TRACER LES SOURCES DE CONTAMINATION FÉCALE DE L'EAU

ANIMAL SOURCE TRACKING BY MITOCHONDRIAL METAGENOMICS: APPLICATION FOR FECAL CONTAMINATION IN SURFACE WATERS

Rose Ragot, doctorante

Richard Villemur, Professeur titulaire

INRS Centre Armand-Frappier Santé Biotechnologie

eDNA profiling of mammals, birds, and fish of surface waters by mitochondrial metagenomics: application for source tracking of fecal contamination in surface waters

Rose Ragot & Richard Villemur ✉

Feces

- Result from food digestion

- Composed of
  - Host cells
  - Microorganisms (bacteria, protozoan, viruses)
  - Digestion remains

- 55% dry weight : microorganisms
  - >500 bacterial species
  - Essential for digestion

- Some microorganisms are pathogens

Fecal contamination

- High potential of dispersion of pathogens
- In surface waters and groundwater
  - Contamination of sources of drinkable water
  - Recreational impacts (e.g. beach closure)
  - Food safety (e.g. Irrigation)
- => Public health and economical issues

- 2 billion people use water points contaminated with feces.
- Recurrent epidemics: Giardiosis, Cryptosporidiosis, Cholera, Gastroenteritis etc.
- Contaminated water responsible for more than 485,000 deaths per year according to the WHO.

Fecal contamination detection

- Impossibility to detect all pathogens
  - Cost
  - Too many
  - Low concentration
  - Detection method Inefficiency for some

- Instead : Monitor fecal indicator bacteria (FIB)
  - Provide indication that fecal contamination occurred
  - May indicate probability of the presence of pathogens

FIB (or viruses)

- Coliformes
- *Bifidobacterium, Enterococcus, Bacteroides*
- Enteric viruses (Adenovirus, enterovirus, norovirus)
- Phages

- The FIB cannot tell about the animal source
  - Most microorganisms are present in feces of animals

Many sources

- Contamination diffuse
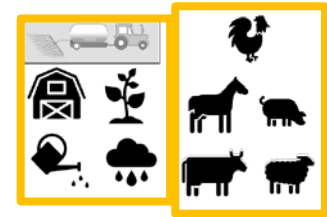- Sampling sites far from the contamination source

Wild

birds
(e.g. geese)

Municipal

- Wastewater treatment plants
- Leaking septic tanks
- Sanitary sewer overflows

Agriculture

Manures

Pathogens

Source tracking markers

- Biochemical markers
  - Fecal sterols
  - Trace of detergent (bleaching agent)
  - Caffeine
- Microbial markers
  - Animal specific microorganisms
- Target specific genetic elements of these microorganisms
  - PCR amplification technology

- Good knowledge of the microbial biota of the animal feces
  - Not obvious with wild animals
  - Can be variable in geographical distribution and could depend of alimentation of the animals
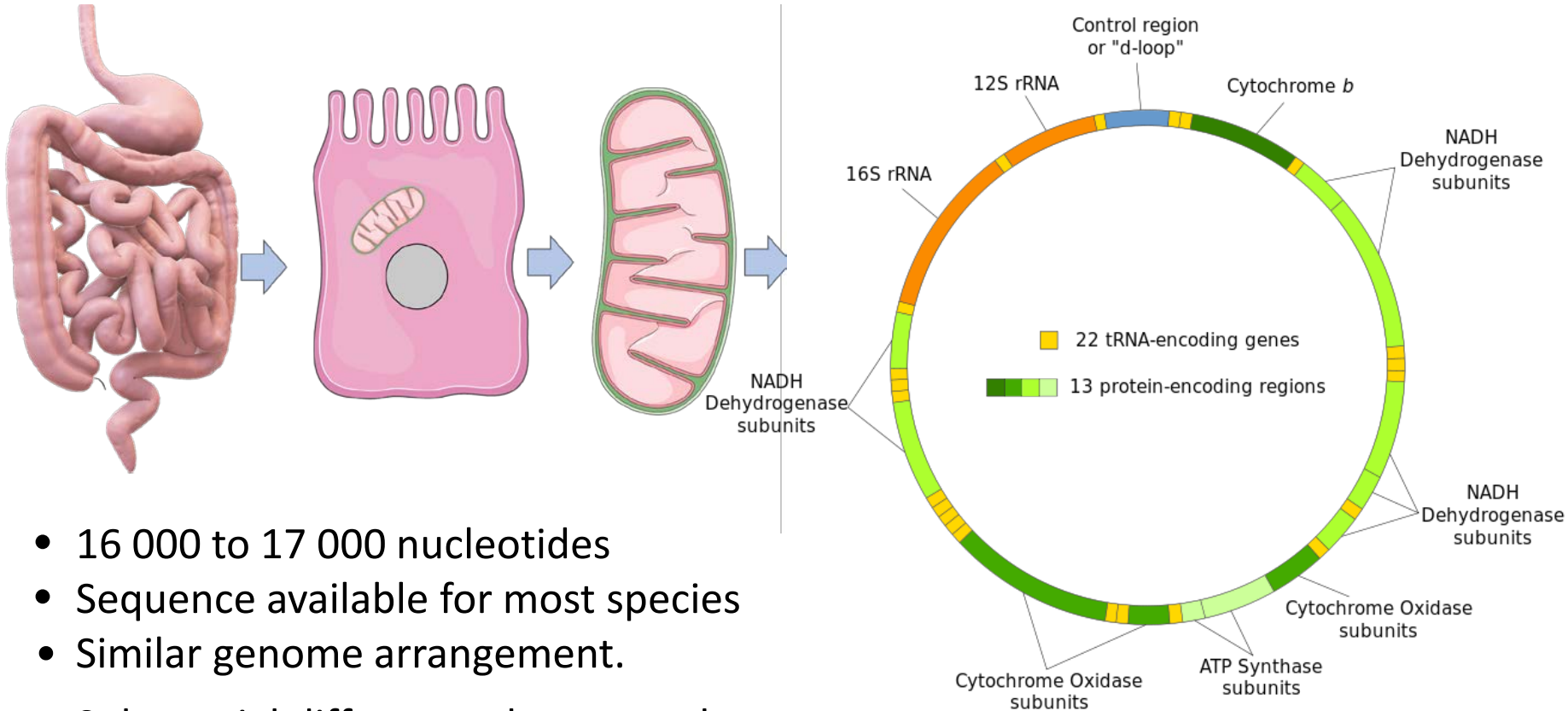
Mitochondrial DNA as Source tracking markers

Gastrointestinal epithelium

- 50 billions of cells
- Renewed each 3-4 days
- 10 billions of cells dejected in the lumen
- Represent 10% of dry weigth feces.
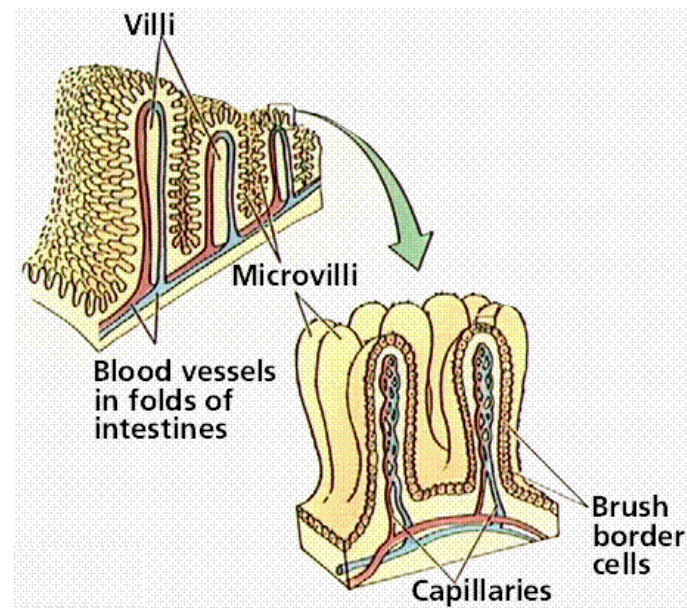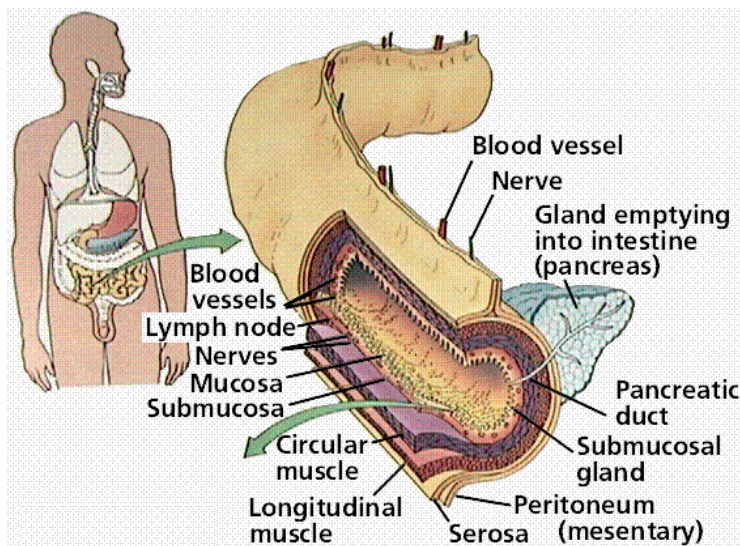
Vertebrate mitochondrial genome (mtDNA)

