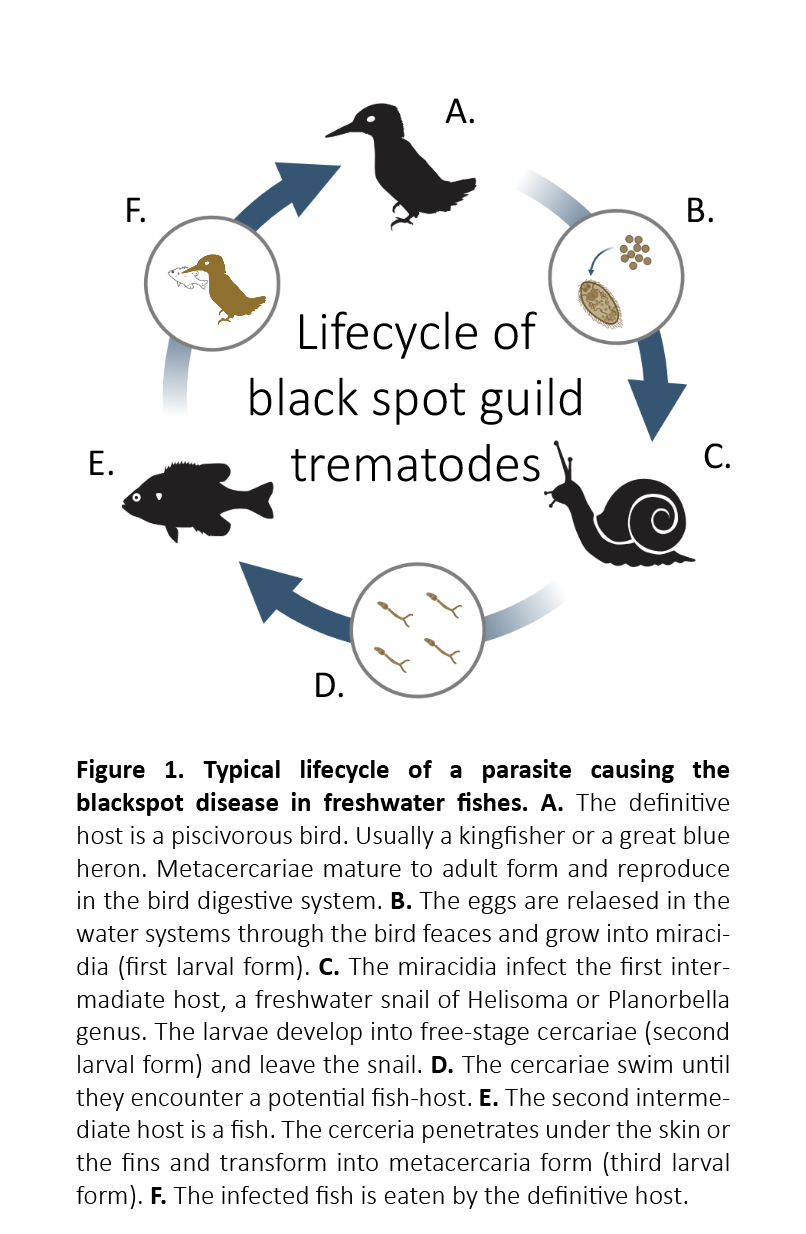
# 2 | METHODS

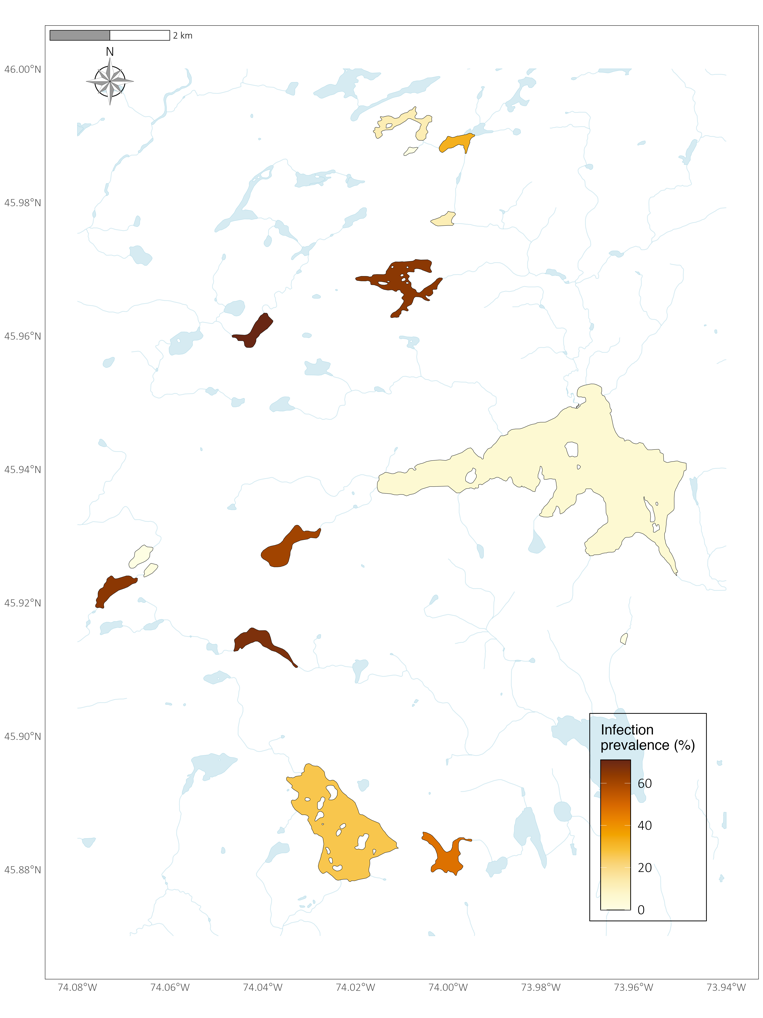
## 2.1. Host-parasite system

The black spot disease is a common infection in marine and freshwater fishes caused by some Digenean trematodes (flukes) (Kurochkin & Biserova, 1996). Some of these parasites (*e.g., Posthodiplostomum cuticula, Uvulifer ambloplitis*, *Crassiphiala bulboglossa* and *Apophallus brevis*) have been studied for a long time and yet, spatiotemporal infection patterns and their relation to environmental conditions are elusive. Black spot causing trematodes all have a similar complex life cycle requiring a snail, a fish and a piscivorous bird as hosts (Figure 1). The fish is the second intermediate host of the parasite where it encysts under the skin, in the fins or the muscles (Hoffman, 1956; Krull, 1932, 1934). Because the parasite is dormant in its fish-host, it can survive many years while it waits to be eating by a piscivorous bird (Hoffman & Putz, 1965) leading to increased parasite load with fish age (Lemly & Esch, 1984). The distinctive black cyst results of the melanin pigment stimulation by the larva penetration in the fish body (Davis, 1967; Lemly & Esch, 1984). Because they have similar requirements and cause the same symptoms, we will refer to black spot trematodes as a guild from now on. Akin to other parasitic diseases, infection parameters can vary across time, space, and species. For example, black spot disease prevalence in Bluegill sunfish (*Lepomis macrochirus*) living in a pond reaches its minimum around April/May and its maximum around September/October (Lemly & Esch, 1984). Because of the snail first intermediate host requirements such as macrophyte and shallow waters, infection levels in fish-host are usually higher is lentic rather lotic systems and in the littoral zone rather than pelagic zone because it maximize the encounter window (Ondrackova et al., 2004). Because black spot infection is caused by a parasite guild (more than one trematode species), a large range of fish can show symptoms of this infection. However, within a system, some species tend to be more heavily infected than others. For instance, Lemly & Esch (1984) found that Bluegills were more heavily infected than Largemouth bass (*Micropterus salmoides*).



