Host specificity in phylogenetic and geographic space

Robert Poulin1, Boris R. Krasnov2 and David Mouillot3,4

1 Department of Zoology, University of Otago, PO Box 56, Dunedin, New Zealand
2 Mitrani Department of Desert Ecology, Institute for Dryland Environmental Research, Jacob Blaustein Institutes for Desert Research, Ben-Gurion University of the Negev, Sede Boqer Campus, 84990 Midreshet Ben-Gurion, Israel
3 ECOSYM, Unité Mixte de Recherche 5119 Centre National de la Recherche Scientifique (CNRS)–UM2–Institut Français de Recherche Pour l’Exploitation de la Mer (IFREMER)–Institut de Recherche pour le Développement (IRD), Université Montpellier 2, Place Eugène Bataillon, 34095 Montpellier Cedex 05, France
4 Australian Research Council Centre of Excellence for Coral Reef Studies, James Cook University, Townsville, QLD 4811, Australia

The measurement of host specificity goes well beyond counting how many host species can successfully be used by a parasite. In particular, specificity can be assessed with respect to how closely related the host species are, or whether a parasite exploits the same or different hosts across its entire geographic range. Recent developments in the measurement of biodiversity offer a new set of analytical tools that can be used to quantify the many aspects of host specificity. We describe here the multifaceted nature of host specificity, summarize the indices available to measure its different facets one at a time or in combination, and discuss their implications for parasite evolution and disease epidemiology.

Host specificity meets biodiversity

Host specificity, in other words the extent to which parasites can exploit different host species, is one of the most fundamental properties of parasitic organisms. For instance, it is a key determinant of how likely parasites are to persist following host extinction [1,2], or to become established following their introduction into new areas [3,4]. Recent genetic and experimental investigations have revealed that for many parasite species, host specificity is often either under- or overestimated when based on field surveys alone [5]. Even when accurate, however, the traditional measurement of host specificity as the mere number of host species that can be used by a parasite greatly limits our ability to understand its evolution and its ecological implications. For instance, can one truly appreciate how specialized a parasite is without considering the relationships between its few host species, or without addressing why it uses different hosts in different geographic areas? Similar questions have been raised regarding the measurement of biodiversity, for which mere species counts are now often seen to be insufficient [6]. Indeed, the concept of biodiversity encompasses many facets, including phylogenetic and functional differences between species and their turnover across space [7]. In this review we develop a hierarchical and multifaceted view of host specificity that parallels the rigorous approach now used for the measurement of biodiversity. First, we dissect host specificity into its various components, each manifested at different scales, to highlight the many nuances of specificity. Second, the indices and analytical approaches that recently became available to measure various facets of host specificity and investigate their associations are summarized. Finally, the evolutionary pressures shaping different aspects of host specificity and their ecological implications are discussed.

The many facets of host specificity

Two parasite species exploiting the same number of host species could appear to have identical host specificity, whereas in fact their plasticity in host use can differ at numerous levels. Similarities among different parasites in the number of host species exploited, or basic host specificity in its traditional sense and as used in many earlier studies, can mask fundamental differences either in the relative ecological importance of each host for the parasite population, in their phylogenetic closeness to each other, or in variation in host use on geographical scales. In the discussion of these various facets of host specificity that follows, we employ the term ‘host’ to refer only to those species in which a parasite can successfully survive, grow and/or reproduce; we ignore ‘accidental’ hosts in which parasites cannot develop, because these do not really belong to the repertoire of the parasite.

Structural specificity

Consider two related parasite species, A and B, each using the same number of host species – for example, four host species. However, their populations are structured differently across their host species; parasite A achieves roughly similar prevalence and abundance in all four of its hosts, whereas B is highly abundant in one host species but rare in the other three. Therefore, parasite B displays higher structural specificity than A because the bulk of its population utilizes a single host species (Figure 1). The uneven distribution of a parasite population between two or more host species is commonly observed in both terrestrial and aquatic ecosystems. For example, the mean abundance of several species of helminths and cephepods differs substantially, often by a factor of 10-fold or more, between different

Corresponding author: Poulin, R (robert.poulin@otago.ac.nz).

