Climate variation influences host specificity in avian malaria parasites

Abstract
Parasites with low host specificity (e.g. infecting a large diversity of host species) are of special interest in disease ecology, as they are likely more capable of circumventing ecological or evolutionary barriers to infect new hosts than are specialist parasites. Yet for many parasites, host specificity is not fixed and can vary in response to environmental conditions. Using data on host associations for avian malaria parasites (Apicomplexa: Haemosporida), we develop a hierarchical model that quantifies this environmental dependency by partitioning host specificity variation into region- and parasite-level effects. Parasites were generally phylogenetic host specialists, infecting phylogenetically clustered subsets of available avian hosts. However, the magnitude of this specialisation varied biogeographically, with parasites exhibiting higher host specificity in regions with more pronounced rainfall seasonality and wetter dry seasons. Recognising the environmental dependency of parasite specialisation can provide useful leverage for improving predictions of infection risk in response to global climate change.

Keywords
avian malaria, climate change, disease ecology, disease emergence, host shifting, host specificity, infectious disease, niche specialisation, parasite specialisation, vector borne disease.

INTRODUCTION
Global disease burdens commonly reflect host range expansions (termed herein as ‘host shifting’) by multi-host parasites (Han et al. 2015; Wells et al. 2018a). Host specificity (e.g. the diversity of host species a parasite is capable of infecting) is a useful metric to describe differences among parasites in their capacity to infect novel hosts or trigger parasitic disease emergence events (Poulin et al. 2011). With the majority of emerging infectious diseases thought to result from shifting host associations by parasites with low host specificity, a major goal in disease ecology is to apply host specificity metrics using observed host association data to identify ‘generalist’ parasites before they cause disease outbreaks (Cooper et al. 2012; Brooks et al. 2014; Dallas et al. 2017). However, mounting experimental and theoretical evidence suggests that host specificity is not a fixed trait (Poulin & Mouillot 2005; Brooks & Hoberg 2007; Agosta et al. 2010; Nylin et al. 2018). Instead, a parasite’s local host specificity (herein termed ‘realized host specificity’) can be considered the product of a hierarchical process involving both regional and evolutionary forces (Wells et al. 2018b). For host range expansions to occur, a parasite must first be exposed to a novel host species. This exposure will be influenced by environmental conditions that determine host community composition, as spatiotemporal variation in host occurrences alters host-parasite contact rates (Canard et al. 2014). Second, adaptation to a new host is required to facilitate transmission. For many parasites, this process is expected to adhere to the principle of ‘ecological fitting’ (Janzen 1985), which states that sharing certain characteristics with previous host species is necessary for successful infection (Brooks et al. 2006; Davies & Pedersen 2008; Poulin et al. 2011; Clark & Clegg 2017). Yet host traits that influence susceptibility, such as clutch size or breeding behaviour, can fluctuate in response to environmental conditions (Moller et al. 2013). Despite an accelerated focus on describing host specificity for a multitude of parasites (Hellgren et al. 2009; de Vienne et al. 2009; Farrell et al. 2013; Clark et al. 2018; Doña et al. 2018; Park et al. 2018), few empirical studies recognise this environmental dependency by treating

1Laboratório de Evolução e Biogeografia, Universidade Federal da Bahia, Rua Barão de Jeremoabo 147, Salvador, BA 40170115, Brazil
2Department of Biosciences, Swansea University, Swansea SA2 8PP, UK
3Department of Biology, University of North Dakota, 1 Campus Drive and Cornell Street, Grand Forks, ND 58202, USA
4Department of Biology, University of North Dakota, 10 Cornell Street, Grand Forks, ND 58202, USA
5Department of Surgery, University of Chicago, 5812 S. Ellis Ave., Chicago, IL 606372, USA
6Integrative Research Center, Field Museum of Natural History, 1400 S. Lake Shore Drive, Chicago, IL 60605, USA

© 2019 John Wiley & Sons Ltd/CNRS
specificity as a geographically labile trait (but see Krasnov et al. 2004a,b; Wells et al. 2018b).

The challenge of assessing variation in realised host specificity is understandable. This requires detailed information about the host distributions of parasites across climatically variable bioregions, which is difficult to acquire for many parasites (Murray et al. 2015). Nevertheless, the lack of a comprehensive assessment of how climate variation influences host specificity presents an impediment for both predicting the emergence of infectious diseases and developing mitigation strategies (Altizer et al. 2013; Brooks et al. 2014). We address this gap by using a hierarchical modelling approach to test whether the realised host specificity of multi-host parasites varies across biogeographical regions. Focusing on a cosmopolitan group of avian blood parasites, we partition variation in realised host specificity into regional and parasite-level effects.