- 16 000 to 17 000 nucleotides
- Sequence available for most species
- Similar genome arrangement.

- Substantial differences between close animal species
  - 9% differences between human and chimpanzee.

Problematic

- Multiple sources of fecal contamination
  - Human activities overlapped (municipal vs agricultural areas)
  - Wild animals also involved
- Multiple specific markers to develop
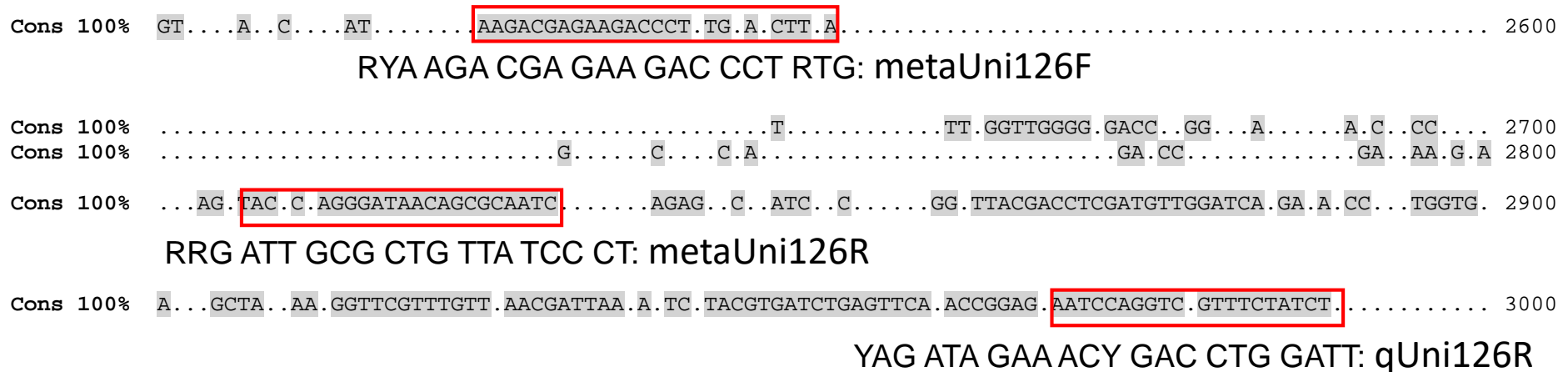- Multiple PCR to carry

New approach

- With high-throughput sequencing methods, it is now possible to sequence a PCR product with a multitude of sequences

- Found PCR primers targeting vertebrate mtDNA

- Discriminate animal species by sequencing

Chosen animals

- Occurrence of mammals and birds in North America
  - human, domestic animals (cat, dog), livestock (swine, bovine, ovine, poultry, and farm-raised exotic animals such as llama, ostrich, and emu).
- Wild terrestrial animals that are commonly encountered on riverbanks
  - Birds (e.g., goose, ducks, and gulls), raccoon, muskrat, beaver, elk, caribou, and deer.
- Other mammalian and bird species to broader the diversity.
- Fish species chosen for their occurrence in the rivers of the Province of Quebec and in aquaculture were also included.
- The inclusion of fishes in our study was to assess the importance of their mtDNA in the river samples
- 126 mitochondrial genomes
  - 46 bird species
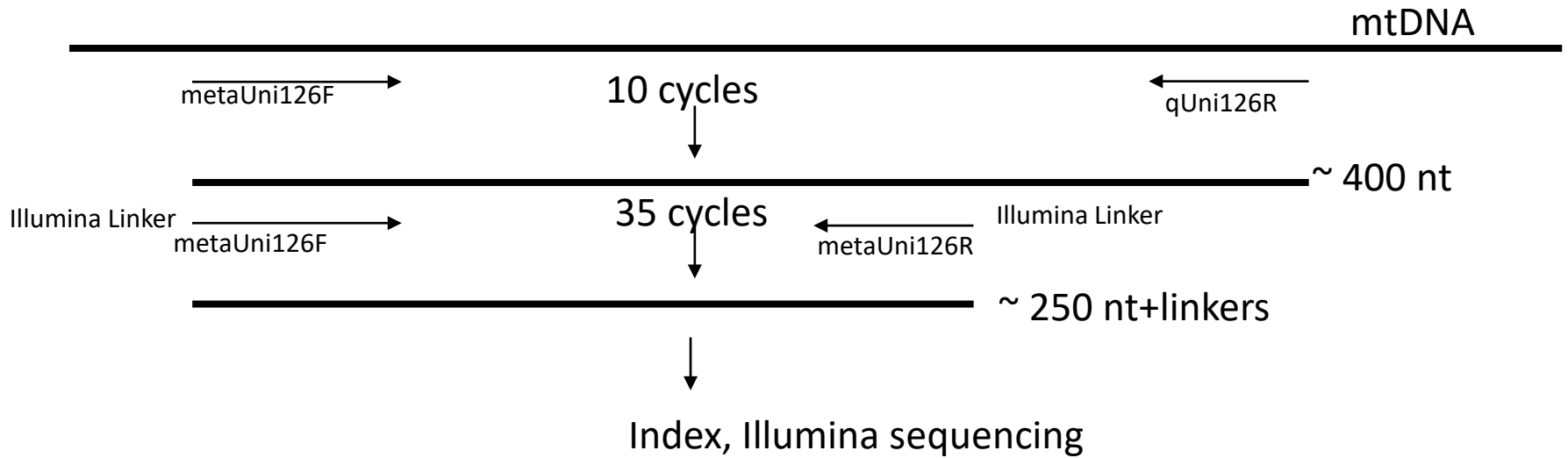 - 62 mammals
  - 12 fish species

PCR primers

- Alignment of the 126 genomes

- One region fit the criteria: mitochondrial 16S ribosomal RNA (rRNA) gene

```
Cons 100%   GT....A..C....AT.......AAGACGAGAAGACCCT.TG.A.CTT.A...........................................  2600
                                   RYA AGA CGA GAA GAC CCT RTG: metaUni126F

Cons 100%   .....................................................T.........TT.GGTTGGGG.GACC..GG...A......A.C..CC....  2700
Cons 100%   .......................................G.......C....C.A........................GA.CC..............GA...AA.G.A  2800

Cons 100%   ...AG.TAC.C.AGGGATAACAGCGCAATC.......AGAG..C..ATC..C......GG.TTACGACCTCGATGTTGGATCA.GA.A.CC...TGGTG.  2900
                RRG ATT GCG CTG TTA TCC CT: metaUni126R

Cons 100%   A...GCTA..AA.GGTTCGTTTGTT.AACGATTAA.A.TC.TACGTGATCTGAGTTCA.ACCGGAG.AATCCAGGTC.GTTTCTATCT...........  3000
                                                                          YAG ATA GAA ACY GAC CTG GATT: qUni126R
```

