Host phenology, geographic range size and regional occurrence explain interspecific variation in damselfly–water mite associations

Julia J. Mlynarek, W. Knee and Mark R. Forbes

In this study, we tested which host species’ characteristics explain the nature and level of parasitism for host damselfly (Coenagrionidae)–water mite (Arrenuridae) parasite associations. Prevalence and intensity of mite parasites, and mite species richness were examined in relation to geographic range size, regional occurrence, relative local abundance, phenology and body size of host damselfly species. A total of 7107 damselfly individuals were collected representing 16 species from 13 sites in southeastern Ontario and southwestern Quebec, Canada. Using comparative methods, differences in prevalence and intensity of parasitism could be predicted by a host species’ geographic range and phenology. Barcoding based on Cytochrome Oxidase I revealed 15 operational taxonomic units (OTUs) for mite species. The number of mite OTUs known to infest a given host species was explained by a host species’ regional occurrence. Our findings demonstrate the need to measure factors at several ecological scales in order to understand the breadth of evolutionary interactions with host–parasite associations and the selective ‘milieu’ for particular species of both hosts and parasites.

Evolutionary history of the association between two species is expected to have an effect on present-day interactions, but so are ecological factors (Thompson 2005). In most cases, that explanatory power of ecological factors depends on the host–parasite system under study, perhaps because of certain life history traits (Poulin 2007). Several ecological factors have been reported to explain variation in parasitism between host species, such as size or landscape configuration of habitat (Tella et al. 1999); size, age, reproductive stage or phenology of the host species (Triplet and Richner 1997, Arneberg et al. 1998, Krasnov et al. 2006, Munoz et al. 2007), and size, virulence and life history of the parasite (Krasnov et al. 2004). Although there are many predictors to explain parasitism, host body size, geographic distribution, and population density, seem to be strong predictors in most host–parasite associations (Kamiya et al. 2014).

Host species’ characteristics such as geographic range size, regional occurrence, local abundance, and phenology have been offered as explanations for interspecific variation in parasitism (Morand and Krasnov 2010). As a host species range increases, so does the number of species it interacts with (Krasnov et al. 2004, Ilvonen et al. 2011) because as a host species range increases, so too does the probability that the host species will come into contact and be parasitized by more parasites (Price et al. 1988). However within its entire range, the host species is not evenly distributed in either space or time (Thompson 2005), because there are habitats where individuals cannot survive and reproduce well, if at all. Thus, not only the size of a species’ geographic range, but also actual representation of host species across the region of interest could be important. It is expected that a regionally well-represented host species will come into contact with more parasite species than host species less well represented regionally (cf. Durrer and Schmid-Hempel 1995). Species are not only unevenly distributed across space, but there are also differences in the representation of a species through time (phenology). These phenological differences between species are expected to influence their interactions, especially between groups that are active during different times of the year (Altizer et al. 2006, Locklin and Vodopich 2010). One expectation is that density of infective stages of parasites would follow density of potential hosts throughout the season of activity (Forbes et al. 2012). More specifically, the density of infective stages of parasites and measures of parasitism itself (e.g. prevalence and intensity) should be highest when potential host individuals of one or more species are most abundant. Local abundance of a host species might influence its interactions with co-occurring species. The intensity of interactions increases with increased host abundance because the probability of encounters with other species increases (Poulin 1991, Arneberg et al. 1998, Krasnov et al. 2002, Patterson et al. 2008). As a host species becomes more abundant in a habitat, it is thus expected to be under higher pressure from the parasites (Arneberg et al. 1998).

Our understanding of parasite associations in insect hosts is not as well documented as that of vertebrate hosts.
In mites associated with ants, mite prevalence was explained by the degree of sociality, the size and the generic identity of the host species (Campbell et al. 2013). Durrer and Schimd-Hempel (1995) did a comprehensive study of bumblebee parasites and found that parasite load was positively correlated with average colony size, local abundance and geographic distribution of the host. The geographic distribution of damselfly hosts explained in part the prevalence and intensity of gregarine endoparasites (Mlynarek et al. 2012). In contrast, host geographic distribution was a poor predictor of \textit{Arrenurus} (Arrenuridae) water mite prevalence and intensity when comparing pairs of closely related damselfly host species from single sites (Mlynarek et al. 2013).

The association between damselflies and \textit{Arrenurus} water mites is nonetheless a good model system for host–parasite studies because of the nature of the obligate interaction characterized by a direct life cycle, the similarity in trophic level of the host species, and the known differences in distribution and expected differences in phenology and regional occurrence of different host species (Forbes and Robb 2008). Larval \textit{Arrenurus} mites contact their future damselfly hosts in the water column. As the formerly larval damselfly ecloses, the mites move from the exuviae onto the teneral adult and embed their gnathosome through the unsclerotized cuticle into the hemolymph (Smith et al. 2010). The mites then form a blind ended feeding tube known as a stylostome and start feeding (Smith et al. 2010). Damselflies can resist mites by grooming loosely attached mites, or through melanisation once a mite has formed a stylostome (Forbes et al. 1999, Leung et al. 1999, Mlynarek et al. 2014).

