# Cronfa - Swansea University Open Access Repository

_____

This is an author produced version of a paper published in:
*Trends in Parasitology*

_____

Cronfa URL for this paper:
http://cronfa.swan.ac.uk/Record/cronfa49872

_____

**Paper:**

Wells, K. & Clark, N. (2019). Host Specificity in Variable Environments. *Trends in Parasitology*
http://dx.doi.org/10.1016/j.pt.2019.04.001

http://www.swansea.ac.uk/library/researchsupport/ris-support/

# Host specificity in variable environments

Konstans Wells[1*], Nicholas J. Clark[2]

*1 Department of Biosciences, Swansea University, Swansea SA2 8PP, Wales, U.K.*

*2 School of Veterinary Science, University of Queensland, Gatton QLD 4343, Australia*

\* Address for correspondence (E-mail: k.l.wells@swansea.ac.uk, phone: +44 (0)1792 298633).

## Highlights

• Increasing interests in emerging infectious diseases and parasite spillover coincide with a rise of studies reporting and comparing host specificity for multihost parasite. Intuitively, higher host specificity means less spillover risk but to date, a systematic consensus on such a relationship is lacking.

• Host specificity can vary in space and time due to changing compositions of potential host species and constraints in environmental conditions. Eco-evolutionary dynamics and contemporary conditions synergistically determine 'realized' host specificity across regional scales.

• Modelling advances to capture spatiotemporal variation in the distributions and biotic interactions of species provide the basis to quantify variation in realized host specificity and progress towards determining how this relates to spillover risk.

## Abstract

Host specificity encompasses the range and diversity of host species that a parasite is capable of infecting and is considered a crucial measure of a parasite's potential to shift hosts and trigger disease emergence. Yet empirical studies rarely consider that regional observations only reflect a parasite's 'realized' host range under particular conditions: the true 'fundamental' range of host specificity is typically not approached. We provide an overview of challenges and directions in modelling host specificity under variable environmental conditions. Combining tractable modelling frameworks with multiple data sources that account for the strong interplay between a parasites' evolutionary history, transmission mode and environmental filters that shape host-parasite interactions will improve efforts to quantify emerging disease risk in times of global change.

**Key words:** disease spread; pathogen spillover; co-speciation; host shifting; host-parasite interactions; realized host specificity

## Host specificity in times of emerging infectious diseases

The diversity of infectious disease affecting humans and animals are strongly determined by parasites capable of infecting multiple host species. Estimates that up to 70% of human parasites are **zoonotic** (i.e. shared by humans and at least one other animal species; see Glossary) showcase that multi-host parasites, as opposed to those that only infect a single host species, are the rule rather than the exception [1, 2]. Frequent identification of **Emerging Infectious Diseases** (EIDs) highlights the global importance of contemporary **host shifting** that can result in infection of novel and often immunologically naïve hosts [3, 4].

**Host specificity**, representing the number and/or diversity of host species a parasite is capable of infecting, is considered a key indicator of its propensity to shift hosts [5]. Host specificity is a topic of considerable interest in the fields of disease ecology and One Health that is quickly becoming a key feature of research agendas (Figure 1). Much of this work is aimed at delineating possible EID **reservoirs** by attempting to identify parasites that may be capable of shifting between wildlife and human hosts [6-8]. Wildlife parasitology research has uncovered a broad spectrum of host specificity 'strategies' for an impressive diversity of parasites. These range from haemosporidian blood parasites and feather mites that associate

60    with only a small number of closely-related host species [9, 10] to rabies viruses that are

61    seemingly capable of infecting virtually any mammalian species they encounter [11].

62

63    Intuitively, one may assume that among multi-host parasites, those that exhibit low host

64    specificity (i.e. high levels of host generalism) are more capable to shift hosts [12]. Indeed,

65    numerous studies have identified so-called 'generalist' parasites that seem capable of

66    infecting a broad spectrum of phylogenetically and sometimes ecologically dissimilar host

67    species [5, 13, 14]. Yet whether low host specificity meaningfully reflects a higher risk for

68    EID-related **spillover** remains unresolved [15, 16]. This is largely because there are multiple

69    facets of host specificity, many of which are not captured by common metrics, and there are

70    idiosyncratic ways in which different host species contribute to a parasite's specificity [17].

71    Moreover, there is a great deal of uncertainty about whether any retrospective summary of

72    observed host ranges into host specificity measures, which basically reflect the accumulation

73    of historical host-parasite interactions, translate into a parasite's potential for contemporary

74    host shifting [18].

75

76    Here, we summarize the state of current research on host specificity and highlight how such

77    work can play a role in advancing our ability to quantify host shifting capacity. We extend

78    previous reviews of how novel species communities may relate to shifting host ranges and

79    variation in parasite transmission dynamics [4, 17, 19] by discussing challenges in the use of

80    host specificity metrics and outlining frameworks that align common data structures to

81    relevant modelling tools.

82

83    **The eco-evolutionary backbone of host specificity**

84    Specialization of species in parasitic or mutualistic interactions is assumed to be tightly

85    linked to the levels of adaptation exhibited by interacting partners [20]. Host-parasite

86    interactions often require highly tuned transmission modes and adaptations by the parasite to

87    survive and reproduce in the host environment [21]. It therefore comes as little surprise that

88    much of the emphasis surrounding investigation of host-parasite interactions has been placed

89    on host-parasite co-evolution. Co-speciation, whereby a parasite species evolves into two

90    distinct species in response to host speciation, has long been used to explain apparent

91    congruence in host and parasite phylogenies. The prevailing reasoning is that a parasite's

92    evolutionary history sets the stage for host specificity by introducing phylogenetic and

93  ecological barriers to host shifting [22, 23]. Host shifting requires parasites to be exposed to

94  new hosts that exhibit certain levels of physiological and/or behavioural overlap with

95  previous hosts (**ecological fitting**), as this allows circumvention of barriers caused by

96  variation in host competence or immunity [22, 24].

97  An expanded line of thought, primarily derived from studying mutualisms such as

98  animal-plant interactions, suggests that host shifting is probabilistic and relies on opportunity

99  for hosts and parasites to interact under variable environmental conditions [4, 25]. A

100 prominent example involves herbivorous insects introduced as biocontrol agents into

101 different environments, which showcases that regional host ranges are largely determined by

102 local environmental conditions such as plant community composition, relative abundance and

103 phenology in response to climate variation [26, 27].

104 By analogy, we argue that similar environmental forces will also be important for

105 shaping host-parasite interactions [4, 28-30]. This idea that a parasite's capacity for host

106 shifting can vary in response to environmental conditions (Figure 2) presents a new forefront

107 of research on spillover risk in times of global change [4, 31-34]. Fortunately, a burst in

108 analytical tools designed to explore spatiotemporal variation in species interactions [35-37]

109 make it possible to characterize how host specificity changes across environmental gradients.

110 Consistent frameworks are now needed to disentangle the evolutionary and ecological aspects

111 of host-parasite interactions that should be considered when judging host specificity.

112

113 **A niche perspective on host specificity**

114 Developing a framework to assess host specificity, and to begin relating specificity to

115 potential spillover risk, relies first on developing a consistent definition for host specificity.

