

MOLECULAR ECOLOGY**Integrating phylogenetic and ecological distances reveals
new insights into parasite host specificity**

Journal:	<i>Molecular Ecology</i>
Manuscript ID	MEC-17-0010.R1
Manuscript Type:	Original Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Clark, Nicholas; University of Queensland, School of Veterinary Sciences; Griffith University School of Environment, Clegg, Sonya; University of Oxford, Department of Zoology
Keywords:	avian malaria, community assembly, habitat niche, host specificity, phylogenetic dispersion, phylogenetic β diversity

Integrating phylogenetic and ecological distances reveals new insights into parasite host specificity

*Nicholas J Clark^{1,2}, Sonya M Clegg³

¹ *School of Veterinary Science, The University of Queensland, Gatton, Queensland 4343*

² *Environmental Futures Research Institute, Griffith University, QLD 4111, Australia*

³ *Edward Grey Institute of Field Ornithology, Department of Zoology, University of Oxford, Oxford, OX1 3PS, UK*

*corresponding author email: nicholas.j.clark1214@gmail.com

Running title: Host specificity in avian blood parasites

Key words: avian malaria, community assembly, habitat niche, host specificity, phylogenetic dispersion, phylogenetic β diversity

ABSTRACT

The range of hosts a pathogen infects (host specificity) is a key element of disease risk that may be influenced by both shared phylogenetic history and shared ecological attributes of prospective hosts. Phyllospecificity indices quantify host specificity in terms of host relatedness, but can fail to capture ecological attributes that increase susceptibility. For instance, similarity in habitat niche may expose phylogenetically unrelated host species to similar pathogen assemblages. Using a recently proposed method that integrates multiple distances, we develop a new index to assess the relative contributions of host phylogenetic and functional distances to pathogen host specificity (functional-phylogenetic host specificity). We apply this index to a dataset of avian malaria parasite (*Plasmodium* and *Haemoproteus* spp.) infections from Melanesian birds to show that multi-host parasites generally use hosts that are closely related, not hosts with similar habitat niches. We also show that host community phylogenetic β diversity ($P\beta d$) predicts parasite $P\beta d$, and that individual host species carry phylogenetically clustered *Haemoproteus* parasite assemblages. Our findings were robust to phylogenetic uncertainty, and suggest that phylogenetic ancestry of both hosts and parasites play important roles in driving avian malaria host specificity and community assembly. However, restricting host specificity analyses to either recent or historical timescales identified notable exceptions, including a 'habitat specialist' parasite that infects a diversity of unrelated host species with similar habitat niches. This work highlights that integrating ecological and phylogenetic distances provides a powerful approach to better understand drivers of pathogen host specificity and community assembly.

26
27

INTRODUCTION

28 How pathogens spread to new hosts and assemble into communities are key
29 questions in ecology (Krasnov *et al.* 2016; Woolhouse *et al.* 2001). Pervasive features
30 of host-pathogen systems are the ability of some pathogens to infect multiple host
31 species (i.e. pathogens with broad host specificity) and the ability of some host
32 species to carry diverse pathogen assemblages (Braga *et al.* 2015; Johnson &
33 Hoverman 2014; Poulin *et al.* 2006). The degree to which both aspects are exhibited
34 has important ramifications, from regulating pathogen dispersal and range
35 expansions to shaping transmission dynamics (Cooper *et al.* 2012; Woolhouse *et al.*
36 2001). However, identifying factors that influence host specificity and the pathogen
37 assemblages carried by hosts remains challenging (Gervasi *et al.* 2015).

38 Infection by the same pathogen is often considered more likely among
39 related hosts due to shared physiological and/or behavioural traits (Schulze-Lefert &
40 Panstruga 2011; Streicker *et al.* 2010; Wells *et al.* 2015). Accordingly, host specificity
41 is widely described in terms of host phylogenetic distances (phylospecificity; Braga *et*
42 *al.* 2015; Huang *et al.* 2014). However, many multi-host pathogens infect unrelated
43 but ecologically similar hosts (Farrell *et al.* 2013; Gómez *et al.* 2013; Lootvoet *et al.*
44 2013). Prominent examples are food-borne bacteria in hosts that feed on similar
45 resources (Woolhouse *et al.* 2001) and bat-borne viruses spilling into co-occurring
46 domestic animals (Plowright *et al.* 2015). Phylospecificity indices can therefore fail to
47 capture key host niche attributes that facilitate infection, particularly those with low
48 phylogenetic signal (Krasnov *et al.* 2010).

49 Accounting for both phylogeny and ecology has parallels in community

ecology, where assembly patterns have phylogenetic and ecological influences (Silvertown *et al.* 2001). For instance, exploring the relative influences of host phylogenetic and ecological turnover on pathogen turnover can uncover key drivers of pathogen assembly (Krasnov *et al.* 2010; 2015). Yet determining how host phylogeny and ecology influence pathogen host specificity is not straightforward (Poulin & Mouillot 2005). Recently, Cadotte *et al.* (2013) proposed a method to disentangle community assembly patterns by nonlinearly integrating functional and phylogenetic distances (functional-phylogenetic distance; *FPDist*). Crucially, phylogenetic and functional weights can be altered to investigate their relative effects on assembly patterns (Cadotte *et al.* 2013). Calculating *FPDist* host specificity at varying weights thus presents a new way to quantify the relative influences of host ecology and phylogeny on pathogen host specificity.

Host-pathogen interactions result from processes that have occurred over both historical and recent evolutionary timescales (Ricklefs *et al.* 2004). However another restraint of phyllospecificity indices is that they only consider mean pairwise host distance (MPD), which is restricted to detecting relatively ancient processes such as ancestral habitat filtering (Kraft *et al.* 2007; Stegen *et al.* 2012). To capture more recently developed patterns, such as sister host species with different infection susceptibilities (Perlman & Jaenike 2003), nearest taxon distance (NTD) is more appropriate (Weiblen *et al.* 2006). This measure only considers the distance between the most closely related taxa (i.e. minimum pairwise distance), and can be used in complement with MPD to explore complex evolutionary patterns (Weiblen *et al.* 2006; Fig. 1). For example, pathogens that infect phylogenetically clustered hosts (i.e. more closely related than expected by chance) at ancient but not recent

timescales may reveal hosts that have escaped from pathogens carried by close relatives (Fig. 1b). This could provide evidence for ‘enemy release’, a mechanism thought to contribute to the success of invasive species (Adlard *et al.* 2015; Clark *et al.* 2015a; Keane & Crawley 2002). In contrast, hosts that are phylogenetically clustered at recent but not ancient timescales may illustrate independent evolution in multiple lineages of traits that increase susceptibility, which are then shared in sister species (Fig. 1c). For example, infection susceptibility to parasitic trematodes and associated pathology are both more similar for amphibian species that have similar developmental rates (Johnson *et al.* 2012).

The dispersion comparisons described above can also be used to explore relationships of pathogen assemblages carried by a particular host species. Phylogenetic variance among a host’s pathogens provides critical insights into a host’s infection susceptibility as well it’s influence on pathogen community assembly (Johnson & Hoverman 2014). Moreover, intra-host assessments can uncover potential evolutionary restrictions to pathogen host shifting. Closely related pathogens are commonly expected to have similar host affinities, as they may have similar abilities to navigate host immune systems (De Vienne *et al.* 2009). However, while it is widely accepted that pathogen diversity and infection rates are uneven among infected host species (Shaw & Dobson 1995), whether multi-pathogen hosts carry particular phylogenetic pathogen assemblages is less often explored (Cooper *et al.* 2012).

Vector-borne parasites offer an exciting opportunity to study host specificity, as vectors can have specific habitat requirements while exhibiting broad feeding preferences that promote host range expansions (Gager *et al.* 2008; Medeiros *et al.*

198 2013; Njabo *et al.* 2011). Avian malaria parasites (genera *Plasmodium* and
199 *Haemoproteus*) are vector-borne blood parasites that infect birds around the globe
100 (Clark *et al.* 2014, 2016a). Because avian malaria parasites are transmitted by vectors
101 which may have specific habitat preferences (Santiago-Alarcon *et al.* 2012), it stands
102 to reason that similarity in host habitat niche may lead to similar exposure to
103 pathogen assemblages. For example, a wide diversity of bat species are susceptible
104 to white nose syndrome, yet infection is often restricted to species that use similar
105 hibernation sites that are more favourable to the pathogen (Zukal *et al.* 2014).
106 However, while host specificity is commonly quantified in avian malaria research,
107 inferences have to date relied on specificity indices that ignore host ecological
108 attributes (Ewen *et al.* 2012; Moens & Pérez-Tris 2015; Olsson-Pons *et al.* 2015; but
109 see Svensson-Coelho *et al.* 2016). Moreover, little is known about the phylogenetic
110 dispersion of avian malaria parasites carried by individual host species, despite the
111 fact that birds are commonly thought to acquire new parasites and leave others
112 behind as their ranges expand and contract (Lauron *et al.* 2014; Ricklefs *et al.* 2004).

