Predictability of helminth parasite host range using information on geography, host traits and parasite community structure

TAD DALLAS[1,2]*, ANDREW W. PARK[1,3] *and* JOHN M. DRAKE[1,3]

[1] *Odum School of Ecology, University of Georgia, Athens, GA 30602, USA*
[2] *Department of Environmental Sciences and Policy, University of California, Davis, CA 95616, USA*
[3] *Center for the Ecology of Infectious Diseases, University of Georgia, Athens, GA 30602, USA*

(*Received 7 May 2016; revised 14 July 2016; accepted 15 July 2016; first published online 20 October 2016*)

SUMMARY

Host–parasite associations are complex interactions dependent on aspects of hosts (e.g. traits, phylogeny or coevolutionary history), parasites (e.g. traits and parasite interactions) and geography (e.g. latitude). Predicting the permissive host set or the subset of the host community that a parasite can infect is a central goal of parasite ecology. Here we develop models that accurately predict the permissive host set of 562 helminth parasites in five different parasite taxonomic groups. We developed predictive models using host traits, host taxonomy, geographic covariates, and parasite community composition, finding that models trained on parasite community variables were more accurate than any other covariate group, even though parasite community covariates only captured a quarter of the variance in parasite community composition. This suggests that it is possible to predict the permissive host set for a given parasite, and that parasite community structure is an important predictor, potentially because parasite communities are interacting non-random assemblages.

Key words: FishPEST, species distribution model, boosted regression tree, parasite niche.

INTRODUCTION

Parasites are ubiquitous in nature, and have diverse life histories, transmission modes and ranges of host specificity (Poulin, 2011). Determining which host species a parasite can infect (referred to herein as the permissive host set or host range) could enable targeted host surveillance (Pilosof *et al.* 2015; Walton *et al.* 2016), estimation of zoonotic risk potential (Han *et al.* 2015) and insight into parasite specialization/generalism (Poulin, 2005). However, host–parasite interactions are complex, and may be influenced by environment (Locke *et al.* 2013), geography (Davies and Pedersen, 2008), coevolutionary history (Krasnov *et al.* 2012) or trait matching between host and parasite (Rohr *et al.* 2016). The relative importance of these constraints is generally not known, impeding prediction of currently undersampled or unknown host–parasite associations. Previous efforts to characterize parasite communities have largely focused on parasite richness (Ezenwa *et al.* 2006), topological network measures (Krasnov *et al.* 2012; Canard *et al.* 2014) and parasite community dissimilarity (Locke *et al.* 2013). However, few studies have focused on predicting which host species a parasite will infect (i.e. the permissive host set). As a result, the ability to predict the identity of permissive hosts (i.e. those that may be infected; host range) for a given parasite is limited and poses an obstructive knowledge gap. The development of predictive models could forecast parasite spillover to humans (Daszak *et al.* 2000) or wildlife (Colautti *et al.* 2004), and could provide an understanding of what factors underlie host–parasite associations.

One major barrier to the development of predictive models of the permissive host set is the low quality of data [but see Strona and Lafferty (2012*a*)], limited tools to access these data [but see Dallas (2016)] and a lack of cross-validated, accurate and predictive models. For instance, the PaNic model (Strona and Lafferty, 2012*b*) attempts to predict suitable fish host species given userweighted constraints (e.g. host trait weights or geographic constraints). While valuable, this tool has a limited feature space (e.g. only five host traits used for prediction), and cannot validate predictions, so there is no way to determine predictive accuracy. In essence, previous researchers asked which host species might be suitable for a parasite (Strona and Lafferty, 2012*a*), whereas we explored the factors that determine host suitability for a given parasite mechanistically, creating cross-validated models capable of predicting both *known* and *potential* permissive host set. Specifically, the *known* permissive host set would correspond to predictive accuracy of trained models on data not used to train the model, and the *potential* permissive host set would correspond to model-predicted suitable hosts without an observed occurrence record.

