Oecologia (2020) 194:723–733
https://doi.org/10.1007/s00442-020-04782-x

ECOSYSTEM ECOLOGY – ORIGINAL RESEARCH

Large-scale disease patterns explained by climatic seasonality and host traits

Antoine Filion[1] · Alan Eriksson[2] · Fátima Jorge[1] · Chris N. Niebuhr[3] · Robert Poulin[1]

Received: 28 January 2020 / Accepted: 9 October 2020 / Published online: 24 October 2020

Abstract

Understanding factors affecting the distribution of vector-borne diseases in space and across species is of prime importance to conservation ecologists. Identifying the underlying patterns of disease requires a perspective encompassing large spatial scales. However, few studies have investigated disease ecology from a macroecological perspective. Hence, we use a global disease database to uncover worldwide infection patterns using avian malaria (*Plasmodium*) as a model for vector-borne disease transmission. Using data on 678 bird species from 442 locations, we show that environmental variables likely to synchronize bird and vector abundance are the key factors dictating infection risk for birds. Moreover, direct effects of host traits on exposure risk as well as potential trade-offs in resource allocation were also shown to affect disease susceptibility, with larger bird species being more prone to infection. Our results suggest that considering evolutionary strategies and factors influencing spatial overlap between hosts and vectors is crucial for understanding worldwide patterns of disease transmission success.

Keywords Abiotic regulation · Avian malaria · Co-evolutionary interplay · Comparative method · Disease macroecology · Host susceptibility · *Plasmodium*

Introduction

Diseases act as a major selective force regulating wildlife populations, exerting pressure at both individual and population levels by influencing reproduction, survival, and/or dispersal of susceptible individuals (Anderson and May 1978; Scott 1988; Hudson and Greenman 1998; Tompkins et al. 2011; McDonald et al. 2017). In the context of current efforts to protect wildlife, there is an urgent need to identify the underlying mechanisms that regulate disease occurrence in wildlife worldwide. Macroecology aims to understand large-scale patterns arising from complex mechanistic ecological processes, with the spatial distribution of organisms being at the core of this approach (Keith et al. 2012), thus providing an ideal framework to investigate large-scale patterns of disease occurrence. Indeed, by analysing global trends and factors associated with the distribution of diseases worldwide, disease macroecology can make it possible to map disease hotspots and identify their drivers, making it an important tool for conservation ecologists (Stephens et al. 2016).

Vector-borne diseases, which rely on invertebrate vectors for transmission to vertebrate hosts (WHO 2017), are an ideal model to investigate disease patterns in the macroecological perspective. Environmental characteristics affecting host and/or pathogen ecology can have direct or indirect effects on the spatial dynamics of vector-borne diseases by modifying transmission and exposure patterns (Shrag and Wiener 1995). Since vector-borne diseases are often fatal to naïve populations (i.e. populations that have not evolved in the presence of a particular disease; Tompkins and Poulin

Communicated by Indrikis Krams.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00442-020-04782-x) contains supplementary material, which is available to authorized users.

✉ Antoine Filion
afilion90@gmail.com

[1] Department of Zoology, University of Otago, 340 Great King St, Dunedin 9016, New Zealand

[2] Programa de Pós-Graduação em Ecologia e Conservação, Universidade Federal de Mato Grosso do Sul, Campo Grande, Mato Grosso do Sul 79070-900, Brazil

[3] Manaaki Whenua – Landcare Research, PO Box 69040, Lincoln 7640, New Zealand

2006; Tompkins et al. 2015), it is therefore important to unravel the main processes affecting the balance between vector-borne disease pathogens and their natural hosts. The processes impacting host and vector abundance and infection prevalence can be summarized into three general categories: (1) intrinsic host exposure to, and susceptibility towards, a pathogen, (2) encounter rates between vectors and hosts modulated by landscape attributes, and (3) regulation of vector abundance or activity by the environment (see Fig. 1). However, it is still unclear whether those processes apply at larger scales to multihost-species systems.

Firstly, host susceptibility towards a pathogen results from a long co-evolutionary process involving complex interactions between the host and the pathogen, which is in large part shaped by host traits. Hosts co-evolving with a disease face a trade-off as resources must be allocated to both immunity and other key functions such as reproduction (Sheldon and Verhulst 1996, Duffield et al. 2017), creating an evolutionary continuum of resource allocation between evolutionary immunocompetence (i.e. ability to minimize fitness costs of an infection; Owens and Wilson 1999) and reproduction. However, this relationship might not be universal among species (Martin et al. 2001). Since ecological and physiological costs are related to immunocompetence (e.g. fighting off a disease can result in a decrease in food intake) (Zuk and Stoehr 2002), host susceptibility in the context of immunocompetence can be measured using specific host traits affecting the trade-off. For example, in populations of great tits (*Parus major*), breeding individuals have a weaker immune response to pathogenic agents compared to non-breeding individuals (Deerenberg et al. 1997), suggesting that reproduction is a key trait in gauging immunocompetence. Similar associations have been reported in multiple populations of the same species (e.g. Duffy et al. 2012), but also among individuals of the same population (see review by Zuk and Stoehr 2002). Another key trait that can relate to both exposure to disease and immunocompetence is body size (Freeland 1983; Moore and Wilson 2002). For instance, all else being equal, in interspecific comparisons larger-bodied species incur greater parasitic infections (Kamiya et al. 2014), while within host species, larger individuals provide greater surface areas for vector attachment and feeding, produce more cues, like CO_2, which can attract vectors and tend to have greater longevity, which exposes them to increased chances of vector encounter (Edman and Scott 1987; Torr et al. 2006). Thus, both inter- and intraspecifically, host traits related to body size and reproductive strategies can play roles in determining exposure and immunity to disease.