## 2.2. Study area and sampling design

We sampled 15 lakes from six watersheds in Saint-Hippolyte, QC, Canada (Figure 2). Corriveau, Croche, Cromwell and Triton lakes are in the protected area of the biology station of the Université de Montréal while the other lakes are surrounding by private properties. Considering their geographical proximity, we assume that all lakes are exposed to the same climatic conditions. The lakes were selected nonrandomly according to accessibility and availability of morphometric data (see Table S1). Field work was restricted to unrainy days to avoid sampling bias due to meteorological effects.



**FIGURE 2**

Map of the study area. The lakes sampled are colored according to the the black spot disease infection prevalence in the local fish communities from the littoral zone. The data used to estimate infection prevalence comes from all the methods combined.

The number of samplings within lakes was determined according to lake area except for minnow traps (Table S2). As it needed less time and manipulations, all the 15 minnow traps were set in each lake to maximize the number of captures.

## 2.3. Data acquisition

Three sampling methods were used to assess prevalence infection and maximize chances of catching species richness. Fishing sampling took place from mid-June to end of July 2022. A seine net was used for daytime sampling and two size of minnow traps were set for dusk sampling (see Table S3 for fishing gear dimensions). Minnow traps were set at approximatively equal distance along the shore, from 4PM to 8PM, to target species that are most active at dusk. Half of the minnow traps were baited (3 large and 5 small traps) to sample diverse feeding preferences and habits. All fishing gear was cleaned between lakes following MFFP recommendations to prevent exotic species contamination.

All fishes were counted, identified to species level, and measured (estimation of the total length to the nearest centimeter) directly after capture and released afterwards. Northern redbelly dace (*Chrosomus eos*) and Finescale dace (*Chrosomus neogaeus*) individuals were grouped into *Chrosomus spp*. naming as the two species hybrid in our system and cannot be distinguished based on morphology in the field (Leung et al., 2017). The presence of black cysts was assessed on the left side of the fish (De Bonville, *in prep*.). Both juveniles and adults were considered in this study as all individuals are vulnerable to black spot infection. Animal handling was approved by the Université de Montréal’s animal care committee (protocol number 22-025) and scientific fishing permit was delivered by the Ministère des Forêts, de la Faune et des Parcs (MFFP) of Québec (2022-05-16-1971-15-S-P).

