

UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

BIOSURVEILLANCE DES PETITS COURS D'EAU : COMPARAISON ENTRE ADN
ENVIRONNEMENTAL ET LES MÉTHODES TRADITIONNELLES BASÉES SUR LES
DIATOMÉES ET LES INVERTÉBRÉS.

MÉMOIRE PRÉSENTÉ
COMME EXIGENCE PARTIELLE DE LA
MAÎTRISE EN SCIENCES DE L'ENVIRONNEMENT

PAR
MÉLISSA GINGRAS-DUBREUIL

NOVEMBRE 2025

Université du Québec à Trois-Rivières

Service de la bibliothèque

Avertissement

L'auteur de ce mémoire, de cette thèse ou de cet essai a autorisé l'Université du Québec à Trois-Rivières à diffuser, à des fins non lucratives, une copie de son mémoire, de sa thèse ou de son essai.

Cette diffusion n'entraîne pas une renonciation de la part de l'auteur à ses droits de propriété intellectuelle, incluant le droit d'auteur, sur ce mémoire, cette thèse ou cet essai. Notamment, la reproduction ou la publication de la totalité ou d'une partie importante de ce mémoire, de cette thèse et de son essai requiert son autorisation.

Remerciements

Ce mémoire n'aurait pu être complété sans l'assistance de nombreuses personnes. Je souhaite ainsi remercier :

- Louis Astorg pour son aide inestimable avec la bio-informatique.
- Mes superviseurs, Vani Mohit et Vincent Fugère, les membres de mon comité d'orientation, Isabelle Lavoie et Valérie Langlois, ainsi que Jérôme Comte pour leurs précieux conseils et pour la relecture de ce mémoire.
- Les membres des équipes terrain : Pascal Côté, Saylena Fay, Audrey Robert, Claudia Nault, Marc-Olivier Maalsnaes et Kristof Mercier. Merci pour votre efficacité et votre capacité à travailler de manière indépendante!
- Le CEAEQ qui m'a accueillie et formée pour les manipulations d'extraction de l'ADN.
- Ma sœur, Aurée Gingras-Dubreuil, pour son assistance avec les nombreuses relectures, les corrections et le support moral.
- Mes chats, Cachou, Quenouille, Astrid et feu Princesse, pour leur réconfort et les pauses forcées.

Liste des figures

Figure 1. Répartition géographique des sites d'échantillonnage et de leur niveau de perturbation agricole.....	14
Figure 2. Analyse en composantes principales montrant la présence d'un gradient de perturbation agricole à travers les sites échantillonnés.	26
Figure 3. Comparaison de deux indices de diversité obtenus par les deux méthodes séparées selon le taxon étudié.....	27
Figure 4. Compositions des communautés par ADNe et par les méthodes traditionnelles représentées par un positionnement multidimensionnel non-métrique (NMDS)	28
Figure 5. Réponse de la composition de la communauté (score sur une NMDS1) au gradient de perturbation agricole selon la méthode.....	29

Liste des tableaux

Tableau 1 Détails méthodologiques des PCRs.....	19
Tableau 2 Paramètres d'élagage du pipeline DADA2	20
Tableau 3 Résumé statistique du pipeline DADA2	21

Résumé

La biodiversité d'eau douce est fortement menacée par la diminution de la qualité de l'eau de surface causée par l'intensification de l'agriculture. Pour mieux comprendre les effets de cette pression sur les communautés aquatiques, il est nécessaire d'évaluer la diversité sur plusieurs niveaux trophiques plutôt que de se concentrer sur un seul groupe d'organismes. Les méthodes traditionnelles d'échantillonnage et d'identification exigent des taxonomistes hautement qualifiés et sont chronophages, ce qui rend l'approche multitrophique très coûteuse.

Ce mémoire a comme objectif de tester l'usage de l'ADN environnemental, extrait d'un seul échantillon d'eau, combiné à une approche multimarqueur (*rbcL* et *COI*) comme alternative à l'échantillonnage traditionnel de deux groupes d'organismes couramment utilisés en biosurveillance : les diatomées et les invertébrés. La comparaison entre les communautés décelées par chaque méthode indique qu'elles diffèrent grandement dans leur composition et leur diversité, mais les approches visant l'ADNe sont tout de même efficaces pour détecter le signal d'un gradient de perturbation agricole.

Table des matières

Remerciements.....	2
Liste des figures	3
Liste des tableaux.....	3
Résumé.....	4
1. Mise en contexte	6
1.1 Effets de l'intensification agricole sur les cours d'eau	6
1.2 Survol du suivi abiotique et biotique de la pollution agricole	8
1.3 L'ADN environnemental comme outils en étude des communautés.....	10
1.4 Objectifs de l'étude.....	12
2. Méthodes.....	13
2.1 Sites à l'étude.....	13
2.2 Plan d'échantillonnage.....	15
2.3 Protocoles terrain	15
2.3.1 Protocole ADNe.....	15
2.3.2 Protocole d'échantillonnage traditionnel	16
2.3.3 Paramètres physicochimiques.....	17
2.4 Usages du territoire	18
2.5 Protocoles de laboratoire.....	18
2.6 Pipelines bio-informatiques	20
2.7 Traitement des données.....	22
2.8 Optimisation des bases de données taxonomiques traditionnelles	22
2.9 Épuration des bases de données taxonomiques d'ADNe.....	24
2.10 Analyses statistiques	25
3. Résultats.....	26
4. Discussion.....	30
4.1 Correspondance entre les méthodes.....	30
4.2 Effet de l'usage du territoire	33
5. Conclusion	35
7. Références.....	36
ANNEXE A	49

1. Mise en contexte

1.1 Effets de l'intensification agricole sur les cours d'eau

L'intensification agricole est une menace majeure à la qualité des eaux intérieures au Québec et ailleurs dans le monde (Giacomazzo et al., 2020; Vörösmarty et al., 2010). Depuis la révolution verte des années 1960s, les exploitations agricoles utilisent de plus en plus d'engrais, de produits agrochimiques, d'instruments mécaniques et de variétés de plantes plus productives; des pratiques aux champs qui visent à accroître les rendements sans pour autant augmenter la superficie cultivée. Au Québec, la superficie de terre dédiée à la culture de maïs/soja (grand interligne) était de 825 700 ha en 2020, ce qui en fait le type de culture le plus important de la province (Giroux, 2022). Ces cultures intensives, souvent composées d'organismes génétiquement modifiés (OGM) résistants à certains herbicides, favorisent un usage plus fréquent de ces produits et ont remplacé d'autres cultures moins intensives, comme les céréales ou le foin (Giroux, 2022). Cette transition vers un type de culture plus polluant est une des dimensions importantes de l'intensification agricole.

Ce changement dans l'usage des terres amène de nombreuses perturbations dans les écosystèmes aquatiques qui reçoivent l'eau des bassins versants agricoles. Les cultures à grande interligne sont très exigeantes en intrants chimiques : engrais pour enrichir le sol, insecticides pour protéger les cultures, herbicides pour éliminer les adventices, etc. (Allan, 2004). Ces produits agrochimiques peuvent être emportés par l'eau de pluie et le vent de la terre vers les milieux aquatiques et déséquilibrer l'écosystème. L'enrichissement en nutriments d'un cours d'eau augmente la productivité de l'écosystème et la biomasse des organismes autotrophes, ce qui modifie l'assemblage de la communauté (Budnick et al., 2019). Cette biomasse supplémentaire meurt et se décompose, ce qui réduit l'oxygène dissous et induit un autre changement dans la communauté en faveur des espèces plus tolérantes à l'hypoxie (Carpenter et al., 1998; Niyogi et al., 2003). Ce phénomène est plutôt associé aux écosystèmes lacustres, mais va aussi affecter les cours d'eau (Blaszczak et al., 2023; Garvey et al., 2007; Mallin et al., 2006). Ces changements à la base de la chaîne alimentaire entraînent un contrôle ascendant dans le réseau trophique, ce qui modifie la biodiversité aquatique de manière générale. Les polluants tels que les pesticides peuvent

avoir des effets plus directs sur les espèces via une toxicité létale ou sublétale, ce qui altère considérablement le réseau trophique (Blann et al., 2009).

Contrairement aux cultures pérennes, comme les plantes fourragères, les cultures à grand interligne laissent le sol à nu à la fin de l'été. Il ne reste donc plus de végétation pour retenir la terre qui est érodée dans les ruisseaux lors des pluies automnales et de la fonte des neiges au printemps, augmentant la turbidité des cours d'eau (Kort et al., 1998). Les particules de sol en suspension dans l'eau bloquent les rayons du soleil, ce qui diminue la production primaire et la qualité de nourriture disponible (Henley et al., 2000). De plus, l'abrasion par les sédiments en mouvement et leur déposition sur le substrat rendent le milieu moins propice pour le périphyton. Il y a également un remplissage des interstices et des crevasses, des habitats essentiels pour les invertébrés, et le recouvrement des substrats pour la fraie des espèces de poissons qui pondent dans le gravier. La déposition des sédiments homogénéise la profondeur du ruisseau, entraînant une diminution de la biodiversité (Walser & Bart Jr., 1999). En somme, les apports en éléments nutritifs, en contaminants agrochimiques et en sédiments en lien avec l'agriculture intensive influencent fortement les composantes abiotiques et biotiques des cours d'eau.

L'intensification agricole apporte également des modifications dans l'hydrologie du bassin versant. La compaction du sol par la machinerie et l'ajout de drainage accélèrent l'écoulement de l'eau de pluie vers les ruisseaux, entraînant des crues importantes et rapides. Cela favorise le processus d'érosion, le transport des sédiments, des nutriments et des contaminants, contribuant à la dégradation du cours d'eau (Blann et al., 2009). Le changement dans le type de végétation réduit généralement la rétention de l'eau du bassin versant par l'évapotranspiration, réduisant le niveau de l'étiage et allongeant sa durée. Les espèces qui dépendent des cours d'eau ont évolué en réponse à des régimes hydriques particuliers et aux conditions environnementales qui en découlent (Lytle & Poff, 2004). Ainsi, certaines espèces sont adaptées à des débits stables alors que d'autres peuvent tolérer des fluctuations extrêmes.

Les racines de la végétation en bordure des cours d'eau contribuent à maintenir la structure de la berge, à retenir les sédiments et les contaminants. Ainsi, le retrait de la bande riveraine souvent associé à l'agriculture diminue la stabilité des berges, ce qui les rend plus vulnérables à l'érosion, et augmente la quantité de sédiments et de contaminants qui sont lessivés ou lixiviés vers le cours d'eau (Allan, 2004). Dus à la réduction de la canopée, plus de rayonnements solaires

atteignent le ruisseau et y pénètrent plus profondément, ce qui augmente la température et favorise la croissance des algues et des plantes aquatiques. La bande riveraine est également une source importante de matières organiques et de débris ligneux. En l'éliminant, la quantité et la qualité de carbone organique particulaire et dissous se rendant à l'écosystème sont altérées (Findlay et al., 2001). La perte des gros débris ligneux dans les ruisseaux amène une diminution d'habitats pour plusieurs organismes qui colonisent ou utilisent ces substrats (Gurnell et al., 1995).

1.2 Survol du suivi abiotique et biotique de la pollution agricole

Les effets de l'utilisation du territoire et de l'intensification agricole sur les eaux douces sont souvent évalués à partir de mesures abiotiques, telles que des mesures chimiques des teneurs en éléments nutritifs et en contaminants, ou d'indices biotiques de qualité du milieu. En plus d'être coûteuses, les analyses chimiques de la qualité de l'eau sont limitées dans leur couverture temporelle. Les niveaux des contaminants et des autres paramètres chimiques sont très variables dans le temps, avec des sommets à différents moments selon des événements, tels que les grosses pluies, la fonte printanière ou l'épandage dans les champs (Giroux, 2022). Cette variabilité temporelle est bien illustrée dans Lavoie et al. (2008), où des pointes de phosphore total ont été enregistrées dans trois jeux de données. Ce qui est détecté dans un échantillon d'eau est représentatif uniquement du moment de la prise, ces sommets peuvent donc passer sous le radar et la qualité de l'eau d'un cours d'eau peut être surestimée. De plus, mesurer chaque contaminant individuellement, en plus d'être très coûteux, ne permet pas de prendre en compte les effets cumulatifs et les possibles synergies entre les différents contaminants dans l'environnement (Relyea, 2009). Les seuils de toxicité auxquels les teneurs chimiques sont comparées sont généralement basés sur des bioessais en laboratoire, ce qui ne représente pas toujours bien les conséquences en milieu naturel où les interactions entre les organismes et leur environnement peuvent moduler les effets toxiques des contaminants.

Pour pallier ces limites, des outils de biosurveillance basés sur la composition en espèces ont été développés (Karr, 1991). Les communautés se bâtissent à travers le temps et intègrent tous les maximums et les effets synergiques inconnus que les analyses chimiques et les bioessais peuvent manquer. Ainsi, un seul inventaire en fin de saison peut donner une bonne évaluation de la qualité de l'eau et de l'habitat sans avoir à réaliser plusieurs analyses chimiques au cours de la saison.

Les indices biotiques utilisés en biosurveillance se concentrent habituellement sur un seul niveau trophique ou taxon. Ainsi, le *Canadian Aquatic Biomonitoring Network* (CABIN) et le Réseau-Benthos du Québec utilisent uniquement les macroinvertébrés benthiques pour évaluer l'intégrité biotique des cours d'eau québécois ou canadiens (Environment and Climate Change Canada, s. d.; MDDEP, 2012). En plus des invertébrés, la composition des assemblages de diatomées est fréquemment utilisée pour évaluer la qualité de l'eau et des habitats des rivières, par exemple l'Indice Diatomées de l'Est du Canada (IDEC) utilisé au Québec. Les communautés de poissons sont utilisées aux États-Unis pour l'*Index of Biotic Integrity* (IBI; Karr, 1991).

Il est rare que plus d'un indice soit utilisé pour évaluer une même rivière. L'usage d'une approche considérant la diversité à plusieurs niveaux de la chaîne trophique (diversité multitrophique ou multitaxa) amène une évaluation plus complète de l'état d'une rivière, puisque chaque groupe d'organismes reflète un aspect différent de l'intégrité écologique (Seibold et al., 2018). Par exemple, les diatomées et les invertébrés comme les larves d'insectes sont des indicateurs de la qualité de l'eau, tandis que les bactéries et les mycètes sont reliés plus étroitement aux processus biogéochimiques (Cordier et al., 2019). D'autres groupes, comme les poissons ou encore certains mollusques, incluent quant à eux la plupart des espèces menacées et représentent donc des taxa importants dans une perspective de conservation de la biodiversité. Les poissons et les invertébrés sont également de bons indicateurs de qualité de l'habitat. Par ailleurs, l'ajout, la disparition ou encore un changement dans l'abondance, la physiologie ou le comportement d'une de ces espèces dans un écosystème peut être lié à une perturbation écologique (Whitfield & Elliott, 2002). En plus de servir au développement d'indices biotiques plus exhaustifs qui considèrent plusieurs facettes de la biodiversité d'un site, la quantification de la diversité multitrophique représente un sujet de pointe en écologie théorique et en écologie des communautés, car elle permettrait de mieux prédire l'impact d'une perte de diversité sur le fonctionnement des écosystèmes (A. D. Barnes et al., 2017; Gossner et al., 2016; Sirami et al., 2019; Soliveres et al., 2016).

