Size matching between fruits and bill gapes differentially affects gulper and masher frugivorous birds

Supplementary Material 2: Models specification and assessment

The consumption rate model was defined as follows:

$$y_i \sim \text{Gamma}(\text{shape} = \phi_{k[i]}^{-1}, \ \text{rate} = (\phi_{k[i]} \ \mu_i)^{-1})$$

$$\mu_i = \frac{\kappa_{k[i]}}{1 + \exp(-\eta_i)}$$

$$\begin{aligned}
\eta_i = \ &\alpha_{k[i]} \ + \ \beta_{k[i]} \ \text{sizediff}_i \ + \\
&\varepsilon_{\text{bird},b[i]} \ + \\
&\varepsilon_{\text{plant},p[i]} \ + \\
&\varepsilon_{\text{source},s[i]}
\end{aligned}$$

$$\alpha_k \sim \text{Normal}(0, 3)$$
$$\beta_k \sim \text{Normal}(0, 10 \ \text{SD}(\text{sizediff})^{-1})$$
$$\kappa_k \sim \text{Unif}(1, 30)$$

$$\varepsilon_{\text{bird},b} \sim \text{Normal}(0, \sigma_{\text{bird},k[b]})$$
$$\varepsilon_{\text{plant},p} \sim \text{Normal}(0, \sigma_{\text{plant}})$$
$$\varepsilon_{\text{source},s} \sim \text{Normal}(0, \sigma_{\text{source}})$$

$$\sigma_{\text{bird},k} \sim \text{Normal}(0, 3) \ \text{T}(0, \infty)$$
$$\sigma_{\text{plant}} \sim \text{Normal}(0, 3) \ \text{T}(0, \infty)$$
$$\sigma_{\text{source}} \sim \text{Normal}(0, 3) \ \text{T}(0, \infty)$$

$$\phi_k \sim \text{Normal}(0, 3) \ \text{T}(0, \infty).$$

y_i is the consumption rate corresponding to the foraging behaviour $k[i] \in \{1 = \text{gulper}, 2 = \text{masher}\}$, bird species $b[i]$, plant species $p[i]$, and source study $s[i]$. μ_i is the mean of the Gamma distribution, $\phi_{k[i]}$ is the dispersion parameter, and the variance is defined as $\text{Var}(Y_i) = \mu_i^2 \ \phi_{k[i]}$. "sizediff" is the difference between the bill gape width and the fruit diameter (mm). ε represents random effects for the bird species, plant species, and source. Note that $\alpha, \beta, \kappa, \phi$, and σ_{bird} vary between mashers and gulpers, indicated by the k suffix. Normal distributions are parameterised with mean and standard deviation, and T(,) indicates the lower and upper bound for truncated distributions.

The model for consumption probability was very similar, assuming a Bernoulli distribution of the response

and not including the source random effect:

$$y_i \sim \text{Bernoulli}(p_i)$$

$$p_i = \frac{\kappa_{k[i]}}{1 + \exp(-\eta_i)}$$

$$\eta_i = \alpha_{k[i]} \ + \ \beta_{k[i]} \ \text{sizediff}_i \ +$$
$$\varepsilon_{\text{bird},b[i]} \ +$$
$$\varepsilon_{\text{plant},p[i]}$$

$$\alpha_k \sim \text{Normal}(0,3)$$
$$\beta_k \sim \text{Normal}(0, 10 \ \text{SD}(\text{sizediff})^{-1})$$
$$\kappa_k \sim \text{Unif}(0.2, 1)$$

$$\varepsilon_{\text{bird},b} \sim \text{Normal}(0, \sigma_{\text{bird},k[b]})$$
$$\varepsilon_{\text{plant},p} \sim \text{Normal}(0, \sigma_{\text{plant}})$$

$$\sigma_{\text{bird},k} \sim \text{Normal}(0,3) \ \text{T}(0,\infty)$$
$$\sigma_{\text{plant}} \sim \text{Normal}(0,3) \ \text{T}(0,\infty).$$

Here, $y_i \in \{0,1\}$ indicated whether each pair of bird-plant species had the consumption rate recorded in their corresponding study.

