**Can migratory** **birds spread avian haemosporidian parasites?**

Daniela de Angeli Dutra¹\*, Antoine Filion¹, Alan Fecchio², Érika Martins Braga³, Robert Poulin¹

**SUPPLEMENTARY MATERIAL**

Supplementary Table 1: Dataset summary detailing information regarding locality ID, total number of birds individuals and host richness per locality, biome, latitude, longitude and data source.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Locality | Total of Birds | Host Richness | Biome | Latitude | Longitude | Year of Collection | Source |
| BSB | 242 | 50 | Brazilian Savanna | -15.5211053 | -47.552045 | 2014/2015 | unpublished |
| Salvador | 161 | 28 | Atlantic Rain Forest | -12.5450889 | -38.207439 | 2015/2016 | unpublished |
| Viracopos | 150 | 24 | Brazilian Savanna | -23.029538 | -47.815246 | 2014 | unpublished |
| Aiu1 | 62 | 24 | Caatinga | -6.6 | -40.116667 | 2012/2013 | Fecchio et al. 2019 |
| Aracruz | 84 | 26 | Atlantic Rain Forest | -19.4910125 | -40.1206373 | 2011/2012 | Lacorte et al. 2013 |
| Base UFMT | 94 | 53 | Pantanal | -15.3633507 | -56.355506 | 2011/2012 | unpublished |
| Bel1 | 320 | 157 | Amazonia | -3.7 | -46.75 | 2013 | Fecchio et al. 2019 |
| Bocaiuva | 295 | 59 | Brazilian Savanna | -17.657674 | -43.4844164 | 2011/2012 | Lacorte et al. 2013 |
| Bra1 | 788 | 52 | Brazilian Savanna | -15.533333 | -47.55 | 2005 a 2009 | Fecchio et al. 2019 |
| Brasilandia de Minas | 133 | 56 | Brazilian Savanna | -16.5630651 | -45.5308804 | 2011/2012 | Lacorte et al. 2013 |
| Bru1 | 53 | 25 | Atlantic Rain Forest | -27.098611 | -48.8925 | 2016 | Fecchio et al. 2019 |
| Busca Vida | 169 | 37 | Atlantic Rain Forest | -12.5143471 | -38.1715148 | 2015/2016 | unpublished |
| Caratinga | 126 | 33 | Atlantic Rain Forest | -19.4724319 | -42.825918 | 2011/2012 | Lacorte et al. 2013 |
| Cha1 | 90 | 36 | Brazilian Savanna | -15.316667 | -55.866667 | 2011/2014/2015 | Fecchio et al. 2019 |
| ChaP Guim | 101 | 32 | Brazilian Savanna | -15.61953 | -55.327475 | 2011/2012 | Ferreira et al. 2018 |
| CLBI | 1800 | 70 | Atlantic Rain Forest | -5.5524107 | -35.103905 | 2014/2015 | unpublished |
| ESEC | 933 | 56 | Caatinga | -6.344364 | -37.152527 | 2013/2014 | unpublished |
| Esp1 | 244 | 63 | Grassland | -30.205564 | -57.495736 | 2014 | Fecchio et al. 2019 |
| Faz St. Maria | 131 | 35 | Brazilian Savanna | -21.1658279 | -47.45862 | 2014 | unpublished |
| Felixlandia | 175 | 40 | Brazilian Savanna | -18.4526633 | -44.5356927 | 2011/2012 | unpublished |
| Gra1 | 127 | 26 | Brazilian Savanna | -15.166667 | -45.733333 | 2013 | Fecchio et al. 2019 |
| Gui1 | 178 | 53 | Amazonia | -0.4 | -64.8 | 2012 | Fecchio et al. 2019 |
| Gui2 | 304 | 77 | Amazonia | -1.35 | -56.366667 | 2011/2012 | Fecchio et al. 2019 |
| Gui3 | 39 | 15 | Amazonia | -2.932778 | -59.972533 | 2014 | Fecchio et al. 2019 |
| Hual1 | 80 | 42 | Peruvian Andes | -6.583333 | -77.55 | 2010 | Fecchio et al. 2019 |
| Hual2 | 167 | 55 | Peruvian Andes | -6.683333 | -77.683333 | 2010/2011 | Fecchio et al. 2019 |
| Hual3 | 157 | 60 | Peruvian Andes | -5.1 | -79.233333 | 2010/2011 | Fecchio et al. 2019 |
| Hual4 | 352 | 101 | Peruvian Andes | -6.583333 | -77.55 | 2010 | Fecchio et al. 2019 |
