Oikos 120:740–747, 2011
doi: 10.1111/j.1600-0706.2010.19036.x

Subject Editor: Rampal S. Etienne. Accepted 7 September 2010

Identifying hotspots of parasite diversity from species–area relationships: host phylogeny versus host ecology

Robert Poulin, François Guilhaumon, Haseeb S. Randhawa, José L. Luque and David Mouillot

R. Poulin (robert.poulin@stonebow.otago.ac.nz) and H. S. Randhawa, Dept of Zoology, Univ. of Otago, PO Box 56, Dunedin 9054, New Zealand. – F. Guilhaumon and D. Mouillot, UMR CNRS-UMII 5119 Ecosystemes Lagunaires, Univ. de Montpellier II, CC093, FR–34095 Montpellier Cedex 5, France. – J. L. Luque, Depto de Parasitologia Animal, Univ. Federal Rural do Rio de Janeiro, Caixa Postal 74.508, CEP 23851-970, Seropédica RJ, Brazil.

Interspecific variation in parasite species richness among host species has generated much empirical research. As in comparisons among geographical areas, controlling for variation in host body size is crucial because host size determines resource availability. Recent developments in the use of species–area relationships (SARs) to detect hotspots of biodiversity provide a powerful way to control for host body size, and to identify 'hot' and 'cold hosts' of parasite diversity, i.e. hosts with more or fewer parasites than expected from their size. Applying SAR modelling to six large datasets on parasite species richness in vertebrates, we search for hot and cold hosts and assess the effect of other ecological variables on the probability that a host species is hot/cold taking body size (and sampling effort) into account. Five non-sigmoid SAR models were fitted to the data by optimisation; their relative likelihood was evaluated using the Bayesian information criterion, before deriving an averaged SAR function. Overall, the fit between the five SAR models and the actual data was poor; there was substantial uncertainty surrounding the fitted models, and the best model differed among the six datasets. These results show that host body size is not a strong or consistent determinant of parasite species richness across taxa. Hotspots were defined as host species lying above the upper limit of the 80% confidence interval of the averaged SAR, and coldspots as species lying below its lower limit. Our analyses revealed (1) no apparent effect of specific ecological factors (i.e. water temperature, mean depth range, latitude or population density) on the likelihood of a host species being a hot or coldspot; (2) evidence of phylogenetic clustering, i.e. hosts from certain families are more likely to be hotspots (or coldspots) than other species, independently of body size. These findings suggest that host phylogeny may sometimes outweigh specific host ecological traits as a predictor of whether or not a host species harbours more (or fewer) parasite species than expected for its size.

The search for the determinants of parasite diversity has a long history in ecology (Gregory et al. 1996, Poulin 1997, Poulin and Morand 2004, Bordes et al. 2009). Variation in parasite species richness among host species provides not only a good model for studies of community diversification, but is also of great interest in the context of predicting disease risk for conservation targets (Nunn et al. 2003, Poulin and Morand 2004). In comparisons among host species, just like in comparisons among geographical areas (Rosenzweig 1995, Rosenzweig and Sandlin 1997), one must control for variation in habitat size in order to identify other important drivers of species diversity. Habitat size, for parasites, can correspond more closely to host body size than to the surface area of the geographic region in which they occur (Poulin and Morand 2004). The average body size of the host species represents the extent of the typical habitat patch in which a parasite lives, and is thus proportional to the amount of resources available (Kuris et al. 1980, Poulin and Morand 2004). Host body size also correlates with host lifespan, with larger hosts providing longer-lived habitats. There have therefore been several arguments (though with some caveats) proposing that, all else being equal, larger-bodied host species should harbour richer parasite faunas than related but smaller-bodied species (Kuris et al. 1980, Poulin 1995, Gregory et al. 1996, Nunn et al. 2003, Poulin and Morand 2004). Comparative evidence provides support for a general link between host body size and parasite species richness across related host species, for a wide range of host and parasite taxa (Poulin 1997, Poulin and Morand 2004).

However, previous comparative analyses have only identified correlations between host body size and parasite species richness, without determining the shape of the relationship linking them. Also, the correlation coefficients between host size and parasite richness are generally not very strong or convincing (Poulin 1997, 2004, Nunn et al. 2003, Poulin and Morand 2004), suggesting that the function linking host body size and parasite richness is more complex than a straight line.

This is where recent developments in related fields can provide a way forward. Areas of unusually high biodiversity, or hotspots where the number of species exceeds what is expected for the size of the area, have long been of great interest to evolution, ecology and conservation biology. In this context, species–area relationships (SARs) are widely

used to quantify the link between the size of an area and the richness of the species it supports (He and Legendre 1996, Lomolino 2000). The shape of these relationships can shed light on the processes generating them, though there is no consensus regarding the best function to describe them (Williams et al. 2009). Recently, SARs have been used to identify hotspots of plant and animal diversity as those areas that support significantly more species than what is predicted by the SAR function best-fitting the data (Veech 2000, Hobohm 2003, Ulrich and Buszko 2005, Fattorini 2006, Guilhaumon et al. 2008). Thus, SARs may prove very useful not only to understand how biodiversity is generated and maintained, but also to pinpoint areas of high conservation priority. With respect to parasite diversity, by substituting area size for host body size, we can use recent methodological developments in the SAR approach to fit a wider range of functions to the relationship between parasite richness and host body size. In addition, this method can identify 'hot hosts' and 'cold hosts' of parasite biodiversity, i.e. host species that have accumulated a very different (higher or lower) number of parasites than what would be expected from their body size.