1471-4922/$ - see front matter © 2011 Elsevier Ltd. All rights reserved. doi:10.1016/j.pt.2011.05.003 Trends in Parasitology, August 2011, Vol. 27, No. 8 355
sympatric flatfish species inhabiting coastal Portuguese waters [8]. Similarly, the mean abundance of the ectoparasitic mite, *Spinturnix myoti*, is about 35% higher on one of its two sibling bat host species, although the prevalence of infection is identical in both hosts [9]. Indices of host specificity have been developed specifically to account for these differences [10], although the widely used indices of species diversity that incorporate both the number of species and the evenness of their abundances, such as the Simpson or Shannon indices [6], would be excellent indices of structural host specificity when applied to data on relative host use.

**Phylogenetic specificity**
Consider now another two related parasite species, A and B, each using four host species and achieving roughly equal abundance in all their hosts. However, the hosts of parasite A belong to distantly related families, whereas those of parasite B all belong to the same genus. Therefore, we can easily argue that parasite B displays higher phylogenetic specificity than A because its hosts are restricted to a narrower phylogenetic spectrum (Figure 2). There are numerous examples of parasites that are either restricted to a few closely related host species, or are capable of exploiting completely unrelated host taxa. For example, the nematode *Trichostrongylus axei* occurs predominantly in ungulates but can also infect mammals from other orders [11], whereas acanthocephalans parasitic in fishes can typically exploit hosts from completely different families [12]. Existing indices of phylogenetic biodiversity [13–16] could easily be applied in such situations to distinguish between parasites with different levels of specificity, and indeed they have been used for that purpose [17]. Alternatively, an index of phylogenetic host specificity has also been proposed [18]. Whether designed for measuring biodiversity or host specificity, essentially all these indices compute either the average or total phylogenetic (or taxonomic in the absence of explicit branch length data) distance between all possible pairs of host species used by a parasite, to estimate their phylogenetic distinctness.

As an extra step, it is relatively easy to combine data on phylogenetic relatedness and data on relative abundances, that is, to combine structural specificity and phylogenetic specificity into a single index of structural phylogenetic specificity. Again, this can be achieved either by applying methods designed for the study of biodiversity [19–21] or by using the related indices developed specifically for the measurement of host specificity [17,22]. Here again the average or total phylogenetic distance is calculated among all host species, but proportionally greater weight is given to host species in which the parasite achieves greater abundance. Thus, if a parasite achieves high abundance in two congeneric hosts but low abundance in two other hosts from another family, its combined phylo-structural specificity will be higher than if, all else being equal, it had reached high abundance in one host from each family.

**Specificity in geographic space**
Finally, it is worth considering host specificity on a geographic scale. Across its geographic range, any parasite species will be exposed to communities of potential host species that change in composition and relative density from one locality to the next [23]. Specificity can thus also be manifested as the consistency in host use across a changing host landscape. For example, imagine that in one area parasite species A and B each use four host species, of similar phylogenetic relatedness, and achieve roughly equal abundance in all their hosts. However, in a different locality, parasite B uses the same four host species whereas A uses four completely different host species. Therefore, although both parasites have similar structural and phylogenetic host specificity, parasite B displays higher host specificity than A across its geographical range, because its use of hosts from the available species pool is clearly more limited both locally and globally. Again, this hypothetical scenario is often seen in nature. For example, consider the two fleas *Mesopsylla hebes* and *Oropsylla itoaiiskii* which, on average, both
exploit only two host species in each locality where they occur. However, across their geographic ranges, the vast majority of hosts harboring the former flea species belong to a single rodent family, Dipodidae, whereas the latter flea is found on rodents belonging to three different families [24].

Research on biodiversity recognizes this spatial dimension. Ecologists now distinguish between local species diversity, or $\alpha$-diversity, and the difference in species composition among localities, or $\beta$-diversity, which captures the turnover of species in geographical space [6,25–27]. Similarly, we can distinguish between $\alpha$-specificity, or the host specificity shown by a parasite in one locality, and $\beta$-specificity, or the spatial turnover in host species used across localities [24]. Thus, particular parasite species can show high specificity at a local scale while being generalists on a global scale, or vice versa. In addition, once again in parallel to the measurement of biodiversity in space, one can compute an overall measure of host specificity, or $\gamma$-specificity, across the entire geographical range of a parasite, based on the total list of host species used across all localities where the parasite occurs.

