egg project

Janna M

8/23/2021

This paper examines the connection between egg mass and avian flight. Past studies have found a link between reproductive strategies in extant non-avian theropod dinosaurs.There have been differences found in the eggshell mass proportions of crown birds and dinosaurs. The differences may be linked to flight but denser sampling was needed . This paper fills that gap by examining the egg shell thickness of 114 dinosaurs and birds and comparing those differences to flight and other life history correlations .

In my data replication i will be generating a phylogenetic tree, by reading in a nexus file and using the plot tree function. ##Plot Tree #plotTree(tree)

I will also be performing a path analysis to assess i bird eggs and reptile eggs are closely related by using the Wilcox signed-rank test to get the P-value . #wilcox.test(log(EggMassMbird), paired=T)

I performed a semi path analysis to describe the relationship between body mass and Egg mass and body mass and eggshell thickness, using ggplot. Results show that egg mass and eggshell thickness are strongly correlated and that the correlation is much stronger in flying taxa.

#ggplot(datadino, aes(log(Egg\_mass), log(Eggshell\_thickness), color=Flight)) + geom\_point(size=2) +

### Loading packages

* I began by downloading all of the library that the author used . I was unable to download the library“(ggtree)” at first.However, i ended up using ggplot so i didn’t need ggtree.

library(nlme)  
library(ape)  
library(ggplot2)  
library(phytools)

## Loading required package: maps

library(RColorBrewer)  
library(plyr)

##   
## Attaching package: 'plyr'

## The following object is masked from 'package:maps':  
##   
## ozone

library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:plyr':  
##   
## arrange, count, desc, failwith, id, mutate, rename, summarise,  
## summarize

## The following object is masked from 'package:nlme':  
##   
## collapse

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

library(evobiR)  
library(AICcmodavg)

##   
## Attaching package: 'AICcmodavg'

## The following object is masked from 'package:evobiR':  
##   
## AICc

library(RRphylo)

## Loading required package: emmeans

##   
## Attaching package: 'RRphylo'

## The following object is masked from 'package:phytools':  
##   
## node.paths

library(car)

## Loading required package: carData

##   
## Attaching package: 'car'

## The following object is masked from 'package:dplyr':  
##   
## recode

library(hrbrthemes)

## NOTE: Either Arial Narrow or Roboto Condensed fonts are required to use these themes.

## Please use hrbrthemes::import\_roboto\_condensed() to install Roboto Condensed and

## if Arial Narrow is not on your system, please see https://bit.ly/arialnarrow

library(phylopath)  
library(caper)

## Loading required package: MASS

##   
## Attaching package: 'MASS'

## The following object is masked from 'package:dplyr':  
##   
## select

## Loading required package: mvtnorm

library(emmeans)  
library(GGally)

## Registered S3 method overwritten by 'GGally':  
## method from   
## +.gg ggplot2

##   
## Attaching package: 'GGally'

## The following object is masked from 'package:emmeans':  
##   
## pigs

I download a more recent version of R then Ran the following code to try to download ggtree

#if(!requireNamespace(“BioManager”,quietly=TRUE)) #install.packages(“BioManager”) #BioManager::install(“ggtree”)

However I could not download the package BioManager, but i think ggplot was installed because i was able to use it .

Next, i tried to load the tree. I’ve done this before in my own research , but at first it would not run.Once I re-imported the data set, it ran.

# Load tree   
library(readr)  
treewholenew\_trees <- read\_csv("~/egg project/dinosaureggshellproject-main/treewholenew.trees.nex")

## Rows: 213 Columns: 1

## -- Column specification --------------------------------------------------------  
## Delimiter: ","  
## chr (1): #NEXUS

##   
## i Use `spec()` to retrieve the full column specification for this data.  
## i Specify the column types or set `show\_col\_types = FALSE` to quiet this message.