Les approches de biosurveillance traditionnelle présentent certaines limites qui nous empêchent d'effectuer des analyses de biodiversité multitrophique. Elles sont très coûteuses en ressources humaines, et elles nécessitent une grande expertise taxonomique pour chaque taxon pris en compte dans l'analyse. Il est donc très rare qu'un programme de suivi puisse se permettre

d'estimer systématiquement la diversité et la composition taxonomique de plusieurs groupes d'organismes provenant des mêmes sites d'échantillonnage. Au Québec, le suivi biotique est coordonné par le ministère de l'Environnement, de la Lutte contre les changements climatiques, de la Faune et des Parcs (MELCCFP). Les diatomées et les macroinvertébrés sont généralement utilisés pour la surveillance des cours d'eau de taille moyenne. En contrepartie, les pêches scientifiques concentrent la majeure partie de leurs activités de suivi sur les lacs, sauf certains suivis en rivières tels que la Yamaska et ses tributaires (Richard, 2014), la Bourlamaque (Richard, 2006) et le ruisseau Saint-George (Richard & Giroux, 2004). À l'exception des cyanobactéries qui représentent une menace directe à la santé humaine, aucune instance gouvernementale n'échantillonne de façon systématique les communautés de microorganismes, tels les bactéries, les mycètes et le phytoplancton, pour en faire le suivi. Ainsi, la diversité multitrophique et les assemblages de communautés des rivières et sa variation en fonction de l'usage des terres demeurent relativement méconnus.

1.3 L'ADN environnemental comme outils en étude des communautés

L'usage de l'ADN environnemental (ADNe) aurait l'avantage de réduire le nombre d'inventaires nécessaires à un échantillon d'eau par site, facilitant les mesures de diversité multitrophique. L'ADNe est le matériel génétique extrait d'un échantillon environnemental, par exemple du sol, des sédiments, de l'eau ou de l'air, sans isolation d'organismes au préalable (Taberlet et al., 2012). Les organismes vivants laissent des traces génétiques dans l'environnement par les fèces, le mucus et différentes cellules qui peuvent se détacher de l'organisme (Creer et al., 2016). L'ADNe est composé d'ADN extracellulaire et intracellulaire de tissus morts. L'ADN extracellulaire est l'ADN qui est échappé des cellules. L'ADNe peut aussi se trouver sous forme intracellulaire, c'est-à-dire provenant d'organismes toujours vivants, comme des bactéries, des algues ou la méiofaune invertébrée, ou bien de cellules mortes intactes. Ainsi, l'eau d'un cours d'eau contient des traces génétiques de tous les êtres vivants qui y sont présents. Séquencer l'ADNe d'un échantillon d'eau permet de déterminer quelles espèces ont laissé des traces génétiques dans le milieu, et ainsi caractériser les communautés biotiques de cet environnement sans avoir besoin de capturer et d'identifier visuellement les organismes qui forment ces communautés (Deiner et al., 2016). De plus, l'évolution des technologies de séquençage, c'est-à-dire la lecture de l'ADN, amène une diminution des coûts d'analyse de l'ADNe (Shokralla et al.,

2012), ce qui rend réaliste l'usage de plusieurs marqueurs génétiques pour détecter différents taxa (Adamowicz et al., 2019; Drummond et al., 2015; Stat et al., 2017; West et al., 2020). C'est ce qu'on appelle une approche par ADNe multimarqueur, où un même échantillon est séquencé à plusieurs reprises, chaque fois avec une amorce spécifique à un type d'organismes différent. Cela revient un peu à faire plusieurs échantillonnages dans une bouteille d'eau plutôt que dans le ruisseau. Il serait donc possible à l'aide de l'ADNe de mesurer une plus grande proportion de la biodiversité totale d'un site à un coût moindre qu'avec des méthodes d'échantillonnage dites "traditionnelles", telles que la pêche expérimentale et le dénombrement au microscope.

Les approches par ADNe présentent plusieurs limitations. Dans les écosystèmes d'eau douce, la détectabilité de l'ADNe dépend de sa production par les organismes, de sa dégradation et de son transport. Selon leur trait d'histoire de vie, chaque organisme relâche des quantités différentes d'ADN dans l'environnement (Stewart, 2019). De même, il existe des variations intraspécifiques dans la production de matériel génétique pouvant être associée à des facteurs physiologiques, tels que le stress et le taux métabolique. De multiples études montrent que les fragments d'ADN propre à l'espèce persisteraient entre quelques jours à un mois dans l'eau (Rees et al., 2014). Plusieurs facteurs biotiques et abiotiques influencent son taux de dégradation. L'augmentation de la température, des rayons UV-B et de l'alcalinité (Strickler et al., 2015), ainsi que de l'activité microbienne (Barnes et al., 2014), accélèrent la dégradation de l'ADN, alors que les particules en suspension vont contribuer à le préserver (Turner et al., 2015). La distance de transport en milieu lotique n'est pas claire, puisque les études présentent des résultats très différents, allant de 5 m à 12,3 km selon l'hydrologie et la morphologie du cours d'eau (Taberlet et al., 2018). La distance parcourue par l'ADN dépend de sa concentration initiale, de la vitesse du courant et de la sensibilité de détection de la méthode utilisée, ce qui explique la variabilité entre les études. En effet, la quantité d'eau filtrée, la taille des pores du filtre, le type de membrane du filtre et l'appareil de filtration peuvent jouer un rôle dans la détection ou non de l'ADNe d'un organisme. En raison de ces complications, il y a un besoin important d'études qui valident "l'outil ADNe" dans différents milieux en comparant la composition des communautés estimée à partir de l'ADNe et celle obtenue à partir de méthodes traditionnelles. De telles études de validation sont particulièrement rares pour la biodiversité multitrophique, où plusieurs groupes d'organismes sont caractérisés à l'aide des deux approches mentionnées plus haut.

Les bases de données de références qui servent à assigner la taxonomie aux séquences d'ADN présentent des lacunes, particulièrement pour les diatomées et les invertébrés en eau douce (Weigand et al., 2019). Ainsi, une grande partie des séquences d'un jeu de données ne peuvent pas être associées à une résolution taxonomique suffisante pour être interprétable, rendant la communauté détectée incomplète. Il est donc possible de manquer un pan complet de la communauté, tel un groupe d'espèces mal représenté dans les bases de données de références, et d'obtenir une interprétation des résultats erronée. Dans un contexte de biosurveillance, une méthode suggérée pour contrer cela est d'utiliser une approche sans taxonomie, qui consiste à analyser directement les séquences d'ADNe sans les associer à des espèces (Apothéloz-Perret-Gentil et al., 2017). Cela permet de détecter des changements dans les structures des communautés, des perturbations écologiques ou de comparer des sites selon leur degré de dégradation, sans nécessiter une identification taxonomique complète (Apothéloz-Perret-Gentil et al., 2017; Tapolczai et al., 2019). Cependant, cette approche ne permet pas une analyse des préférences ou des fonctions écologiques des séquences détectées.

1.4 Objectifs de l'étude

Notre projet cherche ainsi à explorer l'usage de la méthode par ADNe multimarqueur dans les petits cours d'eau afin d'ouvrir la voie à des études concernant les impacts de l'intensification agricole sur la biodiversité aquatique multitrophique. En premier lieu, nous avons analysé les données pour déterminer s'il existe une correspondance entre les communautés de diatomées et d'invertébrés échantillonnées de manière traditionnelle avec celles déterminées par l'ADNe tiré d'un même échantillon d'eau. En second lieu, nous avons vérifié si l'ADNe reflète l'impact de l'usage des terres sur la diversité et la composition des communautés précédemment mentionnées, deux bio-indicateurs reconnus des cours d'eau. Ainsi, répondre à ces objectifs nous permettra de savoir si le contenu génétique d'un échantillon d'eau permet de renseigner sur l'ensemble des stress anthropiques agricoles grâce aux variations dans les communautés aquatiques des ruisseaux.

Les bases de données regroupant les mesures de diversité multitrophique et des paramètres abiotiques sont rares. En ajoutant une dimension ADNe au projet, nous avons une opportunité unique de tester une nouvelle approche de biosurveillance à l'aide d'un outil moléculaire à l'échelle du paysage. Si nos tests confirment que l'ADNe permet de détecter les impacts de l'usage

des terres sur la biodiversité multitrophique, le Québec pourrait développer un nouvel indice d'intégrité écologique multitrophique basé sur l'ADNe. Cet indice serait plus complet et plus accessible en termes de coût et de temps pour les organismes de protection de l'eau que l'échantillonnage exhaustif d'un réseau trophique par les méthodes traditionnelles.

2. Méthodes

2.1 Sites à l'étude

Pour répondre aux objectifs de ce projet de maîtrise, des sites ont été sélectionnés dans le bassin versant du Lac Saint-Pierre (LSP). C'est un lieu de forte biodiversité au Québec, dont la région agricole est passée de cultures majoritairement fourragères à des cultures intensives de maïs/soja entre 1950 et 1997 (Dauphin & Jobin, 2016). Ainsi, les cours d'eau qui l'alimentent ont des bassins versants qui varient dans leur degré de naturalité : forêts et milieux humides naturels, cultures fourragères peu polluantes et monocultures intensives plus polluantes (Simoneau, 2017). La grande diversité d'usage du territoire dans son bassin versant permet d'établir un gradient d'intensité agricole. L'importance économique, récréative et écologique du LSP fait en sorte que le territoire est bien connu et de nombreux suivis sont déjà effectués depuis plusieurs années. Cette connaissance préalable du territoire a facilité la sélection des sites d'échantillonnages pour établir le gradient d'intensification agricole.

En 2021, douze ruisseaux, répartis sur les deux rives du LSP, ont été choisis selon leur niveau d'intensité agricole afin d'obtenir un gradient du plus naturel possible au plus dégradé. Plusieurs critères ont été établis afin de sélectionner les ruisseaux .

1. Une partie des ruisseaux devaient se déverser directement dans le littoral du LSP pour que les informations récoltées puissent contribuer à la préservation de l'habitat du poisson et à la protection de la biodiversité.
2. Les bassins versants devaient tous être approximativement de la même taille, puisqu'il existe une relation forte entre celle-ci et la biodiversité d'un cours d'eau (Vander Vorste et al., 2017).
3. Le ruisseau devait être assez petit pour permettre l'échantillonnage de tous les organismes ciblés dans l'étude, c'est-à-dire pouvoir y circuler à gué, mais suffisamment gros pour ne pas

être intermittents. Un ruisseau qui s’assècherait verrait sa diversité changer par rapport aux autres (Soria et al., 2017), donc ce paramètre devait demeurer le plus constant possible.

4. La composition de l’usage des terres devait être considérée pour établir un gradient de naturalité. Il fallait des bassins versants dont le territoire était majoritairement composé de milieux naturels ou de milieux agricoles et d’autres qui possédaient les deux types de milieux en différentes proportions.
5. Les ruisseaux sélectionnés devaient être d’intérêt particulier pour nos partenaires, soit quatre organismes de bassins versants de la région, de même que le bureau “Environnement et Terre” de la communauté abénaquise Odanak.

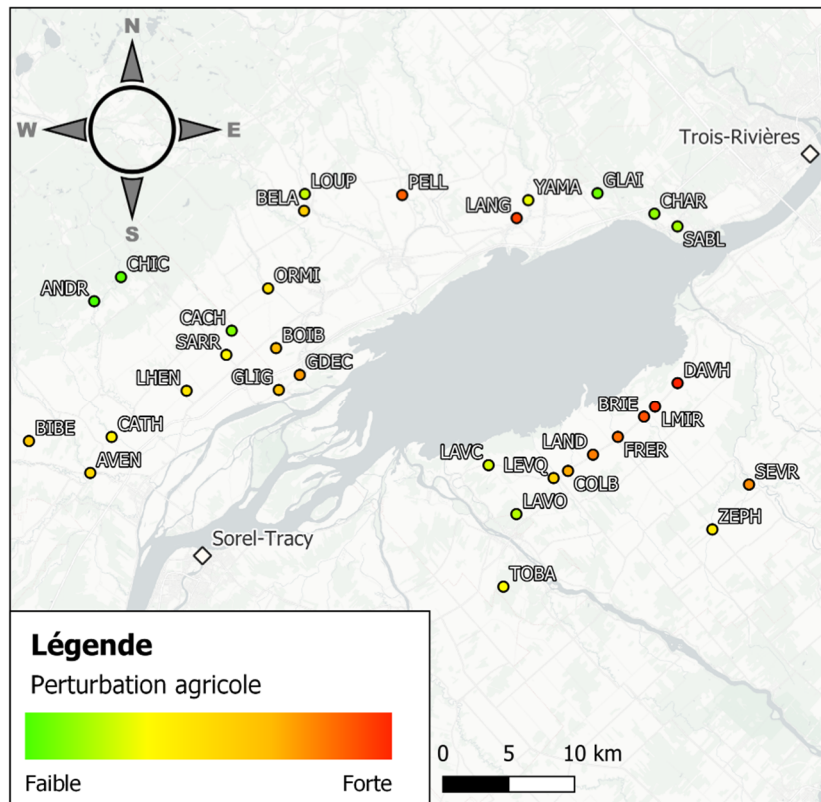


Figure 1. Répartition géographique des sites d’échantillonnage et de leur niveau de perturbation agricole correspondant au score sur le premier axe d’une analyse en composante principale (ACP), tel que présenté à la figure 2.

En 2022, la zone d’étude a été élargie en ajoutant vingt ruisseaux en s’éloignant du LSP, tout en restant dans son bassin versant. Il a donc fallu changer les critères de sélection, puisque certains étaient rendus inapplicables. Ainsi, les deux premiers points de la précédente liste n’ont pas été

appliqués pour la sélection des nouveaux sites de 2022. Les coordonnées et les dates d'échantillonnage des sites sont disponibles dans le tableau de l'annexe A.

2.2 Plan d'échantillonnage

Lors de l'été 2021, chaque site a été visité trois fois, soit une fois au début, au milieu et à la fin de la saison. Les mesures de paramètres physicochimiques et les échantillons d'ADNe étaient récoltés à chacune des visites. À la dernière visite, l'échantillonnage des diatomées et des macroinvertébrés benthiques pour la taxonomie traditionnelle a été ajouté à tous les autres. En 2022, une seule visite par site a été effectuée durant les deux dernières semaines du mois de juillet. Cela a permis l'élargissement de l'échantillonnage en incluant des sites plus loin du LSP, augmentant ainsi le pouvoir statistique de l'étude et amenant une meilleure généralisation des résultats. Pour les deux années, une équipe préterrain passait avant le début des inventaires pour délimiter le tronçon de 100m sur lequel seraient effectués les divers échantillonnages. Le même tronçon a été réutilisé pour les sites échantillonnés sur deux ans.

