library("gplots", quietly = T)
```

```
##
## Attaching package: 'gplots'

## The following object is masked from 'package:stats':
##
##   lowess
```

Analysis of cercarial production

Data loading

```
# load raw data
data_output_raw = read.csv("data/Zea_output_v6.3.csv", header=TRUE)

# inspect data
#summary(data_output_raw)
str(data_output_raw)
```

```
## 'data.frame':   230 obs. of  13 variables:
## $ snail      : Factor w/ 30 levels "Y29","Y30","Y31",...: 1 1 1 1 1 1 1 1 2 2 ...
## $ temperature: int   15 15 15 15 20 20 20 20 15 15 ...
## $ week       : int    5 6 7 8 13 14 15 16 5 6 ...
## $ A          : int   330 398 113 276 580 387 264 889 779 369 ...
## $ B1         : int   139 95 15 104 47 64 38 69 77 31 ...
## $ B2         : int   140 60 39 103 49 71 45 153 99 52 ...
## $ C          : int    39 45 15 44 73 6 9 22 28 62 ...
## $ D          : int    0 1 0 2 0 0 17 1 0 2 ...
## $ E          : int    8 86 29 29 32 5 9 9 2 168 ...
## $ total      : int   656 685 211 558 781 533 382 1143 985 684 ...
## $ tank       : int    1 1 1 1 1 1 1 1 1 1 ...
## $ length     : num   13.9 13.9 13.9 13.9 13.9 13.9 13.9 13.9 13.9 13.9 ...
```

```
## $ infection : int  2 2 2 2 2 2 2 2 1 1 ...
# add up values per snail per temperature
data_output = data.table(data_output_raw)
data_output$temperature = ifelse((data_output$week<10), 15,20)

# make $tank and $week factors
data_output$tank = as.factor(data_output$tank)
data_output$week = as.factor(data_output$week)

data_output$infection = as.factor(data_output$infection)

#select single infection snails
data_output_sing = subset(data_output,infection == '1')
data_output_sing = droplevels(data_output_sing)

#select multiple infection snails
data_output_multi = subset(data_output,infection == '2')
data_output_multi = droplevels(data_output_multi)
```

Models on cercarial output Reformat data for analysis. Then, construct the different models and select most appropriate models on the analysis of the total number of cercariae produced from each host snail.

```
# prepare data for analysis
data_output_glm = data_output_sing
data_output_glm_multi = data_output_multi
data_output_glm_all = data_output

# test for differences between different sized snails
glm_length = glm(total~length, data=data_output_glm, family="quasipoisson")
glm_length_multi = glm(total~length, data=data_output_glm_multi, family="quasipoisson")
glm_length_all = glm(total~length, data=data_output_glm_all, family="quasipoisson")
summary(glm_length)
```

```
##
## Call:
## glm(formula = total ~ length, family = "quasipoisson", data = data_output_glm)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -34.668  -15.702   -2.733    8.107   55.959
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  7.65761    1.57715   4.855 3.66e-06 ***
## length      -0.09603    0.12096  -0.794   0.429
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasipoisson family taken to be 356.5715)
##
##      Null deviance: 40081  on 121  degrees of freedom
## Residual deviance: 39857  on 120  degrees of freedom
## AIC: NA
##
## Number of Fisher Scoring iterations: 5
```

```
summary(glm_length_multi)
```

```
##
## Call:
## glm(formula = total ~ length, family = "quasipoisson", data = data_output_glm_multi)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -34.962  -16.604   -8.373    9.431   67.605
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.9850     2.1592   1.382  0.1697
## length        0.2765     0.1655   1.671  0.0977 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasipoisson family taken to be 485.0491)
##
##      Null deviance: 45765  on 107  degrees of freedom
## Residual deviance: 44420  on 106  degrees of freedom
## AIC: NA
##
## Number of Fisher Scoring iterations: 5
```

```
summary(glm_length_all)
```

```
##
## Call:
## glm(formula = total ~ length, family = "quasipoisson", data = data_output_glm_all)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -35.372  -16.015   -6.062    7.997   75.308
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   5.89682     1.32917   4.436 1.42e-05 ***
## length        0.04568     0.10192   0.448  0.654
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasipoisson family taken to be 443.9767)
##
##      Null deviance: 87023  on 229  degrees of freedom
## Residual deviance: 86934  on 228  degrees of freedom
## AIC: NA
##
## Number of Fisher Scoring iterations: 5
```

```
# test for differences between single and multiple infections
```

```
glm_sing_multi_15 = glm(total~infection, data=subset(data_output_glm_all, temperature=='15'), family="quasipoisson")
glm_sing_multi_20 = glm(total~infection, data=subset(data_output_glm_all, temperature=='20'), family="quasipoisson")
summary(glm_sing_multi_15)
```

```
##
## Call:
## glm(formula = total ~ infection, family = "quasipoisson", data = subset(data_output_glm_all,
##   temperature == "15"))
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -27.362  -11.621   -3.630    8.657   30.456
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  6.17801    0.07631  80.961  <2e-16 ***
## infection2    0.19899    0.10619   1.874   0.0634 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasipoisson family taken to be 179.6389)
##
##      Null deviance: 21751  on 119  degrees of freedom
## Residual deviance: 21120  on 118  degrees of freedom
## AIC: NA
##
## Number of Fisher Scoring iterations: 5
summary(glm_sing_multi_20)
```

```
##
## Call:
## glm(formula = total ~ infection, family = "quasipoisson", data = subset(data_output_glm_all,
##   temperature == "20"))
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -40.261  -20.685   -5.240    9.482   65.873
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  6.6075    0.1159  56.996  <2e-16 ***
## infection2    0.1553    0.1621   0.958   0.34
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasipoisson family taken to be 577.3026)
##
##      Null deviance: 59022  on 109  degrees of freedom
## Residual deviance: 58492  on 108  degrees of freedom
## AIC: NA
##
## Number of Fisher Scoring iterations: 5
glm_sing_multi = glm(total~infection, data=data_output_glm_all, family="quasipoisson")
summary(glm_sing_multi) # reported in paper
```