View(treewholenew\_trees)

## Warning: One or more parsing issues, see `problems()` for details

show\_col\_types = FALSE

# Read tree using Nexus file   
library(ape)  
library(phytools)  
library(tidyverse)

## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --

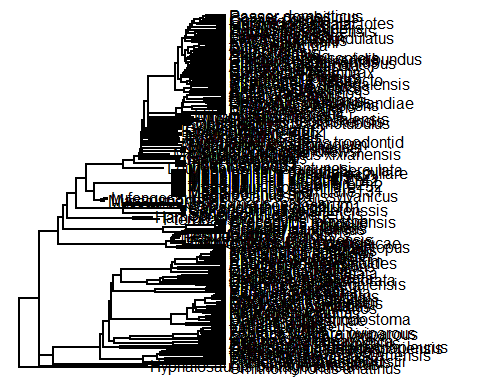
## v tibble 3.1.4 v stringr 1.4.0  
## v tidyr 1.1.3 v forcats 0.5.1  
## v purrr 0.3.4

## -- Conflicts ------------------------------------------ tidyverse\_conflicts() --  
## x dplyr::arrange() masks plyr::arrange()  
## x dplyr::collapse() masks nlme::collapse()  
## x purrr::compact() masks plyr::compact()  
## x dplyr::count() masks plyr::count()  
## x dplyr::failwith() masks plyr::failwith()  
## x dplyr::filter() masks stats::filter()  
## x dplyr::id() masks plyr::id()  
## x dplyr::lag() masks stats::lag()  
## x purrr::map() masks maps::map()  
## x dplyr::mutate() masks plyr::mutate()  
## x car::recode() masks dplyr::recode()  
## x dplyr::rename() masks plyr::rename()  
## x MASS::select() masks dplyr::select()  
## x purrr::some() masks car::some()  
## x dplyr::summarise() masks plyr::summarise()  
## x dplyr::summarize() masks plyr::summarize()

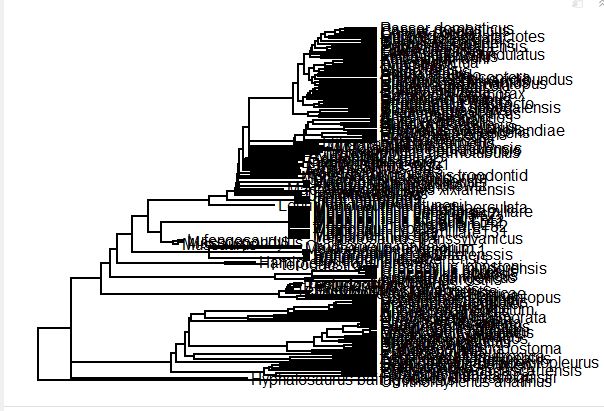
#tree and data  
tree<-read.nexus("~/egg project/dinosaureggshellproject-main/treewholenew.trees.nex")

Next I plotted the tree.

#Plot Tree   
plotTree(tree)



#Authors graph  
knitr::include\_graphics("tree pic.JPG")

 Then i imported the rest of the datasets i found online into the environment manually.

Next , I read in the data table that contains the information for the clade, taxon,eggshell thickness, egg\_mass, body mass,flight,nesting site, Nest type,and precociality.

Then i subset the data because several values for egg mass were missing.

data<-read.table("Datawhole.txt", header=T, stringsAsFactors=T)  
  
mydata<-subset(data, !is.na(Egg\_mass))

Next, i assign the data egg mass equation as egg mass. I calculated the delta between the EggMassMbird. I also calculated the P-value.

EggMass$Mreptile- is the egg of a non-avian reptile,using the equation of Deeming and Ferguson, 1990.

EggMass$Mbird- is the estimated egg mass of a bird , using the equation of Hoyt, 1979.

Delta is the change in the values .