2.3 Protocoles terrain

2.3.1 Protocole ADNe

Le protocole pour l'échantillonnage de l'eau pour l'ADNe en ruisseau agricole est inspiré de celui de la direction générale de la gestion de la faune du MELCCFP (Ministère des Forêts, de la Faune et des Parcs, 2021). Au laboratoire, des bouteilles neuves de 500mL servant à la récolte de l'eau pour l'ADNe étaient lavées trois fois avec une solution de Javel 10%, puis rincées à l'eau ultra-pure pour éradiquer toutes traces d'ADN qui pourraient y être. Cette étape n'a pas été nécessaire en 2022, puisque des bouteilles stériles neuves ont été achetées pour cette campagne terrain. La glacière contenant les bouteilles dédiées à la collecte d'eau était nettoyée avec la solution de Javel 10%.

Toutes les manipulations sur le terrain ont été effectuées avec des gants stériles. Afin d'éviter une contamination possible du ruisseau par de l'ADNe provenant du matériel de l'équipe, l'échantillon d'eau pour l'ADNe était pris en premier, avant qu'un membre de l'équipe entre en contact avec l'eau du ruisseau pour collecter les échantillons traditionnels. L'eau était récoltée en aval du tronçon de 100m, dans une bouteille stérile de 500mL. La bouteille était fixée au bout

d'une perche télescopique pour éviter que la personne effectuant la manipulation entre en contact avec l'eau du ruisseau et contamine l'échantillon. La bouteille était d'abord rincée deux fois avec l'eau du ruisseau avant de récolter l'échantillon. Deux bouteilles étaient récoltées par site, une pour chaque rive du cours d'eau.

La membrane venait encaissée dans une capsule stérile. Cette dernière était connectée à une pompe qui tire l'eau à travers la membrane. La filtration de l'eau a été effectuée directement sur le terrain à l'aide de filtres Smith-Root « auto-préservants » (*self-preserving*) équipés d'une membrane dont la maille est de 1,2 μm . La présente étude porte sur les diatomées et les invertébrés, mais les données servent aussi à l'étude d'autres groupes taxonomiques, dont les poissons. Le choix de la maille pour l'échantillonnage a surtout été fait pour ce dernier groupe. Le choix de la taille de maille dépend du milieu d'échantillonnage et des taxons cibles (Bruce et al., 2021). Pour cette étude, les taxons ciblés sont les diatomées et les macroinvertébrés. Puisque les bactéries ne sont pas ciblées, il n'est pas essentiel d'utiliser une maille fine telle que celle de 0,22 μm ou de 0,45 μm . Pour les macroinvertébrés, il est suggéré d'utiliser une maille de 1 μm ou moins. Les ruisseaux sélectionnés pour l'étude étant souvent turbides, une maille trop petite occasionnerait des obstructions dues aux sédiments en suspensions et il serait très difficile de filtrer suffisamment d'eau pour obtenir une bonne quantité d'ADNe (Bruce et al., 2021). À la lumière de ces informations, le choix d'une maille de 1,2 μm a été effectué comme compromis entre le volume d'eau filtré et le risque de perdre des particules d'ADNe à travers les mailles. Un blanc de terrain a été récolté au dernier site de chaque journée de terrain. Une bouteille stérile contenant de l'eau ultra-pure a été maintenue ouverte dans les airs et refermée après trois secondes. Elle a ensuite été filtrée après les échantillons du site.

De retour au laboratoire à la fin de la journée, tout le matériel a été désinfecté avec une solution d'eau de javel à 10%. Les capsules contenant les filtres et l'ADNe ont été entreposées dans un réfrigérateur à 4°C jusqu'au moment de l'extraction.

2.3.2 Protocole d'échantillonnage traditionnel

Le protocole de l'Indice de Santé du Benthos pour substrat meuble a été suivi pour l'échantillonnage des macroinvertébrés benthiques (Moisan et Pelletier, 2011). Vingt coups de filet troubleau (maille de 500 μm) ont été donnés dans différents habitats (berges, débris ligneux,

macrophytes) selon la proportion relative de ces habitats dans le tronçon de 100 m, en s'assurant de donner dix coups sur chaque rive. Les échantillons ont été triés à l'aide de seaux grillagés (maille de 541 μm), puis préservés avec de l'alcool éthylique à 95%. Les échantillons ont été envoyés au Healthy Headwaters Lab de l'Université de Windsor, où ils ont été analysés selon le protocole de tri et d'identification de niveau 3 décrit dans le Guide de biosurveillance basée sur les macroinvertébrés benthiques d'eau douce du Québec (Ministère du Développement durable, de l'Environnement, de la Faune et des Parcs (MDDEFP), 2013).

Les diatomées ont été récoltées en suivant le protocole d'échantillonnage des diatomées benthiques (Ministère du Développement durable, de l'Environnement et de la Lutte contre les Changements climatiques (MDDELCC), 2018). Lorsque possible, exactement cinq roches ont été prélevées sur le tronçon. 100 cm^2 de surface, répartie entre les roches, était frotté avec une brosse à dents neuve. Le biofilm récolté à chaque site était rincé dans un même flacon de 50ml à l'aide d'eau distillée. Lorsqu'il n'était pas possible de trouver suffisamment de roches, plus de biofilms étaient prélevés sur chacune jusqu'à obtenir la surface requise. Une solution de Lugol a été ajoutée pour conserver l'échantillon. Les échantillons ont été envoyés au Centre d'Expertise en Analyse Environnementale du Québec (CEAEQ), à Québec, où ils ont été analysés selon les démarches décrites dans le Guide d'identification des diatomées des rivières de l'Est du Canada (Lavoie, Hamilton, et al., 2008).

2.3.3 Paramètres physicochimiques

Divers paramètres physicochimiques ont été mesurés lors de l'échantillonnage. Ces données serviront l'établissement du gradient de perturbation agricole pour donner une cote à chaque site selon sa position sur un axe d'une analyse en composante principale sur les paramètres physicochimiques. Le pourcentage d'oxygène dissous, la teneur en oxygène dissous, la température et la turbidité ont été mesurés avec une sonde YSI. Le pH, la conductivité et les solides dissous totaux ont été mesurés avec une sonde Hanna. Un fluoromètre de terrain BBE Benthotorch a servi à prendre des mesures de biomasse de périphyton sur les roches utilisées pour échantillonner les diatomées. Des mesures de profondeur et de largeur moyennes en eau et aux berges ont aussi été prises à chaque site aux points 0m, 50m et 100m à l'aide d'un ruban à mesurer pour les largeurs et d'un bâton de mesure pour la profondeur. Les types de substrats et la composition des berges ont été décrits en suivant les instructions du protocole d'Indices de Santé

du Benthos (ISB) pour substrat meuble (Moisan & Pelletier, 2011). Les types de substrats sont divisés en argile-limon, sable, gravier, galets, blocs et roc, pour lesquels il faut indiquer le pourcentage approximatif du ruisseau entrant dans chaque catégorie. La composition des berges fonctionne de la même manière, divisée en roches ou roc, sol nu, herbacées naturelles, arbustes, arbres, pâturage, fourrage et pelouse, culture, coupe forestière récente et infrastructure de soutien. Pour la présente étude, les types de substrats ont été regroupés en trois catégories: 1) Limon pour inclure l'argile et le limon, 2) Sable, 3) Gravier et plus gros pour les substrats de taille supérieure à l'argile, le limon et le sable. Les variables de composition des berges ont également été regroupées en trois catégories : 1) Espèces ligneuses pour inclure les arbres et arbustes, 2) Espèces herbacées et fourragères pour inclure les herbacées naturelles, pâturage, fourrage et pelouse et 3) Cultures.

2.4 Usages du territoire

La délimitation des bassins versants a été effectuée à l'aide de données LiDAR (Ministère des Ressources Naturelles et des Forêts, 2023) d'une résolution de un mètre, à partir desquelles un modèle numérique de terrain a été généré. Les ponceaux ont d'abord été ajoutés manuellement pour permettre le passage de l'eau sous les routes. La direction de l'écoulement, les limites et les aires de bassins versants ont été calculées automatiquement en utilisant le package « watershed » du logiciel ArcGIS Pro version 3.0.

Un fichier raster contenant des données d'usage des terres et de couvertures du territoire (UTCT) d'une résolution de dix mètres publié par le gouvernement (Ministère de l'Environnement, Lutte contre les changements climatiques, Faune et Parcs, 2018) a été utilisé afin d'évaluer la proportion des différents types d'utilisation du territoire dans les bassins versants des ruisseaux étudiés. Les proportions des 1186 UTCT ont été calculées pour chaque bassin versant, puis regroupées en catégories plus larges : agriculture, anthropogène, forêt, sol nu, coupe forestière, milieu humide et non classifié.

2.5 Protocoles de laboratoire

L'extraction de l'ADNe a été effectuée à l'aide de la trousse DNeasy PowerSoil Pro de Qiagen. Tout le matériel a été passé aux rayons UV pendant 20 minutes avant de commencer les

manipulations et cinq minutes avant chaque série d'extraction. Le sol étant le substrat pour lequel la trousse a été fabriquée, des modifications ont été appliquées au protocole du fabricant afin d'extraire l'ADN des filtres et de l'adapter aux appareils disponibles au laboratoire (Hermans et al., 2018). Ainsi, plutôt que de peser une quantité de sols, les filtres ont été coupés avec des ciseaux préalablement décontaminés avec de l'éthanol 70%, puis les morceaux ont été insérés dans le tube PowerBead Pro. Le Vortex Adapter a été remplacé par un Mixer Mill MM 400 ajusté à une fréquence de 30 hertz pendant dix minutes durant lesquelles le plateau a été retourné après cinq minutes. La quantification de l'ADN pour chaque échantillon a été réalisée à l'aide d'un Qubit4 et la trousse dsDNA HS Assay en suivant le protocole du fabricant avec 2 µl de l'échantillon. Les mesures étaient prises deux fois. Les échantillons ayant donné un résultat trop élevé ont été dilués dix fois avant d'être quantifiés de nouveau.

Tableau 1 Détails méthodologiques des PCRs

Fragment	COI-insectes	<i>rbcL</i>
Nom du fragment	fwhF2/EPTDr2n	Diatrbcl708F1-2-3/R3-1-2
Kit de polymérase (1ère PCR)	Phusion HF	Phusion HF
1ère dilution de PCR	1/100	1/50
Cycles pour la 1ère PCR	30	30
Température d'hybridation	56°C	52°C
Kit de polymérase (2e PCR - ajout des barcodes)	Phusion HF	Kapa HiFi
Cycles pour la 2e PCR	12	12

Le marqueur génétique COI, le gène codant pour la sous-unité I de la cytochrome oxydase mitochondriale, a été utilisé pour les macroinvertébrés, puisqu'il possède des bibliothèques de références plus complètes (Andújar et al., 2018). Récemment, une nouvelle paire d'amorces a permis de cibler des régions du gène COI plus spécifiques aux insectes (Leese et al., 2021). D'après les tests effectués par les auteurs, les amorces fwhF2/EPTDr2n permettent d'obtenir une meilleure précision taxonomique des invertébrés benthiques que les amorces classiques BF2/BR2. La bibliothèque de références pour le gène *rbcL*, qui code pour une sous-unité de l'enzyme Rubisco retrouvée dans les chloroplastes, est la plus complète pour les diatomées, permettant une assignation taxonomique qui se rapproche plus des données morphologiques que le marqueur 18S (Apothéloz-Perret-Gentil et al., 2021; Vasselon et al., 2017). Les PCRs et le séquençage ont été

effectués par le McGill Genome Center sur une plateforme Illumina MiSeq 250. Le protocole pour les PCRs est présenté dans le tableau 1.

2.6 Pipelines bio-informatiques

Les séquences résultantes des amplicons du *rbcL*, et COI ont été traitées par le pipeline *Divisive Amplicon Denoising Algorithm 2* (DADA2) dans R (version 1.16) (Callahan et al., 2016). DADA2 est utilisé pour le traitement du séquençage d’amplicons à haut débit afin d’en inférer des « variants de séquence exacts » (ASVs) à partir des lectures (« reads ») de séquences brutes.

Dans DADA2, les lectures avant (« forward ») et inverses (« reverse ») ont été évaluées pour la qualité, puis élaguées et filtrées. Les séquences ont été débruitées à l’aide de l’algorithme DADA2 (Callahan et al., 2016), puis les lectures avant et inverses ont été combinées, les ASVs ont été identifiés et comptés, et les chimères potentielles éliminées. La taxonomie a été assignée à chaque ASV en utilisant le classifieur bayésien naïf de DADA2 pour les séquences *rbcL* et la banque de données de séquences de références de *Barcode Of Life Data system* (BOLD) en utilisant le programme python Boldigger (Buchner & Leese, 2020) pour les séquences COI.

Tableau 2 Paramètres d’élitage du pipeline DADA2

Amplicon	Année d’échantillonnage	Longueur avant	Longueur inverse
<i>rbcL</i>	2021	210	210
<i>rbcL</i>	2022	210	210
COI (EPTDr2n)	2021	142	142
COI (EPTDr2n)	2022	150	150

Le traitement bio-informatique pour *rbcL* et COI adhère à la chaîne de traitement décrit dans le tutoriel du pipeline DADA2 (version 1.12). Premièrement, le logiciel cutadapt (v4.1) (Martin, 2011) a enlevé les séquences des amorces. Les paramètres de cutadapt ont été réglés pour rejeter les séquences non élaguées et le nombre maximal de nucléotides N permis à 0. La qualité des séquences pour chaque amplicon et année d’échantillonnage a été évaluée pour établir les

paramètres d'élagage (tableau 2). La qualité moyenne des séquences varie d'un lot de séquençage à l'autre, ce qui cause des variations entre les années pour la longueur de l'élagage.

Des paramètres standards de filtration ont été appliqués (maxN=0, truncQ=2, rm.phix=TRUE, maxEE=2), à l'exception d'une augmentation du maxEE à cinq pour les séquences inverses de COI afin d'améliorer la rétention des lectures. Le COI de 2022 a montré une proportion anormalement élevée de lectures chimériques en utilisant les paramètres standards. Afin de remédier à la situation, le paramètre minFoldParentOverAbundance, qui module à quel point l'abondance du parent supposé d'une chimère doit être plus élevée que celle de la chimère pour l'élimination de cette dernière, a été ajusté du deux par défaut à huit. Cela a pour effet de diminuer le nombre d'ASV éliminés à ce stade du processus.