113 Using avian malaria infection data from southern Melanesian birds, we apply
114 community and species level analyses to assess the relative effects of host phylogeny
115 and habitat niche on parasite assembly and host specificity. Analyses are evaluated
116 using MPD and NTD timescales to explore both historical and more recent
117 relationships among infected hosts. If host habitat niche is more important than
118 phylogeny in driving parasite host specificity, we expect community turnover in host
119 habitat niche dissimilarity ($E\beta d$), not host phylogenetic turnover ($P\beta d$), to predict
120 parasite $P\beta d$ across both timescales. We also expect parasites to infect hosts whose
121 recent timescale $FPDist$ distances become clustered as habitat niche weight

122 increases (recent habitat convergence promotes cross-species infection), but whose
 123 historical distances either show no pattern or become overdispersed as phylogenetic
 124 weight increases (parasites are not restricted to hosts from the same clade; Fig.
 125 1c,d). If host phylogeny is important, host $P\beta d$ will predict parasite $P\beta d$, host $FPDist$
 126 distances will either show no pattern or become overdispersed as habitat niche
 127 weight increases (similarity in host habitat niche does not predict host specificity),
 128 and historical host distances will become clustered as phylogenetic weight increases
 129 (parasites are restricted to particular host clades; Fig 1a,b). Depending on the
 130 evolutionary forces at work, we may expect recent timescale host distances to show
 131 one of two patterns as phylogenetic weight increases. If co-divergence or the
 132 exploitation of hosts with highly conserved traits has occurred, recent timescale
 133 distances will become clustered (Fig. 1a), whereas if recent host escape has
 134 occurred, recent timescale distances will become overdispersed (Fig. 1b).

135 We next explore phylogenetic relationships of parasites carried by individual
 136 host species over a spectrum of parasite evolutionary timescales. Ideally, one would
 137 include both ecological and phylogenetic parasite distances to study assemblage
 138 dispersion. However, assessing parasite traits is difficult and subjective (Mouillot *et*
 139 *al.* 2005). We instead explore a spectrum of evolutionary timescales for an in-depth
 140 perspective of parasite dispersion patterns. If avian host species acquire and spread
 141 diverse parasite assemblages, we expect parasite assemblages carried by individual
 142 host species to become more phylogenetically overdispersed at historical timescales.
 143 However if closely related parasites have similar host affinities due to similar abilities
 144 to navigate host immune systems, we would expect individual host species to carry
 145 phylogenetically clustered parasite assemblages at recent timescales.

146

147 **MATERIALS AND METHODS**

148 Avian malaria infection data was gathered from 1755 birds (belonging to 41 species
149 including 39 native and 2 introduced species) sampled on 17 islands in southern
150 Melanesia (Clark *et al.* 2016b; Olsson-Pons *et al.* 2015). Birds were captured with
151 mistnets and blood samples taken to detect malaria parasites (see Olsson-Pons *et al.*
152 2015 for additional sampling details). Avian species varied in their distributional
153 extents, with the number of islands occupied by each sampled species ranging from
154 one to 17 (mean = 9.26, sd = 6.55) and the number of species captured on each
155 island ranging from two to 20 (mean = 9.53, sd = 5.03). Using molecular (PCR)
156 parasite screening (described in Clark *et al.* 2016b), 34 *Haemoproteus* cytochrome-*b*
157 (cyt-*b*) lineages (324 total infections across 20 host species) and 19 *Plasmodium*
158 lineages (139 infections across 14 host species; Fig. S1) were identified. A subset of
159 samples were also subjected to blood smear examination to reduce false positives
160 and identify co-infections (see Clark *et al.* 2016b for details of smear screening). Co-
161 infections accounted for 15.7% of total infections. We refer to each parasite lineage
162 (differing by at least 1 base pair) as a unique parasite, as evidence suggests that
163 lineages differing by as little as one base pair may represent reproductively isolated
164 entities (Bensch *et al.* 2004). The number of parasites found on each island ranged
165 from three to 23 (mean = 8.59, sd = 4.56), while the number of parasites recorded in
166 each individual host species ranged from one to 20 (mean = 3.49, sd = 4.11).

167

168 **Parasite and host molecular phylogenies**

169 Bayesian parasite phylogenies were built using a cyt-*b* fragment (see Supplementary

170 Data for details of phylogenetic reconstruction). From a posterior distribution of
171 32000 parasite trees, we randomly sampled 1000 trees (without replacement) to
172 account for phylogenetic uncertainty. Avian host phylogenies were gathered from
173 the Birdtree.org database (<http://birdtree.org>), which contains a Bayesian posterior
174 distribution of trees generated for 9,993 avian species (Jetz *et al.* 2012). We sampled
175 100 trees from the ‘Ericsson All Species Trees’ dataset for our sampled avian hosts.

176

177 **Host habitat niche dendrograms**

178 Dendrograms were built to describe similarities between host habitat niches. We
179 applied hierarchical clustering to a habitat dissimilarity Gower’s distance matrix
180 (Gower 1971) incorporating the following host attributes:

181 *Habitat breadth* is a multiplicative β diversity metric describing the range of
182 habitats occupied by a species while considering the compositional diversity of co-
183 occurring species (Ducatez *et al.* 2014). This index quantifies the target species’
184 degree of specialization, where a ‘generalist’ occurs in a range of habitats that differ
185 in species composition while a ‘specialist’ uses habitats that contain a consistent
186 collection of other species (Ducatez *et al.* 2014). Thus, a host’s habitat breadth could
187 drive parasite host specificity by determining the diversity of macrohabitats and co-
188 occurring avian species that the host encounters. *Habitat breadth* data were
189 gathered from Ducatez *et al.* (2014) using the within-class index, which only
190 considers the diversity of co-occurring bird species (as opposed to all terrestrial
191 vertebrates). *Habitat breadth* was as a continuous variable, scaled by standard
192 deviation.

193 *Island occupancy*, defined as the presence or absence of each species on each

194 island, was compiled from avian distribution data in Dutson (2012). A host's island
195 occupancy could influence the parasites it is exposed to, with host species that share
196 similar distributions more likely to be exposed to similar parasite assemblages. We
197 included data for 17 sampled islands, four in New Caledonia and 13 in Vanuatu.
198 *Island occupancy* for each species on each island was included as a binomial variable.

199 *Foraging strata preference* represents the proportion of time that species
200 forage in each of five forest strata categories: *ground*, *understory*, *mid-high*, *canopy*
201 and *aerial*. While island occupancy and habitat breadth describe species'
202 distributions and macrohabitat use, foraging strata describes a species' microhabitat
203 use. Avian malaria vector spp. are known to vary in distribution and abundance
204 across forest microhabitats (Černý *et al.* 2011). Therefore, this variable may
205 influence parasite exposure by capturing whether or not certain host species avoid
206 blood-seeking vectors by using particular forest strata. Foraging data was gathered
207 from the *EltonTraits* database of vertebrate ecological traits (Wilman *et al.* 2014).
208 Because species can use multiple foraging strata, *Foraging strata preference* was
209 included as a fuzzy proportional variable to describe the affinity of each species for
210 each strata category (Pavoine *et al.* 2009). Values for each category were scaled by
211 the category maximum.

212 We weighted habitat attributes by the inverse of their phylogenetic
213 autocorrelations to capture variance in host habitat niche that is not already
214 captured by phylogenetic relationships. For quantitative traits (*habitat breadth*,
215 *foraging strata preference*), we used Abouheif's *C*, which is efficient at detecting
216 phylogenetic autocorrelations regardless of tree shape (Abouheif 1999; Pavoine &
217 Ricotta 2013). *Island occupancy* values were given equal weights, using the median

218 of the quantitative trait weights. The distance matrix was built following Pavoine *et*
219 *al.* (2009). Because different clustering algorithms can lead to different conclusions
220 (Kulbicki *et al.* 2013), we generated six habitat niche dendrograms to capture
221 uncertainty. Each dendrogram was built using a different clustering algorithm:
222 (*ward.D2*, *complete*, *average*, *mcquitty*, *median* and *centroid*).