The two equations are similar, but the allometric exponent for for birds: M = 0.00056 \* L \* W^2 and for non-avian reptiles the allometic exponent is : M = 0.000548 \* L \* W^2

#Load and assign egg mass   
  
EggMass <-read.table("Data\_Eggmass\_equations.txt", header=T, stringsAsFactors=T)  
  
  
# Find the change in value or delta   
delta<-log(EggMass$Mreptile) - log(EggMass$Mbird)  
show(delta)

## [1] 0.0216615 0.0216615 0.0216615 0.0216615 0.0216615 0.0216615 0.0216615  
## [8] 0.0216615 0.0216615 0.0216615 0.0216615 0.0216615 0.0216615 0.0216615  
## [15] 0.0216615 0.0216615 0.0216615 0.0216615 0.0216615 0.0216615 0.0216615  
## [22] 0.0216615 0.0216615 0.0216615 0.0216615 0.0216615 0.0216615 0.0216615  
## [29] 0.0216615 0.0216615 0.0216615 0.0216615 0.0216615 0.0216615 0.0216615  
## [36] 0.0216615 0.0216615 0.0216615 0.0216615 0.0216615 0.0216615 0.0216615  
## [43] 0.0216615 0.0216615 0.0216615 0.0216615 0.0216615 0.0216615 0.0216615  
## [50] 0.0216615 0.0216615 0.0216615 0.0216615 0.0216615 0.0216615 0.0216615  
## [57] 0.0216615 0.0216615 0.0216615 0.0216615 0.0216615 0.0216615 0.0216615

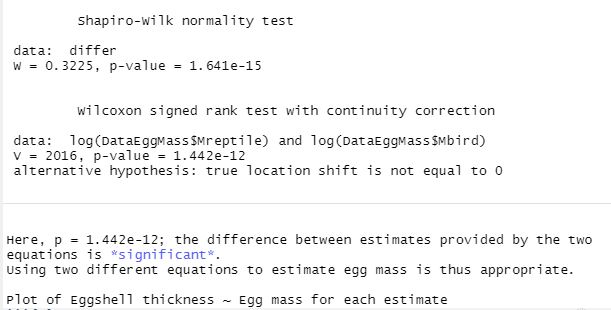
# Wilcoxon signed-rank test to get the P-value   
wilcox.test(log(EggMass$Mreptile), log(EggMass$Mbird), paired=T)

##   
## Wilcoxon signed rank test with continuity correction  
##   
## data: log(EggMass$Mreptile) and log(EggMass$Mbird)  
## V = 2016, p-value = 1.442e-12  
## alternative hypothesis: true location shift is not equal to 0

# test for body mass vs egg mass  
wilcox.test(log(EggMass$Mbird), log(EggMass$Mreptile), paired=T)

##   
## Wilcoxon signed rank test with continuity correction  
##   
## data: log(EggMass$Mbird) and log(EggMass$Mreptile)  
## V = 0, p-value = 1.442e-12  
## alternative hypothesis: true location shift is not equal to 0

#Authors wilcox test   
knitr::include\_graphics("wilcox correct.JPG")

 The results show that the reptile and bird eggs have a high correlation and are statically significant based on the P values.

The Equations are so similar that using one or the other does not influence the result and the delta value shows this.

The author and I got the same P-Values.

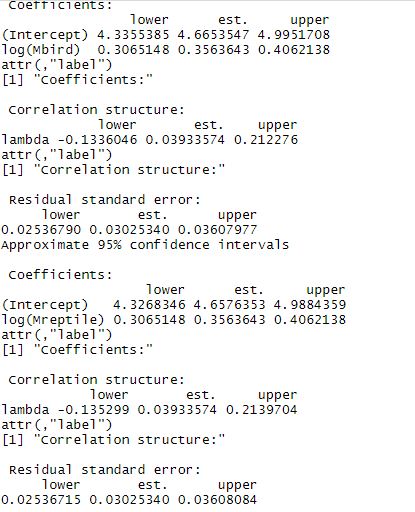
Next i performed a regression describing the relationship between egg mass and eggshell thickness.

#Regression

datadino<-data%>%filter(Clade=="Neognaths"|Clade=="Non-avian\_dinosaurs"|Clade=="Paleognaths"|Clade== "Enantiornithes"|Clade=="Ornithuromorphs")  
regressL<- lm(Eggshell\_thickness~Egg\_mass,data = datadino)  
summary(regressL)

##   
## Call:  
## lm(formula = Eggshell\_thickness ~ Egg\_mass, data = datadino)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1379.5 -322.4 -182.3 280.5 1535.5   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 474.10192 53.64858 8.837 1.58e-14 \*\*\*  
## Egg\_mass 0.38669 0.02939 13.159 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 498.2 on 112 degrees of freedom  
## (1 observation deleted due to missingness)  
## Multiple R-squared: 0.6072, Adjusted R-squared: 0.6037   
## F-statistic: 173.2 on 1 and 112 DF, p-value: < 2.2e-16