We used a large database on helminth infections of fish hosts (Strona *et al.* 2013) to develop predictive

* Corresponding author: University of Georgia, Odum School of Ecology, 140 East Green Street, Athens, GA 30606, USA. E-mail: tdallas@uga.edu

Parasitology (2017), **144**, 200–205. © Cambridge University Press 2016
doi:10.1017/S0031182016001608

https://doi.org/10.1017/S0031182016001608 Published online by Cambridge University Press

Table 1. Variables used to predict parasite occurrences

Variable	Units	Description	Range
Age at maturity	Years	Age at sexual maturity	0·1–34
Growth rate	Years	Rate to approach asymptotic length	0·02–9·87
Lifespan	Years	Estimated maximum age	0–145
Max length	cm	Maximum fish species length	1–2000
Trophic level	–	1 + mean trophic level of food	2–5
Area of occupancy	No. 1 × 1 ° cells	Global host distribution	1–1610
Geographic region	–	Biogeographic region	–
Latitude	max— min °	Latitudinal distribution	1–148
Longitude	max— min °	Longitudinal distribution	1–359
Parasite species richness	#	No. of parasite species of host species	0–89
Principal components	–	PCA axes of host–parasite network	−11·7–9·8
Principal coordinates	–	PCoA axes of taxonomic distance matrix	−1091–634

Row colours correspond to the different variable classes, and are consistent with colours in Figs 1 and 2 (pink for host traits, green for geographic variables, blue for parasite community traits and orange for host taxonomic variables).

models of the permissive host set for a number of parasite species (n = 562), using data on host traits, host taxonomy, geographic variables and parasite community variables. We found that the permissive host community is most constrained by the existing parasite community (i.e. parasite community variables), and much less so by geographic variables, host traits or host taxonomy. This suggests that the parasite community of a host species contains more information for predicting host–parasite associations and the permissive host set of a given parasite than the traditional host-centred approach, which places emphasis on host traits and taxonomy.

METHODS

Fish and parasite data

A global database of fish–parasite associations [FishPest (Strona *et al*. 2013)] was used, consisting of over 38 000 helminth parasite records from three phyla (Acanthocephala, Nematoda, Platyhelminthes), including examples from each major class of parasitic platyhelminthes (Cestoda, Monogenea, Trematoda). Our analyses focused on parasite species with at least 10 host occurrence records (n = 562). Though rare, duplicate host records did exist in the database, such that some parasites specialized on a small number of hosts (e.g. *Bicotylophora trachinoti* was only found on five *Trachinotus* species). Models were trained only on unique host–parasite combinations, which resulted in a range of specialist and generalist parasites with a variety of transmission modes, life histories and distributions.

Fish trait and geographic data were obtained from FishPest (Strona and Lafferty, 2012*a*; Strona *et al*. 2013) and FishBase (Froese and Pauly, 2000), a database of fish life history traits. Host trait variables included host age at sexual maturity, growth rate, lifespan, maximum body length and trophic level. Covariates relating to the geographic distribution of hosts included area of occupancy, biogeographic region (explained further in Supplemental Materials) and latitudinal/longitudinal range. Details of host trait and geographic variable determination are provided in Strona *et al*. (2013) and units are provided in Table 1. Missing covariate values were imputed through the iterative training of a machine learning algorithm (random forests), which imputes missing data based on averaged non-missing values weighted by proximity of observations in the random forest. This procedure was performed using the randomForest R package (Liaw and Wiener, 2002). Parasite community variables were calculated using the first five principal components of the host–parasite interaction network, a binary matrix where host species are rows, parasite species are columns, and known interactions between host and parasite are denoted as 1. This approach captures parasite community structure by attempting to compress the information contained in the entire interaction matrix into a smaller set of explanatory vectors. Other approaches calculate summary statistics of the interaction matrix, such as parasite species richness of each host. We incorporate both approaches, using both principal components and parasite richness per host species as covariates. Principal components were calculated using the princomp function, with the focal parasite removed so as not to introduce bias, meaning that each model used slightly different principal component vectors. The amount of variance explained by the first five principal component (PCA) vectors was not strongly influenced by this (σ = 0·0006). Reducing binary covariates to a small number of features is difficult, and five PCA vectors explained only a portion of the variation in parasite community composition (average variance

https://doi.org/10.1017/S0031182016001608 Published online by Cambridge University Press

explained = 27%). Additional PCA vectors explained little more variation in parasite community composition (see Supplemental Materials), so only the first five PCA vectors were used.

