**4 | DISCUSSION**

Prevalence is simple and common parasitological indice that gives information on the proportion of infected individuals in a given group that must always be provided in parasite survey (Bush et al., 1997; Rózsa et al., 2000). Our goal here was to investigate the estimate of prevalence in fish communities across multiple spatial scales using data on the black spot disease. This approach allowed us to skim over spatial patterns of infection, sample size effect, sampling methods bias and environmental predictors of prevalence of infection in fish communities.

**[Comments on methods]**

Despite fish-host have been the focus of many empirical studies on aquatic parasites, sampling design and methods bias are not usually discussed although they can induce substantial differences in measurements derived from wildlife samples. For example, Wilson et al. (1993) reported difference in parasite infection in pumpkinseed sunfish (*Lepomis gibbosus*) caught by two different fishing methods. They mentioned that black spots were twice as abundant in trapped fish against seined fish, and white grubs twice as abundant in seined fish, supporting method-induce bias. Our findings also support fishing method-induced bias in the measurement of community-level infection prevalence. At landscape-scale, black spots prevalence estimates varied between 19.2% and 35.6 % (Table S12) depending on the method while it varied up to 46.7% between methods at lake-scale (Table S13).

Sampling wildlife is challenging, and all methods comes with their advantages and inconveniences. While fishing methods require a more extensive sampling than field observations, assessments of infection are accurate and precise because the observer can take the necessary time to inspect fish individuals. Regardless, minnow traps are typically selective on active, bold and curious behavior because fish must enter voluntarily in the trap (Biro et al., 2006; Wilson et al., 1993). Even if seine net is not directly a behavior-selective gear, this method comes with important habitat-selection restrictions. Sites must not be deeper than the height of net to avoid fish escaping and, ideally, must not have large obstacles (e.g., branches, trunks, boulders) because the net would get stuck in it, resulting in tearing the net or giving opening and time for fish to escape while researchers unstuck it. Consequently, seine sites have similar habitat features. Since species have different habitat preferences, and that preferences can change in context of competition and/or predation (Fischer, 2000; Mittelbach, 1986; Werner & Hall, 1977), seine net might be indirectly selective on certain fish species and personality. Plus, species interactions are often size-mediated (Ortiz & Arim, 2016), and since older and bigger fish usually have a higher risk of infection than younger and smaller fish (Lemly & Esch, 1984), community and age structure might influence infection prevalence at community-level in our sites. Observational snorkeling transect is not a habitat-restricted nor a behavior-selective methodology. Although, quality of observations underwater relies on, meteorological conditions, water transparency (color, turbidity) and expertise of the observers to identify fish species in movement. As a result, precision of infection assessment can vary between sampling days and, between lakes. Low black spot abundance might be missed because of poor visibility or fish swift movement, leading to underestimation of prevalence estimates. Moreover, since infection assessment is easier in curious and solitary fishes as they usually get closer to the observers, accuracy of prevalence estimates might vary between species.

Sampling effort must be sizeable to measure adequate biological metric in natural systems. However, parasitological indices are often estimated with small sample size because of ethic, time, and monetary constraints. Our results showed that low sampling effort lean toward overestimating the prevalence value at landscape-scale (Figure 6). This evidence presence of infection hotspot that are more often sampled when sampling effort is limited. Parasites have an aggregated nature, where more individuals have low infection intensity, and few individuals are highly infected. This means highly infected individuals might occur in habitats favoring parasite occurrence and transmission. Thus, spatially heterogenous landscape should present infection hotspot where prevalence is high as it correlates with parasite abundance (Šimková et al., 2002). Although, Marques & Cabral (2007) found no significant effect of sample size on prevalence values, they reported that simulated median values tended to over or underestimate the real value at low sample size. Jovani & Tella (2006) also reported that low sample size resulted in greater inaccuracy and suggested an optimal sample size around 15 as a trade-off between losing data and maintaining accuracy. Our results evidence that minimum number of samples needed to adequately estimate landscape prevalence differed between methods. The only sampling method that reached a stable value was the observational snorkeling transects (around 10 samples) although the seine and methods combination showed less than 2% of difference between resampled prevalence and actual prevalence after 35 samples (Table S17). Indeed, snorkeling transects allow observations on more individuals than the fishing methods, explaining why it reaches stability faster and thus, need less sampling effort to measure accurate prevalence. 35 sampling efforts were not exhaustive enough to capture the actual infection landscape prevalence (Table S17). Actually, optimal sample size (or sampling effort) for estimate accurate parasitological indices can vary with parasite aggregation level (Poulin, 2007; Shvydka et al., 2018), host species target, sampling design (e.g., timing) and as we shown, sampling method.

**[Comments on spatial patterns]**

Our results shows that prevalence of infection is spatially heterogenous in our system (Figure 5). Across the landscape, lakes present varying values of prevalence and estimates also vary within lakes. Regardless, we did not find evidence of geographical patterns. Additional lakes might allow to confirm whether infection is spatially random, or patterns could not be detected because of our small lake sample size (N=14). Accordingly, none of the spatial attributes that we tested were good predictors of infection at small spatial scale (site-scale) (Table S18). Frequency distributions of lake’s prevalence varied between the different sampling methods (Figure 5), some suggesting a bimodal distribution tendency (combined methods and transect) and others with no obvious pattern. Bimodal frequency distributions are interesting because they imply that lakes are most often weakly infected or highly infected and so, there must be a threshold where parasites population growth overwhelm natural limitations. Differences in frequency distributions also imply that community prevalence estimate is method-dependent at lake-scale. Maps (Figure 5) clearly show that lakes’ prevalence can be over or underestimated by a given method. For example, in lake Croche, seine net overestimated the community prevalence while in lake Pin rouge, transect method underestimated the community prevalence (see Table S13). This evidence method-biases in field sampling of fish communities. If we take back the previous example, this means that the proportion of infected fishes sampled by seine net in lake Croche was higher than the proportion sampled by the other methods. So, the method might have selected fish species with high host specificity, sampled infection hotspots, sampled older individuals with higher infection exposure or sampled active and bold individuals that have a higher infection risk.