Just as with biodiversity measures [28,29], the basic, structural and phylogenetic components of host specificity can be combined into multidimensional measures, and treated at $\alpha$, $\beta$ and/or $\gamma$ geographical scales (Figure 3). Depending on the completeness of the data available, host specificity can be investigated in one or more dimensions at a time. For example, data on the occurrence of parasite species in various host species could be limited to presences and absences, but they could also include actual abundances. If abundance data are available for numerous localities, it is the relative abundance of a parasite in its different sympatric hosts, and not the absolute abundance, that matters in geographical comparisons – because absolute abundance could vary due to the effects of environmental conditions on infective stages. Indices and measures of host specificity are available for all levels of complexity, whether focused on a single component of specificity or on all of them, and at any spatial scale (Box 1). There are also programs available for the computation of all these measures (Box 2).

There are other facets of host specificity, of course, beyond those discussed above. For instance, in helminths with complex life cycles, different sets of host species are used at different stages of the life cycle, adding an ontogenetic component to host specificity. However, because most interspecific analyses seeking to understand the ecology and evolution of host specificity compare the specificity of different parasite species at a given developmental stage, this ontogenetic component will not be discussed further.

Evolutionary and ecological implications of host specificity

Deconstructing host specificity into its structural, phylogenetic and spatial components opens up several new research possibilities [30]. It will not only allow us to generate more precise and targeted hypotheses about host–parasite ecology and evolution, but also provide the tool-kit to test them. In particular, the measurement of separate facets of host specificity will increase the value of comparative analyses, which have long been powerful approaches in evolutionary biology [31]. These are founded on relationships between variables measured in different localities or across species, to evaluate the contribution of various factors to intra- or interspecific variation, respectively. It will now be possible to precisely match a putative selective factor with the specific facet of host specificity on which it is presumed to act.

The first step, however, will be to determine how different aspects of host specificity are correlated with each other across different parasite species. The subset of species used as hosts within a local host community is not a random subset of those available: typically, parasites infect host species more closely related to each other than if host selection was completely random [32]. However, as more host species are added to the repertoire of a parasite the breadth of their phylogenetic spectrum invariably increases, leading to weak but significant interspecific correlations between basic and phylogenetic measures of host specificity [18,33]. Also, two closely related host species are more likely to harbor similar abundances of a given parasite species than two unrelated hosts, and this is true whether the parasites are helminths [34] or arthropods [35]. Therefore, we might expect a negative relationship between measures of structural and phylogenetic host specificity (Figure 4), if a parasite population is apportioned more unevenly among host species when the latter belong to different genera or families. Similarly, there are correlations between the different measures of biodiversity used in ecology and conservation [7,36]. It will be essential to assess the statistical independence of different facets of host specificity using comparative analyses before their full potential can be realized.

It is important to determine which selective forces shape the different facets of host specificity, and which are most

\[
\text{Figure 3. Host specificity in geographic space. For each of four parasite species, host use in two separate populations (adjacent shaded boxes) from distinct geographical sites is shown. For any level of phylogenetic specificity, which depends on host relatedness, a parasite can show low geographic specificity (i.e. $\beta$-specificity) if the identity of the host species (colored symbols) it uses varies among localities, or high $\beta$-specificity if it exploits the same host species everywhere. In these examples, both basic and structural specificity, based on the number of host species used and the distribution of parasite individuals among host species, respectively, are assumed to be the same in all cases.}
\]
### Box 1. Host specificity indices in phylogenetic and geographic space

Host specificity is inversely proportional to ‘generalism’ in host use, and it is the extent to which parasites are generalists that is measured by most indices. Hence, for the indices we recommend below, low index values correspond to high host specificity, and vice versa.

1. **Basic and structural host specificity**

   Traditionally, host specificity is simply estimated as the number of host species (S) used by a parasite species, but this value can also be corrected for biases arising from the undersampling of rare hosts using the Chao indices [62]. When data on prevalence or abundance can be incorporated into the measurement of structural specificity, we suggest using composite indices such as the Shannon index, or else a ‘pure’ evenness index, independent from the number of host species, such as the Bulla index [63].

   Because $PS_i$ is not totally independent from the number of host species used by a parasite and thus provides information redundant with $S$, two options are possible.

   (i) Estimate the standardized effect size of $PS_i$ or $SPS_i$, using random subsets of potential host species drawn from the regional pool to determine whether the hosts actually used by the parasite are more or less closely related than expected by chance, and thus whether the phylospecificity of parasite $i$ is high or low for a given value of $S$ using:

   \[
   SPS_i = \frac{(PS_i - PS_{i\text{mean}})}{SD(PS_{i\text{mean}})}
   \]  

   where $PS_i$ is the observed phylospecificity of parasite $i$, $PS_{i\text{mean}}$ is the mean phylospecificity of random host subsets, and $SD(PS_{i\text{mean}})$ is the standard deviation of all randomized phylospecificity values.