| Hual5 | 400 | 114 | Peruvian Andes | -6.716667 | -77.416667 | 2010 | Fecchio et al. 2019 |
| Ime1 | 164 | 51 | Amazonia | -0.583333 | -64.916667 | 2010 | Fecchio et al. 2019 |
| Ina1 | 211 | 40 | Amazonia | -4.998 | -62.935 | 2012 | Fecchio et al. 2019 |
| Ina2 | 208 | 53 | Amazonia | -5.72 | -63.217 | 2012 | Fecchio et al. 2019 |
| Ina3 | 297 | 65 | Amazonia | -9.1 | -64.466667 | 2010/2011 | Fecchio et al. 2019 |
| Jau1 | 116 | 37 | Brazilian Savanna | -14.866667 | -59 | 2011 | Fecchio et al. 2019 |
| Jequitinhonha | 166 | 50 | Brazilian Savanna | -16.2559998 | -41.011001 | 2011/2012 | Lacorte et al. 2013 |
| Joi1 | 49 | 18 | Atlantic Rain Forest | -26.3 | -48.883333 | 2014 | Fecchio et al. 2019 |
| Manga | 120 | 42 | Brazilian Savanna | -14.4510092 | -43.5630135 | 2011/2012 | Lacorte et al. 2013 |
| Mata Seca | 461 | 64 | Brazilian Savanna | -14.50911 | -43.59298 | 2013/2014 | Ferreira et al. 2016 |
| Mic1 | 106 | 34 | Atlantic Rain Forest | -13.84 | -39.241 | 2017 | Fecchio et al. 2019 |
| Mil1 | 214 | 64 | Caatinga | -12.90198 | -39.841985 | 2015/2017 | Fecchio et al. 2019 |
| Nat1 | 37 | 16 | Atlantic Rain Forest | -5.916667 | -35.166667 | 2012/2015 | Fecchio et al. 2019 |
| Nova Lima | 164 | 46 | Atlantic Rain Forest | -19.5907472 | -43.5049448 | 2011/2012 | Lacorte et al. 2013 |
| Pan1 | 110 | 10 | Pantanal | -19.566667 | -57.016667 | 2009/2012 | Fecchio et al. 2019 |
| Pan2 | 122 | 57 | Pantanal | -16.25 | -56.366667 | 2009/2017 | Fecchio et al. 2019 |
| ParNa BSB | 258 | 41 | Brazilian Savanna | -15.4018485 | -47.5852374 | 2014/2015 | unpublished |
| Rond1 | 429 | 86 | Amazonia | -4.683333 | -56.633333 | 2012 | Fecchio et al. 2019 |
| Rond2 | 60 | 35 | Amazonia | -5.066667 | -56.85 | 2012 | Fecchio et al. 2019 |
| Rond3 | 176 | 53 | Amazonia | -9.316667 | -64.716667 | 2010/2011 | Fecchio et al. 2019 |
| Rond4 | 117 | 40 | Amazonia | -12.216667 | -60.733333 | 2011 | Fecchio et al. 2019 |
| Rond5 | 136 | 42 | Amazonia | -13.8 | -59.683333 | 2011 | Fecchio et al. 2019 |
| RondIL | 85 | 18 | Amazonia | -4.5 | -56.266667 | 2012 | Fecchio et al. 2019 |
| Saj1 | 55 | 21 | Atlantic Rain Forest | -28.153369 | -49.641754 | 2014/2015 | Fecchio et al. 2019 |
| Salto da Divisa | 196 | 56 | Brazilian Savanna | -15.5954324 | -39.5655673 | 2011/2012 | Lacorte et al. 2013 |
| Seb1 | 76 | 30 | Atlantic Rain Forest | -23.166667 | -44.833333 | 2015 | Fecchio et al. 2019 |
| Ser1 | 178 | 38 | Caatinga | -6.566667 | -37.266667 | 2012/2013 | Fecchio et al. 2019 |
| Sfr1 | 179 | 57 | Atlantic Rain Forest | -29.466667 | -50.166667 | 2017 | Fecchio et al. 2019 |
| Sooretama | 86 | 21 | Atlantic Rain Forest | -19.1130526 | -40.613344 | 2011/2012 | Lacorte et al. 2013 |
| Sta1 | 156 | 30 | Grassland | -36.716667 | -64.283333 | 2015 | Fecchio et al. 2019 |
| Taiama | 44 | 22 | Pantanal | -16.577528 | -55.1027227 | 2011/2012 | unpublished |
| Tap1 | 39 | 22 | Amazonia | -5.1 | -56.433333 | 2012 | Fecchio et al. 2019 |
| Tap2 | 61 | 36 | Amazonia | -5.216667 | -56.916667 | 2012 | Fecchio et al. 2019 |
| Varz Grande | 38 | 13 | Brazilian Savanna | -15.3846493 | -56.757767 | 2011/2012 | unpublished |
| Xing1 | 322 | 104 | Amazonia | -1.95 | -51.6 | 2007 | Fecchio et al. 2019 |
| Total | 13191 | 914 |  |  |  |  |  |