#Authors Regression   
knitr::include\_graphics("Author reg.JPG")

 The R- squared value shows a that there is a relationship between but not a high correlation but the the relationship statistically significant

My values and the authors values were not the same

Next i ran regression on egg mass and eggshell thickness to body mass.

regress1<- lm(Egg\_mass+Eggshell\_thickness~Body\_mass,data = data)  
summary(regress1)

##   
## Call:  
## lm(formula = Egg\_mass + Eggshell\_thickness ~ Body\_mass, data = data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -801.5 -136.9 -54.6 45.1 745.7   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 2.541e+02 3.428e+01 7.412 7.97e-10 \*\*\*  
## Body\_mass 3.505e-02 6.421e-04 54.589 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 249 on 55 degrees of freedom  
## (117 observations deleted due to missingness)  
## Multiple R-squared: 0.9819, Adjusted R-squared: 0.9815   
## F-statistic: 2980 on 1 and 55 DF, p-value: < 2.2e-16

Egg mass and eggshell thickness are highly correlated based on the R-Squared value. For every one gram of egg mass there is a 28.216 unit increase .For ever micrometer of eggshell thickness there is a 27.546 unit increase

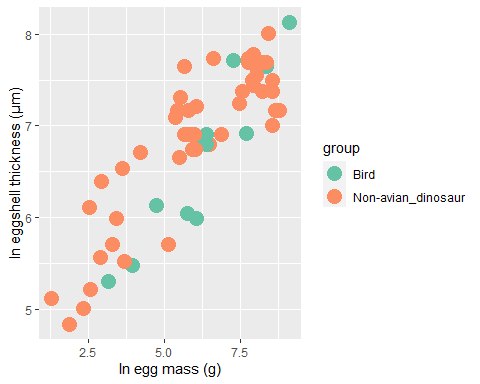
#Data Visualization  
I used ggplot for this graph that represents Ln of egg mass vs Ln of eggshell thickness. I was under the impression it did not load but it did .

Read in table for equations

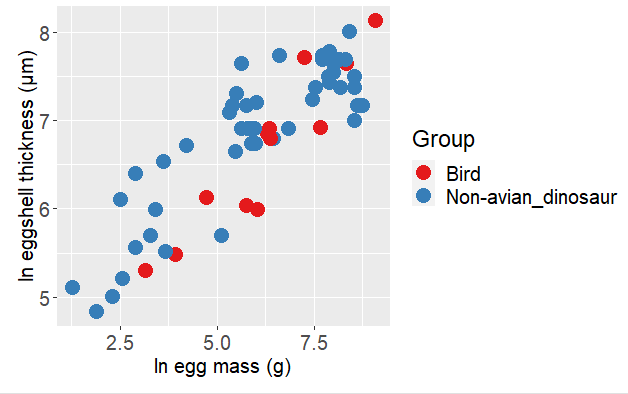
EggMass<-read.table ("Data\_Eggmass\_equations.txt", header=TRUE)

Plot egg mass vs eggshell thickness

ggplot(EggMass, aes(log(Mreptile), log(Eggshell\_thickness), color=Group)) +  
geom\_point(size=5) +  
xlab("ln egg mass (g)") +  
ylab("ln eggshell thickness (µm)") +  
scale\_colour\_brewer("group", palette="Set2")



#Authors graph  
knitr::include\_graphics("pic2.png")



Plot body mass vs egg mass

I used ggplot again to plot body mass vs egg mass. At first the graph looked weird because many of the data was missing body mass.