Host traits may not provide a complete representation of functional differences among hosts that could determine parasite occurrence. Meanwhile, host taxonomy may capture unmeasured host trait variation. To incorporate host taxonomy into our analyses, we calculated the pairwise taxonomic dissimilarity between all host species using the vegan package in R, and then performed a Principal Coordinates Analysis on the taxonomic dissimilarity matrix to obtain a reduced set of features containing information on taxonomic dissimilarity of fish hosts. This procedure attempts to compress information contained in the host taxonomic dissimilarity matrix into a set of host-level covariates. Measures of taxonomic distinctiveness calculate the mean taxonomic distance for a host species to all other species, while our approach better captures the variation within the taxonomic dissimilarity matrix. We used the first five vectors from this analysis, which explained 40·5% of total variation in the host taxonomic dissimilarity matrix, to represent host taxonomic relationships.

The absence of a recorded host–parasite interaction does not rule out the existence of an association. One approach to account for this, originally developed for species distribution modelling, is to sample 'background' interactions, which we define as a random set of possible hosts for a focal parasite. In species distribution modelling, this would include both observed (i.e. occurrences) and unobserved interactions (Elith *et al.* 2011) aimed at characterizing the distribution of predictor variables across space. However, since parasite occurrence data were so limited, and duplicating hosts in the training data could introduce bias to our models, we sampled background data from the subset of host species where the focal parasite was not observed. Background data consisted of a random sample of 20 times the number of recorded parasite occurrences, which maintained a constant proportion of parasite presence values to background points. The use of background data is essential for model development, since the background sample represent the plausible ranges of host covariates (see Table 1), creating a parameter space of host covariates. The number of background hosts considered did not influence model predictions (see Supplementary Materials). Together, our data consisted of parasite occurrence records, which included both known parasite occurrences and background data, and a set of host-level covariates (Table 1) with which to predict parasite occurrence. These data were split into training (70%) and test (30%) sets in order to reduce model overfitting, and to evaluate predictive model accuracy on independent data (i.e. the test set).

Predictive model formulation

Boosted regression trees (BRT), a powerful analysis that bypasses known issues with alternative regression analyses (Elith *et al.* 2008), were trained using R package gbm (Ridgeway, 2015) (learning rate = 0·001; interaction depth = 4). The optimal number of trees was determined using 'out-of-bag' estimation (max trees = 50 000). Boosting is a process where weak learning trees are built iteratively, and each tree is dependent on the residuals of the previous tree. This creates an ensemble of trees, which are then combined to create the final model. In our application, we trained models using geographic covariates, host traits, host taxonomy or parasite community covariates as a means to predict parasite occurrence on hosts in the independent test data.

Relative contribution (R_c) values for each covariate were determined by permuting each predictor (i.e. randomizing to remove any potential influence on resulting model) and determining the resulting change in model performance. Some species could be found in many biogeographic regions, so binary dummy variables were used for this covariate, and the relative contribution value reported here is the sum of the relative contribution value for all binary dummy variables (see Supplementary Materials for further discussion). Accuracy was assessed using AUC (the area under the receiver operating characteristic curve), a measure scaled between 0 (perfectly inaccurate prediction) to 1 (perfect prediction), with a value of 0·5 corresponding to random prediction (Flach, 2010). Here, AUC measures the frequency with which a randomly sampled host species is parasitized would be more highly ranked by the BRT model than a randomly-selected host from the background data. A null model, representing the random prediction case discussed above, was used for comparison to trained models. Here, occurrences were randomly predicted, constrained by the total number of parasite occurrences in the test set (i.e. the number of parasite occurrences was kept constant).

RESULTS

All models performed better than null predictions (mean ± SE; null model = 0·50 ± 0·0003). Trained models predicted the permissive host sets of parasite species with varying levels of accuracy (Fig. 1), depending on whether models were trained with data corresponding to host traits ($\overline{AUC} = 0{\cdot}69 \pm 0{\cdot}007$), host taxonomic distinctiveness ($\overline{AUC} = 0{\cdot}84 \pm 0{\cdot}006$), geography ($\overline{AUC} = 0{\cdot}85 \pm 0{\cdot}005$), parasite community composition ($\overline{AUC} = 0{\cdot}90 \pm 0{\cdot}005$) or all covariates ($\overline{AUC} = 0{\cdot}94 \pm 0{\cdot}003$).

