**1 | INTRODUCTION**

**[PAR1]** Context of the study

* Why parasites are important?
* Why is it important to study them?
* What are the gaps in the literature?

**[PAR2]** Spatial scaling

* Scale-dependence of species distribution
* Problem of single-scales studies

**[PAR3]** Biological scaling – Hierarchical organization

* Hierarchical organization of hosts and parasites
* Association levels that have not been studied

**[PAR4]** Deterministic filtering of infection

* Biotic (host characteristics and local community properties)
* Abiotic (local physical and chemical habitat)
* Spatial (spatial position and organization of the systems)

**[PAR5]** Sampling dependencies

* Sampling methods on metric estimations
* Sampling area

**[PAR6]** Study objectives

* Regional scale : resampling simulations with 3 methods
* Local scale : frequency distribution of lakes prevalence
* Fine scale : ecological predictors

**2 | METHODS**

*2.1. Host-parasite system*

* Trematodes infection in fish
* Life cycle + **Fig.1 (Life cycle)**
* Spatiotemporal variations in infection prevalence

*2.2. Study area and design*

* Sampling area + **Fig.2 (Prevalence map)**
* Proximity – weather
* Lake selection + **Table S1 (Geographic and Morphometric data)**
* Sampling weather
* Sampling effort + **Table S2 (Sampling effort determination)**

*2.3. Data acquisition*

* Temporality
* Fishing methodology + **Appendix S1 – Table S3 (Fishing gear dimensions)**
* Transect methodology
* Habitat description
* Water samples

*2.3. System description*

* Abundance data + **Appendix S2 – Table S1, S2, S3, S4 (ALL, MT, S, T)**
* Lengths + **Appendix S3 – Table S1, S2, S3 (Sp, Lake, Sp x Lake)**
* Infections + **Appendix S4 – Table S1 (References)** and prevalences + **Appendix S4 – Table S2, S3 (Local, fine-scale)**
* QGIS
* Results for water samples and physico-chem + **Appendix S5 – Table S1** **(Results table)**
* Habitat description + **Appendix S5 – Table S2 (Results table)**

*2.4. Statistical methods*

* RStudio
* Regional : Random resampling methodology
* Local : Histograms
* Fine-scale : GAMMs methodology, packages

**3 | RESULTS**

*3.2. Regional scale – random simulations*

* Accumulation curves (infectied, all and prevalence)

*3.3. Local scale – frequency distribution*

* Histogramme lake and transects?
* Map

*3.4. Fine scale – GAMMs*

* Summed plots of partial effects of significant models?
* Tab of all models, r2, p-value, and REML?

[PAR1] Parasites play significant roles in the regulation of host populations and are increasingly recognized as an essential component to understand how global changes will influence future population and community dynamics (REFs). Key local (e.g., abiotic) and regional (e.g., dispersal) predictors of fish parasite prevalence have been identified, using a variety of sampling tools and at different spatial scales (REFs). NEED SOMETHING PUNCHY HERE FOR THE FIRST PARAGRAPH (probably more something about multi-scale assessment of parasite prevalence in landscapes)

[PAR2] One overlooked constraint of using such a multi-scale approach is that the effect of predictors can only be compared across studies and across scales if infection prevalence scales linearly or does not change with sampling area (i.e., scale invariant). Depending on how parasite prevalence scale with sampling area, driving processes could vary with spatial scale, which would seriously hamper generalization on the identified drivers at one specific scale (local or regional). It is thus essential to characterize, with multiple sampling approaches, how parasite prevalence scale-up in sampling area, to know whether predictors of prevalence are scale-free or scale-dependant.

[PAR4] Several scenarios are possible. Infection prevalence is the fraction of the number of infected individuals over the total number of individuals in the sample. The number of individuals is generally assumed to increase linearly with sampling area (REFs), but it’s less clear for the number of infected individuals. In a well-mixed population, cumulative random draws should lead to a linear increase in infected individuals, but this would not be the case if infected individuals tend to be clustered in space. If one of the numerator or denominators increases non-linearly with sampling area, then prevalence will also scale non-linearly. If both increases linearly but not at the same speed, then prevalene will either decrease or increase linearly with sampling area. Finally, if both the numerator and denominator changes at the same speed (same slope), infection prevalence will not change with sampling area….

[PAR4] In this study we aimed to **i)** test how fish parasite prevalence scale with space**, ii)** test whether this relationship hold to different sampling approaches and **iii)** identify key drivers of fish prevalence for future management. Prevalence is measured as a fraction of number of infected on total population or community abundance. As such, only one scenario led to a scale invariant situation: both the numerator and denominator have to scale linearly with increasing sampling area. If this is the case, then we should expect that the relationship between sampling area and prevalence will be null (a flat line), which essentially mean that fish prevalence does not change with spatial scale…..

in the historical indigenous lands of Anishinabewaki ᐊᓂᔑᓈᐯᐗᑭ and Omàmìwininìwag (Algonquin) \*\*https://native-land.ca/\*.