116 This is particularly necessary when considering the staggering diversity of advocated indices

117 and metrics (Table 1). The most common of these focus on the range of hosts a parasite is

118 observed to infect and in which it can persist and/or complete its life cycle. The simplest way

119 to do this is to count the number of host species a parasite infects [10]. But this provides little

120 information about the diversity of host 'habitats' that comprise a parasite's **niche**. Authors

121 have more recently recognized that adapting ecological niche concepts to host specificity can

122 improve understanding of a parasite's host range using concepts from a widely-supported

123 theoretical framework [13, 32]. An important aspect of niche theory distinguishes between

124 **potential resources** (resources that a species could utilise if it encountered them) and

125 surrounding conditions that determine whether resources are available and can be utilized

126 (**environmental filtering**). Recent advances in ecological niche modelling have capitalized

127 on the concept of potential resources to expand our understanding of niche filling by

128 differentiating between a species' **fundamental niche** and its **realized niche**. Here, the

129 fundamental niche refers to resources that can be used in the absence of any restrictive

130 conditions, while the realized niche refers to regional consumer-resource dynamics that are

131 observed under local environmental conditions [38].

132         Translated to multi-host parasites, the fundamental niche describes a parasite's

133 capacity to explore different host species independently of conditions that may restrict its

134 exposure to these potential hosts. The realized niche, in turn, is based on the regional

135 diversity of host species a parasite is actually observed to infect [39](Figure 2).

136 Biogeographic structuring of host assemblages across regional scales will constrain

137 opportunities for host-parasite interactions. In other words, variation in the diversity of

138 potential hosts occurring in regional species pools can ensure that a parasite's fundamental

139 host specificity is not completely realized [40]. Empirical support for this can be derived

140 from a number of field studies from diverse host-parasite systems. Biogeographic structure in

141 host species distributions likely narrows the realized host specificity of chewing lice infecting

142 toucans, particularly when closely related host species are spatially disconnected [41]. Host

143 range expansions by relatively fast evolving (RNA) rabies viruses depend on local

144 compositions of bat assemblages, as different virus lineages often cannot cross species

145 barriers to infect distantly related host species [11]. Regional climate conditions that

146 influence vector habitats are associated with observed host specificities for widespread avian

147 malaria parasites [34]. Experimental studies that artificially increase the host range accessible

148 to a parasite further support the concept of fundamental vs realized host specificity: the set of

149 hosts that can be infected experimentally can be much larger than the actual range observed

150 under natural conditions, even for parasites that are only observed to infect a single host

151 species [42]. This increasing recognition that local variation changes the suite of hosts to

152 which a parasite is exposed and pre-existing capacity enables host shifting upon newly

153 arising opportunities has been raised by a number of recent studies that collectively

154 contribute to a meta-theory called the **Stockholm Paradigm** [32]).

155

## Inferring host specificity

157 Ecological dynamics impact the host specificity of parasites across local and regional scales

158 [43, 44]; this has important ramifications for formulating concepts to gather inferences about

159    explanatory mechanisms. Identifying factors that influence opportunities for novel host-

160    parasite interactions, and how these shape realized host specificity, are vital steps to begin

161    uncovering the true fundamental host specificity (see **Box 1**). This requires an understanding

162    of how host communities are shaped and how resources are utilised, both of which have

163    strong parallels in community ecology. A growing consensus states that a suite of factors

164    such as evolutionary contingencies, speciation, adaptive plasticity, dispersal capacity,

165    environmental filters and biotic interactions [45-48] act in concert with stochastic processes

166    [49] to shape communities. Accounting for plasticity in resource utilisation helps to align

167    ecological modelling approaches to reality for understanding shifts in species' habitat use,

168    geographical ranges, or trophic interactions across environmental gradients [20, 50].

169    Moreover, tractable modelling concepts that can untangle the effects of environmental

170    forcing and species interactions on resource utilisation have recently been developed [51-54].

171    Note that in most situations, sufficient empirical and experimental evidence is

172    necessary to infer aspects of fundamental host specificity (see Box 2 for some considerations

173    about matching data to study questions and models). We now outline a number of important

174    modelling approaches that are available to begin understanding how host specificity is shaped

175    across scales.

176

177    ***Trait-based approaches to host specificity***

178    Trait-based approaches, which group species of interest according to attributes such as

179    phylogenetic relationships, body size, diet, climate tolerance or distributions, are increasingly

180    adopted to study species ecological preferences and their potential responses to global change

181    [55, 56]. In the context of host specificity, there are at least two trait-based approaches that

182    have received considerable attention recently.

183    First, indices of functional and phylogenetic diversity are used to delineate host

184    specificity according to the observed (or estimated) variation in traits exhibited by competent

185    host species [5]. Often, such measures are based on pair-wise distances that are calculated

186    among all possible combinations of viable host species [57, 58]. For instance, phylogenetic

187    trees, which depict evolutionary relationships among host species, can be used to generate

188    indices of phylogenetic host specificity [17, 59, 60]. In a similar way, distance-based

189    diversity measures can be generated using host species' ecological traits [14, 61], with

190    supporting computational algorithms readily available in open-source software such as R

191    [62]. The central aim when using these measures is to determine whether the distribution of

192    pairwise distances between infected host species (observed distances) is different to a

193    distribution of distances between potential host species (expected distances). Here, an

194    observed distance distribution that is statistically smaller than expected indicates that infected

195    hosts are more closely related than expected by chance. It is worth noting that such metrics

196    largely depend on sample size (i.e. the number of host species recorded) and thus are

197    potentially subject to sampling bias [63]. Null model permutation approaches [64] and more

198    recently, hierarchical models [40], have been proposed to account for sampling bias while

199    comparing observed and expected distance distributions. Recent studies have put these ideas

200    into practice by demonstrating that host functional traits are as important as phylogenetic

201    relationships for assessing whether primates share the same parasites [15] and by showing

202    that host phylogenetic relationships appear to strongly shape the host ranges of avian malaria

203    lineages [14]. Moreover, a recent multi-taxa study deciphered that phylogenetic specialization

204    among prospective hosts is more pronounced for helminths and viruses than for other parasite

205    groups [65]. Finally, similar trait-based measures were used to show that fleas with certain

206    traits are more likely to infest the same subsets of phylogenetically and functionally related

207    mammalian host species [66], showcasing that evolutionary history and ecological fitting

208    synergistically drive the realized host specificity of these ectoparasites.

209          Trait-based regression models have also gained popularity for analysing whether the

210    presence-absence of parasites in a suite of host species is linked to host traits [6-8]. Such

211    approaches are of relevance for host specificity measures as model-based estimates of a

212    parasite's associations with particular host traits can enable projections onto unmeasured host

213    species, enabling prediction of unknown interactions [67]. Generalised linear models (GLMs;

214    readily estimated using Maximum Likelihood or Bayesian frameworks through available R

215    packages; [68-70]) or machine-learning regression trees [71] are the methods of choice as

216    they estimate associations using data from a variety of outcome distributions. Prominent

217    examples have found that higher proportions of zoonotic viruses occur in mammals that are

218    closely related to humans [8], whereas studies of rodents have shown that hosts with faster

219    life histories have higher occurrences of zoonotic viruses [6]. Another study found that the

220    intensity of helminth parasite sharing between humans, domestic animals and wildlife

221    appeared to be predominantly driven by dietary traits of wildlife species [7]. At the species

222    level (i.e. presence-absence of a particular parasite in a suite of host species), however, we

223    stress again that such regressions are only useful if the underpinning data include sufficient

224    numbers of both presence and absence records to allow meaningful inference.

225

226 *Network approaches to study host-parasite interactions*

227 Ecological networks summarize biotic interactions among groups of species that live in

228 trophic or symbiotic relationships [72]. Amongst the many useful insights gained from such

229 community-scale analysis, they allow users to draw conclusions about the preference

230 (specialization) of species towards a suite of potential partners [73]. Networks can be

231 constructed as bipartite graphs that denote whether species interact or not (e.g., binary

232 'presence-absence' data of interactions between combinations of host and parasite species) or

233 the frequencies of interactions (e.g., the number/proportion of individuals from a given pool

234 of host species infected with the concerned parasites). A simple measure of a parasite's

235 specialization derived from binary networks can be the proportion of host species infected

236 [74], an index that resembles host specificity without taking link strengths into account. This

237 measure can be linked at the community scale when used in combination with null model

238 analysis [75], allowing users to ask which parasites are more or less specialist. In contrast,

239 many network measures of species-level host specificity, such as the species-specific

240 specialization index $d'$ [76], are based on both the link distributions as well link strengths.