223

224 **Comparing host and parasite community phylogenetic β diversity**

225 To test if parasite $P\beta d$ was predicted by host $P\beta d$ or host $E\beta d$, we generated 10
226 equal-sized communities of infected hosts (randomly sampling infected hosts from
227 the data with replacement). Host and parasite phylogenetic pairwise distance
228 matrices were calculated using community MPD with function '*cd.query*' in the
229 '*phyloMeasures*' package (Tsirogianis & Sandel 2015). For parasites, this resulted in
230 a 10 X 10 matrix of pairwise $P\beta d$ values. For hosts, two separate matrices were
231 created using phylogenies and habitat niche dendrograms (pairwise $P\beta d$ and
232 pairwise $E\beta d$). We repeated the process by calculating community NTD distances
233 using function '*cdnt.averaged.query*'. For each distance measure, we partitioned
234 effects of host $P\beta d$ and $E\beta d$ on parasite $P\beta d$ with a multiple regression on distance
235 matrices, using functions in the package '*ecodist*' (Goslee & Urban 2007; Lichstein
236 2007).

237 A distribution of regression coefficients was drawn from 10,000 iterations. To
238 account for potential bias where small sample sizes (number of hosts in each
239 community sample) could be composed of few abundant host species, we varied
240 sample size across a uniform distribution from min = 20 to max = 1/3 total infected
241 hosts (i.e. max of 47 infected hosts for *Plasmodium* and 108 for *Haemoproteus*

242 samples; Table S1). We considered predictors to explain a significant proportion of
243 parasite $P\beta d$ variance if coefficient 95% credible intervals (CI) did not include zero.

244

245 **Parasite functional-phylogenetic host specificity**

246 Host pairwise $FPDist$ was calculated using phylogenetic and habitat niche distances.

247 Phylogeny and dendrogram branch lengths were scaled (dividing distances by the

248 maximum distance) so that pairwise distances ranged from zero to one. Pairwise

249 phylogenetic and habitat niche distances ($PDist$ and $FDist$, respectively, to follow

250 terminology from Cadotte *et al.* 2013) were used to calculate $FPDist$:

251

$$252 \quad FPDist = (\alpha PDist^p + (1 - \alpha) FDist^p)^{1/p} \quad (\text{eqn 1})$$

253

254 The weight parameter α can vary from zero to one, with values approaching one

255 giving greater weight to $PDist$ and values approaching zero giving greater weight to

256 $FDist$. We set $p = 2$ to calculate squared Euclidean distances.

257 For parasites infecting more than one host species, host $FPDist$ distances

258 were used to calculate a phyllospecificity index (STD^*), following methods in Poulin &

259 Mouillot (2005). We also calculated a nearest taxon STD^* index using only the

260 minimum pairwise host $FPDist$ ($NTSTD^*$). In both indices, $FPDist$ distances were

261 weighted by prevalence to give greater weight to distances between commonly

262 infected hosts and less weight to rare host-parasite interactions (Poulin & Mouillot

263 2005).

264 Null host distributions were created for each parasite by randomly sampling

265 the observed number of infected host species from the regional host pool and

calculating expected STD^* and $NTSTD^*$. We allowed α to vary across a uniform distribution from zero to one to alter relative weights of host phylogeny and habitat niche in each iteration. For a parasite infecting four host species, our approach would entail: (1) integrating host phylogenetic and habitat niche distances into $FPDist$ at the specified α weight; (2) gathering observed pairwise host $FPDist$ distances, using the full distance distribution to calculate STD^* and the minimum distance to calculate $NTSTD^*$; (3) randomly sampling four species from the 41 available host species and calculating expected STD^* and $NTSTD^*$. Expected values were subtracted from observed to yield specificity differentials, which will be negative if hosts are more similar than expected (clustered) and positive if hosts are less similar (overdispersed). This process was repeated 10,000 times to generate distributions of STD^* and $NTSTD^*$ differentials. For parasites observed to infect two host species, only one pairwise distance can be calculated, so STD^* and $NTSTD^*$ will be identical. We thus calculated observed and expected STD^* for all parasites infecting more than one host species, while $NTSTD^*$ was calculated for parasites infecting more than two host species. We conservatively considered STD^* and $NTSTD^*$ differentials significant if 5% and 95% quantiles of the 10,000 iterations did not include zero.

284

285 **Phylogenetic dispersion of parasite assemblages carried by individual host species**

For host species carrying more than one parasite, we explored phylogenetic dispersion of observed parasite assemblages. Following the process outlined above, we randomly sampled the observed number of parasites from the regional pool (34 *Haemoproteus* and 19 *Plasmodium* parasites). We calculated observed and expected

290 parasite phylogenetic distances at seven distance quantiles, resulting in a spectrum
291 of evolutionary intervals ranging from NTD (0% quantile; minimum pairwise
292 distance) to MPD (100% quantile; mean of all pairwise distances). Expected quantile
293 means were subtracted from observed, resulting in dispersion differentials at each
294 quantile. Dispersion differentials were generated for each host species using 10,000
295 iterations and significance assessed as above.

296 Analyses were performed in R (R Development Team 2008). We accounted
297 for uncertainty in evolutionary (host and parasite phylogenies) and habitat niche
298 relationships (host dendrograms) by randomly sampling trees (with replacement) in
299 each iteration. Each parasite genus (*Haemoproteus* and *Plasmodium*) was analysed
300 separately. Host-parasite data, host niche trait data and analysis R code are available
301 from the Dryad Digital Repository: <http://dx.doi.org/10.5061/dryad.50603>.

303 RESULTS

304 Host and parasite phylogenetic community β diversity

305 Host $E\beta d$ did not predict parasite $P\beta d$ for either parasite genus (regression
306 coefficients overlapped with zero at NTD and MPD; Table S1). Instead, host $P\beta d$
307 explained a significant proportion of variation in parasite $P\beta d$, regardless of whether
308 we used community NTD (coefficient 95%CI: 0.02 – 0.04 for *Haemoproteus*; 0.03 –
309 0.05 for *Plasmodium*) or MPD (0.01 – 0.02 for *Haemoproteus*; 0.02 – 0.04 for
310 *Plasmodium*; Fig. 2). The strength of correlations differed between parasite genera,
311 with host $P\beta d$ explaining a greater proportion of variance in *Haemoproteus* $P\beta d$
312 (median $R^2 = 0.16$ at NTD; 0.09 at MPD) compared to *Plasmodium* $P\beta d$ (median $R^2 =$
313 0.05 at NTD; 0.04 at MPD; Fig. 2).

314

315 **Parasite functional-phylogenetic host specificity**316 Three patterns were evident from *STD** differentials for 19 multi-host parasites (10317 *Haemoproteus*; nine *Plasmodium*; Table S2). First, two *Haemoproteus* parasites318 infected hosts that became clustered as ecological weight increased (decreasing *a*)

319 but did not differ from expected or became overdispersed as phylogenetic weight

320 increased (Fig. 3a), suggesting specificity was driven by habitat niche similarity and

321 not by host relatedness. Second, eight parasites (five *Haemoproteus*; three322 *Plasmodium*) were restricted by host phylogeny but not host habitat niche, with

323 differentials either resembling null expectations or becoming overdispersed as

324 ecological weight increased, but becoming clustered as phylogenetic weight

325 increased (Fig. 3a). Finally, four parasites (two *Haemoproteus*; two *Plasmodium*)326 showed clustered differentials at the majority of *a* weights (Fig. 3a), indicating hosts

327 were both more closely related and more ecologically similar than expected.

328 Differentials for the remaining parasites (one *Haemoproteus*; four *Plasmodium*) did329 not differ from null expectations at any *a* weight (Fig. 3a).

330 Among 10 parasites infecting more than two host species (five

331 *Haemoproteus*; five *Plasmodium*; Table S2), *NTSTD** and *STD** indices decreased as332 phylogenetic weight increased for all apart from one (*Haemoproteus* ZOSFLA04),

333 suggesting infected hosts were more phylogenetically than ecologically similar

334 (indices for four parasites, representing the spectrum of patterns observed, are

335 shown in Fig. 3b). Differentials for both *NTSTD** and *STD** did not differ from

336 expected as ecological weight increased, again for all parasites apart from ZOSFLA04

337 (Fig. S2). For ZOSFLA04, *NTSTD** patterns were similar (indices decreased and

338 differentials became clustered as phylogenetic weight increased), but *STD** patterns
339 were opposite to the remaining nine parasites (indices increased and differentials
340 became overdispersed as phylogenetic weight increased; Fig. 3b, Fig. S2).

341 Comparing *NTSTD** and *STD** using only host phylogenetic distances ($\alpha = 1$)
342 identified patterns similar to two of the four evolutionary dispersion scenarios
343 outlined in (Fig. 1). First, five of 10 parasites infected clustered hosts at both *NTSTD**
344 and *STD** (Fig. S2), consistent with patterns expected from coevolution or the
345 exploitation of hosts with highly conserved traits (Fig. 1a). Second, parasite
346 ZOSFLA04 showed clustered *NTSTD** and overdispersed *STD** differentials (Fig. S2),
347 perhaps indicating convergence of host habitat preferences that have increased
348 infection susceptibility (Fig. 1c). In addition, three *Plasmodium* parasites infected
349 hosts that were clustered at *STD** but did not differ from chance at *NTSTD** (Fig. S2),
350 somewhat resembling patterns expected under a scenario of recent host escape (Fig
351 1b).