In this study, we tested the extent to which geographic range, regional occurrence, local abundance, phenology and body size of host species are good predictors of prevalence and intensity of water mite parasitism or of proxies of species richness of mite parasites.

Our study was restricted to damselflies of the family Coenagrionidae, but we expand upon Mlynarek et al. (2013) by examining more host species across many sites and by using parasite species diversity as an additional measure of parasitism. We tested the generality of regional occurrence, local abundance and body size in explaining interspecific variation in parasitism, because they have been shown to be important factors in vertebrate–parasite systems (Poulin 2007).

\section*{Methods}

\subsection*{Sites}

Thirteen sites were chosen and were visited on a weekly basis for eight weeks in June and July 2011 (Fig. 1). Sites were spread along a 400 km longitudinal transect from the Queen’s Univ. Biological Station (QUBS), 60 km northwest of Kingston, Ontario to Johnville Bog 20 km east of the city of Sherbrooke, Quebec to ensure a high number of species interactions occurring (Supplementary material Appendix 2, Table A1). The most western sites (QUBS) were collected at the beginning of the week with the eastern sites collected at the end of any given week.

Some sites were closer together than others, so there could have been a spatial bias and non-independence of the sites. Before running our main analysis, we ran spatial autocorrelation on overall parasite prevalence and intensity with each host species run separately and determined that all the sites were independent from one another. In addition, we interspersed bog and marsh sites across the transect in order to sample as many host species as possible. This likely had the observed effect of reducing potential autocorrelation in species abundance across the transect.

\subsection*{Host sampling}

Damselflies were collected weekly using aerial sweep nets by two collectors. At each site, one collector assessed host species assemblages by collecting individuals, while walking along the length of the shoreline for 20 min. Collecting all damselflies within this allotted time also allowed us to index...
relative local abundance of each damselflies species. After the initial 20 min collecting period, sampling continued for up to three hours or until 30 individuals of each species at each site were collected. All the damselflies were stored in 95% ethanol in individual vials. In the lab, damselflies were re-identified and examined for water mites.

Parasite sampling and measures of parasitism

Damselflies were examined for water mites using a ZEISS SteREO Discovery V8 dissecting microscope. All mites were quantified, collected and stored in 95% ethanol. We used (Rózsa et al. 2000) definitions of prevalence and mean intensity. Prevalence was defined as the proportion of host individuals of a particular species infected by at least one Arrenurus individual over the total number of hosts collected. Intensity was the number of Arrenurus individuals on an infected host. We measured the number of Arrenurus ‘species’ hereafter Operational Taxonomic Units (OTUs) per host species, using a molecular dataset. For most of the parameters, we pooled the data from all the sites because there were no significant differences in prevalence (F = 0.95; DF = 12, 80; p = 0.50) or intensity (F = 0.97; DF = 12, 80; p = 0.48) across sites. Analyses were performed in QP3.0 (Rózsa et al. 2000).

Composite host phylogeny

To examine potential importance of a host species’ evolutionary history to its likelihood of being parasitized by specific OTUs of water mites, we created a hypothesized phylogeny of Coenagrionidae collected in the marshes and bogs of southeastern Ontario and southwestern Quebec, by compiling recent phylogenies (Chippindale et al. 1999, Brown et al. 2000, May 2002, O’Grady and May 2003, Dumont et al. 2010). This approach allowed us to control for evolution of the characters among host species. Because molecular phylogenetic information is lacking for many of the damselfly species collected, taxonomy is a reasonable substitute to phylogeny (Crozier et al. 2005). We assumed that genera are monophyletic. We performed analyses of phylogenetic relatedness within the entire species assemblage by assigning the age of the node splits using BLADJ (branch length adjuster) module of Phylocom software package (Webb 2000). BLADJ places internal nodes where age estimates from phylogenetic trees were known. Age estimates were determined through the literature and discussion with experts. BLADJ also sets the remaining branch lengths by spacing the undated nodes in the tree evenly between those dated nodes. There are assumptions made, such as the true age of the splits, but this approach reduces variance in branch lengths and the nodal distances are based on evolutionary time and are independent of species richness and levels of resolution within clades.

Phylogeny of water mites

To address whether host species of sibling species pairs might share mite OTUs, and to have a preliminary account of host species range of those mite OTUs, a subset of water mites were sent to the Barcode of Life Database, BOLD; <www.boldsystems.org> (Ratnasingham and Hebert 2007) for barcoding using Cytochrome Oxidase I (COI). Randomly chosen mites were sent to the Canadian Centre for DNA Barcoding (CCDB; <www.ccdb.ca>) at the Biodiversity Inst. of Ontario to perform DNA barcoding using standard, high throughput methods (Ivanova et al. 2006). DNA extractions from 285 larval mites were performed. These larval water mites were selected randomly, based principally on host species and site of host collection. This random selection was as follows: if present, ten mites (five from the thorax and five from the abdomen because thoracic and abdominal mites are often different species (B. P. Smith pers. comm.)), from every host species at every site were collected. Each mite was haphazardly chosen from a different host individual if there were enough infected individuals. Twelve of sixteen host species had thoracic mites, and eight of sixteen species also had abdominal mites. In those latter species, both thoracic and abdominal water mites were collected, which is why sample size was initially 285.

