Appendix 1: Additional details about the interaction network simulation approach

Appendix 2: Source code for the mechanistic trait-based seed dispersal model in the R language for statistical computing

Appendix 3: Additional details about the trait-based seed dispersal model and allometric equations used

Table A1: Summary of feeding trial studies for the relationship between avian frugivore body mass and gut passage time

Table A2: Summary of field-based empirical studies for the relationship between avian frugivore body mass and seed dispersal distances

Figure A1: Relationship between body mass and mean gut passage time using data extracted from empirical feeding trials for frugivorous birds

Figure A2: Relationship between body mass and max seed dispersal distance using data extracted from empirical feeding trials for frugivorous birds

Figure A3: Relationship between body mass and mean dispersal distance using data extracted from empirical studies of seed dispersal by frugivorous birds

Appendix 4: Long-distance seed dispersal (LDD) for small, medium, and large fruited plant species

Figure A4: Small, medium, large fruited plant species long-distance seed dispersal

Appendix 5: Sensitivity analysis

Table A3: Model parameters included in the sensitivity analysis

Figure A5: Gut passage time exponent (GPTexp) local sensitivity analysis

Figure A6: Gut passage time variance (GPTvar) local sensitivity analysis

Figure A7: Correction factor (corrfactor) local sensitivity analysis
Appendix 1: Additional details about the interaction network simulation approach

In the network simulations, total interaction frequencies of plants took differences in plant species abundance into account. We assumed a negative relationship between fruit size and interaction frequency (Donoso et al. 2017; González-Castro et al. 2015; Moles et al. 2005):

\[ f_i = \frac{1}{x_i} \]

where \( x_i \) represents the fruit volume value for plant \( i \), and \( f_i \) represents the expected total interaction frequency (Donoso et al. 2017).

Similarly, total interaction frequencies of bird species took differences in bird abundance into account. We assumed a negative relationship between body mass and abundance (Cotgreave 1993; González-Castro et al. 2015); in this case, we assumed undercompensation (i.e. interaction frequency decreases less rapidly than bird size increases) as large birds tend to consume more fruits per individual (García et al. 2014).

\[ g_j = \frac{1}{y_j} + \beta \]

where \( y_j \) is the bird size value for bird \( j \), \( g_j \) is the expected total bird interaction frequency, and \( \beta \) being an undercompensation parameter, set to 10% of the maximum value of \( 1/y \). Donoso et al. 2017 found that results were robust to variation in the value of \( \beta \). Because factors other than size may influence species abundances, we investigated whether imperfect relationships between size and interaction frequencies had an effect on model results. To do this, we predefined an imperfect \( (r^2 = 0.6) \) relationship between interaction frequency and fruit volume, and between interaction frequency and bird size, using the R package \textit{faux} version 0.0.1.0. We found that mean seed dispersal distances, derived from our simulations, were unchanged when implementing these imperfect relationships between size and interaction frequencies.
Appendix 2: Source code for the mechanistic trait-based seed dispersal model in the R language for statistical computing

Explanatory comments (#)

#----------------trait-based seed dispersal distance model----------------#

#The dispsimulation function generates estimated seed dispersal distances for plant-
#bird interactions. This function takes as input an object with disperser body mass
(#kg) for each interaction event in a network.

#Nbird = number of bird species in the community
#obsperbird = number of interaction events for each bird species

dispsimulation <- function (x) {
  dispdist <- rep(NA, obsperbird * Nbird)
  for(i in 1:nrow(x)) {
    #a mean GPT is selected from the allometric equation derived from empirical data
    #presented in this study. [i,3] indicates the column where bird body mass is located,
    #this may not fit with other data structures
    meanGPThour <- 4.5*x[i,1]^0.5

    #convert GPT to seconds (since speed is in m/s)
    meanGPT <- meanGPThour*3600

    #calculate the shape and scale parameters for the gamma distribution using meanGPT
    #and variance (we chose variance = 100241, since this was the average GPT
    #variance calculated across 11 empirical studies in which variance was reported;
    #see Table S2)
    scalevalue <- 75311/meanGPT
    shapevalue <- meanGPT^2/75311