```
##
## Call:
```

```

## glm(formula = total ~ infection, family = "quasipoisson", data = data_output_glm_all)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -36.570  -16.238   -5.395    8.500   72.798
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  6.40517    0.07662  83.601  <2e-16 ***
## infection2    0.17613    0.10691   1.647    0.101
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasipoisson family taken to be 433.2384)
##
##      Null deviance: 87023  on 229  degrees of freedom
## Residual deviance: 85847  on 228  degrees of freedom
## AIC: NA
##
## Number of Fisher Scoring iterations: 5
# test for differences between different temperatures
glm_temp = glm(total~temperature, data=data_output_glm, family="quasipoisson")
glm_temp_multi = glm(total~temperature, data=data_output_glm_multi, family="quasipoisson")
glm_temp_all = glm(total~temperature, data=data_output_glm_all, family="quasipoisson")
summary(glm_temp) # reported in paper

##
## Call:
## glm(formula = total ~ temperature, family = "quasipoisson", data = data_output_glm)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -37.835  -13.568   -3.537    9.204   51.785
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.88957    0.46737  10.462  < 2e-16 ***
## temperature  0.08590    0.02585   3.323  0.00118 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasipoisson family taken to be 299.9874)
##
##      Null deviance: 40081  on 121  degrees of freedom
## Residual deviance: 36716  on 120  degrees of freedom
## AIC: NA
##
## Number of Fisher Scoring iterations: 5
summary(glm_temp_multi)

##
## Call:
## glm(formula = total ~ temperature, family = "quasipoisson", data = data_output_glm_multi)

```

```

##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -40.261  -17.061   -4.822    8.749   65.873
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  5.21961    0.55466   9.410 1.19e-15 ***
## temperature  0.07716    0.03072   2.512  0.0135 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasipoisson family taken to be 448.5621)
##
##      Null deviance: 45765  on 107  degrees of freedom
## Residual deviance: 42896  on 106  degrees of freedom
## AIC: NA
##
## Number of Fisher Scoring iterations: 5
summary(glm_temp_all)

##
## Call:
## glm(formula = total ~ temperature, family = "quasipoisson", data = data_output_glm_all)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -39.351  -15.386   -4.155    8.306   68.922
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  5.05150    0.36516  13.833 < 2e-16 ***
## temperature  0.08162    0.02021   4.038 7.35e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasipoisson family taken to be 377.5566)
##
##      Null deviance: 87023  on 229  degrees of freedom
## Residual deviance: 80773  on 228  degrees of freedom
## AIC: NA
##
## Number of Fisher Scoring iterations: 5
# generate parameter for random slopes
parameter<-as.factor(seq(1:length(data_output_glm$temperature)))

# test for normality total output
shapiro.test(data_output_glm$total) #not normaly distributed

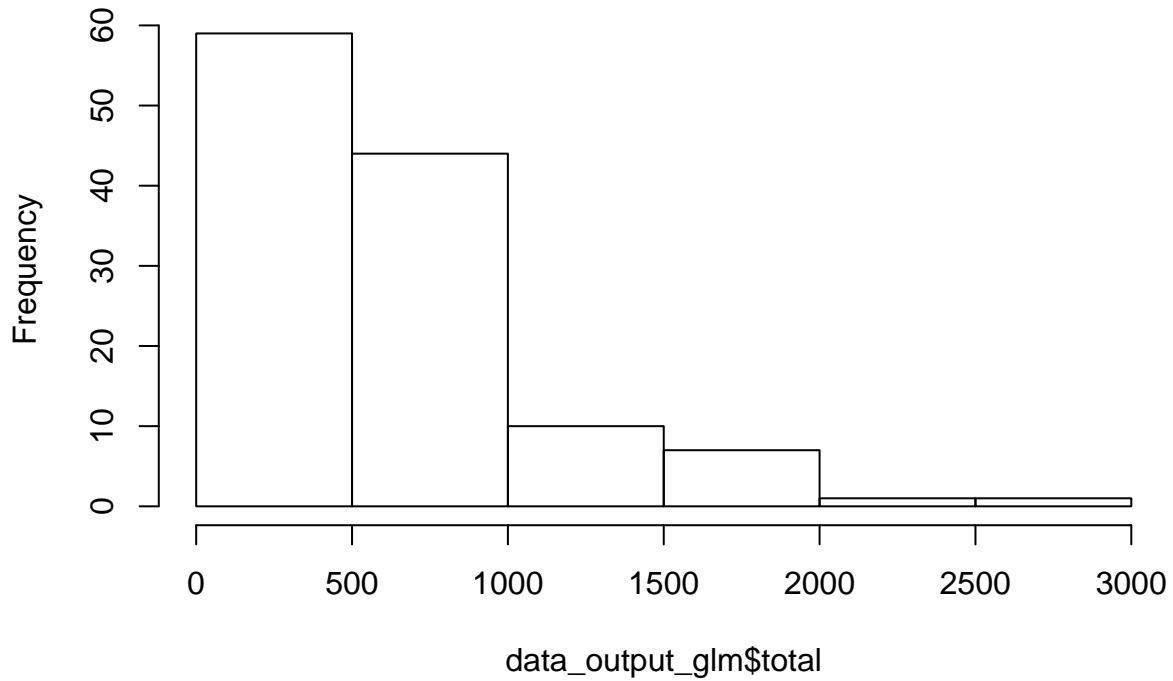
##
## Shapiro-Wilk normality test
##
## data:  data_output_glm$total

```

