***LOCAL ENVIRONMENT AND SAMPLING BIAS DRIVE PARASITE PREVALENCE ESTIMATES IN FRESHWATER FISH COMMUNITIES***

**ABSTRACT**

Parasite occurrence and infection estimates vary through time and space, making understanding the underlying drivers highly complex. Comparative studies based on empirical data must consider the factors of variation involved in estimating infection metrics in natural populations to make appropriate and reliable comparisons. Using a multi-scale approach, we explored the sources of variation in the estimation of infection prevalence, focusing on black spot disease in littoral freshwater fish communities sampled across 15 lakes in Québec, Canada. Method-related sampling biases led to significant variations in prevalence estimates and spatial patterns of disease occurrence. Our results also indicated that low sampling efforts tend to overestimate the prevalence of infection in the landscape, with minimum sampling effort required to estimate an accurate infection prevalence depending on the sampling method employed. Our results showed that infection prevalence is spatially heterogeneous across the landscape with evidence of infection hotspots and coldspots. Physico-chemical characteristics of the sites and local fish community structure were found to be the best drivers of infection at smaller spatial scales. Furthermore, our results suggest dilution effects, due to physical obstruction and compatibility barriers, limit the survival of the free-living cercaria parasite life stage. Several relationships between infection prevalence and environmental drivers revealed non-linearity, suggesting complex interactions. Examining infection prevalence data at various spatial scales revealed method-induced biases, sampling effort effect and environment driven relationships underscoring the importance of context-dependencies and scale-dependencies in empirical studies on host-parasite interactions.

**KEYWORDS**

Disease ecology; distribution patterns; host-parasite interactions; infection predictors, infection parameters; sampling design; sampling methods.

## INTRODUCTION

Despite often being neglected in ecological studies (Morley, 2012; Scholz and Choudhury, 2014; Chrétien et al., 2023), parasites are key components of communities and ecosystems (Dobson and Hudson, 1986; Minchella and Scott, 1991; Poulin, 1999; Marcogliese, 2004; Lafferty et al., 2008; Frainer et al., 2018; Pascal et al., 2020). Parasites can induce physiological and behavioral changes in their hosts (Barber et al., 2000; Iwanowicz, 2011) and are major elements of food webs (Marcogliese and Cone, 1997; Lafferty et al., 2006, 2008). However, studying infection dynamics is challenging since infection patterns within a population or community can vary across taxa, time and/or space (Poulin, 2006; Poulin and Dick, 2007; Thieltges and Reise, 2007; Young and Maccoll, 2017; Villalba-Vasquez et al., 2018; Happel, 2019). Since data collection is constrained by context-dependencies such as spatial scaling, system characteristics and sampling design, it is imperative to understand how these elements influence our understanding of infection patterns to facilitate comparisons among studies, especially in the context of anthropogenic change and conservation management (Cohen et al., 2016). However, parasitological ecology literature concerning these issues is currently limited.

Spatial scale-dependance of ecological processes has long been established (Rietkerk et al., 2002; Viana and Chase, 2019). Thus, single-scale studies tend to overlook the full extent of ecological processes shaping species distribution and interactions, leading to biased conclusions (Levin, 1992; Peterson and Parker, 1998). This issue has also been highlighted in disease ecology. For example, Cohen et al. (2016) found that the relative importance of processes driving the distribution of disease emergence (host richness, abiotic factors, and human population density) varied depending on the spatial extent of the study. Moreover, regional-scale infection patterns can be caused by different underlying local patterns of infection. For instance, infection prevalence is a common parasitological metric that gives information on the proportion of infected individuals in a given group, and can be measured for both populations and communities (Bush et al., 1997; Rózsa et al., 2000). A regional prevalence value of 20% could be caused by all sites having the same prevalence with no variance (i.e., all sites at 20%) or by the average prevalence of all sites being at 20% with large inter-site variance (e.g., half at 0% and half at 40%). Given the scale-dependence of ecological processes (Rietkerk et al., 2002; Viana and Chase, 2019) and the aggregated nature of parasites (Poulin 2007), multi-scale studies are thus essential for understanding how infection is distributed across a landscape.

Investigating a wide range of drivers can help reveal processes underlying infection clusters (i.e., infection hotspots) within a system that appears homogenous (Bolnick et al., 2020). Individual host characteristics such as sex (Zuk and McKean, 1996; Kowalski et al., 2015), age/size (Poulin, 2000; Marcogliese et al., 2001; Kowalski et al., 2015), genetics (Williams-Blangero et al., 2012) and personality traits (Hart, 1990; Barber and Dingemanse, 2010; Gradito et al., 2024) can drive host susceptibility to infection leading to variations in infection metrics across host populations. For example, male threespine stickleback (*Gasterosteus aculeatus*) typically exhibit higher infection prevalence than females (Reimchen and Nosil, 2001). On the other hand, host community metrics such as species richness, diversity and abundance can create a “dilution effect” by reducing a parasite’s encounter rates with target hosts (Dargent et al., 2013; Civitello et al., 2015; Lagrue and Poulin, 2015; Buck and Lutterschmidt, 2017; Ahn and Goater, 2021). For instance, a higher non-host (incompetent) abundance might reduce prevalence estimates in populations and communities because unsuccessful infection attempts by infective stage on non-host species (compatibility barrier) reduce parasite energy budgets, shorten the infective temporal window or kill infective stages (Johnson and Thieltges, 2010). Non-hosts may also induce a behavior change in susceptible hosts that reduces encounter rate with infective parasite life stages. For example, Ahn and Goater (2021) suggested that behavioral changes occur in fathead minnows (*Pimephales promelas*; host) when they co-occur with emerald shiners (*Notropis atherinoides*; non-host) reducing their likelihood of being infected by brain-encysting trematodes. Several local habitat characteristics (e.g., waterbed morphometry, habitat structure, water physico-chemistry) have also been associated with infection patterns. In American eels (*Anguilla rostrata*), parasite diversity decreases when the water pH is below 5.4, with some parasite families such as digenean trematodes being absent below pH 4.7 (Marcogliese and Cone, 1997). Furthermore, spatial features of an ecosystem (e.g., watershed, connectivity) act as large scale filtering by limiting dispersal of both parasites and hosts (Bolnick et al., 2020) thus creating infection clustering at the landscape level. For instance, Bolnick et al. (2020) found that increased distance between waterbodies leads to greater differences in parasite community composition in *G. aculeatus*. Thus, a complete ecological perspective on host-parasite systems must incorporate abiotic, biotic and spatial elements to highlight mechanisms shaping patterns of infection (Cohen et al., 2016; Bolnick et al., 2020).

Although the influence of sampling method and/or sampling effort on the accuracy of population parameter estimates (e.g., abundance, recruitment, age classes, sex-ratio) is widely acknowledged by the scientific community, how these components influence infection metric estimates is less understood and largely overlooked. However, wildlife sampling methods often rely on animals’ risk-driven decision, and devices that require a high level of interaction with the animal for detection or capture it (e.g., a traps versus cameras), tend to sample fewer, bolder individuals introducing a “personality bias” in the sampled population (Johnstone et al., 2021). For example, Biro and Post (2008) noted that bold rainbow trout were caught three times as often as shy individuals in gillnets. Consequently, bias introduced by sampling method may result in inaccurate estimates of community metrics. Nusser et al. (2008) demonstrated that the survey method used to estimate the prevalence of wasting disease in deer can either over or underestimate the true infection parameter. A decrease in sampling effort (i.e., number of sampling attempts) can also influence detection probabilities (de Solla et al., 2005) and the estimation of demographic parameters such as species occurrence and abundance (Symons et al., 2018). Consequently, sampling effort might influence infection metrics, like prevalence of infection, that rely on infection detectability and host-population abundance.

Here, we explore environmental and sampling sources of variation in infection prevalence estimates focusing on trematode parasites in littoral fish communities across 15 lakes varying in morphometric attributes and local biotic and abiotic conditions. We analyze infection prevalence data at the fish-community level (all fish hosts) for context-dependencies across three spatial scales (landscape-, lake- and site-scale) in order to investigate (i) the effect of increasing sampling effort on landscape prevalence estimates across different sampling methods (ii) the distribution lake prevalence estimates across the landscape through different sampling methods and (iii) the importance of ecological drivers on the variation of site-scale prevalence estimates.

At the landscape-scale, if infection prevalences are distributed homogeneously across the landscape, the shape of the curve defining the relationship between prevalence and sampling effort should be flat (Figure 1a). However, if prevalences are patchy with many coldspots and few hotspots distributed in the landscape, an underestimation of prevalences should occur at low sampling effort. Conversely, many hotspots with few coldspots should result in overestimating prevalence at low sampling effort. The sampling method should also influence the number of samples needed to accurately estimate prevalence (Figure 1b). At lake-scale, we expect that prevalence will vary across lakes following a heterogeneous distribution, and that frequency distribution patterns will vary among sampling methods (Figure 1c–f). At site-scale, the shapes and strength of the relationships between infection metrics and drivers will help elucidate which conditions influence prevalence variations. Our results provide insights into how infection prevalence estimates behave across multiple spatial scale with special interest in context-dependencies influencing estimates such as local environment and sampling design.

