

The physiological mechanisms and evolutionary history of diapause strategies in butterflies

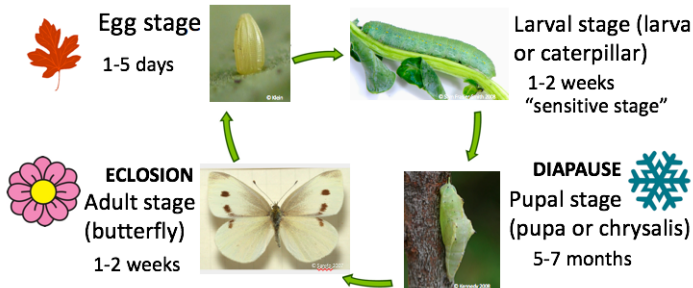
Emily Mikucki

Computational Biology



What is diapause?

- ▶ State of physiological dormancy (overwintering)
- ▶ Developmental arrest
- ▶ Depressed metabolic rate
- ▶ Heightened stress and cold tolerance



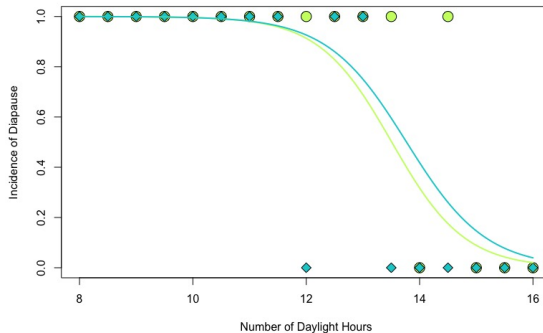
My main research objectives

- ▶ To determine the diapause induction and termination patterns of *Pieris rapae* butterflies under current climatic conditions and to predict potential physiological consequences of winter warming on diapausing individuals
- ▶ To understand the molecular bases of the cold tolerance mechanisms associated with diapausing individuals from different populations
- ▶ To reveal the evolutionary history of diapause strategies across Lepidopterans



The environmental cues associated with diapause induction

- ▶ Photoperiod differentially induces diapause at different latitudes
- ▶ Potential interaction between photoperiod and diapause induction in a Vermont population and a North Carolina population of *Pieris rapae* butterflies







R code for binomial logistic regression

```
#Binomial Logistic Regression
photodata2 <- read.csv("Photoperiod_LogReg2.csv", header = TRUE, sep = ",")
head(photodata2)
# data
xVar <- photodata2$Photoperiod
yVar <- photodata2$Incidence
dataFrame <- data.frame(xVar[1:17],yVar[1:17])
dataFrame1 <- data.frame(xVar[18:34],yVar[18:34])
# models for VT and NC
logRegMod <- glm(yVar ~ xVar,
                 data=dataFrame,
                 family=binomial(link="logit"))
logRegMod1 <- glm(yVar ~ xVar,
                 data=dataFrame1,
                 family=binomial(link="logit"))

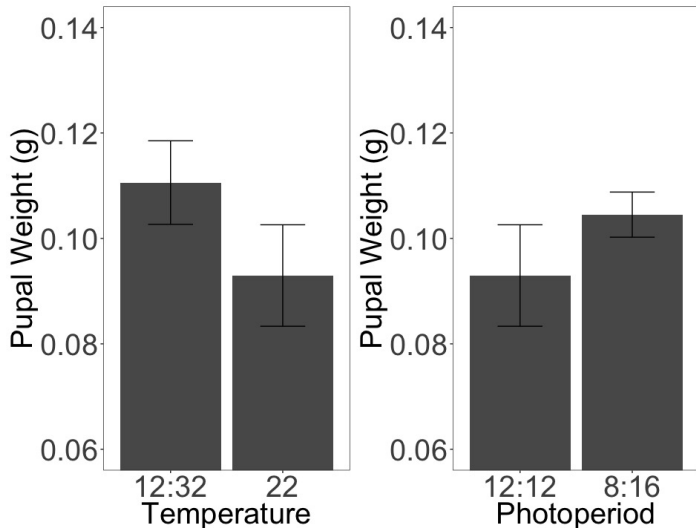
# model outputs
print(logRegMod)
summary(logRegMod)
print(logRegMod1)
summary(logRegMod1)
#Using subset points
VTpoints <- photodata2$Incidence[1:17]
NCpoints <- photodata2$Incidence[18:34]
#Plot with 2 logsitic regression curves
plot(x=xVar, y=yVar, type="n",
     xlab="Number of Daylight Hours", ylab="Incidence of Diapause")
points(x=photodata2$Photoperiod[1:17], y=VTpoints, cex=2, pch=21,bg="darkolivegreen1")
points(x=photodata2$Photoperiod[18:34], y=NCpoints, cex=1.5, pch=23,bg="darkturquoise" )
curve(predict(logRegMod,data.frame(xVar=x),type="response"),add=TRUE,lwd=2, col="darkolivegreen1")
curve(predict(logRegMod1,data.frame(xVar=x),type="response"),add=TRUE,lwd=2, col="darkturquoise")
.frame(xVar=x),type="response"),add=TRUE,lwd=2, col="darkturquoise")
```