241 Such community-scale analysis acknowledges the fact that every single interaction is

242 embedded in a network of species interactions and depicts a parasite's

243 preference/specialization on particular hosts relative to both the overall host availability and

244 the host utilisation by other parasites. Such an approach was recently used to explore network

245 compositions of fish parasites and mammalian fleas across a number of regions [74]. The

246 authors showed that specialist parasites tended to interact with hosts that harboured high

247 richness of parasites, and that hosts with high parasite richness also tended to be more

248 abundant. However, given the fact that network specificity indices are commonly derived

249 from a finite set of community-scale observations, their utility for predicting fundamental

250 host specificity needs to be carefully evaluated. This is because host specificity is a species-

251 specific attribute, and we argue that host specificity is not necessarily shaped by the

252 specificities of other parasites in the community. Moreover, networks assembled under

253 particular regional conditions will only yield measures of realized host specificity and thus do

254 not necessarily provide accurate insights about host specificity under novel (unsampled)

255 environmental conditions. Nevertheless, advances in techniques to model how ecological

256 network properties respond to environmental variation [36, 77] may provide promising

257 opportunities to estimate fundamental host specificity in future research.

258      Ecological network concepts have also been employed to detect the centrality of key

259 host species and/or the modularity of interaction compositions in observed host-parasite

260 networks [78]. These approaches aggregate host-parasite interactions into adjacency matrices

261 to depict focal species that share similar sets of interactions (i.e. two host species that are

262 infected by similar parasites may be connected within the network; [78, 79]). This of course

263 comes at the cost of losing information about species identity, but can nevertheless be helpful

264 for identifying roles that different host species may play for facilitating parasite spread or for

265 understanding whether host-parasite interactions exhibit a modular or nested structure.

266

267 *Identifying environmental filters related to realized host specificity*

268 Despite the examples outlined above, few studies have examined how changes in host

269 specificity relate to spatiotemporal changes in environmental conditions [80, 81]. Capturing

270 the complex ways in which environmental filtering can affect realized host specificity is a

271 looming challenge that calls for integrative approaches to consolidate the synergies between

272 species distributions and biotic interactions [46]. For example, if variation in realized host

273 specificity is linked to changes in regional host composition [40], a comprehensive

274 understanding of how environmental filters impact realized host specificity requires

275 disentangling their effects on host species occurrence and on host-parasite interactions (i.e. by

276 influencing epidemiological factors such as host susceptibility, parasite survival and

277 transmission potential). Some first step towards capturing this process have been taken by

278 applying statistical models that estimate how realized host specificity changes in relation to

279 the variation in host community compositions [40, 82] and environmental filters such as

280 climate [34]. These approaches expand on the trait-based methods described above by

281 comparing suites of ecological and/or phylogenetic distances among infected pairs of host

282 species to distances that describe all potential host species within each region that a parasite

283 occupies. Results have provided promising new insights. For example, Wells *et al.* [40]

284 showed that helminth parasites generally exhibited the lowest phylogenetic host specificity in

285 regional 'hotspots' that showed high variation in prospective host diversity; despite being

286 globally distributed, some parasites still infected less functionally diverse hosts than

287 expected, indicating limited potential to infect hosts from different ecological niches. Fecchio

288 *et al.* [34] showed that avian malaria parasites are more constrained in their capacity to

289 exploit a diversity of host species in regions with pronounced rainfall seasonality and wetter

290 dry seasons. Other recent developments can account for biotic interactions within

291 multivariate community models, which can be helpful to understand how host-parasite

292 interactions may change across environmental gradients [52]. While we are unaware of these

293 models being used to assess changes in realized host specificity, their ability to detect

294   associations among different parasites opens exciting avenues for uncovering how host
295   specificity is shaped across regional scales.
296

297   ***The need to capture uncertainty: probabilistic tools and an outlook on forecasting host***
298   ***shifting***
299   Because observational data dominates the host specificity literature, the above sections stress
300   that taking advantage of contemporary modelling tools for best-possible inference offers
301   significant improvements over simply drawing conclusions from finite observations at hand.
302   We also urge the use of probabilistic methods when applying such tools, as they can
303   distinguish drivers of host-parasite interactions from underlying observation processes to
304   obtain model-based estimates [83-85]. Probabilistic sampling approaches have a number of
305   benefits that make them suitable to host specificity research, including: (1) Capturing
306   uncertainty and sampling bias in infection and host-parasite association data, and allowing for
307   random data imputation/augmentation (i.e. for poorly sampled host species there might be a
308   certain probability this species is infected, even if records of such associations are missing);
309   (2) Utilising a diversity of data sources such as host-parasite association data, trait variables
310   and spatiotemporal environmental data in consistent model frameworks; (3) Capturing the
311   hierarchical nature of realized host specificity by conceptualizing conditional dependencies
312   such as 'the probability a host species contributes to realized host specificity, conditional on
313   its presence and compatibility under regional conditions'. The ability to capture uncertainties
314   is particularly imperative for forecasting the possible infection of a novel host species under
315   future or as yet unexplored environmental conditions, which is the basis for mitigating the
316   public and animal health risks posed by EIDs. Given the importance of plasticity in host
317   specificity, quantitative solutions require developing and applying tractable forecasting tools
318   to answer questions such as 'how likely is a parasite to shift from one host species to another
319   under XYZ regional conditions?'. This can be a challenging task for a large range of parasites
320   that affect only a small number of host species and/or for which retrospective data on host
321   shifting events are rare.
322        Despite these challenges, novel modelling frameworks offer a foundation for
323   prediction about when and in which host species a parasite may occur. Correlative
324   approaches in species distribution and ecological niche modelling, for example, aim to
325   estimate species persistence under a range of surveyed environmental conditions (aiming to
326   describe the pattern but not necessarily the underlying mechanism) and then project species

327  distributions across larger environmental space [86]. The trait-based regression models

328  discussed above fall into this category, and may be useful for prediction as they generally

329  have moderate data needs, including the presence-absence of host-parasite associations and

330  information about host traits and environmental conditions. However, projecting models to

331  estimate a parasite's fundamental host range (and to predict host shifting events) across

332  gradients of substantial spatiotemporal change might significantly violate model assumptions

333  and could be hampered by poor **transferability** [87]. Only for parasites with invariable

334  realized host specificity across environmental gradients would one expect to obtain

335  reasonable accuracy in projected host specificity under unsampled conditions. This raises the

336  need for detailed sampling regimes that cover much (if not all) of the parasite's known

337  distribution. In light of the potential drawbacks of correlative predictions, a key aim for future

338  research could be to use historical data on true host shifting events to see if any patterns of

339  realized host specificity prior to the shift could have been informative for prediction. Using

340  such hindcasting approaches to evaluate our capacity to forecast future EIDs could tell us

341  whether any of our metrics have any real value, or whether other situational aspects (such as

342  rapid changes in contact rates, the emergence of new host-host contacts or the stochastic

343  emergence of new strains) are more important.