Other drivers of parasite biodiversity may account for the unusually high parasite species richness in some host species. In addition to host body size, factors typically correlated with the accumulation of many parasite species include latitude (or other variables associated with environmental conditions), host geographical range, host population density, and host diet, among others (Poulin 1995, 1997, Nunn et al. 2003, Poulin and Morand 2004, Lindenfors et al. 2007, Bordes et al. 2009). Using the SAR approach to correct for any host body size effects as well as identifying host species harbouring extreme numbers of parasites, we can then better assess the effect of these other variables as determinants (or mere correlates) of parasite diversity.

Here, we combine SAR modelling and regression analyses on six large data sets including comparative data on the species richness of parasites in vertebrate hosts, in order to: (1) investigate the strength and shape of the relationship between host body size and parasite species richness, (2) identify 'hot hosts' and 'cold hosts' of parasite diversity, or host taxa harbouring more or fewer parasites than expected from their size, and (3) assess the effect of likely ecological predictors of parasite richness on the probability that a host species is a hotspot/coldspot of parasite diversity taking body size (and sampling effort) into account. The analyses presented here provide strong evidence that host body size has only a weak influence on parasite diversity, and that the latter is generally dependent on complex interactions between the host's phylogenetic affinities and its ecology.

Methods

Datasets

The six datasets used in the present analyses have been previously compiled for studies of parasite diversity (Poulin 1995, Poulin and Mouillot 2004, Luque and Poulin 2007, 2008, Randhawa and Poulin 2010). They all include data on parasite species richness, sampling effort and host body size, along with data on one additional continuous variable that was the most likely covariate of parasite species richness among the ones tested in the original studies (Table 1, Supplementary material Appendix 1 for full datasets). In brief, the datasets are:

(1) 'Metazoan parasites of 338 marine teleost fishes', and (2) 'Metazoan parasites of 259 freshwater teleost fishes'. Data on the species richness of all ecto- and endoparasitic metazoans of teleosts (all in the class Actinopterygii) from the Neotropical region were taken from Luque and Poulin (2007, 2008). Parasite species richness is calculated across each host's geographical range, based on published records. Sampling effort was measured as the total number of publications on each fish species found in the Zoological Record database. Host body size was measured as maximum total body length, based on information in Fish Base (Froese and Pauly 2006). There may be occasional uncertainty regarding entries in Fish Base, e.g. total length confused with standard length; however, as our analyses compare species ranging in size over several orders of magnitude, these small errors do not affect the results. The additional predictor variable examined was the mean water temperature of the distribution area of each fish species.

(3) 'Cestodes of 127 sharks', and (4) 'Cestodes of 172 batoids (skates and rays)'. Data on the species richness of cestode parasites in elasmobranchs were taken from Randhawa and Poulin (2010). Cestode species richness was calculated as above. Sampling effort was measured as the number of publications on each fish species related to parasitism found in the Zoological Record database. Host body size was measured as the maximum length from the tip of the snout to the mid-point of the pelvic fins, based on information in Fish Base (Froese and Pauly 2006) and Compagno et al. (2005). The additional predictor variable examined was the mid-point of the depth range of the host species, though data were not available for all host species.

(5) 'Helminths of 76 birds'. Data on the species richness of helminth parasites of birds were taken from Poulin (1995). Helminth species richness is measured as the highest value recorded in one host population. Sampling effort was measured as the number of individual hosts examined for helminths in that population. Host body size was measured as average body mass for each species (from Dunning 1993). The additional predictor variable examined was the latitude of the sampling locality.

(6) 'Helminths of 110 mammals'. Data on the species richness of helminth parasites of mammals were taken from Poulin and Mouillot (2004). Helminth species richness, sampling effort and host body size were measured as for birds (see Poulin and Mouillot 2004 for data sources for host body masses). The additional predictor variable examined was host population density (taken from Damuth 1987), though data on density were not available for all host species.

Measuring parasite species richness

First, we needed estimates of parasite species richness that were independent of sampling effort, since these two variables are generally positively correlated (Walther et al. 1995, Poulin and Morand 2004). We included only entries with sampling effort higher than 24 host individuals for the

Table 1. Overview of the six host-parasite datasets used in the analyses; see Supplementary material Appendix 1 for further details.