    #select a GPT value for this particular interaction from the GPT gamma distribution
    GPT <- rgamma(1, shape = shapevalue, scale = scalevalue)

    #then select a mean flight speed (calculated used the allometric equation presented
    #in Alerstam et al. 2007)
    meanspeed <- 15.7*x[i,1]^0.17

    #select a flight speed value for this particular interaction using meanspeed and
    #2.078 to parameterize rnorm. 2.078 is the flight speed sd average
    #reported in Alerstam et al. 2007 for those species with body mass lower than 1.77 kg
    #(which is the largest bird species across our 7 Andean communities)
    speed <- rnorm(1, meanspeed, 2.078)

    #calculate the max distance travelled (if flying straight without stopping) given the
    #selected GPT.
    max_distance <- speed*GPT

    #correction factor which accounts for birds resting/not always moving in a straight
    #line.
    distance <- 0.002 * max_distance
  }
}

# NOTE! there may be a few cases where the speed value -selected from rnorm- could have a negative value.
# For these few cases, the negative seed dispersal distance is replaced with NA.
if (distance < 0){
  distance<-NA
  dispdist[i] <- distance
}
return(dispdist)
We developed an allometric equation specific to frugivores. We only included studies that fed natural fruit to birds and excluded studies using artificial seeds or fruits, or marker dyes. We used the search strings “seed or fruit + gut + retention or passage”. For some studies GPT medians were reported instead of means, if means could not be attained via author personal communication or digitisation from presented plots, the study was not included in our analysis. The allometric relationship between body mass and GPT presented by Robbins 1993 included data on 21 bird species across all diet types (including studies using liquid and marker dye to measure GPT). Only 4 of the 21 species were fed fruits. The 37 included species are widely distributed across the weight range of frugivore species found in the seven Andean communities. Generally, standard errors were reported instead of standard deviations; however, if standard errors and sample sizes were both reported we converted standard error to standard deviation.
140 Table A2. Summary of field-based empirical studies for the relationship between avian frugivore
141 body mass and seed dispersal distances.

<table>
<thead>
<tr>
<th>Species</th>
<th>Body mass (g)</th>
<th>Mean dispersal distance (m)</th>
<th>Max dispersal distance (m)</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bycanistes bucinator</td>
<td>635</td>
<td>528</td>
<td>14790</td>
<td>Mueller et al. 2014</td>
</tr>
<tr>
<td>Ceratogymna atrata</td>
<td>1431</td>
<td>1521</td>
<td>6919</td>
<td>Holbrook &amp; Smith 2000</td>
</tr>
<tr>
<td>Ceratogymna cyllindricus</td>
<td>1038</td>
<td>1537</td>
<td>4628</td>
<td>Holbrook &amp; Smith 2000</td>
</tr>
<tr>
<td>Corythaeca cristata</td>
<td>1000</td>
<td>240.5</td>
<td>NA</td>
<td>Sun et al. 1997</td>
</tr>
<tr>
<td>Dicaeum hirundinaceum</td>
<td>9.25</td>
<td>103.67</td>
<td>500</td>
<td>Ward &amp; Paton 2007</td>
</tr>
<tr>
<td>Hemiphaga novaeseelandiae</td>
<td>650</td>
<td>84.7</td>
<td>1469</td>
<td>Wotton &amp; Kelly 2012; Wotton et al. 2008</td>
</tr>
<tr>
<td>Micronectes oleagineus</td>
<td>11.5</td>
<td>26.16</td>
<td>86</td>
<td>Westcott &amp; Graham 2000</td>
</tr>
<tr>
<td>Musophaga johnstoni</td>
<td>250</td>
<td>137.5</td>
<td>NA</td>
<td>Sun et al. 1997</td>
</tr>
<tr>
<td>Myadestes melanops</td>
<td>31.5</td>
<td>84.7</td>
<td>364.7</td>
<td>Murray 1988</td>
</tr>
<tr>
<td>Onychognathus tristramii</td>
<td>119</td>
<td>1168</td>
<td>4800</td>
<td>Spiegel &amp; Nathan 2007</td>
</tr>
<tr>
<td>Phainoptila melanoxantha</td>
<td>58</td>
<td>84.9</td>
<td>504.7</td>
<td>Murray 1988</td>
</tr>
<tr>
<td>Prosthemadera novaeseelandiae</td>
<td>105</td>
<td>222.5</td>
<td>NA</td>
<td>O’Connor 2006</td>
</tr>
<tr>
<td>Pycnonotus xanthopygos</td>
<td>40.5</td>
<td>303</td>
<td>900</td>
<td>Spiegel &amp; Nathan 2007</td>
</tr>
<tr>
<td>Semnornis frantzii</td>
<td>63.25</td>
<td>62.6</td>
<td>215</td>
<td>Murray 1988</td>
</tr>
<tr>
<td>Turaco schuetii</td>
<td>250</td>
<td>149</td>
<td>NA</td>
<td>Sun et al. 1997</td>
</tr>
<tr>
<td>Turdus merula</td>
<td>100</td>
<td>89.48</td>
<td>2220</td>
<td>Breitbach et al. 2012</td>
</tr>
</tbody>
</table>

