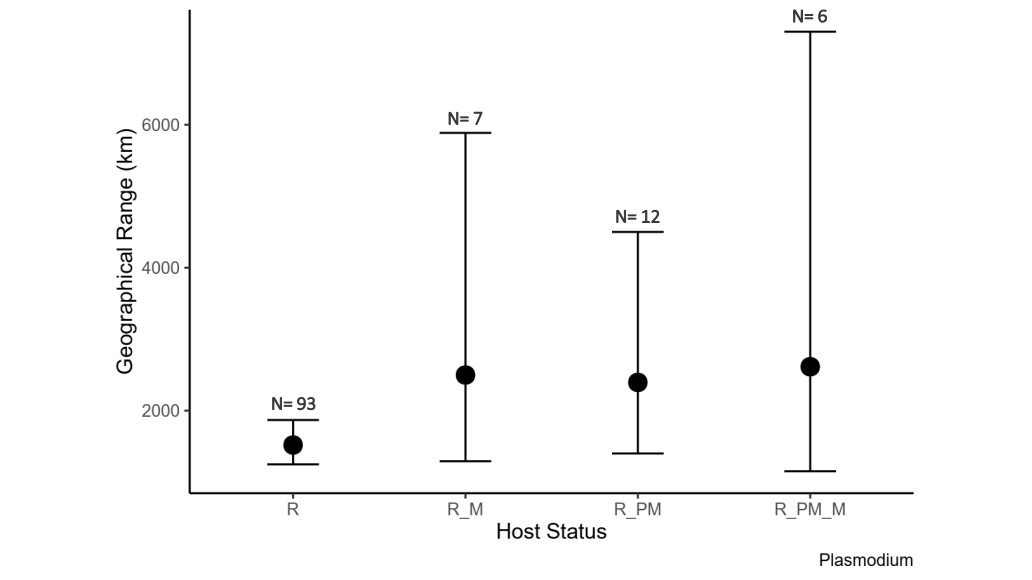
**Can migratory** **birds spread avian haemosporidian parasites?**