Gray: Nucleotides found in the 126 genomes
Dot: Nucleotides specific for each lineage

Strategy: Nested PCR

mtDNA

metaUni126F   10 cycles   qUni126R

~ 400 nt

Illumina Linker   35 cycles   Illumina Linker
metaUni126F   metaUni126R

~ 250 nt+linkers

Index, Illumina sequencing

Possibility of qPCR assay

metaUni126F   qUni126R
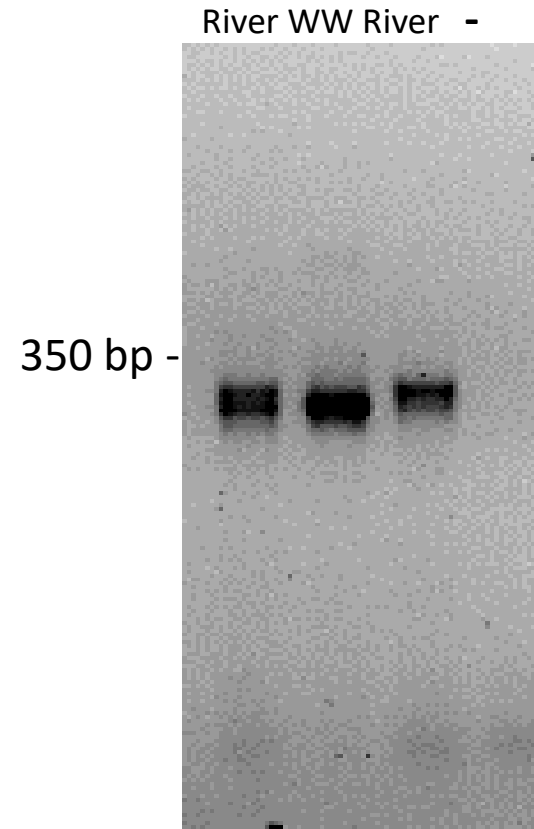
PCR products generated by primers targeting the mitochondrial 16S rRNA genes

- Agarose gel electrophoresis
- No non-specific amplification

River WW River -

350 bp -

Percentage of mitochondrial genomes containing the consensus sequences

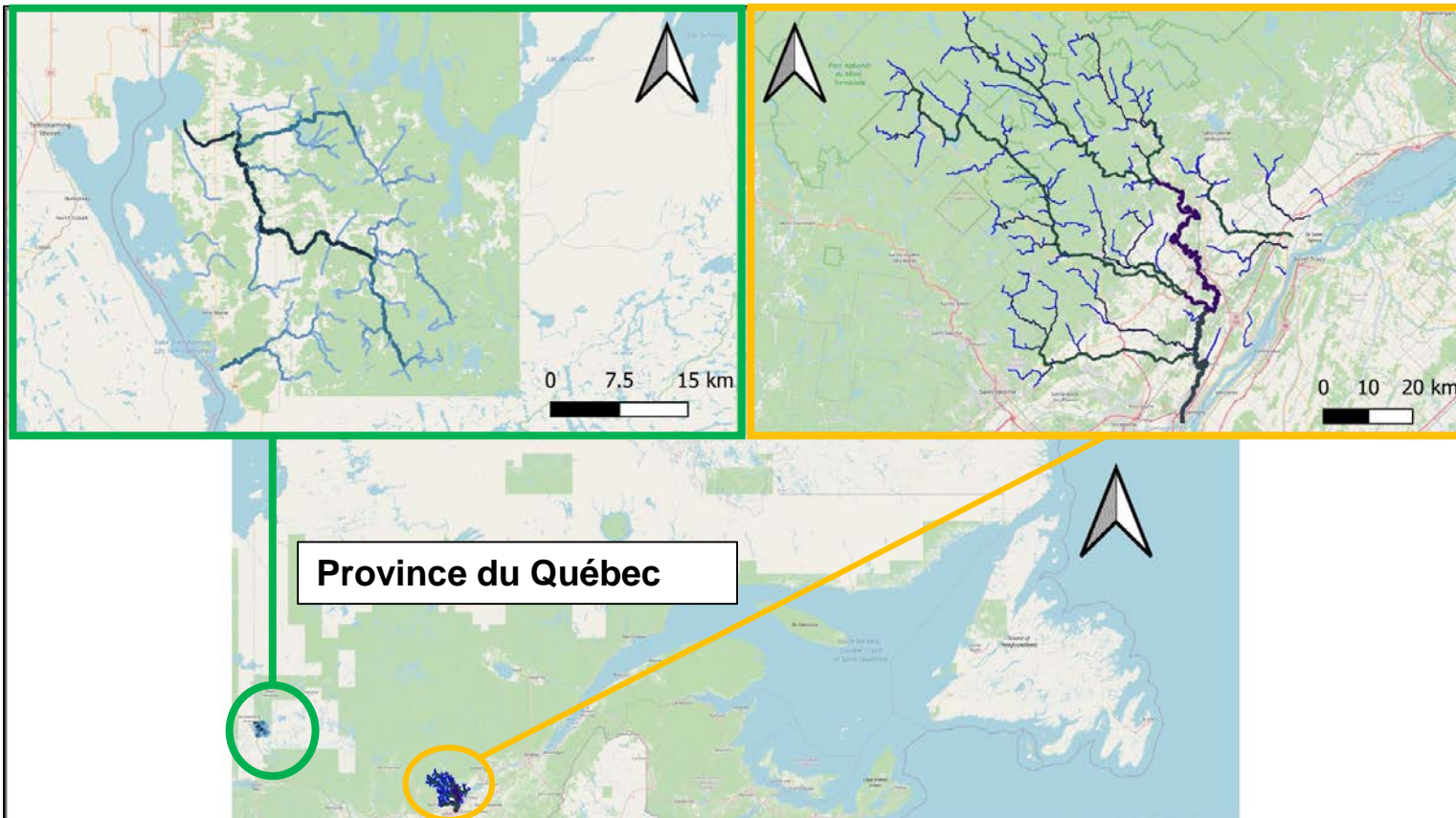| Lineages		metaUni126F	metaUni126R	qUni126R
Mammalia		100.0%	99.9%	99.8%
Lepidosauria	Snake, lizard …	48.5%	21.7%	73.8%
Archelosauria				
Testudines	Turtle	94.9%	96.5%	95.1%
Archosauria	Bird, crocodilian	94.0%	98.0%	94.4%
Amphibia		51.3%	52.1%	88.9%
Actinopterygii	Ray-finned fish			
Euteleosteomorpha	Bony fish	92.6%	93.2%	96.0%
Otomorpha		6.9% [1]	98.7%	98.8%
others		92.5%	100.0%	100.0%
Chondrichthyes	Cartilaginous fish	89.7%	89.3%	90.5%
Cyclostomata	Fish no jaw	84.6%	78.6%	73.3%
Insecta		0.4%	0.1%	0.0%

Sampling process

- Water samplings (100-500 mL)