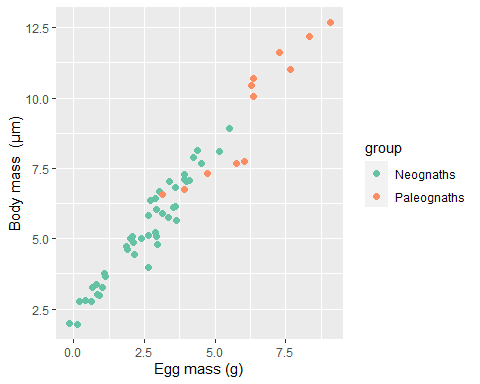
Egg mass was missing from many species i used code to remove missing data for body mass then i re-ran the code and i got a graph similar to that of the author.

datamass<-subset(data, !is.na(data$Body\_mass))

The two graphs look similar

Next i wanted to examine the relationship between Egg mass and body mass. (Extra)

ggplot(datamass, aes(log(Egg\_mass), log(Body\_mass), color=Clade)) + geom\_point(size=2) +  
xlab("Egg mass (g)") +  
ylab("Body mass (µm)") +  
scale\_colour\_brewer("group", palette="Set2")

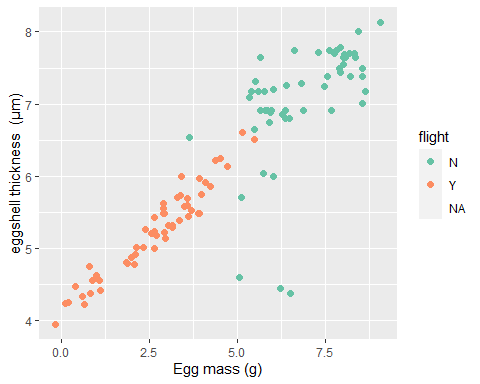


Phylogenetic path analysis is an extremely long process to go through it is both time and computationally intensive. It also requires knowledge that has nothing to do with statistics and more to do with phylogenetics. What I wanted to do Is illustrate the author’s main point which is that there is a difference in the relationship between eggshell thickness and egg mass in taxa that fly and taxa that do not .

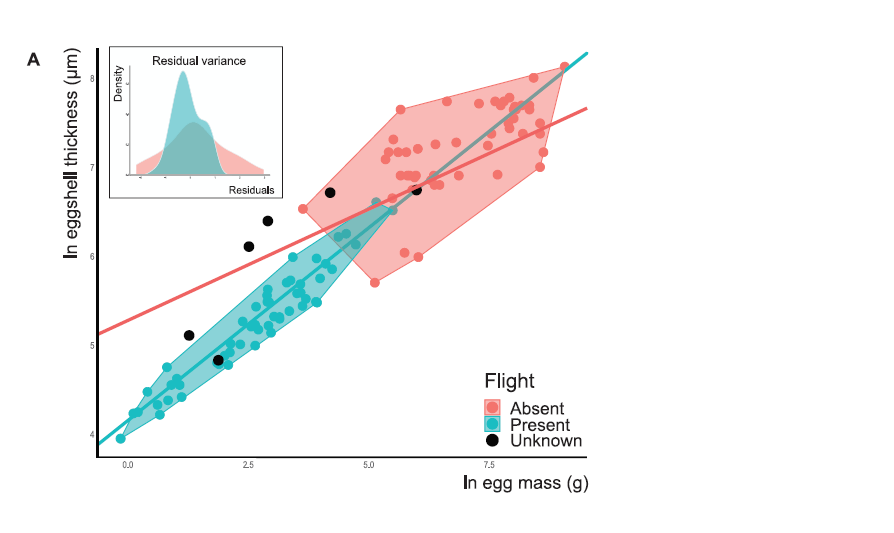
This is my attempt to recreate a path analysis graph I was unable to get the residual variance graph to display on the graph.

ggplot(datadino, aes(log(Egg\_mass), log(Eggshell\_thickness), color=Flight)) + geom\_point(size=2) +  
xlab("Egg mass (g)") +  
ylab("eggshell thickness (µm)") +  
scale\_colour\_brewer("flight", palette="Set2")

## Warning: Removed 7 rows containing missing values (geom\_point).



#Authors graph  
knitr::include\_graphics("final.png")

 ``` Egg mass is a good proxy for body mass.

Non flyers have a higher egg mass and eggshell thickness.

Results show that egg mass and eggshell thickness are strongly correlated and that the correlation is much stronger in flying taxa.