Local host specialization, host-switching, and dispersal shape the regional distributions of avian haemosporidian parasites

Vincenzo A. Ellis[a,b,1], Michael D. Collins[c], Matthew C. I. Medeiros[a], Eloisa H. R. Sari[b], Elyse D. Coffey[a], Rebecca C. Dickerson[a,d], Camile Lugarini[a,e], Jeffrey A. Stratford[f], Donata R. Henry[g], Loren Merrill[h], Alicenzo A. Ellis

across eastern North America and related the distributions of individual parasite lineages to regional climate variation and to the distributions and abundances of their avian hosts. Community dissimilarities between sampling locations based on host assemblage structure (i.e., the relative abundances of potential host species) were positively correlated with those based on parasite

Table 1. Results of partial Mantel tests comparing hypothesized relationships between space (i.e., geographic distance between sites), the environment (climate differences between sites), birds (host community dissimilarity between sites), and parasites (parasite community dissimilarity between sites) identified in Fig. 2

Relationship between	And	Controlling for	r_p	P
Space	Environment	None	0.595	0.005
Birds	Environment	Parasites	0.772	<0.001
Birds	Space	Parasites	0.504	0.012
Birds	Environment	Space	0.720	<0.001
Birds	Space	Environment	0.185	0.137
Parasites	Environment	Birds	0.117	0.277
Parasites	Space	Birds	0.097	0.302
Parasites	Environment	Space	0.303	0.076
Parasites	Space	Environment	0.101	0.300
Birds	Parasites	Environment	0.191	0.144
Birds	Parasites	Space	0.335	0.027

We report the partial Mantel correlation coefficient (r_p) and associated P value. The relationship between space and environment was tested with a standard Mantel test. Bolded values of

similarity does not decline with distance [i.e., parasite distributions were not spatially restricted (35) when controlling for hosts], suggesting that parasites disperse readily across the region within their host populations. These results generally held when the parasite genera were analyzed separately (*SI Appendix*, Table S5) and when using an alternative statistical approach (*SI Appendix*, Table S6).

The host-breadth of a parasite may vary geographically or temporally, and may also be limited by the phylogenetic relatedness of potential host species (13). For example, in the Chicago location, each *Plasmodium* parasite lineage was associated with a single host taxon at the superfamily level (23). To determine the importance of host phylogeny on parasite distributions across the region, we created a phylogenetic distance matrix for all hosts infected at least once by any of the 33 parasite lineages sampled 10 or more times (60 host species). We then calculated a second matrix by computing Bray–Curtis dissimilarities between those hosts based on the number of times each host species was infected with each of the 33 parasite lineages. A Mantel test comparing these two matrices showed a weak, but significant, correlation ($r = 0.28$, $P = 0.002$), indicating that parasite host distribution is constrained to more closely related hosts than expected by chance. Interestingly, this effect varied across locations in the region (*SI Appendix*, Table S7).

To quantify the host-breadth of each parasite, we used the Gini–Simpson index (36), which accounts for the number of infections recorded for each host species (13). We weighted the index by the phylogenetic distance between hosts using the formula for Rao's quadratic entropy [Rao's *QE* (37, 38); see *Materials and Methods* for formula; results did not change qualitatively if phylogenetic distances were not included in these analyses]. Although ecologists often distinguish generalist and specialist parasites, host-breadth in the 33 parasite lineages sampled 10 or more times was continuously distributed (*SI Appendix*, Fig. S2) and did not differ statistically from a unimodal distribution [Hartigan's dip test: $D_{33} = 0.047$, $P = 0.87$ (39)]. Furthermore, we found no difference in the host-breadth of individual parasite lineages between the parasite genera ($t_{31} = -1.1$, $P = 0.28$).

When all years were pooled, parasite lineages recovered at least four times from each of at least four community sampling locations exhibited variation in local host-breadth across the region (Fig. 3). A linear mixed-effects model with parasite lineage as a random effect showed no influence of local phylogenetically weighted bird diversity (Rao's *QE*, using host species infected at least once in the region) on parasite host-breadth ($F_{1,21.4} = 1.26$, $P = 0.27$), suggesting that variation in host-breadth is not simply attributable to the diversity of available hosts. Furthermore, local parasite diversity did not influence parasite host-breadth ($F_{1,21.2} = 2.41$, $P = 0.14$). For example, parasite lineage LA01 (*Haemoproteus* sp.) was recovered exclusively from *Dumetella carolinensis* in Chicago, IL (23/157 *D. carolinensis* hosts infected; years sampled 2006 and 2007); Connecticut (4/45; 2002 and 2003); and Michigan (11/94; 2012). However, in the 2013 Tennessee sample, LA01 was recovered from the hosts *Mimus polyglottos* (like *D. carolinensis*, in the family Mimidae; 2/9 infected), *Cardinalis cardinalis* (1/36), and *Sialia sialis* (1/19), whereas the two *D. carolinensis* hosts sampled in Tennessee were both uninfected. We also recovered LA01 from *D. carolinensis* in the western Chicago location (6/7) in 2014 and from *D. carolinensis* (2/6) and *Toxostoma rufum* (also in the family Mimidae; 1/7) in Champaign, IL, in the same year (although those were not community samples).