352

353 **Phylogenetic dispersion of parasite assemblages**

354 All five host species carrying more than two *Haemoproteus* parasites were infected
355 with more closely related parasites than expected by chance at recent timescales
356 (clustered from 0% to 25% distance quantiles), though it should be noted that four of
357 these species belong to the genus *Zosterops* (Fig. 4; Fig S1). In contrast, none of the
358 five tested host species carried phylogenetically clustered *Plasmodium* assemblages
359 at recent timescales, while two host species carried overdispersed *Plasmodium*
360 assemblages (*Lichmera incana* and *Pachycephala caledonica*; Fig. 4). When assessing
361 ancient timescales (100% distance quantile; MPD), the majority of host species

362 carried parasite assemblages that either did not differ from expected (6 of 15 host
363 species for *Haemoproteus*; 7 of 9 species for *Plasmodium*) or were phylogenetically
364 clustered (7 of 15 for *Haemoproteus*; 2 of 9 for *Plasmodium*; Fig. 4; Fig. S3).

365

366 **DISCUSSION**

367 We show that avian malaria community turnover and host specificity are primarily
368 driven by the phylogenetic relatedness of hosts, not by similarity in host habitat
369 niche. Parasite communities were largely composed of phylogenetic host ‘specialists’
370 that infected related hosts regardless of their habitat associations, suggesting that
371 host phylogenetic composition is a key driver of parasite community assembly.
372 Moreover, host species tended to carry phylogenetically clustered parasite
373 assemblages, showing that related parasites are likely to infect the same host
374 species and that birds may not acquire diverse parasites through time. However,
375 assessing functional-phylogenetic host specificity over multiple timescales allowed
376 us to identify notable exceptions, including parasites that infect ecologically similar
377 but historically unrelated hosts.

378

379 **Host relatedness shapes parasite host specificity and phylogenetic β diversity**

380 Avian malaria parasites tended to infect related hosts regardless of whether those
381 hosts have similar habitat niches. This result rejects the idea that multi-host
382 parasites infect unrelated hosts with similar habitat niches, but is consistent with the
383 finding from a range of host-pathogen systems that related hosts carry similar
384 pathogens (Griekspoor *et al.* 2013; Longdon *et al.* 2011; Clark *et al.* 2015b; Parker *et*
385 *al.* 2015). Interestingly though, some avian malaria studies show differing patterns.

386 These include a lack of congruence between host and parasite phylogenies (Lauron
387 *et al.* 2014; Ricklefs *et al.* 2004), and a finding that sister parasite lineages often
388 infect unrelated hosts (Ricklefs *et al.* 2014). One possible explanation for the
389 discrepancy between these studies and ours is that Melanesia does not experience
390 influxes of migratory forest birds that can transport parasites and perhaps lead to
391 greater host contacts (Dutson 2012), a notable feature in American and African study
392 sites of much previous avian malaria research (Ellis *et al.* 2015; Hellgren *et al.* 2009;
393 Lauron *et al.* 2014). Yet our results are not entirely at odds with the above-
394 mentioned studies, as seven of 19 multi-host parasites infected hosts whose
395 phylogenetic relationships either did not differ from chance or were historically
396 overdispersed. This variation presses the need for a clearer understanding of
397 plasticity in the evolution and expression of host specificity (Little *et al.* 2006), an
398 objective that functional-phylogenetic analyses can help achieve.

399 Our β diversity simulations further indicate that host phylogeny shapes
400 parasite community assembly. This finding has implications at the community level
401 by implying that phylogenetically diverse host assemblages can increase
402 heterogeneity in parasite compositions. Understanding how parasite biodiversity is
403 shaped at multiple scales is a key goal in ecology (Clark *et al.* 2014; Nunn *et al.* 2005),
404 and our evidence provides support for the host diversity begets parasite diversity
405 paradigm (Hechinger & Lafferty 2005; Johnson *et al.* 2016). However, differences
406 were evident between parasite genera, as the relationship between host and
407 parasite phylogenetic turnover was stronger for *Haemoproteus* than for *Plasmodium*
408 parasites, perhaps indicating that *Haemoproteus* assemblages are more restricted by
409 host evolutionary history. In addition to host phylogenetic relationships, our

410 *Plasmodium* results suggest that other factors may drive assembly, perhaps including
 411 environmental variables that restrict vector distributions (e.g. temperature or
 412 rainfall; Sehgal 2015; Sehgal *et al.* 2010; Clark & Clegg 2015). An important
 413 consideration is that we used simulated communities, and so were unable to
 414 account for climatic differences that would be evident across spatially discrete
 415 communities (Ellis *et al.* 2015). Nevertheless, previous work in Melanesia supports
 416 the idea of climate-driven *Plasmodium* communities by showing that community
 417 composition is influenced by both local island effects and by geographic isolation
 418 (Olsson-Pons *et al.* 2015).

420 **Importance of using multiple timescales to explore parasite host specificity**

421 Assessing host specificity at multiple evolutionary timescales uncovered patterns
 422 that would have been missed by common phyllospecificity indices. By infecting
 423 phylogenetically clustered hosts at both *NTSTD** and *STD**, half of the tested multi-
 424 host parasites showed patterns consistent with host-parasite coevolution or the
 425 exploitation of hosts with conserved traits (Brooks *et al.* 2015). Three parasites
 426 infected hosts that were also phylogenetically clustered at *STD** but not at *NTSTD**,
 427 perhaps indicating susceptible hosts that have dispersed to areas outside the
 428 transmission zone (Clark *et al.* 2015a; Heger & Jeschke 2014). For instance,
 429 *Plasmodium* parasites GRW04 and ORW1 infected two widespread *Zosterops* spp.
 430 across two islands in New Caledonia (*Z. lateralis* and *Z. xanthochrous*), but were not
 431 found on the New Caledonian island of Lifou where two local endemic *Zosterops* spp.
 432 occur (*Z. minutus* and *Z. inornatus*). Moreover, neither parasite was observed in
 433 Vanuatu, where *Z. flavifrons* carried a diversity of parasites across nearly all islands.

434 It may therefore be plausible that uninfected *Zosterops* spp. are susceptible, but
435 have avoided infections by occurring in regions where the parasites have not
436 colonised. While recent timescale processes differed, the patterns illustrated above
437 support the hypothesis that historical host relationships help shape parasite host
438 ranges (Cooper *et al.* 2012; Longdon *et al.* 2011).

439 The *Haemoproteus* parasite ZOSFLA04 showed a fascinating trend by
440 infecting phylogenetic sister hosts that share similar habitat niches, but are more
441 distantly related than expected over ancient timescales. This parasite was the only
442 one of 10 tested parasites to show this pattern, indicating that the strategy to infect
443 unrelated but ecologically similar host species is rare in Melanesian avian malaria
444 parasites. Relying only on MPD-based phylospecificity, ZOSFLA04 would likely be
445 classified a 'host generalist' that infects numerous unrelated hosts (Clark & Clegg
446 2015; Moens & Pérez-Tris 2015). Instead, using multiple timescales to assess
447 phylogenetic (clustered *NTSD**, overdispersed *STD**) and ecological dispersion
448 (clustered *STD**) indicated that perhaps this parasite is a 'habitat specialist' capable
449 of infecting historically unrelated hosts that use similar habitats (Zukal *et al.* 2014).
450 This could be facilitated by a filtering mechanism where broad host contacts in a
451 habitat-restricted vector promote wide host ranges (Altizer *et al.* 1998; McCoy *et al.*
452 2013; Medeiros *et al.* 2013), or perhaps by novel parasite adaptations that expand
453 the realised niche (Little *et al.* 2006). The strategy to infect unrelated but ecologically
454 similar hosts is widespread in nature (Gervasi *et al.* 2015; Plowright *et al.* 2015;
455 Woolhouse *et al.* 2001), yet uncovering why this is rare in avian malaria requires
456 further study.

