

# Sexual Dichromatism is associated with the slow life histories in passerine birds

Amy Duclaux

2025-05-08

This script follows the analysis from Duclaux et al testing the relationship between life history traits and Sexual Dichromatism in passerine birds while accounting for potential confounding issues with latitude and body mass.

## Upload packages

To complete this analysis we will need to use the standard `caper`, `phytools` and `TreeTools` packages for handling and plotting the phylogeny and the `MCMCglmm` to run the Bayesian phylogenetic comparative models. As we also want to account for uncertainty in the phylogeny itself so we will use the `mulTree` package to run the each model over 100 randomly drawn trees. We can download `mulTree` from github here (<https://github.com/TGuillerme/mulTree>).

```
library(caper)

library(phytools)

library(MCMCglmm)

library(TreeTools)

if(!require(devtools)) install.packages("devtools")
library(devtools)
install_github("TGuillerme/mulTree", ref = "release")
library(mulTree)
```

## Upload Data

To run the analysis we will use data on the colour intensity fo passerines from Cooney et. al 2022 (<https://www.nature.com/articles/s41559-022-01714-1>), Bird et al 2020 (<https://conbio.onlinelibrary.wiley.com/doi/abs/10.1111/cobi.13486>) and from McQueen et al. 2019 (<https://onlinelibrary.wiley.com/doi/abs/10.1111/ele.13375>).

```
bird_data <- read.csv("ADuclaux birds dichro lat fast-slow mating sys 1may25.csv")
```

For the phylogeny we use Jetz et al. 2012 (<https://www.nature.com/articles/nature11631>)

```
phylo <- read.tree("bird_phylo.tre")
class(phylo)
```

```
## [1] "multiPhylo"
```

We will log body mass, maximum lifespan and age at first reproduction and set latitude to absolute values.

```
dichro_bird_data <- data.frame(animal= bird_data$animal,
                              latin_name= bird_data$latin_name,
                              order= bird_data$order,
                              diff_LociVS= bird_data$diff_LociVS,
                              midpoint_lat= abs(bird_data$midpoint_lat),
                              log_mass= log10(bird_data$mass),
                              survival= bird_data$survival,
                              log_max_lifespan= log10(bird_data$max_lifespan),
                              log_age_first_breed= log10(bird_data$age_first_breed),
                              mating_system= as.factor(bird_data$mating_system_binary))
```

```
dichro_bird_data <- unique(dichro_bird_data)
dichro_bird_data <- na.omit(dichro_bird_data)
```

```
head(dichro_bird_data)
```

```
##           animal          latin_name      order diff_LociVS
## 1  Abroscopus_albogularis  Abroscopus_albogularis Passeriformes  0.2088049
## 5  Abroscopus_superciliaris Abroscopus_superciliaris Passeriformes -1.4542725
## 7  Acanthagenys_rufogularis Acanthagenys_rufogularis Passeriformes -8.1521900
## 9   Acanthisitta_chloris   Acanthisitta_chloris Passeriformes  27.1174346
## 11  Acanthiza_apicalis     Acanthiza_apicalis Passeriformes -3.4947721
## 13  Acanthiza_chrysorrhoa  Acanthiza_chrysorrhoa Passeriformes -0.4284985
## midpoint_lat log_mass survival log_max_lifespan log_age_first_breed
## 1      30.31612 1.576915 0.4967672      0.7858776      -0.088489893
## 5      12.84058 1.868721 0.5087253      0.7984239      -0.046711386
## 7      33.00662 3.867026 0.7179357      1.1326005       0.149555647
## 9      48.85290 1.932970 0.5215184      0.8171323      -0.058030884
## 11     33.69648 2.030776 0.6767431      1.2279219       0.012054488
## 13     35.19802 2.228939 0.6483806      1.1174075      -0.004625051
## mating_system
## 1           0
## 5           0
## 7           0
## 9           0
## 11          0
## 13          0
```

Create a multiphylo object of 100 phylogenies which are matched to our dataset.

```
#Create a dataset of just the names to match
match_data <- dichro_bird_data[,c("latin_name", "order")]

comp_phylo_birds <- list()

for (i in 1:100){

  comp_phylo_birds[[i]] <- comparative.data(phy = phylo[[i]],
                                           data = match_data,
                                           names.col = latin_name,
                                           vcv= TRUE)$phy}

class(comp_phylo_birds)
```

```
## [1] "list"
```

```
#[1] "list"

multiphylo_birds <- as.multiPhylo(comp_phylo_birds)
class(multiphylo_birds)
```

```
## [1] "multiPhylo"
```

```
#[1] "multiPhylo"
```