In both models, the posterior distribution showed negligible influence from the prior.
We assessed model fit by computing randomized cumulative probability residuals (Dunn and Smyth 1996), using the DHARMa R package (Hartig 2022; Fig. A1). Both models showed adequate fit to data. Residuals for the consumption probability model in mashers at high size difference showed a negative trend, but as it was influenced by too few observations, we preferred to keep the logit-linear model.
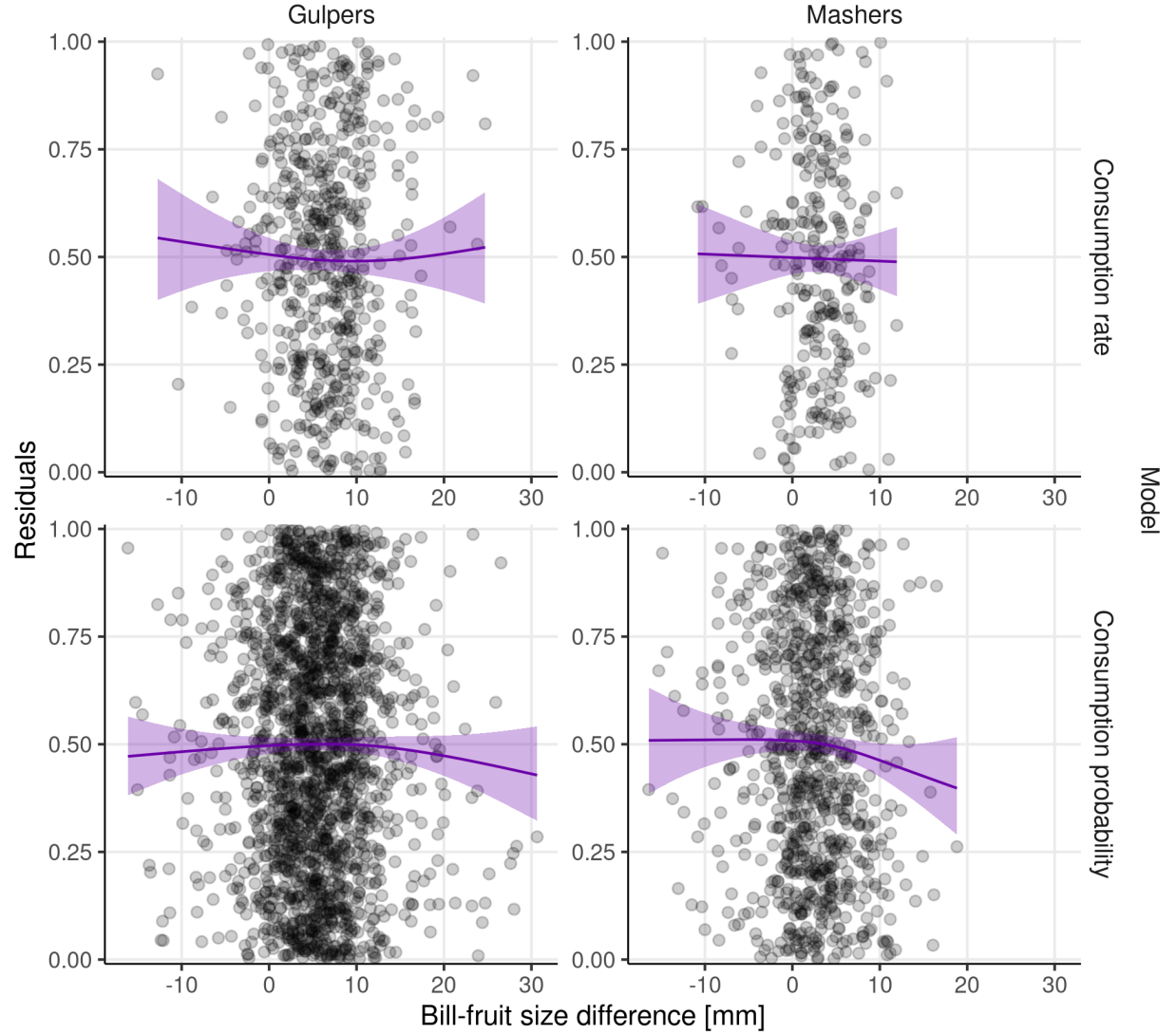
Figure S1. Residuals from the consumption rate and consumption probability models, as a function of explanatory variables. The curves show the prediction from a Beta Generalized Additive Model fitted to residuals, using 19 basis functions apart from the intercept. Under good model fit, these residuals show a Uniform distribution in [0, 1], with no patterns as a function of predictors.

We checked for phylogenetic signal in the data by analysing whether the similarity in pairs of random effects for birds and plants were related to the phylogenetic correlation. Phylogenies were obtained pruned from the VertLife project (Jetz et al. 2012) for birds and using the `V.Phyllomaker` 2.0 R package (Jin and Qian, 2019) for plants. After multichotomies were solved, we used the `ape` R package (Paradis and Schliep, 2019) to get the phylogenetic correlation matrix of the species.