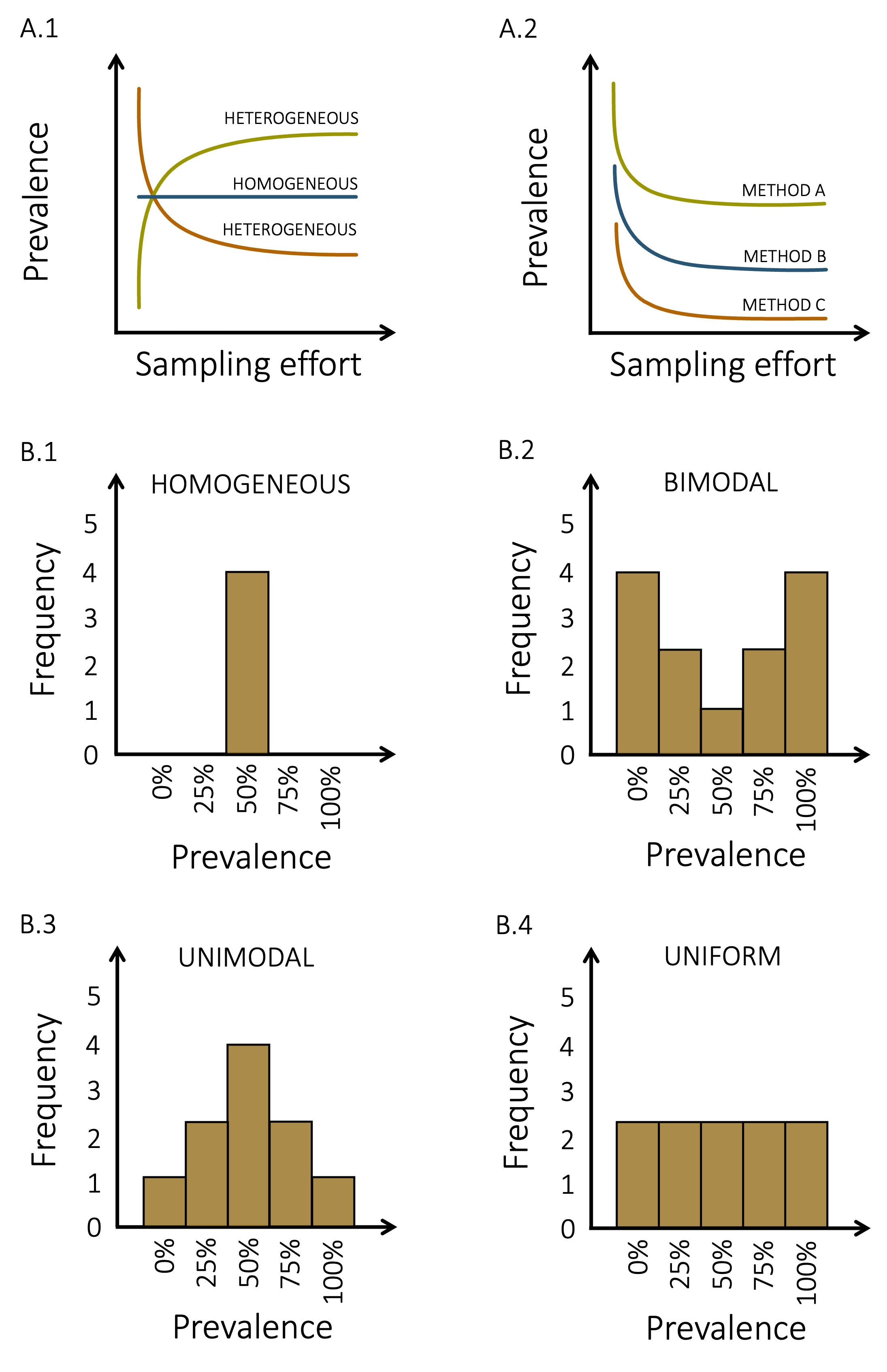


Figure 1. Case scenarios of (A) sampling effort effect on landscape prevalence estimates and (B) frequency distributions of lake prevalence estimates. A.1) Prevalence could be homogeneous (blue line) or heterogeneous (green and brown lines) across the landscape. In heterogeneous scenarios, prevalence is either overestimated (brown line) or under underestimated (green line) before stabilizing around the accurate prevalence. A.2) Sampling methods could lead to different prevalence estimates across the landscape. Prevalence difference between stabilized lines would represent method biases. B.1) All lakes could be infected at the same prevalence level. B.2) Lake’s frequency distribution could follow a bimodal pattern where low and high prevalence are more common across the landscape. B.3) Lake’s frequency distribution could follow a unimodal pattern where mean landscape prevalence is more common. B.4) Lake’s frequency distribution could follow a uniform pattern where all infection levels are equally represented.

## MATERIAL AND METHODS

### Host-parasite system

The black spot disease is a common infection in marine and freshwater fishes caused by Digenean trematodes (flukes) (e.g., *Posthodiplostomum cuticola, Uvulifer ambloplitis*, *Crassiphiala bulboglossa* and *Apophallus brevis*) (Kurochkin and Biserova, 1996). While some aspects of the ecology of these parasite species are relatively well studied (e.g., Hunter, 1933; Miller, 1946; Hoffman, 1956; Hoffman and Putz, 1965; Sinclair, 1972; Ondrackova et al., 2004), spatiotemporal infection patterns and their relation to environmental conditions remain poorly understood. Black spot-causing trematodes all have similar complex life cycles requiring snails, fishes and piscivorous birds as hosts (Figure 2). The parasite encysts under the skin, in the fins or the muscles of the fish host (Krull, 1932, 1934; Hoffman, 1956) where it can survive in a dormant form for several years (Hoffman and Putz, 1965) until the infected fish is eaten by a bird. The distinctive black cyst symptom results from melanin pigment stimulation in the fish tissues induced by trematode cercaria penetration (Davis, 1967; Lemly and Esch, 1984a). Here, we consider black spot-causing trematodes as a guild because of their similar ecological requirements and symptoms caused in fish hosts.

Akin to other parasitic diseases, the infection parameters of black spot disease can vary across time, space, and species. For example, the infection prevalence in bluegill sunfish (*Lepomis macrochirus*) is typically lowest in April/May reaching its maximum around September/October (Lemly and Esch, 1984a, 1984b). This seasonal trend occurs because heavily infected fish are more susceptible to overwintering mortality leading to reduced infection levels in the population in the spring, with new recruits gradually get infected by cercariae released from snails throughout the summer (Lemly and Esch, 1984a, 1984b). Habitat characteristics that favor snail populations such as shallow waters, low water flow and abundant macrophytes such as lentic and/or littoral habitats have also been associated with increasing encounter probability leading to higher abundance and prevalence of the black spot disease in fish hosts (Marcogliese et al. 2001; Ondrackova et al. 2004). A large range of fish species can be infected with some species more heavily infected than others. For instance, Lemly & Esch, (1984a) found that bluegill sunfish were more heavily infected than largemouth bass (*Micropterus nigricans*, syn. *Micropterus salmoides*) because the two species differ in habitat preferences and suitability for *U. ambloplitis*.

### Study area

We sampled 15 lakes from six sub-watersheds around Saint-Hippolyte, QC, Canada (45.92°N, 73.97°W). The lakes Corriveau, Croche, Cromwell and Triton are located in the protected area of the Université de Montréal’s Biology Station of the Laurentians. The remaining lakes are surrounded by private properties and are accessible only upon consent of landowners or municipal associations. Considering their geographic proximity (study area covers less than 100 km2), we assumed that all sampled lakes are exposed to the same climatic conditions. The lakes were selected nonrandomly according to their accessibility (i.e., shore access on private properties granted by landowners) and the availability and variability of morphometric data (Appendix S1: Table S1).

### Fish community sampling

Field work took place from June 15th to August 27th, 2022, and was restricted to days without precipitation to limit sampling bias due to meteorological effects. Three sampling methods (minnow traps, seine nets and snorkeling transects) were used to assess infection prevalence while maximizing fish abundance and species richness. Sampling effort within lakes was determined according to the lake surface area, except for minnow traps (Appendix S1: Table S2) as setting these traps requires less time and manipulation than the other methods. Seine sampling varied between four and eight fishing efforts while the number of snorkeling transect sites varied between two and six per lake. 15 minnow traps were set in each lake to maximize the number of captures. Animal handling and collection was approved by the Université de Montréal’s animal care committee (protocol number 22-025) and a scientific fishing permit granted by the Ministère de l’Environnement, de la Lutte contre les changements climatiques, de la Faune et des Parcs (MELCCFP) of Québec (2022-05-16-1971-15-S-P).

The minnow trap and seine net samplings were both conducted on the same day for each lake (see Appendix S1: Table S3 for fishing gear dimensions). The seine samplings occurred during daytime (usually between 10AM and 4PM) using a beach seine towed on foot or partially deployed from a boat in habitats without large obstacles like trunks and rocks to prevent net obstruction. Two sizes of minnow traps were set at 4PM and pulled four hours later, to target species that are most active at dusk. All 15 traps were set on the bottom substrate (between 0.5 and 4 m deep) and distributed at approximately equal distance along the shore to ensure full lake perimeter coverage. Eight of the traps (3 large and 5 small traps) were baited with bread to sample various feeding preferences and behavioral traits. All the fishing gear was pressured washed and sun dried between lake sampling following MELCCFP recommendations to prevent the spread of exotic species. All captured fishes were counted, identified to the species level, and measured (total length to the nearest centimeter) directly after capture and subsequently released. Northern redbelly dace (*Chrosomus eos*) and finescale dace (*Chrosomus neogaeus*) individuals were identified as *Chrosomus* spp. as the two species hybridize in these lakes and are difficult to distinguish based on morphology in the field (Angers and Schlosser, 2007; Leung et al., 2017). The presence of black cysts was assessed by visual examination of the fish surfaces. If a cyst was observed, the fish was noted as infected. Juveniles and adults were included in this study as both life stages are vulnerable to black spot disease.