- Filtering and DNA extraction

- PCR amplification and sequencing by Illumina technology Miseq PE-250

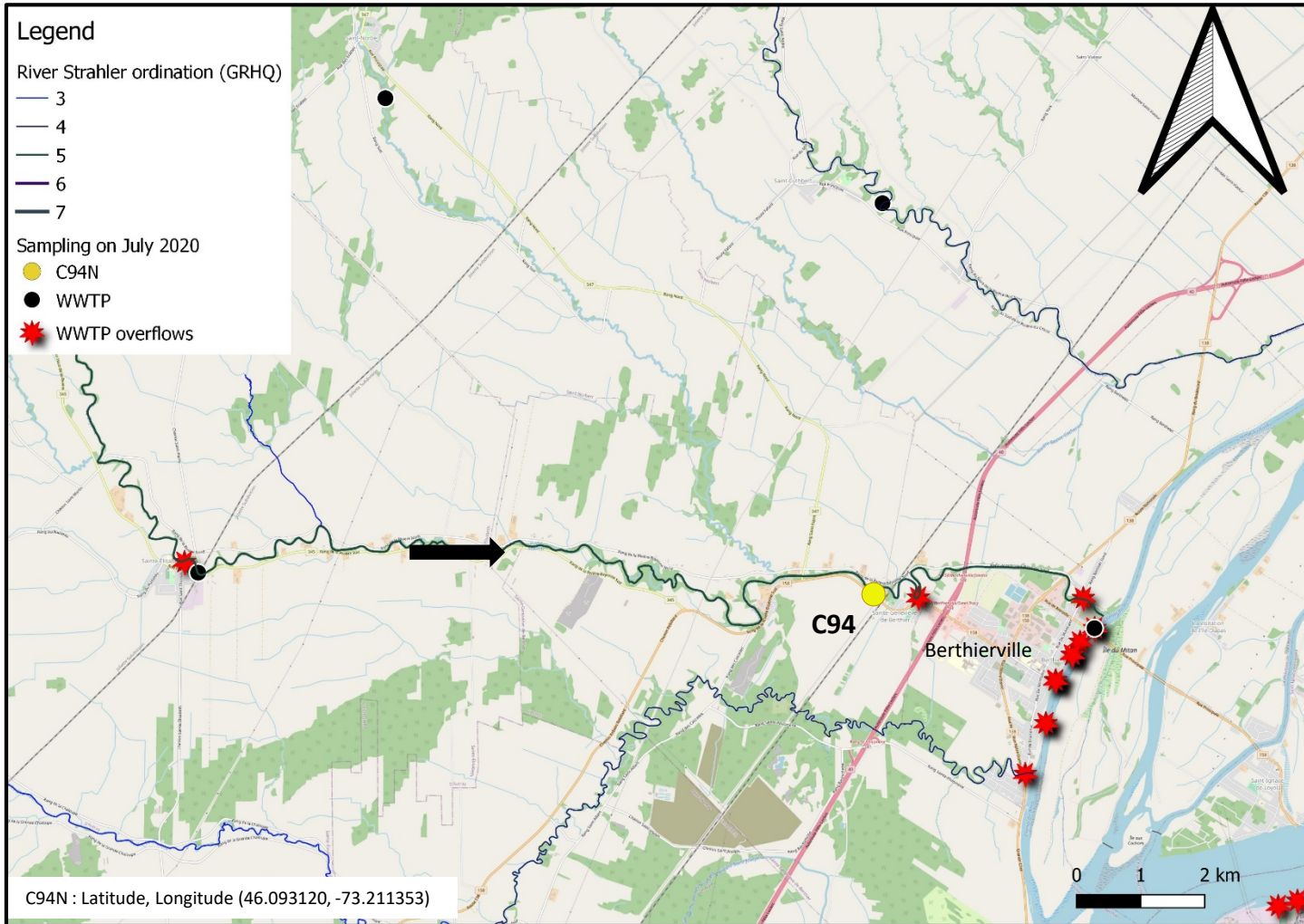- Sequence clustering (dada2 and CD-HIT-EST at 95% identity for species affiliation)

Sampling areas

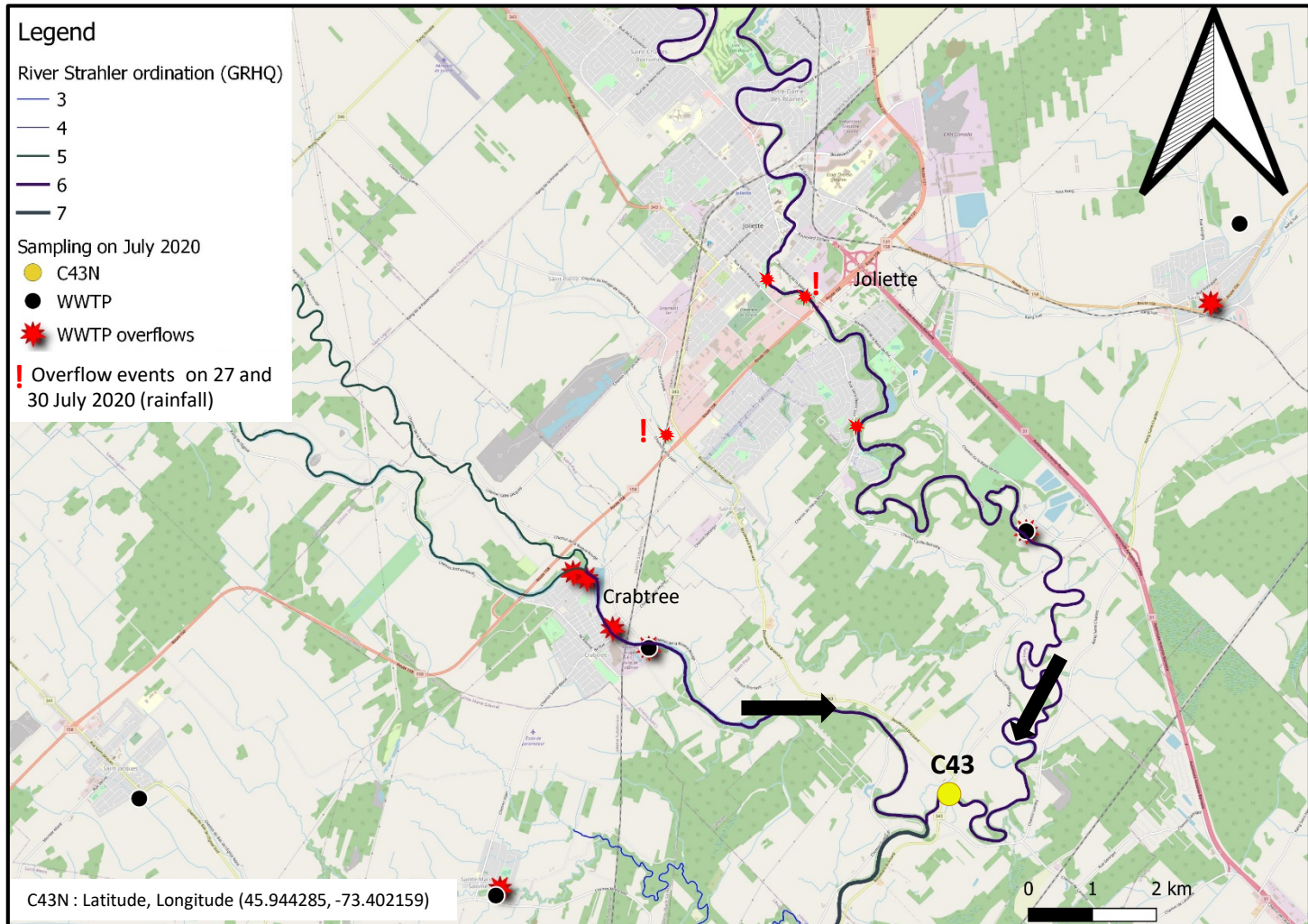**Témiscamingue watershed**
27 samples, November 2020.