142 We included empirical seed dispersal studies which quantified seed dispersal distances by
143 combining gut passage time and frugivore movement data. We did not include studies
144 approximating SDD based on molecular data. We used ordinary least squares (OLS) to fit an
145 allometric equation between bird body mass and mean seed dispersal distance for empirical
146 field-based studies (Table S1). This resulted in the following equation: \( z = 504BM^{0.48} \), where
147 \( z \) is seed dispersal distance and BM is disperser species body mass. The ratio between the
148 allometric constant from the independent expectation (equation 7 in the main text; 504/254340)
149 and the allometric constant from empirical studies presented here was used to calculate the
150 correction factor (0.002; accounting for movements deviation from a straight line and time not
151 moving).
Fig. A1. Relationship between body mass and mean gut passage time using data extracted from empirical feeding trials for frugivorous birds (see detailed information about the studies in Table S2). Body mass is positively related to mean gut passage time ($r^2 = 0.69$, $p < 0.0001$, $n=39$). The grey shaded region indicates the confidence interval for the regression.
Fig. A2. Relationship between **body mass** and **mean dispersal distance** using data extracted from empirical studies of seed dispersal by frugivorous birds (see Table S2 for included studies). Body mass is positively related to mean dispersal distance ($r^2 = 0.4, p = 0.007, n = 16$). The grey shaded region indicates the confidence interval for the regression.
Fig. A3. Relationship between body mass and max seed dispersal distance using data extracted from empirical feeding trials for frugivorous birds (see Table A2). Body mass is positively related to max seed dispersal distance ($r^2 = 0.62$, $p = 0.001$, $n=12$). The grey shaded region indicates the confidence interval for the regression.
Appendix 4.

Fig. A4. Long-distance seed dispersal (LDD) results for (b) small, (c) medium, and (d) large fruited plant species.
Appendix 5: Sensitivity analysis

Morris’s elementary effects method estimates the effect of each factor on the model output repeatedly, while the other factors take on different values from their entire ranges, and then averages these estimates into a measure of overall effect; these effects are called elementary effects. The elementary effects are statistically analysed to measure their relative importance (Thiele et al. 2004). We used the estimated mean of the distribution of the absolute values of the elementary effects, $\mu^*$, as a sensitivity measure to establish the overall impact of a parameter on the output.

We performed the sensitivity analysis on five model parameters ($k$; Table 1), which were varied according to predefined ranges (see Table A3). The number of tested settings is given by $r \times (k + 1)$, where $r$ is the number of elementary effects computed per parameter. As we chose 160 elementary effects, this led to $160 \times (5 + 1) = 960$ model runs. We ran the global sensitivity analysis for both, the mean and the 95% quantile of seed dispersal distances.

We used the following methods to determine the range of the parameter values to be included in the global sensitivity analysis. For $GPT^{exp}$ we used the 95% confidence intervals of the exponent from the fitted allometric equation; for $GPT^{var}$ we used the min and max values from feeding trial studies (Table A1); for $FS^{exp}$ we took the range of 95% confidence intervals of the exponent from those calculated in a similar flight speed allometric equation presented in Alerstam et al. 2007; for $FS^{sd}$ we took the min and max standard deviation values from those reported from empirical flight speed data in Alerstam et al. 2007; for the CorrFactor we simply used a min value that was half of the estimated value and a maximum value that was twice the estimated value.
Table A3. Sensitivity analysis model parameters and results from the Morris screening method.