Fifty-meter snorkeling transects positioned along the shore were conducted to assess black spot infection prevalence in the fish communities. We selected sites that were between approximately 0.5 and 3.0 m deep, not fully covered by macrophytes, not obstructed by docks and with vegetation, rock, or trunk refuges. The sites were selected in late May 2022 and flagged along the shoreline every ten meters. The transects were performed simultaneously by two observers each monitoring a one-meter field of view in front of them and moving forward at a pace of three minutes per ten meters. The fishes entering the field of view from behind the observer were not counted to avoid recounting. All the fishes were identified to species with the exception of cyprinids, which were identified to family level due to the difficulty in accurately identifying these fishes to species level while swimming. Visible black spot infection was noted for each individual.

Following the prevalence transects, we conducted a habitat description. The percentage of four substrate categories (silt, sand, rock, and boulder), the macrophyte coverage, the number of trunks (or large branches), and the mean depth was estimated for each 10-meter transect section. The mean estimate of each category was then calculated for all variables at the transect-scale except for the number or trunks that were summed. Temperature, dissolved oxygen, conductivity, turbidity, and pH were measured at mid-depth (around 0.3 to 2 meters) at the beginning of every transect with a YSI ProDSS Multi-Parameter Water Quality Meter. 1 L of unfiltered water was taken at mid-depth at each transect in previously acid-washed HDPE bottles for carbon and nutrient content analyses. Water samples were placed on ice in a dark cooler and transported back to the field station laboratory. Each water sample was separated into one previously acid-washed 40 mL vial to measure total organic carbon (TOC), and one 500 mL HDPE bottle for total nitrogen (TN) and total phosphorus (TP). TOC samples were refrigerated at 4°C refrigerator and analyzed within a week while TN and TP samples were kept frozen at -20°C until processing at Université de Montréal (MIL campus). TOC samples were measured on a Siervers M5310 C Laboratory Total Organic Carbon Analyzer. TN and TP samples were oxidized with persulfate and autoclaved the day before analysis following EPA methods 353.2 and 365.3 respectively (US EPA, 1978, 1993). TN concentrations were measured on a Lachat QuikChem 8500 analyzer and TP concentrations were analyzed on an Asoria-Pacific Astoria2.

### Statistical analyses

For each sampling method, fish community infection prevalence was estimated as the number of host fishes infected with black spot disease divided by the total number of hosts sampled at a given fish community scale (e.g., lake community prevalence is the result of all infected fishes sampled in the lake divided by all sampled hosts in this same lake). Since we did not identify black spot-causing trematode species, using species-specific infection metric would be inadequate to assess environmental characteristics driving infection risk. Prevalence estimates are less sensitive to host and population identity (Poulin 2006) and are thus, a more appropriate metric to study community-level sources of variation in disease dynamics.

We described the black spot prevalence patterns across multiple scales (landscape, lake and site; with each scale pooling all fishes sampled at the corresponding scale). All data manipulations and analyses were conducted using R (version 4.2.3). Lake Tracy was excluded from the lake-scale analysis because of low fish abundance data obtained across all fishing methods (only one individual sampled). Prevalence distributions were compared between methods using a Kruskal-Wallis rank sum test and Dunn test with a Benjamini-Hochberg correction (Dinno, 2017). Community metrics (species richness, evenness and Gini-Simpson’s diversity) were calculated with the *vegan* package (Oksanen et al., 2024).

For the landscape-scale analysis, we looked at the effect of the sampling design (here sampling method and sampling effort) on the estimation of landscape fish infection prevalence. The sampling effort (i.e., sampling attempt; either one minnow trap pulled out, one seine stroke or one snorkeling transect) associated prevalence was calculated as the number of infected individuals divided by the total number of individuals regardless of the origin lake. We used a resampling approach on the data from the different sampling methods (minnow trap, seine net, transect, and all methods combined) to investigate how the infection prevalence changes along a gradient of increasing sampling effort (N). For each sampling method, we randomly selected N sampling effort from our data to estimate the landscape prevalence (here resampled landscape prevalence). We used a weighted mean to estimate the landscape prevalence across samples to avoid the effect of different fish abundances in samples. The operation was repeated 999 times for a sampling gradient N from 1 to 35 samples. We set the largest sampling effort at 35 because we had a total of 39 transect samples across all lakes, and we did not want to resample the entire data set. Mean and 95% confidence intervals were calculated for each N distribution generated. A local regression (loess) was then fitted to the landscape estimates generated by our simulation at each N value, for visualization of the trend of mean resampled prevalence. We then compared for each method the observed landscape prevalence (estimated with all sampled fish in the landscape) with resampled landscape prevalence at N = 5, 10, 15, 20, 25, 30 and 35 to see when accurate prevalence (i.e., when resampled preference is not different from observed landscape prevalence) is reached.

For the lake-scale analysis, we examined the frequency distribution of the fish community infection prevalence according to the sampling method. The lake infection prevalence was previously calculated for each lake (i.e., all infected fishes divided by all sampled hosts within each lake), and visualization was made with histograms set at six bins to avoid gaps. The distributions were then compared with the maps of the study area showing prevalence gradient in order to look for spatial patterns (e.g., spatial distribution of infection hotspots/coldspots, hydrological connectivity displaying prevalence gradients).

For the site-scale analysis, we modeled the relationships between environmental drivers and the community infection prevalence at the transect level (site infection prevalence). We used the data from the transect method because it is the only method that allows an accurate association of infection prevalence with all selected environmental variables, thus can best reveal the key drivers of infection prevalence. Prior to analysis, we explored the data for outliers in both the response and explanatory variables, for collinearity between explanatory variables, and for non-linearity in the relationships between the response and the explanatory variables following recommendations by Zuur et al. (2009). Since our data showed some non-linear patterns, we applied generalized additive models in a mixed-model framework (GAMMs) using the *mgcv* package (Wood, 2023). We modeled one environmental variable at a time using a cubic spline (smooth function) and used the lake variable as a random effect to account for spatial clustering of sites from the same lake. Maximum likelihood method (method = “ML”) was used for estimating penalties. We used a quasi-binomial family distribution because our response variable (site infection prevalence) is a proportion constrained between 0 and 1 and to account for overdispersion in the data. The optimal amount of smoothing (λ) was internally determined by the modeling function from the *mgcv* package (Wood, 2023) and the number of basis functions (k) was set at 5 for every models to lower chances of overfitting and computational cost (Wood and Augustin, 2002). The deviance explained (D2) was used as a measure of model fit. The model validations were conducted with *gratia* package (Simpson and Singmann, 2024) and visualization of partial effects was made with *gratia* (Simpson and Singmann, 2024) and *ggplot2* (Wickham et al., 2024) packages. Validation of smooth terms significance was done by drawing a horizontal line through the confidence region (Wood and Augustin, 2002).

## RESULTS

We recorded a total of 11 297 individual fishes belonging to 15 species and 8 families in the sampled lakes (Appendix S2: Table S1). The minnow traps caught 1906 individuals from 10 species while seine nets caught 2427 individuals from 14 species. 6964 individuals belonging to four species and one separate family (cyprinids) were observed in the snorkeling transects. The mean total length of fish sampled with our fishing methods was 6 ± 3 cm (mean length ± SD; N = 4333, Appendix S2: Table S2; Table S3).

Scientific literature supports black spot disease susceptibility in all fish species sampled except for *Ameiurus nebulosus* and *Esox masquinongy* (Appendix S2: Table S4). Consequently, both species were excluded from all prevalence estimations as there is no evidence they can host these trematodes. Our field sampling found evidence of black spot infection in 7 species: pumpkinseed sunfish (*Lepomis gibbosus*), rock bass (*Ambloplites rupestris*), yellow perch (*Perca flavescens*), smallmouth bass (*Micropterus dolomieu*), creek chub (*Semotilus atromaculatus*), bluntnose minnow (*Pimephales notatus*) and fathead minnow (*Pimephales promelas*). Details on infection prevalence of fish communities are presented for every sampling method in Appendix S2: Table S5 (landscape-scale); Table S6 (lake-scale); Table S7 (site-scale).

### Landscape-scale

We used a resampling approach to compare how different sampling methods change the mean prevalence estimate through an increasing sampling effort across the landscape (Figure 3). Accumulation curves built for each sampling method tended to overestimate the mean resampled landscape prevalence at low numbers of samplings (N < 10). The estimate of resampled landscape prevalence differed among the sampling methods (Kruskal-Wallis chi-squared = 1288, p-value = 0, Appendix S2: Table S8), varying between 22% and 36% (Appendix S2: Table S9) with confidence intervals decreasing with N (Appendix S2: Table S10). After 35 samples, the transect method generated the highest mean resampled landscape prevalence (36 %) followed by the methods combination (31 %), the minnow trap method (24 %) and the seine method (22 %). However, the minnow trap curve did not stabilize after the 35 random samples with final resampled landscape prevalence (N = 35) higher than observed landscape prevalence (Appendix S2: Table S9). However, relatively few random samples are necessary to estimate a landscape prevalence for the transect method (approximately 10 samples) while a stabilized landscape prevalence value only occurs after 30 samples for the seine method. Even if the method combination curve (in gray) stabilizes around 20 samples, there is variation around the curve because of the variability among the different methods used.