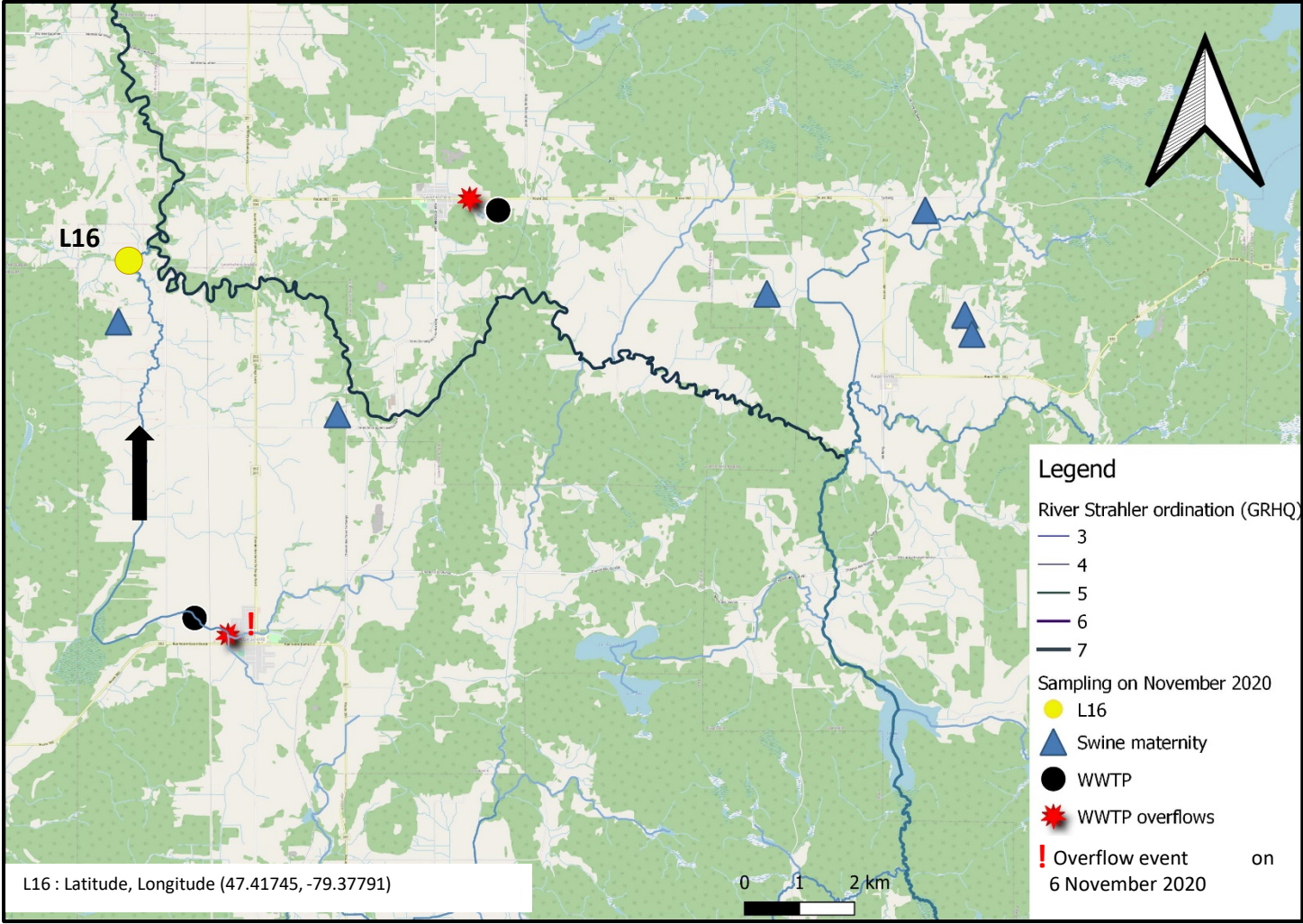**L'Assomption watershed**
59 samples 2019 to 2020.

Province du Québec

Bayonne River (C94)

Legend

River Strahler ordination (GRHQ)
— 3
— 4
— 5
— 6
— 7

Sampling on July 2020
● C94N
● WWTP
✳ WWTP overflows

C94

Berthierville

C94N : Latitude, Longitude (46.093120, -73.211353)

0   1   2 km

L'Assomption River (C43)

Legend

River Strahler ordination (GRHQ)
— 3
— 4
— 5
— 6
— 7

Sampling on July 2020
● C43N
● WWTP
✳ WWTP overflows

❗ Overflow events on 27 and 30 July 2020 (rainfall)

Joliette

Crabtree

C43

C43N : Latitude, Longitude (45.944285, -73.402159)

0   1   2 km

Dumais Stream-Témiscamingue (L16)

L16