**Potential correlates of prevalence and richness**

*Spatial autocorrelation*

All analyses were conducted in R version 4.0 (R Core Team, 2019). To evaluate the potential impact of locality on our variables, we determined whether there was significant spatial autocorrelation among localities for total prevalence (i.e., number of infected hosts/total number of hosts) and parasite richness (total number of parasites lineages) in our dataset by calculating the Moran Index value. Index values vary between -1.0 and +1.0, with 0 indicating no spatial autocorrelation and -1.0 or +1.0 high spatial autocorrelation. In order to estimate this index, we combined the coordinates data into a matrix and employed the function “Moran.I” from the “Ape” package (Paradis & Schliep, 2018).

*Phylogenetic Signal*

In order to estimate the phylogenetic signal (i.e., tendency for phylogenetically closely-related species to resemble each other more than random species of the same tree) among prevalence and richness estimates for the bird species in our dataset, we downloaded the file AllBirdsHackett1.tre from <https://birdtree.org/> website. Using the “treeman” package (Bennett, Sutton, & Turvey, 2017), we created a treeman file containing all trees from the original file. Then, we randomly selected 100 trees. This new file was converted from treeman to a phylo file, from which we extracted one single random tree to account for phylogenetic uncertainty. We grouped our data per species and eliminated all bird species from the phylo tree which were not present in our dataset. Using the “match” function from the “picante” package (Kembel et al., 2010), we matched the species between the tree and our dataset. Then, we calculated Pagel’s lambda (λ) to evaluate the phylogenetic signal among bird species in our dataset, for haemosporidian (all three genera) prevalence and parasite richness (Pagel, 1999). Values of λ can range between 0 and 1, being 1 when the trait has evolved consistently with a Brownian motion. In order to estimate lambda (λ), we applied the “phylosig” function from the “phytools” package (Revell, 2012).

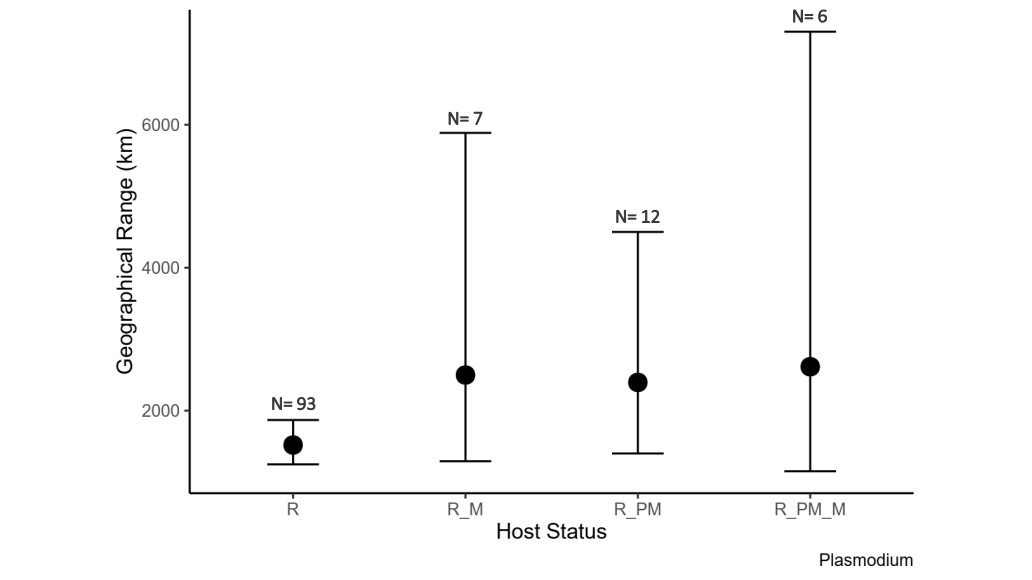
*Climate variables*

We used annual mean precipitation (variable BIO15) and annual mean temperature (variable BIO1, ºC) as predictors in the Bayesian and mixed models since temperature and precipitation are known to impact haemosporidian infections (Clark, Drovetski, & Voelker, 2020; Illera, López, García-Padilla, & Moreno, 2017). We used R to extract these climate variables from the Worlclim database (<https://worldclim.org/version2>). Using the package “raster”, we extracted the data using the “getData” function, then we selected the data from the 63 localities included in our original dataset (10 minute resolution) since climate variables were applied only in the mixed model and the second Bayesian analyses, for which the MalAvi data were not employed.