```
## W = 0.88188, p-value = 2.104e-08
```

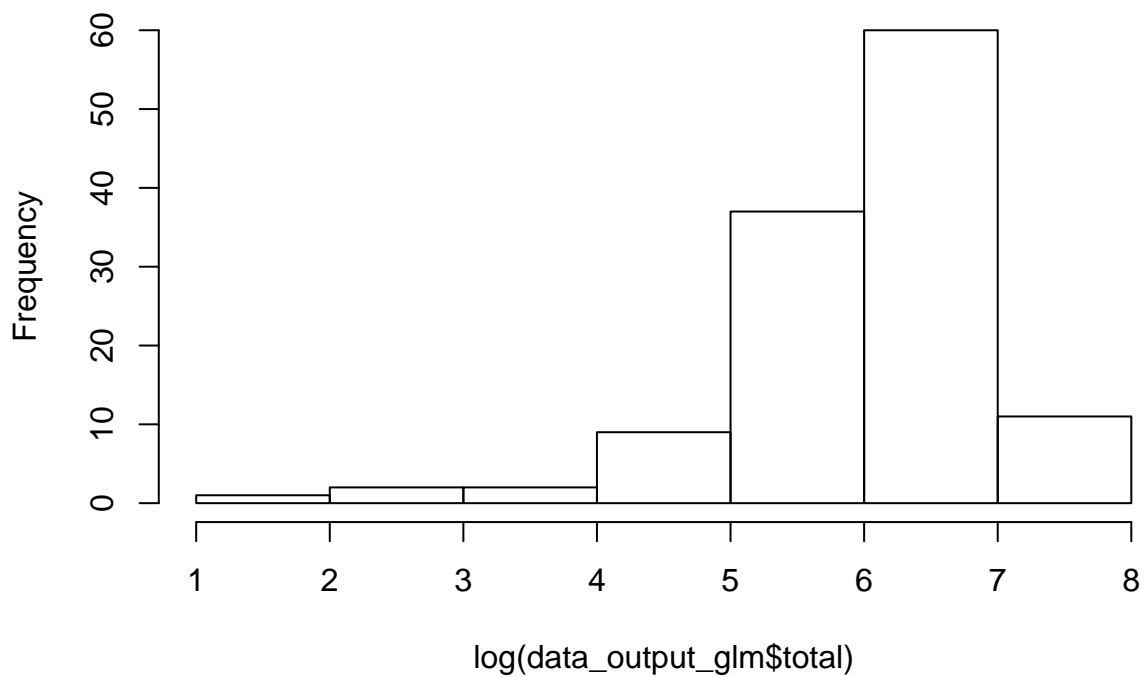
```
hist(data_output_glm$total) #right skewed, p << 0.05
```

Histogram of data_output_glm\$total



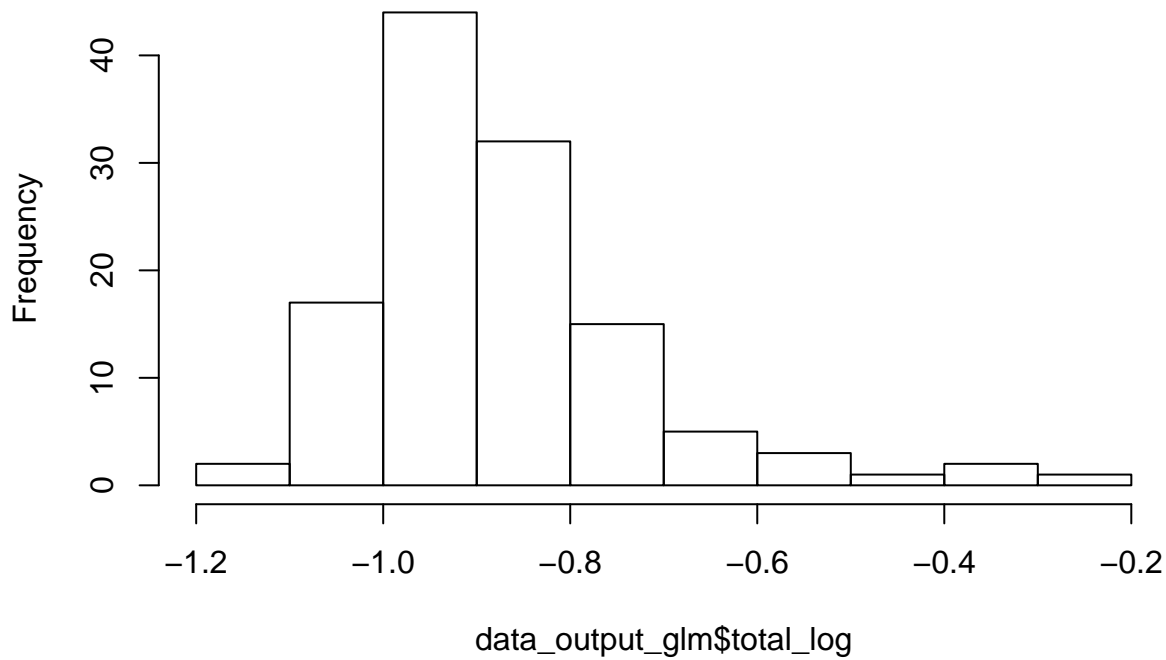
```
hist(log(data_output_glm$total)) #left skewed, p << 0.05
```

Histogram of log(data_output_glm\$total)



```
data_output_glm$total_log=log(data_output_glm$total,.001)
hist(data_output_glm$total_log)
```

Histogram of data_output_glm\$total_log



```
shapiro.test(data_output_glm$total_log)
```

```
##
## Shapiro-Wilk normality test
##
## data: data_output_glm$total_log
## W = 0.88873, p-value = 4.448e-08
# run analysis to compare cercarial production between temperatures with all fixed and random effects i
glm_full = glmer(total~temperature + (temperature|snail) + (1|week) + (1|parameter), data=data_output_g
glm_full2 = glmer(total~temperature + (1|snail) + (1|week) + (1|parameter), data=data_output_glm, famil
anova(glm_full,glm_full2) # reported in paper
```

```
## Data: data_output_glm
## Models:
## glm_full2: total ~ temperature + (1 | snail) + (1 | week) + (1 | parameter)
## glm_full: total ~ temperature + (temperature | snail) + (1 | week) + (1 |
## glm_full: parameter)
##      npar      AIC      BIC logLik deviance Chisq Df Pr(>Chisq)
## glm_full2    5 1773.8 1787.8 -881.88   1763.8
## glm_full     7 1750.6 1770.2 -868.29   1736.6 27.192  2 1.245e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
summary(glm_full) # reported in paper
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
```



```
## Approximation) [glmerMod]
## Family: poisson ( log )
## Formula: total ~ temperature + (temperature | snail) + (1 | week) + (1 |
## parameter)
## Data: data_output_glm
##
##      AIC      BIC   logLik deviance df.resid
## 1750.6   1770.2   -868.3   1736.6     115
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.60422 -0.05826  0.00378  0.03816  0.29221
##
## Random effects:
## Groups      Name      Variance Std.Dev. Corr
## parameter (Intercept) 0.27669  0.5260
## snail      (Intercept) 2.82359  1.6804
##           temperature 0.01665  0.1290  -0.98
## week      (Intercept) 0.10865  0.3296
## Number of obs: 122, groups:  parameter, 122; snail, 16; week, 8
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   5.49280    0.98611   5.570 2.54e-08 ***
## temperature   0.03332    0.05997   0.556   0.578
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              (Intr)
## temperature -0.979
```

Figure 1