344         In contrast to correlative approaches, processed-based methods explicitly model the

345  important processes underlying patterns [88, 89]. In terms of host-parasite interactions, such

346  models may aim to predict and reproduce host shifting events from a suite of eco-

347  epidemiological factors that jointly drive system dynamics, including variation in contact

348  opportunities among host species and spatiotemporal environmental variation. Perhaps a

349  promising move towards predicting fundamental host specificity and forecasting EIDs could

350  be the establishment of so-called hybrid models. Hybrid models synthesise correlative and

351  process-based models by combining static projections from correlative approaches with

352  simulation of key processes; in species distribution modelling, these processes include

353  aspects such as species abundances, the realistic co-occurrence of interacting species and

354  dispersal events [90]. Hybrid models can improve the transferability of correlation-based

355  approaches by more realistically accounting for key processes while avoiding specification of

356  a large number of parameters [91]. For our purposes, such models could aim to capture the

357  most essential dynamics underlying host shifting, such as possible range shifts of key hosts,

358  expected variation in host community composition in relation to climate change [92] or the

359  emergence of species invasion 'hotspots' [93].

360

## Concluding Remarks

A large body of research on human, animal and plant parasites uses host specificity to gauge the risk of EIDs and spillover events. However, in times of global change and the large-scale spread of parasites across former geographic barriers, drawing conclusions about a parasite's host shifting capacity using simple specificity indices may not be suitable to predict such events under novel conditions. Host specificity cannot be considered a fixed trait, as environmental conditions cause considerable variation in realized host specificity. The task of predicting host shifting events must rely on tractable modelling frameworks that sit at the core of ecological forecasting [94]. Ultimately, accounting for plasticity and uncertainty in a parasite's realized host specificity may be a worthy step to better predict disease emergence and host shifting events (see **Outstanding questions**). But identifying the multifaceted processes involved in multi-host parasite transmission is laborious and will require considerable empirical and quantitative research. Along the way towards understanding such complexities, we should not ignore the fact that disease emergence inevitably means parasites often have a hidden potential to infect novel host species. Initial conclusions based on realized host specificity alone need to be carefully revisited as more data becomes available. This will leverage our growing understanding about which parasites are capable of crossing the species barrier and causing unwanted diseases.

## Outstanding questions

• Are there generalities about the plasticity in host specificity in response to environmental conditional for parasites from different taxonomic groups and/or with different transmission modes?

• Do strong host-parasite co-evolutionary histories constrain the plasticity in host specificity and potential for host shifting across environmental gradients?

• Can integrated model frameworks and validation procedures for inferring fundamental host specificity allow us to better predict future host shifting events?

## References

1. Taylor, L.H. et al. (2001) Risk factors for human disease emergence. *Philos. Trans. Royal Soc. B* 356, 983-989.

2. Jones, K.E. et al. (2008) Global trends in emerging infectious diseases. *Nature* 451, 990-994.

3. Woolhouse, M.E.J. and Gowtage-Sequeria, S. (2005) Host range and emerging and reemerging pathogens. *Emerg. Infect. Dis.* 11, 1842-1847.

4. Nylin, S. et al. (2018) Embracing colonizations: a new paradigm for species association dynamics. *Trends Ecol. Evol.* 33, 4-14.

5. Poulin, R. and Mouillot, D. (2003) Parasite specialization from a phylogenetic perspective: a new index of host specificity. *Parasitology* 126, 473-480.

6. Han, B.A. et al. (2015) Rodent reservoirs of future zoonotic diseases. *Proc. Natl. Acad. Sci. U. S. A.* 112, 7039-7044.

7. Wells, K. et al. (2018) Global spread of helminth parasites at the human–domestic animal–wildlife interface. *Global Change Biol.* 24, 3254-3265.

8. Olival, K.J. et al. (2017) Host and viral traits predict zoonotic spillover from mammals. *Nature* 546, 646.

9. Clark, N.J. et al. (2014) A review of global diversity in avian haemosporidians (*Plasmodium* and *Haemoproteus*: Haemosporida): new insights from molecular data. *Int. J. Parasitol.* 44, 329-338.

10. Doña, J. et al. (2017) Host specificity, infrequent major host switching and the diversification of highly host-specific symbionts: The case of vane-dwelling feather mites. *Global Ecol. Biogeogr.* 27, 188-198.

11. Streicker, D.G. et al. (2010) Host phylogeny constrains cross-species emergence and establishment of rabies virus in bats. *Science* 329, 676-679.

423   12. Cleaveland, S. et al. (2001) Diseases of humans and their domestic mammals: pathogen
424          characteristics, host range and the risk of emergence. *Philos. Trans. Royal Soc. B* 356,
425          991-999.

426   13. Krasnov, B.R. et al. (2004) Ectoparasitic "jacks-of-all-trades": relationship between
427          abundance and host specificity in fleas (Siphonaptera) parasitic on small mammals.
428          *Am. Nat.* 164, 506-516.

429   14. Clark, N.J. and Clegg, S.M. (2017) Integrating phylogenetic and ecological distances
430          reveals new insights into parasite host specificity. *Mol. Ecol.* 26, 3074-3086.

431   15. Cooper, N. et al. (2012) Phylogenetic host specificity and understanding parasite sharing
432          in primates. *Ecol. Lett.* 15, 1370-1377.

433   16. Brooks, D.R. et al. (2014) Finding them before they find us: informatics, parasites, and
434          environments in accelerating climate change. *Comp. Parasitol.* 81, 155-164.

435   17. Poulin, R. et al. (2011) Host specificity in phylogenetic and geographic space. *Trends
436          Parasitol.* 27, 355-361.

437   18. Morse, S.S. et al. (2012) Prediction and prevention of the next pandemic zoonosis. *The
438          Lancet* 380, 1956-1965.

439   19. Young, H.S. et al. (2017) Introduced species, disease ecology, and biodiversity-disease
440          relationships. *Trends Ecol. Evol.* 32, 41-54.

441   20. Devictor, V. et al. (2010) Defining and measuring ecological specialization. *J. Appl. Ecol.*
442          47, 15-25.

443   21. Poulin, R. (2007) *Evolutionary ecology of parasites*, Princeton University Press.

444   22. Agosta, S.J. and Klemens, J.A. (2008) Ecological fitting by phenotypically flexible
445          genotypes: implications for species associations, community assembly and evolution.
446          *Ecol. Lett.* 11, 1123-1134.

447   23. Araujo, S.B.L. et al. (2015) Understanding host-switching by ecological fitting. *PLoS
448          ONE* 10, e0139225.

449   24. Woolhouse, M.E.J. et al. (2005) Emerging pathogens: the epidemiology and evolution of
450          species jumps. *Trends Ecol. Evol.* 20, 238-244.

451   25. Thrall, P.H. et al. (2007) Coevolution of symbiotic mutualists and parasites in a
452          community context. *Trends Ecol. Evol.* 22, 120-126.

453   26. Zwölfer, H. and Harris, P. (1971) Host specificity determination of insects for biological
454          control of weeds. *Annu. Rev. Entomol.* 16, 159-178.

27. Van Klinken, R.D. and Edwards, O.R. (2002) Is host-specificity of weed biological control agents likely to evolve rapidly following establishment? *Ecol. Lett.* 5, 590-596.

28. Pedersen, A.B. et al. (2005) Patterns of host specificity and transmission among parasites of wild primates. *Int. J. Parasitol.* 35, 647-657.

29. Janz, N. (2011) Ehrlich and Raven revisited: mechanisms underlying codiversification of plants and enemies. *Annu. Rev. Ecol. Evol. Syst.* 42, 71-89.

30. Brooks, D.R. and Boeger, W.A. (2019) Climate change and emerging infectious diseases: evolutionary complexity in action. *Curr. Opin. Syst. Biol.* 13, 75-81.