Physiological consequences of increased winter warming in *P. rapae* butterflies

- ▶ 1-week +5°C warming event in diapausing pupae
- ▶ Used weight (in g) as a proxy for running through stored lipid reserves

Treatment	14L:10D	12L:12D		8L:16D
Original temp.	25°C	22°C	12°C-32°C	22°C
Experimental temp. (+5°C)	 30°C		 27°C	 22°C (CONTROL)
Number of diapausing pupae (N≈10)				

Temperature affects pupal weight ($p=0.03$), but not photoperiod ($p>0.05$)



R code for ANOVA and ggplots

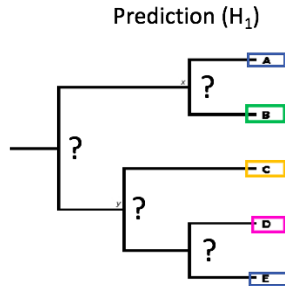
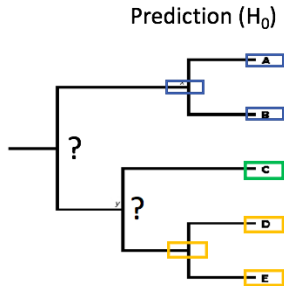
```
library(lme4)
library(lmerTest)
library(ggplot2)
library(Rmisc)
library(gridExtra)
library(grid)
library(lattice)
#ANOVA models
aovweight <- aov(Pupal_Weight ~ Time_Point*Photoperiod*Temperature +
                 Error(Time_Point/Individual_ID), data = Weight)
summary(aovweight)
aovweight <- aov(Pupal_Weight ~ Photoperiod+Temperature+Time_Point +
                 Error(Time_Point/Individual_ID), data = Weight)
summary(aovweight)
aovweight <- aov(Pupal_Weight ~ Photoperiod+Temperature+Time_Point +
                 Individual_ID, data = Weight)
summary(aovweight)
mod1<-aov(Pupal_Weight~Temperature+Time_Point+Error(Time_Point/Individual_ID),data=w2)
summary(mod1)

mod3<-lmer(Pupal_Weight~Temperature+Time_Point+(1|Individual_ID),data=w2)
#grid.arrange = par for ggplot
a <- ggplot(data=tgc,aes(x=Temperature,y=Pupal_Weight))+
  geom_bar(stat="identity")+geom_errorbar(aes(ymin=Pupal_Weight-se,ymax=Pupal_Weight+se),
                                           width=.4,line=3)+coord_cartesian(ylim=c(.06,.14))+
  T+ylabel("Pupal Weight (g)")

b <- ggplot(data=tgc2,aes(x=Photoperiod,y=Pupal_Weight))+
  geom_bar(stat="identity")+geom_errorbar(aes(ymin=Pupal_Weight-se,ymax=Pupal_Weight+se),
                                           width=.4,line=3)+coord_cartesian(ylim=c(.06,.14))+
  T+ylabel("Pupal Weight (g)")
grid.arrange(a, b, ncol=2)
```