```
#calculate mean, SD and SE
data_output_graph = data.table(data_output_sing)
data_output_mean = data_output_graph[, list(mean=mean(total),sd=sd(total), se.l=(mean=mean(total)-(sqrt

#calculate cercarial ratio 15C/20C and rearrange the data
data_output_mean_2 = data.frame(data_output_mean[order(snail),])
data_output_rat = cbind(subset(data_output_mean_2, temperature == 15), subset(data_output_mean_2, tempe
colnames(data_output_rat) = c("snail", "temperature", "mean_15", "sd_15", "se.l_15", "se.u_15", "mean_20
data_output_rat$ratio = data_output_rat$mean_20/data_output_rat$mean_15
data_output_rat2 = data.table(data_output_rat)
data_output_rat = data_output_rat2[order(-ratio),]
data_output_rat$plotorder = c(1:16)

data_output_pbar=data.frame(data_output_rat)
data_output_bar = t(as.matrix(data_output_pbar[,c(3,7)]))
se.l = t(as.matrix(data_output_pbar[,c(5,9)]))
se.u = t(as.matrix(data_output_pbar[,c(6,10)]))

#drawing the graphs
colors = c("dimgrey","gray86")
ticks = c(1:length(data_output_rat$snail))
```

```

labels = as.character(c(1:length(data_output_rat$snail)))

tiff('figures/Figure_1.tiff', width = 1600, height = 1200)

par(mfrow=c(2,1), mar=c(8,11,8,4),mgp = c(5, 2, 0))
plot(ratio~plotorder, data=data_output_rat, xlab="", ylab="", ylim=c(0,6), pch = 21, las=1, bty="n",adj=
mtext(text="Genotype", side=1, line=4,cex=2.5)
mtext(text="Ratio of mean cercarial output\n (20 °C/15 °C)", side=2, line=5,cex=2.5)
mtext(text="A",side = 3,adj=0,line=0.5,cex=2)

axis(side=1, at=ticks, labels=labels,cex.axis=2)

abline(1,0,lty=3)
barplot2(data_output_bar, beside=TRUE, col=colors, xlab="", ylim=c(0,2500), plot.ci=TRUE, ci.l=se.l, ci
mtext(text="Genotype", side=1, line=4,cex=2.5)
mtext(text="Mean cercarial output\n (20 °C/15 °C)", side=2, line=5,cex=2.5)
mtext(text="B",side = 3,adj=0,line=0.5,cex=2)

axis(side=1, at=((ticks*3)-1), cex.axis=2,labels=labels)
legend("topright", c("15 °C", "20 °C"), fill=colors,cex=2)

dev.off()

## pdf
## 2

```

Analyses on cercarial survival/activity

Load the data

```

# load raw data
data_survival_raw = read.csv("data/Zea_survival_v3.0.csv", header=TRUE)

# inspect data
#summary(data_survival_raw)
str(data_survival_raw)

## 'data.frame': 14400 obs. of 9 variables:
## $ snail : Factor w/ 30 levels "Y29","Y30","Y31",...: 11 11 11 11 11 11 11 11 11 11 ...
## $ temperature: int 15 15 15 15 15 15 15 15 15 15 ...
## $ plate : int 1 1 1 1 1 1 1 1 1 1 ...
## $ tank : int 2 2 2 2 2 2 2 2 2 2 ...
## $ week : int 5 5 5 5 5 5 5 5 5 5 ...
## $ replicate : int 1 1 1 1 1 1 1 1 2 2 ...
## $ time : int 2 3 4 5 6 7 8 9 2 3 ...
## $ total : int 30 30 30 30 30 30 30 30 11 11 ...
## $ dead_cc : int 1 1 0 2 5 4 8 10 0 0 ...

# Creating a new total value, which is adjusted for the miscounted totals
data_survival_raw$total2 = data_survival_raw$total
data_survival_raw$total2[which(data_survival_raw$total < data_survival_raw$dead_cc)] = data_survival_raw

# calculating the surviving individuals
data_survival_raw$live = data_survival_raw$total2 - data_survival_raw$dead_cc
data_survival_raw$survPP = data_survival_raw$live/data_survival_raw$total2

```

```
# select replicates containing at least 10 individuals
data_survival = subset(data_survival_raw, data_survival_raw$total2 > 9)
```

```
# inspect subset of data
#summary(data_survival)
str(data_survival)
```

```
## 'data.frame': 8912 obs. of 12 variables:
## $ snail : Factor w/ 30 levels "Y29","Y30","Y31",...: 11 11 11 11 11 11 11 11 11 11 ...
## $ temperature: int 15 15 15 15 15 15 15 15 15 15 ...
## $ plate : int 1 1 1 1 1 1 1 1 1 1 ...
## $ tank : int 2 2 2 2 2 2 2 2 2 2 ...
## $ week : int 5 5 5 5 5 5 5 5 5 5 ...
## $ replicate : int 1 1 1 1 1 1 1 1 2 2 ...
## $ time : int 2 3 4 5 6 7 8 9 2 3 ...
## $ total : int 30 30 30 30 30 30 30 30 11 11 ...
## $ dead_cc : int 1 1 0 2 5 4 8 10 0 0 ...
## $ total2 : int 30 30 30 30 30 30 30 30 11 11 ...
## $ live : int 29 29 30 28 25 26 22 20 11 11 ...
## $ survPP : num 0.967 0.967 1 0.933 0.833 ...
```

```
# add column for fraction surviving cercariae (for each replicate at each time point)
data_survival$survF = data_survival$dead_cc/data_survival$total
```

```
# add column for single or multiple infections
data_survival$infection = ifelse(data_survival$snail=='Y30'
| data_survival$snail=='Y33'
| data_survival$snail=='Y35'
| data_survival$snail=='Y38'
| data_survival$snail=='Y40'
| data_survival$snail=='Y41'
| data_survival$snail=='Y42'
| data_survival$snail=='Y44'
| data_survival$snail=='Y47'
| data_survival$snail=='Y48'
| data_survival$snail=='Y51'
| data_survival$snail=='Y55'
| data_survival$snail=='Y57'
| data_survival$snail=='Y60'
| data_survival$snail=='Y61'
| data_survival$snail=='Y64', '1', '2')
```

```
# select data for time point 8
data_survival_8h = subset(data_survival, time==8)
```