We can now create a mulTree object that contains the 100 phylogenies and the data

```
multree_birds <- as.mulTree(data= dichro_bird_data,
                           tree= multiphylo_birds,
                           taxa= "latin_name")
```

We can now set the number of iterations, the thinning and the burnin we will use for the models. We will also set the non-informative prior we will use throughout for the models.

```
parameters <- c(120000, 50, 20000)

priors <- list(R= list(V= 1, nu= 0.002),
              G = list(G1=list(V= 1,n= 2, nu= 0.002)))
```

## Models

In order to define the slow-fast continuum we run a model for each of the life history traits of maximum lifespan, survival and age at first reproduction as response variables and mass and latitude as explanatory variables. We then use the residuals from these models in a PCA in order to define slow fast continuum.

### Survival model

First we run the survival model, due to the long run time of this model we will not evaluate it here but read in a previous run.

```
#Define the model
survival_resids_formula <- survival ~ log_mass + midpoint_lat

mulTree(mulTree.data = multree_birds,
        formula = survival_resids_formula,
        priors = priors,
        parameters = parameters,
        output = "bird_survival_resids",
        ESS= 1000,
        chains= 2)
```

Read in the survival models and get the overall summary

```
multree_surv_res <- read.mulTree("bird_survival_resids")

summary(multree_surv_res)
```

```
##              Estimates(mode hdr) lower.CI(2.5) lower.CI(25)
## (Intercept)          0.502827990    0.442992008    0.481555803
## log_mass             0.061419134    0.056622645    0.059767231
## midpoint_lat        -0.001587213   -0.001735308   -0.001639999
## phylogenetic.variance 0.005999754    0.005021444    0.005656023
## residual.variance    0.001216694    0.001073160    0.001166624
```

```
##               upper.CI(75) upper.CI(97.5)
## (Intercept)      0.521629715    0.560429337
## log_mass         0.063071036    0.066154587
## midpoint_lat     -0.001530028   -0.001434720
## phylogenetic.variance 0.006343484    0.007083827
## residual.variance  0.001266765    0.001366905
## attr("class")
## [1] "matrix" "mulTree"
```

Check if all the chains converged, as indicated by a convergence value of less than 1.1

```
multree_surv_conv <- read.mulTree("bird_survival_resids", convergence= TRUE)
all(unlist(multree_surv_conv) < 1.1)
```

```
## [1] TRUE
```

*#calculate the residuals*

```
surv_resids <- multree_birds$data$survival -
  c(summary(multree_surv_res)[1]
    + multree_birds$data$log_mass*summary(multree_surv_res)[2]
    + multree_birds$data$midpoint_lat*summary(multree_surv_res)[3])
```

## Max lifespan

Next we run the max lifespan model, due to the long run time of this model we will not evaluate it here but read in a previous run.

```
max_lifespan_resids_formula <- log_max_lifespan ~ log_mass + midpoint_lat
```

```
mulTree(mulTree.data = multree_birds,
  formula = max_lifespan_resids_formula,
  priors = priors,
  parameters = parameters,
  output = "bird_max_lifespan_resids",
  ESS= 1000,
  chains= 2)
```

Read in the max lifespan model

```
multree_max_lifespan_res <- read.mulTree("bird_max_lifespan_resids")
summary(multree_max_lifespan_res)
```

```
##               Estimates(mode hdr) lower.CI(2.5) lower.CI(25)
## (Intercept)      0.688764597    0.599929398    0.658695608
## log_mass         0.079154144    0.072558662    0.076934920
## midpoint_lat     0.001361719    0.001177664    0.001304936
## phylogenetic.variance 0.013988745    0.011996198    0.013155882
## residual.variance  0.001737367    0.001522567    0.001668419
##               upper.CI(75) upper.CI(97.5)
## (Intercept)      0.719470381    0.778003604
## log_mass         0.081442874    0.085808844
## midpoint_lat     0.001441998    0.001579060
## phylogenetic.variance 0.014737269    0.016634933
## residual.variance  0.001827530    0.001999901
## attr("class")
## [1] "matrix" "mulTree"
```

```
multree_max_lifespan_con <- read.mulTree("bird_max_lifespan_resids",
                                         convergence= TRUE)
all(unlist(multree_max_lifespan_con) < 1.1)
```

```
## [1] TRUE
```

```
max_life_resids <- multree_birds$data$log_max_lifespan -
  c(summary(multree_max_lifespan_res)[1]
    + multree_birds$data$log_mass*summary(multree_max_lifespan_res)[2]
    + multree_birds$data$midpoint_lat*summary(multree_max_lifespan_res)[3])
```