Legend

River Strahler ordination (GRHQ)
— 3
— 4
— 5
— 6
— 7

Sampling on November 2020
● L16
▲ Swine maternity
● WWTP
✸ WWTP overflows
! Overflow event   on
6 November 2020

L16 : Latitude, Longitude (47.41745, -79.37791)

0   1   2 km

Bastien Stream-Témiscamingue (L17.3)

L17.3

L17.3 : Latitude, Longitude (47.47426, -79.37558)

Legend

River Strahler ordination (GRHQ)
— 3
— 4
— 5
— 6
— 7

Sampling on November 2020
● L17.3
▲ Swine maternity
● WWTP
✳ WWTP overflows

0   1   2 km
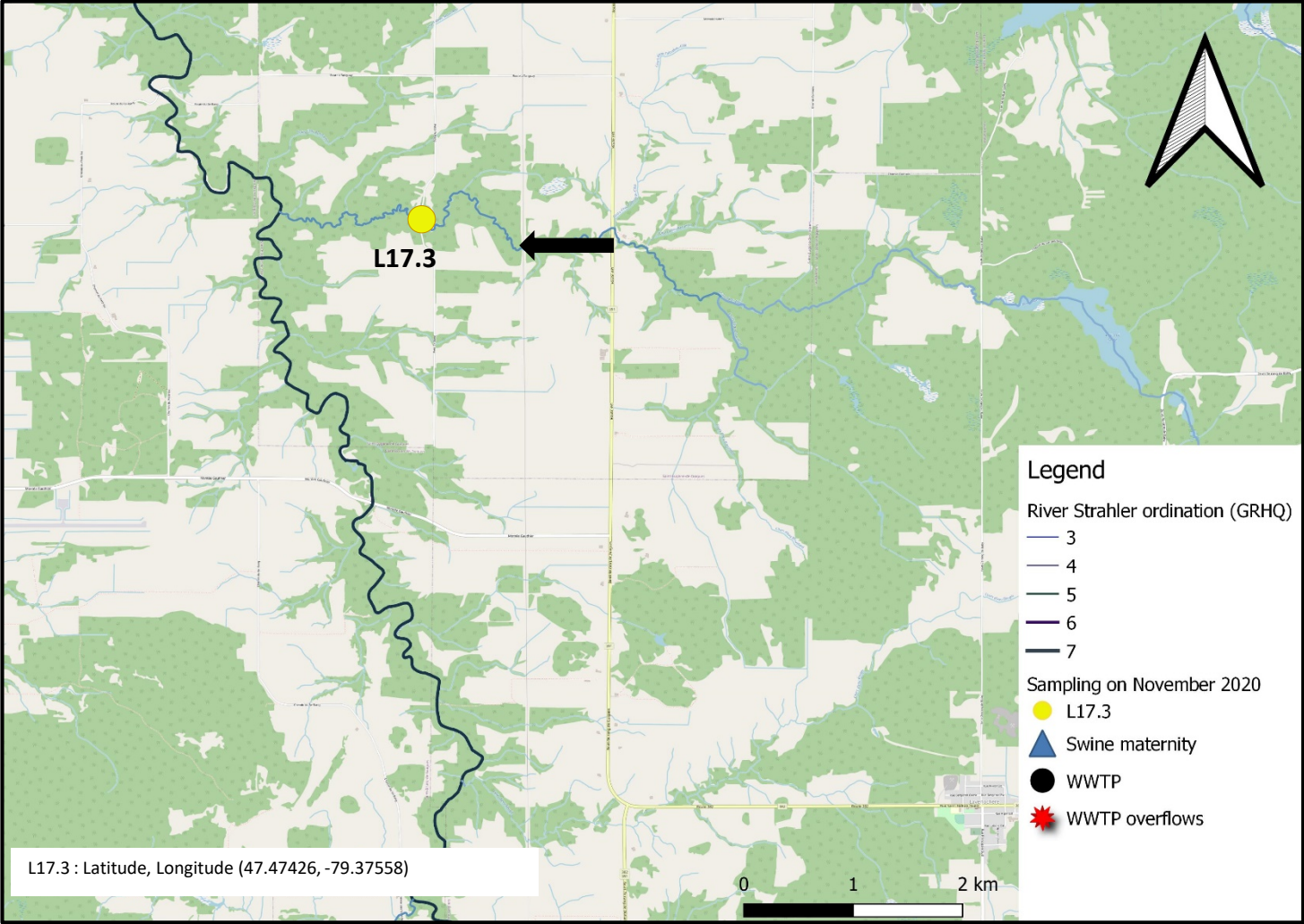
Inflow of a wastewater treatment plant
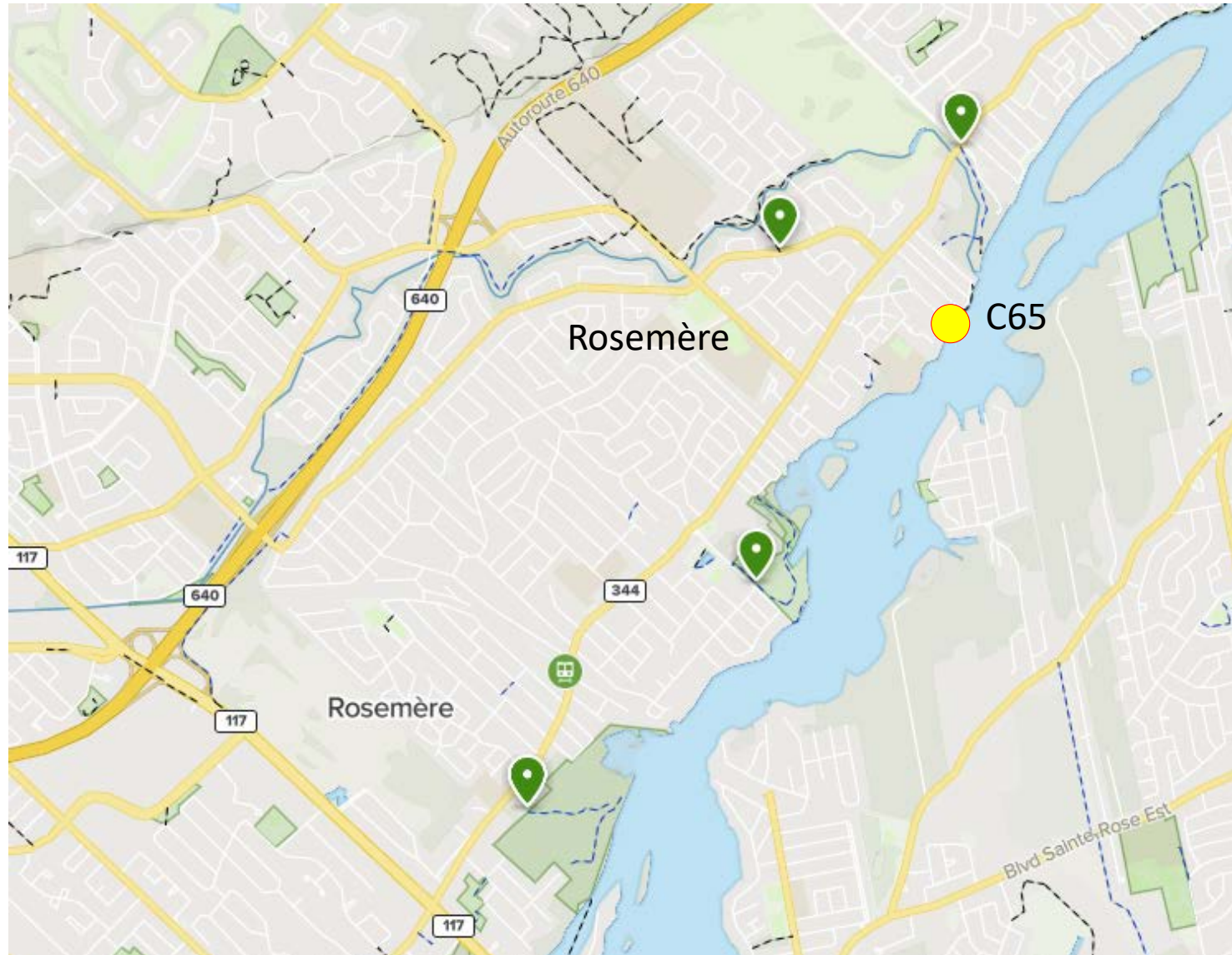
Concentrations of FIB and mtDNA source tracking markers