Avian haemosporidians (including the genera Plasmodium and Haemoproteus, termed hereafter as avian malaria parasites) are protozoan parasites that infect bird blood cells and are transmitted by haematophagous dipteran vectors (Valkiūnas 2005; Santiago-Alarcon et al. 2012a). These parasites are globally distributed, abundant and diverse in most bird clades, and their estimated host specificities are highly variable, ranging from infecting a single host species to many unrelated species (Valkiūnas 2005; Clark et al. 2014; Møens & Pérez-Tris 2016). Previous studies have revealed that the distributions and community compositions of avian malaria parasites are the outcome of host switching events over macroevolutionary timescales (Ricklefs et al. 2014; Fecchio et al. 2018b) and recent ecological forces such as variation in host dispersion capability (Pérez-Tris & Bensch 2005; Ellis et al. 2015; Clark et al. 2017; Fecchio et al. 2018b). Yet mechanisms that contribute to the large observed variation in host specificity are largely unknown, although environmental forces may play substantial roles (Clark et al. 2017). For example, risk of Plasmodium infection in birds is expected to increase with increasing temperatures on a global scale (Garamszegi 2011). Despite not being able to directly link climatic conditions to parasite specialisation, Garamszegi (2011) demonstrated that the impact of climate change on avian Plasmodium prevalence varies on a continental scale, with the strongest effects found for Europe and Africa. Climate variation also influences rates of parasite reproduction and development within vector hosts, which could in turn affect parasite transmission and the exposure of parasites to novel host species (LaPointe et al. 2010; Santiago-Alarcon et al. 2012a).

The search for general processes governing host specificity should assess both ecological and phylogenetic relationships of potential host species in efforts to identify barriers to host range expansions (Poulin & Mouillot 2005; Hoberg & Brooks 2008; Clark et al. 2017). We present our framework using a database comprising infection data for 154 multi-host avian malaria parasites sampled from 15 541 individual birds in South America and the Australia-Pacific. Samples cover 109 sites, which we group into 10 biogeographical regions to delineate communities with similar environmental conditions and avian compositions (Fig. 1 and Table 1). We consider variation in pairwise distances between infected (i.e. observed to carry a parasite) and potential (i.e. hosts occurring in the same region but not found to be infected) host species to represent a signal of realised host specificity. We test whether magnitudes of realised phylogenetic and ecological host specificity vary across regions using a multilevel model that includes parameters reflecting region-level and parasite-level contributions. We then test whether environmental (rainfall and temperature) and biotic (host and parasite species diversity) factors explain regional differences in host specificity.

**MATERIAL AND METHODS**

**Parasite database**

All parasite lineages in our dataset were identified using PCR-based detection methods targeting a 477 cytochrome-b (cyt-b) barcoding fragment of the haemosporidian mitochondrial genome. The majority of observations came from field studies led by the authors from the period of 2005–2016, with remaining observations extracted from published studies that took place in the study region (see Clark et al. 2017 for details). Protocols detailing reactions, reagents, primer names, cycling conditions and how lineages were determined can be found in (Hellgren et al. 2004; Bensch et al. 2009; Bell et al. 2015). As evidence indicates that avian malaria lineages differing by one cyt-b nucleotide may be reproductively isolated entities (Bensch et al. 2004), we use the standard practice of referring to each unique cyt-b lineage as a unique parasite.

**Climate variable extraction and biogeographic region delineation**

Our study region was delineated to represent a diversity of habitats, avian compositions, and climate envelopes so that we had robust statistical power to estimate associations between regional conditions and host ranges of parasites. We extracted 19 climate variables (based on average values from the years 1970–2000) for each site (n = 109) from www.worldclim.org (accessed March 2018; see Appendix S1 in Supporting Information). We chose WorldClim records as such lower resolution climate data are more appropriate for predicting species’ distributions across large bioregions (i.e. the distributions of potential and realised hosts), which higher resolution climate data may fail to detect due to localised weather events or stochastic variation. Records of avian occurrences for sites were extracted from species distribution maps acquired from www.datazone.birdlife.org (BirdLife International and NatureServe, 2017). We grouped sites into 10 biogeographical regions using hierarchical clustering of a Gower’s matrix (Gower 1971) capturing dissimilarity in avian community composition and climate variables (Table 1, Fig. 1, Appendix S1). We chose this method for grouping sites into regions based on mounting evidence that variation in avian composition and long-term climate variables both have major influences on the assembly, prevalence, and host specificity of avian malaria parasites (e.g. Sehgal 2010; Clark et al. 2017; Clark 2018; Fecchio et al. 2018a). Our clustering method therefore presents a data-driven approach designed to delineate regions that are biologically meaningful at the parasite level.
Avian host phylogenetic and ecological relationships