## Age at first reproduction

Next we run the age at first reproduction model, due to the long run time of this model we will not evaluate it here but read in a previous run.

```
first_breed_resids_formula <- log_age_first_breed ~ log_mass + midpoint_lat
```

```
mulTree(mulTree.data = multree_birds,
        formula = first_breed_resids_formula,
        priors = priors,
        parameters = parameters,
        output = "bird_first_breed_resids",
        ESS= 1000,
        chains= 2)
```

Read in the Age at first reproduction models and get the overall summary

```
multree_first_breed_res <- read.mulTree("bird_first_breed_resids")
summary(multree_first_breed_res)
```

```
##              Estimates(mode hdr) lower.CI(2.5) lower.CI(25)
## (Intercept)      -0.0971695314 -0.1621827704 -0.1201841086
## log_mass          0.0783458828  0.0739379619  0.0769431707
## midpoint_lat     -0.0006796277 -0.0008188536 -0.0007358290
## phylogenetic.variance 0.0072477151  0.0062445722  0.0069059851
## residual.variance   0.0006029739  0.0005083972  0.0005665983
##              upper.CI(75) upper.CI(97.5)
## (Intercept)     -0.0761929847 -0.0338752572
## log_mass         0.0800480892  0.0829446178
## midpoint_lat     -0.0006464179 -0.0005570068
## phylogenetic.variance 0.0077327512  0.0085595174
## residual.variance 0.0006393496  0.0007121009
## attr(,"class")
## [1] "matrix" "mulTree"
```

Check if all the chains converged, as indicated by a convergence value of less than 1.1

```
multree_first_breed_conv <- read.mulTree("bird_first_breed_resids", convergence= TRUE)
all(unlist(multree_first_breed_conv) < 1.1)
```

```
## [1] TRUE
```

```
#calculate the residuals
```

```
first_breed_resids <- multree_birds$data$log_age_first_breed -
  c(summary(multree_first_breed_res)[1]
    + multree_birds$data$log_mass*summary(multree_first_breed_res)[2]
```

```
+ multree_birds$data$midpoint_lat*summary(multree_first_breed_res)[3])
```

## PCA

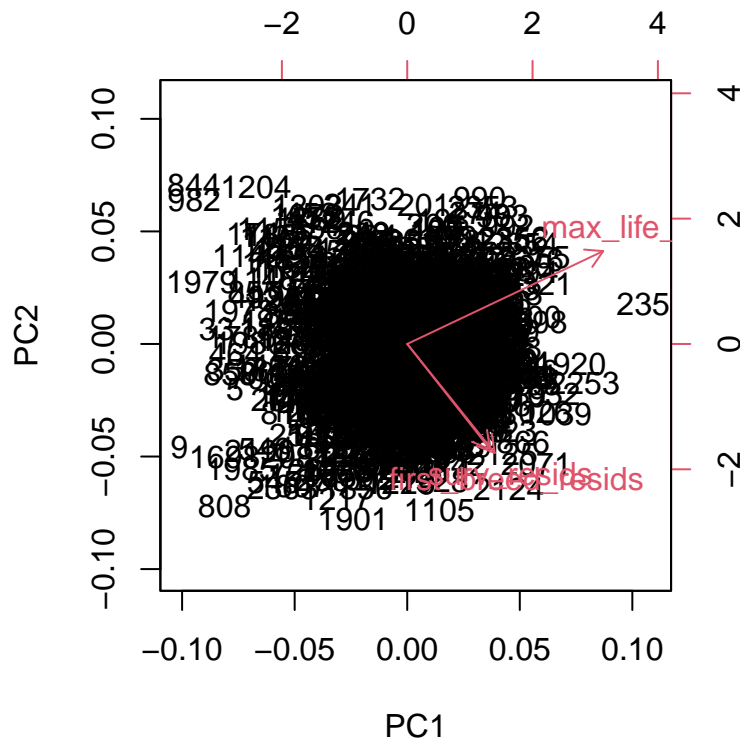
We then create a PCA using the residuals from the above three models. First create a dataset of the residuals.

```
res_fast_slow_bird_data <- data.frame(surv_resids,
                                       max_life_resids,
                                       first_breed_resids)
```

Next run the PCA

```
pca_data <- data.frame(surv_resids= res_fast_slow_bird_data$surv_resids,
                      max_life_resids= res_fast_slow_bird_data$max_life_resids,
                      first_breed_resids= res_fast_slow_bird_data$first_breed_resids)

biplot(prcomp(pca_data))
```



```
prcomp(pca_data)
```

```
## Standard deviations (1, ..., p=3):
## [1] 0.09418904 0.07271252 0.05496058
##
## Rotation (n x k) = (3 x 3):
##           PC1      PC2      PC3
## surv_resids  0.3590443 -0.5912669 -0.722143091
## max_life_resids  0.8522614  0.5230962 -0.004555714
## first_breed_resids 0.3804440 -0.6138190  0.691728705
```

```
#The first axis best defines the slow-fast continuum so we will use that.
pca_data2 <- prcomp(pca_data)
```

```
fast_slow_cont <- pca_data2$x[,1]
```

