Analyses for bed bug sexual dynamics manuscript

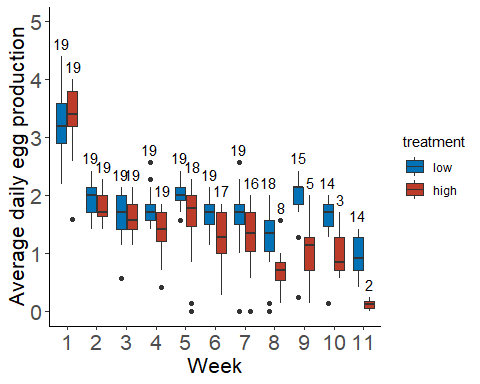
Janice Yan

2023-07-29

Updated analyses for the sexual dynamics manuscript Experiments 1 and 3.

# Experiment 1: The fitness cost of traumatic insemination

## a) Daily egg laying rate



The goal of this analysis is to see whether insemination rate affects the egg production rate of females while they are alive. Therefore, when females died, eggs produced was entered as NA, not zero.

Here I fit a GLMM with treatment and week as fixed factors and arena number as a random factor (nested within treatment because the same numbers are used across treatments). I used number of eggs as the response variable with log(days\_alive) as an offset to account for females dying in the middle of a week and having fewer days to produce eggs as suggested by Ben Bolker in data-lunch. Using total number of eggs with an offset for days instead of egg production rate (eggs laid/days for laying) allows me to use a poisson or negative binomial distribution instead of the standard linear model. I showed the plots for the linear model at data-lunch and both Ben and Jonathan suggested poisson or negative binomial instead.

I ended up using a negative binomial distribution. The residuals vs. predicted plot still isn’t ideal (red lines don’t align with black lines well) but this is the best fit I’ve been able to achieve and better than using the standard linear model or the poisson distribution. I plotted the residuals against my predictors (week and treatment) as recommended by others. Within-group distribution and variance appear to be fairly uniform across the the two treatments when I plot the model residuals against treatment so I think we are overall fine to use this model.

I’ve included the code used to generate all statistical models.

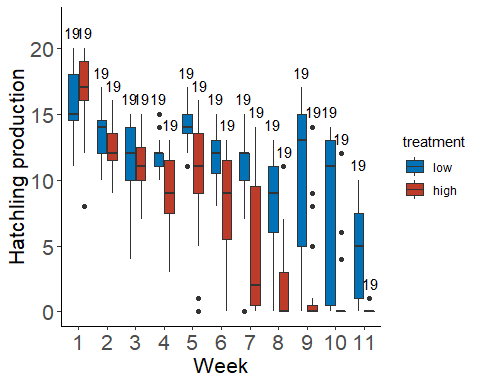
egg\_mod <- glmmTMB(data = weekly\_data, eggs ~ treatment + week + offset(log(days\_alive)) + (1|treatment:arena),  
 family = nbinom1())

Anova(egg\_mod)

## Analysis of Deviance Table (Type II Wald chisquare tests)  
##   
## Response: eggs  
## Chisq Df Pr(>Chisq)   
## treatment 28.027 1 1.196e-07 \*\*\*  
## week 139.144 1 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

**Overall, we get a highly significant result of both treatment and week.**

## b) Lifetime hatchling production



Here we wanted to look at the effect of insemination rate on hatchling production. When females died, hatchlings produced was entered as 0 and not NA. These results are intentionally not independent of differential mortality since we are interested in lifetime fitness here.

For this analysis, I used a linear model with hatchlings produced per week as the response variable. I included treatment and week as fixed factors and arena number as a random factor. I tried fitting a Poisson distribution seeing as this is count data but the model fit appeared to be worse than a standard linear model based on the DHARMa diagnostic plots. When I try to use a poisson distribution, the residuals deviate much more from the model prediction and I also get a warning that says “qu = 0.25, log(sigma) = -2.424222 : outer Newton did not converge fully.” Negative binomial looked better than poisson but not better than a linear model.

hatchling\_mod <- lmer(data = weekly\_data, hatchlings ~ treatment + week + (1|treatment:arena))

Anova(hatchling\_mod)

## Analysis of Deviance Table (Type II Wald chisquare tests)  
##   
## Response: hatchlings  
## Chisq Df Pr(>Chisq)   
## treatment 42.947 1 5.624e-11 \*\*\*  
## week 456.363 1 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

**Again, we get a highly significant result of both treatment and week.**

## c) Egg hatch rates

Lastly, we wanted to examine whether eggs from females of both treatments hatched at different rates as a measure of indirect fitness. Increased egg viability in high-insemination rate females could be an indication that many inseminations provide a genetic benefit to females.

To analyse this, we used cbind(hatched eggs, unhatched eggs) as the response variable in a binomial model as suggested by a reviewer. Again we had treatment and week as fixed factors and arena number as a random factor.

