# List the packages to be used

library(car)

library(emmeans)

library(GGally)

library(lme4)

library(tidyverse)

# Note that 2024\_02\_12\_GCP\_preliminary\_glmer.csv contains only the data from visit 3 (June 3) onwards

# because many of the stems that were clipped in half during visit 1 were not found visit 2 and so they were

# replaced with randomly chosen new clones during visit 2 (May 30) and then data were recorded from visit 3

# onwards for all these stems.

# Note also that these data are recorded in the following ways: 1) All subsequent visits for a stem are recorded as

# zeros for alive1\_dead0 and for height and eggs where there is a note that the plant is dead or when the

# height of the plant was recorded as zero for every subsequent visit; 2) If stems could not be found then they

# are recorded as "NA" for all subsequent visits because we cannot be sure if the stem was eaten, even though

# the most likely scenario is that the stem died (i.e. senesced) but recorded as missing data (i.e. "NA") is

# the most conservative option to make certain we do not inflate the percentage of dead stems.

# Set the working directory (i.e. folder where .csv data files are located)

setwd(choose.dir())

####### First examine survival per stem, averaged per clone (Fig. 1 in manuscript)

AVG\_clones <- read.csv("Poynor\_Dickson\_average\_per\_clone\_June3\_Sept16.csv")

head(AVG\_clones)

# Change variables that are numbers to factors (i.e. categorical variables), as appropriate

AVG\_clones <- within(AVG\_clones, {

visit <- factor(visit)

rep\_ID <- factor(rep\_ID)

clone <- factor(clone)

})

# Complete a repeated-measures mixed-effect Analysis of Variance (i.e. linear mixed-effect analysis)

clone\_survival\_analysis <- lmer(formula = AVGalive1\_dead0 ~ visit \* treatment + (1 | rep\_ID) , data = AVG\_clones)

summary(clone\_survival\_analysis)

anova(clone\_survival\_analysis)

# Histogram of residuals of analysis

hist(resid(clone\_survival\_analysis))

# Examine contrasts between treatments across all visits

pairs(emmeans(clone\_survival\_analysis, "treatment"))

# List of means and SE for treatment \* visit

emmeans(clone\_survival\_analysis, "visit", by = "treatment")

# Use the immediately previous emmeans output to make Fig. 1 in the manuscript (can use visit rather than date\_labels or user can manually enter date\_labels)

survival\_graph <- ggplot(data = emmeans\_clone\_survival\_analysis,

mapping = aes(x = date\_labels)) +

geom\_ribbon(mapping = aes(ymin = (CON\_lower.CL),

ymax = (CON\_upper.CL)),

alpha = 0.2) +

geom\_line(mapping = aes(y = CON\_emmean),

color = 'black',

linewidth = 1.5) +

geom\_line(mapping = aes(y = FL\_emmean),

color = 'royalblue',

linewidth = 1.5) +

geom\_line(mapping = aes(y = EX\_emmean),

color = 'violet',

linewidth = 1.5) +

geom\_text(data = emmeans\_clone\_survival\_analysis %>% filter(visit == 25),

mapping = aes(x = date\_labels,

y = CON\_emmean,

label = "Control"),

color = 'black',

size = 6,

hjust = 0,

vjust = -0.3) +

geom\_text(data = emmeans\_clone\_survival\_analysis %>% filter(visit == 25),

mapping = aes(x = date\_labels,

y = FL\_emmean,

label = "Floral clipping"),

color = 'royalblue',

size = 6,

hjust = 1.3,

vjust = -0.3) +

geom\_text(data = emmeans\_clone\_survival\_analysis %>% filter(visit == 25),

mapping = aes(x = date\_labels,

y = EX\_emmean,

label = "Stem clipping"),

color = 'violet',

size = 6,

hjust = 1.3,

vjust = -1.1) +

labs( x = "Date",

y = "Common milkweed stem survival") +

coord\_cartesian(ylim = c(0, 1)) +

theme\_light() +

theme(axis.text=element\_text(size=12),

axis.title=element\_text(size=14,face="bold"))

# Print the graph using abbreviated month/Year(capital letter means no abbreviation)

survival\_graph + scale\_x\_date(date\_breaks = "1 month", date\_labels = "%B")

####### Next examine the number of monarch eggs per stem, averaged per clone (Fig. 2 in manuscript)