The relative importance of covariates in submodels (models trained on only one covariate class) approximately retained their rank order position

https://doi.org/10.1017/S0031182016001608 Published online by Cambridge University Press

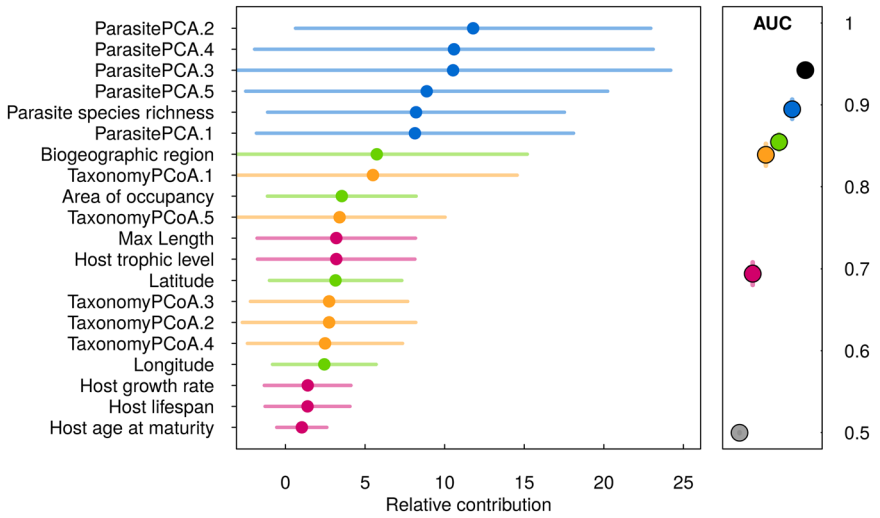


Fig. 1. Relative contribution values (mean ± S.D.) from the full model (left panel; black point in right panel) demonstrate the importance of parasite community variables (blue) relative to host traits (pink), taxonomy (orange) and geographic covariates (green). Numbers following host taxonomic (TaxonomyPCoA) and parasite community (ParasitePCA) covariates refer to principal component axis number. Other variable definitions and units are available in Table 1. Accuracy of submodels (mean ± 2 × S.D.) trained on covariate groups (right panel) all performed better than a null expectation (grey point).

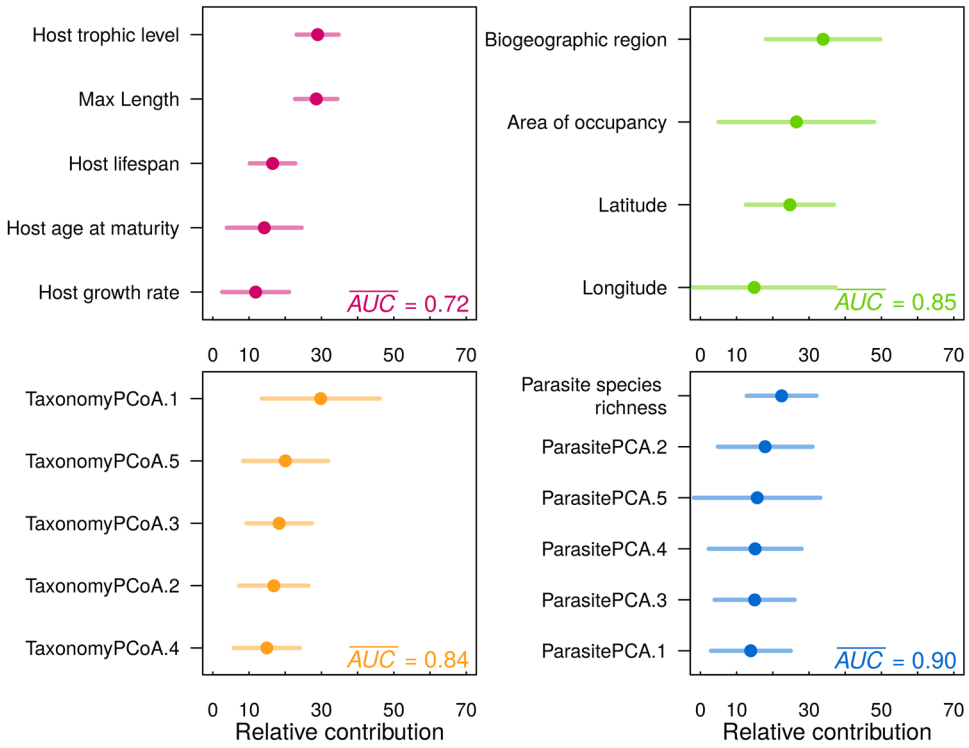


Fig. 2. Relative importance (mean ± S.D.) and average model accuracy for each submodel trained on either host traits (top left), geographic covariates (top right), host taxonomic covariates (bottom left) or parasite community covariates (bottom right). Numbers following host taxonomic (TaxonomyPCoA) and parasite community (ParasitePCA) covariates refer to principal component axis number.