**Geographical Range**

In order to estimate the geographical range of each parasite lineage, we used the R package “GeoRange” and chose the variable minimum spanning tree distance (i.e. shortest total distance of all lines connecting each locality). Initially, using the function “create.matrix” from the “fossil” package in R, we created a matrix of species and localities. After generating the occurrence matrix, we filtered our data to account only for the parasites present in two or more localities and used the function “GeoRange\_MultiTaxa” to calculate the minimum spanning tree distance for each parasite lineage.

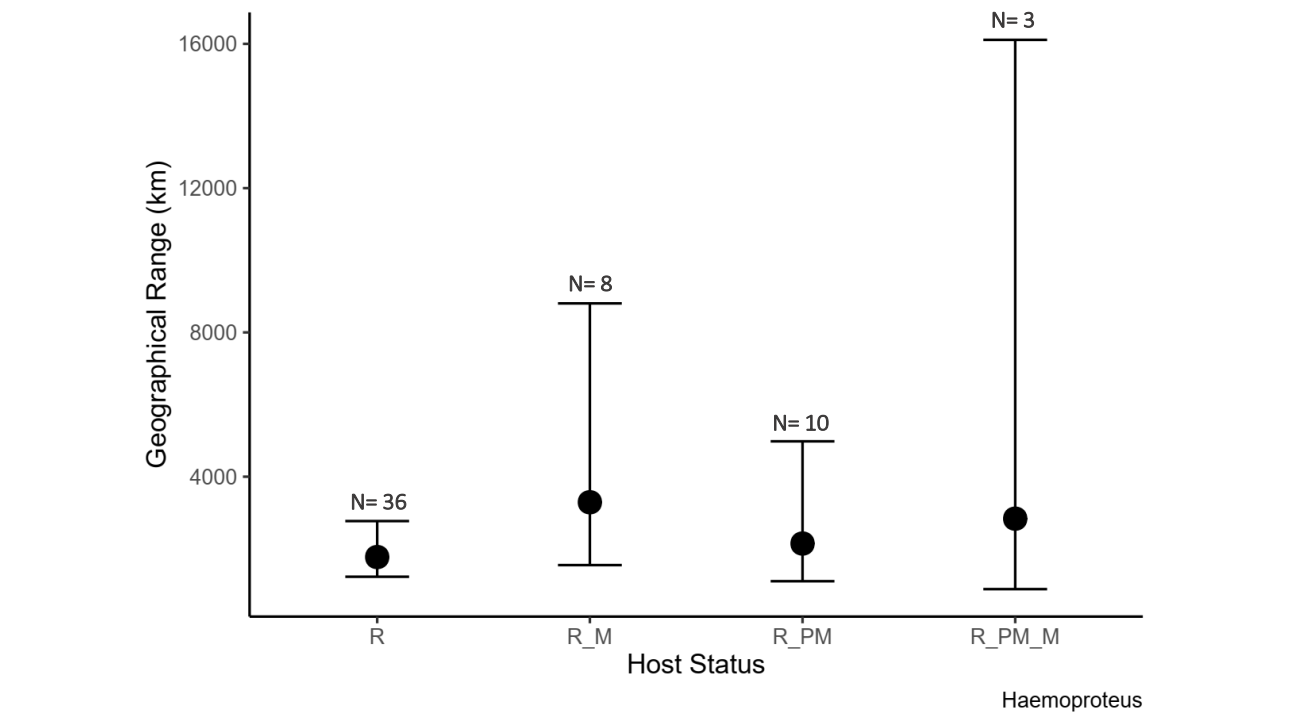
**Supplementary Results**



Supplementary Figure 1: Mean (±confidence intervals) geographical range in kilometers in which *Plasmodium* lineages are detected according to the type of birds in which they are found. R = resident, R\_M = resident and full migratory, R\_PM = resident and partial migratory and R\_PM\_M = resident, partial migratory and full migratory. Number of lineages in each of the four categories are shown on the graph.

Supplementary Table 2: Parameter estimates, standard errors, and confidence intervals for the Bayesian model testing the differences in the geographical range of *Plasmodium* lineages among those that occur in migratory and/or resident avian host species. (Residents only = reference category)