457

458 **Host species carry phylogenetically clustered parasite assemblages**

459 Why host species did not typically carry clustered *Plasmodium* assemblages is an
 460 intriguing question. One explanation could be that *Plasmodium* parasites utilise
 461 rapidly shifting surface proteins to evade detection by host immune systems
 462 (Magowan *et al.* 1988), and evolution of these protein-coding genes may allow
 463 parasites to shift among hosts that do not share conserved immune system traits
 464 (Hellgren *et al.* 2013). In contrast, individual host species typically carried
 465 phylogenetically clustered *Haemoproteus* parasites across recent timescales,
 466 suggesting that closely related parasites have similar host affinities and supporting
 467 evidence that range expansions to unrelated hosts may not be common in Pacific
 468 *Haemoproteus* parasites (Beadell *et al.* 2004; Clark *et al.* 2015a,b, 2016a). In
 469 addition, a lack of historical timescale clustering among *Haemoproteus* parasites
 470 does not necessarily discount an influence of parasite phylogeny, but may rather be
 471 evidence of hosts carrying multiple clades of specialist parasites. In Melanesia,
 472 *Zosterops* species commonly carry lineages of *Haemoproteus zosteropsis* and *H.*
 473 *killangoi*, both of which appear globally restricted to Zosteropidae hosts (Clark &
 474 Clegg 2015; Clark *et al.* 2016b; Valkiūnas 2005). Yet due to the fact that these two
 475 parasites sit in different clades, historical timescale analyses failed to uncover
 476 clustering of *Haemoproteus* parasites carried by *Zosterops* hosts. Hosts may be
 477 locally coevolving with multiple clades of parasites, stressing the utility of exploring
 478 dispersion patterns across a range of timescales (Parker *et al.* 2015).

479 While the majority of host species carried parasite assemblages that either
 480 did not differ from random expectations or were phylogenetically clustered, we did

481 find evidence that at least some host species carry overdispersed parasite
482 assemblages. Two host species carried overdispersed *Plasmodium* assemblages
483 across a range of evolutionary timescales, indicating that these hosts are susceptible
484 to a range of unrelated parasites. Thus, while we provide evidence that malaria
485 parasite evolutionary history plays a role determining which avian hosts are infected,
486 our findings do not discount the idea that some bird species spread diverse
487 pathogens as they disperse (Gregoir *et al.* 2015; Viana *et al.* 2016).

488

489 **Conclusions**

490 Consideration of pathogen host specificity and of pathogen assemblages carried by
491 host species provides complementary angles to broaden our understanding of host-
492 pathogen interactions. By uncovering distinct dispersion patterns in the relationships
493 of infected hosts, our findings highlight the utility of incorporating host ecology to
494 assess drivers of host specificity and pathogen community assembly. Moreover,
495 these patterns were robust to phylogenetic uncertainty, an element that improves
496 host specificity inferences (Garamszegi 2009; Szoellosi *et al.* 2011) yet is often
497 ignored in evolutionary ecology (Rangel *et al.* 2015). We hope that our approach to
498 consider phylogenetic uncertainty, together with our use of global databases
499 (Ducatez *et al.* 2014; Jetz *et al.* 2012; Wilman *et al.* 2014) to access phylogenetic and
500 trait data, will facilitate application of our methods to other host-pathogen systems.
501 Finally, the role of vectors is often less well understood than that for vertebrate
502 hosts, and an improved understanding will be necessary to disentangle roles of host
503 habitat niche and vector ecology in driving parasite host specificity (Žiegytė &
504 Valkiūnas 2015; Goulding *et al.* 2016).

505

506 **FIGURE LEGENDS**

507 Fig. 1: Hypothetical patterns of infected host phylogenetic relationships and
508 potential causal mechanisms. Trees represent a phylogeny of potential host species,
509 with grey shading indicating uninfected hosts and black shading indicating hosts
510 infected by a multi-host pathogen. Clustered hosts indicate hosts that are more
511 closely related than expected, while overdispersed hosts are more distantly related
512 than expected. (Full page figure, 169mm)

513

514 Fig. 2: Density scatterplots showing relationships between parasite phylogenetic β
515 diversity ($P\beta d$) and host $P\beta d$, controlling for host ecological β diversity. 10,000
516 similarity matrices were generated by randomly drawing samples of infected hosts
517 and calculating phylogenetic community nearest taxon distance (NTD) and
518 phylogenetic community mean pairwise distance (MPD) between sample pairs.
519 Regression lines represent 95% credible intervals of variance explained (R^2) from
520 multiple regressions on distance matrices. Individual host and parasite pairwise
521 similarity points were scaled prior to plotting. (Full page figure, 169mm)

522

523 Fig. 3: Host specificity of avian malaria parasites. (a) Differentials between observed
524 and expected host functional-phylogenetic specificity (STD^*) at varying α weights.
525 Weights approaching 0 give higher weight to host ecological distance; values
526 approaching 1 give higher weight to host phylogenetic distance. Negative
527 differentials indicate clustered hosts (more similar than expected); positive values
528 indicate overdispersed hosts (more dissimilar than expected). Differentials were
529 generated from 10,000 iterations, using randomly sampled host phylogenies and
530 habitat dendrograms in each iteration. Boxplots show differential medians (lines
531 within boxes), 25% and 75% quartiles (hinges) for individual parasites. Asterisks (*)
532 indicate significant differences from 0. (b) Relationship between α weights and
533 specificity indices for select parasites infecting ≥ 3 host species. For clarity, four
534 parasites were chosen to showcase the range of index patterns observed, with
535 colours matching corresponding boxplots in part (a). Indices were generated using
536 only nearest taxon functional-phylogenetic distance ($NTSTD^*$; top graph) or mean
537 pairwise distances (STD^*). Higher indices indicate more generalist parasites, while
538 low scores indicate specialists. (Full page figure, 169mm).

539

540 Fig. 4: Dispersion in parasite phylogenetic distances for host species carrying ≥ 3
541 parasites. Numbers in parentheses indicate the number of parasites observed
542 infecting each host species. Observed and expected parasite distances were
543 generated from 10,000 iterations, using a randomly sampled parasite phylogeny in
544 each iteration. Negative differentials indicate phylogenetically clustered parasites
545 (more closely related than expected by chance) while positive differentials indicate
546 overdispersed parasites (more distantly related than expected). Presented are
547 quantiles of parasite distances ranging from zero (nearest taxon distance; NTD) to 1
548 (mean pairwise distance; MPD). Boxplots show differential medians (lines within
549 boxes) as well as 25% and 75% quartiles (hinges). Asterisks (*) indicate significant
550 differences from 0. (Two thirds page figure, 112mm)

551

For Review Only

552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567

ACKNOWLEDGEMENTS

We thank D. Treby, K. Lowe, J. LeBreton, S. Oghino, F. Cugny, W. Waheoneme, T. Read, Association pour la Sauvegarde de la Biodiversité d’Ouvéa, and Société Calédonienne d’Ornithologie for field and/or logistical assistance. We express our gratitude to O. Hébert and A. Rouquié for assistance with Loyalty Island fieldwork. Permits were provided by Direction de l’Environnement Province Sud (Grand Terre South Province Permit N° 3177-2013/ARR/DENV) and Direction Du Developpement Economique (Loyalty Islands Permit N° 6161-43/PR). We thank G. Kakue for issuing Loyalty Island permits. We thank K. Wells and two anonymous reviewers for comments that improved the manuscript. Fieldwork and collection of avian blood samples was conducted Griffith University ethics clearance ENV/01/12/AEC. Funding was provided by a Griffith University New Researcher Grant and a National Geographic Society Committee for Research and Exploration Grant (#9383-13) to SMC.

For Review Only

REFERENCES

- Abouheif E (1999) A method for testing the assumption of phylogenetic independence in comparative data. *Evolutionary Ecology Research* **1**, 895-909.
- Adlard RD, Miller TL, Smit NJ (2015) The butterfly effect: parasite diversity, environment, and emerging disease in aquatic wildlife. *Trends in Parasitology* **31**, 160-166.
- Altizer SM, Thrall PH, Antonovics J (1998) Vector behavior and the transmission of anther-smut infection in *Silene alba*. *The American Midland Naturalist* **139**, 147-163.
- Beadell JS, Gering E, Austin J, *et al.* (2004) Prevalence and differential host-specificity of two avian blood parasite genera in the Australo-Papuan region. *Molecular Ecology* **13**, 3829-3844.
- Bensch, S., Pérez-Tris, J., Waldenström, J. & Hellgren, O. (2004) Linkage between nuclear and mitochondrial DNA sequences in avian malaria parasites: Multiple cases of cryptic speciation? *Evolution* **58**, 1617-1621.
- Braga MP, Razzolini E, Boeger WA (2015) Drivers of parasite sharing among Neotropical freshwater fishes. *Journal of Animal Ecology* **84**, 487-497.
- Brooks DR, Hoberg EP, Boeger WA (2015) In the eye of the cyclops: the classic case of cospeciation and why paradigms are important. *Comparative Parasitology* **82**, 1-8.
- Cadotte M, Albert CH, Walker SC (2013) The ecology of differences: assessing community assembly with trait and evolutionary distances. *Ecology Letters* **16**, 1234-1244.
- Clark NJ, Clegg SM (2015) The influence of vagrant hosts and weather patterns on the colonization and persistence of blood parasites in an island bird. *Journal of Biogeography* **42**, 641-651.
- Clark NJ, Clegg SM, Klaassen M (2016a) Migration strategy and pathogen risk: non-breeding distribution drives malaria prevalence in migratory waders. *Oikos* **125**, 1358-1368.
- Clark NJ, Clegg SM, Lima MR (2014) A review of global diversity in avian haemosporidians (*Plasmodium* and *Haemoproteus*: Haemosporida): new insights from molecular data. *International Journal for Parasitology* **44**, 329-338.
- Clark NJ, Olsson-Pons S, Ishtiaq F, Clegg SM (2015a) Specialist enemies, generalist weapons and the potential spread of exotic pathogens: malaria parasites in a highly invasive bird. *International Journal for Parasitology* **45**, 891-899.
- Clark, NJ, Adlard, RD, Clegg, SM (2015b) Molecular and morphological characterization of *Haemoproteus* (*Parahaemoproteus*) *ptilotis*, a parasite infecting Australian honeyeaters (Meliphagidae), with remarks on prevalence and potential cryptic speciation. *Parasitology Research* **114**, 1921-1928.
- Clark NJ, Wells K, Dimitrov D, Clegg SM (2016b) Co-infections and environmental conditions drive the distributions of blood parasites in wild birds. *Journal of Animal Ecology* **85**, 1461-1470.
- Cooper N, Griffin R, Franz M, Omotayo M, Nunn CL (2012) Phylogenetic host specificity and understanding parasite sharing in primates. *Ecology Letters* **15**,