Secondly, encounter rates between a vector and its host are also modulated by geographical and landscape factors that influence host–disease interactions. Indeed, two species must overlap spatially to interact (Keeling 1999). A suitable host found outside the geographical range of a disease and therefore a lack of contact between disease and host can

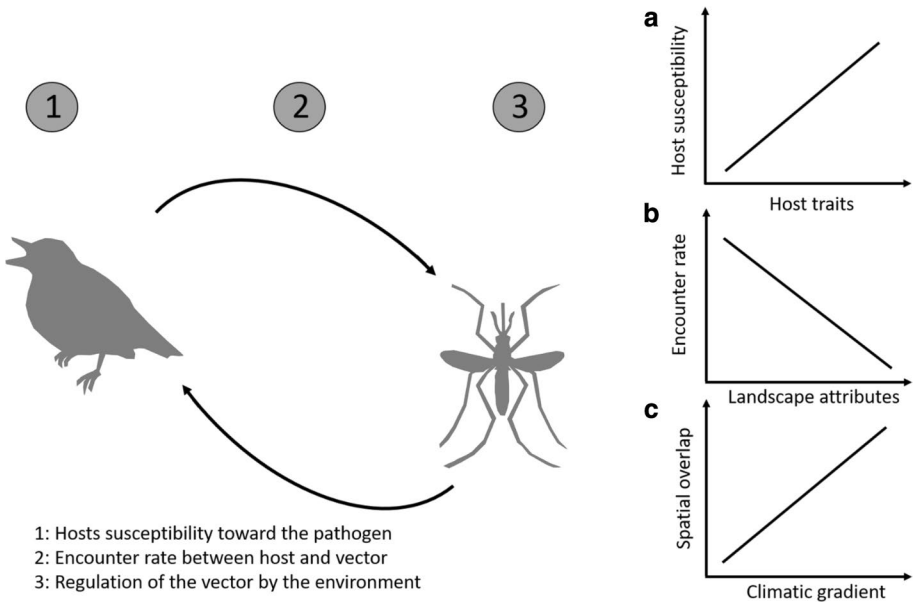

Fig. 1 Potential main processes, along with the expected directions of their effects, underlying vector-borne disease transmission in wildlife: **a** increasing value of hosts traits, such as body size and clutch size, should increase host susceptibility towards a disease according to size-dependent exposure and immunocompetence trade-offs; **b** Increasing elevation and landscape/vegetation heterogeneity should result in a lower probability of encounter rate between the vector and potential hosts of avian malaria; and **c** increasing values of climatic variables, such as temperature and precipitation, should increase the probability of spatial overlap between a vector and its host, increasing disease transmission

limit disease dispersal, thus creating a spatial patchwork of hot and cold spots where both species co-occur (Thompson and Cunningham 2002). Two main mechanisms can explain this spatial overlap: landscape attributes and climatic filtering. These bring into play the third category of processes determining prevalence of disease, that is, the regulation of vectors by the environment. On the one hand, landscape composition is a well-known regulator of spatial co-occurrence of species because it affects metacommunity dynamics (Giraudoux et al. 2003). Geographical barriers across the landscape can influence the rate of encounter between hosts and pathogens in the environment, therefore altering infection prevalence between populations (Smith et al. 2002; Reisen 2010). Moreover, by affecting spatial processes of diseases, vegetation density is known to affect vector-borne diseases by regulating presence or abundance of vectors and hosts (Woolhouse et al. 1997, Ferraguti et al. 2016).

On the other hand, climatic filtering can modify the exposure patterns between host and vector. The climatic niche acts as a spatial filter, limiting the dispersal of vector-borne diseases by restricting the development of the vector or by promoting the transmission of the pathogen through climatic conditions favourable to the vector species (Reisen 2010). For example, cases of Saint-Louis encephalitis virus are synchronized with temperature and rainfall, which promote not only the emergence and oviposition of the mosquito vector, but also their active transmission of the disease towards its definitive host (Day et al. 1990; Reeves et al. 1994; Day 2001). Moreover, stability of the environment is known to influence disease patterns. Indeed, wild populations in seasonally stable areas, such as the tropics, are more prone to vector-borne disease because of faster disease development rate and contact with suitable hosts all over the year (Lindgren et al. 2000; Altizer et al. 2006; Ogden et al. 2006). Therefore, climatic conditions and seasonality can act as important drivers of disease prevalence in natural systems.

Certain diseases have received widespread attention from researchers due to their impacts on wild populations. These provide ideal model systems for disease ecologists to test the general processes described above. One such system is avian malaria (genus *Plasmodium*), a vector-borne disease affecting a wide range of bird species worldwide, except in Antarctica (Grimaldi et al. 2015). Extensive global data are available on the occurrence of this disease. *Plasmodium* parasites use a mosquito (family Culicidae) vector and a bird host to complete their life cycle (Valkiūnas 2005; Fig. 1). Avian malaria can have huge impacts on bird communities. For example, naïve populations of native birds in Hawaii have either gone extinct or shifted their distribution to higher elevations in order to avoid encountering introduced mosquito vectors of *Plasmodium* (Atkinson and Lapointe 2009; Atkinson and Samuel 2010). Numerous studies have investigated the effects of multiple drivers of avian malaria and their potential interplay in natural systems, albeit only at local and regional scales. At those limited scales, climatic drivers that influence vector abundance and parasite development such as temperature (Loiseau et al. 2013), rainfall (Galardo et al. 2009) and seasonality of climate (Pascual et al. 2009), as well as host traits such as body size (Scheuerlein and Ricklefs 2004; Santiago-Alarcon et al. 2015) and reproductive output (Podmokla et al. 2014), have been linked with susceptibility to malaria. However, previous studies have shown that host susceptibility may in part or totally be influenced by a species effect. Indeed, macro-ecological co-evolutionary history of host–parasite interactions has led to increased susceptibility in some host species, resulting from parasite phylogenetic restrictions towards certain hosts (Pulgarin-R et al. 2018) due to species-specific intrinsic factors (Medeiros et al. 2013; Ellison et al. 2015). Species-specific susceptibility of certain hosts to disease can cause subsequent changes in community assembly (Holt 1977; Holt and Bonsall 2017). Recently, Barrow et al. (2019) showed that the influence of some host traits, climatic and landscape factors on avian malaria infection across bird species in the Andes was overshadowed by a strong phylogenetic signal, indicating that species in some avian families are intrinsically more susceptible to infection. It is therefore essential to disentangle the effect of phylogenetic conservatism from those of other drivers in interspecific comparative analyses of susceptibility to disease. However, it remains unclear whether any of these factors operates at higher taxonomic and spatial levels, since these findings have yet to be extended to more complex, and truly global-scale systems.

Only a few extensive datasets allow for large-scale investigations of disease patterns (Stephens et al. 2016), thus the link between macroecology and disease ecology awaits empirical strengthening. Here, we take advantage of an open access, large-scale dataset, the MalAvi database, which contains information on avian malaria infections in wild bird populations (Bensch et al. 2009), to shed light on infectious disease macroecology and to test whether local avian malaria patterns are similar to those observed on a truly global scale. We predict that: (1) bird species with a larger clutch size are more susceptible to avian malaria due to a trade-off with immunocompetence; (2) bird species with larger body sizes have a higher prevalence of avian malaria due to greater exposure to the vectors; (3) populations of birds at lower elevations have a higher prevalence of infection due to greater encounter rates with vectors; and (4) areas with higher rainfall, temperature, landscape/vegetation homogeneity and climate stability (e.g. with low seasonality) are associated with higher malaria prevalence as these conditions should increase mosquito abundance and disease transmission (see Fig. 1). We test these predictions using a phylogenetically controlled comparative analysis comprising 678 bird species, and uncover patterns at a global scale

that are not always aligned with those reported at local or regional scales.