(Fig. 2) when compared with the full model (Fig. 1). Principal component (ParasitePCA) and principal coordinate (TaxonomyPCoA) axes change rank in submodels relative to the full model, likely because of the small difference in relative variable importance, and the broad overlap in error bars. This means that parasites responded differentially to PCA axes, suggesting that unique information is contained in each axis. Model performance did not differ by parasite taxonomic group (Supplementary Figure A2) or host specificity (Supplementary Figure A3).

https://doi.org/10.1017/S0031182016001608 Published online by Cambridge University Press

DISCUSSION

Parasite community covariates were the most important variables to the full model, and had the highest submodel accuracy, despite the limited amount of variance explained by the decomposition of parasite community composition (i.e. the five PCA vectors). This suggests that parasite species cooccurrence data contain valuable information, either on unmeasured host trait variation, or on forces underlying parasite community composition among hosts (e.g. parasite competition, facilitation or community assembly). It is unlikely that the parasite community vectors could relate to unmeasured host trait variation, as the selected host traits have previously been related to parasite richness (Lindenfors *et al.* 2007) and occurrence probability (Strona and Lafferty, 2012*b*). Further, unmeasured trait variation accounted for by parasite community covariates would have be entirely separate from other covariates (i.e. collinearity between unmeasured variation and measured host traits), which seems unlikely given known scaling relationships in host traits (Killen *et al.* 2010; Rall *et al.* 2012). On the other hand, if parasite community covariates were able to capture interactions among parasites, structural constraints to infection or the identity of coinfecting parasites, it is plausible to conclude that helminth parasites of fish form non-random, interacting communities. Whether this is the case or not, parasite community composition was key in predicting host community composition for a large number of parasites.

Previous studies have found host traits to be associated with parasite diversity (Locke *et al.* 2014) and parasite community similarity (Davies and Pedersen, 2008). Further, host taxonomy and geography have been found to be associated with parasite occurrence probability (Strona *et al.* 2013). Our findings do not disagree with these previous studies. Models trained using data on geographic covariates, host taxonomy and host traits still performed much better than our null model, suggesting that these groups of variables contain important information. Our findings do suggest, however, that parasite community structure may be more important than both host attributes and geography. We posit two hypotheses to explain why parasite community variables are important in predicting which hosts a given parasite infects. First, similar parasites may infect similar host species, resulting in the parasite community variables capturing parasite trait information. Second, parasites interact over long timescales, resulting in quantifiable non-random co-occurrence patterns capable of accurate host community prediction. The testing of these hypotheses is an undertaking that will require data on parasite traits and evolutionary histories, in conjunction with large databases of host–parasite interactions, such as the London Natural History Museum's helminth database (Gibson *et al.* 2005) and the Global Mammal Parasite Database (Nunn and Altizer, 2005).

There is mixed evidence for the predictability of parasite community structure in fish parasites, despite the existence of an extensive body of literature on the subject (Holmes, 1990; Kennedy, 1990, 2009; Poulin and Rohde, 1997; Sasal *et al.* 1999; Poulin and Valtonen, 2002; Rohde, 2002). As it is possible to predict the permissive host set using parasite community composition data, it may be possible to predict parasite community composition from the permissive host set. Further, given that parasite community composition was important, it is plausible that information on parasite traits, such as transmission mode, feeding behaviour or host tissue infected may increase the accuracy with which the permissive host set is determined. Interestingly, we found no difference in model accuracy as a function of parasite taxonomic group or number of hosts infected (see Supplementary Materials). This is intriguing, as it suggests that prediction of the host community (i.e. the permissive host set) may be possible even for specialized parasites or those with complex life histories. The application of our analytical approach to a broader range of host and parasite taxa, considering different spatial scales and incorporating information on more host and parasite traits may help discern when the permissive host set is predictable, and identify potential likely hosts for a given parasite species. Lastly, the compilation and curation of parasite trait databases, in the style of current host trait databases (Froese and Pauly, 2000), is necessary for the accurate forecasting of parasite occurrences on novel hosts.