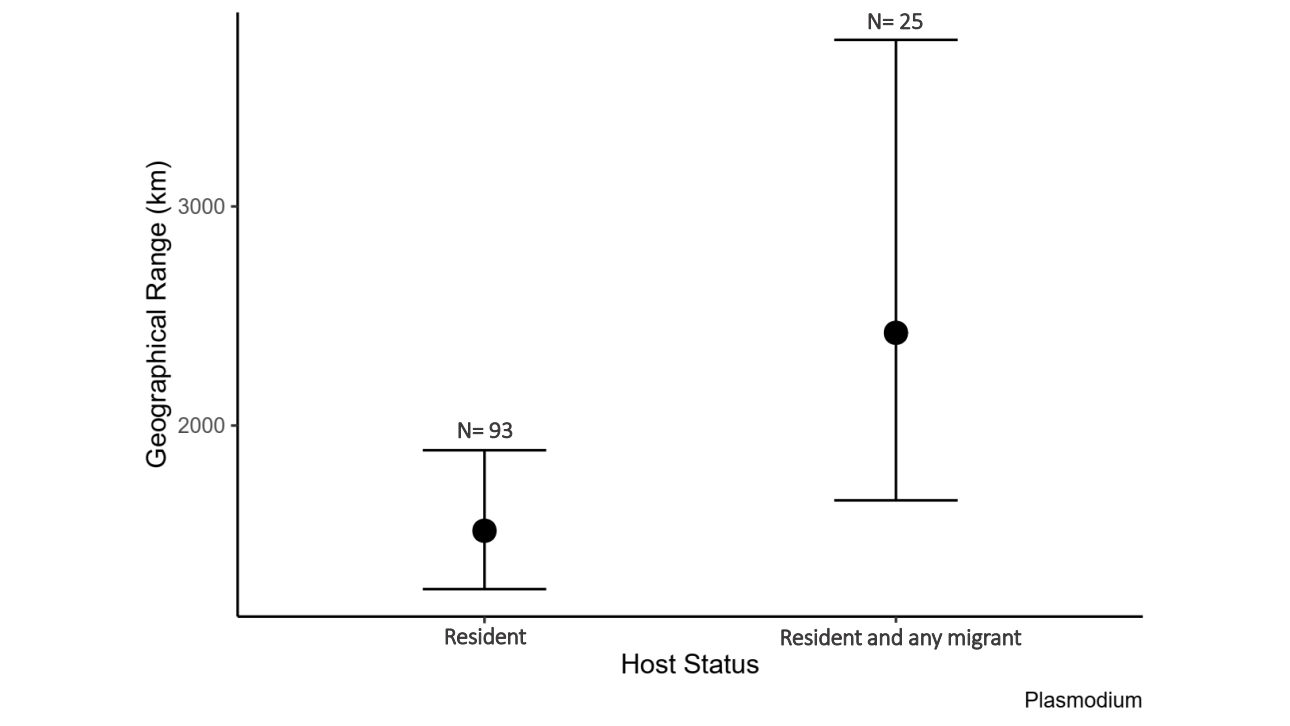
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Estimate** | **Std. error** | **Conf. Inter (95%)** | |
| Intercept | 6.96 | 0.14 | 6.69 | 7.23 |
| Resident and full migrant | 0.52 | 0.40 | -0.20 | 1.39 |
| Resident and partial migrant | 0.47 | 0.32 | -0.13 | 1.13 |
| Resident, partial and full migrant | 0.57 | 0.49 | -0.32 | 1.60 |
| Number of bird individuals | -0.02 | 0.02 | -0.06 | 0.03 |
| Number of host species per lineage | 0.11 | 0.04 | 0.02 | 0.19 |



Supplementary Figure 2: Mean (±confidence intervals) geographical range in kilometers in which *Haemoproteus* lineages are detected according to the type of birds in which they are found. R = resident, R\_M = resident and full migratory, R\_PM = resident and partial migratory and R\_PM\_M = resident, partial migratory and full migratory. Number of lineages in each of the four categories are shown on the graph.

Supplementary Table 3: Parameter estimates, standard errors, and confidence intervals for the Bayesian model testing the differences in the geographical range of *Haemoproteus* lineages among those that occur in migratory and/or resident avian host species. (Residents only = reference category)