Samples		Coliform CFU/100 mL	Enterococci CFU/100 mL	qPCR	
				Human mtDNA copies/100 mL	Bovine mtDNA copies/100 mL
	River				
C43	L'Assomption	>30 000	900	2350	230
C94	Bayonne	>30 000	400	900	<LOD
	Streams				
L16	Dumais	151	ND	<LOD	<LOD
L17.3	Bastien	3	ND	<LOD	<LOD
Raw wastewater (inflow of a wastewater treatment plant)					
C65	Rosemère	>30 000	19250	2 440 000	4050

Fecal contamination — River (C43, C94)

Clean — Streams (L16, L17.3)

<LOD : below the limit of detection

ND: not done

L'Assomption and Bayonne Rivers

*Rhinichthys cataractae** longnose dace   17.8%
Micropterus dolomieu Smallmouth bass   14.5%
Etheostoma olmstedi Tessellated darter   11.4%
*Catostomus commersonii** white sucker   9.6%
Etheostoma flabellare Fantail darter   8.1%
*Ictalurus punctatus** Channel catfish   6.0%
*Percopsis omiscomaycu*s Trout-perch   3.4%
Ambloplites rupestris Rock bass   2.9%
Perca flavescens American yellow perch   2.7%
Esox Lucius Northern pike   2.4%
*Semotilus atromaculatus** Creek chub   2.0%
*Pimephales notatus** Bluntnose minnow   1.7%
Sander vitreus Walleye   1.1%
*Cyprinella spiloptera** Spotfin shiner   0.92%
*Aplodinotus grunnie*ns Freshwater drum   0.83%
Hiodon tergisus Mooneye   *0.65%*
*Ameiurus nebulosus** Brown bullhead   0.62%
*Cyprinus carpio** Common carp   0.56%
*Pimephales promelas** Fathead minnow   0.39%
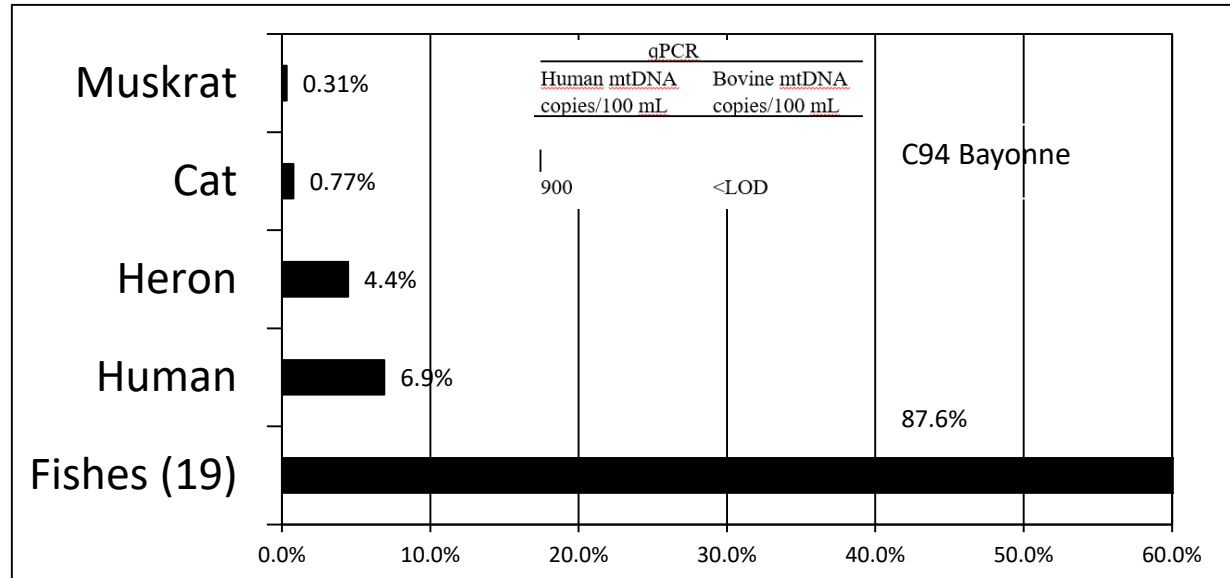
*Micropterus dolomieu S*mallmouth bass   23.2%
*Moxostoma anisurum** silver redhorse   6.3%
Etheostoma nigrum Johnny darter   3.4%
Percina caprodes Logperch   2.7%
*Cyprinus carpio** Common carp   2.5%
*Notropis volucellus** Mimic shiner   2.1%
*Rhinichthys cataractae** longnose dace   1.8%
Sander vitreus Walleye   1.5%
Esox masquinongy Muskellunge   1.4%
*Catostomus commersonii** white sucker   1.10%
*Cyprinella spiloptera** Spotfin shiner   1.04%
*Hybognathus regius** silvery minnow   1.03%
*Noturus flavus** Stonecat   0.99%
***Poecilia reticulata* Guppy   0.63%**
*Ictalurus punctatus** Channel catfish   0.40%
Etheostoma olmstedi Tessellated darter   0.17%
*Semotilus corporalis** Fallfish   0.084%
Perca fluviatilis Perch   0.077%
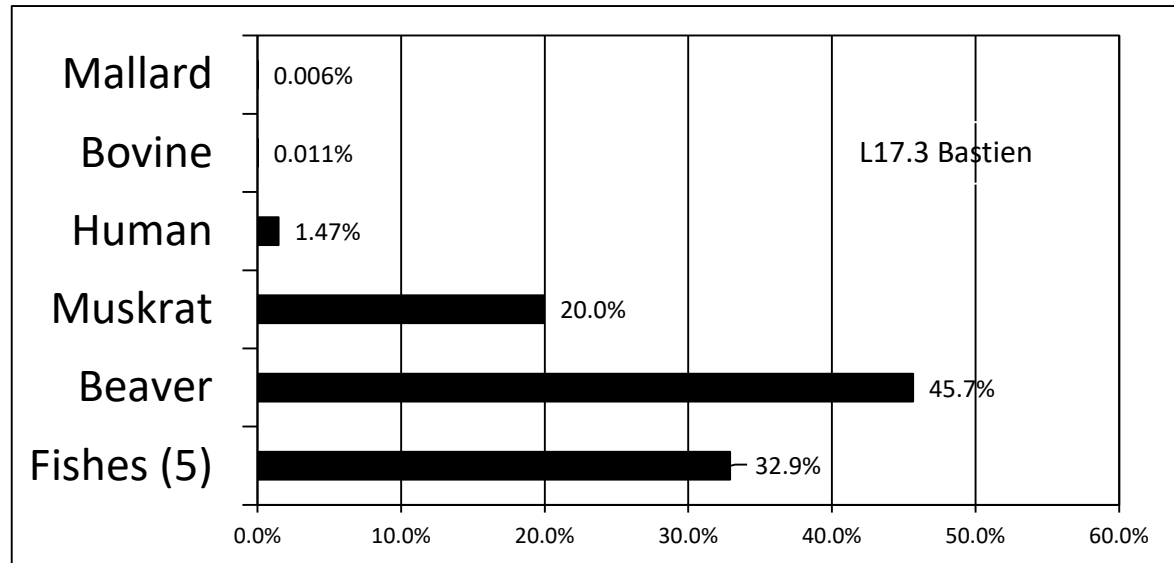*Luxilus chrysocephalus** striped shiner   0.010%
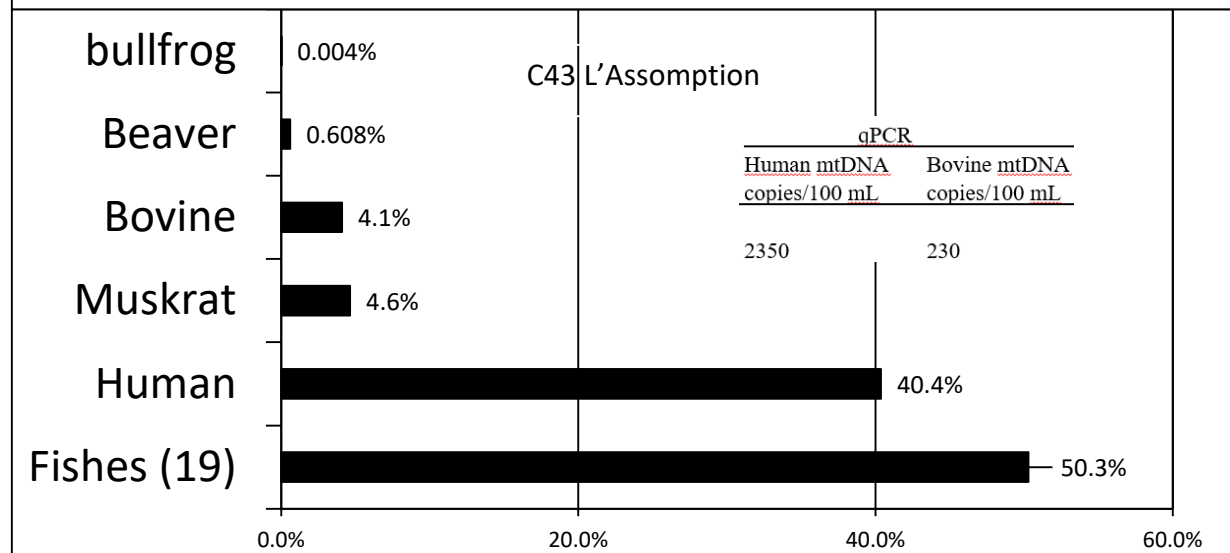
Warning!!!!

- Values are given in relative level (%).
- Tell nothing about the absolute concentration
- qPCR is required

Bastien and Dumais streams

Semotilus atromaculatus* creek chub 18.6%
Rhinichthys cataractae* longnose dace 12.1%
Etheostoma nigrum Johnny darter 1.4%
Luxilus chrysocephalus* striped shiner 0.58%
Catostomus commersonii* white sucker 0.21%