31. Morand, S. and R., K.B. (2010) *The biogeography of host-parasite interactions*, Oxford University Press.

32. Hoberg, E.P. and Brooks, D.R. (2015) Evolution in action: climate change, biodiversity dynamics and emerging infectious disease. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 370, 20130553.

33. Wells, K. et al. (2015) The importance of parasite geography and spillover effects for global patterns of host–parasite associations in two invasive species. *Diversity Distrib.* 21, 477-486.

34. Fecchio, A. et al. (2019) Climate variation influences host specificity in avian malaria parasites. *Ecol. Lett.* 22, 547-557.

35. Thorson, J.T. et al. (2016) Joint dynamic species distribution models: a tool for community ordination and spatio-temporal monitoring. *Global Ecol. Biogeogr.* 25, 1144-1158.

36. Dormann, C.F. et al. (2017) Identifying causes of patterns in ecological networks: opportunities and limitations. *Annu. Rev. Ecol. Evol. Syst.* 48, 559-584.

37. Delmas, E. et al. (2018) Analysing ecological networks of species interactions. *Biol. Rev.* 94, 16-36.

38. Soberón, J. (2007) Grinnellian and Eltonian niches and geographic distributions of species. *Ecol. Lett.* 10, 1115-1123.

39. Calatayud, J. et al. (2016) Geography and major host evolutionary transitions shape the resource use of plant parasites. *Proc. Natl. Acad. Sci. U. S. A.* 113 (35), 9840-9845.

40. Wells, K. et al. (2018) Global patterns in helminth host specificity: phylogenetic and functional diversity of regional host species pools matter. *Ecography* 42, 416-427.

41. Weckstein, J.D. (2004) Biogeography explains cophylogenetic patterns in toucan chewing lice. *Syst. Biol.* 53, 154-164.

489  42. Perlman, S.J. and Jaenike, J. (2003) Infection success in novel hosts: an experimental and
490      phylogenetic study of Drosophila-parasitic nematodes. *Evolution* 57, 544-557.

491  43. Penczykowski, R.M. et al. (2015) Understanding the ecology and evolution of host–
492      parasite interactions across scales. *Evol. Appl.* 9, 37-52.

493  44. Plowright, R.K. et al. (2015) Ecological dynamics of emerging bat virus spillover. *Proc.*
494      *R. Soc. B Biol. Sci.* 282 (1798).

495  45. Gravel, D. et al. (2006) Reconciling niche and neutrality: the continuum hypothesis. *Ecol.*
496      *Lett.* 9, 399-409.

497  46. Wisz, M.S. et al. (2013) The role of biotic interactions in shaping distributions and
498      realised assemblages of species: implications for species distribution modelling. *Biol.*
499      *Rev.* 88, 15-30.

500  47. Viana, D.S. et al. (2016) Migratory birds as global dispersal vectors. *Trends Ecol. Evol.*
501      31, 763-775.

502  48. Clark, N.J. et al. (2017) Climate, host phylogeny and the connectivity of host
503      communities govern regional parasite assembly. *Diversity Distrib.* 24, 13-23.

504  49. Chase, J.M. and Myers, J.A. (2011) Disentangling the importance of ecological niches
505      from stochastic processes across scales. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 366,
506      2351-2363.

507  50. Guisan, A. and Zimmermann, N.E. (2000) Predictive habitat distribution models in
508      ecology. *Ecol Modell.* 135, 147-186.

509  51. Warton, D.I. et al. (2015) So many variables: joint modeling in community ecology.
510      *Trends Ecol. Evol.* 30, 766–779.

511  52. Ovaskainen, O. et al. (2017) How to make more out of community data? A conceptual
512      framework and its implementation as models and software. *Ecol. Lett.* 20, 561-576.

513  53. Clark, N.J. et al. (2016) Co-infections and environmental conditions drive the
514      distributions of blood parasites in wild birds. *J. Anim. Ecol.* 85, 1461–1470

515  54. Dormann, C.F. et al. (2018) Biotic interactions in species distribution modelling: 10
516      questions to guide interpretation and avoid false conclusions. *Global Ecol. Biogeogr.*
517      27, 1004-1016.

518  55. McGill, B.J. et al. (2006) Rebuilding community ecology from functional traits. *Trends*
519      *Ecol. Evol.* 21, 178-185.

520  56. Webb, C.T. et al. (2010) A structured and dynamic framework to advance traits-based
521      theory and prediction in ecology. *Ecol. Lett.* 13, 267-283.

522    57. Villéger, S. et al. (2008) New multidimensional functional diversity indices for a
523        multifaceted framework in functional ecology. *Ecology* 89, 2290–2301.

524    58. Laliberté, E. and Legendre, P. (2010) A distance-based framework for measuring
525        functional diversity from multiple traits. *Ecology* 91, 299-305.

526    59. Mouchet, M. et al. (2008) Towards a consensus for calculating dendrogram-based
527        functional diversity indices. *Oikos* 117, 794-800.

528    60. Tucker, C.M. et al. (2017) A guide to phylogenetic metrics for conservation, community
529        ecology and macroecology. *Biol. Rev.* 92, 698-715.

530    61. Cadotte, M. et al. (2013) The ecology of differences: assessing community assembly with
531        trait and evolutionary distances. *Ecol. Lett.* 16, 1234-1244.

532    62. Kembel, S.W. et al. (2010) Picante: R tools for integrating phylogenies and ecology.
533        *Bioinformatics* 26.

534    63. Plas, F. et al. (2017) Sensitivity of functional diversity metrics to sampling intensity.
535        *Methods Ecol. Evol.* 8, 1072-1080.

536    64. Webb, C.O. et al. (2002) Phylogenies and community ecology. *Annu. Rev. Ecol. Syst.* 33,
537        475-505.

538    65. Park, A.W. et al. (2018) Characterizing the phylogenetic specialism–generalism spectrum
539        of mammal parasites. *Proc. R. Soc. B Biol. Sci.* 285, 20172613.

540    66. Krasnov, B.R. et al. (2016) Trait-based and phylogenetic associations between parasites
541        and their hosts: a case study with small mammals and fleas in the Palearctic. *Oikos*
542        125, 29-38.

543    67. Morales-Castilla, I. et al. (2015) Inferring biotic interactions from proxies. *Trends Ecol.*
544        *Evol.* 30, 347-356.

545    68. Hadfield, J.D. (2010) MCMC methods for multi-response generalized linear mixed
546        models: The MCMCglmm R package. *J. Stats. Softw.* 33, 1-22.

547    69. Bates, D. et al. (2015) Fitting linear mixed-effects models using lme4. *J. Stats. Softw.* 67,
548        48.

549    70. Goodrich, B. et al. (2018) rstanarm: Bayesian applied regression modeling via Stan. R
550        package version 2.17.4, http://mc-stan.org/.

551    71. Elith, J. et al. (2008) A working guide to boosted regression trees. *J. Anim. Ecol.* 77, 802-
552        813.

553    72. Ings, T.C. et al. (2009) Ecological networks - beyond food webs. *J. Anim. Ecol.* 78, 253-
554        269.

555    73. Poulin, R. (2010) Network analysis shining light on parasite ecology and diversity.
556         *Trends Parasitol.* 26, 492-498.

557    74. Vázquez, D.P. et al. (2005) Species abundance and the distribution of specialization in
558         host-parasite interaction networks. *J. Anim. Ecol.* 74, 946-955.

559    75. Poisot, T. et al. (2012) A comparative study of ecological specialization estimators.
560         *Methods Ecol. Evol.* 3, 537-544.

561    76. Blüthgen, N. and Menzel, F. (2006) Measuring specialization in species interaction
562         networks. *BMC Ecology* 6, 9.