Distributions for a total of 5450 avian species overlapped our sample area according to BirdLife species range maps. We extracted phylogenetic and ecological data for these species to generate estimates of historical and functional relationships of potential host species. Note that only species sampled for avian malaria parasites were considered as potential hosts (957 species). Phylogenetic distances were calculated as mean pairwise distance across 100 phylogenetic trees sampled from a global avian supertree distribution (Jetz et al. 2012; accessed at http://birdtree.org/subsets/). We extracted species’ proportional use of ten diet categories and seven foraging habitats (traits likely to impact parasite exposure) from EltonTraits v1.0 (Wilman et al. 2014). We quantified pairwise ecological distances using a Gower’s distance matrix (Gower 1971) following methods in Pavoine et al. (2009). Host phylogenetic and ecological distance matrices were scaled (dividing by the maximum for each matrix; see Appendix S2).

Statistical analysis

Parasite- and region-specific host specificity

Lists of potential avian host species (i.e. species sampled for avian malaria parasites) were generated for parasites in each region where the parasite was recorded. This resulted in parasite- and region-specific potential host pools for which associations were recorded as binary variables (i.e. ‘1’ if the

Table 1 Sample sizes (sites, individual birds, host species and parasite lineages sampled/modelled), bird species richness and host specificity coefficients (Phylo $\beta_{region}$ and Eco $\beta_{region}$) for biogeographical regions

<table>
<thead>
<tr>
<th>Bioregion</th>
<th>Latitude (mean)</th>
<th>Longitude (mean)</th>
<th>Phylo Beta (mean)</th>
<th>Host phylo diversity</th>
<th>Eco Beta (mean)</th>
<th>Host eco diversity</th>
<th># sample sites</th>
<th># samples</th>
<th># host species</th>
<th># parasites</th>
<th># parasites (modelled)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia N/PNG</td>
<td>−15</td>
<td>145.97</td>
<td>−0.02</td>
<td>62.48</td>
<td>−0.02</td>
<td>70.38</td>
<td>6</td>
<td>1336</td>
<td>115</td>
<td>117</td>
<td>23</td>
</tr>
<tr>
<td>Australia SE</td>
<td>−31.19</td>
<td>148.1</td>
<td>0.01</td>
<td>40.61</td>
<td>−0.05</td>
<td>46.74</td>
<td>10</td>
<td>3926</td>
<td>73</td>
<td>43</td>
<td>24</td>
</tr>
<tr>
<td>Australia W</td>
<td>−28.02</td>
<td>117.65</td>
<td>−0.06</td>
<td>5.18</td>
<td>−0.03</td>
<td>6.34</td>
<td>2</td>
<td>191</td>
<td>10</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>Brazil Amazonia</td>
<td>−2.75</td>
<td>−56.28</td>
<td>0.04</td>
<td>188.73</td>
<td>0.1</td>
<td>172.68</td>
<td>14</td>
<td>2251</td>
<td>308</td>
<td>206</td>
<td>35</td>
</tr>
<tr>
<td>Brazil Central/E</td>
<td>−13.71</td>
<td>−48.15</td>
<td>−0.03</td>
<td>102.75</td>
<td>−0.06</td>
<td>128.93</td>
<td>12</td>
<td>1757</td>
<td>210</td>
<td>99</td>
<td>23</td>
</tr>
<tr>
<td>Malaysia</td>
<td>−2.85</td>
<td>103.45</td>
<td>0.01</td>
<td>19.17</td>
<td>0.08</td>
<td>20.59</td>
<td>2</td>
<td>143</td>
<td>36</td>
<td>19</td>
<td>3</td>
</tr>
<tr>
<td>Melanesia</td>
<td>−17.36</td>
<td>167.63</td>
<td>0.1</td>
<td>20.42</td>
<td>0.1</td>
<td>22.19</td>
<td>28</td>
<td>1947</td>
<td>44</td>
<td>54</td>
<td>18</td>
</tr>
<tr>
<td>New Zealand</td>
<td>−40.2</td>
<td>174.25</td>
<td>−0.03</td>
<td>15.43</td>
<td>−0.02</td>
<td>16.7</td>
<td>23</td>
<td>2544</td>
<td>29</td>
<td>12</td>
<td>4</td>
</tr>
<tr>
<td>Peru</td>
<td>−5.87</td>
<td>−77.27</td>
<td>0.03</td>
<td>154.22</td>
<td>0.03</td>
<td>170.78</td>
<td>4</td>
<td>1174</td>
<td>270</td>
<td>72</td>
<td>15</td>
</tr>
<tr>
<td>Philippines</td>
<td>12.57</td>
<td>121.79</td>
<td>−0.06</td>
<td>17.31</td>
<td>−0.12</td>
<td>21.4</td>
<td>7</td>
<td>245</td>
<td>37</td>
<td>59</td>
<td>7</td>
</tr>
</tbody>
</table>