   (ii) Estimate phylospecificity as the average phylogenetic distinctiveness, $SPD_i$, between all pairs of host species [18], which is independent from how many host species are used by a parasite:

   \[
   SPD_i = \frac{\sum \sum_{j < k} \omega_{jk}}{S(S - 1)}
   \]

   where $\omega_{jk}$ is the phylogenetic distance between host species $j$ and $k$ used by parasite $i$, or, when the phylogeny is not fully resolved, the number of taxonomic steps required to reach a node common to both.

   The double summation is over the set $\{k = 1, \ldots, S_i; j = 1, \ldots, S\}$, such that $j < k$ in order to consider all host species pairs.

3. **Geographic host specificity, or $\beta$-specificity**

   Measuring the turnover of host species used by a parasite among different localities, in other words $\beta$-specificity or $B_S$, involves estimating the dissimilarity in host species identities between localities. Most dissimilarity or $\beta$-specificity indices have been designed for two samples only [65]. Ideally, estimates of $\beta$-specificity across space should include several samples (i.e. data from different localities [24]). We suggest using the extension of the Sørensen dissimilarity index for multiple-sites [66] to measure $\beta$-specificity:

   \[
   BS = 1 - \frac{T}{T-1} \left(1 - \sum_{i,j} S_i^j \right)
   \]

   where $T$ is the number of samples or localities, $S_i^j$ is the number of host species used in locality $i$ and $S_j^i$ the total number of host species used by parasite species across all localities (i.e. the regional host pool). If parasite $i$ exploits the same host species across all localities, then $S_i^j = S_j^i$ and $BS = 0$. If parasite $i$ uses totally different host species from one locality to the next, then $S_i^j = S_j^i$ and $BS = 1$.

4. **Combining phylogenetic and geographic specificity: phylobatspecficity**

   Information about the phylogenetic relatedness of host species and their different use across localities can be combined into a single index of phylogenetic $\beta$-specificity, or $PBS$. This corresponds to the phylogenetic turnover of host species used by parasite $i$ over geographic space. For this, we can use an extension of the Sørensen index [66] to branches instead of species following the principle underlying the construction of the Phylosor index [67]:

   \[
   PBS = 1 - \frac{T}{T-1} \left(1 - \frac{T}{\sum_t PD_t} \right)
   \]

   where $T$ is the number of samples or localities, $PD_i$ is the phylogenetic diversity of host species used by the parasite in locality $i$, and $PD_t$ the phylogenetic diversity of all host species used by parasite species $i$ across all $T$ localities. If parasite $i$ exploits the same host species across all localities, we obtain $PD_i = PD_t$ and $PBS = 0$. If parasite $i$ uses different host species from one locality to the next, then the less phylogenetically related those hosts are, the higher the PBS value.

   We opted to present a common statistical framework using the Sørensen index because it can cope with multiple localities [66] and can incorporate phylogenetic diversity [67]. If prevalences are known for each host species used by a parasite, an alternative framework could be developed to estimate specificity over geographic and phylogenetic space while also incorporating structural host specificity (see main text), based for instance on the Rao index [7,68,69] or the Shannon entropy index [60]; however, this is beyond the scope of the present summary.

---

important at each spatial scale. For instance, we might expect that unstable host communities, characterized by population densities that fluctuate substantially over time because of seasonal migration or reproduction, should favor parasites with low phylogenetic specificity but no particular pattern of structural specificity, because parasite populations would need to use hosts in proportion to their current densities in order to persist. Current work on such unstable host–parasite systems [37] has yet to consider host specificity beyond the simple number of host species used, and thus such predictions remain untested. In addition, we might expect that selected host or parasite properties, in particular dispersal abilities, could drive patterns of geographical specificity. Host mobility has already been shown to homogenize parasite community composition across spatially distant localities [38,39], but its impact upon various facets of host specificity, especially $\alpha$- and $\beta$-specificity, remains unknown.