563    77. Tylianakis, J.M. and Morris, R.J. (2017) Ecological networks across environmental
564         gradients. *Annu. Rev. Ecol. Evol. Syst.* 48, 25-48.

565    78. Gómez, J.M. et al. (2013) Centrality in primate–parasite networks reveals the potential for
566         the transmission of emerging infectious diseases to humans. *Proc. Natl. Acad. Sci. U.*
567         *S. A.* 110, 7738-7741.

568    79. Dallas, T. and Cornelius, E. (2015) Co-extinction in a host-parasite network: identifying
569         key hosts for network stability. *Sci. Rep.* 5, 13185.

570    80. Altizer, S. et al. (2013) Climate change and infectious diseases: from evidence to a
571         predictive framework. *Science* 341, 514-519.

572    81. Wright, M.G. and Bennett, G.M. (2018) Evolution of biological control agents following
573         introduction to new environments. *BioControl* 63, 105-116.

574    82. Drake, J.M. and Richards, R.L. (2018) Estimating environmental suitability. *Ecosphere* 9
575         (9), e02373.

576    83. Wells, K. and O'Hara, R.B. (2013) Species interactions: estimating per-individual
577         interaction strength and covariates before simplifying data into per-species ecological
578         networks. *Methods Ecol. Evol.* 4, 1-8.

579    84. Coblentz, K.E. et al. (2017) The application of Bayesian hierarchical models to quantify
580         individual diet specialization. *Ecology* 98, 1535-1547.

581    85. Graham, C.H. and Weinstein, B.G. (2018) Towards a predictive model of species
582         interaction beta diversity. *Ecol. Lett.* 21, 1299-1310

583    86. Peterson, A.T. et al. (2015) Mechanistic and correlative models of ecological niches.
584         *European J. Ecol.* 1, 28-38.

585    87. Vaughan, I.P. and Ormerod, S.J. (2005) The continuing challenges of testing species
586         distribution models. *J. Appl. Ecol.* 42, 720-730.

587    88. Grimm, V. et al. (2005) Pattern-oriented modeling of agent-based complex systems:
588         lessons from ecology. *Science* 310, 987-991.

589  89. Kearney, M. and Porter, W. (2009) Mechanistic niche modelling: combining
590      physiological and spatial data to predict species' ranges. *Ecol. Lett.* 12, 334-350.

591  90. Ehrlén, J. and Morris, W.F. (2015) Predicting changes in the distribution and abundance
592      of species under environmental change. *Ecol. Lett.* 18, 303-314.

593  91. Dormann, C.F. et al. (2012) Correlation and process in species distribution models:
594      bridging a dichotomy. *J. Biogeogr.* 39, 2119-2131.

595  92. Santini, L. et al. (2016) A trait-based approach for predicting species responses to
596      environmental change from sparse data: how well might terrestrial mammals track
597      climate change? *Global Change Biol.* 22, 2415-2424.

598  93. Dawson, W. et al. (2017) Global hotspots and correlates of alien species richness across
599      taxonomic groups. *Nat. Ecol. Evol.* 1, 0186.

600  94. Mouquet, N. et al. (2015) Predictive ecology in a changing world. *J. Appl. Ecol.* 52,
601      1293-1310.

602  95. Weldon, C. et al. (2004) Origin of the amphibian chytrid fungus. *Emerg. Infect. Dis.* 10,
603      2100-2105.

604  96. Holt, R.D. et al. (2003) Parasite establishment in host communities. *Ecol. Lett.* 6, 837-
605      842.

606  97. Johnson, P.T.J. et al. (2015) Why infectious disease research needs community ecology.
607      *Science* 349, 1259504.

608  98. Cleaveland, S. et al. (2000) Serological and demographic evidence for domestic dogs as a
609      source of canine distemper virus infection for Serengeti wildlife. *Vet. Microbiol.* 72,
610      217-227.

611  99. Clark, N.J. et al. (2018) Parasite spread at the domestic animal - wildlife interface:
612      anthropogenic habitat use, phylogeny and body mass drive risk of cat and dog flea
613      (*Ctenocephalides* spp.) infestation in wild mammals. *Parasites Vectors* 11, 8.

614  100. Romig, T. et al. (2017) Ecology and life cycle patterns of *Echinococcus* species. In
615      *Advances in Parasitology* (Thompson, R.C.A. et al. eds), pp. 213-314, Academic
616      Press.

617  101. Telfer, S. et al. (2010) Species interactions in a parasite community drive infection risk
618      in a wildlife population. *Science* 330, 243-246.

619  102. Susi, H. et al. (2015) Co-infection alters population dynamics of infectious disease. *Nat.*
620      *Commun.* 6, 5975.

621  103. Parola, P. et al. (2008) Warmer weather linked to tick attack and emergence of severe
622      Rickettsioses. *PLoS Negl. Trop. Dis.* 2, e338.

623    104. Kurtenbach, K. et al. (2006) Fundamental processes in the evolutionary ecology of
624          Lyme borreliosis. *Nat. Rev. Microbiol.* 4, 660-669.

625    105. Ogden, N.H. et al. (2006) Climate change and the potential for range expansion of the
626          Lyme disease vector *Ixodes scapularis* in Canada. *Int. J. Parasitol.* 36, 63-70.

627    106. Vesk, P.A. et al. (2010) How many hosts? Modelling host breadth from field samples.
628          *Methods Ecol. Evol.* 1, 292-299.

629    106. Dallas, T. et al. (2017) Estimating parasite host range. *Proc. R Soc. B* 284, 20171250.

630    107. Dallas, T. et al. (2017) Predicting cryptic links in host-parasite networks. *PLoS Comp.*
631          *Biol.* 13, e1005557.

632    108. Hellgren, O. et al. (2009) A jack-of-all-trades and still a master of some: Prevalence and
633          host range in avian malaria and related blood parasites. *Ecology* 90, 2840-2849.

634    109. Morgan, E.R. et al. (2007) Parasite transmission in a migratory multiple host system.
635          *Ecol Modell.* 200, 511-520.

636    110. Biek, R. et al. (2012) Whole genome sequencing reveals local transmission patterns of
637          *Mycobacterium bovis* in sympatric cattle and badger populations. *PLOS Pathogens* 8,
638          e1003008.

639    111. Rohde, K. (1980) Host specificity indices of parasites and their application. *Experientia*
640          36, 1369-1371.

641    112. Lozupone, C. et al. (2006) UniFrac - An online tool for comparing microbial community
642          diversity in a phylogenetic context. *BMC Bioinformatics* 7.

643    113. Krasnov, B.R. et al. (2011) Beta-specificity: The turnover of host species in space and
644          another way to measure host specificity. *Int. J. Parasitol.* 41, 33-41.

645    114. Poulin, R. and Mouillot, D. (2005) Combining phylogenetic and ecological information
646          into a new index of host specificity. *J. Parasitol.* 91, 511–514.

647    115. Legendre, P. et al. (2002) A statistical test for host-parasite coevolution. *Syst. Biol.* 51,
648          217-234.

649    116. Gervasi, S.S. et al. (2015) The context of host competence: a role for plasticity in host–
650          parasite dynamics. *Trends Parasitol.* 31, 419-425.

651    117. LoGiudice, K. et al. (2003) The ecology of infectious disease: effects of host diversity
652          and community composition on Lyme disease risk. *Proc. Natl. Acad. Sci. U. S. A.* of
653          the United States of America 100, 567-571.

654

655

656

# Glossary

● **Ecological fitting**: Species association enabled by pre-existing capacity without genetic change. (e.g. a certain suite of potential host traits enables infection by a parasite upon encounter).

