**Supplementary material**

This is an annotated example of the statistical analyses performed in the R environment for the paper “Host specificity and the reproductive strategies of parasites” by Doherty, J.‐F., Milotic, M., Filion, A., and Eriksson, A. in *Parasitology*. This example was run in R version 4.0.0.

Objective of the paper: Investigate the effect of host specificity on the reproductive strategies of parasitic copepods of fishes and corals.

####R Script####

#This annotated example of Bayesian multilevel modelling tests for the effect of host specificity on the clutch size of fish-infecting copepods, with copepod family as a random effect. To run the analyses for coral-infecting copepods, just change the “fishcop” database to “coralcop” database.

#Packages required

library(brms)

#Note that the latest version of “rtools” must be installed in your computer.

#Load the database

cop <- read.csv(choose.files(), sep = ",", dec = ".", h = T) #Select the dataset provided in the Supplementary Material (the sheet named “data” in the “Dataset\_full.xlsx” file).

#Before using the following code, make sure that all variables meet model assumptions (e.g., if a continuous variable is loaded as an integer in your database, the model will not work).

#We first check the structure of the database.

str(cop)

#Modify the data if needed using the following line of code as an example.

cop$female\_size\_numeric <- as.numeric(paste(cop$female\_size))

#Create separate databases for fish- and coral-infecting copepods.

fishcop <- cop[cop$host\_type == "fish",]

coralcop <- cop[cop$host\_type == "coral",]

#Correcting host specificity for study effort.

prior <- bf(host\_nb ~ host\_ref, family = negbinomial())

get\_prior(prior, data = fishcop)

correct\_hostnb = brm(host\_nb ~ host\_ref, family = negbinomial(), data = fishcop,

chains = 2, cores = 4, iter = 4000,

prior = c(

prior(student\_t(3, 1.1, 2.5), "Intercept"),

prior(gamma(0.01,0.01), "shape")

))

summary(correct\_hostnb)

#Priors refer to the distribution of your variables in the database. By setting priors, we take into account the a priori knowledge we have of our data, making the sampling algorithm more efficient. “Chains”, “cores”, and “iter” are model parameters (more information can be found on the brms help page).

#We now add the predicted estimates to the dataframe.

pp <- predict(correct\_hostnb, newdata = fishcop)

pp1 <- as.data.frame(pp)

fishcop1 <- cbind(fishcop, pp1$Estimate)

names(fishcop1)[18] <- "Pred\_host"

#Here, we run the models (first the null model, then the full model with the effect variable of interest).

#Null model – we consider the model with female size alone as the null model.

prior <- bf(clutch\_size ~ scale(log(female\_size)) + (1|host\_family), family = Gamma(link = "log"))

get\_prior(prior, data = fishcop)

mod1\_fishclutch <- brm(clutch\_size ~ scale(log(female\_size)) + (1|host\_family), family = Gamma(link = "log"), data = fishcop1,

chains = 2, cores = 4, iter = 4000,

prior = c(

prior(student\_t(3, 5.3, 2.5), "Intercept"),

prior(student\_t(3, 0, 2.5), "sd"),

prior(gamma(0.01, 0.01), “shape”)

))

summary(mod1\_fishclutch)

#Final model with the predictor (fish-infecting copepods only).

prior <- bf(clutch\_size ~ scale(log(female\_size)) + scale(log(Pred\_host)) + (1|host\_family), family = Gamma(link = "log"))

get\_prior(prior, data = fishcop1)

mod2\_fishclutch <- brm(clutch\_size ~ scale(log(female\_size)) + scale(log(Pred\_host)) + (1|host\_family), family = Gamma(link = "log"), data = fishcop1,

chains = 2, cores = 4, iter = 4000, control = list(adapt\_delta = 0.95),

prior = c(

prior(student\_t(3, 5.3, 2.5), "Intercept"),

prior(student\_t(3, 0, 2.5), "sd"),

prior(gamma(0.01, 0.01), "shape")

))

summary(mod2\_fishclutch)

conditional\_effects(mod2\_fishclutch)

#Visualising the model predictions in a graph (fish-infecting copepods only).

p <- plot(conditional\_effects(mod2\_fishclutch))

p[[2]] + ggplot2::theme\_bw()+ ggplot2::theme(text = element\_text(size = 20), axis.title.y = element\_text(vjust = 2), axis.text.x = element\_text(size = 15), panel.grid.major = element\_blank(), panel.grid.minor = element\_blank(), panel.border = element\_rect(size = 2)) +

ggplot2::labs(x = "Host number", y ="Clutch size") + ylim(0,2300)