Tableau 3 Résumé statistique du pipeline DADA2

Amplicon	Année d'échantillonnage	Total des lectures non-chimériques	% de lectures brutes après l'élimination des chimères	Nombre d'ASV après l'élimination des contaminants
<i>rbcL</i>	2021	9 841 194	71.64	5195
<i>rbcL</i>	2022	5 649 331	81.92	4432
COI (EPTDr2n)	2021	3 922 682	52.53	10965
COI (EPTDr2n)	2022	5 727 413	83.46	8787

L'assignation taxonomique de *rbcL* a été effectuée avec les fonctions assignTaxonomy (minBoot = 75) ainsi que assignSpecies du package DADA2, utilisant diat.barcode, la base de données de références la plus à jour et la mieux maintenue pour les diatomées (Rimet et al., 2019). Pour le marqueur COI, la base de données de références BOLD a été choisie pour son exhaustivité (autour de 23M entrées) et son processus de traitement (Ratnasingham et al., 2024). Cependant, aucune base de données formatée et complète de BOLD n'est maintenue pour les fonctions de classifieurs disponibles dans le pipeline DADA2. Il a donc fallu utiliser le programme BOLDigger, qui emploie le moteur d'identification BOLD. Dans BOLDigger, l'option de pipeline JAMP a été utilisée afin de trouver le meilleur résultat du top 20 des résultats rapportés par le moteur d'identification BOLDigger. Après l'assignation taxonomique, la fonction isContaminant du

package R *decontam* a servi à identifier les séquences pouvant provenir de la contamination des échantillons en examinant la fréquence et l'abondance des ASVs dans les échantillons de contrôles négatifs, c'est-à-dire tous les blancs disponibles. Les contaminants ainsi identifiés ont ensuite été éliminés de la base de données pour tous les marqueurs analysés utilisant le pipeline DADA2. Un résumé statistique du pipeline DADA2 se trouve dans le tableau 3.

2.7 Traitement des données

Comme le protocole prévoit deux filtrations d'eau, une pour la rive gauche et une pour la droite, ainsi que deux PCR pour chaque filtre, chaque site possédait quatre réplifications. Puisque certains sites avaient très peu de lectures, nous avons choisi d'additionner les répliques pour améliorer ce paramètre.

Il est pratique courante dans le traitement des données d'ADNe de vérifier si le nombre de lectures et la richesse en espèces ou en ASV sont corrélés. Une corrélation de Pearson a déterminé que c'était le cas pour le marqueur COI (EPTDr2n). Ainsi, la matrice d'espèce provenant de COI (EPTDr2n) a été raréfiée avec la fonction *rarefy* du package R *vegan* (Oksanen et al., 2018). Le même traitement a été appliqué à la matrice *rbcL* malgré l'absence de corrélation entre le nombre de lectures et la richesse en espèces afin de conserver une démarche rigoureuse.

Pour ramener les données sur des échelles comparables, une transformation de Hellinger a été appliquée sur les matrices d'espèces récoltées par ADNe et de manière traditionnelle. Cette transformation utilise la racine carrée de l'abondance relative, ce qui a pour effet de mettre les données sur la même échelle et de réduire l'impact des valeurs extrêmes ainsi que des zéros dans les jeux de données de communautés (Legendre & Gallagher, 2001). Elle a été appliquée avec la fonction *decostand* du package R *vegan*.

2.8 Optimisation des bases de données taxonomiques traditionnelles

Puisqu'une partie de l'étude consiste à comparer des données qui proviennent de différentes sources, c'est-à-dire des comptes d'espèces obtenus par des méthodes traditionnelles et du séquençage d'ADNe, il est nécessaire de s'assurer que les taxa sont identifiés sous les mêmes noms dans tous les jeux de données et d'y uniformiser la classification afin de les faire correspondre adéquatement dans les analyses. Ainsi, pour les données provenant de

l'échantillonnage traditionnel, les méthodes suivantes ont été appliquées selon le groupe d'organismes afin de rapprocher la taxonomie et la classification des taxa le plus possible de celles assignées aux séquences ADNe dans les bases de données de références.

Pour les données traditionnelles des diatomées, nous nous sommes servis de la liste des espèces du guide d'identification de l'IDEC. La première étape a été de nettoyer les noms. Cela comprend d'enlever tout ce qui suit l'espèce dans le nom scientifique, c'est-à-dire garder uniquement les deux premiers mots. Les spp., aff., fo. et cf. ont été remplacés par sp., puisque l'étude se limite aux espèces identifiées avec certitude. Les chiffres ont été éliminés, de même que le contenu des parenthèses. De plus, les groupements d'espèces ont été ramenés au genre uniquement.

La liste IDEC a été comparée avec les espèces contenues dans diat.barcode, dont les noms d'espèces ont été nettoyés de la même manière que ceux de la liste IDEC, afin de vérifier la présence des espèces de la liste IDEC dans diat.barcode. Les espèces qui n'ont pas été trouvées possèdent potentiellement des erreurs dans le nom, ce qui peut être la cause d'une absence de correspondance avec diat.barcode. Ainsi, la fonction `gnr_resolve` du package R `taxize` a été appliquée sur les espèces non trouvées afin de parcourir diverses bases de données en ligne et de trouver les correspondances les plus proches. Toutes les espèces dont le nom soumis différait de la correspondance trouvée ont été minutieusement examinées pour repérer les erreurs et vérifier la validité du nouveau nom. Les corrections ont été effectuées à partir de R pour garder une trace des changements appliqués.

La classification de la liste IDEC a ensuite été extraite de la base de données de diat.barcode. De cette manière, la classification des espèces trouvées de manière traditionnelle et par ADNe est identique. Les espèces qui n'ont pas trouvé de correspondance dans diat.barcode ont vu leur classification complétée d'abord en utilisant le genre. Celles qui n'ont pu être complétées ont été passées dans NCBI, la seule base de données utilisée par diat.barcode qui est publique. Par la suite, la classification des espèces restantes a été puisée dans GBIF, qui a permis d'effectuer la finalisation de la classification pour le restant de la liste sans avoir recours à d'autres sources. Ces dernières étapes ont été effectuées avec la fonction `classification` du package R `taxize` (Chamberlain et al., 2025).

Comme pour les diatomées, la première étape pour les invertébrés a été de nettoyer les noms. Puisque les analyses pour les invertébrés se font au genre, les noms d'espèces n'ont pas besoin d'être conservés. Le nettoyage a consisté à séparer les divers regroupements, enlever tous les sp., spp., etc., de même que les symboles et les chiffres, afin de ne garder que le nom de taxon. De plus, certaines spécifications non essentielles à l'étude, telles que groupe, adulte, famille, etc., ont aussi dû être filtrées et éliminées.

Le package R *taxize* possède une série de fonctions qui permettent d'utiliser le système BOLD à des fins de classification. La fonction *bold_search* a permis de vérifier la présence du taxon dans la base de données et d'en retourner un tableau avec certaines informations de base. Il a ainsi été possible d'établir une liste de taxons qui n'ont pas été trouvés dans BOLD et de la passer dans *gnr_resolve* pour trouver et corriger les erreurs de la même manière que pour les diatomées, en vérifiant minutieusement l'orthographe et les synonymes qui ont pu être employés. Les corrections ont été effectuées à partir de R pour garder une trace des changements appliqués. La classification BOLD peut être obtenue directement par la fonction *classification* du package R *taxize*. Aucune autre source n'a été nécessaire pour compléter la classification des taxons des invertébrés.

2.9 Épuration des bases de données taxonomiques d'ADNe

Une autre difficulté pour faire correspondre les bases de données traditionnelles avec celles issues de l'ADNe réside dans la présence de taxa improbables dans le contexte des ruisseaux agricoles étudiés, par exemple lorsqu'une erreur de séquençage introduit artificiellement une espèce tropicale. Il y a aussi la présence de taxa qui ne sont pas pris en compte dans notre étude, tels que les animaux terrestres dont l'ADN peut se retrouver dans l'échantillon d'eau récolté. Ce faisant, les méthodes qui suivent ont été appliquées après avoir assigné la taxonomie aux séquences afin d'élaguer le jeu de données pour ne garder que les taxa probables qui concernent l'étude.

Pour *rbcL*, seul le phylum *Bacillariophyta* a été conservé. Il est difficile de trier plus loin que cela, car des espèces de diatomées d'eau saumâtre peuvent être récoltées dans des ruisseaux très dégradés où la conductivité est élevée (Stenger-Kovács et al., 2023). Étant donné que les genres de diatomées sont cosmopolites (Vanormelingen et al., 2009), les distinctions géographiques sont ardues, même entre les continents. Les séquences qui n'étaient pas identifiées

à l'espèce ont été éliminées. 3 622 séquences ont été retirées avec cette méthode, ce qui nous a laissé 182 espèces de diatomées identifiées à l'aide de l'ADNe.

Pour le marqueur COI, une liste de taxa à chaque rang taxonomique, déjà observés physiquement dans des cours d'eau du Québec, a été compilée à partir de notre base de données traditionnelle ainsi que de celle de Réseaux Benthos. Cette liste a été comparée avec celle de nos données ADNe. Chaque taxon qui n'a pas été trouvé dans notre liste de présence officielle a été vérifié dans WoRMS, ou une autre base de données disponibles, telles que Catalogue of Life, afin de déterminer sa pertinence pour notre étude. Toute séquence assignée à un taxon marin, terrestre, microscopique, végétal ou vertébré a été retirée. Ce tri a été effectué jusqu'à la famille. Ensuite, la diversité de taxa était trop grande pour que ce soit justifiable de passer autant de temps à les vérifier : nous avons considéré que l'épuration à la famille serait suffisante pour garder majoritairement les genres qui sont probables dans nos échantillons. Les séquences qui n'étaient pas identifiées à la famille ont été éliminées. 3 212 séquences ont été retirées avec cette méthode, ce qui nous a laissé 131 genres d'invertébrés.

2.10 Analyses statistiques

Afin de répondre aux objectifs, il a d'abord fallu confirmer la présence du gradient de perturbation agricole. Pour ce faire, une analyse en composante principale (ACP) avec les variables environnementales a été effectuée avec la fonction `prcomp` de R de base (R Core Team, 2024). Les scores sur le premier axe deviendront une variable de perturbation agricole.

Deux indices de biodiversité, soit l'indice de diversité de Shannon et la richesse en espèces pour les diatomées et en genres pour les invertébrés, ainsi que les structures des communautés obtenues par chacune des méthodes pour chaque site, ont été choisis pour vérifier la correspondance entre les méthodes. L'indice de diversité de Shannon et la richesse ont été calculés avec la fonction `diversity` du package R `vegan`. La structure des communautés a été déterminée par un positionnement multidimensionnel non métrique (NMDS) en utilisant la fonction `metaMDS` du package R `vegan`. L'indice de dissimilarité utilisée est la distance de Chord. Les scores sur le premier axe ont servi de variable de structure de communauté pour les régressions linéaires en lien avec le gradient de perturbation agricole.

Les graphiques ont été réalisés à l'aide de diverses combinaisons des packages R ggplot2 (Wickham et al., 2025), ggrepel (Slowikowski et al., 2024) et patchwork (Pedersen, 2025).

3. Résultats

L'ACP ayant servi à déterminer la variable de perturbation agricole est présentée à la figure 2. Les deux premiers axes expliquent 44.2% de la variance dans le jeu de données. Les proportions de terres en agriculture et en forêt dans les bassins versants des ruisseaux échantillonnés, ainsi que les variables physico-chimiques de conductivité, taux de sédimentation et turbidité, vont dans le sens de l'axe 1. Les pourcentages de terres cultivées et forestières sont opposés l'un à l'autre, avec les variables physico-chimiques mentionnées précédemment suivant le pourcentage de terres agricoles. Ainsi, l'axe 1 semble bien représenter la perturbation des ruisseaux par les activités

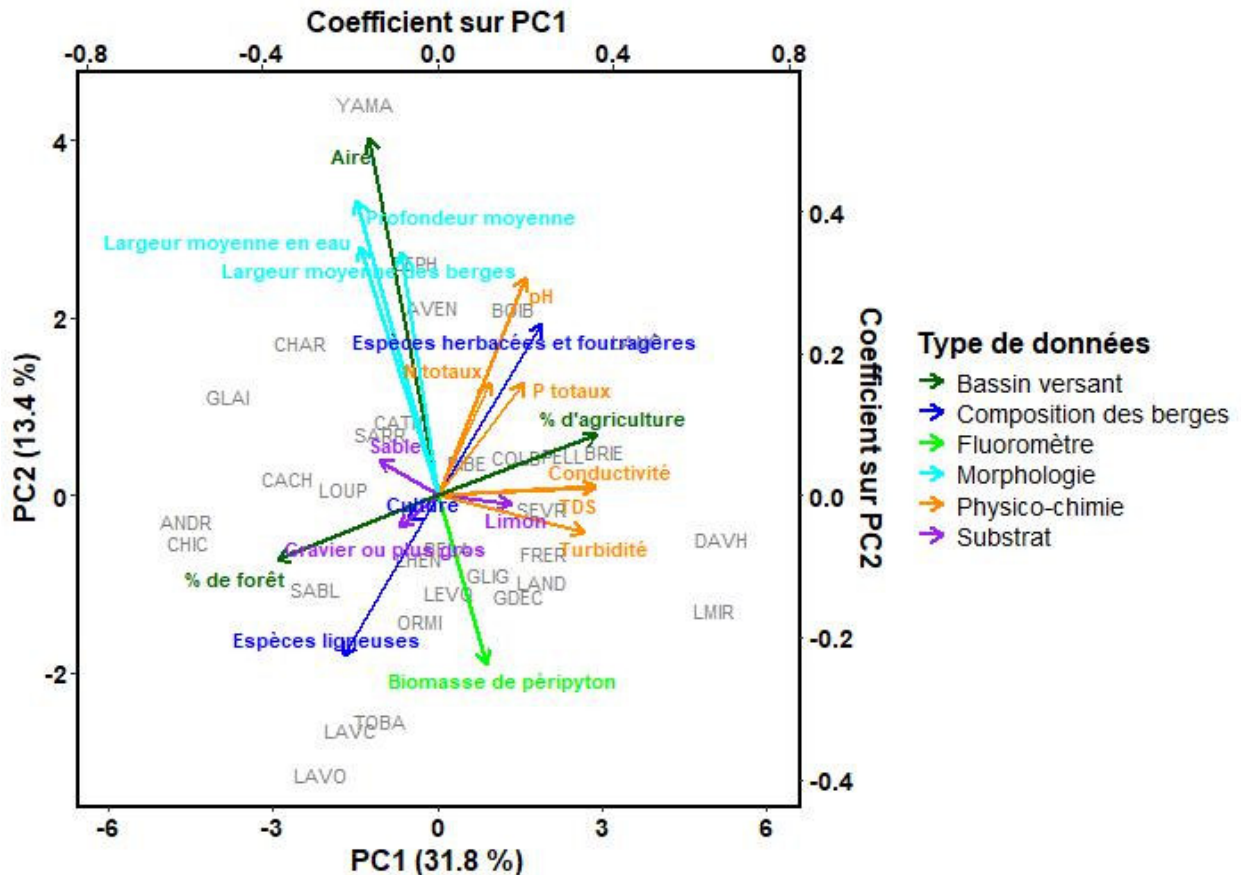


Figure 2. Analyse en composantes principales montrant la présence d'un gradient de perturbation agricole à travers les sites échantillonnés. La couleur des flèches représente le type de donnée utilisé.

agricoles, avec à gauche les sites moins perturbés et à droite les sites plus affectés. Les scores sur le premier axe ont été utilisés comme variable pour représenter le gradient de perturbation agricole dans les analyses subséquentes. L'axe 2 est plus corrélé avec des paramètres de taille du ruisseau et possiblement de débit. En effet, l'aire du bassin versant va dans la même direction que les mesures de largeurs et de profondeurs du ruisseau, en opposition avec la biomasse de périphyton présent sur les roches.