## Main model

We now add the fast-slow data to the multree data to set up for the main model.

```
fast_slow_bird_data <- data.frame(multree_birds$data, fast_slow_cont)
head(fast_slow_bird_data)
```

```
##           animal                sp.col      order diff_LociVS
## 1  Abroscopus_albogularis Abroscopus_albogularis Passeriformes  0.2088049
## 5  Abroscopus_superciliaris Abroscopus_superciliaris Passeriformes -1.4542725
## 7  Acanthagenys_rufogularis Acanthagenys_rufogularis Passeriformes -8.1521900
## 9    Acanthisitta_chloris    Acanthisitta_chloris Passeriformes 27.1174346
## 11   Acanthiza_apicalis    Acanthiza_apicalis Passeriformes -3.4947721
## 13   Acanthiza_chrysorrhoa Acanthiza_chrysorrhoa Passeriformes -0.4284985
## midpoint_lat log_mass survival log_max_lifespan log_age_first_breed
## 1      30.31612 1.576915 0.4967672      0.7858776      -0.088489893
## 5      12.84058 1.868721 0.5087253      0.7984239      -0.046711386
## 7      33.00662 3.867026 0.7179357      1.1326005       0.149555647
## 9      48.85290 1.932970 0.5215184      0.8171323      -0.058030884
## 11     33.69648 2.030776 0.6767431      1.2279219       0.012054488
## 13     35.19802 2.228939 0.6483806      1.1174075      -0.004625051
## mating_system fast_slow_cont
## 1              0      0.07349190
## 5              0      0.07162546
## 7              0     -0.11783293
## 9              0      0.07502000
## 11             0     -0.35083955
## 13             0     -0.21598031
```

```
nrow(fast_slow_bird_data)
```

```
## [1] 2371
```

```
#[1] 2371
```

```
fast_slow_bird_data <- data.frame(animal= fast_slow_bird_data$animal,
                                  latin_name= fast_slow_bird_data$sp.col,
                                  diff_LociVS= fast_slow_bird_data$diff_LociVS,
                                  midpoint_lat= fast_slow_bird_data$midpoint_lat,
                                  log_mass= fast_slow_bird_data$log_mass,
                                  mating_system= fast_slow_bird_data$mating_system,
                                  fast_slow_cont= fast_slow_bird_data$fast_slow_cont)
```

```
multree_birds_fast_slow <- as.mulTree(data= fast_slow_bird_data,
                                       tree= multiphylo_birds,
                                       taxa= "latin_name")
```

Running multree model with diff\_LociVS residual fast\_slow\_cont, midpoint\_latitude, log10(mass) and mating\_system. As its a long run we will read in a previous run.

```
multree_fsc_diff_LociVS <- diff_LociVS ~ fast_slow_cont + midpoint_lat
+ mating_system + log_mass
```

```

mulTree(mulTree.data = multtree_birds_fast_slow,
        formula = multtree_fsc_diff_LociVS,
        priors = priors,
        parameters = parameters,
        output = "fsc_diff_LociVS_June_4",
        ESS= 1000,
        chains= 2)

```

Read in a previous run of the model

```

multtree_fsc_diff_LociVS_model <- read.mulTree("fsc_diff_LociVS")

```

```

summary(multtree_fsc_diff_LociVS_model, use.hdr = FALSE)

```

```

##              Estimates(median) lower.CI(2.5) lower.CI(25)
## (Intercept)          27.3795167    -4.89862443    16.35882443
## fast_slow_cont        17.6707578     1.35949840    12.06353544
## midpoint_lat           0.1229018     0.04349461     0.09554608
## mating_system1        18.1580820    11.11327531    15.74276735
## log_mass              -7.3005046    -9.88806305    -8.18871079
## phylogenetic.variance 1832.6191972 1321.32070865 1648.75584155
## residual.variance      346.5909930  255.02987658  314.42878086
##              upper.CI(75) upper.CI(97.5)
## (Intercept)          38.4269425     59.732273
## fast_slow_cont        23.3094549     34.051877
## midpoint_lat           0.1505099     0.202658
## mating_system1        20.5602373     25.147483
## log_mass              -6.4175794     -4.733986
## phylogenetic.variance 2030.5923452  2403.599238
## residual.variance      378.7651226     443.185901
## attr(,"class")
## [1] "matrix" "mulTree"

```