# Complete a repeated-measures mixed-effect Analysis of Variance (i.e. linear mixed-effect analysis)

clone\_eggs\_analysis <- lmer(formula = AVGeggs ~ treatment \* visit + (1 | rep\_ID), data = AVG\_clones)

summary(clone\_eggs\_analysis)

anova(clone\_eggs\_analysis)

# Histogram of residuals of analysis

hist(resid(clone\_eggs\_analysis))

# Examine contrasts between treatments across all visits

pairs(emmeans(clone\_eggs\_analysis, "treatment"))

# List of means and SE for treatment \* visit

emmeans(clone\_eggs\_analysis, "visit", by = "treatment")

# Use the immediately previous emmeans output to make Fig. 2 in the manuscript (can use visit rather than date\_labels or user can manually enter date\_labels)

eggs\_stem0\_graph <- ggplot(data = emmeans\_clone\_eggs\_analysis,

mapping = aes(x = date\_labels)) +

geom\_ribbon(mapping = aes(ymin = (CON\_lower.CL),

ymax = (CON\_upper.CL)),

alpha = 0.2) +

geom\_line(mapping = aes(y = CON\_emmean),

color = 'black',

size = 1.5) +

geom\_line(mapping = aes(y = FL\_emmean),

color = 'royalblue',

size = 1.5) +

geom\_line(mapping = aes(y = EX\_emmean),

color = 'violet',

size = 1.5) +

geom\_text(data = emmeans\_clone\_eggs\_analysis %>% filter(visit == 21),

mapping = aes(x = date\_labels,

y = CON\_emmean,

label = "Control"),

color = 'black',

size = 6,

hjust = 0.12,

vjust = -0.8) +

geom\_text(data = emmeans\_clone\_eggs\_analysis %>% filter(visit == 24),

mapping = aes(x = date\_labels,

y = FL\_emmean,

label = "Floral clipping"),

color = 'royalblue',

size = 6,

hjust = -0.1,

vjust = 0.8) +

geom\_text(data = emmeans\_clone\_eggs\_analysis %>% filter(visit == 22),

mapping = aes(x = date\_labels,

y = EX\_emmean,

label = "Stem clipping"),

color = 'violet',

size = 6,

hjust = 0.5,

vjust = -0.3) +

labs( x = "Date",

y = "Monarch eggs per common milkweed stem") +

coord\_cartesian(ylim = c(0, 0.4)) +

theme\_light() +

theme(axis.text=element\_text(size=12),

axis.title=element\_text(size=14,face="bold"))

# Print the graph using abbreviated month/Year(capital letter means no abbreviation)

eggs\_stem0\_graph + scale\_x\_date(date\_breaks = "1 month", date\_labels = "%B")

####### Next examine the number of monarch eggs per stem, NOT averaged per clone and ignoring dead (i.e. senesced) stems (Figs. 3 and S3 in manuscript)

# Read and inspect the data for visit 3-21 (June3 - August5)

stems\_deadNA <- read.csv("Poyner\_Dickson\_all\_data\_June3\_Aug5\_senesced=NA.csv")

head(stems\_deadNA)

ggpairs(stems\_deadNA, mapping = aes(colour = treatment), columns = 4:9,

lower = list(continuous = "smooth"))

# Change variables that are numbers to factors (i.e. categorical variables), as appropriate

stems\_deadNA <- within(stems\_deadNA, {

visit <- factor(visit)

treatment\_num <- factor(treatment\_num)

rep\_ID <- factor(rep\_ID)

clone <- factor(clone)

stem <- factor(stem)

alive1\_dead0 <- factor(alive1\_dead0)

})

# Complete generalized linear mixed models using only living stems (note that a warning message appears that a

# large eigenvalue ratio is present and that stem height should be rescaled). We complete the analysis

# immediately below to create a graph of the raw stem heights (Fig. 3), but we calculate P-values for the analysis

# and create Fig. S3 from the generalized\_eggs\_per\_stem\_sc model below that scales stem height to a Z-score.

# Force the glmer to run more iterations to increase the likelihood of the model being identifiable by including the following:

# glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 1000000))

generalized\_eggs\_per\_stem <- glmer(formula = R\_eggs ~ visit + P\_height\_mm + (1 | rep\_ID),

data = stems\_deadNA, family = poisson(link = log),

control=glmerControl(optimizer="bobyqa", optCtrl = list(maxfun = 1000000)))

# Estimate the trendline for P\_height\_mm explanatory variable and add values to the data frame under new "pi.hat" variable

summary(generalized\_eggs\_per\_stem)