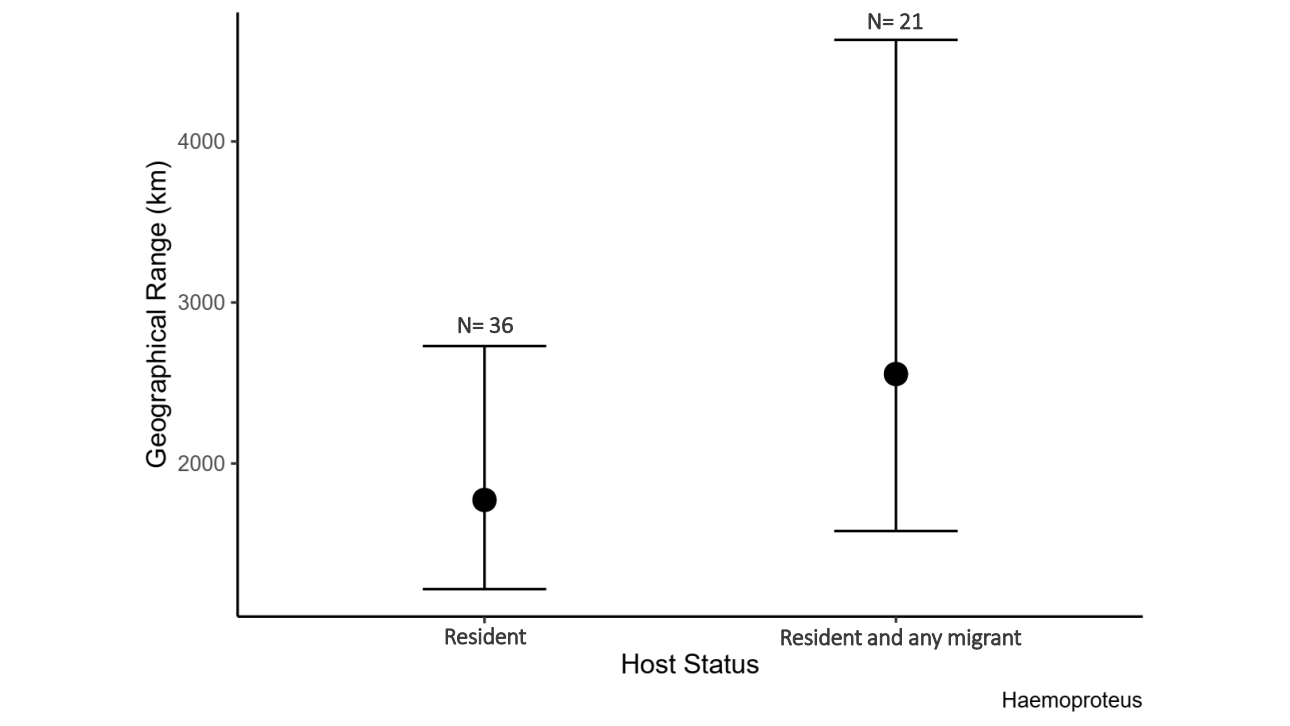
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Estimate** | **Std. error** | **Conf. Inter (95%)** | |
| Intercept | 7.35 | 0.23 | 6.91 | 7.80 |
| Resident and full migrant | 0.64 | 0.50 | -0.27 | 1.70 |
| Resident and partial migrant | 0.21 | 0.45 | -0.65 | 1.13 |
| Resident, partial and full migrant | 0.57 | 0.79 | -0.81 | 2.24 |
| Number of bird individuals | 0.01 | 0.02 | -0.02 | 0.06 |
| Number of host species per lineage | 0.01 | 0.04 | -0.07 | 0.10 |



Supplementary Figure 3: Mean (±confidence intervals) geographical range in kilometers in which *Plasmodium* lineages are detected according to the type of birds in which they are found. Number of lineages in each of the two categories are shown on the graph.

Supplementary Table 4: Parameter estimates, standard errors, and confidence intervals for the Bayesian model testing the differences in the geographical range of *Plasmodium* lineages among those that occur in migratory and/or resident avian host species. (Residents only = reference category)

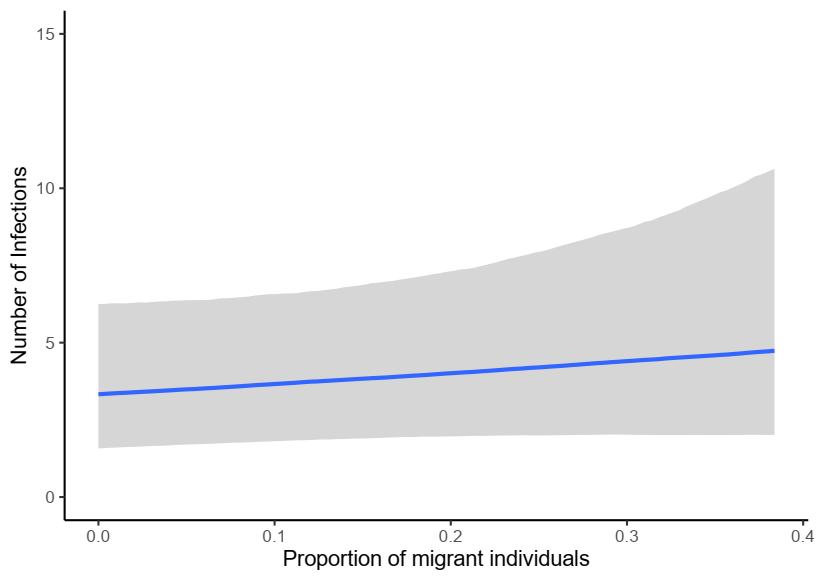
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Estimate** | **Std. error** | **Conf. Inter (95%)** | |
| Intercept | 6.95 | 0.13 | 6.71 | 7.21 |
| Resident and any migrant | 0.47 | 0.24 | 0.01 | 0.95 |
| Number of bird individuals | -0.02 | 0.02 | -0.06 | 0.03 |
| Number of host species per lineage | 0.11 | 0.04 | 0.02 | 0.19 |