● **Emerging infectious disease (EID)**: Infectious disease that recently appeared in a population or is recognized as a 'novel' disease with rapid spread.

● **Environmental filtering**: Environmental conditions that constrain resource utilisation by a species (e.g. the use of different host species by a parasite).

● **Fundamental niche (e.g. fundamental host range)**: For parasites, the set of all host species, whether known to be infected or not, that would serve a parasite as hosts under any environmental condition. The overall fundamental host range cannot necessarily be determined empirically, as only the existing (realized) host range across the accessible host species pool can be surveyed.

● **Host shifting**: The event of colonizing of a novel host species by a parasite, involving host range expansion (here defined as colonization of a novel host without losing the ancestral host opposed to considerations that parasites may shift from one host to another without range expansion).

● **Host specificity**: Measures of the number and/or diversity of host species a parasite is capable of infecting.

● **Niche**: Broadly and indistinctly defined environmental space suitable for a species to survive and reproduce. For parasites, the host range broadly refers to the main component of their niche.

● **Potential resources**: The range of resources that a species could utilise if it encountered them. In terms of host species, this is equivalent to fundamental host specificity, representing the full (unknown) range of hosts species a parasite is capable of infecting.

● **Realized niche (e.g. realized host range)**: A set of host species observed to be infected by a parasite in a specific regional and spatiotemporal context.

● **Spillover**: Cross-species transmission of a parasite into a host population not previously infected. In contrast to 'host shifting', 'spillover' often refers to infection of novel *populations* but not necessarily novel *species* (i.e. no host range expansion). The term appears to be most commonly used to describe cross-species transmission from wildlife to humans.

● **Stockholm Paradigm**: Hypothetical concept arguing that host range expansions by parasites result from the interplay between novel host-parasite opportunities in response to

690 shifting community assembly and phylogenetic and/or ecological barriers that limit parasite

691 spread to novel hosts.

692 • **Transferability (models)**: Whether a model (and its parameter estimates) derived from a

693 particular set of spatiotemporal conditions is transferable to other conditions and can be

694 generalized. In terms of resource utilisation, a transferable model would allow accurate

695 predictions of resource use from a model built using data from elsewhere.

696 • **Transmission**: The transfer of a parasite between different host individuals or other entities

697 (such as relevant vectors). A term typically used in epidemiological studies.

698 • **Zoonosis (zoonotic)**: An infectious disease of humans caused by parasites acquired from an

699 animal reservoir (host individual/ population/ species infected with a parasite and acting as a

700 source for further infection and parasite spillover).

701

702

## Box 1: Drivers of contemporary host specificity

704 Host community composition is influenced by various biotic and environmental filters that

705 can collectively lead to dramatic variation in a parasite's realized host specificity. Here we

706 outline a number of these situations, though it is important to note that this list is by no means

707 exhaustive.

708 • **Anthropogenic invasion of key hosts:** Changes in host community composition

709 caused by anthropogenic invasions can lead to important changes in parasite realized

710 host specificities. For example, the rapid expansion of chytrid fungus, the parasite

711 responsible for threatening declines of many amphibians globally, largely occurred

712 along wildlife trade routes [95]. In addition, invasive commensal rats have been

713 crucial for the global spread of parasitic helminths that have 'hitch-hiked' their way to

714 encountering novel host species [33], ultimately shaping parasite biogeographic

715 distributions [31].

716 • **Changing community compositions leading to new transmission dynamics:** The

717 population structure of a parasite depends on the diversities and relative abundances

718 of different host species infected [17]; host abundance is itself a plastic trait that

719 typically varies across environmental gradients, strongly contributing to plasticity in

720 realized host specificity [96, 97]. Following the examples above, introduced species

721 may not only serve as potential vehicles for introduced parasites, but may also directly

alter existing local host-parasite interactions by changing the abundance of endemic host species [19], hence inducing cascading effects on host-parasite interactions.

- **Habitat encroachment and the human-wildlife interface**: As a consequence of habitat conversion and fragmentation, humans and domestic species are in frequent contact with wildlife species. Cat fleas, intestinal helminths, and canine distemper virus are among the increasing number of parasites observed to expand their host ranges to include a diversity of wild mammals following such novel human-wildlife encounters [7, 98, 99].

- **Expanding dietary range of a key host**: Biotic interactions among key host species play a large role in driving plasticity in host specificity. For example, within its exotic range in Australia, the presence of dingos and dingo/dog hybrids that feed on a large range of endemic wildlife has enabled the establishment of stable transmission cycles of the tapeworm *Echinococcus granulosus* through wild dogs and endemic wildlife [100], illustrating how host shifting into novel communities may be facilitated by particular regional conditions.

- **Exposure of parasites to competitors or facilitators**: Parasites themselves can also exhibit important biotic interactions. Antagonistic and synergistic effects in multi-host, multi-parasite systems affect both the co-occurrence of co-infecting parasites within the same host individuals [53, 101] as well as eco-epidemiological transmission dynamics [102].

- **Climate-related changes in the host affinity of vectors:** Climate-driven changes in the feeding patterns of important vectors may facilitate opportunities for vector-borne parasites to contact novel host species. For instance, warming climate influences the human-feeding habits of rickettsiae-vectoring tick species, leading to human spillover events [103]. Climate in combination with habitat changes can also affect the host range of tick-borne *Borrelia* bacteria, the cause of Lyme disease [104, 105].

## Box 2: A data primer on host shifting and specificity

Any conclusion on fundamental host specificity depends on a sufficiently large number of individuals and potential host species examined for robust inference. At their most basic level, host specificity analyses rely on binary vectors describing the confirmed presence or absence of infection by a particular parasite from a diversity of sampled host species. Ideally,

755 the scope of the study should be narrow enough that the sampled species will all be suitable

756 potential hosts for the selected parasite (occurring within the parasite's geographical

757 distribution and exhibiting some level of possible exposure). Yet even with good foresight

758 and selection of possible hosts, these presence-absence vectors are surprisingly difficult to

759 acquire. For example, if data from multiple studies are compiled to represent a suite of known

760 host-parasite associations, such data are usually strongly susceptible to bias. This is because

761 such databases typically contain presence-only records. For host specificity inferences

762 beyond simple diversity metrics of the observed host species, the absence records

763 (representing species not found to be host after a reasonable sampling effort) are just as

764 important as presence records and should be included where possible. An option to make

765 presence-only data accessible to analysis could be the utilisation of pseudo presence-absence

766 data, in which infected host species are recorded as 'viable' and uninfected species (i.e. those

767 species present in a pool of potential host species but not recorded to be infected) as 'non-

768 viable' hosts.

769      Without detailed information on sampling efforts such as the number of host

770 individuals screened for a parasite, there is uncertainty whether parasites are truly absent

771 from a host species that is reported to be uninfected, challenging the estimate of host range

772 [106, 107]. Such absences can represent 'false zeros' (missing observations of interactions)

773 when small sample sizes and a naturally low parasite prevalence result in limited detection

774 probability. A simple proxy of sampling bias could be a measure of research effort, such as

775 the number of scientific publications linked to a particular potential host species (see e.g.

776 [99]). However, conservative interpretation is warranted as indices of research effort are only

777 coarse proxies of the true underlying sampling bias (i.e. the number of publications does not

778 necessarily reflect the true sampling efforts of how many host individuals have been surveyed

779 for a parasite).

780      Preferably, individual-level data (i.e. detailed data on the number of individuals

781 examined and infected) will be available so that biological patterns and processes can be

782 distinguished from sampling bias arising from unequal and small sample sizes [83, 85, 108].