- 1370-1377.
- De Vienne DM, Hood ME, Giraud T (2009) Phylogenetic determinants of potential host shifts in fungal pathogens. *Journal of Evolutionary Biology* **22**, 2532-2541.
- Ducatez S, Tingley R, Shine R (2014) Using species co-occurrence patterns to quantify relative habitat breadth in terrestrial vertebrates. *Ecosphere* **5**, 1-12.
- Dutson G (2012) *Birds of Melanesia: Bismarcks, Solomons, Vanuatu, and New Caledonia* A&C Black, New Jersey, USA.
- Ellis VA, Collins MD, Medeiros MC, *et al.* (2015) Local host specialization, host-switching, and dispersal shape the regional distributions of avian haemosporidian parasites. *Proceedings of the National Academy of Sciences* **112**, 11294-11299.
- Ewen JG, Bensch S, Blackburn TM, *et al.* (2012) Establishment of exotic parasites: the origins and characteristics of an avian malaria community in an isolated island avifauna. *Ecology Letters* **15**, 1112-1119.
- Farrell MJ, Berrang-Ford L, Davies TJ (2013) The study of parasite sharing for surveillance of zoonotic diseases. *Environmental Research Letters* **8**, 015036.
- Gager AB, Loaiza JDR, Dearborn DC, Bermingham E (2008) Do mosquitoes filter the access of *Plasmodium* cytochrome b lineages to an avian host? *Molecular Ecology* **17**, 2552-2561.
- Garamszegi LZ (2009) Patterns of co-speciation and host switching in primate malaria parasites. *Malaria Journal* **8**, 1.
- Gervasi SS, Civitello DJ, Kilvitis HJ, Martin LB (2015) The context of host competence: a role for plasticity in host-parasite dynamics. *Trends in Parasitology* **9**, 419-425.
- Gómez JM, Nunn CL, Verdú M (2013) Centrality in primate-parasite networks reveals the potential for the transmission of emerging infectious diseases to humans. *Proceedings of the National Academy of Sciences* **110**, 7738-7741.
- Goslee SC, Urban DL (2007) The ecodist package for dissimilarity-based analysis of ecological data. *Journal of Statistical Software* **22**, 1-19.
- Goulding, W, Adlard, RD, Clegg, SM, Clark, NJ (2016) Molecular and morphological description of *Haemoproteus* (*Parahaemoproteus*) *bukaka* (species nova), a haemosporidian associated with the strictly Australo-Papuan host subfamily Cracticinae. *Parasitology Research* **115**, 3387-3400.
- Gower JC (1971) A general coefficient of similarity and some of its properties. *Biometrics*, 857-871.
- Gregoir AF, Hablützel PI, Vanhove MP, *et al.* (2015) A link between host dispersal and parasite diversity in two sympatric cichlids of Lake Tanganyika. *Freshwater Biology* **60**, 323-335.
- Griekspoor P, Colles FM, McCarthy ND, *et al.* (2013) Marked host specificity and lack of phylogeographic population structure of *Campylobacter jejuni* in wild birds. *Molecular Ecology* **22**, 1463-1472.
- Hechinger RF, Lafferty KD (2005) Host diversity begets parasite diversity: bird final hosts and trematodes in snail intermediate hosts. *Proceedings of the Royal Society B: Biological Sciences* **272**, 1059-1066.
- Heger T, Jeschke JM (2014) The enemy release hypothesis as a hierarchy of hypotheses. *Oikos* **123**, 741-750.

- 662 Hellgren O, Pérez J, Bensch S (2009) A jack-of-all-trades and still a master of some:
663 prevalence and host range in avian malaria and related blood parasites.
664 *Ecology* **90**, 2840-2849.
- 665 Hellgren, O, Kutzer, M, Bensch, S, Valkiūnas, G & Palinauskas, V (2013) Identification
666 and characterization of the merozoite surface protein 1 (msp1) gene in a
667 host-generalist avian malaria parasite, *Plasmodium relictum* (lineages SGS1
668 and GRW4) with the use of blood transcriptome. *Malaria Journal* **12**, 381.
- 669 Huang S, Bininda-Emonds OR, Stephens PR, Gittleman JL, Altizer S (2014)
670 Phylogenetically related and ecologically similar carnivores harbour similar
671 parasite assemblages. *Journal of Animal Ecology* **83**, 671-680.
- 672 Jetz W, Thomas G, Joy J, Hartmann K, Mooers A (2012) The global diversity of birds in
673 space and time. *Nature* **491**, 444-448.
- 674 Johnson, PT, Rohr, JR, Hoverman, JT, *et al.* (2012) Living fast and dying of infection:
675 host life history drives interspecific variation in infection and disease risk.
676 *Ecology Letters* **15**, 235-242.
- 677 Johnson PT, Wood CL, Joseph MB, *et al.* (2016) Habitat heterogeneity drives the
678 host-diversity-begets-parasite-diversity relationship: evidence from
679 experimental and field studies. *Ecology Letters* **19**, 752-761
- 680 Johnson PTJ, Hoverman JT (2014) Heterogeneous hosts: how variation in host size,
681 behaviour and immunity affects parasite aggregation. *Journal of Animal*
682 *Ecology* **83**, 1103-1112.
- 683 Keane RM, Crawley MJ (2002) Exotic plant invasions and the enemy release
684 hypothesis. *Trends in Ecology & Evolution* **17**, 164-170.
- 685 Kraft NJ, Cornwell WK, Webb CO, Ackerly DD (2007) Trait evolution, community
686 assembly, and the phylogenetic structure of ecological communities. *The*
687 *American Naturalist* **170**, 271-283.
- 688 Krasnov BR, Mouillot D, Shenbrot GI, *et al.* (2010) Similarity in ectoparasite faunas of
689 Palaearctic rodents as a function of host phylogenetic, geographic or
690 environmental distances: which matters the most? *International Journal for*
691 *Parasitology* **40**, 807-817.
- 692 Krasnov BR, Shenbrot GI, Khokhlova IS, Degen AA (2016) Trait-based and
693 phylogenetic associations between parasites and their hosts: a case study
694 with small mammals and fleas in the Palearctic. *Oikos* **125**, 29-38.
- 695 Krasnov BR, Shenbrot GI, Khokhlova IS, *et al.* (2015) Assembly rules of ectoparasite
696 communities across scales: combining patterns of abiotic factors, host
697 composition, geographic space, phylogeny and traits. *Ecography* **38**, 184-197.
- 698 Kulbicki M, Parravicini V, Bellwood DR, *et al.* (2013) Global biogeography of reef
699 fishes: a hierarchical quantitative delineation of regions. *Plos One* **8**, e81847.
- 700 Lauron EJ, Loiseau C, Bowie RCK, *et al.* (2014) Coevolutionary patterns and
701 diversification of avian malaria parasites in African sunbirds (Family
702 Nectariniidae). *Parasitology* **142**, 635-647.
- 703 Lichstein JW (2007) Multiple regression on distance matrices: a multivariate spatial
704 analysis tool. *Plant Ecology* **188**, 117-131.
- 705 Little TJ, Watt K, Ebert D, Koella J (2006) Parasite-host specificity: experimental
706 studies on the basis of parasite adaptation. *Evolution* **60**, 31-38.
- 707 Longdon B, Hadfield JD, Webster CL, Obbard DJ, Jiggins FM (2011) Host phylogeny
708 determines viral persistence and replication in novel hosts. *PLoS Pathogens* **7**,