Chromatograms were edited and contiguous sequences were assembled using Sequencher ver. 4.7 (Gene Codes, Ann Arbor, MI, USA). COI sequences were aligned manually in Mesquite ver. 2.74 (Maddison and Maddison 2011) according to the translated amino acid sequence. Sequences have been submitted to GenBank (KJ709142-KJ709343). Homologous sequences from four Arrenurus sp. individuals collected from Nehalennia gracilis and N. irene damselflies (Mlynarek et al. 2014) and four Enallagma ebrrium (Mlynarek et al. 2013) were included in the COI alignment, and seven water mite species were selected from GenBank to serve as outgroups (AB530314, JX838402, JX836526, JX835088, JN018105, JN018109, AB530311).

Pairwise distances were calculated using neighbour-joining analysis with the uncorrected (p’) model in PAUP* ver. 4.0b10 (Swofford 2003). Phylogenetic analysis of the COI dataset was performed using Bayesian inference (BI) in MrBayes ver. 3.1.2 (Hueslenbeck and Ronquist 2001, Ronquist and Huelsenbeck 2003). The best-fit model of molecular evolution was determined to be GTR + I + G, using MrModeltest ver. 2.3 (Nylander 2004). Bayesian analysis was performed in MrBayes with a Markov Chain Monte Carlo (MCMC) method, two independent runs, with nucmodel = 4by4, Nst = 6, rates = invgamma, samplefreq = 1000, four chains = one cold and three heated, 10 million generations. Excluding a 10% burn-in, the remaining trees were used to generate a majority-rule consensus tree.

Host parameters

We chose five host species’ characteristics that have been reasonably strong explanatory variables of other host–parasite associations (Krasnov et al. 2004, Poulin 2007).

Geographic range size (Range) was assessed using the Odonata Central database (Abbott 2007). The records in this database are verified and managed regularly by experts. To determine the size of a host’s geographic range, a 2° latitude–longitude grid was superimposed on the geographic data. Presence–absence was then observed for each cell. The area of the species distribution was determined by adding the
Area of each cell where that species was present (see Hassall 2012 for details).

Regional occurrence (RO) of a host species was determined simply as the number of sites a host species was collected at within our sampling season. Since we collected at 13 sites, the maximum regional occurrence a species can have is 13. Stated otherwise, if the regional occurrence of a host species is 13, that host was collected at all the sites (if 12, at 12 of 13 potential sites, etc.).

We assessed relative local abundance by collecting damselflies at a site for 20 min. From this, we could gauge how rich the host species assembly was at a given site and the relative abundance of each host species there. Local abundance for our study is the proportion of a particular host species over the number of host species collected within 20 min. Because we collected at each site eight times, the relative local abundance is the average local abundance at each site for eight weeks. Our final measure for analysis is the mean local abundance (MLA). This is the grand mean of all the relative local abundances from all the sites.

We assessed phenology of a host species using species accumulation curves (Supplementary material Appendix 1, Fig. A1). We observed the week where we collected 50% of the total individuals of that species. Once the peak of activity for all the species was assessed, we determined peak of damselfly activity by calculating the mean of all the weeks where 50% of the population of a species. We then took, for each host species, the absolute value of the difference between the 50% week and the mean of all the weeks. Through this, we could assess whether the peak flight season of a host species was near the peak of damselfly activity across all species.

Using digital calipers, we measured wing length from the base to the tip of the right forewing, as a proxy for host body size. In this case, our measure of wing length is the mean of wing length for each damselfly species from all individuals of that species across sites.

Statistics

We evaluated the data using phylogenetically generalized least squares (PGLS), which was implemented using the package caper in R (Orme et al. 2012). Fixed effect covariates included size of host distribution (Range), regional occurrence (RO), phenology (Phenology), grand mean local abundance (MLA) and wing length (WingL). Model selection was based on Akaike's information criterion corrected for sample size (AICc) (Burnham and Anderson 2002). Regression coefficients (i.e. β-values on a logit scale) were estimated by model averaging and statistical significance was judged on the basis of sign (positive or negative relationship between the predictor and response variables) and the precision of estimates (wherein values with 95% confidence intervals that did not overlap with zero were considered important predictors). Our candidate set of models included all additive combinations of the predictor variables. To avoid over-parameterization, we did not evaluate interactions between variables.

For our analyses of prevalence and intensity, we excluded the species with sample sizes less than 50 individuals because of potential sampling bias. Based on power analysis, 30–50 individuals of a host species provided appropriate measures of parasitism for statistical comparison.

Results

Measures of general Arrenurus parasitism

A total of 7107 damselflies were collected (Table 1). Host abundance per species per site per week ranged from 61 damselfly individuals collected in Osprey Marsh to 219 host individuals collected in Stony Swamp. There was variation in the cumulative date of emergence of the damselfly species across the eight weeks (Supplementary material Appendix 1, Fig. A1). Based on our analysis, early species included Coenagrion interrogatum, Coenagrion resolutum and Enallagma cyathigerum. We collected at least 50% of the individuals of these species in the first week of collecting. There are two species, Enallagma aspersum and Enallagma geminatum that emerged late in our collecting. No individuals of these species were collected until the sixth week of collecting (Supplementary material Appendix 1, Fig. A1).