```
# set variables as factor
data_survival_8h$week = as.factor(data_survival_8h$week)
data_survival_8h$plate = as.factor(data_survival_8h$plate)
data_survival_8h$replicate = as.factor(data_survival_8h$replicate)
data_survival_8h$tank = as.factor(data_survival_8h$tank)
data_survival_8h$temperature = as.factor(data_survival_8h$temperature)
data_survival_8h$infection = as.factor(data_survival_8h$infection)
```

```
# exclude snail Y35 (because it only had four data points left after removing replicates with fewer than 10 individuals)
```

```
data_survival_8h = subset(data_survival_8h, snail!='Y35')
```

```
# subset for single infections
```

```
SurvSing8 = subset(data_survival_8h, infection == '1')
```

```
SurvSing8 = droplevels(SurvSing8)
```

```
# subset for multiple infections
```

```
SurvAll8 = data_survival_8h
```

```
SurvAll8 = droplevels(SurvAll8)
```

Models

Models including all snails

```
# Compare activity (survival) after 8h for both temperatures.
```

```
model_correlation_8h_15 = glmer(cbind(live, dead_cc) ~ tank * infection + (1 | snail) + (1 | week), data = data_survival_8h_15)
```

```
model_correlation_8h_20 = glmer(cbind(live, dead_cc) ~ tank * infection + (1 | snail) + (1 | week), data = data_survival_8h_20)
```

```
summary(model_correlation_8h_15) # reported in paper
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: cbind(live, dead_cc) ~ tank * infection + (1 | snail) + (1 |
## week)
## Data: subset(SurvAll8, temperature == "15")
##
##      AIC      BIC   logLik deviance df.resid
##  3289.0   3323.9  -1636.5   3273.0     571
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -3.9818 -1.0289 -0.0922  0.9019  4.2708
##
## Random effects:
##  Groups Name            Variance Std.Dev.
##  snail  (Intercept)  0.30563   0.5528
##  week   (Intercept)  0.04618   0.2149
## Number of obs: 579, groups:  snail, 29; week, 4
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -0.36150    0.30149  -1.199   0.231
## tank2          0.45240    0.36568   1.237   0.216
## tank3         -0.47351    0.38360  -1.234   0.217
## infection2     0.08011    0.37724   0.212   0.832
## tank2:infection2 -0.53610    0.52911  -1.013   0.311
## tank3:infection2  0.35813    0.52227   0.686   0.493
##
## Correlation of Fixed Effects:
##              (Intr) tank2  tank3  infct2 tnk2:2
## tank2        -0.720
## tank3        -0.687  0.566
## infection2    -0.697  0.575  0.548
## tnk2:nfctn2   0.497 -0.691 -0.391 -0.713
## tnk3:nfctn2   0.504 -0.415 -0.734 -0.722  0.515
```

```
summary(model_correlation_8h_20) # reported in paper
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: cbind(live, dead_cc) ~ tank * infection + (1 | snail) + (1 |
## week)
## Data: subset(SurvAll8, temperature == "20")
##
##      AIC      BIC    logLik deviance df.resid
## 3083.9    3118.1  -1534.0   3067.9     523
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -4.2040 -1.0805 -0.0310  0.9076  4.3630
##
## Random effects:
## Groups Name             Variance Std.Dev.
## snail   (Intercept) 0.37322  0.6109
## week    (Intercept) 0.07425  0.2725
## Number of obs: 531, groups:  snail, 29; week, 4
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -0.2614    0.3391  -0.771  0.4409
## tank2          0.9892    0.4016   2.463  0.0138 *
## tank3          0.3980    0.4197   0.948  0.3430
## infection2     -0.3326    0.4282  -0.777  0.4373
## tank2:infection2 -0.2232    0.5912  -0.377  0.7058
## tank3:infection2  0.1514    0.5825   0.260  0.7949
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              (Intr) tank2  tank3  infct2  tnk2:2
## tank2         -0.708
## tank3         -0.677  0.572
## infection2     -0.664  0.560  0.536
## tnk2:nfctn2    0.481 -0.678 -0.388 -0.724
## tnk3:nfctn2    0.488 -0.412 -0.721 -0.735  0.533
```

```
#test for effect of tank on survival of cercariae after 8 hours for both temperatures
```

```
tank_15_all = glmer(cbind(live,dead_cc)~tank + (1|snail), family="binomial", data=subset(SurvAll8, temp
summary(tank_15_all)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: cbind(live, dead_cc) ~ tank + (1 | snail)
## Data: subset(SurvAll8, temperature == "15")
##
##      AIC      BIC    logLik deviance df.resid
## 3372.6    3390.1  -1682.3   3364.6     575
##
```

```

## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -3.7881 -0.9985 -0.0956  0.9475  4.6998
##
## Random effects:
##   Groups Name      Variance Std.Dev.
##   snail  (Intercept) 0.3318  0.5761
## Number of obs: 579, groups:  snail, 29
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -0.3711     0.1950  -1.904   0.057 .
## tank2         0.2413     0.2711   0.890   0.373
## tank3        -0.2301     0.2705  -0.851   0.395
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##      (Intr) tank2
## tank2 -0.719
## tank3 -0.721  0.518

tank_20_all = glmer(cbind(live,dead_cc)~tank + (1|snail), family="binomial", data=subset(SurvAll8, temp
summary(tank_20_all)

## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: cbind(live, dead_cc) ~ tank + (1 | snail)
## Data: subset(SurvAll8, temperature == "20")
##
##      AIC      BIC    logLik deviance df.resid
##  3235.9   3253.0  -1613.9   3227.9     527
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -4.8196 -0.9591  0.0830  1.0938  4.4149
##
## Random effects:
##   Groups Name      Variance Std.Dev.
##   snail  (Intercept) 0.4082  0.6389
## Number of obs: 531, groups:  snail, 29
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -0.4993     0.2229  -2.240  0.02512 *
## tank2         0.9239     0.3045   3.034  0.00242 **
## tank3         0.4972     0.3035   1.638  0.10137
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##      (Intr) tank2
## tank2 -0.732
## tank3 -0.734  0.537

```