### Lake-scale

We compared the frequency distribution of the lake’s prevalence for each sampling method to see how the landscape prevalence is distributed among lakes and in their geographic context to identify spatial patterns (Figure 4). Fish community infection prevalence estimates measured at the lake level are not constant across the landscape. Both frequency distributions and prevalence maps suggest differences among methods leading to biases in infection estimates. The lake prevalence frequency distributions for the combined methods and for the transect method showed a bimodal distribution, with the landscape composed of many low-prevalence and high-prevalence lakes. However, neither method yielded prevalence estimates over 75% (Appendix S2: Table S6). The distribution patterns for the two fishing methods (seine net and minnow trap) are less clear. These methods show a right-skew distribution representing more low-prevalence lakes. Accordingly, seine and minnow traps had similar estimates of landscape-scale prevalence (20% and 19% respectively, Appendix S2: Table S5). Although spatial patterns vary across sampling methods, heavily infected and less infected lakes do not appear to be clustered in space at the regional scale (see maps in Figure 4). Moreover, close and connected lakes do not appear to follow a spatial infection gradient. However, lake-scale geographic patterns were not statistically tested.

### Site-scale

The relationships between the potential predictors and the site-scale infection prevalence (transect-level prevalence) were assessed with generalized additive mixed effects models (see Appendix S2: Table S11 for model estimates). The environmental characteristics of transect sites used as model predictors are presented in Appendix S2: Table S12.

The models for TN:TP, TN, TP, TOC, substrate type (silt, sand, rock and boulder), mean transect depth, number of trunks, turbidity, dissolved oxygen, lake perimeter, lake maximum depth, lake mean depth, water residence time, drainage area, lake elevation, distance to the nearest lake, species richness, fish community diversity and evenness were rejected based on significance of smooth term (Appendix S2: Table S11). We did not find any relationships between black spot site prevalence and any spatial attributes of the lakes. However, water physico-chemistry and metrics of community abundance were good predictors of site infection prevalence of fish communities.

The non-host abundance model (here abundance of cyprinids) had the best fit (D2 = 87.49%) and showed a negative relationship with site prevalence suggesting that sites with more unsuitable fishes having a lower prevalence of infection (Figure 5a). Total fish abundance model presented the same pattern (Figure 5b). Site infection prevalence increased proportionally with both water temperature and pH with prevalence being highest in warmer, more alkaline sites (Figure 5c–d). The conductivity and site prevalence had a non-linear, unimodal relationship peaking around 80 (μS/cm) (Figure 5e). For the macrophyte coverage, we found that low macrophyte cover correlates with high infection prevalence in site-scale fish communities (Figure 5f). A parabolic curve is observed in the area:perimeter and surface area models although the variation interval increases at high ratio values (Figure 5g–h).

**DISCUSSION**

Our main goal was to investigate context-dependencies in infection prevalence estimates of black spot disease in freshwater fish communities across three spatial scale levels (landscape, lake, site). First, we built resampled accumulation curves to explore the influence of sampling effort and sampling methods in landscape prevalence estimation, then we compared the frequency distribution of the lake’s prevalence across the landscape for each sampling method in order to assess spatial dynamics. Lastly, we identified ecological drivers explaining variations of site-scale prevalence estimates and suggested potential mechanisms explaining observed relationships. This approach allowed us to investigate overlooked issues in parasite ecology such as spatial patterns of infection, sampling effort effect, sampling methods biases and environmental drivers of infection prevalence in fish communities. Our results show substantial differences in prevalence estimates across methods suggesting major sampling bias in infection estimates. We also show that minimal sampling effort must also consider the sampling method as they present different outcomes in resampled landscape prevalence accumulation curves. Local environmental characteristics such as water physico-chemistry and community structure were found to be strong drivers of infection prevalence at small spatial scales.

Study designs in ecology typically aim to generate accurate and precise data that are representative of natural systems while minimizing stress on the environment and maximizing statistical power. However, researchers often compromise aspects of the ideal sampling design to overcome other challenges potentially biasing the collected data. Surveys of disease prevalence among populations and communities often do not consider the importance of sampling design and methods bias despite the potential to substantially influence the estimates derived from wildlife samples collected using different techniques (Biro and Dingemanse, 2009). For example, Wilson et al. (1993) found that sunfish infected with black spot disease were twice as abundant in trapped fish compared to seined fish, whereas white grubs were twice abundant in seined fish, suggesting method-induced bias in he estimates of parasite communities. In our study, estimates of prevalence among sampling methods varied up to 17% at the landscape-scale (Appendix S2: Table S5) and 74% at lake-scale (Appendix S2: Table S6) suggesting that method-induced bias can be substantial in community-level infection prevalence estimation. However, while transects estimated the highest prevalence and minnow traps the lowest prevalence at the landscape-scale, this trend was not maintained at the lake-scale. In fact, at the lake-scale, methods that led to an overestimation or underestimation of the infection prevalence change from one lake to another. Moreover, method choice did not always lead to strong variation between prevalence estimates (e.g., prevalence variation is 0% in lake Beaver, see Appendix S2: Table S6). This suggests that method-induced bias is highly dependent on local community structure and composition.

Different sampling methods introduce their own sources of bias that might affect prevalence estimates. Minnow traps typically select active, bold and curious individuals because fishes must voluntarily enter the trap (Wilson et al., 1993; Biro et al., 2006). Individual behavioral differences can also lead to variation in infection risk (Hart, 1990), where more active and bold individuals are sometimes associated with higher infection risk (Wilson et al., 1993; Boyer et al., 2010; Santicchia et al., 2019; Gradito et al., 2024). Conversely, parasite infection can alter host behavior including habitat use and risk taking (Moore, 2002; Gradito et al., 2024), making them more vulnerable to capture (i.e., “trappability”) (Barber and Dingemanse, 2010; Thambithurai et al., 2022). The direction of causality between infection and behavior remains unclear, particularly as it relates to sampling (Dubois and Binning, 2022; Gradito et al., 2024).

Although seine nets do not directly select host behavior, this method has important habitat-selection restrictions which may influence the fishes likely to be sampled. Sites must not be deeper than the height of the net to avoid fish escaping and, ideally, must not have large obstacles (e.g., branches, trunks, boulders) which might damage or impair the net function. Consequently, seine sites have similar habitat features that facilitate deployment and capture success. Therefore, seine nets might indirectly select for fish species or personality types with similar habitat preferences with according encounter risk with parasites. For example, Biro et al. (2006) found that bold trout were more frequently caught in open risk habitats than shy trout, corroborating habitat use segregation based on personality differences. Although the seine nets sampled more rare species that the other methods, our data did not show consistent species-specific selection across lakes (Appendix S2: Table S1). However, habitat use can change in the context of competition and/or predation (Werner and Hall, 1977; Mittelbach, 1986; Fischer, 2000). Therefore, community composition must be considered when inferring species selection bias based on habitat use.

Observational snorkeling transect are a less habitat and behavior-selective methodology. However, the quality of the observations relies on appropriate meteorological conditions, water transparency (color, turbidity) and expertise of the observers to identify fishes in movement. As a result, precision of infection assessment can vary among sampling days and lakes. For example, infection signs might be missed during times of poor visibility, leading to underestimation of prevalence estimates on cloudy days or following intense rainfall. Moreover, since infection assessment is easier in slower swimming, curious species who tend to approach observers, accuracy of prevalence estimates might also vary among fish species. Detectability of infection is also easier when intensity of infection is high (a fish with only a few black spots could easily be missed by the observer) and thus, prevalence estimation might be more accurate in lakes where mean infection intensity is high no matter the prevalence value.

Finally, since older and/or bigger fishes have a higher risk of infection than younger/smaller ones, population age structures and growth curves might influence infection estimates (Lemly and Esch, 1984a). Further analyses would be needed to determine how species composition of samples, personality types and age structure influence community prevalence estimates, and how these vary across spatial scales.

Few studies have addressed the relationship between infection prevalence estimates in fish communities and sampling effort. Our results showed that low sampling efforts tend to overestimate community infection prevalence at the landscape-scale (Figure 3). This case happens when most samples provide a resampled landscape prevalence value above the observed landscape prevalence and very low estimates pull down the average with increasing sampling effort. By sampling fewer sites, there is a high chance of sampling a higher-than-average prevalence thus overestimating the mean resampled landscape prevalence. With increasing sampling effort, the presence of few samples with zero or very low prevalence bring down the mean resampled landscape prevalence until observed landscape prevalence is reached. Presence of high and zero prevalence values show that infection is spatially heterogeneous with the existence of spatial infection clusters (infection hotspots and coldspots) which are reflected in our data regardless to lake size. Parasites tend to be aggregated in their hosts, where a few individuals in a population or community host the majority of the infection (Shaw and Dobson, 1995; Poulin, 2007). This pattern can be driven by the environmental aggregation of infective parasite stages leading to heterogenous host exposure risk across a landscape (Karvonen et al., 2004). Here, the data showed many higher-than-averaged-prevalence sites with some uninfected sites which is to be expected for common generalist parasites. However, our data does not allow investigating aggregation patterns and abundance-occupancy relationships further since parasite abundance in hosts was not measured.