Notes: Regions were delineated based on dissimilarities in avian community composition and climate variables. Diversity metrics (phylo diversity, eco diversity) were calculated by multiplying host species richness by posterior modes of regression intercepts ($\mu_{region}$), which represent the average pairwise distance between potential host species in a region. PNG, Papua New Guinea.
potential host was infected, ‘0’ if uninfected). Vectors of potential host pairwise distances were response variables in

\[
distance \sim \mathcal{N}(\mu_{\text{region}} + \beta_{\text{region}-\text{parasite}} \cdot \text{host.pair}, \sigma^2)
\]

\[
\mu_{\text{region}} \sim \mathcal{N}(H_0, \sigma^2_0); \beta_{\text{region}-\text{parasite}} \sim \mathcal{N}(\mu_{lb} + \rho_{\text{parasite}} + \beta_{\text{region}}, \sigma^2_\beta)
\]

\[\mathcal{N}(\mu, \sigma^2)\] denotes normal distributions with mean \(\mu\) and variance \(\sigma^2\). \(\mu_{\text{region}}\) denotes regional averages (corresponding to the intercept of linear models) of either ecological or phylogenetic pairwise distances (distance) for potential host pools, drawn from a hyperprior \(H_0\) with Gaussian error \(\sigma^2_0\), representing ‘global’ averages. Coefficients \(\beta_{\text{region}-\text{parasite}}\) represent parasite- and region-specific estimates of differences between observed and potential host distances (i.e. the binary indicator variable host.pair where ‘1’ indicates the pair of potential host species that is infected; ‘0’ indicates that they are uninfected). This was modelled with intercept \(\mu_{lb}\) and coefficients \(\beta_{\text{parasite}}\) and \(\beta_{\text{region}}\) to capture expectations that host specificity is a function of both parasite identity and environmental conditions.

We used Gibbs variable selection (GVS) to pull \(\beta_{\text{parasite}}\) and \(\beta_{\text{region}}\) estimates towards zero when support was limited (O’Hara & Sillanpää 2009). We sampled Bernoulli indicator variables, in this case \(l\), to control whether the effect \(l\) was included in the model. By specifying a low prior probability of drawing 1 for indicator variable \(l\), we only estimated \(l\) if sufficient data existed to warrant its inclusion. If \(l = 0\), indicating little support for sampling \(l\) according to likelihood estimates, we sampled instead from a ‘pseudo-prior’ that resulted in zero-effects. This ensured avoidance of over-parameterisation (Wells et al. 2016). \(\beta_{\text{parasite}}\) estimates were sampled from normal hyperpriors \((H_{\text{genus}})\), which were based on the average specificity for the parasite’s respective genus (\(Plasmodium\) or \(Haemoproteus\)), using parasite-specific variance components \((\sigma^2_{\text{parasite}})\). Estimates for \(\beta_{\text{region}}\) were drawn from a ‘global’ normal distribution. Parameters were estimated independently for phylogenetic and ecological specificity (Appendix S3).

We estimated \(\beta\) coefficients for each parasite and each region using Markov chain Monte Carlo (MCMC) sampling based on the Gibbs sampler in the open-source software JAGS (Plummer 2003). Priors for coefficients were specified with \(H \sim \mathcal{N}(0, 10)\) and \(\sigma \sim \text{dexp}(0.5)\). We ran two MCMC chains for 50 000 iterations for parameter adaptation and sampled 1000 posterior estimates. Mixing of chains was inspected visually and with the Gelman-Rubin diagnostic (all values < 1.2). We compared magnitudes of \(\beta_{\text{region}}\) and \(\beta_{\text{parasite}}\) coefficients to gather evidence that particular parasites and/or regions showed different host specificities in comparison to other parasites/regions. Distances between infected host species that differ from draws from potential host pools indicate specificity; lower values (i.e. 95% credible intervals < 0) indicate higher similarity between observed hosts than expected; values > 0 suggest that parasites infect more distantly related hosts than expected (Clark & Clegg 2017).

Many parasites were only recorded infecting a single host species \((n = 468\) single-host parasites) and for some hosts we sampled only a few individuals \((289\) host species). Because our estimate of host specificity is based on pairwise distances from potential and realised host pools, detecting significant effects is only possible with reasonable sample sizes. We filtered the dataset by keeping (1) host species with at least eight samples in each region and (2) parasites that infected at least one of the included host species and were recorded at least three times overall. This allowed us to assess host species that have been sufficiently sampled to detect relatively rare parasites (i.e. a sample size of eight translates to an \(\sim 80\%\) probability of detecting a parasite with a true prevalence of 20%) and to assess parasites for which we have adequate information on observed host ranges. This dataset included 154 parasite lineages \((71\ Plasmodium, 83\ Haemoproteus;\ Appendix S3 and S4), which were recorded in 2–24 different avian host species and across 1–4 different biogeographical regions (Fig. 2). A total of 289 avian species were included as potential host species across the final dataset. We did not record whether avian species were native or introduced, as their occurrence within a region (regardless of how they came to be there) should still make them suitable as potential host species.