SUPPLEMENTARY MATERIAL

The supplementary material for this article can be found at https://doi.org/10.1017/S0031182016001608. Data and R code to reproduce the analyses is provided at https://doi.org/10.6084/m9.figshare.3795330.

ACKNOWLEDGEMENTS

We thank Kevin Lafferty for providing comments on an earlier manuscript draft. The Macroecology of Infectious Disease Research Coordination Network (funded by NSF DEB 131223) provided useful discussions and support for this work.

FINANCIAL SUPPORT

T. A. D., A. W. P., and J. M. D. were supported by the Odum School of Ecology. T. A. D. is currently supported by the University of California at Davis.

REFERENCES

Canard, E., Mouquet, N., Mouillot, D., Stanko, M., Miklisova, D. and Gravel, D. (2014). Empirical evaluation of neutral interactions in host–parasite networks. *The American Naturalist* **183**, 468–479.

https://doi.org/10.1017/S0031182016001608 Published online by Cambridge University Press

Colautti, R. I., Ricciardi, A., Grigorovich, I. A. and MacIsaac, H. J. (2004). Is invasion success explained by the enemy release hypothesis? *Ecology Letters* **7**, 721–733.

Dallas, T. (2016). helminthR: an R interface to the London Natural History Museum's host–parasite database. *Ecography* **39**, 391–393.

Daszak, P., Cunningham, A. A. and Hyatt, A. D. (2000). Emerging infectious diseases of wildlife–threats to biodiversity and human health. *Science* **287**, 443–449.

Davies, T. J. and Pedersen, A. B. (2008). Phylogeny and geography predict pathogen community similarity in wild primates and humans. *Proceedings of the Royal Society B: Biological Sciences* **275**, 1695–1701.

Elith, J., Leathwick, J. R. and Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology* **77**, 802–813.

Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E. and Yates, C. J. (2011). A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions* **17**, 43–57.

Ezenwa, V. O., Price, S. A., Altizer, S., Vitone, N. D. and Cook, K. C. (2006). Host traits and parasite species richness in even and odd-toed hoofed mammals, Artiodactyla and Perissodactyla. *Oikos* **115**, 526–536.

Flach, P. A. (2010). ROC analysis. In *Encyclopedia of Machine Learning* (ed. Sammut, C. and Webb, G. I.), pp. 869–875. Springer US.

Froese, R. and Pauly, D. (ed.) (2000). *FishBase 2000: Concepts, Design and Data Sources*. ICLARM, Los Baños, Laguna, Philippines. 344 p.

Gibson, D., Bray, R. and Harris, E. (2005). Host–parasite database of the natural history museum, London. http://www.nhm.ac.uk/research-curation/scientific-resources/taxonomy-systematics/host-parasites/database/index.jsp

Han, B. A., Schmidt, J. P., Bowden, S. E. and Drake, J. M. (2015). Rodent reservoirs of future zoonotic diseases. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 7039–7044.

Holmes, J. C. (1990). Helminth communities in marine fishes. In *Parasite Communities: Patterns and Processes* (ed. Esch, G. W., Bush, A. O. and Aho, J. M.), pp. 101–130. Springer, Netherlands.

Kennedy, C. (1990). Helminth communities in freshwater fish: structured communities or stochastic assemblages? In *Parasite Communities: Patterns and Processes* (ed. Esch, G. W., Bush, A. O. and Aho, J. M.), pp. 131–156. Springer, Netherlands.

Kennedy, C. (2009). The ecology of parasites of freshwater fishes: the search for patterns. *Parasitology* **136**, 1653–1662.

Killen, S. S., Atkinson, D. and Glazier, D. S. (2010). The intraspecific scaling of metabolic rate with body mass in fishes depends on lifestyle and temperature. *Ecology Letters* **13**, 184–193.