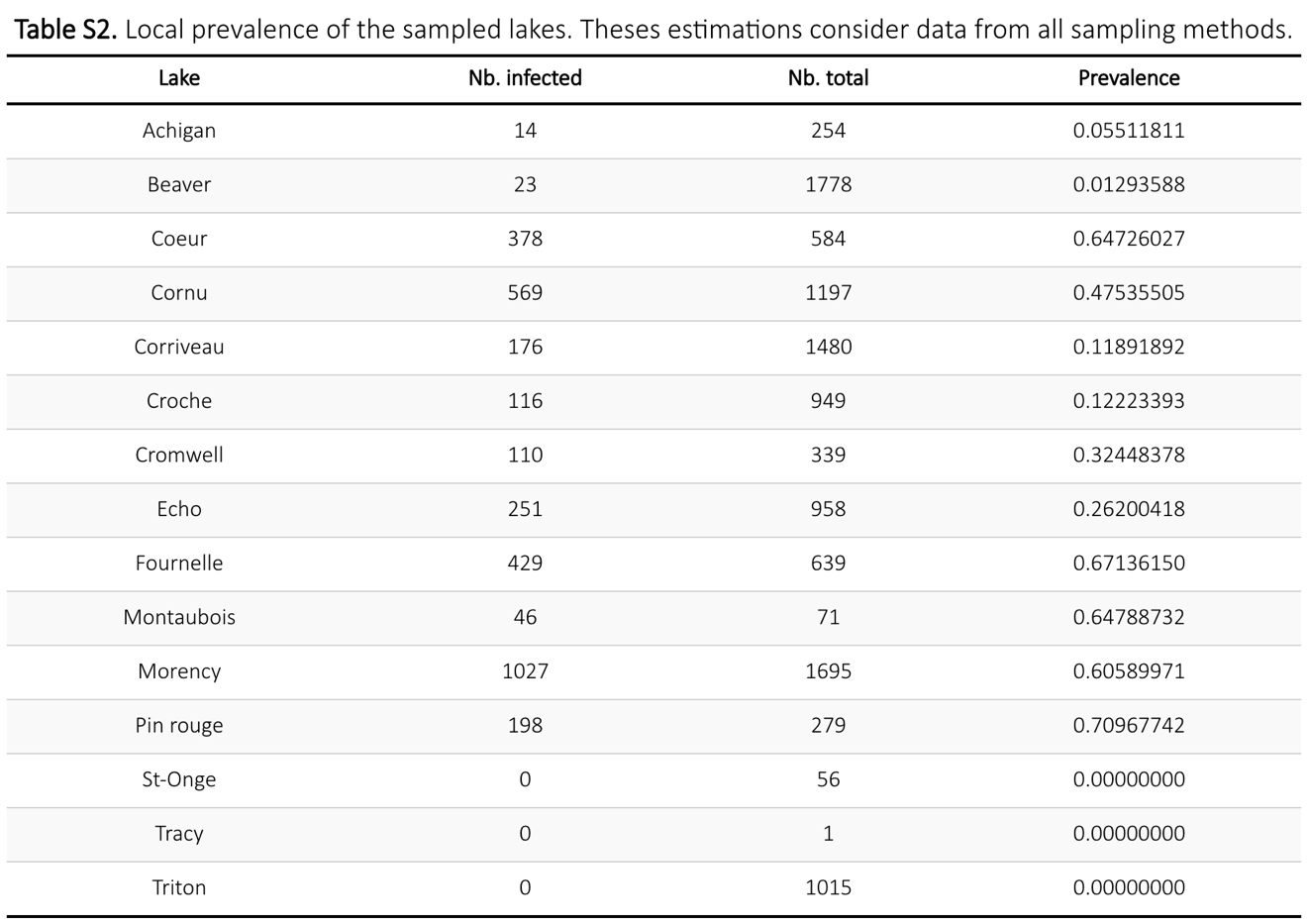
50 meters snorkeling transects along the lakes shore were made to assess black spot infection in the fish community in August 2022. Due to time constraints, no transects were made in lake Beaver, Montaubois, Tracy and St-Onge were excluded of transect sampling because of null infection level or low fish abundance in fishing sampling. For the site selection, we considered emplacements that were approximately between 0.5 and 3 meters deep, not fully covered by macrophytes, not obstructed by docks and preferably with some vegetal, rock or trunk refuges for the fishes. The sites were first chosen and flagged at every 10 meters in end-May. Transects were done by two observers at the time covering each a field of view of 1 meter radius and moving forward at a pace of 3 minutes by 10 meters for a total time of 15 minutes by transect. The fishes coming in the field of view from behind the observer were not counted. All fishes were identified to the lower taxonomic level possible and black spot infection was noted. All cyprinids were grouped into the same category as they are more difficult to identify to species level and usually move faster. Following the prevalence assessment, the description of habitat characteristics was made for every transect. For each 10 meters, the percentage of substrate category (silt, sand, rock and metric block), the coverage of macrophyte, the number of trunk (or large branch), and the mean depth was estimated by each observer. Temperature, dissolved oxygen, conductivity and pH were measured at mid-depth at the beginning of every transect with a YSI ProDSS Multi-Parameter Water Quality Meter. 1L of unfiltered water was taken in the field at mid-depth in previously acid-washed HDPE bottle for carbon and nutrients content analysis. Samples were placed in a dark cooler until brought back to the laboratory where the 1L sample was separated in previously acid-washed 40 mL vials for total organic carbon (TOC), and 500 mL HDPE bottle for total nitrogen (TN) and total phosphorus (TP). TOC samples were placed in 4°C refrigerator until analyze within the week while TN and TP samples were kept in -20°C freezer until processing.