Hosts	n	Parasites	Body size measure (range)	Study effort measure	Other ecological predictor	Source
Marine fishes	338	all metazoans	maximum total length, cm (7.5–500)	hits on Zoological Record*	local water temperature	Luque and Poulin 2007, 2008
Freshwater fishes	259	all metazoans	maximum total length, cm (2.5–450)	hits on Zoological Record*	local water temperature	Luque and Poulin 2007, 2008
Sharks	127	cestodes	maximum body length, cm (14–577)	hits on Zoological Record*	mid-point of depth range**	Randhawa and Poulin 2010
Batoids (skates and rays)	172	cestodes	maximum body length, cm (12.5–265)	hits on Zoological Record*	mid-point of depth range**	Randhawa and Poulin 2010
Birds	76	helminths	average body mass, g (10–5811)	number of hosts examined	latitude	Poulin 1995
Mammals	110	helminths	average body mass, g (4–3 500 000)	number of hosts examined	population density**	Poulin and Mouillot 2004

*number of publications on each fish species found in a search of the electronic database.
**data not available for all host species.

bird and mammal datasets, and higher than nine published records for the other datasets. These threshold values are somewhat arbitrary; they were chosen based on the distribution of sampling effort values as a compromise between the needs to eliminate poorly-studied host species and retain sufficient points for subsequent analyses. Preliminary tests indicate that our results are very robust to changes in these threshold values. Using the remaining values, for each dataset we regressed log-transformed species richness against log-transformed sampling effort, and used residual richness in subsequent analyses. This is justified given the expected saturation of richness with increasing sampling effort, and because this procedure yields residuals whose distribution is more suitable for model-fitting.

However, when plotting parasite richness against host body size, richness residuals include negative values that cannot be handled by the SAR model-fitting algorithms. They were therefore 'positivised' as follows:

$$Y_{new} = Y_{res} + \left|(\min(Y_{res}))\right| + 0.1$$

where Y_{res} are the residuals from the above regressions and Y_{new} are the values used for model-fitting. This results in a vertical translation of residual values along the *y*-axis, so that they are all positive. This has no effect on the structural parameters of the models (i.e. their slope) or on model selection results.

SAR modelling

Our general approach follows that proposed by Guilhaumon et al. (2008) to cope with uncertainties in SAR modelling. All analyses described below were implemented within the R statistical programming environment (R Development Core Team 2009) using the "mmSAR" package (Guilhaumon et al. 2010).

Given the convex shape of all scatterplots of parasite richness, i.e. modified residuals, versus host body size, we investigated only non-sigmoid (both asymptotic and non asymptotic) SAR models: power, exponential, negative exponential, Monod and rational functions. Each of these five models was fitted by minimizing the residual sum of squares in non-linear regressions using the unconstrained Nelder-Mead optimisation algorithm (Dennis and Schnabel 1983, Guilhaumon et al. 2008). We used r^2-values that compare the fit of non-linear regression models with that of a linear intercept-only model (Kvalseth 1985), as indicators of the proportion of the total variation among host species in parasite species richness that is explained (accounted for) by regressions against host body size.

We discriminated among the five above SAR models within a model selection framework (Burnham and Anderson 2002). Using the Bayesian information criterion (BIC), we evaluated and compared the relative support of each parameterized function for each dataset. The lower the BIC associated with a model, the better this model is at explaining the data. Like the Akaike information criterion (AIC), the BIC is widely used in model selection (Burnham and Anderson 2002, Johnson and Omland 2004). Our results are robust to the choice of criterion used for model selection and averaging; we opted for BIC because it is more penalising with respect to the number of model parameters. BIC weights, normalized across the set of models to sum to one, were derived to evaluate the relative likelihood of each of the five SAR models.

Model selection uncertainty may arise when several different models are equally supported by the data, an outcome that invalidates relying only on the best model. To avoid uncertainty, a multimodel inference is preferable (Burnham and Anderson 2002). Therefore, we used model averaging and considered the weighted average of model predictions with respect to model BIC weights to derive an averaged SAR function (Guilhaumon et al. 2008). We checked model averaging results for bias inherent to the violation of either the normality or homoscedasticity of residuals (Lilliefors extension of the Kolmogorov normality test and Pearson's product moment correlation coefficient with host body size, respectively).

Hotspot/coldspot analyses

In conservation biology, hotspots of biodiversity have been identified using SAR, based on departures from the regression line. Residuals have been used repeatedly (Veech 2000, Hobohm 2003, Fattorini 2006) but they do not provide a criterion to distinguish between true hotspots and other

Table 2. Results of model fitting procedure for five SAR models, and their multimodel average, in each of the six parasite-host datasets. ΔBIC are differences in BIC (Bayesian information criterion) between a given model and the best one for that dataset, and r^2 are coefficients of determination.