The number of sampled hosts in a study may also influence infection prevalence estimate (Marques and Cabral, 2007). Jovani and Tella (2006) reported that small host sample sizes resulted in greater inaccuracy in prevalence estimates and suggested an optimal sample size of around 15 individuals as a trade-off between discriminating samples and maintaining accuracy. Gregory and Woolhouse (1993) found contrasting results where simple size has no effect on parasite prevalence values. Conversely, our results showed that the sampling effort influences parasite prevalence values on one hand, and that the minimum number of samples needed to adequately estimate landscape prevalence differs among methods on the other. The only sampling method that reached a stable value following our sampling effort was the observational snorkeling transects (around 10 samples) although the seine and the combination of sampling methods showed less than a 2% difference between resampled landscape prevalence and observed landscape prevalence after 35 samples (Appendix S2: Table S9). Indeed, snorkeling transects allow observations on more individuals than the fishing methods (6964 versus 4333 observations), explaining why it reaches stability faster and thus, needs less sampling effort to reach a precise estimate. Conversely, fishing methods require more extensive sampling in terms of area and time to estimate infection prevalence. Accordingly, for minnow traps, 35 sampling efforts (out of a data set of 225 sampling efforts) were not exhaustive enough to capture the observed landscape infection prevalence (Appendix S2: Table S9). However, assessments of infection intensity and prevalence are more reliable with fishing because the observer can take the necessary time to inspect individuals for signs of disease. Unfortunately, there is no magic threshold of sampling effort for estimating accurate prevalence (Poulin 2007; Shvydka et al. 2018). Although, our results should be used as an empirical tool to guide future determination of sampling effort to assess infection prevalence in natural systems.

Our results show that prevalence of infection is spatially heterogenous in our system. Prevalence estimates varied both among and within lakes. Regardless, we did not find evidence of geographical patterns. Accordingly, none of the spatial attributes that we tested were good predictors of infection at small spatial scales (site-scale) (Appendix S2: Table S11). Future studies should increase the number of lakes sampled to confirm whether infection is spatially random, or if indeed there are other patterns at play.

Frequency distributions are a useful tool to investigate mechanisms behind the spatial occurrence of species (McGeoch and Gaston, 2002). Frequency distributions of prevalence of black spot disease infection in fish communities showed that prevalence is heterogeneously distributed across lakes and that the distribution patterns differed between sampling methods: the transects and the combination of methods tended to have a bimodal distribution (Figure 4a–b) whereas traps and seine net fishing showed no obvious pattern (Figure 4c–d). While binomial distributions are commonly observed in frequency occupation distributions, this pattern may be an artifact of a small sample size (Brown, 1984). However, for our purpose, bimodal distributions revealed that lake communities tend to display either high (>60%) or low (<15%) infection prevalence. Since we assumed that the distribution of trematodes is not limited by the final hosts’ movements (birds) across our study area, these high- and low-prevalence groups likely represent infection hotspots and coldspots characterized by environmental conditions that favor (or restrict) parasite abundance and transmission from snails to fish hosts. Accordingly, map comparisons (Figure 4) clearly show that lake fish community prevalence can be over or underestimated according to the sampling method used. The different methods might have selected for fish species varying in parasite compatibility, environment varying in quality for infective stage recruitment, specific age classes or sampled different personality traits that led to biased prevalence estimates. Although, our approach revealed method biases in prevalence estimates, but it did not allow us to establish the mechanism behind this relationship. Moreover, method biases seem to be lake-dependent implying influence of the local characteristics in method-related biases in prevalence estimation.

At the site-scale, we used GAMMs to inspect relationships between the site prevalence estimate and environmental predictors (Figure 5; Appendix S2: Table S11) in order to look into what may qualify spatial cluster of infection prevalence. Our results showed that small-scale parameters such as water physico-chemistry and community metrics are better predictors of site prevalence estimates than larger scale metrics (e.g., lake- or landscape-scale parameters; Appendix S2: Table S11) suggesting the local scale is more appropriate for management and conservation strategies. Conductivity, area:perimeter and surface area models must, however, be carefully interpreted because of gaps in the corresponding environmental gradient (Figure 5). However we did not account for parasite identity and phylogenetic information of hosts (Martins and Poulin, 2024), our findings suggest that some environmental drivers of prevalence may transcend species identity.

3 out 5 physico-chemistry parameters models (temperature, conductivity, and pH) were better predictors of infection prevalence at the site-scale (Figure 5c–e) than the null model (Appendix S2: Table S11). Given the simplicity of sampling water physico-chemistry measurements, these metrics might be useful as proxies of parasite infection for monitoring and managing strategies. Consequently, we encourage researchers to investigate these relationships in other parasite species, geographic localities, and types of habitats as our data provide only a snapshot the complexity of fish-parasite associations in natural systems.

Surprisingly, most habitat structure features were not good predictors of prevalence of infection in site-scale fish communities from the littoral zone. However, we found a strong (D2 = 84.15%) negative relationship between the macrophyte cover and the prevalence of infection in transect sites (Figure 5f). This was somewhat surprising as previous studies suggested heavier infection is associated with more vegetated zones (Hartmann and Nümann 1977; Marcogliese et al. 2001; Ondrackova et al. 2004). However, trematode cercariae have a short free-living stage (24-72h depending on species and environmental conditions) (Galaktionov and Dobrovolskij, 2003; Morley, 2012) and thus physical barriers created by aquatic vegetation or other debris might prevent the cercaria from encountering a suitable host (Bartoli and Boudouresque, 1997; Welsh et al., 2014). Another explanation is that the snail species releasing cercaria in our system may not be as reliant on macrophytes as some other aquatic gastropods. For example, the mud Amnicola (*Amnicola limosus*) is the second intermediate host to at least one species of the black spot trematode guild in our lake system (L’Heureux, personal communications). This freshwater snail species live partially in the lake sediment from the littoral zone (Pinel-Alloul and Magnin, 1973) and often grazes on periphyton (Kesler, 1981). Consequently, we would not expect a strong relationship with macrophyte cover and cercaria presence for infections transmitted by this species.

Our results suggest that fish community structure is a driver of black spot disease prevalence at small spatial scales. We found significant negative relationships between site infection prevalence and, non-host and total fish abundance (Figure 5a–b). These results support the dilution effect hypothesis. Since total fish abundance and non-host abundance are positively correlated (Appendix S2: Figure S1), we assume similar mechanisms explaining why lower prevalence was measured in transects with higher fish abundance. Several mechanisms, such as of community susceptibility and behavioural change in hosts (Ahn and Goater, 2021), might explain how non-host abundance might reduce prevalence estimates in fish communities. In our lake system, *L. gibbosus* dominates littoral communities and is the species with the highest infection prevalence across the landscape (Appendix S2: Table S13). The presence of competitor species (e.g., *P. flavescens*) in a lake or a site might decrease the overall prevalence of infection in the community by constraining pumpkinseeds sunfish’s behavior thus reducing encounter with cercariae. Indeed, Dargent et al. (2013) found that the presence and abundance of *Gyrodactylus* spp. (monogenean) were lower in hosts when they occur in mixed-species groups.

## CONCLUSIONS

Understanding the processes explaining parasitic disease dynamics across scales is fundamental to predicting future infection and extinction risk in natural systems. Here, we investigated context-dependencies of fish community prevalence estimates of the black spot disease across multiple scales and discussed the roles of sampling effort, sampling method bias, spatial occurrence patterns and environmental predictors in estimating host infection prevalence. Results suggested that prevalence is distributed in a non-random, heterogeneous way across the landscape with small sampling effort largely overestimating regional prevalence. We provide evidence for inconsistent method-induced bias in prevalence estimates at both the lake and landscape-scale, sometimes leading to high variation in prevalence estimates depending on the sampling method used. The fish sampling method also influenced the sampling effort needed to reach accurate landscape prevalence estimates, with observational snorkeling transects requiring the least amount to achieve accurate estimates. The best predictors of site-scale prevalence were local metrics such as water physico-chemistry and community structure metrics, suggesting that small-scale monitoring and management might be appropriate for future conservation strategies. Our findings also suggest that encounter dilution effects by host decoys or physical obstruction reduce infection prevalence in fish hosts. However, we observed non-linearity in these effects suggesting complex interactions that we could not disentangle with our approach. Therefore, we encourage more research across larger spatial and temporal scales to deepen our understanding of the processes linking these predictors to infection parameters for future integration into sampling surveys and conservation plans.

## SPECULATIONS

Authors are aware that lumping, potentially, different parasite species together as “black spot causing trematodes” is an approach that comes with many challenges since phylogenetics is unaccounted for and that environment may influence parasite species in different ways. Patterns identified here should not be used to make predictions but rather stimulate reflections on the implications of context-dependencies in the estimation of infection metrics from field-based data sets. We speculate that this approach could be applied to other host-parasite systems to determine whether our findings are broadly applicable or, once again, context-dependent to this particular system.