There was also a lot of variation of Arrenurus prevalence and intensity between the host species. Prevalence of parasitism varied between 0% in Amphigargion saucium, Chromagrion conditum, Coenagrion interrogatum and Enallagma geminatum to 37.4% in Enallagma ebrarium (Table 1). For the infected species, mean intensity varied between 1.3 Arrenurus water mites in Enallagma carunculatum and 13.3 in Ischnura verticalis (Table 1).

Host phylogeny

We collected 16 damselfly species across all sites which were included in our phylogeny of the Coenagrionidae in marshes and bogs of southeastern Ontario and southwestern Quebec (Fig. 2), of which four species (25%) were collected only at one site and two species (12.5%) occurred at all sites (Table 1). All subsequent statistical analyses were run using this composite phylogeny to control for a potential evolutionary signal.

Parasite phylogeny

A 683bp fragment of COI was amplified from 217 Arrenurus specimens, a total of 327 characters were parsimony informative and 356 were uninformative.

Bayesian inference was performed on the COI dataset for 10 million generations. After burn-in there were 18002 trees, which were summarized in a 50% majority rule consensus tree (TL = 1370, CI = 0.396, RI = 0.927). This tree was well supported, with most nodes having moderate to high posterior probabilities (Fig. 3). The ingroup was divided into 15 well-supported clades, OTUs. Average interclade divergence was 10.7% ± 5.8, and average intraclade divergence was 0.7% ± 0.5. Average divergence between OTUs 1 and 2 was 16.6% ± 0.6, divergence between OTUs 3, 4, 5 and 6 was 7.9% ± 3.2, between OTUs 8, 9, and 10 was 7.5% ± 1.6, and OTUs 11, 12, 13, 14 and 15 show 9.0% ± 3.2 divergence. Considering the high level of divergence between
Table 1. Host parameters of the 16 damselfly species that were collected during the reproductive season of 2011 from 13 sites in southeastern Ontario and southwestern Quebec. N = sample size of each species; RLA = relative local abundance; RO = regional occurrence; OTUs = number of operational taxonomic units; WingL = wing length.