Methods

Data selection

We base this study on all articles present in the MalAvi database (n = 341 as of 22 October 2018; Bensch et al. 2009), which includes studies published between 2005 and 2018. We retained only studies that met the following selection criteria: (1) studies had to include at least one individual bird host infected by a lineage of *Plasmodium* and provide the exact number of individuals screened for each bird species; (2) studies had to have been conducted on wild bird individuals; and (3) studies had to provide coordinates of the site where sampling was conducted. This left us with a total of 136 articles. We then examined each remaining article and manually compiled data on the number of infected individuals and the total number of individuals examined, separately for each bird species and each site sampled per study; we also included host species that were screened but for which no infection was found. In the analyses that follow, each data point corresponds to the prevalence, i.e. the number of individuals infected by *Plasmodium* out of the total number examined, per population of a given species per site sampled. Geographic coordinates and the accuracy of bird species taxonomy were checked and corrected whenever appropriate.

The MalAvi database incorporates information on the number of individuals of a single species of the same population found infected by each lineage of *Plasmodium*, which means that the same population (e.g. having the same number of individuals tested) of the same species can appear twice (or more) for the same study, subsequently biasing real estimates of prevalence. We therefore combined each line in the dataset that had the same number of individuals screened for malaria per species per site, thus giving unbiased infected/sampled populations. Afterwards, we extracted mean annual temperature, mean annual precipitation, elevation, temperature seasonality and precipitation seasonality data at a resolution of 2.5° using the Worldclim database ranging from 1970 to 2000 (Hijmans et al. 2005) for each set of geographic coordinates (i.e. each sampled site) in the database (see Online Supplementary Material for range of variables). We then used the Amniote project database to incorporate the most relevant life history traits of bird hosts into our database (Myhrvold et al. 2015), namely body weight (hereafter referred to as body size) and clutch size (the number of eggs laid in a single brood). Afterwards, we downloaded NDVI index values (Normalized Difference Vegetation Index) as a measure of vegetation density. For this, data were obtained using "VIPPHEN_NDVI: Vegetation Index and Phenology (VIP) Phenology NDVI Yearly Global 0.05Deg CMG V004" provided by USGS/NASA (Didan and Barreto 2016). The yearly average NDVI values were then extracted for each sampling site from the year 2000 to 2014 using the following packages: *raster*, *rgdal*, *sp* and *gdalUtils* (Pebesma and Bivand 2005, Bivand et al. 2019, Hijmans and van Etten 2019, Greenberg and Mattiuzzi 2020) in the R computing environment (R core team 2018). We then used the mean of these values as a parameter in the subsequent analysis.

The initial dataset included potential sources of bias. First, the data entries were taxonomically unbalanced, with 86% representing passerines. Secondly, many (54%) local populations had a sample size of less than five individuals. These two factors could potentially impact the inferences made from statistical analyses. To account for this, data were partitioned into two data subsets, one including only passerines (hereafter referred to as the Passerine dataset) and the other one including all populations for which we had a sample size of at least ten individuals (hereafter referred to as the Reduced dataset). This threshold sample size provides a good balance between the number of species that could be retained in the Reduced dataset, and some confidence in the accuracy of the prevalence estimate (Gregory and Blackburn 1991). Overall, 678 bird species from 137 families sampled in 442 sites were included in the Complete dataset. The Passerine dataset consisted of 504 species from 85 families and 391 sites, whereas the Reduced dataset consisted of 216 species from 70 families sampled at 361 sites.

Data analyses

Firstly, we downloaded 1000 avian phylogenetic trees from BirdTree.org using the backbone tree from Hackett et al. (2008) and used a random sample of 100 trees to account for phylogenetic uncertainty. We then pruned the phylogenetic trees to include only the species found in our datasets using the ape package (Paradis and Schliep 2018). A covariance matrix was then constructed (MCMCglmm package; Hadfield 2010) to be used in the phylogenetically corrected models using the Bayesian statistics package brms (Bürkner 2017). To determine if malaria prevalence exhibits phylogenetic structure among bird species, we calculated the phylogenetic signal for Pagel's lambda (λ) using the *phylosig* function in the *phytools* package (Revell et al. 2012) by mapping the average prevalence across all sites for each species onto the phylogeny. We then used the *ContMap* function to visualize the phylogenetic influence on malaria prevalence.

To visualize the distribution and the potential collinearity between our population-level variables, and that of hosts traits with latitude, we used a pair plot correlation with the Pearson correlation (*Psych* package; Revelle 2018, see

Online material). We then based our collinearity decision on the threshold of $R^2 \geq 0.7$, as population-level effects above this threshold are known to produce biased estimates in modelling (Dormann et al. 2012). Using this indicator, we found multiple correlations exceeding this threshold. Overall, temperature seasonality, mean annual temperature, and mean annual precipitation were strongly correlated in the complete and the Passerine datasets, with temperature seasonality being also correlated with clutch size in those datasets; nevertheless, we decided to use all of them in the construction of the models, albeit never together, since they each capture different environmental processes affecting disease transmission.

To make sure that the results we observed were not dependent on site relatedness, we calculated the spatial autocorrelation signal of *Plasmodium* prevalence among our sites using the Moran I test. Values of Moran I vary between -1 and 1, with values close to 0 representing a weak spatial autocorrelation and more extreme values representing a stronger spatial autocorrelation signal. If the observed value is above the inverse of the number of locations in the dataset -1 (hereafter referred to as the expected value), then there is spatial autocorrelation in the dataset. In our case, the observed value of spatial autocorrelation of avian malaria prevalence among our sites was of 0.08 and the expected value was of 0.002. While the p value indicates that avian malaria prevalence is more clustered among sites than what would be expected by chance ($p<0.001$), the very low observed value indicates that spatial autocorrelation is weak. We therefore decided not to include a spatial autocorrelation structure in our models, but we took into account stochastic effects of geographical characteristics by including sampling location as a random effect in the models (since in many cases, multiple populations were sampled from the same location).