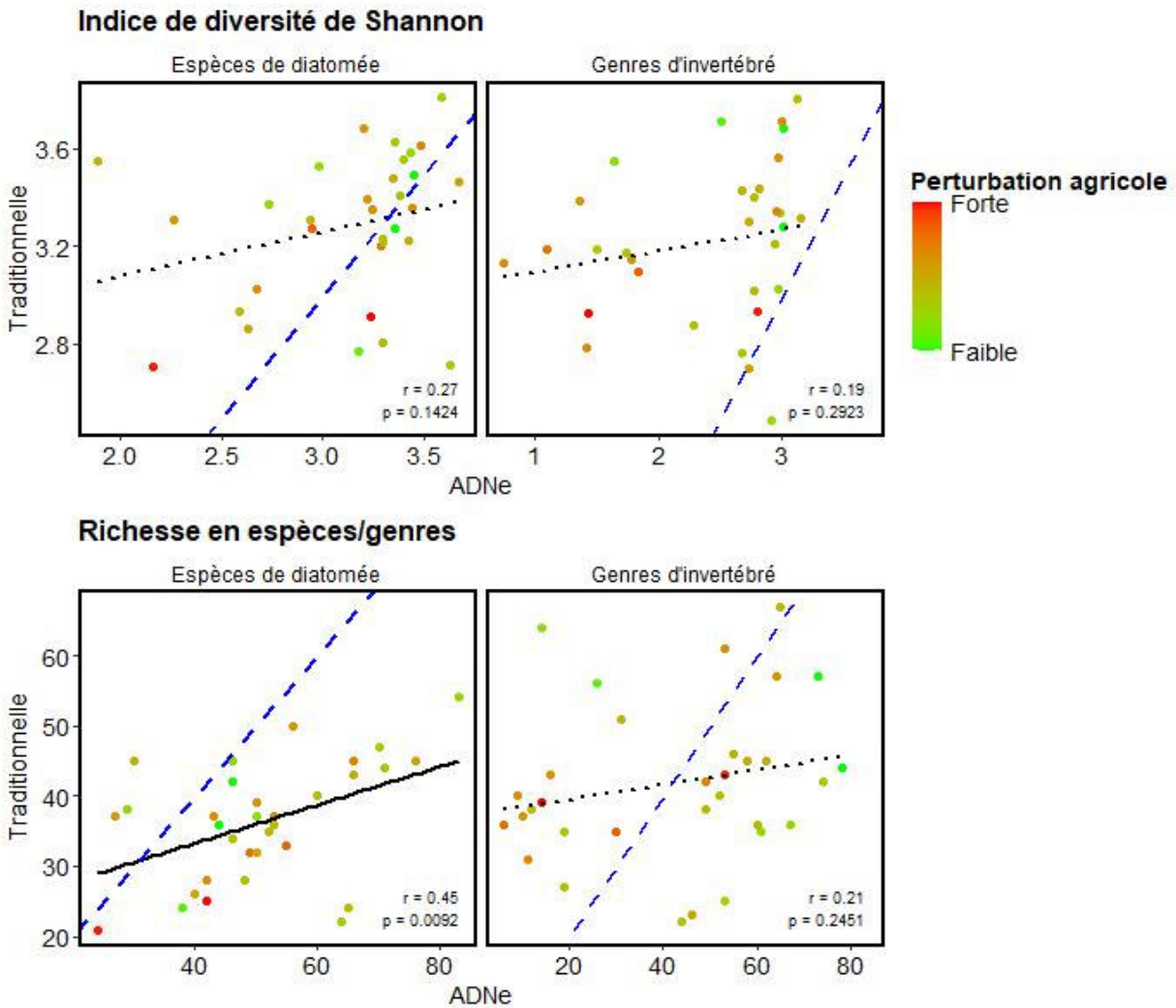


Figure 3. Comparaison de deux indices de diversité obtenus par les deux méthodes séparées selon le taxon étudié. Chaque point représente un site. La droite en tireté bleue représente la ligne d'identité ($y = x$). La droite noire est une régression linéaire. Si elle est pleine, la relation est significative ($p < 0.05$). Si elle est pointillée, la relation n'est pas significative ($p > 0.05$).

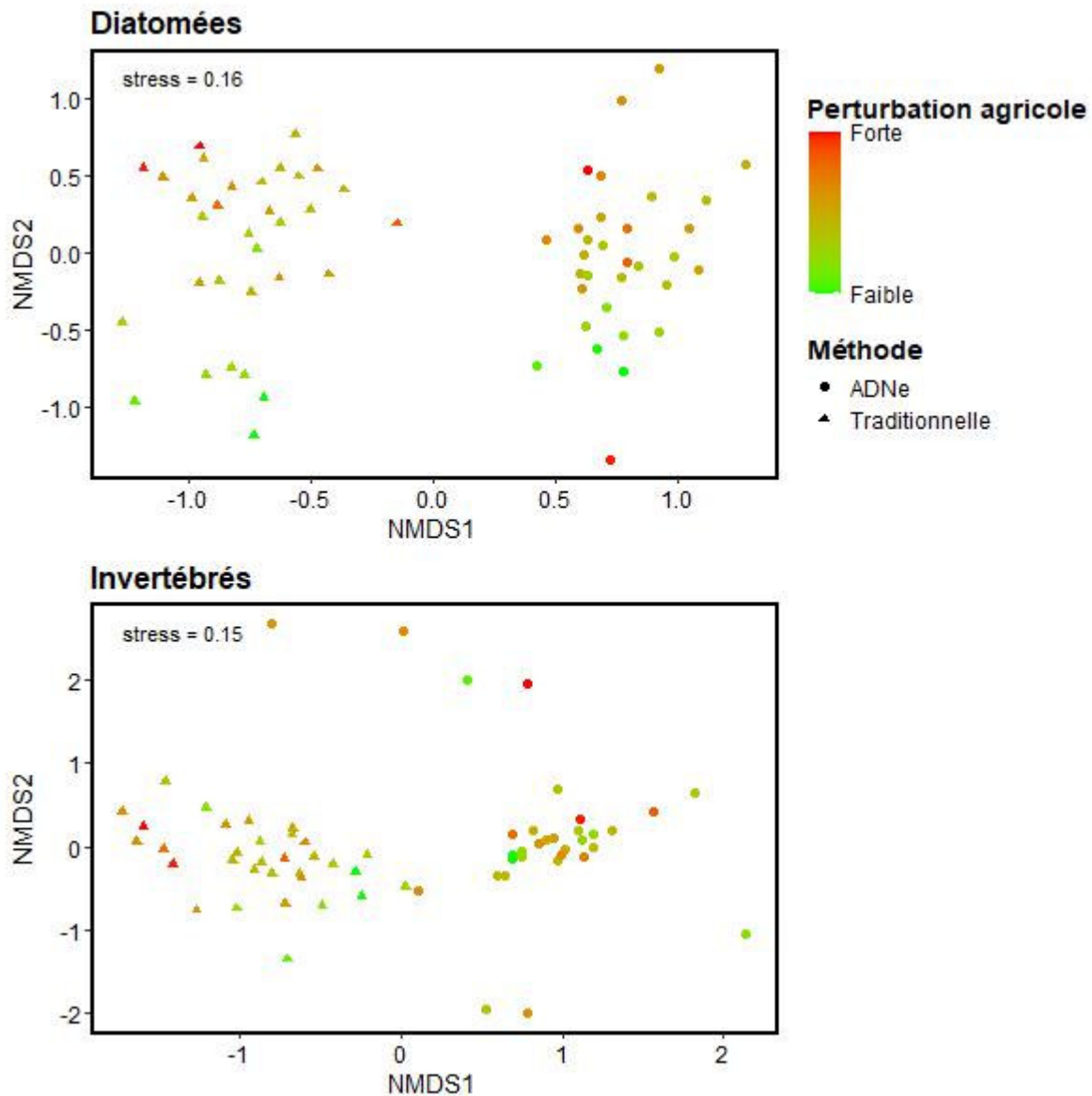


Figure 4. Compositions des communautés par ADNe et par les méthodes traditionnelles représentées par un positionnement multidimensionnel non-métrique (NMDS).

Il n'y a pas de corrélation entre les méthodes traditionnelles et par ADNe pour l'indice de diversité de Shannon (figure 3). Cet indice donne des résultats plus élevés pour les invertébrés échantillonnés de manière traditionnelle. Chez les diatomées, plus les indices de Shannon des deux méthodes sont élevés, plus ils se rapprochent de la ligne d'identité. La richesse en espèces des diatomées présente une relation significative entre les méthodes, avec des résultats généralement plus élevés pour la méthode par ADNe. La richesse en genres des invertébrés ne présente aucun patron. Les points sont distribués de part et d'autre de la ligne d'identité. La composition de la

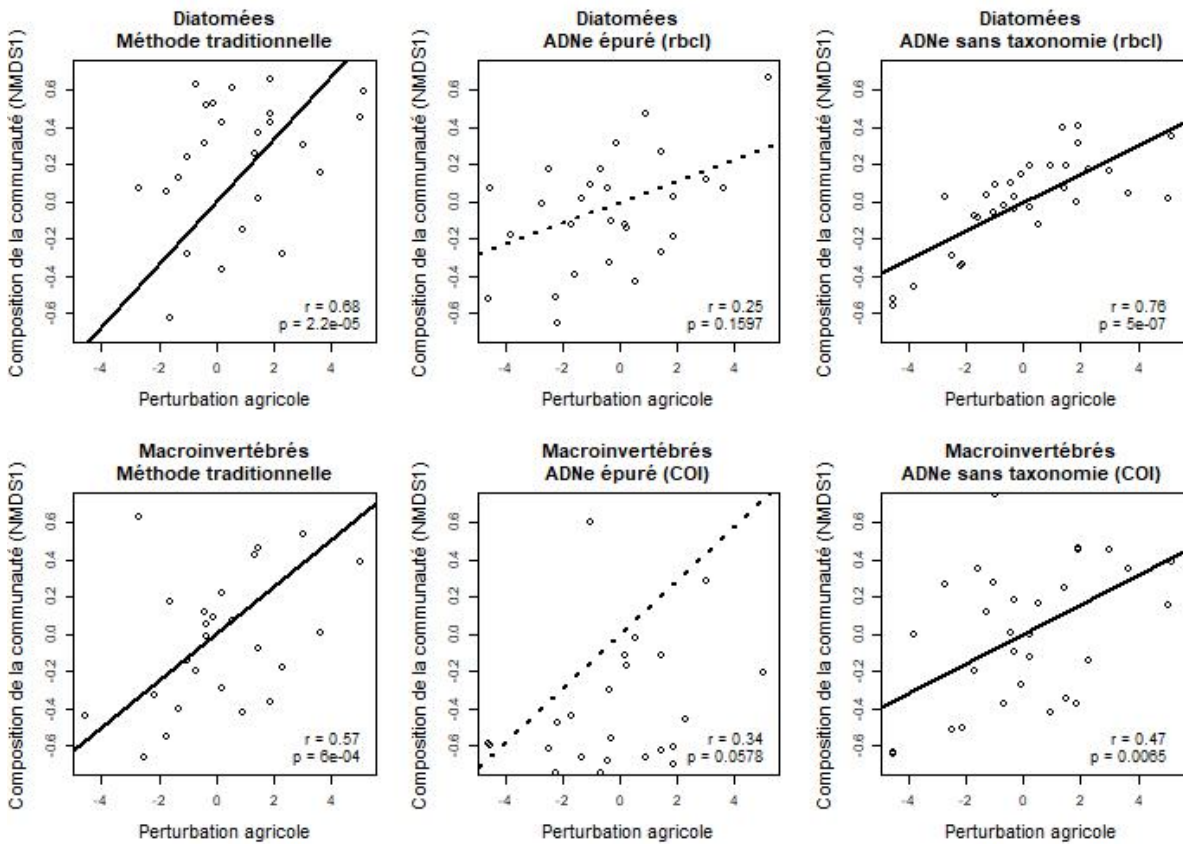


Figure 5. Réponse de la composition de la communauté (score sur une NMDS1) au gradient de perturbation agricole selon la méthode. L'ADNe sans taxonomie correspond à une épuration minimale des données où seulement les grands groupes cibles ont été conservés, c'est-à-dire *Bacillariophyta* pour les diatomées, *Annelida* et *Arthropoda* pour les macroinvertébrés.

communauté diffère grandement selon la méthode pour les deux groupes d'organismes (figure 4). En effet, les NMDS présentent des nuages de points bien définis selon la méthode d'échantillonnage. Pour les diatomées, les espèces dominantes détectées diffèrent fortement entre les deux méthodes. Par exemple, par la méthode traditionnelle, les trois espèces dominantes, lorsque l'on combine tous les échantillons, sont *Nitzschia inconspicua*, *Eolimna minima* et *Achnantheidium minutissimum*, alors qu'avec l'ADNe, il s'agit plutôt de *Nitzschia palea*, *Cyclotella meneghiniana* et *Navicula rostellata*. Chez les invertébrés, les deux méthodes présentent le même groupe dominant, la famille des *Chironomidae*. Cependant, les deuxième et troisième dominants diffèrent. Pour l'ADNe, ce sont *Cypridopsis*, des ostracodes, et *Baetis*, des éphéméroptères, tandis

que pour la méthode traditionnelle, ce sont *Caecidotea*, des isopodes, et *Gammarus*, des gammares.

La composition de la communauté pour les méthodes traditionnelles varie en fonction du gradient de perturbation agricole (figure 5). Les régressions linéaires avec la méthode par ADNe où les données ont été épurées et regroupées ne sont pas significatives, indiquant une réponse faible au gradient. Nous avons choisi d'ajouter à la comparaison la composition des communautés ADNe avec l'approche sans taxonomie. Pour ce faire, nous avons tout de même effectué une sélection des séquences associées aux cibles de nos amorces, c'est-à-dire *Bacillariophyta* pour garder les diatomées avec *rbcL* et *Annelida* et *Arthropoda* pour les invertébrés avec COI. Par la suite, la taxonomie a été abandonnée pour ne garder que les séquences. Les communautés obtenues par cette approche répondent mieux à la perturbation agricole. La méthode ADNe sans taxonomie pour les diatomées possède une relation plus forte avec le gradient de perturbation que les diatomées échantillonnées de manière traditionnelle, alors que le contraire est observé pour les macroinvertébrés.

4. Discussion

Les objectifs de l'étude étaient de 1) déterminer la correspondance entre la méthode par ADNe et les méthodes traditionnelles et 2) vérifier la capacité de l'ADNe à refléter l'impact de l'usage des terres sur les communautés de diatomées et d'invertébrés dans les petits ruisseaux.