<table>
<thead>
<tr>
<th>Host species</th>
<th>N  (N inf.)</th>
<th>Prevalence (%)</th>
<th>Intensity</th>
<th>Phenology</th>
<th>Range (10^6 km^2)</th>
<th>RLA</th>
<th>RO</th>
<th>WingL</th>
<th>OTUs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amphiagrion saucium</td>
<td>8 (0)</td>
<td>0.0 (0–36.9)</td>
<td>0.0 (NA)</td>
<td>0.44</td>
<td>1.81</td>
<td>0.05</td>
<td>1</td>
<td>16.03</td>
<td>0</td>
</tr>
<tr>
<td>Chromagrion conditum</td>
<td>133 (0)</td>
<td>0.0 (0–2.7)</td>
<td>0.0 (NA)</td>
<td>1.44</td>
<td>2.21</td>
<td>0.09</td>
<td>2</td>
<td>21.62</td>
<td>0</td>
</tr>
<tr>
<td>Coenagrion interrogatum</td>
<td>52 (0)</td>
<td>0.0 (0–6.8)</td>
<td>0.0 (NA)</td>
<td>2.44</td>
<td>2.35</td>
<td>0.05</td>
<td>1</td>
<td>18.90</td>
<td>0</td>
</tr>
<tr>
<td>Coenagrion resolutum</td>
<td>542 (15)</td>
<td>2.6 (1.4–4.3)</td>
<td>2.1 (1.4–2.9)</td>
<td>2.44</td>
<td>6.13</td>
<td>0.04</td>
<td>11</td>
<td>18.21</td>
<td>1</td>
</tr>
<tr>
<td>Enallagma aspersum</td>
<td>11 (4)</td>
<td>36.4 (10.9–69.2)</td>
<td>3.5 (2.3–5.3)</td>
<td>1.56</td>
<td>2.58</td>
<td>0.05</td>
<td>1</td>
<td>19.87</td>
<td>3</td>
</tr>
<tr>
<td>Enallagma boreale</td>
<td>242 (32)</td>
<td>12.8 (8.9–17.7)</td>
<td>5.9 (4.5–7.7)</td>
<td>2.44</td>
<td>6.63</td>
<td>0.04</td>
<td>6</td>
<td>20.00</td>
<td>3</td>
</tr>
<tr>
<td>Enallagma carunculatum</td>
<td>123 (3)</td>
<td>2.4 (0.5–7.0)</td>
<td>1.3 (1–1.7)</td>
<td>2.56</td>
<td>5.20</td>
<td>0.10</td>
<td>3</td>
<td>20.56</td>
<td>?</td>
</tr>
<tr>
<td>Enallagma cyathigerum</td>
<td>128 (2)</td>
<td>1.6 (0.2–5.5)</td>
<td>3.0 (1–4)</td>
<td>3.44</td>
<td>6.50</td>
<td>0.04</td>
<td>6</td>
<td>19.81</td>
<td>3</td>
</tr>
<tr>
<td>Enallagma ebrum</td>
<td>1000 (374)</td>
<td>37.4 (34.3–40.5)</td>
<td>10.5 (9.2–11.9)</td>
<td>0.56</td>
<td>4.35</td>
<td>0.14</td>
<td>12</td>
<td>18.09</td>
<td>7</td>
</tr>
<tr>
<td>Enallagma geminatum</td>
<td>100 (0)</td>
<td>0.0 (0–70.8)</td>
<td>0.0 (NA)</td>
<td>2.56</td>
<td>2.98</td>
<td>0.04</td>
<td>1</td>
<td>18.48</td>
<td>0</td>
</tr>
<tr>
<td>Enallagma hageni</td>
<td>257 (51)</td>
<td>19.8 (15.1–25.3)</td>
<td>5.8 (4.3–8.3)</td>
<td>0.56</td>
<td>4.25</td>
<td>0.05</td>
<td>8</td>
<td>17.81</td>
<td>4</td>
</tr>
<tr>
<td>Enallagma signatum</td>
<td>27 (5)</td>
<td>18.5 (6.3–38.1)</td>
<td>5.2 (2.8–7.2)</td>
<td>2.56</td>
<td>3.23</td>
<td>0.07</td>
<td>2</td>
<td>20.25</td>
<td>2</td>
</tr>
<tr>
<td>Ischnura posita</td>
<td>336 (18)</td>
<td>5.4 (3.2–8.3)</td>
<td>1.7 (1.3–2.5)</td>
<td>1.56</td>
<td>3.13</td>
<td>0.05</td>
<td>9</td>
<td>15.92</td>
<td>3</td>
</tr>
<tr>
<td>Ischnura verticalis</td>
<td>349 (85)</td>
<td>24.4 (19.9–29.2)</td>
<td>13.3 (9.9–18.4)</td>
<td>0.44</td>
<td>4.36</td>
<td>0.06</td>
<td>13</td>
<td>16.27</td>
<td>6</td>
</tr>
<tr>
<td>Nehalennia gracilis</td>
<td>820 (64)</td>
<td>7.8 (6.1–9.9)</td>
<td>2.2 (1.8–2.8)</td>
<td>0.56</td>
<td>1.51</td>
<td>0.25</td>
<td>5</td>
<td>14.88</td>
<td>2</td>
</tr>
<tr>
<td>Nehalennia irene</td>
<td>3076 (395)</td>
<td>12.8 (11.7–14.5)</td>
<td>3.3 (3.0–3.6)</td>
<td>0.56</td>
<td>4.03</td>
<td>0.48</td>
<td>13</td>
<td>15.60</td>
<td>4</td>
</tr>
</tbody>
</table>

Figure 2. Composite phylogeny of the Coenagrionidae found at the thirteen sites of southeastern Ontario and southwestern Quebec including the number of Arelnurus OTUs collected from each host species.
Figure 3. Majority rule consensus tree of 18002 trees generated by Bayesian MCMC analysis (10 million generations) of 683 bp fragment of COI from 224 water mites, 217 ingroup and 7 outgroup specimens (TL = 1370, CI = 0.396, RI = 0.927).
OTUs and the low divergence within each clade, as well as the strong support for each clade in the phylogenetic reconstruction, it appears that each OTU may be distinct. We are not implying they are new species and the possibility still exists that there are many more species present if speciation (and adaptive radiation) has occurred recently.

**Host species characteristics as predictors**

Our PGLS examining variation in the prevalence of water mite parasitism in relation to host species characteristics indicated that Range and Phenology were important predictor variables \((\log(L) = -38.31, K = 3, AICc = 84.61, \omega_i = 0.37)\); several other models also were supported by the data \((\Delta AICc < 4)\; \text{Table 2})

Model averaged regression coefficients (Supplementary material Appendix 2, Table A2) showed a negative association between water mite prevalence and phenology \((\beta_{\text{Phenology}} = -10.10, 95\% \text{ CI: } -14.91 \text{ to } -5.29; \text{Fig. 4a})\) and a unimodal association with range \((\beta_{\text{Range}} = 2.75, 95\% \text{ CI: } 0.82 \text{ to } 4.68; \text{Fig. 4b})\). Host body size and local abundance had little predictive power (Supplementary material Appendix 2, Table A2).

The best fitting model in our candidate set of PGLS investigating water mite intensity indicated that again range and phenology were significant predictor variables \((\log(L) = -27.35, K = 3, AICc = 63.71, \omega_i = 0.31)\).

Additionally, other models were supported by the data \((\text{AICc} < 4)\; \text{Table 3})

Model averaged regression coefficients (Supplementary material Appendix 2, Table A3) indicated a negative association with phenology \((\beta_{\text{Phenology}} = -2.46, 95\% \text{ CI: } -3.81 \text{ to } -1.11; \text{Fig. 5a})\) and a unimodal association with range \((\beta_{\text{Range}} = 1.05, 95\% \text{ CI: } 0.39 \text{ to } 1.71; \text{Fig. 5b})\). Host body size and local abundance had little predictive power (Supplementary material Appendix 2, Table A3).