We used phylogenetic Bayesian multilevel modeling (MLM) (*brms* package; Bückner 2017) as those models can incorporate phylogenetic distance matrices as covariates into multilevel models and sample posterior distributions very efficiently using the Hamiltonian Monte Carlo (HMC) method with STAN algorithm. We incorporated both the sampling site and the species as group-level intercept factors since our data structure involved bird populations nested within a species itself nested within a site. We used MLM with zero-inflated negative binomial distribution to model the number of infected individuals for the larger datasets (Complete and Passerine). We selected this distribution to take into account the overdispersion of our data and to explain the overabundance of zeros (46% and 41% for the Complete dataset and the Passerine dataset, respectively), thus avoiding biased estimates of parameters and associated standard error (Zuur et al. 2009, 2010). We used a negative binomial distribution for the Reduced dataset to account for overdispersion in the data. Number of sampled individuals was included as a covariable in all models to account for sampling effort (Gregory and Blackburn 1991). A set of candidate models with increasing complexity was built using different combinations of population-level factors, including Body size (log-transformed scaled value), Clutch size (log-transformed scaled value), Temperature (scaled value), Precipitation (scaled value), Elevation (log-transformed scaled value), Temperature seasonality (scaled value), Precipitation seasonality (scaled value) and NDVI index (scaled value). We used log transformation to normalize population-level effects with non-normal distributions.

Model building and selection

Each model was built with priors obtained from the *get_prior* function in the *brms* package and with 2 chains of 4000 iterations (2000 for warm-up, 2000 for sampling). At first, we built simple models incorporating only one factor in addition to sampling effort. If the variable had a significant effect (e.g. 95% credible interval not crossing 0), we included it into more complex models, otherwise, we considered the variable as non-significant (see Online Supplementary material for ELPD score). We made sure every parameter in the model converged by checking the potential scale reduction factor on split chains (Rhat) indicator (at convergence, Rhat is equal to one). Afterwards, we used a leave-one-out measure to assess model fit relying on the expected log predictive density (ELPD) difference criterion as implemented in the *loo* package (Vehtari et al. 2017) to find out which models were competing. Thereafter, we excluded every model that had a ΔELPD difference over 4 compared to the best model. To ensure we selected the best model possible, we used model averaging to measure the weight of each competing model for each dataset using stacking and pseudo-BMA method for the LOO criterion (see Table 1). Since around 5% of our data had a pareto-k value over 0.7 and it was not possible to correctly fit all of the data with k-fold=10, we did not use the WAIC criterion as it is known to produce biased model weights (Vehtari et al. 2017).

Results

After Passeriformes (86%), the orders contributing the most species to the Complete dataset were Columbiformes (3%), Piciformes (2%) and Apodiformes (2%), with the remaining 9% of species distributed among 24 orders. Overall, 18% (6,938 out of 37,915) of birds examined were found to be infected by avian malaria in the Complete dataset, with similar proportions in the Passerine dataset (6,556 out of 32,507; 20%) and in the Reduced dataset

Table 1 Model averaging of models competing within four ELDP using stacking and pseudo-BMA weighting method

Dataset	Population level effects	Stacking weight	Pseudo-BMA+ weight
Complete (n=678)	Tested + Temperature seasonality + Body size	0.634	0.474
	Tested + Precipitation + Clutch size + Body size	0.366	0.245
Passerine (n=504)	Tested + Elevation	0.378	0.290
	Tested + Temperature seasonality + Body size	0.418	0.364
	Tested + Temperature seasonality + Body size + NDVI	0.204	0.345
Reduced (n=216)	Tested + Body size	1.0	0.737
	Tested + Temperature seasonality	0	0.207
	Tested + Mean precipitation	0	0.057

Sum of weights for each dataset and each method is equal to 1. Models with high weighting values have a higher probability of being rightfully selected as the most probable one. "Tested" population-level effect refers to the number of individuals sampled and tested for malaria presence per species per population

(6,026 out of 34,002; 18%). Among all our models, the effect size of the number of individuals sampled was 0.01 with a standard error (SE) of 0.00. The estimated best models were similar among all the datasets, with climatic drivers (mainly temperature seasonality) and body size being selected in almost every model as the best variables explaining variation of avian malaria worldwide.

Phylogeny effect versus phylogenetic structure

Overall, we show that there is a relatively weak phylogenetic signal in all our datasets, with values for the phylogenetic signal, λ, for the mean proportion of infected individuals per species across the phylogenetic trees not varying much between the three datasets. We only present the phylogenetic signal tree of the Reduced dataset (n=216) as a clear example (Fig. 2). We provide a reproductible example of our complete phylogenetic tree (n=678) as supplementary material (see Online supplementary material). The Passerine dataset has the largest signal (λ=0.274), followed by the Reduced dataset (λ=0.242) and the Complete dataset (λ=0.237).

Analysis of the complete dataset

The best model explaining the variation in the number of individuals infected by avian malaria in the Complete dataset included body size (effect size= 0.18, SE=0.09) and temperature seasonality (effect size=0.38, SE=0.08) (Fig. 3). No other model presented a plausible alternative to the selected model (LOO weights for stacking was over 0.6, pseudo-BMA+ weight close to 0.5; Table 1; see Online Supplementary material for ELPD score).

Analysis of the passerine dataset

The best model explaining the variation in the number of individuals infected by avian malaria in the Passerine dataset included body size (effect size=0.14, SE=0.07) and temperature seasonality (effect size=0.38, SE=0.08) (Fig. 3). The two other competing models included either elevation alone (elevation model: pseudo-BMA weighting=0.290, stacking weighting=0.378; effect size=− 0.03, SE=0.05) or the same selected variables together with NDVI (NDVI model: pseudo-BMA weighting=0.345, stacking weighting=0.204; effect size=- 0.08, SE=0.07). Nevertheless, as credible intervals for both effects were crossing zero, we decided not to include those parameters as plausible explanations of the processes regulating avian malaria transmission worldwide (Table 1; see Online Supplementary material for ELPD score).

Analysis of the reduced dataset

The best model explaining the variation in the number of individuals infected by avian malaria in the Reduced dataset included only body size (effect size=0.18, SE=0.15) (Fig. 3). However, most models were competing in the range of the 4 ELPD criterion, which makes model averaging difficult in this case. Nevertheless, Bayesian stacking method, useful when most models are competing in those ranges (Yao et al. 2018), always pointed towards body size being the most accurate predictor of avian malaria prevalence for this dataset. (Table 1; see Online Supplementary material for ELPD scores).