- 709 e1002260.
- 710 Lootvoet A, Blanchet S, Gevrey M, *et al.* (2013) Patterns and processes of alternative
711 host use in a generalist parasite: insights from a natural host–parasite
712 interaction. *Functional Ecology* **27**, 1403–1414.
- 713 Magowan, C., Wollish, W., Anderson, L. & Leech, J. (1988) Cytoadherence by
714 *Plasmodium falciparum*-infected erythrocytes is correlated with the
715 expression of a family of variable proteins on infected erythrocytes. *Journal*
716 *of Experimental Medicine* **168**, 1307–1320.
- 717 McCoy K, Léger E, Dietrich M (2013) Host specialization in ticks and transmission of
718 tick-borne diseases: a review. *Frontiers in Cellular and Infection Microbiology*
719 **3**, 57.
- 720 Medeiros MC, Hamer GL, Ricklefs RE (2013) Host compatibility rather than vector-
721 host-encounter rate determines the host range of avian *Plasmodium*
722 parasites. *Proceedings of the Royal Society B: Biological Sciences* **280**,
723 20122947.
- 724 Moens MA, Pérez-Tris J (2015) Discovering potential sources of emerging pathogens:
725 South America is a reservoir of generalist avian blood parasites. *International*
726 *Journal for Parasitology* **46**, 41–49.
- 727 Mouillot D, Šimková A, Morand S, Poulin R (2005) Parasite species coexistence and
728 limiting similarity: a multiscale look at phylogenetic, functional and
729 reproductive distances. *Oecologia* **146**, 269–278.
- 730 Njabo KY, Cornel AJ, Bonneaud C, *et al.* (2011) Nonspecific patterns of vector, host
731 and avian malaria parasite associations in a central African rainforest.
732 *Molecular Ecology* **20**, 1049–1061.
- 733 Nunn CL, Altizer SM, Sechrest W, Cunningham AA (2005) Latitudinal gradients of
734 parasite species richness in primates. *Diversity and Distributions* **11**, 249–256.
- 735 Olsson-Pons S, Clark NJ, Ishtiaq F, Clegg SM (2015) Differences in host species
736 relationships and biogeographic influences produce contrasting patterns of
737 prevalence, community composition and genetic structure in two genera of
738 avian malaria parasites in southern Melanesia. *Journal of Animal Ecology* **84**,
739 985–998.
- 740 Parker IM, Saunders M, Bontrager M, *et al.* (2015) Phylogenetic structure and host
741 abundance drive disease pressure in communities. *Nature* **520**, 542–544.
- 742 Pavoine S, Ricotta C (2013) Testing for phylogenetic signal in biological traits: the
743 ubiquity of cross-product statistics. *Evolution* **67**, 828–840.
- 744 Pavoine S, Vallet J, Dufour AB, Gachet S, Daniel H (2009) On the challenge of treating
745 various types of variables: application for improving the measurement of
746 functional diversity. *Oikos* **118**, 391–402.
- 747 Perlman SJ, Jaenike J (2003) Infection success in novel hosts: an experimental and
748 phylogenetic study of *Drosophila*-parasitic nematodes. *Evolution* **57**, 544–557.
- 749 Plowright RK, Eby P, Hudson PJ, *et al.* (2015) Ecological dynamics of emerging bat
750 virus spillover. *Proceedings of the Royal Society of London B: Biological*
751 *Sciences* **282**, 20142124.
- 752 Poulin R, Krasnov BR, Shenbrot GI, Mouillot D, Khokhlova IS (2006) Evolution of host
753 specificity in fleas: is it directional and irreversible? *International Journal for*
754 *Parasitology* **36**, 185–191.
- 755 Poulin R, Mouillot D (2005) Combining phylogenetic and ecological information into

- a new index of host specificity. *Journal of Parasitology* **91**, 511-514.
- R Development Team (2008) R: A language and environment for statistical computing. R Development Core Team, Vienna, Austria.
- Rangel TF, Colwell RK, Graves GR, *et al.* (2015) Phylogenetic uncertainty revisited: implications for ecological analyses. *Evolution* **69**, 1301-1312.
- Ricklefs RE, Fallon SM, Bermingham E (2004) Evolutionary relationships, cospeciation, and host switching in avian malaria parasites. *Systematic Biology* **53**, 111-119.
- Ricklefs RE, Outlaw DC, Svensson-Coelho M, *et al.* (2014) Species formation by host shifting in avian malaria parasites. *Proceedings of the National Academy of Sciences* **111**, 14816-14821.
- Santiago-Alarcon, D, Palinauskas, V, Schaefer, HM (2012) Diptera vectors of avian Haemosporidian parasites: untangling parasite life cycles and their taxonomy. *Biological Reviews* **87**, 928-964.
- Schulze-Lefert P, Panstruga R (2011) A molecular evolutionary concept connecting nonhost resistance, pathogen host range, and pathogen speciation. *Trends in plant science* **16**, 117-125.
- Sehgal RN (2015) Manifold habitat effects on the prevalence and diversity of avian blood parasites. *International Journal for Parasitology: Parasites and Wildlife* **4**, 421-430.
- Sehgal RNM, Buermann W, Harrigan RJ, *et al.* (2010) Spatially explicit predictions of blood parasites in a widely distributed African rainforest bird. *Proceedings of the Royal Society B-Biological Sciences* **278**, 1025-1033.
- Shaw D, Dobson A (1995) Patterns of macroparasite abundance and aggregation in wildlife populations: a quantitative review. *Parasitology* **111**, S111-S133.
- Silvertown J, Dodd M, Gowing D (2001) Phylogeny and the niche structure of meadow plant communities. *Journal of Ecology* **89**, 428-435.
- Stegen JC, Lin X, Konopka AE, Fredrickson JK (2012) Stochastic and deterministic assembly processes in subsurface microbial communities. *The ISME Journal* **6**, 1653-1664.
- Streicker DG, Turmelle AS, Vonhof MJ, *et al.* (2010) Host phylogeny constrains cross-species emergence and establishment of rabies virus in bats. *Science* **329**, 676-679.
- Svensson-Coelho, M, Loiselle, BA, Blake, JG, Ricklefs, RE (2016) Resource predictability and specialization in avian malaria parasites. *Molecular Ecology* **25**, 4377-4391.
- Szoellosi E, Cichon M, Eens M, *et al.* (2011) Determinants of distribution and prevalence of avian malaria in blue tit populations across Europe: separating host and parasite effects. *Journal of Evolutionary Biology* **24**, 2014-2024.
- Tsirogiannis C, Sandel B (2015) PhyloMeasures: a package for computing phylogenetic biodiversity measures and their statistical moments. *Ecography* **39**, 707-714.
- Valkiūnas G (2005) *Avian malaria parasites and other Haemosporida* CRC Press, Boca Raton, FL.
- Viana DS, Santamaría L, Figuerola J (2016) Migratory birds as global dispersal vectors. *Trends in Ecology & Evolution* **31**, 763-775.
- Weiblen GD, Webb CO, Novotny V, Basset Y, Miller SE (2006) Phylogenetic dispersion

803 of host use in a tropical insect herbivore community. *Ecology* **87**, 62-75.
804 Wells K, O'Hara RB, Morand S, Lessard J-P, Ribas A (2015) The importance of parasite
805 geography and spillover effects for global patterns of host–parasite
806 associations in two invasive species. *Diversity and Distributions* **21**, 477-486.
807 Wilman H, Belmaker J, Simpson J, *et al.* (2014) EltonTraits 1.0: Species-level foraging
808 attributes of the world's birds and mammals. *Ecology* **95**, 2027-2027.
809 Woolhouse ME, Taylor LH, Haydon DT (2001) Population biology of multihost
810 pathogens. *Science* **292**, 1109-1112.
811 Žiegytė, R, Valkiūnas, G (2015) Recent advances in vector studies of avian
812 haemosporidian parasites. *Ekologija* **60**, 73-83.
813 Zukal J, Bandouchova H, Bartonicka T, *et al.* (2014) White-nose syndrome fungus: a
814 generalist pathogen of hibernating bats. *PLoS One* **9**, e97224.
815
816