Conversely, we can ask what phenomena are likely to result from certain combinations of structural, phylogenetic and geographic host specificity. Both the evolution of parasite virulence [40,41] and the strength of local adaptation [42,43] are known from theoretical and empirical work to relate to host specificity. However, previous studies have considered host specificity in its most basic sense, in other words the number of host species used, and the importance of different components of host specificity for these evolutionary processes is likely to vary. One can imagine that local adaptation of parasites to their hosts, whatever the number of host species, would be facilitated
by high phylogenetic specificity, and/or by high structural specificity, because these would involve a single host species driving parasite population dynamics. Along those lines, some components of host specificity appear to be more relevant than others to the assumed trade-off between how many hosts a parasite exploits and its performance in any of them or its average performance across all of them, captured by the phrase ‘jack of all trades but master of none’ [44]. Indeed, although physiological incompatibilities between host and parasite can greatly affect this trade-off [45], at the macroecological scale if the trade-off exists, it can be more easily revealed using a measure of phylogenetic specificity than the simpler basic measure [46]. Thus, the newly available measures of all facets of host specificity now allow more sophisticated tests of old evolutionary hypotheses.

Several epidemiological processes are also closely linked with one or more facets of host specificity. However, which components of host specificity are the most important for parasite population dynamics, and which are more important as a hedge against extinction? Structural specificity, in other words data on relative abundance in different host species, is used to estimate the probable number of host plants exploited by phytophagous insects and identify those species most at risk of extinction [47]; the same analysis could be carried out for animal parasites. In the context of managing diseases, which components of host specificity are most relevant for the risk of parasite establishment and emerging infectious disease following biological invasions or climate change? Parasites can re-adjust their host range after changes in geographical distribution or local ecology, so as to align themselves with the new set of suitable hosts they encounter [48]. This process of ecological fitting provides a framework for host-switching and the establishment of new host–parasite associations in the face of environmental change [49,50]. However, the deconstruction of host specificity into quantifiable facets gives it more predictive power. It would be possible, for example, to measure precisely the current phylogenetic specificity of a parasite to determine the likelihood that it would spread to other hosts if introduced to a new area, by measuring the phylogenetic distances between potential hosts and the current host suite. Similarly, measuring structural specificity before local changes in relative host abundances would allow more robust predictions of the likelihood of local parasite persistence.

Concluding remarks

As discussed above, taking a multifaceted look at host specificity creates exciting new opportunities for ecological and evolutionary parasitology. At a time when host specificity is recognized as a key factor determining the spread of parasitic diseases in the face of biological invasions, loss of natural habitats and climate change [51], its dissection into distinct and measurable facets paves the way toward a better understanding of the processes that control which hosts a parasite can use. Experimental studies of model species will always remain important to determine, at a proximate level, how given environmental features affect host use by generalist parasites [52,53]. The analytical tools presented here will empower large-scale studies addressing the causes of variation in host specificity among parasite species. At the moment, however, interspecific patterns revealing how different aspects of host specificity covary with environmental factors or other parasite properties have only been investigated for very few parasite taxa, predominantly avian malaria [17,54] and fleas parasitic for mammals [44,55,56]. Large comparative datasets on other parasite groups are urgently needed to explore the several levels on which pathogens can either specialize or turn into generalists.

Another important future direction will be the further expansion of our outlook on host specificity. In recent years, ecologists and conservation biologists have started to consider biodiversity not only from a numerical or phylo-
nomic perspective, but also from a functional one [57,58]. Several indices have been developed to measure the functional diversity of species assemblages, that is, the diversity of ecological or phenotypic traits that characterize them [59,60]. To borrow once more from biodiversity research, we could envisage the measurement of functional host specificity, by quantifying the similarities among the hosts of a parasite in terms of the functional traits they share rather than their phylogenetic relationships. Furthermore, most indices that can be used to measure phylogenetic specificity can be adapted to measure functional specificity [29] because they are based on distances between species pairs and trees built from traits [61]. For studies of generalist parasites in taxon-rich host communities, this might allow more explicit links between host ecology and host-switching by parasites. The take-home message here is that specialization by parasites can occur at several levels, and that whatever the question we pose, we must make sure to use the proper analytical tools to answer it.

Acknowledgments
We thank David Thiéltges, Charles Criscione and two anonymous reviewers for useful comments on an earlier draft. D.M. was supported by a Marie Curie International Outgoing Fellowship (FISHECO; agreement number IOP-GA-2009-236318). This publication is no. 730 of the Mitrani Department of Desert Ecology.

References
34 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
53 Koebler, A.V. et al. (2011) Genetic and environmental determinants of host use in the trematode Maritrema novaezelandensis (Microphallidae). Parasitology 138, 100–106