Dataset	No. host species	Power		Exponential		Negative exponential		Monod		Rational		Multimodel average
		ΔBIC	r^2	ΔBIC	r^2	ΔBIC	r^2	ΔBIC	r^2	ΔBIC	r^2	r^2
Metazoans/marine fish	182	6.29*	0.00	6.29*	0.00	0*	0.03	3.23*	0.02	8.03*	0.02	0.03
Metazoans/freshwater fish	118	1.58	0.01	1.46	0.01	0.29	0.02	0	0.03	5.67	0.02	0.02
Cestodes/sharks	32	4.64	0.05	4.24	0.07	0	0.18	2.02	0.13	5.11	0.14	0.16
Cestodes/batoids	31	0	0.23	0.02	0.23	1.39	0.20	0.38	0.23	3.50	0.23	0.23
Helminths/birds	64	0	0.14	0.16	0.14	5.43*	0.06	3.44	0.09	4.47*	0.13	0.14
Helminths/mammals	89	0	0.17	0.82	0.16	16.57	0.00	15.06	0.02	0.29	0.21	0.19

*model that failed to respect the regression assumptions (normality of residuals and/or homoscedasticity).

areas of high species richness (Ulrich and Buszko 2005), and can lead to the unsettling conclusion that functions with the poorest fit with the data are the best at identifying hotspots (Veech 2000). Here, we used instead the position of host species relative to the confidence interval of the multimodel averaged SAR. Confidence intervals were devised to take into account model selection and parameter estimation uncertainties by using a nonparametric bootstrapping procedure (Guilhaumon et al. 2008). For each dataset, we defined hotspots as those host species lying above the upper limit of the 80% confidence interval; these correspond to host species with a parasite species richness much higher than the value expected from the averaged SAR. Similarly, we defined coldspots as those host species falling below the lower limit of the 80% confidence interval, a subset showing much lower species richness than expected for their body size.

Whether a host species is a hotspot or a coldspot of parasite richness may depend either on its phylogenetic origins or its ecological characteristics, or both. Since many parasites are inherited by daughter host species from their ancestors through cospeciation, related host species are expected to harbour similar parasite faunas (Vickery and Poulin 1998, Poulin and Morand 2004), just as unrelated host species may accumulate similar numbers of parasites because of shared ecological characteristics. For each dataset, we tested for the effect of host taxonomy, host ecology, and their interaction on the likelihood of a host species being a hotspot/coldspot using binomial ANCOVAs. Three sets of analyses were performed, one in which host species were categorised as either hotspots or non-hotspots, one in which they were categorised as either coldspots or non-coldspots, and the other in which only hotspots versus coldspots were considered; all analyses were run with a binomial error structure and logit link. Host taxonomy was initially considered at two different levels: order and family. In preliminary analyses, there were very few significant effects when 'order' was used as a factor, therefore here we only report the outcome of analyses using host family as factor. For each dataset, the host ecological characteristic entered in the ANCOVAs was the one identified as the best predictor of parasite species richness in the study from which the data were taken (Table 1); all of them are continuous variables. Significance levels are based on the deviance explained by each factor and their interaction, based on χ^2-statistics.

Results

All six original datasets are available in the Supplementary material Appendix 1. The number of host species included in analysis of each of the six data sets after exclusion of poorly-sampled species are shown in Table 2. For the mammal dataset, an additional host species was excluded: the pilot whale *Globicephala melaena*. This host is five times larger than the second-heaviest mammal species in the dataset, and represented an extreme outlier; excluding it was necessary to achieve a more homogeneous distribution of body sizes along the x-axis.

The results of model fitting are shown in Table 2 and those of model selection in Table 3. For metazoan parasites of marine fishes, all models failed to produce normally distributed residuals, as did the negative exponential and rational functions for helminths of birds. However, this lack of fit did not affect model averaging results because in the case of

Table 3. Results of model selection among five SAR models in each of the six parasite-host datasets. The values correspond to model weights, based on BIC (Bayesian information criterion), and are equivalent to the probabilities of each model providing the best fit to the data; combined weights of non-asymptotic models (power and exponential) and asymptotic models (negative exponential, monod and rational) are also shown.

Dataset	Power	Exponential	Non-asymptotic models	Negative exponential	Monod	Rational	Asymptotic models
Metazoans/marine fish	0.03*	0.03*	0.06	0.77*	0.15*	0.02*	0.94
Metazoans/freshwater fish	0.16	0.17	0.33	0.30	0.35	0.02	0.67
Cestodes/sharks	0.06	0.07	0.13	0.60	0.22	0.05	0.87
Cestodes/batoids	0.29	0.28	0.57	0.14	0.24	0.05	0.43
Helminths/birds	0.44	0.41	0.85	0.03*	0.08	0.04*	0.15
Helminths/mammals	0.40	0.26	0.66	0.00	0.00	0.34	0.34

*model that failed to respect the regression assumptions (normality of residuals and/or homoscedasticity).

metazoan parasites of marine fishes, the empirical distribution of residuals was bell-shaped and box-plot analyses did not reveal any skewness, and for helminths of birds the BIC weights of poor models were negligible (Table 3).

