**R code and model output for initial male competition assay**

This document presents the code and output from the first round of the male competition assay from the dates 5th, 6th and 7th of May 2018. This is part of the experiments focussing on the M lines 5, 8, 16, 19, 24, and 95 who have had either male or female benefit fruitless alleles introgressed into a deletion stock background. This assay assessed the competitive ability, and hence fitness, of males when in competition with a competitor male. 1 focal male and 1 competitor male where kept together overnight and in the morning a female from the same stock as the competitor was added. The flies were then left for 90 minutes. After this time, both the males were removed and the female was left to lay eggs. Due to the genotype of the two males, we can tell which male fathered the off spring of the female by looking at the phenotype of the pupa. If all the pupa display the ‘tubby’ phenotype then the competitors male was successful. If the pupa are roughly 50:50 tubby and normal then the focal male was successful. The results were recorded as either ‘tubby’ for all tubby, or ‘norm’ for some normal pupa. This result was transformed to give a proportion of trials that were won by the focal male.

setwd("/Users/michaeljardine/Desktop/DTP/Datasets")

## PACKAGES

library(Rmisc)

library(lme4)

library(aod)

library(lsmeans)

library(ggplot2)

## read in data and check

comp<-read.csv("malecompassayALT.csv")

> str(comp)

'data.frame': 243 obs. of 6 variables:

$ Block : int 1 1 1 1 1 1 1 1 1 1 ...

$ Round : int 1 1 1 1 1 1 1 1 1 1 ...

$ Allele : Factor w/ 2 levels "F","M": 2 2 2 2 2 2 2 2 2 2 ...

$ Line : int 5 5 5 5 5 5 5 5 8 8 ...

$ Genotype: Factor w/ 2 levels "B","D": 1 1 1 1 1 1 1 1 1 1 ...

$ Pupa : Factor w/ 2 levels "Norm","Tub": 1 2 1 2 1 2 1 1 2 2 ...

### line = the number of the line e.g. M5 or M24 etc

### genotype = the chromosomal genotype of the line. Along with the chromosome possessing the introgressed piece of DNA, all flies have an accompanying chromosome that has either a deletion at the fruitless locus, or a balancer chromosome which results in the tubby phenotype.

### allele = refers to the fruitless allele that was introgressed into the ancestral stock, either Male (M) or female (F) beneficial allele.

### pupa = the type of pupa observed

### round = the experiment was repeated over several days with males of the same age on each day.

> ## change Line to a factor

> comp$Line <- factor(comp$Line)

> ## change Round to a factor

> comp$Round <- factor(comp$Round)

> ## change Block to a factor

> comp$Block <- factor(comp$Block)

> str(comp)

'data.frame': 243 obs. of 6 variables:

$ Block : Factor w/ 1 level "1": 1 1 1 1 1 1 1 1 1 1 ...

$ Round : Factor w/ 3 levels "1","2","3": 1 1 1 1 1 1 1 1 1 1 ...

$ Allele : Factor w/ 2 levels "F","M": 2 2 2 2 2 2 2 2 2 2 ...

$ Line : Factor w/ 6 levels "5","8","16","19",..: 1 1 1 1 1 1 1 1 2 2 ...

$ Genotype: Factor w/ 2 levels "B","D": 1 1 1 1 1 1 1 1 1 1 ...

$ Pupa : Factor w/ 2 levels "Norm","Tub": 1 2 1 2 1 2 1 1 2 2 ...

> ## remove NA, some trials failed. Either because neither male succeed in mating or because the female died soon after the trial began.