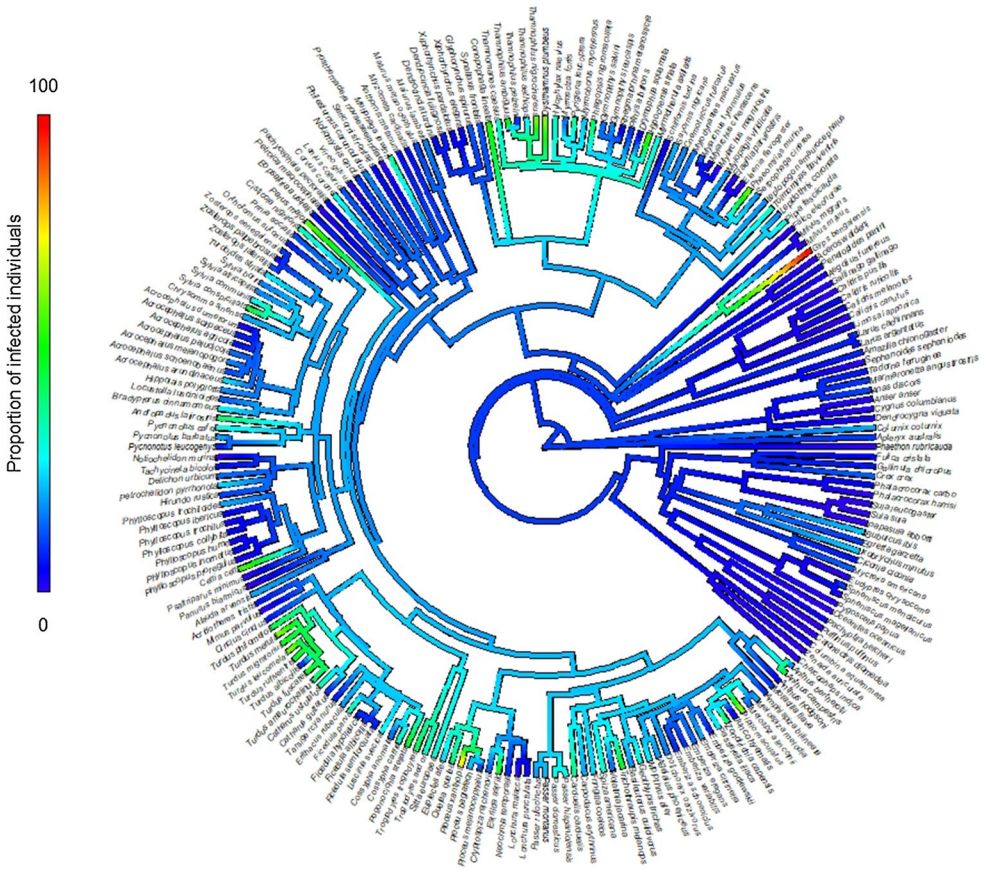

Fig. 2 Phylogenetic distribution of proportion of infected individuals across all bird species comprised in the Reduced dataset (n>10 individuals sampled) shown here for illustrative purpose. The tree includes 216 bird species

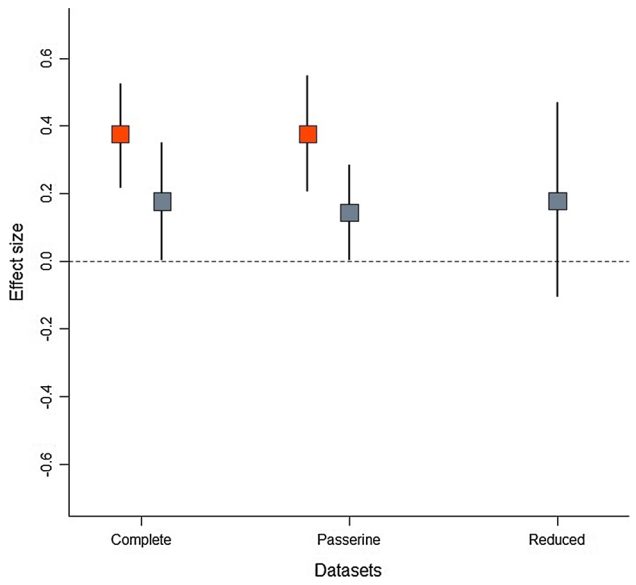

Fig. 3 Effect size and 95% credible intervals predicted for all factors selected in the best models explaining the number of infected bird individuals for the three datasets using the *brms* package (Bücknner 2017). Factors selected are Temperature seasonality (orange), and body size (grey). Variables are considered to have an effect in a model when their credible intervals do not overlap with the zero line

Discussion

Our study shows that the prevalence of *Plasmodium* in wild bird populations around the world is positively affected by both host traits, namely body size, and by temperature seasonality, all pointing in the same direction for two biggest datasets but remaining somewhat unclear for the most reduced one. In contrast, we found no evidence of an effect of precipitation seasonality, mean annual temperature, mean annual precipitation, elevation, clutch size and NDVI index on worldwide avian malaria prevalence. Our statistical approach accounted for host phylogeny, spatial stochasticity, and sample size. To the best of our knowledge, our results are therefore free of the influence of those factors.

Among the three datasets, the effect size of drivers selected with our model averaging approach remained constant, even with increasing model complexity when working on larger datasets. Temperature seasonality and host body size seemed to be the consistent drivers of *Plasmodium* infection, with model selection always pointing towards those variables in the three datasets. Lack of effect of clutch size on avian malaria prevalence could be explained by the high seasonality of this disease, with birds laying larger clutches being more infected at the beginning of the breeding period and subsequently dying off, therefore lowering

the probability of sampling of infected individuals. Moreover, the lack of effect of elevation could be a data artefact, as capturing birds at higher elevation usually requires more effort and yields lower sample sizes, thus leading to the limited influence of this parameter among our datasets. As for vegetation density, its lack of effect could also be due to a data artefact, with fewer populations sampled in areas with lower vegetation density, such as farmlands.

We found a relatively weak host phylogenetic signal, albeit roughly consistent (λ= ~ 0.24) across the three datasets. Interestingly, the strength of the phylogenetic signal we found is very similar to that reported in Barrow et al. (2019) when considering the three main genera of malaria (*Plasmodium*, *Haemoproteus* and *Leucocytozoon*). However, the same authors reported a value of λ=0.35 (95% CI 0.06–0.61) for *Plasmodium* only. The difference may be due to the extended taxonomic range of our study, which included about 200 more bird species than in Barrows et al. (2019). The weak phylogenetic signal we observed implies that susceptibility to malaria infection is roughly homogeneously distributed among the bird phylogenetic tree, with no clade showing consistently high or low prevalence. This may reflect mostly the taxonomic selection of study species in the past, with bird taxa a priori likely to be infected being more prone to be studied than those thought not to be. Nevertheless, the relatively weak signal we found suggests that ecological and environmental forces are stronger determinants of infection than a bird's evolutionary history.