Supplementary Figure 4: Mean (±confidence intervals) geographical range in kilometers in which *Haemoproteus* lineages are detected according to the type of birds in which they are found. Number of lineages in each of the two categories are shown on the graph.

Supplementary Table 5: Parameter estimates, standard errors, and confidence intervals for the Bayesian model testing the differences in the geographical range of *Haemoproteus* lineages among those that occur in migratory and/or resident avian host species. (Residents only = reference category)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Estimate** | **Std. error** | **Conf. Inter (95%)** | |
| Intercept | 7.35 | 0.22 | 6.92 | 7.80 |
| Resident and any migrant | 0.37 | 0.36 | -0.33 | 1.08 |
| Number of bird individuals | 0.01 | 0.02 | -0.03 | 0.06 |
| Number of host species per lineage | 0.01 | 0.04 | -0.06 | 0.10 |



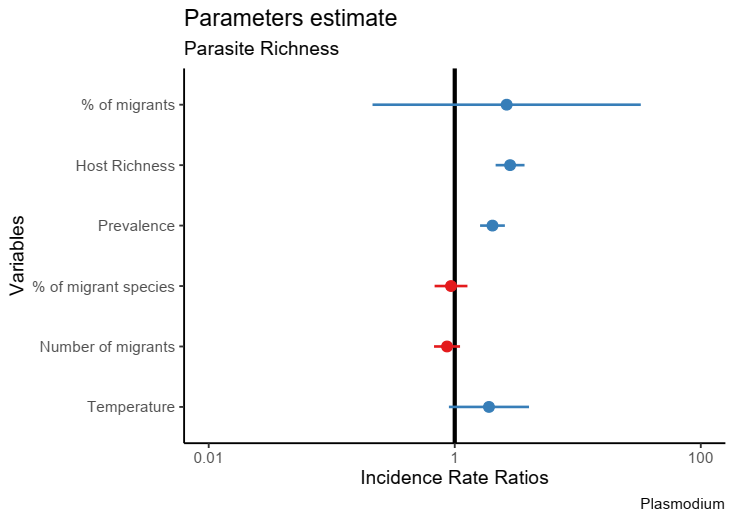
Supplementary Figure 5: : Predicted model relationship (±95% confidence intervals) between local number of bird individuals positive for haemosporidian and proportion of migratory host individuals per locality. We observed a no effect of the proportion of migratory individuals on number of infected birds.

Supplementary Table 6: Parameter estimates, standard errors, confidence intervals for the Bayesian model testing the variation of local number of birds infected by haemosporidian as a function of the proportion of migratory all individual birds sampled per locality and parasite richness.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Estimate** | **Std. error** | **Conf. Inter (95%)** | |
| Intercept | 0.38 | 0.40 | -0.44 | 1.13 |
| Proportion of migrant individuals | 0.91 | 1.05 | -1.09 | 3.03 |
| Parasite richness | 0.02 | 0.01 | 0.00 | 0.03 |

Supplementary table 7: AIC values for mixed models. Below all models tested are shown with all fixed effects and AIC test value.