\* Frequency distributions Poulin 2002 (large-scale)

**[Comments on predictors]**

At the site-scale, we used GAMMs to inspect relationships between the prevalence estimate and environmental predictors (Figure 7, Table S18). Our results evidence some non-linear patterns supporting the use of flexible approach to model empirical (data-driven) infection dynamics. All physico-chemistry parameters (turbidity, temperature, dissolved oxygen, conductivity, and pH) were good predictors of infection prevalence at site-scale (Figure 7.A, B, C, D, E). Comparatively to infection assessment of fish populations and communities by field survey (or biological samples), water physico-chemistry measurements do not require a lot of equipment, time and are non-invasive to organisms. Thus, water physico-chemistry might be use as proxies of parasite infection for future monitoring and managing strategies. Consequently, we encourage researchers to investigate these relationships in other parasite species, geographical localities, and types of habitats as our data only snapshot the complexity of fish-parasite associations in natural systems.

Surprisingly, 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 = 84.17%) negative relationship between the macrophyte cover and the prevalence of infection in our transect sites (Figure 7.G). Contrary to our findings, heavier infection have been measured in fishes associated to vegetated zones in previous studies (Hartmann & Nümann, 1977; Marcogliese et al., 2001; Ondrackova et al., 2004). On one hand, because cercariae have a short life-span (Combes et al., 1994; Pietrock & Marcogliese, 2003), physical barriers might prevent the free-living stage from encountering a suitable host. Macroalgae have been shown to reduce the number of trematode cercariae (Bartoli & Boudouresque, 1997; Welsh et al., 2014) thus interfering with the parasite transmission to the next host. For instance, Prinz et al., (2009) demonstrated that inert seaweed can obstruct transmission of trematodes cercariae to their mussel-host. Similarly, Buss et al., (2022) suggested that microplastic fibers might reduce success of infection by impeding cercariae to encounter their amphibian-host. A treatment of complex shell surface of oyster also lead to reduction of 44% of cercariae in a experimental study (Welsh et al., 2014). In short, structural barriers might impede transmission by trapping cercariae, damaging them or consuming time of their short infective window. On the other hand, infected snails releasing cercariae in our lake system might not live or feed on macrophyte, explaining why we did not find a positive relationship between macrophyte cover and prevalence of infection in fishes. Indeed, if the association between macrophyte and snails doesn’t stand, there is no reason to think that increasing macrophyte would lead to an increase in cercariae release and accordingly prevalence of infection in the fish community. For example, the mud Amnicola (*Amnicola limosus*) is second intermediate host to at least one species of the black spot trematode guild in our lake system (Levet, unpublished). This freshwater snail species live partially in the lake the lake sediment from the littoral zone (Pinel-Alloul & Magnin, 1973) and often graze on periphyton (Kesler, 1981). Consequently, we would not expect a strong relationship with macrophyte cover.

Our results suggest that community structure is a driving factor of prevalence of infection by the black spot disease at small spatial scale. We found negative relationships for the total fish abundance, the non-host abundance and the Simpson’s diversity index (Figure 7.J, K, L). These results support the dilution effect hypothesis. Since total fish abundance and non-host abundance are positively correlated (Figure S1), we assume similar mechanism explaining why lower prevalence were measured in transect with higher fish abundance. A higher non-host fish abundance might reduce prevalence estimate in fish communities by an unsuccessful infection attempt on a non-host individual, by encountering individuals playing a structural barrier role (explained above) or by inducing a behavior change that reduce encounter rate with fish-hosts. For example, Ahn & Goater (2021) showed that abundance of brain worms decreased in fathead minnows (*Pimephales promelas*) when mixed with non-host emerald shinners (*Notropis atherinoides*), but that there were no difference when mixed with other non-host species, suggesting a behavioral change when *P. promelas* and *N. atherinoides* co-occur. This emphases the role of fish species identity in dilution effect mechanisms. Accordingly, diverse communities in our system tend to be less infected than communities with dominant species. In our lake system, *L. gibbosus* is a dominant species of littoral communities and the species with the higher infection prevalence. The presence of competitor species (e.g., *P. flavescens*) might then decrease the overall prevalence in the community by modify pumpkinseeds sunfish’s behavior that favor encounter with cercariae. Indeed, Dargent et al. (2013) found that presence and abundance of *Gyrodactylus* spp. (monogenean) was lower in hosts when they occur in mixed-species groups. Although it was not considered in this study, local diversity of other non-host organisms can influence the infection in fish. For instance, Lagrue & Poulin, (2015) found a negative association of non-host benthic invertebrates diversity with infection prevalence in second intermediate host. In fact, non-host species can reduce infection success of cercariae by actively or passively feeding on them (Thieltges et al., 2013). This has been observed in many taxon in aquatic ecosystems including zooplankton, Bivalvia, crustacean macroinvertebrates, cnidarian, and small fish (Kaplan et al., 2009; E. Mironova et al., 2019; K. Mironova et al., 2020; Schotthoefer et al., 2007; Vielma et al., 2019).