817

818

DATA ACCESSIBILITY

819

820

821

822

823

824

825

AUTHOR CONTRIBUTIONS

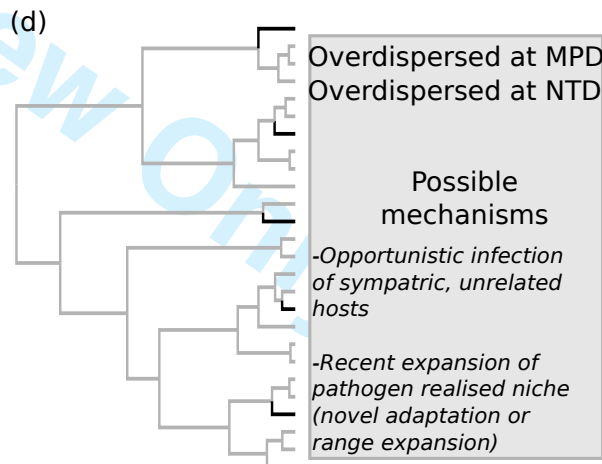
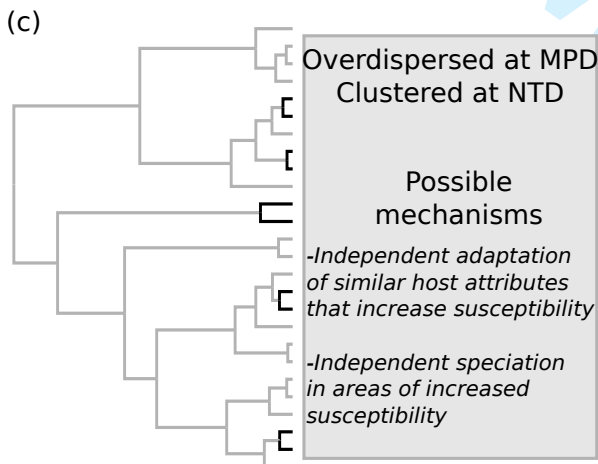
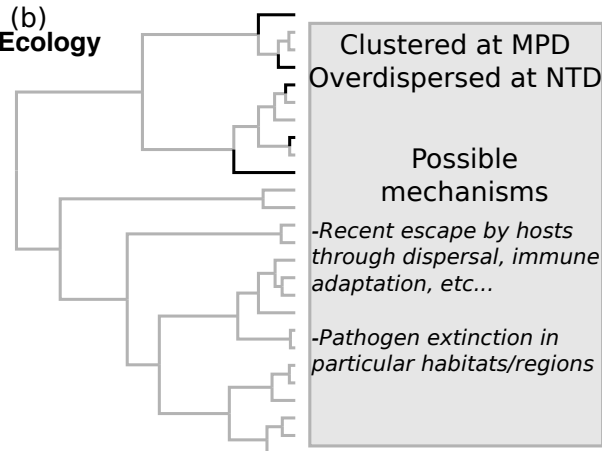
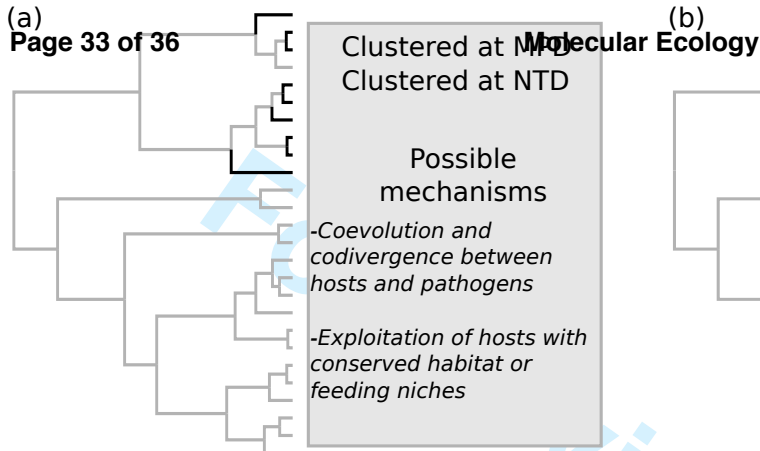
826

827

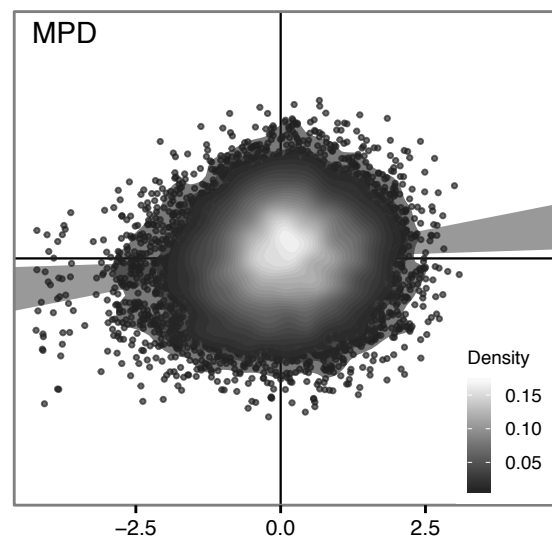
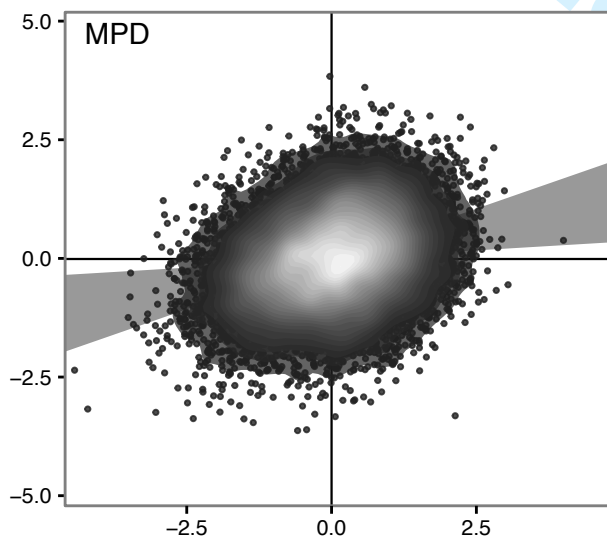
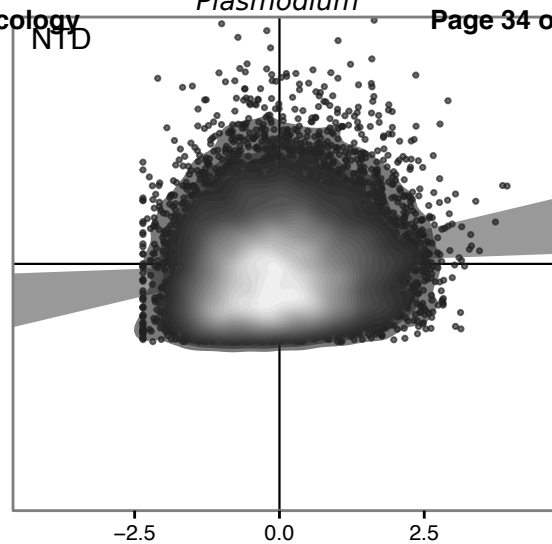
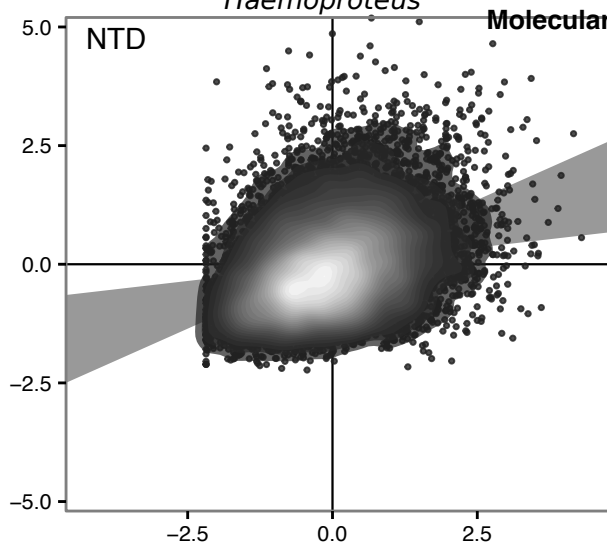
Newly recorded parasite cytochrome *b* lineages are deposited in GenBank (accessions KY671236-KY671244) and the avian malaria database MalAvi (<http://mbio-serv2.mbioekol.lu.se/Malavi/>). Analysis R code, host-parasite infection data and host niche trait data are available from the Dryad Digital Repository: <http://dx.doi.org/10.5061/dryad.50603t>.

N.J.C. devised the study and performed analyses. N.J.C and S.M.C. gathered field data and wrote the manuscript.

For Review Only



MPD (Mean Pairwise Distance): compares mean of all observed host pairwise distances to null expectations
NTD (Nearest Taxon Distance): compares minimum observed host pairwise distance to null expectations

Scaled parasite P β d

Density

0.15

0.10

0.05

Scaled host P β d