> comp <- na.omit(comp)

> ## change pupa to a 1/0 score to see who the winner is

> comp$winner <- ifelse(comp$Pupa %in% c("Tub"), 0, 1)

## winner = competitor = 0 (Tub), focal = 1 (Norm)

### we have 6 lines, each with 2 different genotype for 12 different combinations

### combine these two factors to reflect this

## combine the line and the genotype to create 5B and 5D

comp$Linegenotype <- interaction(comp$Line, comp$Genotype)

str(comp)

## some summary statistics for means and errors

> summarySE(comp, measurevar=c("winner"), groupvars=c("Allele"))

Allele N winner sd se ci

1 F 100 0.5900000 0.4943111 0.04943111 0.09808204

2 M 121 0.4958678 0.5020619 0.04564199 0.09036796

> summarySE(comp, measurevar=c("winner"), groupvars=c("Line"))

Line N winner sd se ci

1 5 40 0.4750000 0.5057363 0.07996393 0.1617423

2 8 42 0.6190476 0.4915074 0.07584124 0.1531645

3 16 39 0.3846154 0.4928641 0.07892141 0.1597680

4 19 39 0.5384615 0.5050354 0.08087038 0.1637135

5 24 31 0.7096774 0.4614144 0.08287247 0.1692482

6 95 30 0.5333333 0.5074163 0.09264111 0.1894723

> summarySE(comp, measurevar=c("winner"), groupvars=c("Genotype"))

Genotype N winner sd se ci

1 B 109 0.5596330 0.4987242 0.04776911 0.09468667

2 D 112 0.5178571 0.5019268 0.04742762 0.09398100

> summarySE(comp, measurevar=c("winner"), groupvars=c("Linegenotype"))

Linegenotype N winner sd se ci

1 5.B 23 0.5217391 0.5107539 0.1064996 0.2208666

2 8.B 24 0.6250000 0.4945354 0.1009466 0.2088240

3 16.B 22 0.4090909 0.5032363 0.1072903 0.2231225

4 19.B 21 0.4285714 0.5070926 0.1106567 0.2308258

5 24.B 10 0.7000000 0.4830459 0.1527525 0.3455502

6 95.B 9 1.0000000 0.0000000 0.0000000 0.0000000

7 5.D 17 0.4117647 0.5072997 0.1230382 0.2608294

8 8.D 18 0.6111111 0.5016313 0.1182356 0.2494554

9 16.D 17 0.3529412 0.4925922 0.1194712 0.2532675

10 19.D 18 0.6666667 0.4850713 0.1143324 0.2412203

11 24.D 21 0.7142857 0.4629100 0.1010153 0.2107141

12 95.D 21 0.3333333 0.4830459 0.1054093 0.2198799

# these numbers can be used to make basic plots of the data

### MODELS ###

## the first model simply looks at the if the winner of a particular contest varies between the 12 line and genotype combinations and if there is an interaction between the two

> GLMM2 <- glmer(winner ~ Line\*Genotype + (1 | Round), data=comp, family=binomial(link=logit))

Warning message:

In checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :

Model is nearly unidentifiable: large eigenvalue ratio

- Rescale variables?

> summary(GLMM2)

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']

Family: binomial ( logit )

Formula: winner ~ Line \* Genotype + (1 | Round)

Data: comp

AIC BIC logLik deviance df.resid

304.2 348.4 -139.1 278.2 208

Scaled residuals:

Min 1Q Median 3Q Max

-1.58114 -0.83666 0.00015 0.79772 1.41421

Random effects:

Groups Name Variance Std.Dev.

Round (Intercept) 0 0

Number of obs: 221, groups: Round, 3

> anova(GLMM2)

Analysis of Variance Table

Df Sum Sq Mean Sq F value

Line 5 11.7501 2.35001 2.3500

Genotype 1 0.0295 0.02947 0.0295

Line:Genotype 5 2.7498 0.54996 0.5500

> wald.test(b=fixef(GLMM2), Sigma=vcov(GLMM2), Terms = 2, df=5)

Wald test:

Chi-squared test:

X2 = 0.51, df = 1, P(> X2) = 0.48

F test:

W = 0.51, df1 = 1, df2 = 5, P(> W) = 0.51

> wald.test(b=fixef(GLMM2), Sigma=vcov(GLMM2), Terms = 3, df=1)

Wald test:

Chi-squared test:

X2 = 0.57, df = 1, P(> X2) = 0.45

F test:

W = 0.57, df1 = 1, df2 = 1, P(> W) = 0.59

> wald.test(b=fixef(GLMM2), Sigma=vcov(GLMM2), Terms = 4, df=5)

Wald test:

Chi-squared test:

X2 = 0.38, df = 1, P(> X2) = 0.54

F test:

W = 0.38, df1 = 1, df2 = 5, P(> W) = 0.56

### Warning message in model! Result of no variation in 95B??

### should we remove the 95 lines and run again??

### should we have a more complete model?? With allele as well???

### MODEL REMOVING THE 95 LINES ###

> lines5 <- glmer(winner ~ Line\*Genotype + (1 | Round), data=fivelines, family=binomial(link=logit))

## at the least the warning message had gone!!

> summary(lines5)

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']

Family: binomial ( logit )

Formula: winner ~ Line \* Genotype + (1 | Round)

Data: fivelines

AIC BIC logLik deviance df.resid

273.5 309.2 -125.7 251.5 180

Scaled residuals:

Min 1Q Median 3Q Max

-1.5811 -0.8660 0.6325 0.7977 1.3540

Random effects:

Groups Name Variance Std.Dev.

Round (Intercept) 0 0

Number of obs: 191, groups: Round, 3

> anova(lines5)

Analysis of Variance Table

Df Sum Sq Mean Sq F value

Line 4 8.7707 2.19267 2.1927

Genotype 1 0.0295 0.02947 0.0295

Line:Genotype 4 2.7497 0.68743 0.6874

> wald.test(b=fixef(lines5), Sigma=vcov(lines5), Terms = 2, df=4)

Wald test:

Chi-squared test:

X2 = 0.51, df = 1, P(> X2) = 0.48

F test:

W = 0.51, df1 = 1, df2 = 4, P(> W) = 0.51

> wald.test(b=fixef(lines5), Sigma=vcov(lines5), Terms = 3, df=1)

Wald test:

Chi-squared test:

X2 = 0.57, df = 1, P(> X2) = 0.45

F test:

W = 0.57, df1 = 1, df2 = 1, P(> W) = 0.59

> wald.test(b=fixef(lines5), Sigma=vcov(lines5), Terms = 4, df=4)

Wald test:

Chi-squared test:

X2 = 0.38, df = 1, P(> X2) = 0.54

F test:

W = 0.38, df1 = 1, df2 = 4, P(> W) = 0.57

### However this seems to have had no overall effect on the outcome of the model – no effect on the outcome of male competition due to its line or genotype.

### MODEL LOOKING AT THE EFFECT OF FRUITLESS ALLELE ###

### a simple model just including allele to see if this effects the version of the introgressed fruitless allele has an effect on male success

> alleleonly <- glmer(winner ~ Allele + (1 | Round), data=comp, family=binomial(link=logit))

> summary(alleleonly)

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']

Family: binomial ( logit )

Formula: winner ~ Allele + (1 | Round)

Data: comp

AIC BIC logLik deviance df.resid

309.1 319.3 -151.6 303.1 218

Scaled residuals:

Min 1Q Median 3Q Max

-1.1996 -0.9918 0.8336 1.0083 1.0083

Random effects:

Groups Name Variance Std.Dev.

Round (Intercept) 0 0

Number of obs: 221, groups: Round, 3

Fixed effects:

Estimate Std. Error z value Pr(>|z|)

(Intercept) 0.3640 0.2033 1.790 0.0734 .

AlleleM -0.3805 0.2728 -1.395 0.1630

> anova(alleleonly)

Analysis of Variance Table

Df Sum Sq Mean Sq F value

Allele 1 1.9459 1.9459 1.9459

> wald.test(b=fixef(alleleonly), Sigma=vcov(alleleonly), Terms = 2, df=1)

Wald test:

Chi-squared test:

X2 = 1.9, df = 1, P(> X2) = 0.16

F test:

W = 1.9, df1 = 1, df2 = 1, P(> W) = 0.4

### appear that there may be some difference but not significantly so

### where do we go from here?

### other models we can do?

### or have I done something wrong?