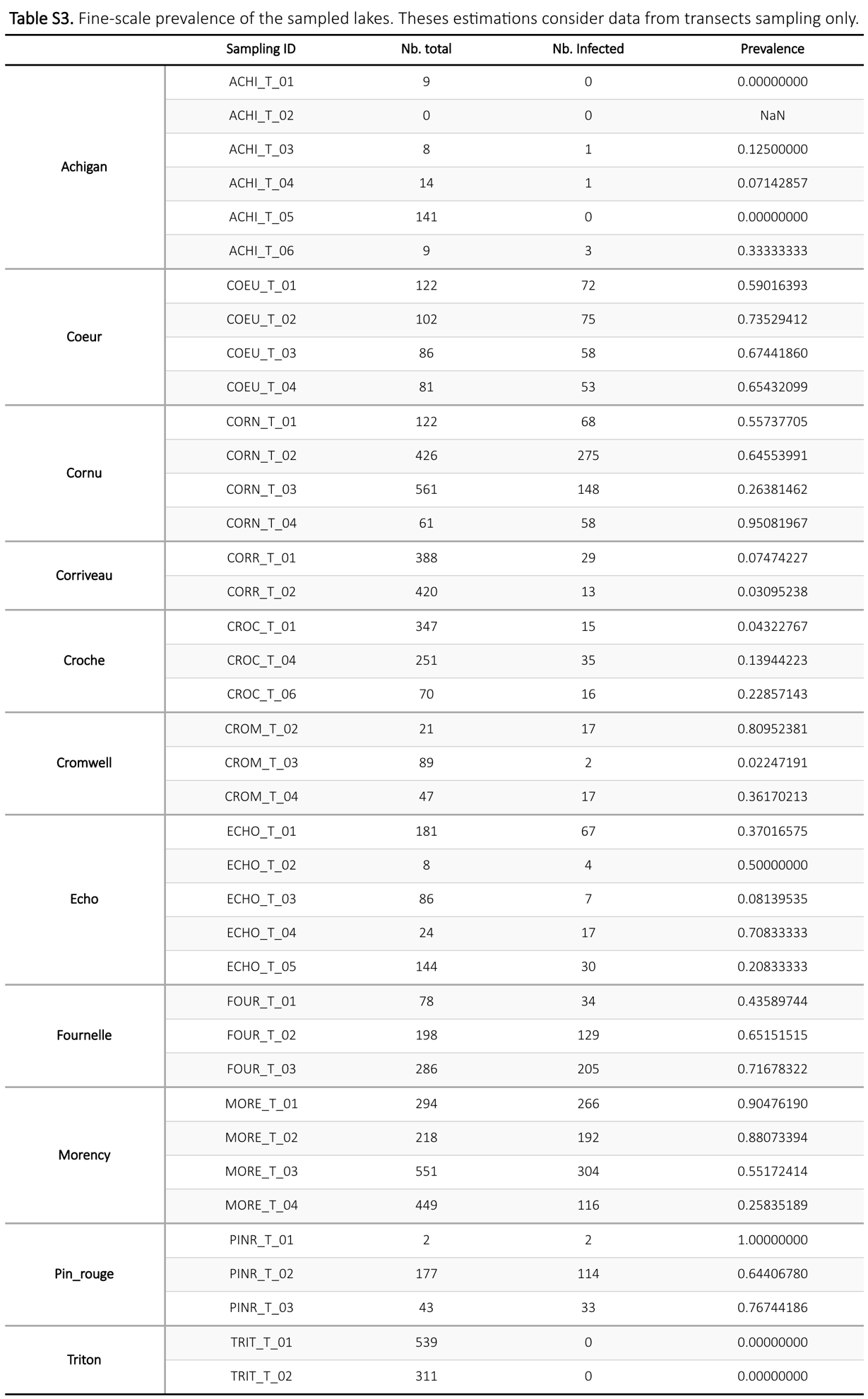
## 2.3. System description

We recorded a total of 11 295 individuals belonging to 15 species for this study (Table S4).