The worldwide prevalence of avian malaria is influenced by traits of the bird hosts. Indeed, our results suggest that larger bird species are more susceptible to infection by avian malaria parasites then their smaller counterparts. This agrees with our prediction, and with previous studies at local-scale which have shown that larger bird hosts have higher prevalence of avian malaria (Scheuerlein and Ricklefs 2004, Santiago-Alarcon et al. 2015). Indeed, this trait is known to be a driver of multiple other vector-borne diseases, such as western-Nile virus (Banerjee et al. 2017) and Lyme disease (Barbour et al. 2015), and we report a similar relation, albeit at a worldwide scale, for birds infected by avian malaria. This relation can be explained by multiple components of the disease life cycle. From the vector's point of view, infecting larger birds might be easier than smaller ones. On the one hand, larger bird hosts produce more chemical cues, such as CO2, that attract vectors (Klowden and Zweibel 2005). On the other hand, larger birds might have a greater susceptibility (e.g. likelihood that a parasite infects a host depending on host antiparasitic defenses; Downs et al. 2019) towards avian malaria as they have more available biting area for the mosquitoes to latch on (de Brooke et al. 1999) as well as a lighter plumage (Peters 1983). From the bird hosts' point of view, this suggests a trade-off in resource allocation, with larger birds being more competent in foraging and competing for resources (Peters 1983, Hin and de Roos 2019) at the cost of lower immunocompetence towards vector-borne diseases, or even higher infection rates due to physiological processes affecting the reproduction of infected cells in larger hosts (see Banerjee et al. 2017). We therefore postulate that geographically separated species facing the same threat of disease show similar evolutionary responses with respect to life history strategies that maximizes their fitness. This is in contradiction to the results of Barrow et al. (2019), as we found that similarity in life history traits, driven by regional host–parasite interactions, is more important than host phylogeny at a global scale.

A central finding of our study is that climatic variables, namely temperature seasonality, also influence worldwide disease patterns of vector-borne disease, with bird populations in areas with high variation in temperature being more prone to infection. This goes against our original prediction, as we expected more seasonally stable areas to have higher infection of avian malaria due to consistent transmission rates all over the year. While this result is unexpected, as *Plasmodium* is known to be constrained by low temperature due to its maturation within the mosquito vector (Lapointe et al. 2010), it follows a similar trend as that found by Fecchio et al. (2019a), although these authors studied a different malarial parasite (*Leucocytozoon*). An explanation for this result may be related to the dynamics of organism abundance in unstable areas (e.g. areas with high seasonality dynamics). Indeed, Fecchio et al. (2019a) hypothesized that temporal concordance between pulses of resources due to pronounced seasonality, opening up niche availability (Tonkin et al. 2017???), and breeding activities of bird (e.g. at the beginning of the summer season; Karr 1976) could concentrate all organisms in the same area, thus enhancing the potential for avian malaria transmission between its hosts for a brief period of time but with high intensity. In contrast, in temporally stable areas with low seasonality and more consistent temperature, low levels of transmission spread over longer periods may not achieve similarly high prevalence of infection.

In conclusion, our results show that worldwide patterns of disease prevalence in wild populations are the result of a complex co-evolutionary interplay of host–disease interactions and climate seasonality. Other factors may potentially influence malaria infection, of course. For example, migratory behaviour can affect bird exposure and immunocompetence. We chose not to include migratory behaviour as a population-level effect term in the models since we do not know the distance that those birds travel or if the sampling occurred in the resident areas or those occupied for shorter migratory periods. We do note that the effect size of the variables selected in our models are relatively small; however, we were nevertheless able to extract complex patterns of host–disease interactions derived from very crude indicators.

More specifically, we provide evidence of traits influencing host susceptibility to diseases at a global scale, playing a more important role than host evolutionary history, and also highlight the importance of considering climatic variables influencing spatial overlap between vectors of a disease and its hosts. Considering that infectious diseases are on the rise worldwide (Jones et al. 2008), and in the context of global climate change, we bring empirical evidence of current avian malaria patterns to conservation ecologists wanting to predict and mitigate future outbreaks.

Acknowledgments The authors would like to thank Lucas Deschamps and Jean-François Doherty for their comments on an earlier version of the manuscript. We would also like to thank all the authors who forwarded original datasets to help complete the analysis, especially Alan Fecchio and Erika Martins Braga. A. Filion is supported by a University of Otago Doctoral Scholarship, A. Eriksson by a CAPES Grant (process number 88881.187634/2018-01) and F. Jorge by a grant from the Marsden Fund (to R. Poulin).

Author contribution statement RP conceived the idea. AF, AE, CNN, FJ and RP designed the study. AF collected the data and wrote the first draft of the manuscript. All authors were involved in data analyses, contributed critically to the drafts, and gave final approval for publication.

Data accessibility statement Data will be deposited in public repository (Dryad or others) upon acceptance of the article.

References

Altizer S, Dobson A, Hosseini P, Hudson P, Pascual M, Rohani P (2006) Seasonality and the dynamic of infectious diseases. Ecol Lett 9:467–484

Anderson RM, May RM (1978) Regulation and stability of host-parasite population interactions: I. Regulatory processes. J Anim Ecol 47:219–247

Atkinson CT, Lapointe DA (2009) Introduced avian diseases, climate change, and the future of Hawaiian honeycreepers. J Avian Med Surg 23:53–63

Atkinson CT, Samuel MD (2010) Avian malaria *Plasmodium relictum* in native Hawaiian forest birds: epizootiology and demographic impacts on ʻApapane *Himatione sanguinea*. J Avian Biol 41:357–366

Banerjee S, Perelson AS, Moses M (2017) Modelling the effects of phylogeny and body size on within-host pathogen replication and immune response. J R Soc Interface 14:20170479

Barbour AG, Bunikis J, Fish D, Hanincová K (2015) Association between body size and reservoir competence of mammals bearing *Borrelia burgdorferi* at an endemic site in the northeastern United States. Parasites Vectors 8:299

Barrow LN, McNew SM, Mitchell N, Galen SP, Lutz HL, Skeen H et al (2019) Deeply conserved susceptibility in a multi-host, multi-parasite system. Ecol Lett 22:987–998