Some points are noteworthy. Firstly, r^2 values are generally low, in particular for small host body sizes, where parasite species richness is often either much lower or higher than predicted by the models (Fig. 1). The lack of fit between the five SAR models and the actual data is especially apparent for marine and freshwater fish: in these hosts, the highest values of parasite species richness are observed in small-to-mid-sized host species (Fig. 1). Second, based on BIC weights, there is generally substantial uncertainty surrounding the fitted models, and the best model differs among the six datasets (Table 2, 3). When comparing the combined weights of non-asymptotic models (power and exponential) with those of asymptotic models (negative exponential, Monod and rational), neither class of models consistently outperforms the other. Asymptotic models provide a better fit to three datasets, and non-asymptotic models provide a better fit to the other three (Table 3), suggesting that there is no universal tendency toward saturation of parasite species richness in large-bodied host species. All this is strong justification for the use of a multimodel SAR for coldspot and hotspot detection.

In some datasets, data on the chosen ecological variable were not available for all host species, and thus analyses of the factors associated with hotspots included fewer species than the preceding SAR analyses (compare Table 2 and 4). Whether looking at hotspots versus non-hotspots, coldspots versus non-coldspots, or at hotspots versus coldspots, the binomial ANCOVAs produced results that are generally consistent (Table 4). One feature of the results from these analyses, also apparent when considering marginal effects ($p < 0.10$), is that there is evidence of phylogenetic clustering of hotspots and coldspots in half of the datasets. In other words, host species from the same family are more likely to be hotspots (or coldspots) of parasite diversity than species from different families, independently of their body sizes. There are exceptions, however, that become apparent when contrasting the lists of hotspots and coldspots from the same dataset (Supplementary material Appendix 1). For instance, two species of the bird genus *Larus* are hotspots of helminth parasite species richness, whereas one species from that genus is a coldspot. A similar scenario is seen in the tropical freshwater fish genera *Cichlasoma* and *Astyanax*, among others.

Another key outcome of the analyses is the clear demonstration that host ecological traits have no influence on whether or not a host species is a hotspot or coldspot of parasite diversity. In all datasets, the host ecological trait (i.e. water temperature, mean depth range, latitude or population density, depending on the dataset) chosen because of its expected effect on parasite species richness had no significant effect (at the $\alpha = 0.05$ level) on the likelihood of a host species being a hotspot or a coldspot (Table 4). In some cases, there was a significant interaction between host family and the ecological variable, indicating that ecological influences are restricted to some host families. For example, among shark species, the likelihood that a host species is a hotspot of cestode diversity is influenced by mean depth range in different ways depending on what family the hosts belong to (Table 4).

Discussion

The search for the underlying cause of variation in parasite species richness among host species has a long history (Poulin