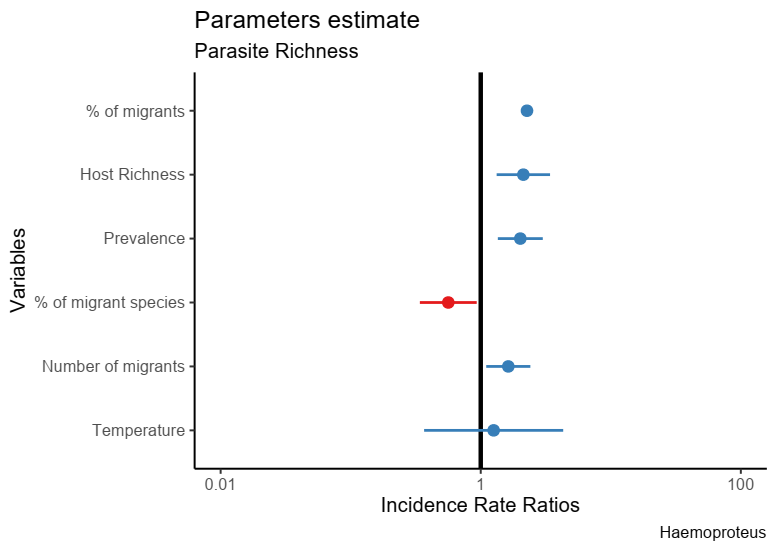
|  |  |  |  |
| --- | --- | --- | --- |
|  | **Variables** | **AIC value** | Δ**AIC** |
| **Model1** | No fixed variables | 470.20 | 79.32 |
| **Model2** | Host Richness | 437.94 | 47.06 |
| **Model3** | Prevalence | 449.75 | 58.87 |
| **Model4** | Host Richness + Prevalence | 399.67 | 8.79 |
| **Model5** | Proportion of Migrants | 470.85 | 79.97 |
| **Model6** | Host Richness + Prevalence + Proportion of Migrants | 397.50 | 6.62 |
| **Model7** | Number of Migrants | 458.68 | 67.8 |
| **Model8** | Host Richness + Prevalence + Proportion of Migrants + Number of Migrants | 399.08 | 8.2 |
| **Model9** | Temperature | 460.94 | 70.06 |
| **Model10** | Host Richness + Prevalence + Proportion of Migrants + Number of Migrants + Temperature | 390.88 | 0 |
| **Model11** | Precipitation | 460.43 | 69.55 |



Supplementary Figure 6: : Parameter estimates relating to their influence on *Plasmodium* richness. No correlation was found between the proportion of migratory individuals and *Plasmodium* richness.

Supplementary Table 8: Parameter estimates, standard errors, z and p values for the mixed model testing the variation of local *Plasmodium* richness as a function of the proportion of migratory individuals out of all individual birds sampled per locality, as well as other potential predictors.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Estimate** | **Std. error** | **Z** | **P** |
| Intercept | -6.92 | 2.06 | -3.35 | <0.001 |
| Proportion of migrant individuals | 0.97 | 1.28 | 0.76 | 0.45 |
| Host richness | 1.04 | 0.14 | 7.51 | <0.001 |
| Prevalence | 0.71 | 0.12 | 5.97 | <0.001 |
| Proportion of migrant species | -0.07 | 0.16 | -0.44 | 0.66 |
| Number of migrants | -0.14 | 0.12 | -1.16 | 0.24 |
| Temperature | 0.64 | 0.38 | 1.67 | 0.09 |



Supplementary Figure 7: : Parameter estimates relating to their influence on *Haemoproteus* richness. No correlation was found between the proportion of migratory individuals and *Haemoproteus* richness.

Supplementary Table 9: Parameter estimates, standard errors, z and p values for the mixed model testing the variation of local *Haemoproteus* richness as a function of the proportion of migratory individuals out of all individual birds sampled per locality, as well as other potential predictors.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Estimate** | **Std. error** | **Z** | **P** |
| Intercept | -5.09 | 3.34 | -1.52 | 0.13 |
| Proportion of migrant individuals | 0.82 | 2.07 | 0.40 | 0.69 |
| Host richness | 0.75 | 0.24 | 3.13 | 0.002 |
| Prevalence | 0.70 | 0.20 | 3.45 | <0.001 |
| Proportion of migrant species | -0.57 | 0.26 | -2.24 | 0.02 |
| Number of migrants | 0.49 | 0.20 | 2.45 | 0.01 |
| Temperature | 0.23 | 0.63 | 0.36 | 0.72 |

**References**

Bennett, D. J., Sutton, M. D., & Turvey, S. T. (2017). Treeman: An R package for efficient and intuitive manipulation of phylogenetic trees. *BMC Research Notes*, *10*(1), 1–10. https://doi.org/10.1186/s13104-016-2340-8

Clark, N. J., Drovetski, S. V., & Voelker, G. (2020). Robust geographical determinants of infection prevalence and a constrasting latitudinal diversity gradient for haemosporidian parasites in Western Palearctic birds. *Molecular Ecology*, 0–1. https://doi.org/10.1111/mec.15545

Illera, J. C., López, G., García-Padilla, L., & Moreno, Á. (2017). Factors governing the prevalence and richness of avian haemosporidian communities within and between temperate mountains. *PLoS ONE*, *12*(9), 1–22. https://doi.org/10.1371/journal.pone.0184587

Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D., … Webb, C. O. (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinformatics*, *26*(11), 1463–1464. https://doi.org/10.1093/bioinformatics/btq166

Mark Pagel. (1999). Inferring historical patterns of biological evolution. *Nature*, *401*(October 1999), 877–884.

Paradis, E., & Schliep, K. (2018). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, *35*, 526–528.

Revell, L. (2012). phytools: An R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol*, *3*, 217–223.