To determine whether local host-breadth differed from a random expectation, we restricted our dataset to infected individuals of those five potential host species of LA01. We then shuffled all parasite lineages infecting those hosts within sampling locations and recalculated randomized host-breadths for LA01 (9,999 randomizations) and compared observed host-breadths to the distribution of randomized host breadths. In Chicago, the host-breadth of LA01 was lower than expected by chance ($P < 0.001$), whereas in Tennessee, this lineage's host-breadth was higher than expected by chance ($P = 0.019$). The host-breadth of LA01 did not differ from random in Connecticut and Michigan because there were no potential alternative hosts in either location. Lineage Ozarks 06 (OZ06) (*Plasmodium* sp.) also varied with respect to host-breadth (Fig. 3). The host-breadth of OZ06 was lower than expected based on a random distribution (again shuffling infections among potential hosts) in Michigan ($P = 0.003$), Indiana ($P < 0.001$), and Tennessee ($P = 0.030$) but did not differ from random in Chicago ($P = 0.76$) and the Ozarks ($P = 0.94$).

Because locations were sampled in different years, some variation in host-breadth between localities might reflect temporal change within localities. Within particular years, parasite lineages sampled more than three times at multiple locations mostly showed little variation in host-breadth. However, in 2013, OZ14 (*Plasmodium* sp.) infected three hosts in Pennsylvania (6/12 *Melospiza melodia* infected, also 1/3 *Pipilo erythrophthalmus*, and 1/1 *Pheucticus ludoviciana*) but infected a larger variety of species in Tennessee (6/50 *Passerina cyanea* individuals infected and 12

importance of host-switching in determining parasite distributions across the region.

Finally, although theoretical (48) and empirical (49) studies suggest that parasites may often limit host population size, the distributions of correlations between host and parasite populations across the region did not differ from random, suggesting that haemosporidian parasites do not impact the population densities of their hosts in eastern North America. Our analyses suggest that populations of haemosporidian parasites are largely structured by populations of their hosts, although parasite lineages change between nearby localities within host species distributions and over short intervals within localities.

We captured birds with mist-nets at 13 locations across eastern North America (Fig. 1) during summer months (primarily late May to August, with minimal sampling in April and September; removal of April and September samples did not qualitatively change results) from 1999 to 2014 (SI Appendix, Table S2). We took a small (approximately 10-µL) blood sample from the brachial vein of each bird and stored the blood in Puregene or Longmire's (50) lysis buffer. We collected all samples under appropriate state and federal permits and Institutional Animal Care and Use Committee (IACUC) protocols.

We extracted DNA from blood samples using an ammonium acetate-isopropanol precipitation protocol (51). We screened DNA samples for haemosporidian parasites using a PCR protocol designed to amplify a small section of parasite mitochondrial DNA (52). We then amplified a portion of the cytochrome b gene in positive samples using several primer pairs and protocols (15, 40, 53, 54). We identified unique parasite lineages based on their cytochrome b sequences and on their host and geographic distributions (55, 56). Multiple infections were separated by phasing (57) where possible. GenBank Accession numbers for all lineages can be found in SI Appendix, Table S1.

All analyses were performed in R v3.1.2 (58), and we report two-tailed P values for all tests. We calculated Bray–Curtis dissimilarities between locations with the "vegdist" function in the vegan package (59). Bray–Curtis dissimilarity between two sampling locations (1, 2) is calculated by

$$D = \frac{\sum_{j=1}^{p} \left| y_{1j} - y_{2j} \right|}{\sum_{j=1}^{p} \left(y_{1j} + y_{2j} \right)},$$

where y represents the number (or frequency) of individuals sampled of species j, and p represents the total number of species sampled over both locations (34).

We created a geographic distance matrix between locations with the "rdist.earth" function in the fields package (60) in R. We compared distance matrices with Mantel and partial Mantel tests using functions "mantel" and "mantel.partial"

11. Ricklefs RE (2010) Host-pathogen coevolution, secondary sympatry and species diversification. Philos Trans R Soc Lond B Biol Sci 365(1543):1139–1147.
12. Combes C (2001) Parasitism: The Ecology and Evolution of Intimate Interactions (Univ of Chicago Press, Chicago).
13. Poulin R, Krasnov BR, Mouillot D (2011) Host specificity in phylogenetic and geographic space. Trends Parasitol 27(8):355–361.