Our PGLS examining OTU richness in host species in relation to host species characteristics indicated that regional occurrence as the only significant predictor variable \((\log(L) = -27.29, K = 2, AICc = 59.71, \omega_i = 0.29)\); several other models also were supported by the data \((\Delta AICc < 4); \text{Table 4})

Model averaged regression coefficients (Supplementary material Appendix 2, Table A4) indicated a positive association between regional occurrence and OTU richness \((\beta_{\text{RO}} = 0.362, 95\% \text{ CI: } 0.19 \text{ to } 0.53; \text{Fig. 6})\). Host body size and local abundance had little predictive power (Supplementary material Appendix 2, Table A4).

**Discussion**

The goal of this study was to determine what host species' characteristics predict the levels of parasitism (prevalence and intensity) and the diversity of parasite interactions. Based on our results, it is clear that host distribution characteristics

---

**Table 2.** Model selection results for PGLSs evaluating water mite prevalence in relation to host distribution attributes. Table includes the global model and all models with AICc < 4. Calculations are described in the footnote1.

<table>
<thead>
<tr>
<th>Model</th>
<th>Log(L)</th>
<th>K</th>
<th>AICc</th>
<th>Δi</th>
<th>Model likelihood</th>
<th>Model prob.</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range + Phenology</td>
<td>-38.31</td>
<td>3</td>
<td>84.61</td>
<td>0.00</td>
<td>1.000</td>
<td>0.365</td>
<td>0.66</td>
</tr>
<tr>
<td>Range + MLA + Phenology</td>
<td>-37.35</td>
<td>4</td>
<td>86.34</td>
<td>1.73</td>
<td>0.421</td>
<td>0.154</td>
<td>0.67</td>
</tr>
<tr>
<td>RO + Phenology</td>
<td>-39.46</td>
<td>3</td>
<td>86.91</td>
<td>2.29</td>
<td>0.317</td>
<td>0.116</td>
<td>0.67</td>
</tr>
<tr>
<td>Phenology</td>
<td>-41.15</td>
<td>2</td>
<td>87.22</td>
<td>2.61</td>
<td>0.271</td>
<td>0.099</td>
<td>0.60</td>
</tr>
<tr>
<td>Range + Phenology + Wingl.</td>
<td>-38.05</td>
<td>4</td>
<td>84.77</td>
<td>3.13</td>
<td>0.209</td>
<td>0.076</td>
<td>0.63</td>
</tr>
<tr>
<td>RO + Range + Phenology</td>
<td>-38.28</td>
<td>4</td>
<td>88.19</td>
<td>3.58</td>
<td>0.167</td>
<td>0.061</td>
<td>0.61</td>
</tr>
<tr>
<td>RO + Phenology + Wingl.</td>
<td>-38.36</td>
<td>4</td>
<td>88.35</td>
<td>3.74</td>
<td>0.154</td>
<td>0.056</td>
<td>0.61</td>
</tr>
<tr>
<td>RO + Range + MLA + Phenology + Wingl.</td>
<td>-37.22</td>
<td>6</td>
<td>103.24</td>
<td>17.62</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>0.57</td>
</tr>
</tbody>
</table>

1K is the number of estimable parameters; \(\Delta i = \text{AICc} - \text{AICcmin}^{\text{Global model}}\); the model likelihood given the data is \(\exp(-0.5 \times \Delta \text{AICc})\); model probability estimated as \(\omega_i = \frac{\exp(-\Delta i)}{\sum \exp(-\Delta i)}\); \(\text{R}^2\) estimated from PGLS.

---

Figure 4. Summary of the important predicting factors explaining *Arrenurus* spp. prevalence. (a) Phenology and (b) range in relation to *Arrenurus* spp. prevalence. Cc = *Chromagrion conditum*; Ci = *Coenagrion interrogerum*; Cr = *C. resolutum*; Eb = *Enallagma boreale*; Ec = *E. carunculatum*; Ecy = *E. cyathigerum*; Ee = *E. ebrarium*; Eh = *E. hageni*; Ip = *Ichthura positus*; Iv = *I. verticalis*; Ng = *Nehalennia gracilis*; Ni = *N. irene*.
can predict damselfly–water mite associations. Differences in prevalence and intensity of *Arrenurus* were explained mostly by two host characteristics, range and phenology, as measured in this study. The differences in number of OTUs can be predicted by a host species’ regional occurrence.