**Predictors of regional variation in host specificity**

We explored potential predictors of regional variation in host specificity using posterior modes of \(\beta_{\text{region}}\) coefficients as response variables in a multiple linear regression with assumed Gaussian error distribution. We tested six climate variables (all related to temperature and precipitation) as continuous covariates. Indices of region-specific host phylogenetic and ecological diversity were calculated using a metric that captured host species richness and average pairwise distances within each region \((\mu_{\text{region}}\) estimates\). These were also included as covariates to assess whether increased host diversity (in terms of either phylogenetic or functional diversity) leads to increased parasite specialisation. To account for sampling bias, we included parasite richness, the number of birds screened (sample size), and the GVS support for \(\beta_{\text{region}}\) estimates as covariates. Collinearity was accounted for by removing the more highly correlated variable (i.e. the variable that showed a higher number of strong pairwise correlations) from those pairs with Pearson correlations > 0.7. Remaining covariates were: minimum rainfall of the driest quarter, maximum rainfall of the wettest quarter, rainfall seasonality, minimum temperature of the coldest quarter, temperature seasonality, parasite species richness, sample size, and host diversity. We used LASSO variable selection (where the important predictors are retained by iteratively regularising coefficients for less important predictors towards zero) and leave-one-out cross-validation to test within-sample model fit (Friedman et al. 2010). This was repeated 1000 times to minimise cross-validated error and identify important predictors (i.e. those retained in at least 90% of cross-validation runs). We calculated proportions of explained variance for retained predictors following Nakagawa & Schielzeth (2013; see Appendix S3).

Analyses were conducted in R version 3.3.3 (R Core Team 2017) and primarily used functions in packages ade4 (Dray & Dufour 2007), dplyr (Wickham et al. 2017), ggmplot (Friedman et al. 2010), readxl (Wickham & Bryan 2017), and sjstats (Plummer 2016). Tutorials to replicate analyses are included in the Appendices. GenBank accession numbers for the 154 modelled parasites are presented in Appendix S4.
RESULTS

Our hierarchical regression is formulated to test the extent to which parasite identity ($\beta_{\text{parasite}}$) and regional environmental conditions ($\beta_{\text{region}}$) contribute to a parasite’s realised host specificity. Applying this framework to observed host association data for 154 multi-host avian malaria parasites, we find that realised host specificity varies across biogeographical regions (Fig. 1). Patterns are similar regardless of whether we assess ecological or phylogenetic $\beta_{\text{region}}$ specificity, suggesting the presence of general biogeographical forces influencing the host ranges of avian malaria parasites (Fig. 1). Cross-validated linear regressions to explore environmental predictors of host specificity variation show a strong influence of precipitation heterogeneity: regions with more pronounced rainfall seasonality harbor more specialised parasites (smaller $\beta_{\text{region}}$ estimates), with the coefficient of rainfall variation accounting for 53% of explained variance in $\beta_{\text{region}}$ estimates ($t = -0.56$; Appendix S3). Although seasonality is important, rainfall in the dry season also correlates with variation in host specificity: minimum rainfall of the driest quarter accounted for a further 35% of explained variance, with parasites becoming more specialised in regions with wetter dry seasons ($t = -0.45$; Fig. 3, Appendix S3).

Inferences on climate-driven effects were robust to potential sampling bias, which we accounted for by focusing on adequately sampled hosts and parasites to minimise underestimates of host ranges (see ‘Parasite- and region-specific host specificity’ in Material and Methods and Appendix S3 for details). Moreover, parameters capturing variation in sample sizes, the diversity of sampled avian hosts and numbers of recovered parasites in a region all had little influence on realised host specificity. This identification of important climate predictors allows delineation of biogeographical areas with greater potential for ongoing host range expansions by generalist parasites. For instance, sites in Brazilian Amazonia, Peruvian Andes, and tropical/sub-tropical islands in Malaysia and Melanesia contain very distinct avian communities (Holt et al. 2013) and exhibit considerable variation in sampling effort and diversity of recovered parasites (Table 1). Yet these regions contained some of the least specialised parasite communities in our dataset, in correspondence with relatively low levels of average rainfall in the dry season (Fig. 3). In contrast, parasite communities in New Zealand, The Philippines, and southeastern Australia were more specialised than expected according to potential host species pools (Table 1, Figs 1 and 3).