Krasnov, B. R., Fortuna, M. A., Mouillot, D., Khokhlova, I. S., Shenbrot, G. I. and Poulin, R. (2012). Phylogenetic signal in module composition and species connectivity in compartmentalized host–parasite networks. *The American Naturalist* **179**, 501–511.

Liaw, A. and Wiener, M. (2002). Classification and regression by randomForest. *R News* **2**, 18–22.

Lindenfors, P., Nunn, C. L., Jones, K. E., Cunningham, A. A., Sechrest, W. and Gittleman, J. L. (2007). Parasite species richness in carnivores: effects of host body mass, latitude, geographical range and population density. *Global Ecology and Biogeography* **16**, 496–509.

Locke, S. A., McLaughlin, J. D. and Marcogliese, D. J. (2013). Predicting the similarity of parasite communities in freshwater fishes using the phylogeny, ecology and proximity of hosts. *Oikos* **122**, 73–83.

Locke, S. A., Marcogliese, D. J. and Valtonen, E. T. (2014). Vulnerability and diet breadth predict larval and adult parasite diversity in fish of the Bothnian bay. *Oecologia* **174**, 253–262.

Nunn, C. L. and Altizer, S. M. (2005). The global mammal parasite database: an online resource for infectious disease records in wild primates. *Evolutionary Anthropology: Issues, News, and Reviews* **14**, 1–2.

Pilosof, S., Morand, S., Krasnov, B. R. and Nunn, C. L. (2015). Potential parasite transmission in multi-host networks based on parasite sharing. *PLoS ONE* **10**, e0117909.

Poulin, R. (2005). Relative infection levels and taxonomic distances among the host species used by a parasite: insights into parasite specialization. *Parasitology* **130**, 109–115.

Poulin, R. (2011). Interactions between species and the parasite niche. In *Evolutionary Ecology of Parasites*, pp. 188–207. Princeton University Press, New Jersey, USA.

Poulin, R. and Rohde, K. (1997). Comparing the richness of metazoan ectoparasite communities of marine fishes: controlling for host phylogeny. *Oecologia* **110**, 278–283.

Poulin, R. and Valtonen, E. T. (2002). The predictability of helminth community structure in space: a comparison of fish populations from adjacent lakes. *International Journal for Parasitology* **32**, 1235–1243.

Rall, B. C., Brose, U., Hartvig, M., Kalinkat, G., Schwarzmüller, F., Vucic-Pestic, O. and Petchey, O. L. (2012). Universal temperature and body-mass scaling of feeding rates. *Philosophical Transactions of the Royal Society B* **367**, 2923–2934.

Ridgeway, G. (2015). *gbm: Generalized Boosted Regression Models*. R package version 2.1.1. https://CRAN.R-project.org/package=gbm

Rohde, K. (2002). Ecology and biogeography of marine parasites. *Advances in Marine Biology* **43**, 1–83.

Rohr, R. P., Naisbit, R. E., Mazza, C. and Bersier, L. F. (2016). Matching–centrality decomposition and the forecasting of new links in networks. *Proceedings of the Royal Society B: Biological Sciences* **283**, 20152702.

Sasal, P., Trouvé, S., Müller-Graf, C. and Morand, S. (1999). Specificity and host predictability: a comparative analysis among monogenean parasites of fish. *Journal of Animal Ecology* **68**, 437–444.

Strona, G. and Lafferty, K. D. (2012*a*). FishPEST: an innovative software suite for fish parasitologists. *Trends in Parasitology* **28**, 123.

Strona, G. and Lafferty, K. D. (2012*b*). How to catch a parasite: parasite Niche Modeler (PaNic) meets Fishbase. *Ecography* **35**, 481–486.

Strona, G., Palomares, M. L. D., Bailly, N., Galli, P. and Lafferty, K. D. (2013). Host range, host ecology, and distribution of more than 11 800 fish parasite species. *Ecological Archives: E094-045* **94**, 544–544.

Walton, L., Marion, G., Davidson, R. S., White, P. C., Smith, L. A., Gavier-Widen, D., Yon, L., Hannant, D. and Hutchings, M. R. (2016). The ecology of wildlife disease surveillance: demographic and prevalence fluctuations undermine surveillance. *Journal of Applied Ecology*. doi: 10.1111/1365-2664.12671.

https://doi.org/10.1017/S0031182016001608 Published online by Cambridge University Press