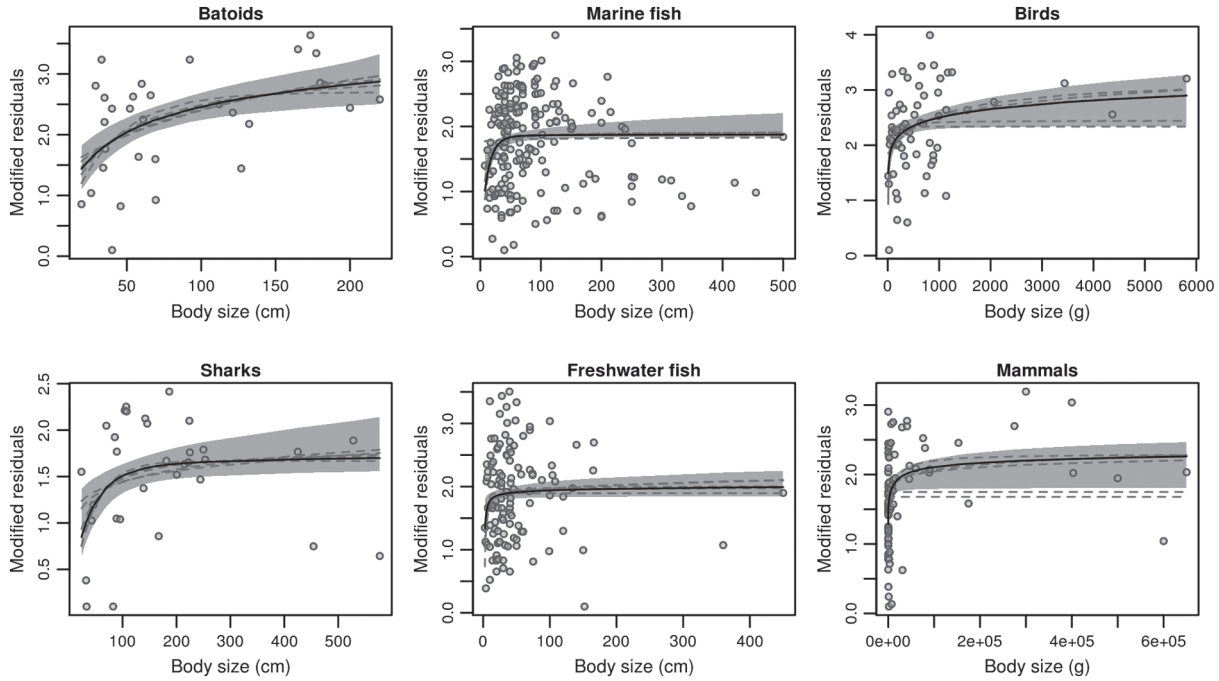

Figure 1. Fitting of five SAR models to the interspecific relationship between parasite species richness and host body size, in six datasets. Parasite species richness is shown as the modified ('positivised') residual of log richness regressed against log sampling effort. Broken lines represent each of the five models, the solid line represents the averaged multimodel SAR based on weights, and the shaded area indicates the 95% confidence interval obtained by a nonparametric bootstrapping procedure.

Table 4. Results of logistic ANCOVAs testing the effects of host taxonomy and ecology on the likelihood of a host species being a hotspot or coldspot of parasite species diversity in each of the six parasite–host datasets. The significance of the deviance explained by each factor is shown.

		Hotspots vs non-hotspots		Coldspots vs non-Coldspots		Hotspots vs coldspots	
Dataset (no. host species)	Factor	DF	p	DF	p	DF	p
Metazoans/marine fish (182)	water temperature	1	0.923	1	0.592	1	0.802
	host family	56	0.217	56	0.121	55	0.179
	interaction	26	0.229	26	0.153	26	0.230
Metazoans/freshwater fish (118)	water temperature	1	0.182	1	0.090*	1	0.108
	host family	29	0.059*	29	0.010**	28	0.026**
	interaction	14	0.999	14	1	14	0.065*
Cestodes/sharks (31)	midpoint of depth range	1	0.878	1	0.867	1	0.844
	host family	14	0.232	14	0.153	9	0.115
	interaction	4	0.009**	4	0.251	2	0.011**
Cestodes/batoids (26)	midpoint of depth range	1	0.276	1	0.912	1	0.648
	host family	6	0.186	6	0.347	5	0.114
	interaction	4	0.508	4	0.276	2	0.496
Helminths/birds (64)	latitude	1	0.364	1	0.928	1	0.639
	host family	26	0.088*	26	0.041**	22	0.015**
	interaction	10	0.549	10	0.074*	8	0.346
Helminths/mammals (49)	population density	1	0.328	1	0.173	1	0.175
	host family	17	0.481	17	0.022**	13	0.075*
	interaction	7	0.999	7	0.599	3	0.830

*, $p < 0.10$; **, $p < 0.05$.

and Morand 2004). This research has been driven both by the possibility of using host–parasite systems as models for studies of biodiversity, and by the more pressing need to understand the factors driving the risk of disease emergence in wildlife. The importance of host body size has persisted as a basis for much of the research to date, despite the fact that apparent parallels with island biogeography are often flawed (Kuris et al. 1980) and that the predictive power of host body size has generally been very weak in past analyses (Poulin 1997, 2004, Poulin and Morand 2004). The size of the host's body is seen as an overall measure of habitat dimension, niche diversity and food supply; in epidemiological terms, it should also correlate with encounter rates with parasite infective stages either via ingestion or contact. Here, we show that host body size is not a good universal predictor of parasite species richness and that exceptional parasite diversity shows a stronger association with host phylogeny than with host ecology.

The scatter of points in the plots of parasite species richness against host body size suggests a positive association between the two variables (Fig. 1), though the relationship is clearly more complicated than the linear function implicitly assumed in many earlier studies. Applying a wide range of models previously used for species–area relationships (SARs), we showed that none of these convex models, whether asymptotic or not, consistently provided a good fit to the data. In all cases, the proportion of variance in parasite species richness explained by host body size was low. The multimodel average SAR gave a fit in which host body size explained between 14 and 23% of the variance in parasite richness for cestodes parasitic in batoids or sharks, and helminths parasitic in birds and mammals. However, for metazoan parasites parasitic in either marine or freshwater teleost fish, these values were 3% or lower; visual inspection of the data suggests either no relationship at all or a peak in richness for small-to-mid-sized fish species. The two data sets on metazoan parasites of fish combined a greater taxonomic range of parasites than the other four datasets, i.e. they included both ectoparasitic crustaceans, leeches and monogeneans as well as internal parasites. Earlier studies of interspecific variation in parasite species richness among teleost fish have failed to agree on the influence of body size, whether focusing on ectoparasites (contrast Poulin 1995 with Rohde et al. 1995) or endoparasites (contrast Gregory et al. 1996 with Sasal et al. 1997). Combining different types of parasites in the same dataset may have obscured the role of host body size by pooling parasite taxa whose transmission and persistence do not depend on host body size to the same extent. Visually (Fig. 1), the clearest trends were seen for cestodes in elasmobranchs, where a single parasite taxonomic group was considered. Future compilations of parasite species richness should therefore avoid pooling different types of parasites. In addition, the measures of total body size used are perhaps not always the most relevant to the parasites; for example, body surface area matters for ectoparasites, whereas the volume of the gut might be more relevant for endoparasites. Nevertheless, our attempts to fit SARs to several different datasets tend to suggest that host body size is not a strong or consistent determinant of parasite species richness across taxa. A glance at the scatterplots in Fig. 1 suffices to show that host body size is, in most cases, a poor predictor of which host species harbour unusually rich parasite communities.