Assessing host specialisation components at the parasite level ($\beta_{\text{parasite}}$) indicates whether parasites are infecting clustered subsets of available hosts. If host range expansions are predominately driven by vector feeding patterns, parasites should infect hosts that are more ecologically similar (i.e. occupying more similar habitats). However, we would also expect parasites to show some level of host phylogenetic specialisation, as different physiological characteristics among unrelated hosts can impose barriers to parasite transmission or within-host development. These mechanisms are not mutually exclusive. By estimating parasite-level specificity components, we find that parasite specialisation was generally driven by host phylogeny, not by host ecological similarity. Phylogenetic $\beta_{\text{parasite}}$ estimates were consistently negative for both parasite genera, indicating that most parasites infected hosts that were phylogenetically clustered within the community (Fig. 4). Ecological $\beta_{\text{parasite}}$ estimates generally centred around zero.

To account for possible influences of biogeographical region delineation on our inferences, we tested the robustness of our results by repeating the analysis using a second grouping scheme (grouping into nine regions rather than 10 and giving more weight to avian composition and climate variables; see Appendix S6 for details). Results were broadly equivalent (strong influences of minimum rainfall in the driest quarter and rainfall seasonality on $\beta_{\text{region}}$ estimates), with the exception that ecological $\beta_{\text{parasite}}$ estimates were also generally negative.

DISCUSSION

Niche specialisation for a multitude of organisms is not fixed but is predicted to vary in response to environmental heterogeneity (Dobzhansky 1950; Janz & Nylin 2008; Schemske et al. 2009). A growing body of anecdotal and theoretical evidence suggests parasites are no exception (Agosta et al. 2010; Araujo et al. 2015; Hoberg & Brooks 2015; Nylin et al. 2018). Using a hierarchical model, we provide empirical evidence that the magnitude of realised host specificity for multi-host parasites varies in response to environmental conditions. While most avian malaria parasites generally infect phylogenetically clustered subsets of available hosts, realised host specificity
specificity increases in regions with higher rainfall during the dry season and more pronounced rainfall seasonality. This may reflect pulses in vector feeding activities or local host contact rates acting as selective barriers to host range expansions. These findings underscore the importance of treating host specificity as a geographically labile trait, contingent on both historical host–parasite interactions and environmental conditions (Hoberg & Brooks 2015). Climate change may have unforeseen consequences on the emergence potential of multi-host pathogens.

Influences of precipitation heterogeneity on realised host specificity
Climate change and biotic homogenisation are major forces acting on the distributions of species (Wilson et al. 2016; Poisot et al. 2017). Efforts to determine how such forces influence distributions of parasites, and the ranges of host species they infect, are needed to understand and predict disease emergence (Poulin et al. 2011; Altizer et al. 2013; Brooks et al. 2014; Wells et al. 2015, 2018a; Dallas et al. 2017). We show that pronounced seasonality in rainfall and higher rainfall during the dry season correlate with increased host specificity for multi-host avian malaria parasites. This link with seasonality goes against expectations of increased specialisation under stable conditions (Futuyma & Moreno 1988). An understanding of vector-vertebrate host interactions is necessary to explain this discrepancy. Successful host range expansions by parasites will predominantly be driven by variation in opportunity (exposure to novel host species) and host-parasite compatibility (driven by ecological fitting; Janzen 1968; Araujo et al. 2015). For the incredible diversity of vector-transmitted parasites, including avian malaria, opportunistic contact with novel host species depends on vector feeding patterns. Birds in seasonal areas typically breed near the start of the wet season, relying on energy reserves accrued during the dry season (Sinclair 1978; Rubenstein & Lovette 2007). Vector reproduction and larval development, both of which affect parasite transmission, are positively correlated with rainfall and synchronised with vertebrate reproduction (LaPointe et al. 2010; Santiago-Alarcon et al. 2012a). Seasonality drives pulses in food, water, and habitat availability, which increases breeding densities and perhaps concentrates organisms near available water bodies (Chesson et al. 2004; Tonkin et al. 2017). This is especially true for birds, which often concentrate in high densities at the start of the breeding season (Karr 1976; Levey 1988). These water bodies could therefore serve as source locations for parasite transmission, as has been shown for West Nile virus transmission to greater sage-grouse (Centrocercus urophasianus) (Zou et al. 2006; Walker et al. 2007). Concordance between avian breeding behaviour and peak vector activity in concentrated areas could also direct parasites to concentrated sets of ecologically similar avian species. This may impose selective pressure towards vertebrate specialisation. In such an environment, where vectors are concentrated and host–vector encounter rates and resource competition are high, one would expect parasites that are more specialised to be more successful. In contrast, if vertebrate hosts are scattered throughout the environment (which may occur in less seasonal environments) indiscriminant vector feeding could increase opportunities for novel host-parasite interactions and perhaps lead to less specialised parasites.