Host geographic range and phenology are important in predicting prevalence and intensity. A positive relationship between size of host geographic range and the prevalence and intensity of parasitism has previously been reported. Price et al. (1988) proposed the geographic range hypothesis with the prediction that more widespread hosts will be under higher parasite pressure. Tella et al. (1999) supported this hypothesis in a study of diurnal birds of prey parasitized by hematozoan blood parasites where they proposed that the pattern was potentially due to the evolutionary history of specific host–parasite associations. Similarly, the prediction of the geographic range hypothesis was supported in a study comparing two species of *Calopteryx* in Finland (Ilvonen et al. 2011). In contrast, the geographic range hypothesis was not supported in two studies comparing gregarine and water mite parasitism in several pairs of closely related coenagrionid damselflies where the less widespread species of the species pairs had higher parasitism levels (Mlynarek et al. 2011, 2013). In the current study based on an increased number of non-phylogenetically paired host species, we find some support for the geographic range hypothesis. In our results, there is a trend towards a unimodal pattern where the species with the smallest and largest geographic distributions have the lowest prevalence and intensity of *Arrenurus* parasitism. This trend is mostly driven by two species, *I. verticalis* and *E. ebrium*, these are species that are very common at all sites and have the highest numbers of parasite species.

Geographic range alone was not able to predict the prevalence and intensity of parasitism. Host phenology was a strong predictor of prevalence and intensity in *Arrenurus* spp. parasitism. In systems where hosts demonstrate cyclical life history traits, such as the yearly life cycle of many insects, infection levels are expected to vary with time within that cycle (Forbes et al. 2012). This has been often demonstrated in infectious diseases (Altizer et al. 2006), bees infected by internal gregarines (Zuk 1987, Locklin and Vodopich 2010). As previous studies (Gillespie 2010), a seasonal unimodal pattern of parasitism was demonstrated in infection levels. This seasonal pattern could be linked to host density where host density may be either the size of the population of one host species (Forbes et al. 2012) or the aggregate density of all hosts in the species assemblage (Gillespie 2010). There is the expectation that parasites time their life cycle to coincide with that of their hosts to increase their chances of finding a host; in this case, the generalist parasite will track when there is the highest density of potential hosts or the peak of the host flight season.

Even if generalist parasites can survive on many host species they demonstrate preferences in host use (Triplet and Richner 1997). In this study, generalist parasite species richness or number of OTUs infecting each damselfly species, was not predicted by host range size or phenology but only by

Table 3. Model selection results for PGLSs evaluating water mite mean intensity in relation to host distribution attributes. Table includes the global model and all models with AICc < 4. Calculations are described in the footnote of Supplementary material Appendix 2, Table A1.

<table>
<thead>
<tr>
<th>Model</th>
<th>Log(L)</th>
<th>K</th>
<th>Δi</th>
<th>Model likelihood</th>
<th>Model prob.</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range + Phenology</td>
<td>-27.35</td>
<td>3</td>
<td>63.71</td>
<td>0.000</td>
<td>1.000</td>
<td>0.307</td>
</tr>
<tr>
<td>Range + MLA + Phenology</td>
<td>-25.21</td>
<td>4</td>
<td>64.14</td>
<td>0.436</td>
<td>0.804</td>
<td>0.247</td>
</tr>
<tr>
<td>RO</td>
<td>-29.90</td>
<td>2</td>
<td>65.13</td>
<td>1.417</td>
<td>0.492</td>
<td>0.151</td>
</tr>
<tr>
<td>RO + MLA</td>
<td>-28.89</td>
<td>3</td>
<td>66.78</td>
<td>3.068</td>
<td>0.216</td>
<td>0.066</td>
</tr>
<tr>
<td>RO + Phenology</td>
<td>-29.30</td>
<td>3</td>
<td>67.60</td>
<td>3.889</td>
<td>0.143</td>
<td>0.044</td>
</tr>
<tr>
<td>RO + Range + MLA + Phenology</td>
<td>-25.07</td>
<td>6</td>
<td>78.93</td>
<td>15.23</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Figure 5. Summary of the important predicting factors explaining *Arrenurus* spp. mean intensity. (a) Phenology and (b) range in relation to *Arrenurus* spp. mean intensity. *Cc* = *Chromagrion conditum*; *Ci* = *Coenagrion interrogation*; *Cr* = *C. resolutum*; *Eb* = *Enallagma boreale*; *Ec* = *E. carunculatum*; *Ecy* = *E. cyathigerum*; *Ee* = *E. ebrium*; *Eb* = *E. hageni*; *Ip* = *Ischnura posita*; *Iv* = *I. verticalis*; *Ng* = *Nehalennia gracilis*; *Ni* = *N. irene.*
Table 4. Model selection results for PGLs evaluating water mite species richness (number of OTUs per host species) in relation to host distribution attributes. Table includes the global model and all models with AICc < 4. Calculations are described in the footnote of Supplementary material Appendix 2, Table A1.