783 If detailed information from empirical field surveillances (such as the number of infected and

784 uninfected host individuals captured) are available, the prevalence of parasites in different

785 sympatric host species can be estimated. If combined with further information on host species

786 occurrence and density (which are often available or can be estimated from trap or survey

787 data), such estimates provide valuable information on the relative importance of different

788 host species as parasite reservoirs. These relative importances can be used to weight the

789 contributions of different host species to a parasite's realized host specificity, aligning to the

790 concept that host species may have different reservoir capacities. Incorporating measures on

791 host presence and abundance might be of particular relevance if host abundances are subject

792 to strong fluctuations [13, 109] and/or migration that drives the connectivity of

793 geographically disparate host assemblages [48, 110].

794      For parasites with complex life cycles, it can also be important to consider details of

795 parasite life histories. For some helminths, for example, different sets of host species are

796 utilised to complete different parts of the life cycle (i.e. predatory carnivorous species that

797 serve as definite host versus herbivorous species that serve as intermediate hosts). In addition,

798 detailed molecular data can provide valuable insight into whether different sympatric host

799 species share the same strains or populations of a parasite [111], eventually narrowing down

800 the pool of host species relevant for analysis.

801

802

803

804

805

806

807

808

809

810

811

812

813

814

815

816

817     **Table 1**. Overview of selected components and methods to define host specificity.

| Method | Methodological approach | Example/Reference |
|---|---|---|
| Number of host species infected | Count of the number of infected host species. This basic count ignores host species attributes such as phylogenetic or ecological relationships. | [112] |
| Diversity indices capturing variation in host community composition (e.g. Shannon-Wiener, phylogenetic diversity, UniFrac) | Diversity measures based on the abundance and/or attributes attributed to the range of observed host species (i.e. phylogenetic diversity). | [5], [113] |
| Geographic specificity and host range turnover (β-diversity) | Measures of the dissimilarity of a parasite's host ranges in different regions, resembling β-diversity measures. | [114], [17] |
| Distance-based phylogenetic and/or functional specificity | Measures of the distances between pairs of host species in terms of phylogenetic or functional relationships. Distance measures can be weighted by prevalence to give greater weight to commonly infected host species. | [115], [15] |
| Network indices of specialization (d') | Calculated from bipartite host-parasite interaction networks, these indices measure of a parasite's interactions with a range of | [76] |

| | potential host species (i.e. the sampled host species pool), weighted relative to the host interactions displayed by other parasites in the community. | |
|---|---|---|
| Degree of matching between host and parasite phylogenies | Measures of the matching between host and parasite phylogenies, used for depicting community-level patterns of possible co-evolution. Specificity inferences are drawn based on how tightly parasite evolution is linked to host evolution. | [116] |
| Host competence heterogeneity | The spread of parasites through host assemblages can be largely determined by their variation in potential hosts' competence and reservoir potential. Such measures have been rarely used to measure host specificity to date, but could be especially useful to express plasticity in host specificity. | [117], [118] |

818

819

820

821

822

823
824
825
826
827
828
829



830
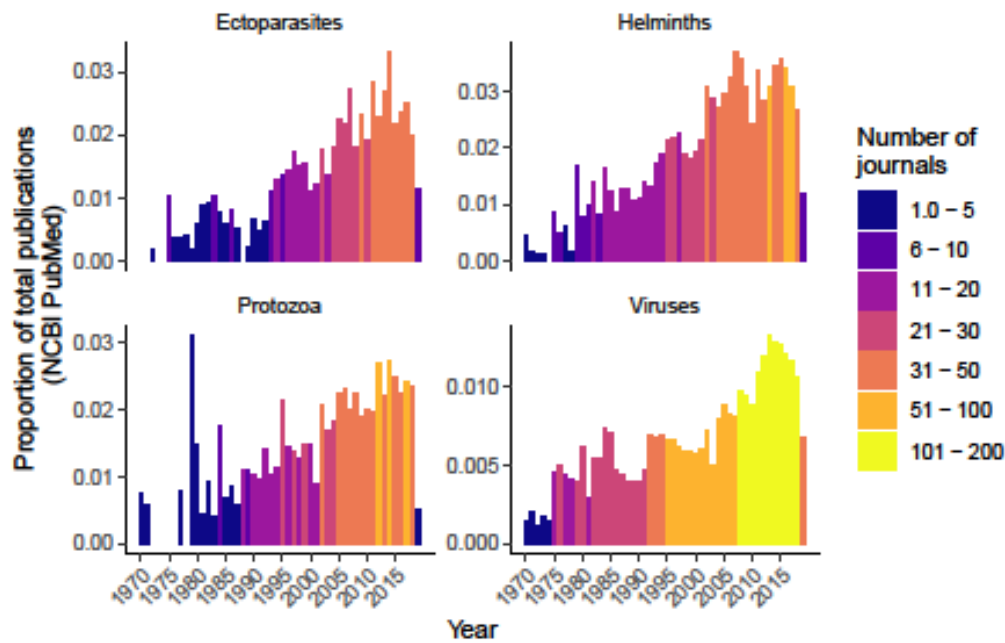
**Figure 1**. Rise in the proportion of peer-reviewed research articles mentioning 'host specificity' for select groups of parasites over time. Colours of bars represent the total number of unique peer-reviewed journals mentioning 'host specificity' for each parasite group in each year. Articles were accessed by searching the NCBI PubMed database on 2$^{nd}$ April 2019.

836
837
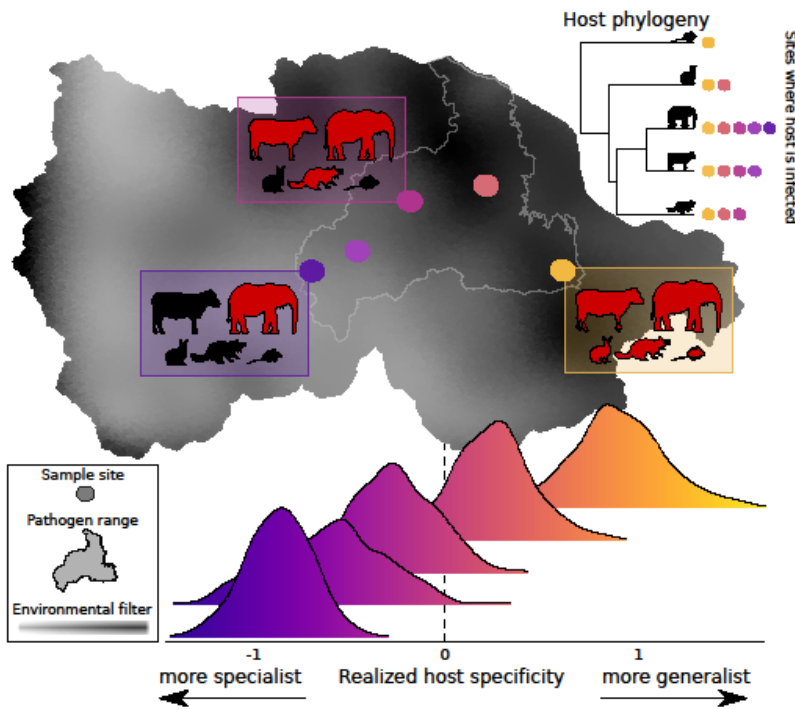
**Figure 2**. Illustration of how environmental filters across a parasite's geographical range may impact the parasite's realized host specificity. Hosts are sampled for the parasite across an environmental gradient that influences opportunity for the parasite to interact with each potential host species. Density plots depict the probabilities that observed host pairwise phylogenetic distances differ from those expected (representing the entire host pool that occurs at a given location) for each site. More negative measures (darker purple tones) indicate a parasite infects hosts that are more closely related than expected, indicating specialism; more positive measures (warmer yellow tones) indicate generalism.