Bensch S, Hellgren O, Pérez-Tris J (2009) MalAvi: a public database of malaria parasites and related heamosporidians in avian hosts based on mitochondrial cytochrome b lineage. Mol Ecol Resour 9:1353–1358

Bivand R, Keitt T and Rowlingson B (2020) rgdal: Bindings for the 'Geospatial' Data Abstraction Library. R package version 1.5-8. https://CRAN.R-project.org/package=rgdal

Brooke ML, Hanley S, Laughlin SB (1999) The scaling of eye size with body mass in birds. Proc. Biol Sci 266:405–412

Bürkner PC (2017) brms: an R package for Bayesian multilevel models using STAN. J Stat Softw 80:1–28

Day JF (2001) Predicting St. Louis encephalitis virus epidemics: lessons from recent, and not so recent, outbreaks. Annu Rev Entomol 46:111–138

Deerenberg C, Arpanius V, Daan S, Bos N (1997) Reproductive effort decreases antibody responsiveness. Proc Royal Soc B 26:1021–1029

Didan K, Barreto A (2016) NASA MEaSUREs Vegetation Index and Phenology (VIP) Phenology NDVI Yearly Global 0.05Deg CMG. In: NASA EOSDIS L. Process. DAAC

Dormann C, Elith J, Bacher S, Buchmann C, Carl G, Carré G et al (2012) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 35:1–20

Downs CJ, Schoenle LA, Han BA, Harrison JF, Martin LB (2019) Scaling of host competence. Trends Parasitol 35:182–192

Duffield KR, Bowers EK, Sakaluk SK, Sadd BM (2017) A dynamic threshold model for terminal investment. Behav Ecol Sociobiol 71:185–202

Duffy MA, Ochs JH, Penczykowski RM, Civitello DJ, Klausmeir CA, Hall SR (2012) Ecological context influences epidemic size and parasite-driven evolution. Science 335:1636–1638

Edman JD, Scott TW (1987) Host defensive behaviour and the feeding success of mosquitoes. J Insect Sci 8:617–622

Ellison AR, Tunstall T, DiRenzo GV, Hughey MC, Rebollar EA, Belden LK et al (2015) More than skin deep: functional genomic basis for resistance to amphibian chytridiomycosis. Genome Biol Evol 7:286–298

Fecchio A, Wells K, Bell JAS, Tkach VV, Lutz HL, Weckstein JD et al (2019a) Climate variation influences host specificity in avian malaria parasites. Lett, Ecol. https://doi.org/10.1111/ele.13215

Fecchio A, Bell JA, Bosholn M, Vaughan JA, Tkach VV et al (2019b) An inverse latitudinal gradient in infection probability and phylogenetic diversity for *Leucocytozoon* blood parasites in New World birds. J Animal Ecol. https://doi.org/10.1111/1365-2656.13117

Ferraguti M, de la Puente JM, Roiz D, Ruiz S, Soriguer R, Figuerola J (2016) Effects of landscape anthropization on mosquito community composition and abundance. Sci Rep 6:29002

Freeland WJ (1983) Parasites and the coexistence of animal host species. Am Nat 121:223–236

Galardo AKR, Zimmerman RH, Loubinos LP, Young LJ, Galardo CD, Arruda M et al (2009) Seasonal abundance of anopheline mosquitoes and their association with rainfall and malaria along the Matapi River, Amapi, Brazil. Med Vet Entomol 23:335–349

Giraudoux P, Craig PS, Delattre P, Bao G, Bartholomot B, Harraga S et al (2003) Interactions between landscape changes and host communities can regulate *Echinococcus multilocularis* transmission. Parasitology 127:121–131

Greenberg JA and Mattiuzzi M (2020) gdalUtils: Wrappers for the Geospatial Data Abstraction Library (GDAL) Utilities. R package version 2.0.3.2

Gregory RD, Blackburn TM (1991) Parasite prevalence and host sample size. Parasitol. Today 7:316–318

Grimaldi WW, Seddon PJ, Lyver PO, Nakagawa S, Tompkins DM (2015) Infectious diseases of Antarctic penguins: current status and future threats. Polar Biol 38:591–606

Hackett SJ, Kimball RT, Reddy S, Bowie RCK, Braun EL, Braun MJ et al (2008) A phylogenomic study of birds reveals their evolutionary history. Science 320:1763–1768

Hadfield JD (2010) MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. J Stat Softw 33:1–22

Hijmans RJ and van Etten J (2019) raster: geographic analysis and modeling with raster data. Rpackage version 2.9-5