**REFERENCES**

Ahn, S., Goater, C.P., 2021. Nonhost species reduce parasite infection in a focal host species within experimental fish communities. Ecol. Evol. 11, 10155–10163. https://doi.org/10.1002/ece3.7823

Angers, B., Schlosser, I.J., 2007. The origin of *Phoxinus eos-neogaeus* unisexual hybrids. Mol. Ecol. 16, 4562–4571. https://doi.org/10.1111/j.1365-294X.2007.03511.x

Barber, I., Dingemanse, N.J., 2010. Parasitism and the evolutionary ecology of animal personality. Philos. Trans. R. Soc. B Biol. Sci. 365, 4077–4088. https://doi.org/10.1098/rstb.2010.0182

Barber, I., Hoare, D., Krause, J., 2000. Effects of parasites on fish behaviour: a review and evolutionary perspective. Rev. Fish Biol. Fish. 10, 131–165. https://doi.org/10.1023/A:1016658224470

Bartoli, P., Boudouresque, C.-F., 1997. Transmission failure of parasites (Digenea) in sites colonized by the recently introduced invasive alga *Caulerpa taxifolia*. Mar. Ecol. Prog. Ser. 154, 253–260.

Biro, P.A., Abrahams, M.V., Post, J.R., Parkinson, E.A., 2006. Behavioural trade-offs between growth and mortality explain evolution of submaximal growth rates. J. Anim. Ecol. 75, 1165–1171. https://doi.org/10.1111/j.1365-2656.2006.01137.x

Biro, P.A., Dingemanse, N.J., 2009. Sampling bias resulting from animal personality. Trends Ecol. Evol. 24, 66–67. https://doi.org/10.1016/j.tree.2008.11.001

Biro, P.A., Post, J.R., 2008. Rapid depletion of genotypes with fast growth and bold personality traits from harvested fish populations. Proc. Natl. Acad. Sci. U. S. A. 105, 2919–2922. https://doi.org/10.1073/pnas.0708159105

Bolnick, D.I., Resetarits, E.J., Ballare, K., Stuart, Y.E., Stutz, W.E., 2020. Scale-dependent effects of host patch traits on species composition in a stickleback parasite metacommunity. Ecology 101, e03181. https://doi.org/10.1002/ecy.3181

Boyer, N., Réale, D., Marmet, J., Pisanu, B., Chapuis, J.-L., 2010. Personality, space use and tick load in an introduced population of Siberian chipmunks *Tamias sibiricus*. J. Anim. Ecol. 79, 538–547. https://doi.org/10.1111/j.1365-2656.2010.01659.x

Brown, J.H., 1984. On the Relationship between Abundance and Distribution of Species. Am. Nat. 124, 255–279. https://doi.org/10.1086/284267

Buck, J.C., Lutterschmidt, W.I., 2017. Parasite abundance decreases with host density: evidence of the encounter-dilution effect for a parasite with a complex life cycle. Hydrobiologia 784, 201–210. https://doi.org/10.1007/s10750-016-2874-8

Bush, A.O., Lafferty, K.D., Lotz, J.M., Shostak, A.W., 1997. Parasitology meets ecology on its own terms: Margolis et al. revisited. J. Parasitol. 83, 575–583.

Chrétien, E., De Bonville, J., Guitard, J., Binning, S.A., Melis, É., Kack, A., Côté, A., Gradito, M., Papillon, A., Thelamon, V., Levet, M., Barou-Dagues, M., 2023. Few studies of wild animal performance account for parasite infections: A systematic review. J. Anim. Ecol. 92, 794–806. https://doi.org/10.1111/1365-2656.13864

Civitello, D.J., Cohen, J., Fatima, H., Halstead, N.T., Liriano, J., McMahon, T.A., Ortega, C.N., Sauer, E.L., Sehgal, T., Young, S., Rohr, J.R., 2015. Biodiversity inhibits parasites: Broad evidence for the dilution effect. Proc. Natl. Acad. Sci. 112, 8667–8671. https://doi.org/10.1073/pnas.1506279112

Cohen, J.M., Civitello, D.J., Brace, A.J., Feichtinger, E.M., Ortega, C.N., Richardson, J.C., Sauer, E.L., Liu, X., Rohr, J.R., 2016. Spatial scale modulates the strength of ecological processes driving disease distributions. Proc. Natl. Acad. Sci. 113, E3359–E3364. https://doi.org/10.1073/pnas.1521657113

Dargent, F., Torres-Dowdall, J., Scott, M.E., Ramnarine, I., Fussmann, G.F., 2013. Can Mixed-Species Groups Reduce Individual Parasite Load? A Field Test with Two Closely Related Poeciliid Fishes (Poecilia reticulata and Poecilia picta). PLoS ONE 8, e56789. https://doi.org/10.1371/journal.pone.0056789

Davis, H.S., 1967. Culture and Diseases of Game Fishes. University of California Press.

de Solla, S.R., Shirose, L.J., Fernie, K.J., Barrett, G.C., Brousseau, C.S., Bishop, C.A., 2005. Effect of sampling effort and species detectability on volunteer based anuran monitoring programs. Biol. Conserv. 121, 585–594. https://doi.org/10.1016/j.biocon.2004.06.018

Dinno, A., 2017. dunn.test: Dunn’s Test of Multiple Comparisons Using Rank Sums. –<https://cran.r-project.org/web/packages/dunn.test/index.html>

Dobson, A.P., Hudson, P.J., 1986. Parasites, disease and the structure of ecological communities. Trends Ecol. Evol. 1, 11–15. https://doi.org/10.1016/0169-5347(86)90060-1

Dubois, F., Binning, S.A., 2022. Predation and parasitism as determinants of animal personalities. J. Anim. Ecol. 91, 1918–1928. https://doi.org/10.1111/1365-2656.13781

Fischer, P., 2000. Test of Competitive Interactions for Space Between Two Benthic Fish Species, Burbot *Lota lota*, and Stone Loach *Barbatula barbatula*. Environ. Biol. Fishes 58, 439–446. https://doi.org/10.1023/A:1007631107521

Frainer, A., McKie, B.G., Amundsen, P.-A., Knudsen, R., Lafferty, K.D., 2018. Parasitism and the Biodiversity-Functioning Relationship. Trends Ecol. Evol. 33, 260–268. https://doi.org/10.1016/j.tree.2018.01.011

Galaktionov, K., Dobrovolskij, A.A., 2003. The Biology and Evolution of Trematodes An Essay on the Biology, Morphology, Life Cycles, Transmissions, and Evolution of Digenetic Trematodes.

Gradito, M., Dubois, F., Noble, D., Binning, S., 2024. Double trouble: host behaviour influences and is influenced by co-infection with parasites. Anim. Behav. 215, 31–44. https://doi.org/10.1016/j.anbehav.2024.06.016

Gregory, R.D., Woolhouse, M.E.J., 1993. Quantification of parasite aggregation: A simulation study. Acta Trop. 54, 131–139. https://doi.org/10.1016/0001-706X(93)90059-K

Happel, A., 2019. A volunteer-populated online database provides evidence for a geographic pattern in symptoms of black spot infections. Int. J. Parasitol. Parasites Wildl. 10, 156–163. https://doi.org/10.1016/j.ijppaw.2019.08.003

Hart, B.L., 1990. Behavioral adaptations to pathogens and parasites: Five strategies. Neurosci. Biobehav. Rev. 14, 273–294. https://doi.org/10.1016/S0149-7634(05)80038-7

Hartmann, J., Nümann, W., 1977. Percids of Lake Constance, a Lake Undergoing Eutrophication. J. Fish. Res. Board Can. 34, 1670–1677. https://doi.org/10.1139/f77-231

Hoffman, G.L., 1956. The Life Cycle of *Crassiphiala bulboglossa* (Trematoda: Strigeida). Development of the Metacercaria and Cyst, and Effect on the Fish Hosts. J. Parasitol. 42, 435–444. https://doi.org/10.2307/3274528

Hoffman, G.L., Putz, R.E., 1965. The Black-Spot (Uvulifer ambloplitis: Trematoda: Strigeoidea) of Centrarchid Fishes. Trans. Am. Fish. Soc. 94, 143–151. https://doi.org/10.1577/1548-8659(1965)94[143:TBUASO]2.0.CO;2

Hunter, G.W., 1933. The Strigeid Trematode, Crassiphiala ambloplitis (Hughes 1927). Parasitology 25, 510–517. https://doi.org/10.1017/S0031182000019752

Iwanowicz, D., 2011. Overview On The Effects Of Parasites On Fish Health.