```
contrasts(SurvAll8$tank) = contr.treatment(levels(SurvAll8$tank), base = 2)
tank_15_all = glmer(cbind(live,dead_cc)~tank + (1|snail), family="binomial", data=subset(SurvAll8, temp
summary(tank_15_all)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: cbind(live, dead_cc) ~ tank + (1 | snail)
## Data: subset(SurvAll8, temperature == "15")
##
##      AIC      BIC   logLik deviance df.resid
##  3372.6   3390.1  -1682.3   3364.6     575
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -3.7881 -0.9985 -0.0956  0.9475  4.6998
##
## Random effects:
## Groups Name          Variance Std.Dev.
## snail (Intercept) 0.3318   0.5761
## Number of obs: 579, groups:  snail, 29
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -0.1299     0.1883  -0.690  0.4904
## tank1         -0.2413     0.2711  -0.890  0.3734
## tank3         -0.4714     0.2657  -1.774  0.0761 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##      (Intr) tank1
## tank1 -0.695
## tank3 -0.709  0.492
```

```
tank_20_all = glmer(cbind(live,dead_cc)~tank + (1|snail), family="binomial", data=subset(SurvAll8, temp
summary(tank_20_all)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: cbind(live, dead_cc) ~ tank + (1 | snail)
## Data: subset(SurvAll8, temperature == "20")
##
##      AIC      BIC   logLik deviance df.resid
##  3235.9   3253.0  -1613.9   3227.9     527
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -4.8196 -0.9591  0.0830  1.0938  4.4149
##
## Random effects:
## Groups Name          Variance Std.Dev.
## snail (Intercept) 0.4082   0.6389
```

```
## Number of obs: 531, groups: snail, 29
##
## Fixed effects:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.4246    0.2074   2.047  0.04062 *
## tank1       -0.9239    0.3045  -3.034  0.00242 **
## tank3       -0.4267    0.2924  -1.459  0.14449
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##      (Intr) tank1
## tank1 -0.681
## tank3 -0.710  0.484
```

```
contrasts(SurvAll8$tank) = contr.treatment(levels(SurvAll8$tank), base = 1)
```

Models for single infected snails

Compare survival between temperature treatments after 8 hours

tests without random effects

```
model_survival_01 = glm(cbind(live, dead_cc) ~ temperature, data = SurvSing8, family = "binomial") # re
summary(model_survival_01)
```

```
##
## Call:
## glm(formula = cbind(live, dead_cc) ~ temperature, family = "binomial",
##      data = SurvSing8)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -6.0224  -1.3977  -0.0265   1.3095   7.0167
##
## Coefficients:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.40656    0.02796  -14.54  <2e-16 ***
## temperature20 0.57544    0.03611   15.94  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 2914.0  on 567  degrees of freedom
## Residual deviance: 2656.8  on 566  degrees of freedom
## AIC: 4465.8
##
## Number of Fisher Scoring iterations: 4
```

```
SurvSing8$temperatureF = as.factor(seq(1:length(SurvSing8$temperature)))
```

```
model_survival_02 = glmer(cbind(live, dead_cc) ~ temperature * tank + (temperature | snail) + (1 | week)
model_survival_03 = glmer(cbind(live, dead_cc) ~ temperature * tank + (temperature | snail) + (1 | week)
```

```
contrasts(SurvSing8$tank) = contr.treatment(levels(SurvSing8$tank), base = 2)
```

```
model_survival_04 = glmer(cbind(live, dead_cc) ~ temperature * tank + (temperature | snail) + (1 | week)
```



```

contrasts(SurvSing8$tank) = contr.treatment(levels(SurvSing8$tank), base = 1)

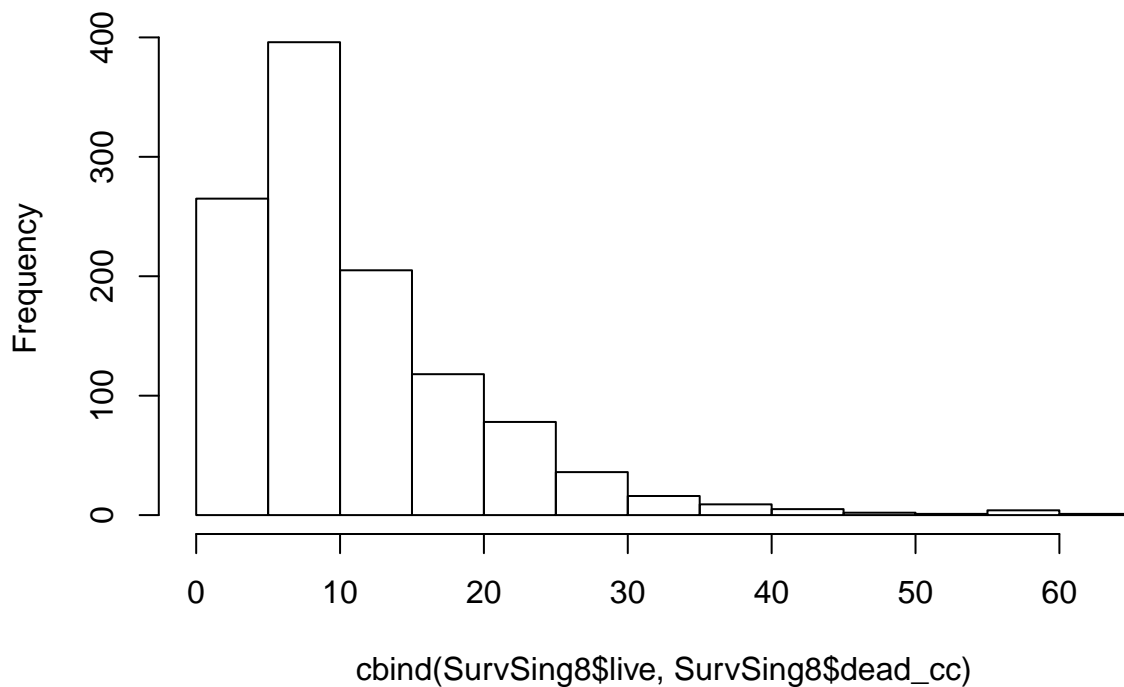
# test for normality
shapiro.test(cbind(SurvSing8$live, SurvSing8$dead_cc)) #not normaly distributed