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. Int J Climatol 25:1965–1978
Hin V, de Roos AM (2019) Evolution of size-dependent intraspecific competition predicts body size scaling of metabolic rate. Funct Ecol 33:479–490
Holt RD (1977) Predation, apparent competition, and the structure of prey communities. Theor Popul Biol 12:197–229
Holt RD, Bonsall MB (2017) Apparent competition. Annu Rev Ecol Evol Syst 48:447–471
Hudson P, Greenman J (1998) Competition mediated by parasite: biological and theoritical progress. Trends Ecol Evol 13:387–390
Jones KE, Patel NG, Levy MA, Storeygard A, Balk D, Gittleman JL et al (2008) Global trends in emerging infectious diseases. Nature 451:990–993
Kamiya T, O'Dwyer K, Nakagawa S, Poulin R (2014) What determines species richness of parasitic organisms? A meta-analysis across animal, plant and fungal hosts. Biol Rev 89:123–134
Karr JR (1976) Seasonality, resource availability, and community diversity in tropical bird communities. Am Nat 100:973–994
Keeling MJ (1999) The effects of local spatial structure on epidemiological invasions. Proc Royal Soc B 266:859–867
Keith SA, Webb T, Böhning-Gaese K, Conolly SA, Dulvy NK, Eigenbrod F et al (2012) What is macroecology? Biol Lett 8:904–906
Klowden MJ, Zweibel LJ (2005) Vector olfaction and behavior. In: Marquardt WC (ed) Biology of disease vectors, 2nd edn. Elsevier Academic Press, San Diego, pp 277–287
LaPointe DA, Goff ML, Atkinson CT (2010) Thermal constraints to the sporogonic development and altitudinal distribution of avian malaria *Plasmodium relictum* in Hawaii. J Parasitol 96:318–324
Lindgren E, Tälleklint L, Polfeldt T (2000) Impact of climatic change on the northern latitude limit and population density of the disease-transmitting European tick *Ixodes ricinus*. Environ Health Perspect 108:119–123
Loiseau C, Harrigan RJ, Bichet C, Julliard R, Garnier S, Lendvai AZ et al (2013) Prediction of avian *Plasmodium* expansion under climate change. Sci Rep. https://doi.org/10.1038/srep01126
Martin TE, Møller AP, Merino S, Clobert J (2001) Does clutch size evolve in response to parasites and immunocompetence? PNAS 98:2071–2076
McDonald JL, Robertson A, Silk MJ (2017) Wildlife disease ecology from the individual to the population: insights from a long-term study of a naturally infected European badger population. J Animal Ecol 87:101–112
Medeiros MCI, Hamer GL, Ricklefs RE (2013) Host compatibility rather than vector host-encounter rate determines the host range of avian *Plasmodium* parasites. Proc. Royal Soc. B. 280:20122947
Moore SL, Wilson K (2002) Parasites as a viability cost of sexual selection in natural population of mammals. Science 297:2015–2018
Myhrvold NP, Baldridge E, Chan B, Sivam D, Freeman DL, Morgan Ernest SK (2015) An amniote life-history database to perform comparative analysis with birds, mammals, and reptiles. Ecology 96:3109–3109
Ogden NH, Maarouf A, Bigras-Poulin IK, Lindsay M, Morshed MG et al (2006) Climate change and the potential for range expansion of the Lyme disease vector *Ixodes scapularis* in Canada. Int J Parasitol 36:63–70
Owens IPF, Wilson K (1999) Immunocompetence: a neglected life history trait or conspicuous red herring? Trends Ecol Evol 14:170–172
Paradis E, Schliep K (2018) ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35:526–528
Pascual M, Dobson AP, Bouma MJ (2009) Underestimating malaria risk under variable temperatures. PNAS 106:13645–13646
Pebesma EJ and Bivand RS (2005) Classes and methods for spatial data in R. R News 5 (2), https://cran.r-project.org/doc/Rnews/
Peters RH (1983) The ecological implication of body size. Cambridge University Press, Cambridge, p 329
Podmokla E, Dubiec A, Drobniak SM, Arct A, Gustafsson A, Chicon M (2014) Avian malaria is associated with increased reproductive investment in the blue tit. J Avian Biol 4:219–224
Pulgarin-R PC, Gomez JP, Robinson S, Ricklefs RE, Cadena CD (2018) Host species, and not environment, predicts variation in blood parasite prevalence, distribution, and diversity along a humidity gradient in northern South America. Ecol Evol 8:3800–3814
R Core Team (2018) R: A language and environment for statistical computing. R foundation for statistical computing, Vienna
Reeves WC, Hardy JL, Reisen WK, Milby MM (1994) Potential effect of global warming on mosquito-borne arboviruses. J Med Entomol 31:323–332
Reisen WK (2010) Landscape epidemiology of vector-borne diseases. Annu Rev Entomol 55:461–483
Revell LJ (2012) phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol Evol 3:217–223
Revelle W (2018) psych: Procedures for Personality and Psychological Research. Northwestern University, Evanston
Santiago-Alarcon D, MacGregor-Fors I, Kühnert K, Segelbacher G and Schaefer HM (2015) Avian haemosporidian parasites in an urban forest and their relationship to bird size and abundance. Urban Ecosyst. https://doi.org/10.1007/s11252-015-0494-0
Scheuerlein A, Ricklefs RE (2004) Prevalence of blood parasites in European passeriform birds. Proc Royal Soc B 271:1363–1370
Scott ME (1988) The impact of infection and disease on animal populations: implications for conservation biology. Conserv Biol 2:40–56
Sheldon BC, Verhulst S (1996) Ecological immunology: costly parasite prevalence defenses and trade-offs in evolutionary ecology. Trends Ecol Evol 11:317–321
Shrag SJ, Wiener P (1995) Emerging infectious disease: what are the relative roles of ecology and evolution? Trends Ecol Evol 10:319–324
Smith DL, Lucey B, Waller LA, Childs JE, Real LA (2002) Predicting the spatial dynamics of rabies epidemics on heterogeneous landscapes. PNAS 99:3668–3672
Stephens PR, Altizer S, Smith KF, Aguirre AA, Brown JH, Budischak SA et al (2016) The macroecology of infectious diseases: a new perspective on global-scale drivers of pathogen distributions and impacts. Ecol Lett 19:1159–1171
Thompson JN, Cunningham BM (2002) Geographic structure and dynamics of coevolutionary selection. Nature 417:735–738
Tompkins DM, Poulin R (2006) Parasites and biological invasions. Biol. Invasions New Zealand 186:67–84
Tompkins DM, Dunn AM, Smith MJ, Telfer S (2011) Wildlife diseases: from individuals to ecosystems. J Anim Ecol 80:19–38
Tompkins DM, Carver S, Jones ME, Krkošek M, Skerratt LF (2015) Emerging infectious diseases of wildlife: a critical perspective. Trends Parasitol 31:149–159
Tonkin JD, Bogan MT, Bonada N, Rios-Touma B, Lytle DA (2017) Seasonality and predictability shape temporal species diversity. Ecology 98:1201–1216
Torr SJ, Mangwiro TN, Hall DR (2006) The effects of host physiology on the attraction of tsetse (Diptera: Glossinidae) and Stomoxys (Diptera: *Muscidae*) to cattle. Bull Entomol Res 96:71–84
Valkiūnas G (2005) Avian malaria parasites and other haemosporidia. CRC Press, New York, p 934
Vehtari A, Gelman A, Gabry J (2017) Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Stat Comput 27:1413–1432
Woolhouse MEJ, Dye C, Etard JF, Smith T, Charlwood JD, Garnett GP et al (1997) Heterogeneities in the transmission of infectious

agents: implications for the design of control programs. PNAS 94:338–342

World health organisation (2017) Vector-borne diseases. Available at: https://www.who.int/news-room/fact-sheets/detail/vector-borne-diseases. Accessed Aug 08, 2019

Yao Y, Vehtari A, Simpson D, Gelman A (2018) Using stacking to average Bayesian predictive distributions. Bayesian Anal 13:917–1007

Zuk M, Stoehr AM (2002) Immune defense and host life history. Am Nat 160:9–22

Zuur AF, Ieno EN, Walker NJ, Saveliev AA, Smith GM (2009) Mixed effects models and extensions in ecology with R. Springer, New York, p 574

Zuur AF, Ieno EN, Elphick CS (2010) A protocol for data exploration to avoid common statistical problems. Methods Ecol Evol 1:3–14