Culaea inconstans brook stickleback 27.9%
Semotilus atromaculatus* creek chub 6.3%
Catostomus commersonii* white sucker 4.6%
Luxilus chrysocephalus* striped shiner 0.98%
Pimephales notatus* bluntnose minnow 0.24%
Cottus cognatus slimy sculpin 0.20%
Lithobates clamitans bronze frog 0.14%
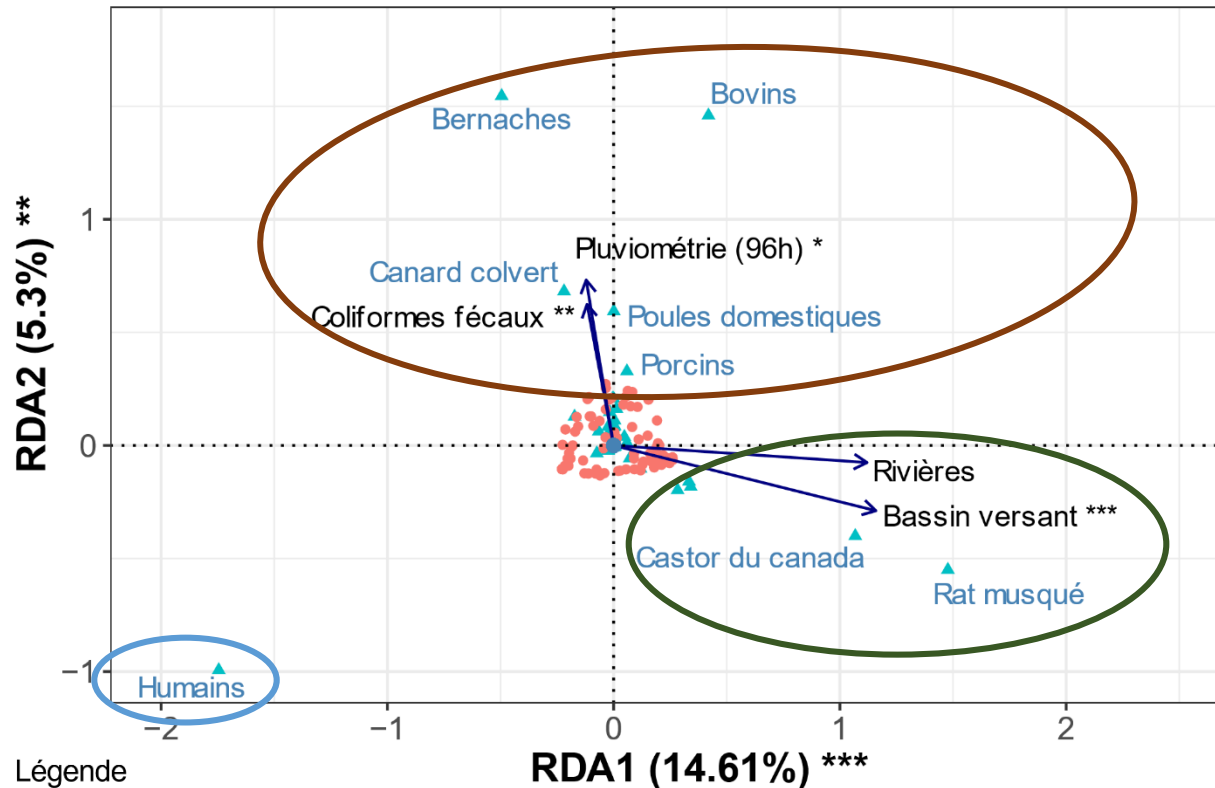Moxostoma anisurum* silver redhorse 0.009%
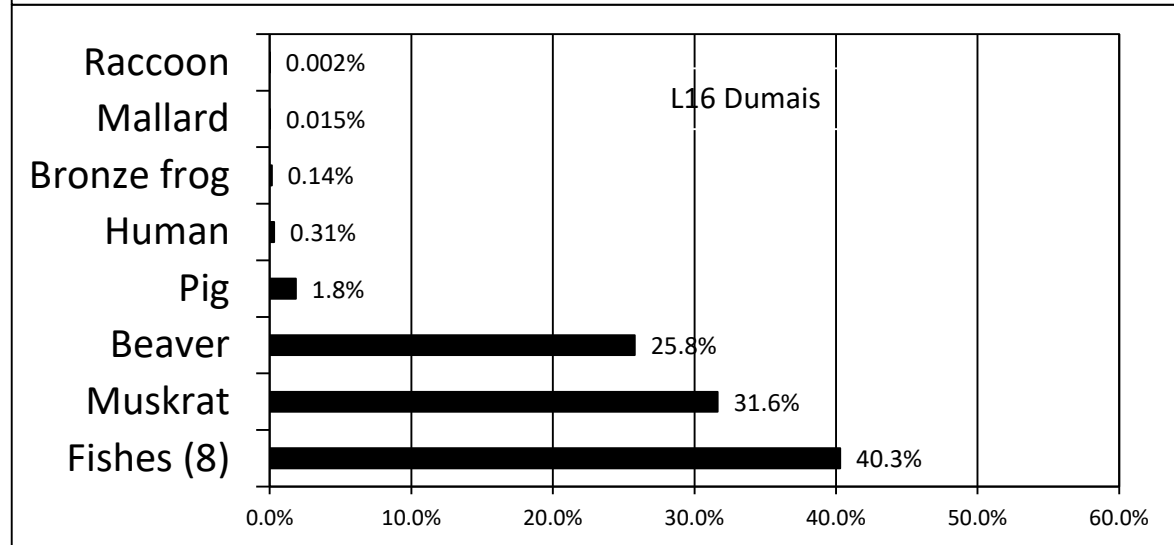
Wastewater
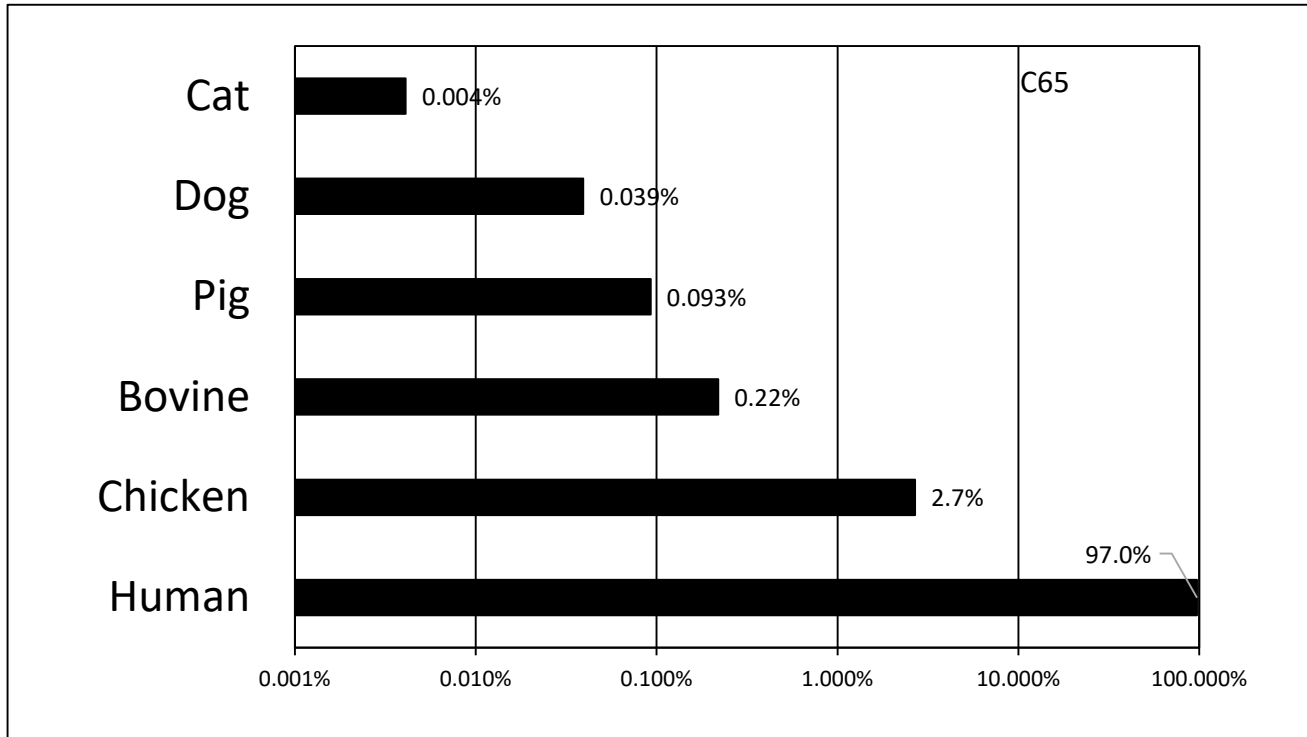
- Other animals: domestic animals (cat and dog), meat waste and undigested meat.
- May have an incidence on the animal profile in contaminated water

Covariation: species vs environmental factors

- Redundancy analyses (RDA) 86 samples

- Coliforms, rainfall and watershed are variables that significantly explain the occurrence and abundance of the different animals.

- 3 bird species, cattle and pigs strongly covary with coliforms and precipitation.

- Muskrat and beaver are geographically dependent (high proportion in Témiscamingue) and do not covary with coliforms.

- Human occurred in 88% of samples and clustered apart from the other animals.

- This suggests that other factors influenced the occurrence of human is water (e.g. presence of beaches, treated water outlets, type of treatment, number of inhabitants, etc.)

Conclusions

- We designed new PCR primers to amplify mtDNA from mammals, birds and fish (in some extent amphibian) from environmental DNA

- Sequencing the amplicon by Illumina and clustering analysis, this provides an powerfull tool to derive the profiles of these animals in watershed (or whatever the environment).

- This allows to assess the potential source of fecal contamination, which could occur from different animals.

- Such identification can allow to develop better strategies by the watershed management authorities in mitigating the contamination at their sources.

Conclusions

- Regarding fish profile, this can provide indication of its environmental health or anthropic pressures.

- Knowing the proportion of mtDNA from specific fish species can be useful for the temporal follow up of these species in a given river.

- Our approach has the potential to survey rapidly and repeatedly the composition of fishes in the rivers or lakes, and also survey the terrestrial animals surrounding these waters.

- This information can be valuable in lake and river management for the evolution of invasive species for instance, or for recreational purpose (e.g. fishing activities).

Remerciements

- Rose Ragot, doctorante
- Florence Lessard, Fondation Rivières
- André Bélanger, Fondation Rivières
- Yves Grafteaux, OBVT
- Lawrence Gervais, OBVT
- CRSNG et la Fondation Rivières pour le financement