<table>
<thead>
<tr>
<th>Model</th>
<th>Log(L)</th>
<th>K</th>
<th>AICc</th>
<th>Δi</th>
<th>Model likelihood</th>
<th>Model prob.</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>RO</td>
<td>-27.39</td>
<td>2</td>
<td>59.71</td>
<td>0.00</td>
<td>1.000</td>
<td>0.291</td>
<td>0.55</td>
</tr>
<tr>
<td>RO + Phenology</td>
<td>-26.19</td>
<td>3</td>
<td>60.38</td>
<td>0.67</td>
<td>0.715</td>
<td>0.208</td>
<td>0.61</td>
</tr>
<tr>
<td>RO + Range</td>
<td>-26.60</td>
<td>3</td>
<td>61.19</td>
<td>1.48</td>
<td>0.477</td>
<td>0.139</td>
<td>0.59</td>
</tr>
<tr>
<td>RO + Phenology + WingL</td>
<td>-25.23</td>
<td>4</td>
<td>62.09</td>
<td>2.37</td>
<td>0.304</td>
<td>0.089</td>
<td>0.62</td>
</tr>
<tr>
<td>RO + WingL</td>
<td>-27.23</td>
<td>3</td>
<td>62.46</td>
<td>2.75</td>
<td>0.252</td>
<td>0.073</td>
<td>0.52</td>
</tr>
<tr>
<td>RO + MLA</td>
<td>-27.30</td>
<td>3</td>
<td>62.61</td>
<td>2.89</td>
<td>0.235</td>
<td>0.068</td>
<td>0.52</td>
</tr>
<tr>
<td>RO + Range + Wingl</td>
<td>-25.61</td>
<td>4</td>
<td>62.85</td>
<td>3.13</td>
<td>0.208</td>
<td>0.061</td>
<td>0.63</td>
</tr>
<tr>
<td>RO + Range + MLA + Phenology + WingL</td>
<td>-24.60</td>
<td>6</td>
<td>70.54</td>
<td>10.83</td>
<td>0.004</td>
<td>0.001</td>
<td>0.60</td>
</tr>
</tbody>
</table>

host regional occurrence. Host species present at more sites were infected by more OTUs. Durrer and Schmid-Hempel (1995) found similar results when comparing bumblebee species infected by all their parasite taxa. One explanation could be the fact that a species is present at more sites it will encounter more parasite species. We propose that the influence of host regional occurrence may be explained by the fact that widespread hosts have not adapted species-specific recognition at particular sites and resistance or evasion does not occur. If a host species is cosmopolitan there may be more gene flow between its open populations, which could maintain their susceptibility to parasitism (Forbes and Baker 1991). There are other studies that show the opposite; there was a lack of support for a relationship between ectoparasites infecting flying squirrels and host size (Perez-Orella and Schulte-Hostedde 2005). Similarly in our study, host size did not predict parasitism levels. The other factor that has little effect was host local abundance. This was also unexpected because a relationship between local abundance and parasite levels has been previously reported (Arneberg et al. 1998, Krasnov et al. 2002). We would expect that more locally abundant species would be stronger targets of selection by the parasites. Local adaptation is expected to evolve in the parasites of locally abundant hosts (Arneberg et al. 1998). Even though host abundance and host size were expected to predict host parasitism, in this system, the better predictors are ones that are measured at larger scales.

Even though hosts were collected at multiple sites within a region, we superimposed the parasitism over the entire range considering it a species’ characteristic, which is common in evolutionary ecology studies (Heff erman et al. 2014). The importance of scale has long been understood to be critical in ecology (Wiens 1989, Levin 1992) but it has been tested only infrequently in empirical studies of parasitism (Tack et al. 2014). Different measures of species’ characteristics and species’ interactions might depend on different geographic or temporal scales (Heff erman et al. 2014) thus underscoring the need of measuring host species’ characteristics of different scales for a better understanding of the relationship between the host and parasite species assemblages.

In this study, we primarily looked at Coenagrionidae–Arrenuridae host–parasite associations from the perspective of host species to explain three measures of parasitism; prevalence, intensity and diversity. As in other host–parasite associations, host characteristics do matter (Poulin 2007). This study is one of the few in invertebrate host systems that elucidate the reasons for the strength of host–parasite associations, based on host characteristics measured at multiple scales. Further research into host–parasite associations in this context including understanding turnover of principal host species and host ranges across space and time would be logical next steps in understanding the evolution of the associations between hosts and their parasites.

Figure 6. Summary of the important predicting factor explaining number of OTUs in each damselfly species. Host regional occurrence in relation to Arrenurus OTUs. As = Amphigrion saucium; Cc = Chromagrion conditum; Ci = Coenagrion interruptum; Cr = C. resolutum; Ea = Enallagma aopersum; Eb = E. boreale; Ec = E. carunculatum; Ey = E. cyathigerum; Ec = E. ebrium; Eg = Enallagma geminatum; Eh = E. hageni; Es = Enallagma signatum; Ip = Ischnura posita; Iv = I. verticalis; Ng = Nebalennia gracilis; Ni = N. irene.
Acknowledgements – We thank L. Nagel, A. Morrill and S. Iverson for discussions on this and related projects, V. Levesque-Beaudin for aid in the barcoding process and M. McPeek and G. Chelby for helping with the reconstruction of the Coenagrionidae Phylogenetic tree. This study was funded by a NSERC CGS awarded to JJM and a NSERC Discovery Grant to MRF. JJM and MRF designed the study. JJM collected and analysed the statistical data. JJM and WK analysed the molecular data. JJM wrote the paper. JJM, WK and MRF edited the paper.

References
Nylander, J. A. A. 2004. MrModeltest v2. – Program distributed by the author, Evolutionary Biology Centre, Uppsala Univ.