##
## Shapiro-Wilk normality test
##
## data: cbind(SurvSing8$live, SurvSing8$dead_cc)
## W = 0.85712, p-value < 2.2e-16

hist(cbind(SurvSing8$live, SurvSing8$dead_cc)) #right skewed, p << 0.05

```

Histogram of cbind(SurvSing8\$live, SurvSing8\$dead_cc)

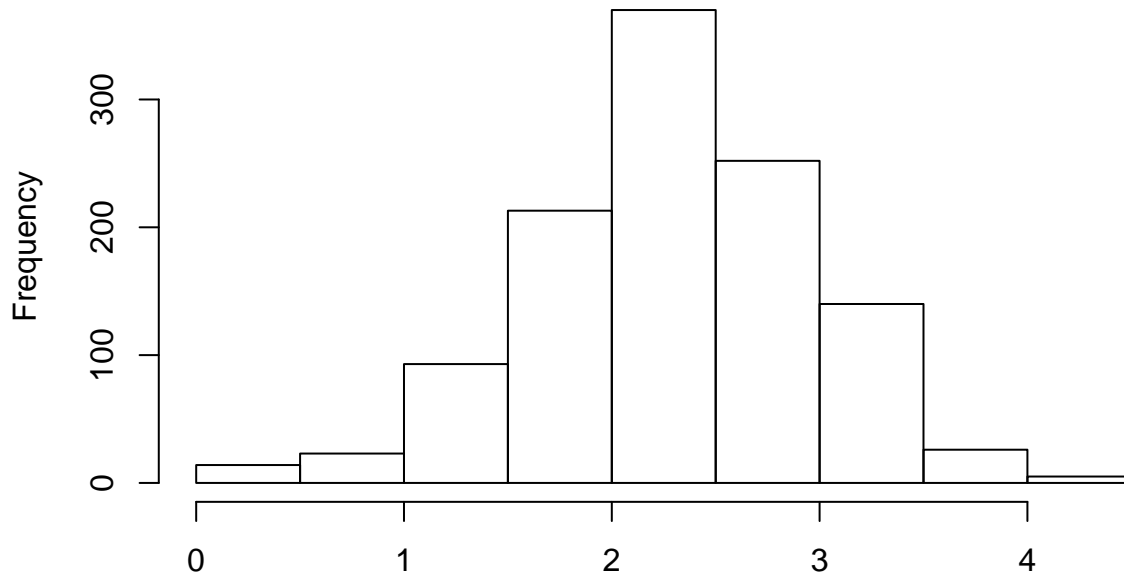


```

hist(cbind(log1p(SurvSing8$live), log1p(SurvSing8$dead_cc))) # looks normal

```

Histogram of cbind(log1p(SurvSing8\$live), log1p(SurvSing8\$dead_cc))



cbind(log1p(SurvSing8\$live), log1p(SurvSing8\$dead_cc))

```
shapiro.test(cbind(log1p(SurvSing8$live), log1p(SurvSing8$dead_cc))) # tells you still not normal, but l
```

```
##
## Shapiro-Wilk normality test
##
## data: cbind(log1p(SurvSing8$live), log1p(SurvSing8$dead_cc))
## W = 0.98056, p-value = 3.322e-11
```

```
slp_nslp = anova(model_survival_03, model_survival_02)
slp_nslp # reported in paper
```

```
## Data: SurvSing8
## Models:
## model_survival_02: cbind(live, dead_cc) ~ temperature * tank + (temperature | snail) +
## model_survival_02:      (1 | week)
## model_survival_03: cbind(live, dead_cc) ~ temperature * tank + (temperature | snail) +
## model_survival_03:      (1 | week) + (1 | temperatureF)
##
##          npar    AIC    BIC logLik deviance Chisq Df Pr(>Chisq)
## model_survival_02    10 3142.6 3186.0 -1561.3   3122.6
## model_survival_03    11 2943.2 2990.9 -1460.6   2921.2 201.44  1 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary(model_survival_03) # reported in paper
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: cbind(live, dead_cc) ~ temperature * tank + (temperature | snail) +
##          (1 | week) + (1 | temperatureF)
## Data: SurvSing8
```

```
##
##      AIC      BIC    logLik deviance df.resid
##  2943.2   2990.9  -1460.6   2921.2     557
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.22175 -0.44403  0.01158  0.45487  2.21888
##
## Random effects:
##   Groups      Name      Variance Std.Dev. Corr
## temperatureF (Intercept)  0.23294  0.4826
## snail         (Intercept)  0.30058  0.5482
##              temperature20 0.47354  0.6881  -0.51
## week          (Intercept)  0.04236  0.2058
## Number of obs: 568, groups:  temperatureF, 568; snail, 15; week, 8
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -0.41683    0.30288  -1.376   0.169
## temperature20     0.09752    0.38968   0.250   0.802
## tank2             0.44017    0.37095   1.187   0.235
## tank3            -0.38720    0.39125  -0.990   0.322
## temperature20:tank2  0.59527    0.47118   1.263   0.206
## temperature20:tank3  0.80462    0.49618   1.622   0.105
##
## Correlation of Fixed Effects:
##              (Intr) tmpr20 tank2  tank3  tm20:2
## temperatr20 -0.545
## tank2        -0.722  0.372
## tank3        -0.687  0.354  0.560
## tmprtr20:t2  0.377 -0.711 -0.527 -0.292
## tmprtr20:t3  0.359 -0.675 -0.293 -0.534  0.558
```

Figure 2

```
#calculate mean and sd's to plot
fig2_mean_15 = tapply(subset(data_survival$survPP*100, data_survival$temperature == '15'), subset(data_survival$survPP*100, data_survival$temperature == '15'), subset(data_survival$survPP*100, data_survival$temperature == '20'), subset(data_survival$survPP*100, data_survival$temperature == '20'), subset(data_survival$survPP*100, data_survival$temperature == '20'))
fig2_sd_15 = tapply(subset(data_survival$survPP*100, data_survival$temperature == '15'), subset(data_survival$survPP*100, data_survival$temperature == '15'), subset(data_survival$survPP*100, data_survival$temperature == '20'), subset(data_survival$survPP*100, data_survival$temperature == '20'))
fig2_sd_20 = tapply(subset(data_survival$survPP*100, data_survival$temperature == '20'), subset(data_survival$survPP*100, data_survival$temperature == '20'), subset(data_survival$survPP*100, data_survival$temperature == '20'))

#create dataframe for plotting
fig2_xval = c(1.965:8.965,2.035:9.035)
fig2_treat = rep(1:2,each=8)
fig2_dummy = cbind(fig2_xval,c(fig2_mean_15,fig2_mean_20),fig2_treat)

#set parameters for errorbars
ang = 90
len = 0.05

col1 = "dimgrey"
col2 = "gray86"