Anova(hatch\_rate\_mod)

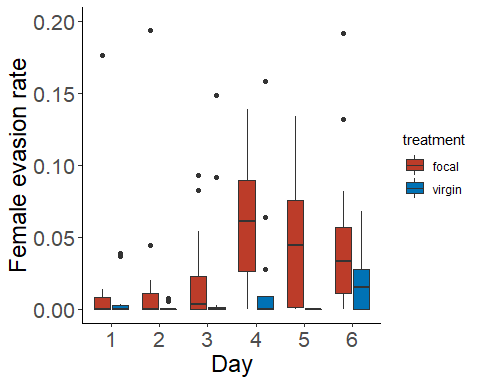
## Analysis of Deviance Table (Type II Wald chisquare tests)  
##   
## Response: cbind(hatchlings, (eggs\_no\_na - hatchlings))  
## Chisq Df Pr(>Chisq)  
## treatment 1.9993 1 0.1574  
## week 0.1838 1 0.6681

**Hatch rate did not significantly differ between groups. We did not intend on including a figure in the manuscript for this which is why I haven’t generated one here.**

# Experiment 3: Responses to consecutive daily inseminations

In this experiment, we tracked change in insemination duration, female evasion rate (proportion of trial spent evading males), and insemination latency as females received one insemination per day for six consecutive days. For all of these models, we controlled for day effects by subtracting the average raw score for the 13 virgin females tested on a given day from the raw score of each of the 13 focal females. For example, the average insemination duration for virgins on day 1 was 74.749 seconds so we subtracted that value from all the day 1 insemination durations for focal females. We then added a positive integer to all the focal values across all days so that all values were positive making it easy to transform our response variables to achieve better model fits if needed. When picking what integer to add, we chose the smallest one possible that would ensure no negative or zero values.

## a) Change in evasion rate as a function of prior consecutive inseminations



I first looked at proportion of trial duration females spent evading males. I fit a LMM with the log of evasion rate, after controlling for day effects, as the response variable and day (which directly corresponds to # of inseminations) as a fixed effect. Arena was included as a random effect. I logged the response variable to meet model assumptions because had I not, diagnostic plots showed several issues in both the QQ residual plot and residuals as a function of categorical predictor plot.

I did get a singular fit warning meaning the estimation of the random effect variance appears to be zero. I’ve left the random effect in the model but I believe it is effectively doing nothing and makes no difference?

evade\_mod <- lmer(data = focal\_data, log(con\_prop\_run\_2) ~ day +   
 (1|arena))

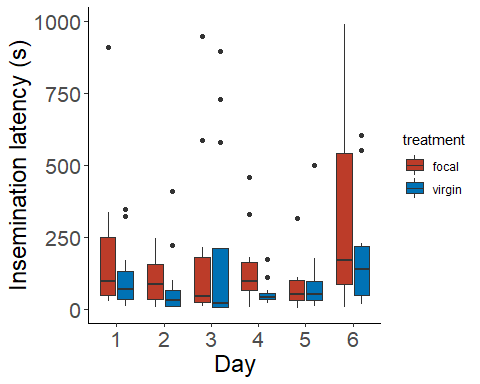
## boundary (singular) fit: see ?isSingular

#plot(simulateResiduals(running\_away\_mod))  
Anova(evade\_mod)

## Analysis of Deviance Table (Type II Wald chisquare tests)  
##   
## Response: log(con\_prop\_run\_2)  
## Chisq Df Pr(>Chisq)   
## day 5.3473 1 0.02075 \*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

**Overall, number of prior inseminations does significantly affect female evasion rate.**

## b) Change in insemination latency as a function of prior consecutive inseminations



I then analyzed insemination latency with day as a fixed effect and arena as a random effect. I once again logged the response variable which was insemination latency in seconds minus the average day effect score for each day plus an integer to avoid negative values. This helped me achieve a better fit for the model. Even after logging, I still get a warning for Levene’s homogeneity of variance test when generating the diagnostic plots but the plot itself doesn’t look too bad so I think this is a reasonable model to go with.

latency\_mod <- lmer(data = focal\_data, log(con\_insem\_lat\_2) ~ day + (1|arena))

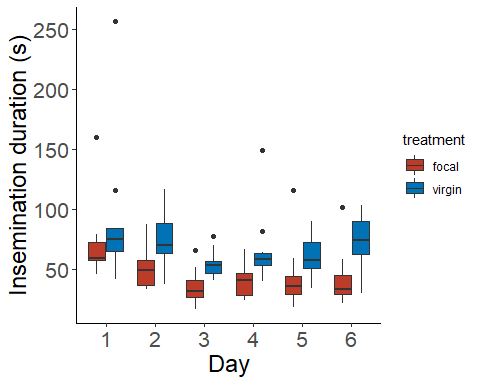
## boundary (singular) fit: see ?isSingular

#plot(simulateResiduals(latency\_mod))  
Anova(latency\_mod)

## Analysis of Deviance Table (Type II Wald chisquare tests)  
##   
## Response: log(con\_insem\_lat\_2)  
## Chisq Df Pr(>Chisq)  
## day 1.1246 1 0.2889

**Number of prior inseminations did not affect insemination latency.**

## c) Change in insemination duration as a function of prior consecutive inseminations



Lastly, for the insemination duration model, we fit a LMM with insemination duration as the response variable, day as a fixed effect, and arena as a random effect. We controlled for day effects the same way we did for insemination latency. The DHARMa residual plot looks completely fine.

# Insemination duration model  
duration\_mod <- lmer(data = focal\_data, con\_insem\_dur ~ day + (1|arena))  
  
Anova(duration\_mod)

## Analysis of Deviance Table (Type II Wald chisquare tests)  
##   
## Response: con\_insem\_dur  
## Chisq Df Pr(>Chisq)   
## day 6.653 1 0.009899 \*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

**Overall, insemination duration significantly decreases as females receive more consecutive inseminations.**