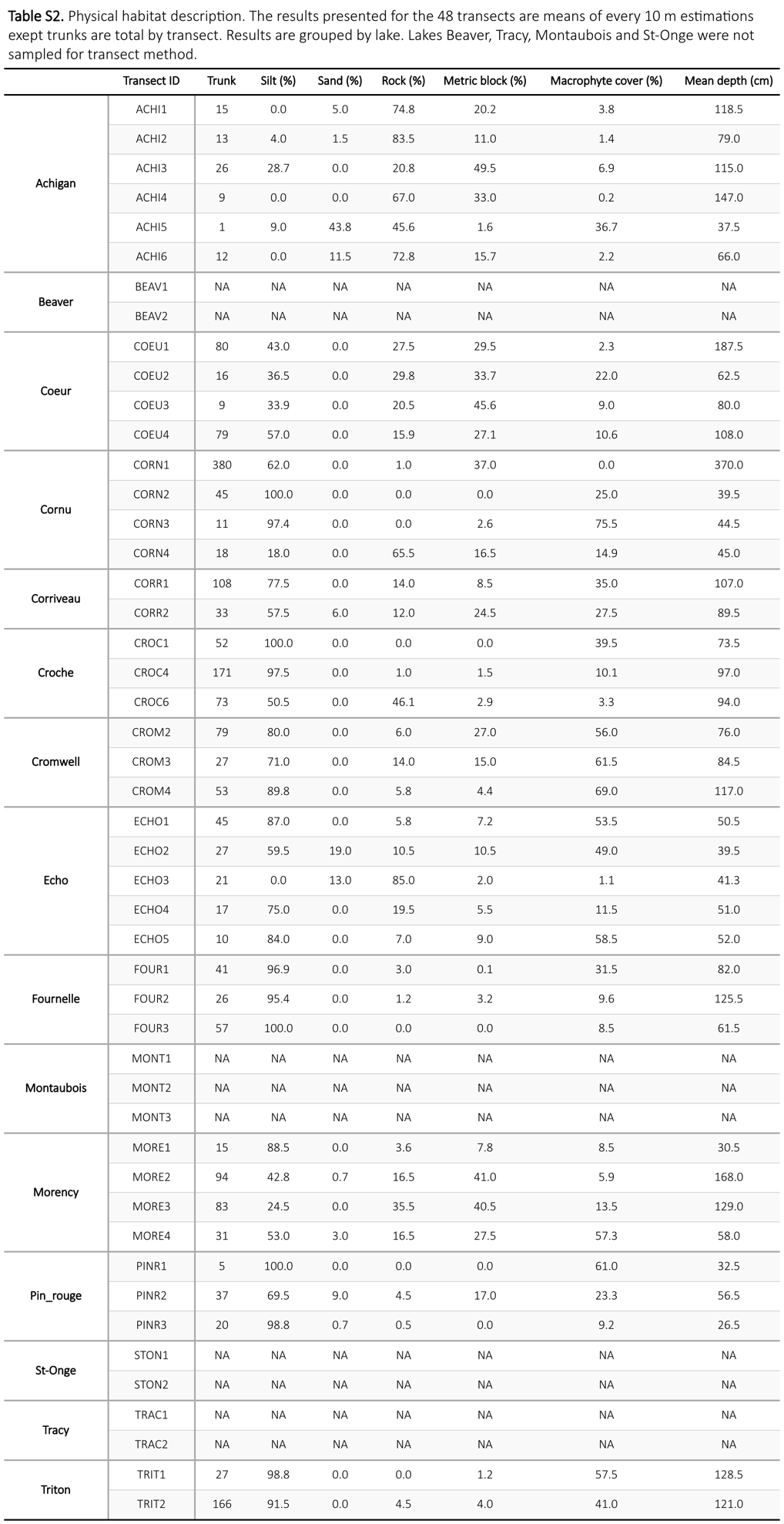
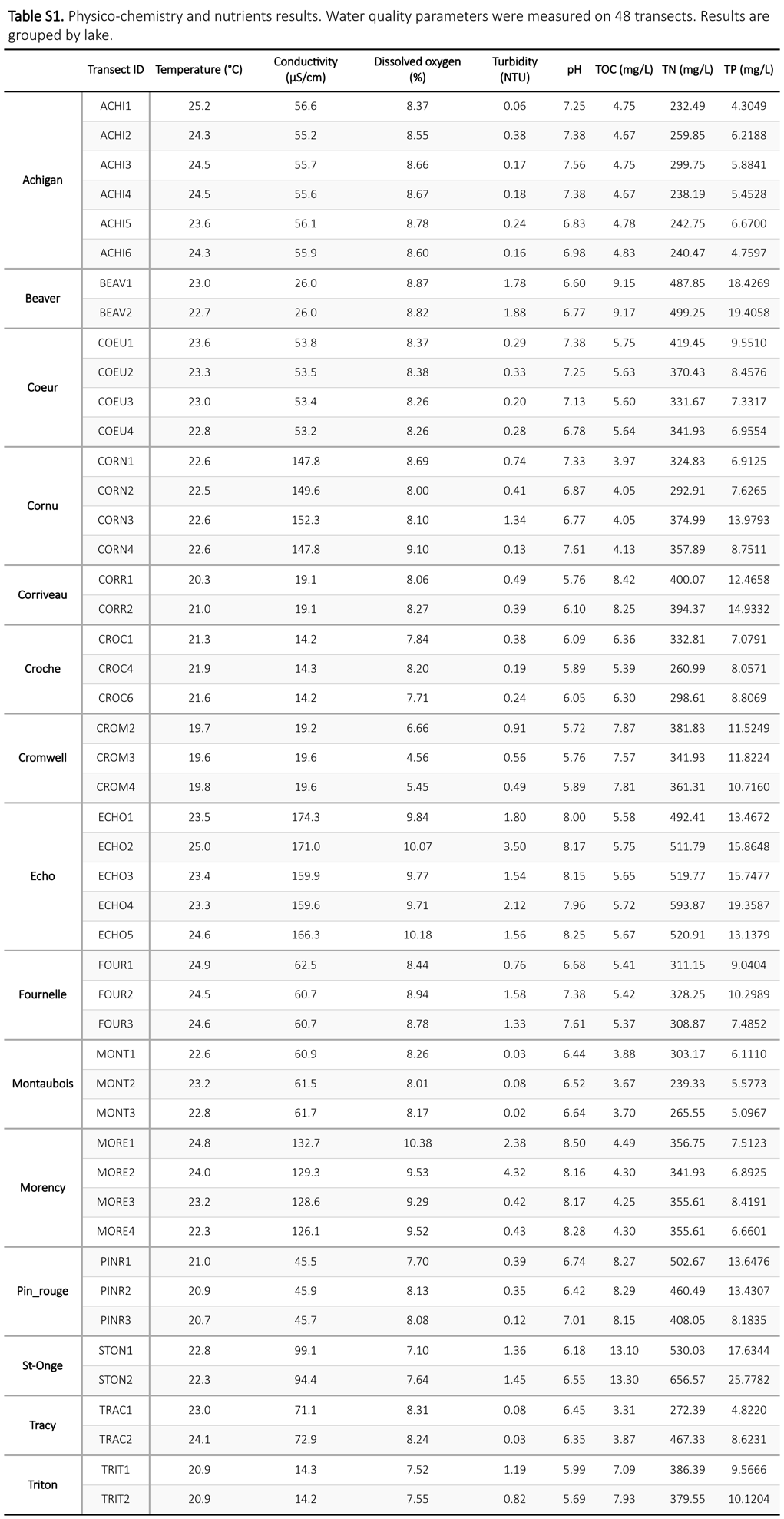
The minnow traps caught 1904 individuals from 10 species while seine nets caught 2427 individuals from 14 species (Table S5 and S6). 6964 individuals belonging to 5 taxonomic groups (4 species and 1 family) were observed in snorkeling transects (Table S7). The mean length of all fish captured through fishing methods was 5.59 ± 2.96 cm (N = 4333). Overall mean length for each lake, species and, species within each lake are presented in Appendix S3 (Table S8, S9 and S10).