Our method to identify 'hot hosts' and 'cold hosts' of parasite diversity follows recent advances and overcomes earlier concerns regarding hotspot detection (Veech 2000, Brummitt and Lughadha 2003, Hobohm 2003, Ovadia 2003, Ulrich and Buszko 2005, Fattorini 2006, 2007). Keeping in mind the relatively poor fit of the multimodel average SAR, several host species lie well outside the 80% confidence intervals, either above (hotspots) or below (coldspots). This

is not an artefact of uneven sampling effort, which has been accounted for in our analyses. Our results indicate that ecological characteristics of host species identified in previous studies as likely determinant of parasite species richness could not reliably predict which host species were parasite hotspots or coldspots. For instance, host population density, identified previously as a correlate of helminth richness in mammals (Morand and Poulin 1998, Poulin and Mouillot 2004), failed to predict which mammal species harbour unusually high or unusually low numbers of parasite species. When host ecological features did play a small role, it was dependent on the host family involved, as indicated by a few significant family-by-ecology interactions. Thus, in sharks, the depth range at which a species lives, which emerged previously as a global correlate of cestode richness (Randhawa and Poulin 2010), may predict which species are cestode hotspots, but this depends entirely on the shark family involved.

Indeed, the taxonomic or phylogenetic affiliations of host species came out of our analyses as the most general predictors of which species are parasite hotspots. Most previous studies have used comparative analytical methods to 'neutralise' the effect of host phylogeny when investigating the role of ecology as a cause of the observed variation in parasite diversity among host species (Poulin 1995, Gregory et al. 1996, Sasal et al. 1997, Nunn et al. 2003, Bordes et al. 2009). However, the phylogenetic position of a species captures much of its ecological (habitat, diet, etc.) and immunological characteristics as well as its past history (biogeographic area of origin, etc.). The combined information conveyed by a species' phylogenetic position possibly makes it a much better predictor of how many parasite species have been acquired over evolutionary time by a particular host lineage than any ecological variable on its own. Close inspection of the species included in the list of 'hot hosts' reveals no obvious shared ecological attribute beyond those already investigated (Supplementary material Appendix 1). However, in all six datasets, there are families that are over-represented among hotspots and families belonging to the same order that are over-represented among coldspots (Supplementary material Appendix 1). It is therefore challenging to pinpoint what exactly makes certain host lineages accumulate an inordinate number of parasite species.

Furthermore, all species within a given host family do not necessarily cluster well above (or below) the multimodel SAR confidence intervals, however. Sometimes congeneric host species show very different parasite species richness, one being a richness hotspot and the other a coldspot, even after our exhaustive corrections for sampling effort and host body size. Changes in the composition and extent of an animal's parasite fauna can occur rapidly, for instance following range expansion (Poulin and Morand 2004), possibly accounting for these exceptions. This may also explain why, in our analyses, host family was not always a significant predictor of whether or not a host species is a parasite hotspot, and why, when it was significant, it was not a particularly strong predictor. Overall, though, our findings suggest that host phylogeny may often outweigh specific host ecological traits as a predictor of whether or not a particular host species harbours more (or fewer) parasite species than expected for its size. In particular, given the poor predictive performance of host body size, across different datasets and using a range of SAR functions, it may be time to abandon it as a general explanation for interspecific variation in parasite diversity.

Acknowledgements –Partial support for this work came via a grant to RP from the Marsden Fund (New Zealand), and a research fellowship to JLL from CNPq (Conselho Nacional de Pesquisa e Desenvolvimento Tecnológico, Brazil).

References

Bordes, F. et al. 2009. Home range and parasite diversity in mammals. – Am. Nat. 173: 467–474.

Brummitt, N. and Lughadha, E. 2003. Biodiversity: where's hot and where's not. – Conserv. Biol. 17: 1442–1448.

Burnham, K. P. and Anderson, D. R. 2002. Model selection and multimodel inference: a practical information-theoretic approach. – Springer.

Compagno, L. J. V. et al. 2005. Sharks of the world. – Princeton Univ. Press.

Damuth, J. 1987. Interspecific allometry of population density in mammals and other animals: the independence of body mass and population energy-use. – Biol. J. Linn. Soc. 31: 193–246.

Dennis, J. E. and Schnabel, R. B. 1983. Numerical methods for unconstrained optimization and nonlinear equations. Classics in applied mathematics 16. – Prentice-Hall.

Dunning, J. B. Jr. 1993. CRC Handbook of avian body masses. – CRC Press.