# The pi.hat predicted trendline uses the summary coefficients to solve the equation y = exp(intercept) \* exp(slope\*P\_height\_mm)

stems\_deadNA$pi.hat <- ((exp(-5.6588618)) \* (exp(0.0041263 \* stems\_deadNA$P\_height\_mm)))

# The minimum and maximum confidence intervals below subtract and add 1.96\*(SE of the slope) to the slope estimate

stems\_deadNA$pi.hat\_min <- ((exp(-5.6588618)) \* (exp((0.0041263 - (1.96\*0.0007445)) \* stems\_deadNA$P\_height\_mm)))

stems\_deadNA$pi.hat\_max <- ((exp(-5.6588618)) \* (exp((0.0041263 + (1.96\*0.0007445)) \* stems\_deadNA$P\_height\_mm)))

# The following code uses ggplot to graph data and predicted line for Fig. 3 (pi.hat calculated above)

height\_eggs\_graph <- ggplot(data = stems\_deadNA,

aes(x = P\_height\_mm, y = jitter(R\_eggs),

color = factor(treatment\_num))) +

geom\_ribbon(mapping = aes(ymin = (pi.hat\_min),

ymax = (pi.hat\_max)),

color = "black",

alpha = 0.2) +

geom\_line(mapping = aes(y = pi.hat),

color = "black",

linewidth = 1.25) +

geom\_point(show.legend = FALSE, size = 1.75) +

scale\_color\_manual(values = c("black", "royalblue", "violet")) +

labs( x = "Stem height (mm)",

y = "Monarch eggs per common milkweed stem") +

theme\_light() +

theme(axis.text=element\_text(size=12),

axis.title=element\_text(size=14,face="bold"))

height\_eggs\_graph

# We used methods from the following website to remove warning messages about very large eigenvalue in the generalized\_eggs\_per\_stem model:

# https://rstudio-pubs-static.s3.amazonaws.com/33653\_57fc7b8e5d484c909b615d8633c01d51.html

# To remove the very large eigenvalue warning message, the following code scales the predictor ("^P\\\_") variables

numcols <- grep("^P\\\_",names(stems\_deadNA))

stems\_deadNA\_scaled <- stems\_deadNA

stems\_deadNA\_scaled[,numcols] <- scale(stems\_deadNA\_scaled[,numcols])

# Force the glmer to run more iterations to increase the likelihood of the model being identifiable by including the following:

# glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 1000000))

generalized\_eggs\_per\_stem\_sc <- glmer(formula = R\_eggs ~ visit + treatment + P\_height\_mm + visit\*treatment + treatment\*P\_height\_mm + (1 | rep\_ID),

data = stems\_deadNA\_scaled, family = poisson(link = log),

control=glmerControl(optimizer="bobyqa", optCtrl = list(maxfun = 1000000)))

summary(generalized\_eggs\_per\_stem\_sc)

######### Now we're going to look for the model where we remove predictor variables from generalized\_eggs\_per\_stem\_sc to get the model with the lowest AIC

# The full model (generalized\_eggs\_per\_stem\_sc) has AIC=1134.6

# The only decreases in AIC occur by first removing the interactions (AIC=1100.2)

# then also removing "treatment" leads to the lowest AIC=1097.4 of any combination of fixed predictor variables (see Table 1 of manuscript).

generalized\_eggs\_per\_stem\_sc\_aic <- glmer(formula = R\_eggs ~ visit + P\_height\_mm + (1 | rep\_ID),

data = stems\_deadNA\_scaled, family = poisson(link = log),

control=glmerControl(optimizer="bobyqa", optCtrl = list(maxfun = 1000000)))

summary(generalized\_eggs\_per\_stem\_sc\_aic)

Anova(generalized\_eggs\_per\_stem\_sc\_aic, test.statistic="Chi")

# Now calculating the trendline for scaled stem height from the output of summary(generalized\_eggs\_per\_stem\_sc\_aic)

# The pi.hat predicted trendline uses the summary coefficients to solve the equation y = exp(intercept) \* exp(slope\*P\_height\_mm)

stems\_deadNA\_scaled$pi.hat <- ((exp(-3.30007)) \* (exp(0.72163 \* stems\_deadNA\_scaled$P\_height\_mm)))

# The minimum and maximum confidence intervals below subtract and add 1.96\*(SE of the slope) to the slope estimate

stems\_deadNA\_scaled$pi.hat\_min <- ((exp(-3.30007)) \* (exp((0.72163 - (1.96\*0.12617)) \* stems\_deadNA\_scaled$P\_height\_mm)))

stems\_deadNA\_scaled$pi.hat\_max <- ((exp(-3.30007)) \* (exp((0.72163 + (1.96\*0.12617)) \* stems\_deadNA\_scaled$P\_height\_mm)))