The top three most influential parameters for median seed dispersal distances are bolded in black; the top three most influential parameters for the 95% quantile of seed dispersal are bolded in orange. $\mu^*$ is an estimate of the overall influence of a factor on the model output (including interactions with other factors), and $\sigma$ is an estimate of how much the influence of a factor depended on interactions and stochasticity.

<table>
<thead>
<tr>
<th>parameter</th>
<th>description</th>
<th>range</th>
<th>$\mu^*$</th>
<th>$\sigma$</th>
<th>$\mu^*$</th>
<th>$\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>gut passage time:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$GPT^{exp}$</td>
<td>exponent of the GPT Eq. 3</td>
<td>0.39–0.62</td>
<td><strong>0.31</strong></td>
<td><strong>0.58</strong></td>
<td>0.16</td>
<td>0.31</td>
</tr>
<tr>
<td>$GPT^{var}$</td>
<td>variance of the GPT gamma distribution, $s^2$ in Eq. 5 and 6</td>
<td>2613–931509</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.98</td>
</tr>
<tr>
<td>bird movement:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$FS^{exp}$</td>
<td>exponent of the FS Eq. 4</td>
<td>0.13–0.21</td>
<td>0.12</td>
<td>0.23</td>
<td>0.01</td>
<td>0.007</td>
</tr>
<tr>
<td>$FS^{sd}$</td>
<td>standard deviation of the FS gaussian distribution</td>
<td>0–4.7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$CorrFactor$</td>
<td>$fc$ in Eq. 7</td>
<td>0.001–0.004</td>
<td><strong>0.58</strong></td>
<td><strong>0.96</strong></td>
<td>0.55</td>
<td>1</td>
</tr>
</tbody>
</table>
**Fig. A5.** The relationship between network specialisation ($H_2'$) and median community seed dispersal distances ($TDK_{community}$) when using the a) min $GPT^{exp}$ value, and b) max $GPT^{exp}$. $GPT^{exp}$ is included in the top three most influential parameters for: mean and the 95% quantile of seed dispersal distances. Results show the same hump-shaped pattern between $H_2'$ and community-wide median seed dispersal distances. Absolute distance values for both the mean (min$GPT^{exp}$: peak in seed dispersal = 98 m; max$GPT^{exp}$: peak in seed dispersal = 53 m) and the 95% quantile values (min$GPT^{exp}$: peak in seed dispersal = 209 m; max$GPT^{exp}$: peak in seed dispersal = 140 m) are different. Please note different scales of the y-axes.
Fig. A6. The relationship between network specialisation ($H_2'$) and mean community seed dispersal distances ($\text{TDK}_{\text{community}}$) when using the a) min $GPT\text{var}$ value, and b) max $GPT\text{var}$. $GPT\text{var}$ is included in the top three most influential parameters for: mean, and 95% quantile seed dispersal distances. All figures show the same hump-shaped pattern between $H_2'$ and mean or LDD community-wide seed dispersal distances. c), and d) report results from the 95% quantile. Absolute seed dispersal distance values were very different for the mean (min$GPT\text{var}$: peak in seed dispersal = 2.5 m; max$GPT\text{var}$: peak in seed dispersal = 909) and 95% quantile of seed dispersal distances (min$GPT\text{var}$: peak in seed dispersal = 6 m; max$GPT\text{var}$: peak in seed dispersal = 2157 m).
Fig. A7. The relationship between network specialisation ($H'_2$) and community seed dispersal distances ($\text{TDK}_{\text{community}}$) when using the a) min $\text{CorrFactor}$ value, and b) max $\text{CorrFactor}$. $\text{CorrFactor}$ is included in the top three most influential parameters for: mean, and 95% quantile seed dispersal distances. c), and d) report results from the 95% quantile. All figures show the same hump-shaped pattern between $H'_2$ and median or LDD community-wide seed dispersal distances. Absolute seed dispersal distance values were longer under the max $\text{CorrFactor}$ scenario. Please note different scales of the y-axes.
Supplementary References