4.1 Correspondance entre les méthodes

Nos résultats montrent une divergence entre les méthodes traditionnelles et ADNe pour toutes les métriques de diversité et de communauté. Un obstacle fréquemment mentionné dans les études sur l'ADNe est le manque de séquences dans les bibliothèques de références. Ainsi, de nombreuses séquences ne peuvent pas être assignées à une espèce et se retrouvent rejetées dans les études qui utilisent la taxonomie. Cela va affecter les métriques de communautés et de diversité, puisque les espèces présentes sur le site, mais absentes des bibliothèques de référence, ne seront pas prises en compte dans les données ADNe. Par exemple, dans nos données *rbcL*, ce sont 2669 séquences qui ont été attribuées au phylum *Bacillariophyta* et seuls 666 de ses séquences ont pu être associées à une espèce, donc utilisées dans l'étude, pour un total de 182 espèces. Compte tenu du fait que moins de la moitié des ASVs de diatomées ont servi dans les analyses, il n'est pas

démessuré de penser qu'une grande partie de la diversité ADNe a été mise de côté et que les résultats auraient été bien différents si une plus grande quantité d'ASVs avaient été identifiés à l'espèce. Cela est également soutenu par la figure 5 où l'ADNe sans taxonomie est notablement plus efficace à détecter le gradient de perturbation agricole que l'ADNe avec taxonomie. Éliminer les séquences selon la précision taxonomique élimine également des informations pertinentes dans la structure des communautés qui renseignent sur l'état du ruisseau.

Des études sur des communautés de diatomées simulées, c'est-à-dire constituées d'espèces et d'abondances connues, indiquent la possibilité de faux négatifs causés par des biais dans l'extraction de l'ADN et l'amplification par PCR (Kermarrec et al., 2013; Marinchel et al., 2023). Ce n'est donc pas uniquement le manque de séquences dans les bases de données de références qui limite la détection des espèces, mais également les étapes du metabarcoding. La conception de la présente étude ne permet cependant pas de vérifier ce type de biais dans les données.

Certains taxa peuvent être difficiles à différencier précisément avec les critères morphologiques utilisés dans les méthodes traditionnelles (Mann et al., 2010; Pfrender et al., 2010). On se retrouve ainsi souvent avec des regroupements plus larges, tels que les *Chironomidae* tous ensemble chez les invertébrés, alors que l'ADNe peut les identifier de manière plus précise (Ruppert et al., 2019). Il est possible de reprendre l'exemple des *Chironomidae* avec les données. En effet, alors que l'identification morphologique traditionnelle les regroupe à la famille, l'ADNe a pu identifier 42 genres différents de *Chironomidae*. Ainsi, des sites riches en taxa difficiles à identifier morphologiquement peuvent voir leur richesse en espèce augmenter drastiquement avec la méthode par ADNe comparé à des sites où il y en a peu. Notamment, l'ADNe a sorti 30 genres de *Chironomidae* pour le site LOUP, alors que BOIB n'en possède qu'un seul.

Les méthodes d'échantillonnages traditionnelles sont plus ciblées envers un type du groupe d'organismes. Entre autres, frotter des roches avec une brosse à dents récolte les espèces de diatomées qui y sont fixées, alors que puiser un échantillon d'eau collecte les espèces pélagiques flottantes dans la colonne d'eau en plus de leur ADNe et de celui des espèces benthiques (Schulte et al., 2024). Par exemple, la deuxième espèce la plus captée par ADNe à travers tous les sites est *Cyclotella meneghiniana*, une espèce pélagique. La richesse en espèces des diatomées de la figure 2 illustre bien ce phénomène, puisque l'ADNe va chercher une plus grande quantité d'espèces à presque chaque site. Une bonne part de cette richesse supplémentaire peut ne pas être attribuable

à de l'ADNe, puisque ce sont les organismes vivants qui ont été pris sur le filtre. Une analyse des données selon les traits fonctionnels, tel que pélagiques VS benthiques, pourrait fournir de plus amples renseignements sur l'importance de ce phénomène. Les coups de filet dans le fond des ruisseaux récoltent les invertébrés benthiques, tandis que l'ADNe dans l'eau peut également aller chercher un signal plus fort provenant de la colonne d'eau et du milieu terrestre adjacent lessivé dans le cours d'eau. L'eau échantillonnée pour l'approche par ADNe ne provient pas uniquement du site, mais aussi de plus en amont. Elle est également passée par les terres du bassin versant, où elle a récupéré des traces génétiques d'organismes terrestres (Deiner et al., 2016). Il est donc probable d'y retrouver des espèces qui ne sont pas présentes sur le tronçon échantillonné par l'approche traditionnelle, créant une discordance entre l'approche par ADNe et traditionnelle. Dans une étude plus localisée, cela peut être une source importante de biais de ne pas savoir si l'espèce est dans le site ou en amont. Cependant, ce peut être un avantage intéressant si le but est d'avoir une vue d'ensemble du ruisseau et de son bassin versant. En effet, en conservant les séquences de toutes origines avec l'approche ADNe sans taxonomie pour les invertébrés, on obtient une meilleure corrélation avec le gradient de perturbation agricole que lorsque les données sont épurées pour ne garder que les taxons aquatiques (figure 5). Les séquences de taxons terrestres fournissent donc des informations importantes sur l'usage du territoire d'un bassin versant.

L'indice de diversité de Shannon prend en compte non seulement la richesse de la communauté, mais aussi l'équitabilité, c'est-à-dire la répartition des abondances entre les espèces dans la communauté. Ainsi, une communauté qui possède la même quantité de toutes les espèces aurait un score élevé. Au contraire, une communauté dont une espèce domine et toutes les autres n'ont que quelques individus obtiendrait un score faible. Donc, en plus des points mentionnés précédemment qui agissent sur la richesse, l'indice de Shannon peut aussi être influencé par des facteurs qui biaisent l'équitabilité. En effet, certaines espèces vont posséder moins de copies du gène ciblé et seront sous-représentées, alors que celles en possédant de nombreuses copies seront surreprésentées (Krehenwinkel et al., 2017; Vasselon et al., 2018). De plus, certains taxa d'invertébrés et de diatomées libèrent plus d'ADNe que d'autres selon leurs traits (Elbrecht & Leese, 2015; Godhe et al., 2008; Hajibabaei et al., 2019; Vasselon et al., 2017). Puisque cet indice de diversité prend en compte l'équitabilité, ces biais modifient sa valeur en modifiant les écarts entre les espèces abondantes et les espèces rares. Ce semble être particulièrement le cas pour les invertébrés, où tous les sites, à l'exception d'un, possède un indice de Shannon plus élevé pour la

méthode traditionnelle. Il faut ainsi être prudent dans l'interprétation des indices de biodiversité qui prennent en compte les abondances.

4.2 Effet de l'usage du territoire

Les communautés échantillonnées de manière traditionnelle répondent bien au gradient de perturbation déterminé par l'ACP. Puisque ce sont des organismes reconnus pour y être sensibles, cela confirme que la méthode pour établir la variable de perturbation a bien fonctionné.

Le gradient de perturbation est mieux détecté par l'ADNe sans taxonomie que celui où les séquences ont été épurées. En prenant tous les ASVs des groupes d'organismes cibles sans en rejeter, la précision devient très grande. Cela rejoint les résultats de Smucker et al. (2020) qui ont pu quantifier, à l'aide de l'ADN de diatomées benthiques, les changements dans l'assemblage des communautés le long de gradients de phosphores et de nitrates sans utiliser la taxonomie. Il est intéressant de noter que l'usage de l'ADN de diatomées benthiques et celle d'ADNe dans l'eau sont capables de capter un signal de perturbation agricole sans avoir à passer par la taxonomie. Cependant, un ASV ne renseigne pas sur l'espèce à laquelle l'ADNe appartient. En effet, toutes les connaissances écologiques accumulées sont portées par l'espèce. Sans étiquette taxonomique, il n'est pas possible de faire une interprétation écologique des résultats. Par exemple, il n'est pas possible de savoir avec les résultats sans taxonomie quelles sont les caractéristiques qui font qu'une séquence est plus abondante dans un environnement riche en nutriments par rapport à une séquence différente qui en serait absente. Ainsi, si l'objectif est simplement d'établir le degré de perturbation d'un cours d'eau, une approche sans taxonomie semble tout indiquée. En revanche, si des chercheurs souhaitent étudier les changements dans les communautés en fonction des perturbations agricoles, mais ne peuvent pas identifier les espèces, l'ADNe perd de son utilité, d'où l'importance d'élargir le nombre de séquences contenues dans les bases de données de références.

Lorsque les séquences qui ont été éliminées à cause d'une précision taxonomique trop faible ou de traits ne correspondant pas à ce qu'on pourrait récupérer dans un échantillon traditionnel, c'est-à-dire les données épurées, le signal du gradient de perturbation est très affaibli. Pour les diatomées, cela peut être attribuable à l'exclusion de taxa de diatomées, particulièrement ceux qui sont plus abondants, de la matrice d'espèces va diminuer la sensibilité de cette dernière

aux changements écologiques (Lavoie et al., 2009). Les bibliothèques de références n'ayant pas les séquences pour plusieurs des espèces, ces dernières n'ont pas été comptabilisées dans la matrice épurée. Par conséquent, des espèces abondantes ont pu y être ignorées. Il faut ajouter à cela la perte de résolution taxonomique due aux méthodes d'optimisation et d'épuration, établies pour favoriser la correspondance entre les approches. En effet, ces étapes n'élimineraient pas autant de séquences et de variants intraspécifiques si ce n'était pas de tenter d'optimiser la correspondance. L'évaluation de l'intégrité d'un site à l'aide des macroinvertébrés benthiques ne requiert pas une identification à l'espèce (Jones, 2008). Par contre, une augmentation de la résolution taxonomique permettrait d'aller chercher les variations plus subtiles dans l'état du ruisseau (Turunen et al., 2021). Ainsi, en théorie, les communautés d'invertébrés détectées par ADNe devraient mieux refléter le gradient de perturbation, notamment grâce à une précision taxonomique plus fine. Or, ce n'est pas ce que l'on observe. Du moins, cette méthode ne devrait pas être désavantagée par une éventuelle baisse de précision. Il est possible que l'amélioration de la résolution taxonomique pour certains groupes, tels que les *Chironomidae*, ne suffise pas à compenser la perte de diversité observée chez d'autres taxa. Par exemple, les mollusques passent de 17 taxa identifiés par la méthode traditionnelle à seulement trois dans les données COI filtrées.

La meilleure résolution de l'ADNe sans taxonomie permet de détecter plus fortement le signal agricole que la méthode traditionnelle pour les diatomées. Une étude révèle que la variation génétique intraspécifique, qui est nommée en anglais « concealed diversity » de par le fait qu'elle n'est pas identifiable au microscope, cache des tolérances différentes pour certains génotypes, telles que la réponse au taux de calcium et à la conductivité qui s'opposent (Pérez-Burillo et al., 2021; Zimmermann et al., 2015). Les espèces pélagiques pourraient également contribuer à la force de la corrélation en apportant une dimension de plus à la communauté.

Pour les invertébrés, la présence dans l'ADNe de taxa terrestres et d'invertébrés trop petits pour être comptabilisés par les méthodes traditionnelles peut être à l'origine de la meilleure correspondance au gradient de l'ADNe à l'aveugle comparativement aux données épurées. De plus, les taxa difficiles à identifier à l'espèce morphologiquement sont différenciés avec l'ADNe, ce qui permet à des groupes importants, comme les *Chironomidae*, qui sont regroupés tous ensemble avec la méthode traditionnelle, d'être séparés et associés à différents degrés de perturbation. Cependant, c'est la méthode traditionnelle qui répond le mieux à la perturbation

agricole pour les invertébrés. En effet, celle-ci ne comptabilise pas uniquement les annélides et les arthropodes, mais également plusieurs mollusques ainsi que les nématodes, plathelminthes et cnidaires, ce qui peint un portrait plus complet de la communauté d'invertébrés et de l'impact de l'usage du territoire. Les amorces utilisées dans la présente étude sont très efficaces pour détecter les arthropodes en limitant l'amplification de taxa non ciblés, mais ne sont pas plus efficaces dans la détection d'autres groupes, comme les mollusques (Leese et al., 2021).

5. Conclusion

L'ADNe est un outil prometteur qui possède le potentiel de révolutionner les études en écologie des communautés et la biosurveillance, notamment en simplifiant les échantillonnages. Notre étude montre qu'il est effectivement possible, avec l'ADNe extrait d'un même échantillon d'eau, de capter le signal de perturbation agricole dans les petits ruisseaux pour les diatomées et les invertébrés. Cependant, les biodiversités détectées par les deux méthodes divergent.

Actuellement, les techniques moléculaires ne peuvent pas remplacer les méthodes traditionnelles et peut-être n'y arriveront-elles jamais. Cependant, leurs usages en parallèle peuvent se compléter l'un l'autre et créer des jeux de données qui pourront élargir nos connaissances sur les communautés aquatiques et sur les facteurs qui influencent les assemblages d'ADNe. L'outil ADNe a effectivement besoin d'être validé sur une durée plus longue et sur un territoire plus grand avant d'être utilisé à des fins de biosurveillance. Des efforts doivent également être fournis pour accroître le nombre de séquences des bibliothèques de référence avec des séquences correspondant aux espèces que l'on trouve dans les cours d'eau du Québec. Il est aussi envisageable de faire une étude similaire avec des sédiments pour les organismes benthiques, qui donnerait possiblement des résultats plus similaires à ceux obtenus de manière traditionnelle pour les diatomées benthiques et les invertébrés, puisque c'est un substrat qui est plus près de ces organismes.

De nombreuses avenues n'ont toujours pas été explorées avec les données récoltées lors de l'étude. Des analyses plus poussées sont à considérer pour mieux comprendre ce qui fait varier les communautés détectées par l'ADNe dans les petits ruisseaux, tel que des caractéristiques physiques du ruisseau, comme la taille et le débit, ou les traits fonctionnels, et confirmer les raisons pour lesquelles elles ne correspondent pas avec les méthodes traditionnelles. Il serait intéressant

d'approfondir l'étude pour mieux examiner les différences entre les espèces détectées et leurs abondances et, ainsi, déterminer les raisons précises de la divergence entre les communautés ADN et les méthodes traditionnelles.

7. Références

Adamowicz, S. J., Boatwright, J. S., Chain, F., Fisher, B. L., Hogg, I. D., Leese, F., Lijtmaer, D. A., Mwale, M., Naaum, A. M., Pochon, X., Steinke, D., Wilson, J.-J., Wood, S., Xu, J., Xu, S., Zhou, X., & van der Bank, M. (2019). Trends in DNA barcoding and metabarcoding. *Genome*, 62(3), v-viii. <https://doi.org/10.1139/gen-2019-0054>

Allan, J. D. (2004). Landscapes and Riverscapes: The Influence of Land Use on Stream Ecosystems. *Annual Review of Ecology, Evolution, and Systematics*, 35, 257-284.

