2.3 Statistical Analyses

The spatial and temporal autocorrelation analyses revealed there was no substantial effect of time or space on parasite richness, however, for prevalence, we observed a Moran Index effect of 0.15, and for this reason, locality was used as a random effect in our second mixed model to control for idiosyncratic characteristics of localities. Likewise, considerable phylogenetic signals were observed among bird species for prevalence (0.49) and parasite richness (0.17). Considering this, phylogenetic covariation was added in Bayesian analyses and we analysed the prevalence using species as a fixed factor in the second mixed model.

*Bayesian model*

In order to determine whether migratory birds spread parasite lineages along their migratory routes and to evaluate the parasite connectivity among localities due to migratory behavior, we used multi-level modeling (MLM) with the “brms” package (Bürkner 2017) to evaluate the percentage of localities in which haemosporidian lineages occurred depending on whether they were found only in resident birds, only in partial migrant and fully migrant birds, or in both residents and migrants. We decided to use this approach as it allows to control for host phylogeny and to statistically estimate the percentage of localities among which lineages are distributed according to their host status.

Firstly, using the “ape” package (Paradis and Schliep 2018), we computed the phylogenetic expected variances and covariances from our bird species and incorporated this to control for phylogenetic effects in our Bayesian model. Secondly, we applied the “get\_priors” function to fit the priors for our model. We considered as independent and dependent variables bird migratory categories and percentage of localities in which each lineage was present, respectively, and res as reference level. As our Moran Index value for spatial autocorrelation of parasite richness among localities was low (-0.0008), we did not consider locality as a variable in our model and also did not use model correction for locality coordinates. Thus, we ran the model applying the “Beta” family, 4 chains with 2000 total iterations per chain and 1000 of warmup interactions. The model results were plotted using the “conditional\_effects” function to visualize the predictions of the population-level effects. We ran three models: one for all three parasite genera combined, one for *Plasmodium* lineages only, and one for *Haemoproteus* lineages only. Finally, we ran the “loo\_model\_weights“ to account for the effect of host richness (here, meaning the number of host species infected by each haemosporidian lineage) and number of hosts infected per lineage in our dataset.

*Mixed models*

Two mixed models were performed to estimate whether localities with more migratory birds have greater prevalence and richness of haemosporidian lineages. We employed the “glmer” function from the “lme4” package (Bates et al. 2015) applying Poisson and binomial distributions, respectively. In both models, we firstly created previous models including all variables that presented significant correlation with our dependent variable, and then selected the best model among them using “AIC” function in R. We tested the following variables as fixed factors: local host richness, local parasite richness, local prevalence across all birds sampled, local number of migrant species (based on birds caught), local number of migrant individuals, local temperature and precipitation.

In the first model, we considered parasite richness as the dependent variable and percentage of migratory bird individuals (i.e., percentage of migratory individuals out of all individual birds sampled in a locality) as the independent variable. According to our previous analyses, we employed local host richness (i.e., number of bird species sampled per locality), prevalence across all birds sampled, percentage of migratory species and number of migrant individuals as fixed variables. Biome was set as random variable. In this model, we did not use data from the MalAvi database, but only our dataset described above since it provides more information regarding the localities, such as prevalence data and host richness. We ran three models: one for all three parasite genera combined, one for *Plasmodium* lineages only, and one for *Haemoproteus* lineages only.

In the second model, we analysed the prevalence of infection in each bird species among localities. For this, we considered local prevalence in each bird species as our dependent variable and local percentage of migratory bird individuals as our independent variable. Following our previous analyses, only temperature was retained as a fixed factor. Further, we considered locality as a random variable. In this model, we filtered our data in order to include only species with 10 or more bird individuals analysed. For this second model we again used only our dataset described above and excluded data from the MalAvi database, since the latter presents only positive and sequenced samples. Again, we ran three models: one for all three parasite genera combined, one for *Plasmodium* lineages only, and one for *Haemoproteus* lineages only.

3. Results

Our Bayesian model analyses revealed the lineages shared by resident and migratory species are the most widespread spatially, as they are found in a higher percentage of localities (Figure 2, Table 1). When considering all haemosporidian genera together, we observed that the lineages shared by all three categories (resident, partial migrant and full migrant) are the most widespread, followed by those shared between residents and either type of migratory species. Nevertheless, despite the fact lineages shared by migratory species and residents are more widely distributed, lineages present in only residents, migratory or partially migratory species presented similar spatial distribution according to our model. We also observed that host richness (number of host species infected by each haemosporidian lineage) had the highest weight in parasite dispersal, followed by the number of birds infected and then host status (0.807, 0.102, 0.091).

When repeating these analyses separately for the two main parasite genera, we observed differences in the pattern of distribution between *Plasmodium* and *Haemoproteus*. For *Plasmodium* parasites, we observed a much greater spatial distribution of lineages shared by all three host categories, followed by the lineages shared by migrant or partial migrant and residents (Figure 3, Table 2). *Plasmodium* spp. lineages occurring in the three bird categories were present in 12.6% (SE = ±1.2%) of localities, a much higher value than for other lineages. However, for *Haemoproteus* lineages, we observed greater spatial distribution of lineages shared only by migrant or partial migrant and resident birds. The lineages shared by all three bird categories and those occurring in only one bird category had similar distributions among localities (Figure 4, Table 3).

Our mixed models examining the influence of migrants on local parasite richness and prevalence of infections also revealed differences depending on whether we considered both haemosporidian genera together or separately. Our first null model revealed that there is a positive correlation between the percentage of migratory bird individuals per locality and local parasite richness (p = 0.002, Figure 5, Table 4). We also observed a positive effect of the percentage of migratory bird individuals on local parasite richness for *Plasmodium*, but not for *Haemoproteus* infections when the two genera were treated separately (p = 0.004, p = 0.15, respectively; Figure S1 and S2, Table S1 and S2). Nevertheless, in all models we observed significant effects on parasite richness of the other two predictors: local host richness and overall local prevalence.

For the second model, in which we analysed the relationship between local prevalence per bird species and local percentage of migratory bird individuals, we observed a positive correlation between the relative occurrence of migrants and prevalence of haemosporidian parasites per species (p = 0.03, Figure 6, Table 5). However, when we repeated the analysis separately for only *Plasmodium* or *Haemoproteus* lineages, we observed no relation between local percent of migrants and prevalence per host species (p = 0.26, p = 0.65, Figure S3 and S4, Table S3 and S4). Temperature had no significant effect on prevalence per bird species, whether when considering all haemosporidian lineages (Table 5), or only *Plasmodium* or *Haemoproteus* lineages (Tables S3 and S4).