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

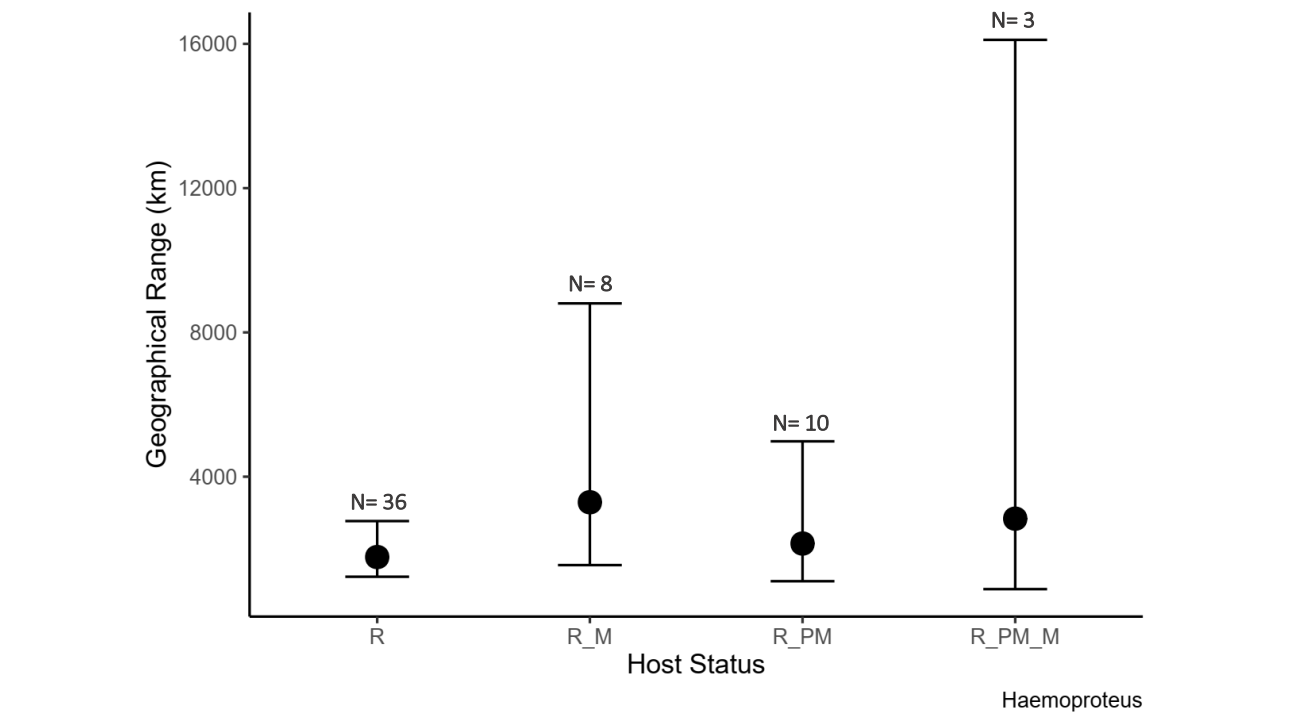
**SUPPLEMENTARY MATERIAL**



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.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **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.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **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 |

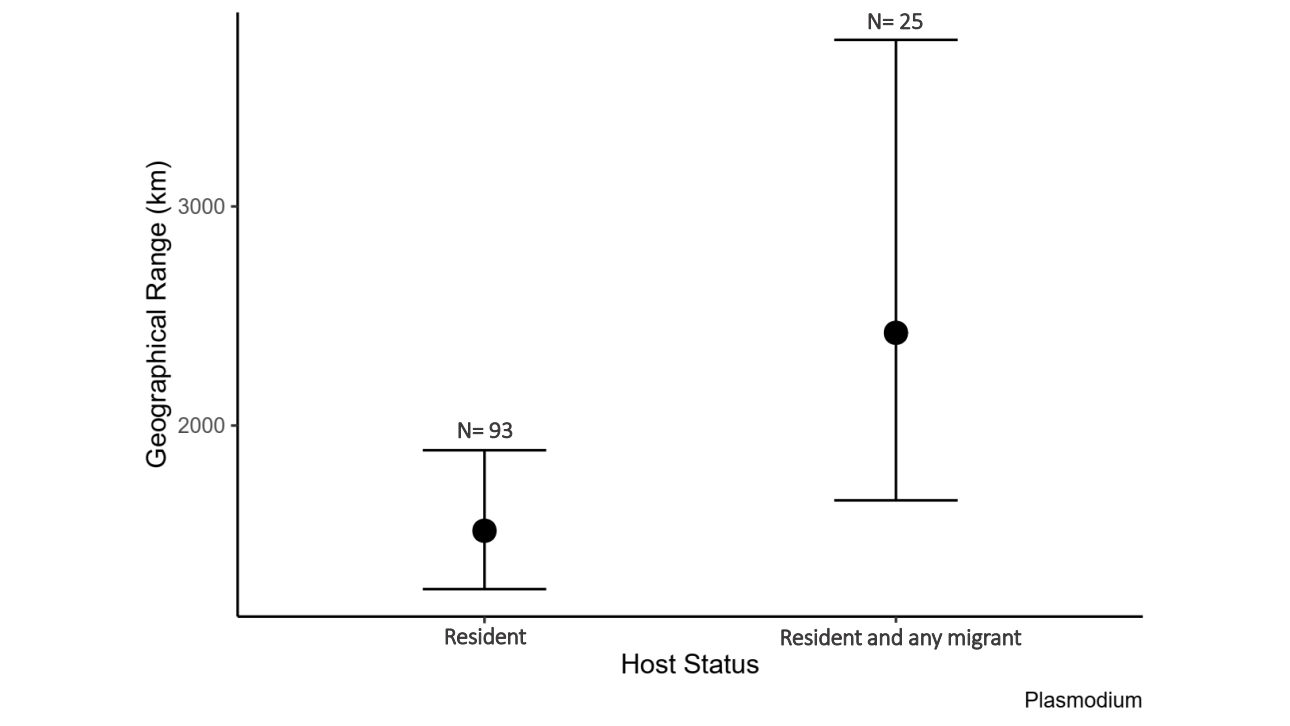


Figure 3: Mean (±confidence intervals) geographical range in kilometers in which *Plasmodium* lineages are detected according to the type of birds in which they are found. Number of lineages in each of the four categories are shown on the graph.