Andújar, C., Arribas, P., Yu, D. W., Vogler, A. P., & Emerson, B. C. (2018). Why the COI barcode should be the community DNA metabarcode for the metazoa. *Molecular Ecology*, 27(20), 3968-3975. <https://doi.org/10.1111/mec.14844>

Apothéloz-Perret-Gentil, L., Bouchez, A., Cordier, T., Cordonier, A., Guéguen, J., Rimet, F., Vasselon, V., & Pawlowski, J. (2021). Monitoring the ecological status of rivers with diatom eDNA metabarcoding: A comparison of taxonomic markers and analytical approaches for the inference of a molecular diatom index. *Molecular Ecology*, 30(13), 2959-2968. <https://doi.org/10.1111/mec.15646>

Apothéloz-Perret-Gentil, L., Cordonier, A., Straub, F., Iseli, J., Esling, P., & Pawlowski, J. (2017). Taxonomy-free molecular diatom index for high-throughput eDNA biomonitoring. *Molecular Ecology Resources*, 17(6), 1231-1242. <https://doi.org/10.1111/1755-0998.12668>

Barnes, A. D., Allen, K., Kreft, H., Corre, M. D., Jochum, M., Veldkamp, E., Clough, Y., Daniel, R., Darras, K., Denmead, L. H., Farikhah Haneda, N., Hertel, D., Knohl, A., Kotowska, M. M., Kurniawan, S., Meijide, A., Rembold, K., Edho Prabowo, W., Schneider, D., ... Brose, U. (2017). Direct and cascading impacts of tropical land-use change on multi-trophic biodiversity. *Nature Ecology & Evolution*, 1(10), Article 10. <https://doi.org/10.1038/s41559-017-0275-7>

- Barnes, M. A., Turner, C. R., Jerde, C. L., Renshaw, M. A., Chadderton, W. L., & Lodge, D. M. (2014). Environmental Conditions Influence eDNA Persistence in Aquatic Systems. *Environmental Science & Technology*, 48(3), 1819-1827. <https://doi.org/10.1021/es404734p>
- Blann, K. L., Anderson, J. L., Sands, G. R., & Vondracek, B. (2009). Effects of Agricultural Drainage on Aquatic Ecosystems : A Review. *Critical Reviews in Environmental Science and Technology*, 39(11), 909-1001. <https://doi.org/10.1080/10643380801977966>
- Blaszczak, J. R., Koenig, L. E., Mejia, F. H., Gómez-Gener, L., Dutton, C. L., Carter, A. M., Grimm, N. B., Harvey, J. W., Helton, A. M., & Cohen, M. J. (2023). Extent, patterns, and drivers of hypoxia in the world's streams and rivers. *Limnology and Oceanography Letters*, 8(3), 453-463. <https://doi.org/10.1002/lol2.10297>
- Bruce, K., Blackman, R., Bourlat, S. J., Hellström, A. M., Bakker, J., Bista, I., Bohmann, K., Bouchez, A., Brys, R., Clark, K., Elbrecht, V., Fazi, S., Fonseca, V., Hänfling, B., Leese, F., Mächler, E., Mahon, A. R., Meissner, K., Panksep, K., ... Deiner, K. (2021). A practical guide to DNA-based methods for biodiversity assessment. *Advanced Books*, 1, e68634. <https://doi.org/10.3897/ab.e68634>
- Buchner, D., & Leese, F. (2020). BOLDigger – a Python package to identify and organise sequences with the Barcode of Life Data systems. *Metabarcoding and Metagenomics*, 4, e53535. <https://doi.org/10.3897/mbmg.4.53535>
- Budnick, W. R., Lebourer, T., Belliard, J., Soininen, J., Lavoie, I., Pound, K., Jamoneau, A., Tison-Rosebery, J., Tales, E., Pajunen, V., Campeau, S., & Passy, S. I. (2019). Local and regional drivers of taxonomic homogenization in stream communities along a land use gradient. *Global Ecology and Biogeography*, 28(11), 1597-1609. <https://doi.org/10.1111/geb.12976>
- Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2 : High-resolution sample inference from Illumina amplicon data. *Nature Methods*, 13(7), 581-583. <https://doi.org/10.1038/nmeth.3869>
- Carpenter, S. R., Caraco, N. F., Correll, D. L., Howarth, R. W., Sharpley, A. N., & Smith, V. H. (1998). Nonpoint Pollution of Surface Waters with Phosphorus and Nitrogen. *Ecological*

Applications, 8(3), 559-568. [https://doi.org/10.1890/1051-0761\(1998\)008%255B0559:NPOSWW%255D2.0.CO;2](https://doi.org/10.1890/1051-0761(1998)008%255B0559:NPOSWW%255D2.0.CO;2)

Chamberlain, S., Szoezs, E., Foster, Z., Arendsee, Z., Boettiger, C., Ram, K., Bartomeus, I., Baumgartner, J., O'Donnell, J., Oksanen, J., Tzovaras, B. G., Marchand, P., Tran, V., Salmon, M., Li, G., Grenié, M., & rOpenSci. (2025). *taxize : Taxonomic Information from Around the Web* (Version 0.10.0) [Software]. <https://cran.r-project.org/web/packages/taxize/index.html>

Cordier, T., Lanzén, A., Apothéloz-Perret-Gentil, L., Stoeck, T., & Pawlowski, J. (2019). Embracing Environmental Genomics and Machine Learning for Routine Biomonitoring. *Trends in Microbiology*, 27(5), 387-397. <https://doi.org/10.1016/j.tim.2018.10.012>

Creer, S., Deiner, K., Frey, S., Porazinska, D., Taberlet, P., Thomas, W. K., Potter, C., & Bik, H. M. (2016). The ecologist's field guide to sequence-based identification of biodiversity. *Methods in Ecology and Evolution*, 7(9), 1008-1018. <https://doi.org/10.1111/2041-210X.12574>

Dauphin, D., & Jobin, B. (2016). Changements de l'occupation du sol dans la plaine inondable du lac Saint-Pierre entre les années 1950 et 1997. *Le Naturaliste canadien*, 140(1), 42-52. <https://doi.org/10.7202/1034097ar>

Deiner, K., Fronhofer, E. A., Mächler, E., Walser, J.-C., & Altermatt, F. (2016). Environmental DNA reveals that rivers are conveyor belts of biodiversity information. *Nature Communications*, 7(1), Article 1. <https://doi.org/10.1038/ncomms12544>

Drummond, A. J., Newcomb, R. D., Buckley, T. R., Xie, D., Dopheide, A., Potter, B. C., Heled, J., Ross, H. A., Tooman, L., Grosser, S., Park, D., Demetras, N. J., Stevens, M. I., Russell, J. C., Anderson, S. H., Carter, A., & Nelson, N. (2015). Evaluating a multigene environmental DNA approach for biodiversity assessment. *GigaScience*, 4(1), s13742-015-0086-1. <https://doi.org/10.1186/s13742-015-0086-1>

Elbrecht, V., & Leese, F. (2015). Can DNA-Based Ecosystem Assessments Quantify Species Abundance? Testing Primer Bias and Biomass—Sequence Relationships with an Innovative Metabarcoding Protocol. *PLOS ONE*, 10(7), e0130324. <https://doi.org/10.1371/journal.pone.0130324>

Environment and Climate Change Canada. (s. d.). *Water : CABIN - Canada.ca*. Consulté 18 avril 2022, à l'adresse <https://www.canada.ca/en/environment-climate-change/services/canadian-aquatic-biomonitoring-network.html>

Findlay, S., Quinn, J. M., Hickey, C. W., Burrell, G., & Downes, M. (2001). Effects of land use and riparian flowpath on delivery of dissolved organic carbon to streams. *Limnology and Oceanography*, 46(2), 345-355. <https://doi.org/10.4319/lo.2001.46.2.0345>

Garvey, J. E., Whiles, M. R., & Streicher, D. (2007). A hierarchical model for oxygen dynamics in streams. *Canadian Journal of Fisheries and Aquatic Sciences*, 64(12), 1816-1827. <https://doi.org/10.1139/f07-144>

Giacomazzo, M., Bertolo, A., Brodeur, P., Massicotte, P., Goyette, J.-O., & Magnan, P. (2020). Linking fisheries to land use : How anthropogenic inputs from the watershed shape fish habitat quality. *Science of The Total Environment*, 717, 135377. <https://doi.org/10.1016/j.scitotenv.2019.135377>

Giroux, I. (2022). *Présence de pesticides dans l'eau au Québec : Portrait et tendances dans les zones de maïs et de soya—2018-2020*. ministère de l'Environnement et de la Lutte contre les changements climatiques.

Godhe, A., Asplund, M. E., Härnström, K., Saravanan, V., Tyagi, A., & Karunasagar, I. (2008). Quantification of Diatom and Dinoflagellate Biomasses in Coastal Marine Seawater Samples by Real-Time PCR. *Applied and Environmental Microbiology*. <https://doi.org/10.1128/AEM.01298-08>

Gossner, M. M., Lewinsohn, T. M., Kahl, T., Grassein, F., Boch, S., Prati, D., Birkhofer, K., Renner, S. C., Sikorski, J., Wubet, T., Arndt, H., Baumgartner, V., Blaser, S., Blüthgen, N., Börschig, C., Buscot, F., Diekötter, T., Jorge, L. R., Jung, K., ... Allan, E. (2016). Land-use intensification causes multitrophic homogenization of grassland communities. *Nature*, 540(7632), Article 7632. <https://doi.org/10.1038/nature20575>

Gurnell, A. M., Gregory, K. J., & Petts, G. E. (1995). The role of coarse woody debris in forest aquatic habitats : Implications for management. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 5(2), 143-166. <https://doi.org/10.1002/aqc.3270050206>

- Hajibabaei, M., Porter, T. M., Robinson, C. V., Baird, D. J., Shokralla, S., & Wright, M. T. G. (2019). Watered-down biodiversity? A comparison of metabarcoding results from DNA extracted from matched water and bulk tissue biomonitoring samples. *PLOS ONE*, *14*(12), e0225409. <https://doi.org/10.1371/journal.pone.0225409>
- Henley, W. F., Patterson, M. A., Neves, R. J., & Lemly, A. D. (2000). Effects of Sedimentation and Turbidity on Lotic Food Webs : A Concise Review for Natural Resource Managers. *Reviews in Fisheries Science*, *8*(2), 125-139. <https://doi.org/10.1080/10641260091129198>
- Hermans, S. M., Buckley, H. L., & Lear, G. (2018). Optimal extraction methods for the simultaneous analysis of DNA from diverse organisms and sample types. *Molecular Ecology Resources*, *18*(3), 557-569. <https://doi.org/10.1111/1755-0998.12762>
- Jones, F. C. J. C. (2008). Taxonomic sufficiency : The influence of taxonomic resolution on freshwater bioassessments using benthic macroinvertebrates. *Environmental Reviews*. <https://doi.org/10.1139/A07-010>
- Karr, J. R. (1991). Biological Integrity : A Long-Neglected Aspect of Water Resource Management. *Ecological Applications*, *1*(1), 66-84. <https://doi.org/10.2307/1941848>
- Kermarrec, L., Franc, A., Rimet, F., Chaumeil, P., Humbert, J. F., & Bouchez, A. (2013). Next-generation sequencing to inventory taxonomic diversity in eukaryotic communities : A test for freshwater diatoms. *Molecular Ecology Resources*, *13*(4), 607-619. <https://doi.org/10.1111/1755-0998.12105>
- Kort, J., Collins, M., & Ditsch, D. (1998). A review of soil erosion potential associated with biomass crops. *Biomass and Bioenergy*, *14*(4), 351-359. [https://doi.org/10.1016/S0961-9534\(97\)10071-X](https://doi.org/10.1016/S0961-9534(97)10071-X)
- Krehenwinkel, H., Wolf, M., Lim, J. Y., Rominger, A. J., Simison, W. B., & Gillespie, R. G. (2017). Estimating and mitigating amplification bias in qualitative and quantitative arthropod metabarcoding. *Scientific Reports*, *7*(1), 17668. <https://doi.org/10.1038/s41598-017-17333-x>
- Lavoie, I., Campeau, S., Darchambeau, F., Cabana, G., & Dillon, P. J. (2008). Are diatoms good integrators of temporal variability in stream water quality? *Freshwater Biology*, *53*(4), 827-841. <https://doi.org/10.1111/j.1365-2427.2007.01935.x>

- Lavoie, I., Dillon, P. J., & Campeau, S. (2009). The effect of excluding diatom taxa and reducing taxonomic resolution on multivariate analyses and stream bioassessment. *Ecological Indicators*, 9(2), 213-225. <https://doi.org/10.1016/j.ecolind.2008.04.003>
- Lavoie, I., Hamilton, P. B., Campeau, S., Grenier, M., & Dillon, P. J. (2008). *Guide d'identification des diatomées des rivières de l'est du Canada*. Presses de l'Université du Québec.
- Leese, F., Sander, M., Buchner, D., Elbrecht, V., Haase, P., & Zizka, V. M. A. (2021). Improved freshwater macroinvertebrate detection from environmental DNA through minimized nontarget amplification. *Environmental DNA*, 3(1), 261-276. <https://doi.org/10.1002/edn3.177>
- Legendre, P., & Gallagher, E. D. (2001). Ecologically meaningful transformations for ordination of species data. *Oecologia*, 129(2), 271-280. <https://doi.org/10.1007/s004420100716>
- Lytle, D. A., & Poff, N. L. (2004). Adaptation to natural flow regimes. *Trends in Ecology & Evolution*, 19(2), 94-100. <https://doi.org/10.1016/j.tree.2003.10.002>
- Mallin, M. A., Johnson, V. L., Ensign, S. H., & MacPherson, T. A. (2006). Factors contributing to hypoxia in rivers, lakes, and streams. *Limnology and Oceanography*, 51(1part2), 690-701. https://doi.org/10.4319/lo.2006.51.1_part_2.0690
- Mann, D. G., Sato, S., Trobajo, R., Vanormelingen, P., & Souffreau, C. (2010). *DNA barcoding for species identification and discovery in diatoms*. Cryptogamie, Algologie.
- Marinchel, N., Marchesini, A., Nardi, D., Girardi, M., Casabianca, S., Vernesi, C., & Penna, A. (2023). Mock community experiments can inform on the reliability of eDNA metabarcoding data : A case study on marine phytoplankton. *Scientific Reports*, 13(1), 20164. <https://doi.org/10.1038/s41598-023-47462-5>
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*, 17(1), Article 1. <https://doi.org/10.14806/ej.17.1.200>
- Ministère de l'Environnement, Lutte contre les changements climatiques, Faune et Parcs. (2018). *Utilisation du territoire* [Dataset]. Données Québec. [<https://www.donneesquebec.ca/recherche/dataset/utilisation-du-territoire>]

Ministère des Forêts, de la Faune et des Parcs. (2021). *Protocole standardisé des procédures de stérilisation et d'échantillonnage d'eau afin de déterminer la présence d'espèces fauniques dans les milieux hydriques par l'analyse d'ADNe au Québec*. Gouvernement du Québec. <https://mffp.gouv.qc.ca/nos-publications/protocole-standardise-sterilisation-echantillonnage-adne/>

Ministère des Ressources Naturelles et des Forêts. (2023). *Lidar—Modèles numériques (terrain, canopée, pente, courbe de niveau)* [Dataset]. Données Québec. <https://www.donneesquebec.ca/recherche/dataset/produits-derives-de-base-du-lidar>

Ministère du Développement durable, de l'Environnement, de la Faune et des Parcs (MDDEFP). (2013). *Guide de surveillance biologique basée sur les macroinvertébrés benthiques d'eau douce du Québec—Cours d'eau peu profonds à substrat grossier, 2013*. Direction du suivi de l'état de l'environnement.