In the context of this study, the infection prevalence is calculated as the number of individuals infected individuals a species of the black spot guild, divided on the total number of individuals in a given fish community. All fish species in our lake system, except for *Ameiurus nebulosus* and *Esox masquinongy* are susceptible to the black spot disease (see Table S11). Based on data from all sampling methods, the regional community prevalence is 29.54 %, local (i.e. lakes) prevalence vary between 0% and 71% (Table S12) and fine-scale prevalence (i.e. transects) vary between 0% and 100% (Table S13). Our survey corroborates black spot infection in *Lepomis gibbosus*, *Ambloplites ruspestris*, *Perca flavescens*, *Micropterus dolomieui*, *Semotilus atromaculatus* and *Pimephales promelas* (see Table Sx for more details).





Lake perimeter and distance to nearest lake (centroid to centroid) were measured with QGIS 3.28 Firenze (Appendix S1 – Table S1). TOC samples contents were measured on … autosampler. TN and TP samples were oxidized with persulfate into the autoclave the day before analysis (EPA353.2; EPA365.3) TN concentration was measured on a Lachat QuikChem 8500 analyzer (EPA353.2) and TP concentration were analyzed on a Asoria-Pacific Astoria2 (EPA365.3). Results for TN and TP concentrations are presented in Appendix S5 – Table S1 along with physico-chemistry measurements. Mean habitat descriptions are presented in Appendix S5 – Table S2 for each transect.



## 2.4. Statistical methods

To describe the black spot prevalence pattern across multiple scales, we employed the following procedures. All data manipulation and analysis were conducted on RStudio (Version 2023.06.2+561). Lake Tracy was excluded from analysis because of low abundance data (N = 1).

For the regional-scale analysis, we wanted to look at the effect of the sampling design (here sampling method and number of samplings) on the estimation of a regional prevalence. We used a resampling approach on the data from different sampling methods minnow trap, seine nets, transects and all methods combined) to visualize the accumulation of i) the infected individuals, ii) the total individuals and, ii) the infection prevalence through an increasing sampling gradient. Number of samplings (N) were set at N = c(1, 2, 3, 5, 7, 10, 15, 20, 25, 35) to not exceed the actual number of sampling (smallest to be 39 sampling for the transect method). Abundance data (infected and total individuals) were pooled together, and prevalence was calculated for every sampling. The data set was then split according to the sampling method. N random samples were drafted from the pooled data and, for the infected and tot individuals, the sum was calculated while the mean was calculated for the prevalence. The operation was replicated 999 times for each N and sampling methods. A linear regression was then fitted to the mean of every value sampled for each number of samplings.

For the local-scale analysis, we examined the frequency distribution of the lakes’ prevalence according to the sampling method. The mean infection prevalence was previously calculated, and visualization was made with histograms.

# For the fine-scale analysis, we modeled the relation between the environmental variables and the infection prevalence on a transect level. We used generalized additive mixed-effects models (GAMMs) from the *mgcv* package with a quasibinomial distribution, a maximum likelihood estimation and the lake as a random factor. Because of the small sample size (N = 39), we included only one environmental variable at the time to save degrees of freedom. The model validations were conducted with *gratia* package. The partial effects visualizations were made with *gratia* and *ggplot*2 package.