Johnson, P.T.J., Thieltges, D.W., 2010. Diversity, decoys and the dilution effect: how ecological communities affect disease risk. J. Exp. Biol. 213, 961–970. https://doi.org/10.1242/jeb.037721

Johnstone, K.C., McArthur, C., Banks, P.B., 2021. Behavioural drivers of survey bias: interactive effects of personality, the perceived risk and device properties. Oecologia 197, 117–127. https://doi.org/10.1007/s00442-021-05021-7

Jovani, R., Tella, J.L., 2006. Parasite prevalence and sample size: misconceptions and solutions. Trends Parasitol. 22, 214–218. https://doi.org/10.1016/j.pt.2006.02.011

Karvonen, A., Hudson, P.J., Seppälä, O., Valtonen, E.T., 2004. Transmission dynamics of a trematode parasite: exposure, acquired resistance and parasite aggregation. Parasitol. Res. 92, 183–188. https://doi.org/10.1007/s00436-003-1035-y

Kesler, D.H., 1981. Periphyton grazing by *Amnicola limosa*: An enclosure-exclosure experiment. J. Freshw. Ecol. 1, 51–59. https://doi.org/10.1080/02705060.1981.9664016

Kowalski, K., Bogdziewicz, M., Eichert, U., Rychlik, L., 2015. Sex differences in flea infections among rodent hosts: is there a male bias? Parasitol. Res. 114, 337–341. https://doi.org/10.1007/s00436-014-4231-z

Krull, 1932. Studies on the development of *Cercaria bessiae* Cort and Brooks, 1928. J. Parasitol. 19, 1934.

Krull, W.H., 1934. *Cercaria bessiae* Cort and Brooks, 1928, an Injurious Parasite of Fish. Copeia 1934, 69–73. https://doi.org/10.2307/1435795

Kurochkin, I.V., Biserova, L.I., 1996. The etiology and diagnosis of “black spot disease” of fish. Parazitologiia 30, 117–125.

Lafferty, K.D., Allesina, S., Arim, M., Briggs, C.J., De Leo, G., Dobson, A.P., Dunne, J.A., Johnson, P.T.J., Kuris, A.M., Marcogliese, D.J., Martinez, N.D., Memmott, J., Marquet, P.A., McLaughlin, J.P., Mordecai, E.A., Pascual, M., Poulin, R., Thieltges, D.W., 2008. Parasites in food webs: the ultimate missing links. Ecol. Lett. 11, 533–546. https://doi.org/10.1111/j.1461-0248.2008.01174.x

Lafferty, K.D., Dobson, A.P., Kuris, A.M., 2006. Parasites dominate food web links. Proc. Natl. Acad. Sci. U. S. A. 103, 11211–11216. https://doi.org/10.1073/pnas.0604755103

Lagrue, C., Poulin, R., 2015. Local diversity reduces infection risk across multiple freshwater host-parasite associations. Freshw. Biol. 60, 2445–2454. https://doi.org/10.1111/fwb.12677

Lemly, A.D., Esch, G.W., 1984a. Population Biology of the Trematode Uvulifer ambloplitis (Hughes, 1927) in Juvenile Bluegill Sunfish, *Lepomis macrochirus*, and Largemouth Bass, *Micropterus salmoides*. J. Parasitol. 70, 466–474. https://doi.org/10.2307/3281394

Lemly, A.D., Esch, G.W., 1984b. Population Biology of the Trematode *Uvulifer ambloplitis* (Hughes, 1927) in the Snail Intermediate Host, *Helisoma trivolvis*. J. Parasitol. 70, 461. https://doi.org/10.2307/3281393

Leung, C., Duclos, K.K., Grünbaum, T., Cloutier, R., Angers, B., 2017. Asymmetry in dentition and shape of pharyngeal arches in the clonal fish *Chrosomus eos-neogaeus*: Phenotypic plasticity and developmental instability. PLOS ONE 12, e0174235. https://doi.org/10.1371/journal.pone.0174235

Levin, S.A., 1992. The Problem of Pattern and Scale in Ecology: The Robert H. MacArthur Award Lecture. Ecology 73, 1943–1967. https://doi.org/10.2307/1941447

Marcogliese, D.J., 2004. Parasites: Small Players with Crucial Roles in the Ecological Theater. EcoHealth 1, 151–164. https://doi.org/10.1007/s10393-004-0028-3

Marcogliese, D.J., Cone, D.K., 1997. Food webs: a plea for parasites. Trends Ecol. Evol. 12, 320–325. https://doi.org/10.1016/S0169-5347(97)01080-X

Marcogliese, D. J., Cone, D.K., 1997. Parasite communities as indicators of ecosystem stress. Parassitologia 39, 227–232.

Marcogliese, D.J., Dumont, P., Gendron, A.D., Mailhot, Y., Bergeron, E., McLaughlin, J.D., 2001. Spatial and temporal variation in abundance of *Diplostomum* spp. in walleye (*Stizostedion vitreum*) and white suckers (*Catostomus commersoni*) from the St. Lawrence River. Can. J. Zool. 79, 355–369. https://doi.org/10.1139/z00-209

Marques, J.F., Cabral, H.N., 2007. Effects of sample size on fish parasite prevalence, mean abundance and mean intensity estimates. J. Appl. Ichthyol. 23, 158–162. https://doi.org/10.1111/j.1439-0426.2006.00823.x

Martins, P.M., Poulin, R., 2024. Universal versus taxon-specific drivers of helminth prevalence and intensity of infection. Proc. R. Soc. B Biol. Sci. 291, 20241673. https://doi.org/10.1098/rspb.2024.1673

McGeoch, M.A., Gaston, K.J., 2002. Occupancy frequency distributions: patterns, artefacts and mechanisms. Biol. Rev. 77, 311–331. https://doi.org/10.1017/S1464793101005887

Miller, M.J., 1946. The cercaria of *Apophallus brevis*. Can. J. Res. 24, 27–29. https://doi.org/10.1139/cjr46d-003

Minchella, D.J., Scott, M.E., 1991. Parasitism: A cryptic determinant of animal community structure. Trends Ecol. Evol. 6, 250–254. https://doi.org/10.1016/0169-5347(91)90071-5

Mittelbach, G., 1986. Predator-mediated habitat use: some consequences for species interactions. Environ. Biol. Fishes 16, 159–169. https://doi.org/10.1007/BF00005168

Moore, J., 2002. Parasites and the Behavior of Animals.

Morley, N.J., 2012. Cercariae (Platyhelminthes: Trematoda) as neglected components of zooplankton communities in freshwater habitats. Hydrobiologia 691, 7–19. https://doi.org/10.1007/s10750-012-1029-9

Nusser, S.M., Clark, W.R., Otis, D.L., Huang, L., 2008. Sampling Considerations for Disease Surveillance in Wildlife Populations. J. Wildl. Manag. 72, 52–60. https://doi.org/10.2193/2007-317

Oksanen, J., Simpson, G.L., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O’Hara, R.B., Solymos, P., Stevens, M.H.H., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Carvalho, G., Chirico, M., Caceres, M.D., Durand, S., Evangelista, H.B.A., FitzJohn, R., Friendly, M., Furneaux, B., Hannigan, G., Hill, M.O., Lahti, L., McGlinn, D., Ouellette, M.-H., Cunha, E.R., Smith, T., Stier, A., Braak, C.J.F.T., Weedon, J., 2024. vegan: Community Ecology Package. – <https://cran.r-project.org/web/packages/vegan/index.html>

Ondrackova, M., Bartosova, S., Valova, Z., Jurajda, P., Gelnar, M., 2004. Occurrence of black-spot disease caused by metacercariae of *Posthodiplostomum cuticola* among juvenile fishes in water bodies in the Morava River Basin. Acta Parasitol. 49, 222-227.

Pascal, L., Grémare, A., de Montaudouin, X., Deflandre, B., Romero-Ramirez, A., Maire, O., 2020. Parasitism in ecosystem engineer species: A key factor controlling marine ecosystem functioning. J. Anim. Ecol. 89, 2192–2205. https://doi.org/10.1111/1365-2656.13236

Peterson, D., Parker, V., 1998. Ecological Scale: Theory and Applications.

Pinel-Alloul, B., Magnin, E., 1973. Observations sur le cycle vital et la croissance d’*Amnicola limosa* (Say) (Mollusca, Gastropoda, Prosobranchia) du lac Saint-Louis près de Montréal. Can. J. Zool. 51, 311–313. https://doi.org/10.1139/z73-043

Poulin, R., 2007. Are there general laws in parasite ecology? Parasitology 134, 763–776. https://doi.org/10.1017/S0031182006002150

Poulin, Robert, 2007. Evolutionary Ecology of Parasites, 2nd edition. ed. Princeton University Press.

Poulin, R., 2006. Variation in infection parameters among populations within parasite species: Intrinsic properties versus local factors. Int. J. Parasitol. 36, 877–885. https://doi.org/10.1016/j.ijpara.2006.02.021

Poulin, R., 2000. Variation in the intraspecific relationship between fish length and intensity of parasitic infection: biological and statistical causes. J. Fish Biol. 56, 123–137. https://doi.org/10.1111/j.1095-8649.2000.tb02090.x

Poulin, R., 1999. The functional importance of parasites in animal communities: many roles at many levels? Int. J. Parasitol. 29, 903–914. https://doi.org/10.1016/S0020-7519(99)00045-4

Poulin, R., Dick, T.A., 2007. Spatial Variation in Population Density across the Geographical Range in Helminth Parasites of Yellow Perch *Perca flavescens*. Ecography 30, 629–636.