A strong phylogenetic signal would make pairs of species with closer phylogenetic relations have more similar random effects in the models. Namely, the pairwise absolute difference in random effects would decrease as the pairwise phylogenetic correlation increases. We only found a small phylogenetic effect in the plants random effects from the consumption rate model (Fig. A2). However, the decreasing trend in the pairwise random effects difference was driven by a small proportion of plant species pairs with zero correlation (93 of 4371 species pairs, 2.13 %). Most species pairs showed relatively high correlation (above 0.55) and the random effects pairwise difference did not vary within this correlation range. Hence, despite there seems to be a small phylogenetic signal in plants, fitting a correlation function to this dataset would be troublesome, as there is no correlation in the bulk of the dataset. Moreover, despite our results might underestimate uncertainty due to phylogenetic signal, as it is reflected only in a small proportion of the dataset, its effect would be negligible. For these reasons we kept the model with independent random effects.
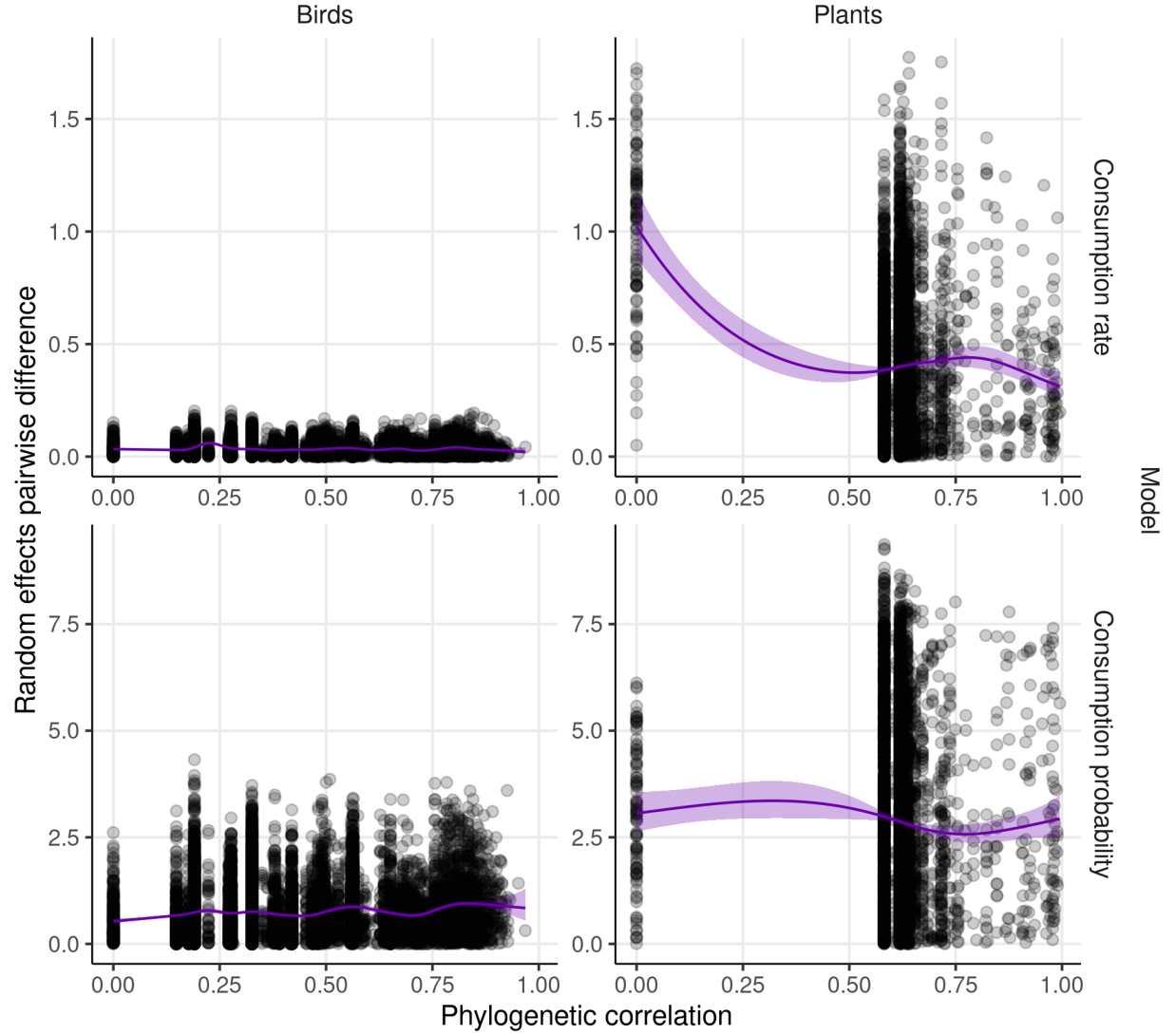
Figure S2. Phylogenetic signal assessment for the consumption rate and probability models. Absolute pairwise difference in plant and bird species random effects as a function of pairwise phylogenetic correlation. Random effects differences are represented by the posterior means. The curves show the prediction from a Generalized Additive Model fitted to points, assuming a Gamma distribution with log link and using 19 basis functions apart from the intercept. Under a strong phylogenetic signal, the pairwise difference should decrease as correlation increases.

References

- Dunn, P. K., & Smyth, G. K. (1996). Randomized quantile residuals. Journal of Computational and graphical statistics, 5(3), 236-244.

- Hartig, F. (2022). DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models. R package version 0.4.6, https://CRAN.R-project.org/package=DHARMa.

- Jetz, W., G. H. Thomas, J. B. Joy, K. Hartmann, & A. O. Mooers (2012). The global diversity of birds in space and time. Nature 491:444–448. http://www.nature.com/nature/journal/v491/n7424/abs/nature11631.html

- Jin, Y., & Qian, H. (2019). V. PhyloMaker: an R package that can generate very large phylogenies for vascular plants. Ecography, 42(8), 1353-1359.

- Paradis E., & Schliep K. (2019). "ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R." *Bioinformatics*, *35*, 526-528. doi:10.1093/bioinformatics/bty633 https://doi.org/10.1093/bioinformatics/bty633.