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.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **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 |

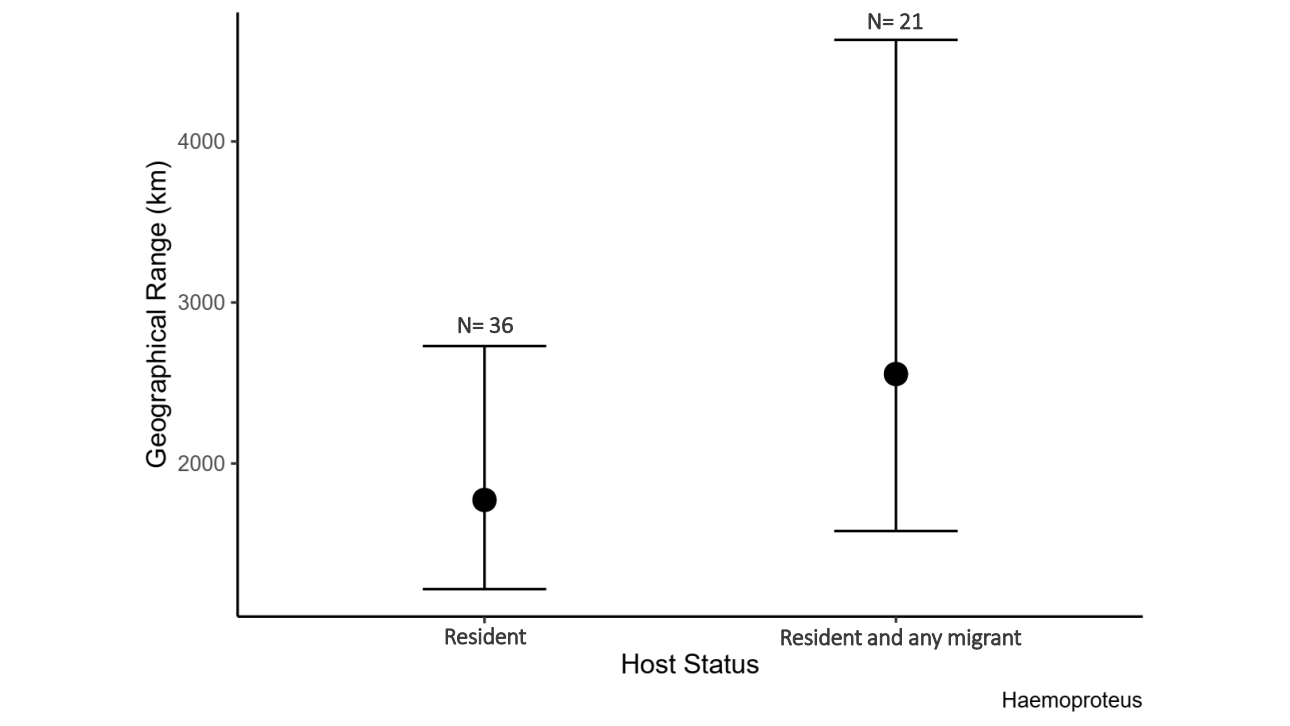
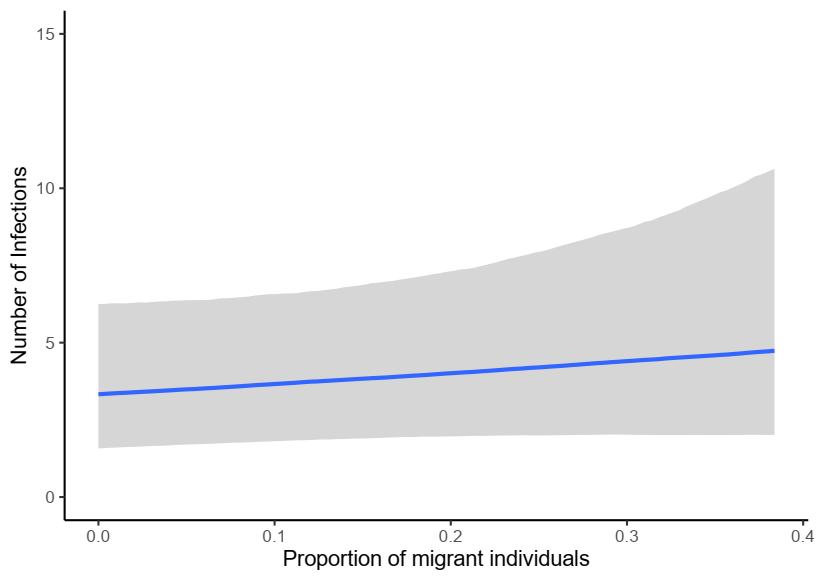


Figure S4: Mean (±confidence intervals) geographical range in kilometers in which *Haemoproteus* lineages are detected according to the type of birds in which they are found. Number of lineages in each of the four categories are shown on the graph.