#plot graph
tiff('figures/Figure_2.tiff', width=1600, height=1200)
```

```

colors = c(col1,col2)
par(mfrow=c(1,1), mar=c(8,8,4,4),mgp = c(5, 2, 0))
plot(fig2_dummy, ylim=c(0,100), col=ifelse(fig2_treat == '1',col1, col2), pch=ifelse(fig2_treat == '1',
mtext(text="Hours after emergence", line=4, side=1, cex=2.5)
mtext(text="Percentage of cercariae active", line=4, side=2,cex=2.5)
legend("topright", c("15 °C", "20 °C"), fill=colors,cex=2.5)
arrows(fig2_xval[1:8], fig2_mean_15-1, fig2_xval[1:8], fig2_mean_15-fig2_sd_15, angle=ang, length=len,
arrows(fig2_xval[1:8], fig2_mean_15+fig2_sd_15, fig2_xval[1:8], fig2_mean_15+1, angle=ang, length=len,
arrows(fig2_xval[9:16], fig2_mean_20-1, fig2_xval[9:16], fig2_mean_20-fig2_sd_20, angle=ang, length=len,
arrows(fig2_xval[9:16], fig2_mean_20+fig2_sd_20, fig2_xval[9:16], fig2_mean_20+1, angle=ang, length=len

dev.off()

```

```

## pdf
## 2

```

Figure 3

```

data_survival_figure = SurvSing8
data_survival_figure$plotorder = ifelse(data_survival_figure$snail=="Y60",1,
ifelse(data_survival_figure$snail=="Y64",2,
ifelse(data_survival_figure$snail=="Y61",3,
ifelse(data_survival_figure$snail=="Y41",4,
ifelse(data_survival_figure$snail=="Y47",5,
ifelse(data_survival_figure$snail=="Y38",6,
ifelse(data_survival_figure$snail=="Y37",7,
ifelse(data_survival_figure$snail=="Y36",8,
ifelse(data_survival_figure$snail=="Y35",9,
ifelse(data_survival_figure$snail=="Y34",10,
ifelse(data_survival_figure$snail=="Y33",11,
ifelse(data_survival_figure$snail=="Y32",12,
ifelse(data_survival_figure$snail=="Y31",13,
ifelse(data_survival_figure$snail=="Y30",14,
ifelse(data_survival_figure$snail=="Y29",15,
ifelse(data_survival_figure$snail=="Y28",16,
ifelse(data_survival_figure$snail=="Y27",17,
ifelse(data_survival_figure$snail=="Y26",18,
ifelse(data_survival_figure$snail=="Y25",19,
ifelse(data_survival_figure$snail=="Y24",20,
ifelse(data_survival_figure$snail=="Y23",21,
ifelse(data_survival_figure$snail=="Y22",22,
ifelse(data_survival_figure$snail=="Y21",23,
ifelse(data_survival_figure$snail=="Y20",24,
ifelse(data_survival_figure$snail=="Y19",25,
ifelse(data_survival_figure$snail=="Y18",26,
ifelse(data_survival_figure$snail=="Y17",27,
ifelse(data_survival_figure$snail=="Y16",28,
ifelse(data_survival_figure$snail=="Y15",29,
ifelse(data_survival_figure$snail=="Y14",30,
ifelse(data_survival_figure$snail=="Y13",31,
ifelse(data_survival_figure$snail=="Y12",32,
ifelse(data_survival_figure$snail=="Y11",33,
ifelse(data_survival_figure$snail=="Y10",34,
ifelse(data_survival_figure$snail=="Y9",35,
ifelse(data_survival_figure$snail=="Y8",36,
ifelse(data_survival_figure$snail=="Y7",37,
ifelse(data_survival_figure$snail=="Y6",38,
ifelse(data_survival_figure$snail=="Y5",39,
ifelse(data_survival_figure$snail=="Y4",40,
ifelse(data_survival_figure$snail=="Y3",41,
ifelse(data_survival_figure$snail=="Y2",42,
ifelse(data_survival_figure$snail=="Y1",43)

data_survival_figure$plotorder = as.integer(data_survival_figure$plotorder)

tiff(filename="figures/Figure_3.tiff", width=2000, height=2000)

par(mfrow = c(2,1), mar=c(8,8,4,4),mgp = c(5, 2, 0))
boxplot((live*100/total2)~plotorder, data = subset(data_survival_figure, temperature == 15), ylim=c(0,
mtext(text="A",side = 3,adj=0,line=0.5,cex=4.5)
mtext(text="Genotype", side = 1, line=4,cex=2.5)
mtext(text="Percentage of cercariae active", side = 2, line=5,cex=2.5)
legend("topright", c("15 °C"), fill=col1,cex=2.5)
abline(a=50, b=0, col="#000000")
abline(v=c(5.5,10.5),lty=3)
boxplot((live*100/total2)~plotorder, data = subset(data_survival_figure, temperature == 20), ylim=c(0,
mtext(text="B",side = 3,adj=0,line=0.5,cex=4.5)
mtext(text="Genotype", side = 1, line=4,cex=2.5)

```

```
## pdf
## 2
```

Extract the data

21

```

corr = lm(survival~output, data=data_survival_corr_05)
summary(corr) # reported in paper

##
## Call:
## lm(formula = survival ~ output, data = data_survival_corr_05)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -30.597 -14.166  -2.277   15.286   29.039
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  57.874287   6.496937   8.908 1.16e-09 ***
## output       -0.007030   0.008847  -0.795   0.434
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 17.46 on 28 degrees of freedom
## Multiple R-squared:  0.02205,    Adjusted R-squared:  -0.01288
## F-statistic: 0.6314 on 1 and 28 DF,  p-value: 0.4335

```

Figure 4 Figure displaying the absence in correlation between cercarial output and survival of cercariae after 8 hours.

```

col1 = "dimgrey"
col2 = "gray86"
colors = c(col1,col2)

#plotting
tiff('figures/Figure_4.tiff', width=1600, height=1600)

par(mfrow=c(1,1),mar=c(10,10,4,4),mgp = c(5, 2, 0))
plot(data_survival_corr_05[,3]~data_survival_corr_05[,1], pch=16, col=ifelse(data_survival_corr_05$temp,
    bty="n", las=1, ylab="",xlab="",
    cex=2.5,cex.axis=2.5,cex.lab=2.5)
mtext(text="Mean cercarial output", side=1, line=4,cex=2.5)
mtext(text="Percentage of cercariae active 8 h after emergence", side=2, line=5,cex=2.5)
legend("topright", c("15 °C", "20 °C"), fill=colors,cex=2.5)
dev.off()

## pdf
## 2

```