# The following code uses ggplot to graph data and predicted line for Fig. S3 (pi.hat calculated above)

# Note this is the graph for scaled stem height data based on the fixed predictor variables with

# the lowest model AIC (i.e. visit + P\_height\_mm)

height\_eggs\_graph\_scaled <- ggplot(data = stems\_deadNA\_scaled,

aes(x = P\_height\_mm, y = jitter(R\_eggs),

color = factor(treatment\_num))) +

geom\_ribbon(mapping = aes(ymin = (pi.hat\_min),

ymax = (pi.hat\_max)),

color = "black",

alpha = 0.2) +

geom\_line(mapping = aes(y = pi.hat),

color = "black",

linewidth = 1.25) +

geom\_point(show.legend = FALSE, size = 1.75) +

scale\_color\_manual(values = c("black", "royalblue", "violet")) +

labs( x = "Stem height (scaled Z-score)",

y = "Monarch eggs per common milkweed stem") +

theme\_light() +

theme(axis.text=element\_text(size=12),

axis.title=element\_text(size=14,face="bold"))

height\_eggs\_graph\_scaled

####### Next examine the average height of stems per clone, ignoring dead (i.e. senesced) stems (Fig. 4 in manuscript)

# First create a new dataframe that is mean height per stem within each clone at each visit from June3-August5

clones\_deadNA <- stems\_deadNA %>% group\_by(visit,treatment,rep\_ID) %>% summarise(AVG\_height\_mm = mean(P\_height\_mm,na.rm=T)) #%>% merge(siteyear)

head(clones\_deadNA)

# Complete a repeated-measures mixed-effect Analysis of Variance for the average height per clone (i.e. linear mixed-effect analysis)

clone\_height\_analysis <- lmer(formula = AVG\_height\_mm ~ visit \* treatment + (1 | rep\_ID), data = clones\_deadNA)

summary(clone\_height\_analysis)

anova(clone\_height\_analysis)

# Histogram of residuals of analysis

hist(resid(clone\_height\_analysis))

# Examine contrasts between treatments across all visits

pairs(emmeans(clone\_height\_analysis, "treatment"))

# List of means and SE for treatment \* visit

emmeans(clone\_height\_analysis, "visit", by = "treatment")

# Use the immediately previous emmeans output to make Fig. 4 in the manuscript (can use visit rather than date\_labels or user can manually enter date\_labels)

height\_stemNA\_graph <- ggplot(data = emmeans\_height\_deadNA,

mapping = aes(x = date\_labels)) +

geom\_ribbon(mapping = aes(ymin = (CON\_lower.CL),

ymax = (CON\_upper.CL)),

alpha = 0.2) +

geom\_line(mapping = aes(y = CON\_emmean),

color = 'black',

linewidth = 1.5) +

geom\_line(mapping = aes(y = FL\_emmean),

color = 'royalblue',

linewidth = 1.5) +

geom\_line(mapping = aes(y = EX\_emmean),

color = 'violet',

linewidth = 1.5) +

geom\_text(data = emmeans\_height\_deadNA %>% filter(visit == 21),

mapping = aes(x = date\_labels,

y = CON\_emmean,

label = "Control"),

color = 'black',

size = 6,

hjust = 0,

vjust = -0.6) +

geom\_text(data = emmeans\_height\_deadNA %>% filter(visit == 21),

mapping = aes(x = date\_labels,

y = FL\_emmean,

label = "Floral clipping"),

color = 'royalblue',

size = 6,

hjust = 0,

vjust = 0.6) +

geom\_text(data = emmeans\_height\_deadNA %>% filter(visit == 21),

mapping = aes(x = date\_labels,

y = EX\_emmean,

label = "Stem clipping"),

color = 'violet',

size = 6,

hjust = 0,

vjust = 0) +

labs( x = "Visit",

y = "Milkweed height with senesced stems ignored (mm)") +

coord\_cartesian(ylim = c(0, 700)) +

theme\_light() +

theme(axis.text=element\_text(size=12),

axis.title=element\_text(size=14,face="bold"))

# Print the graph using abbreviated month/Year(capital letter means no abbreviation)

height\_stemNA\_graph + scale\_x\_date(date\_breaks = "1 month", date\_labels = "%B")