Figure 3 Density distributions of regional host specificity ($β_{\text{region}}$) coefficients, arranged by decreasing minimum rainfall of the driest quarter. Lower $β_{\text{region}}$ values indicate higher similarity between infected host species than expected by chance, suggesting higher host specificity in a region (relative to remaining regions). Higher $β_{\text{region}}$ values suggest parasites infect more distantly related host species than expected, indicating a greater tendency towards host generalism. Minimum rainfall variation accounted for 44% of explained variance in $β_{\text{region}}$ values, with more specialised parasites in regions with wetter dry seasons. Polygon colours correspond to region names in Table 1.

Figure 4 Parasite host specificity regression coefficients ($β_{\text{parasite}}$) presented as 95% highest posterior density credibility intervals. Each vertical bar indicates a parasite species’ ecological (upper panel) and phylogenetic (lower panel) specificity, respectively. Negative (i.e. not overlapping with zero) $β_{\text{parasite}}$ values indicate that pairs of host species tend to be more similar than by chance according to regional host species pools.
Variation in transmission rates may also occur under seasonal conditions. The supposition that disease outbreaks are more prominent in seasonal environments than in constant ones has received strong theoretical and empirical support (Altizer et al. 2006; Lisovski et al. 2017; Huber et al. 2018). A number of explanatory mechanisms have been proposed, including seasonal variation in host sociality, breeding behaviour or immune investment (Altizer et al. 2006). Regardless of underlying processes, higher frequencies of disease outbreaks suggest parasites in seasonal areas may benefit from increased infection prevalence. This could also select against range expansions to phylogenetically or functionally distant potential host species, which require costly adaptation to new defences but may be necessary when overall transmission rates are low (Poulin 1998).

Importantly, we did not test for associations between specialisation and prevalence or infection intensity here, and the idea that vectors are the limiting step in avian malaria distributions or specialisation has received mixed support. For example, some work demonstrates preferential feeding of vectors on certain avian species (Apperson et al. 2004) and tight evolutionary links between Haemoproteus lineages and vector species (Martinez-de la Puente et al. 2011), both of which support our idea that exposure of parasites to new hosts could be limited in seasonal environments. Other studies provide conflicting evidence by suggesting that vector feeding specificity is not important in structuring haemosporidian communities, particularly for Plasmodium parasites (Njabo et al. 2010; Medeiros et al. 2013). Furthermore, a recent work suggests that although a parasite lineage may be found infecting a wide diversity of hosts, they are actually better adapted to key host species as indicated by their infection intensities (Huang et al. 2018). Collectively, this evidence could indicate that other forces besides vector feeding may limit rates of novel host encounters for parasites. Assessing whether vector feeding specificity or activity rates change across regions with differing seasonality patterns would help interpret our findings and generate future research directions.

**Phylogenetic barriers to host range expansions**

Many parasites and pathogens can disperse widely across geographical realms and infect distantly related host species, and avian malaria parasites are no exception (Pérez-Tris & Bensch 2005; Hellgren et al. 2007; Ellis et al. 2015; Ricklefs et al. 2017; Fecchio et al. 2018a,b). Global distributions of several common and potentially invasive Plasmodium lineages (Bensch et al. 2009; Clark et al. 2015; Marzal et al. 2015; Ellis et al. 2018) could be interpreted as evidence that these parasites are indiscriminant host-generalists capable of infecting an enormous diversity of host species in any given environment. We challenge this assertion by showing that multi-host avian malaria parasites, even those that infect a high number of avian host species, generally infect phylogenetically clustered subsets of available hosts. This has important ramifications for our understanding of how host range expansions occur. Local co-occurrence of primary host species is sometimes necessary to facilitate survival of parasites that have encountered novel host species but have not yet locally co-adapted to the new host’s immune defenses (Fox et al. 1997; Best et al. 2010). For avian malaria parasites, phylogenetic relationships between primary and potential avian host species clearly play a central role in determining host associations and community assembly (Ellis et al. 2015; Clark et al. 2017; Fecchio et al. 2018a), despite their reliance on arthropod vectors that may feed on a diversity of avian species (Santiago-Alarcon et al. 2012b). However, although our study agrees with suggestions that breaking phylogenetic host barriers is an evolutionarily rare event (Hellgren et al. 2007; Agosta et al. 2010), this must nevertheless be a key process for generating parasite biodiversity. Host switching is a major macroevolutionary event shaping avian malaria evolution and community turnover (Ricklefs et al. 2014; Alcala et al. 2017; Fecchio et al. 2018b). Importantly, we here only study contemporary host ranges of parasites, rather than inferring patterns of historical host switching. However, our findings could indicate that rainfall seasonality plays a role in the likelihood of host switching over evolutionary timescales. Climate variation should be jointly considered with historical factors in understanding the ecology and evolution of vector-borne pathogens.