Table 2: 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.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **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 |



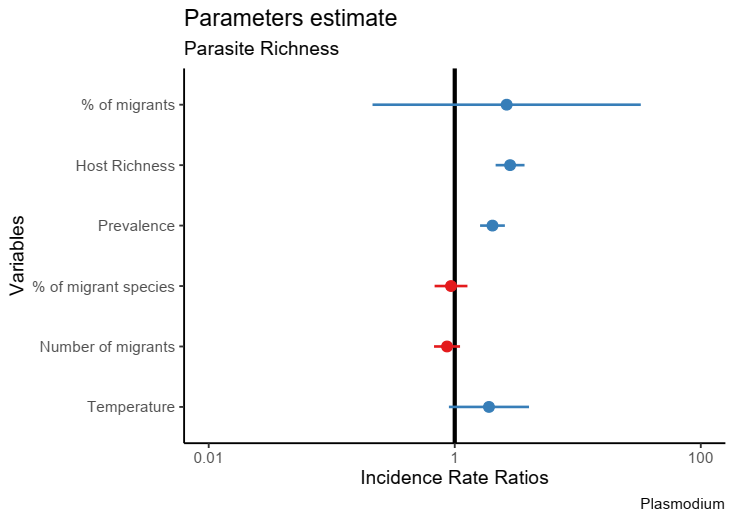
Supplementary Figure 5: Correlation between local number of bird individuals positive for haemosporidian and proportion of migratory host individuals per locality. We observed a no effect of the proportion of migratory individuals on parasite prevalence.

Supplementary Table 6: Parameter estimates, standard errors, confidence intervals for the Bayesian model testing the variation of local haemosporidian prevalence per species 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 model. Below all models tested are shown with all fixed effects and AIC test value.

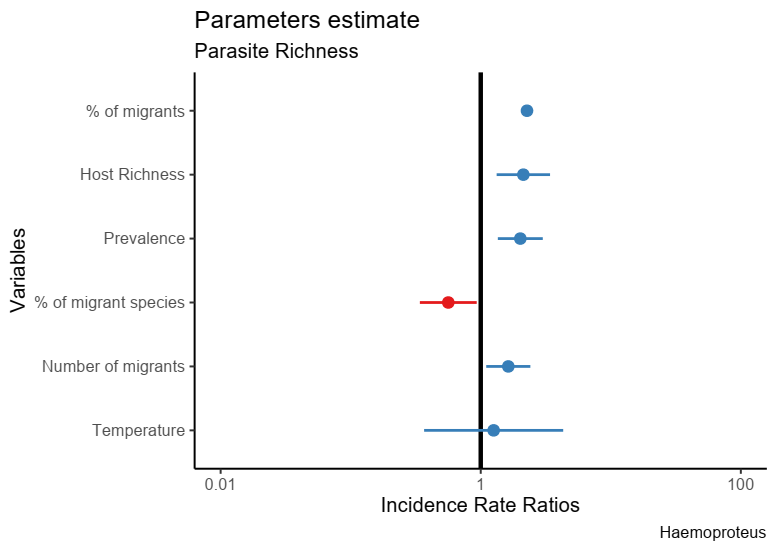
|  |  |  |  |
| --- | --- | --- | --- |
|  | **Variables** | **AIC value** | Δ**AIC** |
| **Model1** | No fixed variables | 470.20 | 79.32 |
| **Model2** | Host Richness | 437.94 | 47.06 |
| **Model3** | Prevalence | 449.75 | 58.87 |
| **Model4** | Host Richness + Prevalence | 399.67 | 8.79 |
| **Model5** | 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: Standardized estimates as a function of *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: Standardized estimates as a function of *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 |

**Potential correlates of prevalence and richness**

*Spatial autocorrelation*

All analyses were conducted in R version 4.0 (R Core Team, 2019). Aiming to evaluate the potential impact of locality on our samples, 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 varies between -1.0 and +1.0, being 0 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 of more phylogenetically similar species to resemble 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 (Mark 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 mixed model and the second Bayesian analyses, for which the MalAvi data were not employed.

**Geographical Range**

Aiming to estimate the geographical range of the parasite lineages we used the R package “GeoRange” and choose the variable minimum spanning tree distance (i.e. total distance of all lines connecting each locality minimized) to account for it in our Bayesian models. Initially, using the function “create.matrix” from “fossil” package in R, we created a matrix of species and localities. After we obtained 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.

**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.