Reimchen, T.E., Nosil, P., 2001. Ecological causes of sex-biased parasitism in threespine stickleback. Biol. J. Linn. Soc. 73, 51–63. https://doi.org/10.1111/j.1095-8312.2001.tb01346.x

Rietkerk, M., van de Koppel, J., Kumar, L., Langevelde, H.H.T., Prins, 2002. The ecology of scale. Ecol. Model. 149, 1–4. https://doi.org/10.1016/S0304-3800(01)00510-5

Rózsa, L., Reiczigel, J., Majoros, G., 2000. Quantifying parasites in samples of hosts. J. Parasitol. 86, 228–232. https://doi.org/10.1645/0022-3395(2000)086[0228:QPISOH]2.0.CO;2

Santicchia, F., Romeo, C., Ferrari, N., Matthysen, E., Vanlauwe, L., Wauters, L.A., Martinoli, A., 2019. The price of being bold? Relationship between personality and endoparasitic infection in a tree squirrel. Mamm. Biol. 97, 1–8. https://doi.org/10.1016/j.mambio.2019.04.007

Scholz, T., Choudhury, A., 2014. Parasites of Freshwater Fishes In North America: Why So Neglected? J. Parasitol. 100, 26–45. https://doi.org/10.1645/13-394.1

Shaw, D.J., Dobson, A.P., 1995. Patterns of macroparasite abundance and aggregation in wildlife populations: a quantitative review. Parasitology 111, S111–S133. https://doi.org/10.1017/S0031182000075855

Shvydka, S., Sarabeev, V., Estruch, V.D., Cadarso-Suárez, C., 2018. Optimum Sample Size to Estimate Mean Parasite Abundance in Fish Parasite Surveys. Helminthologia 55, 52–59. https://doi.org/10.1515/helm-2017-0054

Simpson, G.L., Singmann, H., 2024. gratia: Graceful ’ggplot’-Based Graphics and Other Functions for GAMs Fitted Using “mgcv.” – <https://cran.r-project.org/web/packages/gratia/index.html>

Sinclair, N.R., 1972. Studies on the heterophyid trematode *Apophallus brevis*, the “sand-grain grub” of yellow perch (*Perca flavescens*). II The metacercaria: position, structure, and composition of the cyst; hosts; geographical distribution and variation. Can. J. Zool. 50, 577–584. https://doi.org/10.1139/z72-079

Symons, J., Sprogis, K.R., Bejder, L., 2018. Implications of survey effort on estimating demographic parameters of a long-lived marine top predator. Ecol. Evol. 8, 10470–10481. https://doi.org/10.1002/ece3.4512

Thambithurai, D., Lanthier, I., Contant, E., Killen, S.S., Binning, S.A., 2022. Fish vulnerability to capture by trapping is modulated by individual parasite density. Proc. R. Soc. B Biol. Sci. 289, 20221956. https://doi.org/10.1098/rspb.2022.1956

Thieltges, D.W., Reise, K., 2007. Spatial heterogeneity in parasite infections at different spatial scales in an intertidal bivalve. Oecologia 150, 569–581. https://doi.org/10.1007/s00442-006-0557-2

US EPA, 1993. Method 353.2, Revision 2.0: Determination of Nitrate-Nitrite Nitrogen by Automated Colorimetry.

US EPA, 1978. Method 365.3: Phosphorous, All Forms (Colorimetric, Ascorbic Acid, Two Reagent).

Viana, D.S., Chase, J.M., 2019. Spatial scale modulates the inference of metacommunity assembly processes. Ecology 100, e02576. https://doi.org/10.1002/ecy.2576

Villalba-Vasquez, P.J., Violante-González, J., Monks, S., Marino-Romero, J.U., Ibáñez, S.G., Rojas-Herrera, A.A., Flores-Garza, R., Rosas-Guerrero, V., 2018. Temporal and spatial variations in the metazoan parasite communities of the Panama spadefish, *Parapsettus panamensis* (Pisces: Ephippidae), from the Pacific coast of Mexico. Invertebr. Biol. 137, 339–354. https://doi.org/10.1111/ivb.12232

Welsh, J.E., Meer, J. van der, Brussaard, C.P.D., Thieltges, D.W., 2014. Inventory of organisms interfering with transmission of a marine trematode. J. Mar. Biol. Assoc. U. K. 94, 697–702. https://doi.org/10.1017/S0025315414000034

Werner, E.E., Hall, D.J., 1977. Competition and Habitat Shift in Two Sunfishes (Centrarchidae). Ecology 58, 869–876. https://doi.org/10.2307/1936222

Wickham, H., Chang, W., Henry, L., Pedersen, T.L., Takahashi, K., Wilke, C., Woo, K., Yutani, H., Dunnington, D., Brand, T. van den, Posit, PBC, 2024. ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics. – <https://cran.r-project.org/web/packages/ggplot2/index.html>

Williams-Blangero, S., Criscione, C.D., VandeBerg, J.L., Correa-Oliveira, R., Williams, K.D., Subedi, J., Kent, J.W., Williams, J., Kumar, S., Blangero, J., 2012. Host genetics and population structure effects on parasitic disease. Philos. Trans. R. Soc. B Biol. Sci. 367, 887–894. https://doi.org/10.1098/rstb.2011.0296

Wilson, D.S., Coleman, K., Clark, A.B., Biederman, L., 1993. Shy-bold continuum in pumpkinseed sunfish (*Lepomis gibbosus*): An ecological study of a psychological trait. J. Comp. Psychol. 107, 250–260. https://doi.org/10.1037/0735-7036.107.3.250

Wood, S., 2023. mgcv: Mixed GAM Computation Vehicle with Automatic Smoothness Estimation. – <https://cran.r-project.org/web/packages/mgcv/index.html>

Wood, S.N., Augustin, N.H., 2002. GAMs with integrated model selection using penalized regression splines and applications to environmental modelling. Ecol. Model. 157, 157–177. https://doi.org/10.1016/S0304-3800(02)00193-X

Young, R.E., Maccoll, A.D.C., 2017. Spatial and temporal variation in macroparasite communities of three-spined stickleback. Parasitology 144, 436–449. https://doi.org/10.1017/S0031182016001815

Zuk, M., McKean, K.A., 1996. Sex differences in parasite infections: Patterns and processes. Int. J. Parasitol. 26, 1009–1024. https://doi.org/10.1016/S0020-7519(96)80001-4

Zuur, A.F., Ieno, E.N., Walker, N., Saveliev, A.A., Smith, G.M., 2009. Mixed effects models and extensions in ecology with R, Statistics for Biology and Health. Springer, New York, NY. https://doi.org/10.1007/978-0-387-87458-6

**FIGURE CAPTIONS**

**Figure 1.** Case scenarios of sampling effort effect on landscape prevalence estimates (a–b) and frequency distributions of lake prevalence estimates (c–f). Prevalence could be homogeneous (blue line) or heterogeneous (green and brown lines) across the landscape. In heterogeneous scenarios, prevalence is either overestimated (brown line) or under underestimated (green line) before stabilizing around the accurate prevalence (a). Sampling methods could lead to different prevalence estimates across the landscape. Prevalence difference between stabilized lines would represent method biases (b). All lakes could be infected at the same prevalence level (c). Lake’s frequency distribution could follow a bimodal pattern where low and high prevalence are more common across the landscape (d). Lake’s frequency distribution could follow a unimodal pattern where mean landscape prevalence is more common (e). Lake’s frequency distribution could follow a uniform pattern where all infection levels are equally represented (f).

**Figure 2.** The typical lifecycle of a parasite species causing the black spot disease in North American freshwater fishes. The definitive host is a piscivorous bird, usually a kingfisher (*Megaceryle alcyon*) or a great blue heron (*Ardea herodias*). The metacercariae mature to its adult form and reproduce in the bird digestive system (a). The eggs are released in the water through the bird feces and grow into miracidia (first larval form) (b). The miracidia infect the first intermediate host, a freshwater snail (typically of Planorbidae or Amnicolidae families). The larvae develop into free-living stage cercariae (second larval form) and leave the snail (c). The cercariae swim until they encounter a potential fish host (d). The second intermediate host is a fish. The cercariae penetrate under the skin or the fins and transform in the metacercaria form (e). The fish is eaten by a definitive host (f).

**Figure 3.** Method comparison of landscape prevalence estimates through an increasing random sampling effort. The data points indicate the mean prevalence for a given number of samples (N) randomly selected. 95% confidence intervals of resampled means are presented in Appendix S2: Table S10. The pool data used for the resampling with the minnow traps, the seine nets, the transects and all methods combined respectively contained 225, 75, 39 and 339 samples.

**Figure 4.** Method comparison of the lakes’ local community prevalence estimates across the landscape. Frequency distributions are given in the top right corner of each map. Lakes in grey were not sampled except for lake Tracy that was excluded because of insufficient data (one fish sampled through all methods). Buildings are represented by the grey dots. All the sampling methods combined (weighted mean regional prevalence is 29.62 %) (a). The transect method (weighted mean regional prevalence is 35.55 %) (b). The seine nets (weighted mean regional prevalence is 20.45 %) (c). The minnow traps (weighted mean regional prevalence is 19.46 %) (d).

**Figure 5.** Relationships between the site-scale community infection prevalence and selected environmental drivers. All the models presented are univariate binomial generalized additive mixed models with the lake as a random effect on the intercept. The partial effects of the environmental variables on the prevalence are presented for the significant models. The ticks on the x-axes indicate a data point. Non-host abundance (a). Fishes abundance (b). Temperature (c). pH (d). Conductivity (e). Macrophyte coverage (f). Area:Perimeter ratio (g). Surface area (h). Area:Perimeter ratio is a lake attribute while other parameter are site-scale measurements.

**FIGURES**