Phylogenetic predictions for evolutionary history of diapause in butterflies

- ▶ H0: Phylogenetic inertia \rightarrow diapause strategy
- ▶ Predict closely related species would have the same strategy
- ▶ H1: Environmental selection \rightarrow diversification of diapause strategy
- ▶ Predict species will have different strategies influenced by their life history traits



Using a phylogenetic subset of a large-scale, comprehensive phylogenetic study of Lepidopterans (Regier et al. 2013)

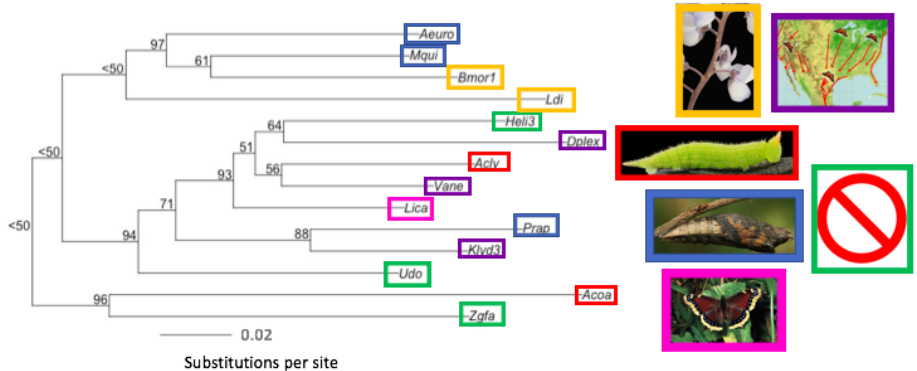
Phylogenetic study:

- 483 species of Lepidopterans
- 19 nuclear protein coding genes
- Open-source data

Determine life history traits:

- Diapause strategy
- Hostplant
- Geographic range
- Number of generations per year (voltinism)

Phylogenetic subset of Regier et al. (2013) with known diapause strategy



R code for phylogenetic tree

```
#install.packages("phytools")
library(phytools)
## phylogeny manipulation
tree1 <- read.tree("phylogenetic_tree_03272017") #exported fig tree files as newick file
tree1$tip.label
tiplabels<-read.csv("phylo_tip_manipulation.csv",header=FALSE)
tree1$tip.label<-as.character(tiplabels[,1])
### data set manipulation
speciesdata<-read.csv("Processed_data_with_tiplabels_03282017.csv",header=TRUE)
str(speciesdata)
speciesdata$Phylo_tip
dim(subset(speciesdata, !is.na(Adult)))
keepdat<-subset(speciesdata, !is.na(Adult))
dim(subset(speciesdata, is.na(Adult))) # want this list to drop the tips
dim(speciesdata)
class(speciesdata$Genus)
speciesdata$Genus<-as.character(speciesdata$Genus)
speciesdata$species<-as.character(speciesdata$species)
dropdat<-subset(speciesdata, is.na(Adult))$Phylo_tip
dim(setdiff(tree1$tip.label,keepdat$Gspec))
bootstraps<-c("<50", "96", "<50", "94", "71", "88", "93",
              "51", "56", "64", "<50", "97", "61") # adding bootstraps
par(mar=c(5,5,2,2))
plot(tree2)
nodelabels(bootstraps,bg="white",frame="none",adj=c(1.1,-0.2))
add.scale.bar()
```

Questions?

Polygonia interrogationis = the question mark butterfly