Fattorini, S. 2006. Detecting biodiversity hotspots by species–area relationships: a case study of Mediterranean beetles. – Conserv. Biol. 20: 1169–1180.

Fattorini, S. 2007. To fit or not to fit? A poorly fitting procedure produces inconsistent results when the species–area relationship is used to locate hotspots. – Biodiv. Conserv. 16: 2531–2538.

Froese, R. and Pauly, D. 2006. FishBase. – <www.fishbase.org>

Gregory, R.D. et al. 1996. Helminth parasite richness among vertebrates. – Biodiv. Conserv. 5: 985–997.

Guilhaumon, F. et al. 2008. Taxonomic and regional uncertainty in species–area relationships and the identification of richness hotspots. – Proc. Natl Acad. Sci. USA 105: 15458–15463.

Guilhaumon, F. et al. 2010. mmSAR: an R-package for multimodel species–area relationship inference. — Ecography 33: 420–424.

He, F. and Legendre, P. 1996. On species-area relations. – Am. Nat. 148: 719–737.

Hobohm, C. 2003. Characterization and ranking of biodiversity hotspots: centres of species richness and endemism. – Biodiv. Conserv. 12: 279–287.

Johnson, J. B. and Omland, K. S. 2004. Model selection in ecology and evolution. – Trends Ecol. Evol. 19: 101–108.

Kuris, A. M. et al. 1980. Hosts as islands. – Am. Nat. 116: 570–586.

Kvalseth, T. O. 1985. Cautionary note about r^2. – Am. Stat. 39: 279–285.

Lindenfors, P. et al. 2007. Parasite species richness in carnivores: effects of host body mass, latitude, geographical range and population density. – Global Ecol. Biogeogr. 16: 496–509.

Lomolino, M. V. 2000. Ecology's most general, yet protean pattern: the species–area relationship. – J. Biogeogr. 27: 17–26.

Luque, J. L. and Poulin, R. 2007. Metazoan parasite species richness in Neotropical fishes: hotspots and the geography of biodiversity. – Parasitology 134: 865–878.

Luque, J. L. and Poulin, R. 2008. Linking ecology with parasite diversity in Neotropical fishes. – J. Fish Biol. 72: 189–204.

Morand, S. and Poulin, R. 1998. Density, body mass and parasite species richness of terrestrial mammals. – Evol. Ecol. 12: 717–727.

Nunn, C. L. et al. 2003. Comparative tests of parasite species richness in primates. – Am. Nat. 162: 597–614.
Ovadia, O. 2003. Ranking hotspots of varying sizes : a lesson from the nonlinearity of the species-area relationship. – Conserv. Biol. 17: 1440–1441.
Poulin, R. 1995. Phylogeny, ecology, and the richness of parasite communities in vertebrates. – Ecol. Monogr. 65: 283–302.
Poulin, R. 1997. Species richness of parasite assemblages: evolution and patterns. – Annu. Rev. Ecol. Syst. 28: 341–358.
Poulin, R. 2004. Macroecological patterns of species richness in parasite assemblages. – Basic Appl. Ecol. 5: 423–434.
Poulin, R. and Morand, S. 2004. Parasite biodiversity. – Smithsonian Inst. Press.
Poulin, R. and Mouillot, D. 2004. The evolution of taxonomic diversity in helminth assemblages of mammalian hosts. – Evol. Ecol. 18: 231–247.
Randhawa, H. S. and Poulin, R. 2010. Determinants of tapeworm species richness in elasmobranch fishes: untangling environmental and phylogenetic influences. – Ecography doi: 10.1111/j.1600-0587.2010.06169.x.

Supplementary material (available online as Appendix o19036 at <www.oikosoffice.lu.se/appendix>.) Appendix 1.

Rohde, K. et al. 1995. Aspects of the ecology of metazoan ectoparasites of marine fishes. – Int. J. Parasitol. 25: 945–970.
Rosenzweig, M. L. 1995. Species diversity in space and time. – Cambridge Univ. Press.
Rosenzweig, M. L. and Sandlin, E. A. 1997. Species diversity and latitudes: listening to area's signal. – Oikos 80: 172–176.
Sasal, P. et al. 1997. Determinants of parasite species richness in Mediterranean marine fish. – Mar. Ecol. Progr. Ser. 149: 61–71.
Ulrich, W. and Buszko, J. 2005. Detecting biodiversity hotspots using species–area and endemics–area relationships: the case of butterflies. – Biodiv. Conserv. 14: 1977–1988.
Veech, J. A. 2000. Choice of species–area function affects identification of hotspots. – Conserv. Biol. 14: 140–147.
Vickery, W. L. and Poulin, R. 1998. Parasite extinction and colonisation and the evolution of parasite communities: a simulation study. – Int. J. Parasitol. 28: 727–737.
Walther, B. A. et al. 1995. Sampling effort and parasite species richness. – Parasitol. Today 11: 306–310.
Williams, M. R. et al. 2009. Species–area functions revisited. – J. Biogeog. 36: 1994–2004.