**Study limitations**

Some limitations of our modelling approach should be recognised. First, we concentrate only on multi-host avian malaria parasites. This ignores the many parasite lineages that only infect a single host species, which may limit our ability to draw conclusions on the biogeography of realised host specificity. Our estimates of realised host specificity rely on adequate support from the data, meaning that precisely estimating coefficients for parasites occurring in a small number of hosts will, in many cases, be limited. Delineating larger biogeographical regions can improve sample sizes, albeit at the cost of resolution. For example, our sensitivity analysis, which used only nine rather than 10 regions, identified a greater tendency for ecological specialisation among parasites. This suggests that the added sample sizes within groups may have provided the extra data necessary to tease apart ecological specialists. Finally, because we constrain estimates with insufficient support to the overall average (through hyperprior specifications and Bayesian variable selection), effects can be considered conservative and should be revisited following acquisition of additional data.

**Extending our models to other host-parasite systems under the emerging Stockholm Paradigm**

Our findings can broadly be interpreted under principles of The Stockholm Paradigm, which postulates that host range expansions by parasites are the product of an interplay between (1) novel host-parasite opportunities occurring across dynamic host landscapes and (2) phylogenetic and/or ecological barriers that limit adaptation by parasites to these opportunistic hosts (Araujo et al. 2015; Hoberg & Brooks 2015). Multi-host parasites exhibit a Sloppy Fitness Space whereby realised host ranges are a subset of larger potential host
ranges, including the full diversity of host species that a parasite is capable of infecting (Hoberg & Klassen 2002; Agosta & Klemens 2008). Our findings suggest that variation in the realised host specificity of avian malaria parasites follows a hierarchical process consisting first of heterogeneity in potential host pools (occurring most notably across regions characterised by different precipitation patterns) and evolutionarily conserved host traits or behaviours that limit successful infection (Wells et al. 2018b). Biogeographical structure in host specificity likely reflects prominent roles of vector feeding patterns or shifts in host compositions in response to regional climatic conditions.

Recognising that host specificity is not fixed, as we have shown here, provides new leverage for outlining region-specific predictions of infectious disease risk by emerging parasites, particularly in areas undergoing rapid climate change. Given that an enormous diversity of macro- and micro-parasites depend on external climate conditions during at least part of the life cycles (Patz et al. 2000; Brooks & Hoberg 2007), our approach can provide new insights into host association patterns for many host–parasite systems. Related models have already been successfully used to uncover global variation in realised host specificity for important zoonotic helminth parasites (Wells et al. 2018b). We have extended the flexibility of these models by incorporating group-level hyperpriors to capitalise on the added power that partial pooling can provide in mixed effects regressions (Gelman & Hill 2007). Used in combination with the increasing availability of remote-sensed environmental variables and host–parasite association datasets (Wardeh et al. 2015; Stephens et al. 2017), our approach can play a key role in determining whether the magnitude of parasite specialisation varies in response to climate patterns. For example, incorporating data on host migration patterns (to provide finer estimates of local host composition) or landscape features (to more adequately describe regional ecological variation) could be a valuable next step to ground-truthing our models for other systems. Improving surveillance regimes and the spatial resolutions of open-source host–parasite databases will enhance our ability to disentangle biological signals of host specificity from inherent noise associated with low resolution data. This is imperative to identify which biotic and abiotic conditions increase risks for parasitic disease emergence and pathogen spillover events.

ACKNOWLEDGEMENTS

Our work would not have been possible without the field assistance of several local residents and ornithologists. We thank the curators from the following museums who loaned part of the samples used in this study: Instituto Nacional de Pesquisas da Amazônia, Museu Paraense Emílio Goeldi, Field Museum of Natural History, and Academy of Natural Sciences of Drexel University. Eric Schall and Therese Catanauch carried out GIS work to extract host occurrence lists. We thank governmental agencies which provided all permits necessary for collection and exportation of tissue samples. We also thank four anonymous referees for their comments on earlier versions of the manuscript. This work was funded in part by US National Science Foundation grants DEB-1503804 to JDW and DEB-1120734 to VVT and a National Geographic Society Committee for Research and Exploration Grant (9383-13) to SMC. During the project, AF was supported by a postdoctoral fellowship (PNPD scholarship) from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

AUTHORSHIP

AF and NJC designed research and wrote the first draft; NJC, and KW analysed data; AF, JAB, VVT, HLL, JDW, SMC and NJC conducted field/lab research; all authors interpreted results and contributed to writing.

REFERENCES


SUPPORTING INFORMATION
Additional supporting information may be found online in the Supporting Information section at the end of the article.

Editor, Peter Thrall
Manuscript received 16 October 2018
First decision made 13 November 2018
Manuscript accepted 28 November 2018