Ministère du Développement durable, de l'Environnement et de la Lutte contre les Changements climatiques (MDDELCC). (2018). *Protocole d'échantillonnage des diatomées benthiques—Procédure à des fins de biosuivi de l'intégrité biotique des cours d'eau*. Direction de l'information sur les milieux aquatiques. <http://www.mddelcc.gouv.qc.ca/eau/flrivlac/Protocole.pdf>

Ministère du Développement durable, de l'Environnement et des Parcs (MDDEP). (2012). *Portrait de la qualité des eaux de surface au Québec 1999-2008*. Direction du suivi de l'état de l'environnement.

Moisan, J., & Pelletier, L. (2011). *Protocole d'échantillonnage des macroinvertébrés benthiques d'eau douce du Québec, Cours d'eau peu profonds à substrat meuble 2011*. Direction du suivi de l'état de l'environnement, Ministère du Développement durable, de l'Environnement et des Parcs.

Niyogi, D. K., Simon, K. S., & Townsend, C. R. (2003). Breakdown of tussock grass in streams along a gradient of agricultural development in New Zealand. *Freshwater Biology*, 48(9), 1698-1708. <https://doi.org/10.1046/j.1365-2427.2003.01104.x>

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlenn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., H. Stevens, M. H., Szoecs, E., & Wagner, H. (2018).

vegan : Community Ecology Package. <https://cran.r-project.org>,
<https://github.com/vegandevs/vegan>

Pedersen, T. L. (2025). *patchwork: The Composer of Plots* (Version 1.3.1) [Software].
<https://cran.r-project.org/web/packages/patchwork/index.html>

Pérez-Burillo, J., Trobajo, R., Leira, M., Keck, F., Rimet, F., Sigró, J., & Mann, D. G. (2021). DNA metabarcoding reveals differences in distribution patterns and ecological preferences among genetic variants within some key freshwater diatom species. *Science of The Total Environment*, 798, 149029. <https://doi.org/10.1016/j.scitotenv.2021.149029>

Pfrender, M. E., Hawkins, C. P., Bagley, M., Courtney, G. W., Creutzburg, B. R., Epler, J. H., Fend, S., Jr, L. C. F., Hartzell, P. L., Jackson, S., Larsen, D. P., Lévesque, C. A., Morse, J. C., Petersen, M. J., Ruitter, D., Schindel, D., & Whiting, M. (2010). Assessing Macroinvertebrate Biodiversity in Freshwater Ecosystems : Advances and Challenges in DNA-based Approaches. *The Quarterly Review of Biology*. <https://doi.org/10.1086/655118>

R Core Team. (2024). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>

Ratnasingham, S., Wei, C., Chan, D., Agda, J., Agda, J., Ballesteros-Mejia, L., Boutou, H. A., El Bastami, Z. M., Ma, E., Manjunath, R., Rea, D., Ho, C., Telfer, A., McKeowan, J., Rahulan, M., Steinke, C., Dorsheimer, J., Milton, M., & Hebert, P. D. N. (2024). BOLD v4: A Centralized Bioinformatics Platform for DNA-Based Biodiversity Data. *Methods in Molecular Biology (Clifton, N.J.)*, 2744, 403-441. https://doi.org/10.1007/978-1-0716-3581-0_26

Rees, H. C., Maddison, B. C., Middleditch, D. J., Patmore, J. R. M., & Gough, K. C. (2014). REVIEW : The detection of aquatic animal species using environmental DNA – a review of eDNA as a survey tool in ecology. *Journal of Applied Ecology*, 51(5), 1450-1459. <https://doi.org/10.1111/1365-2664.12306>

Relyea, R. A. (2009). A cocktail of contaminants: How mixtures of pesticides at low concentrations affect aquatic communities. *Oecologia*, 159(2), 363-376. <https://doi.org/10.1007/s00442-008-1213-9>

Richard, Y. (2006). *L'intégrité biotique de la rivière Bourlamaque : État des communautés benthiques et piscicoles* (p. 22). ministère du Développement durable, de l'Environnement et des Parcs, Direction du suivi de l'état de l'environnement.

Richard, Y. (2014). *Évolution de l'intégrité biotique des communautés piscicoles du bassin versant de la rivière Yamaska entre 1995 et 2011*, (p. 18). ministère du Développement durable, de l'Environnement et des Parcs.

Richard, Y., & Giroux, I. (2004). *Impact de l'agriculture sur les communautés benthiques et piscicoles de ruisseau Saint-Georges (Québec, Canada)* (No. ENV/2004/0226; collection n°QE/148, p. 28). ministère de l'Environnement, Direction du suivi de l'état de l'environnement.

Rimet, F., Gusev, E., Kahlert, M., Kelly, M. G., Kulikovskiy, M., Maltsev, Y., Mann, D. G., Pfannkuchen, M., Trobajo, R., Vasselon, V., Zimmermann, J., & Bouchez, A. (2019). Diat.barcode, an open-access curated barcode library for diatoms. *Scientific Reports*, 9(1), Article 1. <https://doi.org/10.1038/s41598-019-51500-6>

Ruppert, K. M., Kline, R. J., & Rahman, M. S. (2019). Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding : A systematic review in methods, monitoring, and applications of global eDNA. *Global Ecology and Conservation*, 17, e00547. <https://doi.org/10.1016/j.gecco.2019.e00547>

Schulte, N. O., Craine, J. M., Leopold, D. R., Devitt, J. K., & Fierer, N. (2024). Diversity of freshwater algal assemblages across the United States as revealed by DNA metabarcoding. *Environmental DNA*, 6(1), e477. <https://doi.org/10.1002/edn3.477>

Seibold, S., Cadotte, M. W., MacIvor, J. S., Thorn, S., & Müller, J. (2018). The Necessity of Multitrophic Approaches in Community Ecology. *Trends in Ecology & Evolution*, 33(10), 754-764. <https://doi.org/10.1016/j.tree.2018.07.001>

Shokralla, S., Spall, J. L., Gibson, J. F., & Hajibabaei, M. (2012). Next-generation sequencing technologies for environmental DNA research. *Molecular Ecology*, 21(8), 1794-1805. <https://doi.org/10.1111/j.1365-294X.2012.05538.x>

Simoneau, M. (2017). *Qualité de l'eau des tributaires du lac Saint-Pierre : Évolution temporelle 1979-2014 et portrait récent 2012-2014* [Publication gouvernementale]. Ministère du

Développement durable, de l'Environnement et de la Lutte contre les changements climatiques, Direction générale du suivi de l'état de l'environnement. <https://belsp.uqtr.ca/id/eprint/1204/>

Sirami, C., Gross, N., Baillod, A. B., Bertrand, C., Carrié, R., Hass, A., Henckel, L., Miguet, P., Vuillot, C., Alignier, A., Girard, J., Batáry, P., Clough, Y., Violle, C., Giralt, D., Bota, G., Badenhassner, I., Lefebvre, G., Gauffre, B., ... Fahrig, L. (2019). Increasing crop heterogeneity enhances multitrophic diversity across agricultural regions. *Proceedings of the National Academy of Sciences*, 116(33), 16442-16447. <https://doi.org/10.1073/pnas.1906419116>

Slowikowski, K., Schep, A., Hughes, S., Dang, T. K., Lukauskas, S., Irisson, J.-O., Kamvar, Z. N., Ryan, T., Christophe, D., Hiroaki, Y., Gramme, P., Abdol, A. M., Barrett, M., Cannoodt, R., Krassowski, M., Chirico, M., Aphalo, P., & Barton, F. (2024). *ggrepel*: Automatically Position Non-Overlapping Text Labels with « ggplot2 » (Version 0.9.6) [Software]. <https://cran.r-project.org/web/packages/ggrepel/index.html>

Smucker, N. J., Pilgrim, E. M., Nietch, C. T., Darling, J. A., & Johnson, B. R. (2020). DNA metabarcoding effectively quantifies diatom responses to nutrients in streams. *Ecological Applications*, 30(8), e02205. <https://doi.org/10.1002/eap.2205>

Soliveres, S., van der Plas, F., Manning, P., Prati, D., Gossner, M. M., Renner, S. C., Alt, F., Arndt, H., Baumgartner, V., Binkenstein, J., Birkhofer, K., Blaser, S., Blüthgen, N., Boch, S., Böhm, S., Börschig, C., Buscot, F., Diekötter, T., Heinze, J., ... Allan, E. (2016). Biodiversity at multiple trophic levels is needed for ecosystem multifunctionality. *Nature*, 536(7617), Article 7617. <https://doi.org/10.1038/nature19092>

Soria, M., Leigh, C., Datry, T., Bini, L. M., & Bonada, N. (2017). Biodiversity in perennial and intermittent rivers : A meta-analysis. *Oikos*, 126(8), 1078-1089. <https://doi.org/10.1111/oik.04118>

Stat, M., Huggett, M. J., Bernasconi, R., DiBattista, J. D., Berry, T. E., Newman, S. J., Harvey, E. S., & Bunce, M. (2017). Ecosystem biomonitoring with eDNA : Metabarcoding across the tree of life in a tropical marine environment. *Scientific Reports*, 7(1), Article 1. <https://doi.org/10.1038/s41598-017-12501-5>

- Stenger-Kovács, C., Béres, V. B., Buczkó, K., Tapolczai, K., Padišák, J., Selmeczy, G. B., & Lengyel, E. (2023). Diatom community response to inland water salinization: A review. *Hydrobiologia*, 850(20), 4627-4663. <https://doi.org/10.1007/s10750-023-05167-w>
- Stewart, K. A. (2019). Understanding the effects of biotic and abiotic factors on sources of aquatic environmental DNA. *Biodiversity and Conservation*, 28(5), 983-1001. <https://doi.org/10.1007/s10531-019-01709-8>
- Strickler, K. M., Fremier, A. K., & Goldberg, C. S. (2015). Quantifying effects of UV-B, temperature, and pH on eDNA degradation in aquatic microcosms. *Biological Conservation*, 183, 85-92. <https://doi.org/10.1016/j.biocon.2014.11.038>
- Taberlet, P., Bonin, A., Zinger, L., & Coissac, E. (2018). *Environmental DNA : For Biodiversity Research and Monitoring*. Oxford University Press.
- Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., & Willerslev, E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology*, 21(8), 2045-2050. <https://doi.org/10.1111/j.1365-294X.2012.05470.x>
- Tapolczai, K., Keck, F., Bouchez, A., Rimet, F., Kahlert, M., & Vasselon, V. (2019). Diatom DNA Metabarcoding for Biomonitoring : Strategies to Avoid Major Taxonomical and Bioinformatical Biases Limiting Molecular Indices Capacities. *Frontiers in Ecology and Evolution*, 7. <https://doi.org/10.3389/fevo.2019.00409>
- Turner, C. R., Uy, K. L., & Everhart, R. C. (2015). Fish environmental DNA is more concentrated in aquatic sediments than surface water. *Biological Conservation*, 183, 93-102. <https://doi.org/10.1016/j.biocon.2014.11.017>
- Turunen, J., Mykrä, H., Elbrecht, V., Steinke, D., Braukmann, T., & Aroviita, J. (2021). The power of metabarcoding: Can we improve bioassessment and biodiversity surveys of stream macroinvertebrate communities? *Metabarcoding and Metagenomics*, 5, e68938. <https://doi.org/10.3897/mbmg.5.68938>
- Vander Vorste, R., McElmurray, P., Bell, S., Eliason, K. M., & Brown, B. L. (2017). Does Stream Size Really Explain Biodiversity Patterns in Lotic Systems? A Call for Mechanistic Explanations. *Diversity*, 9(3), Article 3. <https://doi.org/10.3390/d9030026>

- Vanormelingen, P., Verleyen, E., & Vyverman, W. (2009). The diversity and distribution of diatoms : From cosmopolitanism to narrow endemism. Dans W. Foissner & D. L. Hawksworth (Éds.), *Protist Diversity and Geographical Distribution* (p. 159-171). Springer Netherlands. https://doi.org/10.1007/978-90-481-2801-3_12
- Vasselon, V., Bouchez, A., Rimet, F., Jacquet, S., Trobajo, R., Corniquel, M., Tapolczai, K., & Domaizon, I. (2018). Avoiding quantification bias in metabarcoding : Application of a cell biovolume correction factor in diatom molecular biomonitoring. *Methods in Ecology and Evolution*, 9(4), 1060-1069. <https://doi.org/10.1111/2041-210X.12960>
- Vasselon, V., Rimet, F., Tapolczai, K., & Bouchez, A. (2017). Assessing ecological status with diatoms DNA metabarcoding : Scaling-up on a WFD monitoring network (Mayotte island, France). *Ecological Indicators*, 82, 1-12. <https://doi.org/10.1016/j.ecolind.2017.06.024>
- Vörösmarty, C. J., McIntyre, P. B., Gessner, M. O., Dudgeon, D., Prusevich, A., Green, P., Glidden, S., Bunn, S. E., Sullivan, C. A., Liermann, C. R., & Davies, P. M. (2010). Global threats to human water security and river biodiversity. *Nature*, 467(7315), Article 7315. <https://doi.org/10.1038/nature09440>
- Walser, C. A., & Bart Jr., H. L. (1999). Influence of agriculture on in-stream habitat and fish community structure in Piedmont watersheds of the Chattahoochee River System. *Ecology of Freshwater Fish*, 8(4), 237-246. <https://doi.org/10.1111/j.1600-0633.1999.tb00075.x>
- Weigand, H., Beermann, A. J., Čiampor, F., Costa, F. O., Csabai, Z., Duarte, S., Geiger, M. F., Grabowski, M., Rimet, F., Rulik, B., Strand, M., Szucsich, N., Weigand, A. M., Willassen, E., Wyler, S. A., Bouchez, A., Borja, A., Čiamporová-Zaťovičová, Z., Ferreira, S., ... Ekrem, T. (2019). DNA barcode reference libraries for the monitoring of aquatic biota in Europe : Gap-analysis and recommendations for future work. *Science of The Total Environment*, 678, 499-524. <https://doi.org/10.1016/j.scitotenv.2019.04.247>
- West, K. M., Stat, M., Harvey, E. S., Skepper, C. L., DiBattista, J. D., Richards, Z. T., Travers, M. J., Newman, S. J., & Bunce, M. (2020). eDNA metabarcoding survey reveals fine-scale coral reef community variation across a remote, tropical island ecosystem. *Molecular Ecology*, 29(6), 1069-1086. <https://doi.org/10.1111/mec.15382>

Whitfield, A. K., & Elliott, M. (2002). Fishes as indicators of environmental and ecological changes within estuaries : A review of progress and some suggestions for the future. *Journal of Fish Biology*, 61(sA), 229-250. <https://doi.org/10.1111/j.1095-8649.2002.tb01773.x>

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. (2025). *ggplot2 : Create Elegant Data Visualisations Using the Grammar of Graphics* (Version 3.5.2) [Software]. <https://cran.r-project.org/web/packages/ggplot2/index.html>

Zimmermann, J., Glöckner, G., Jahn, R., Enke, N., & Gemeinholzer, B. (2015). Metabarcoding vs. Morphological identification to assess diatom diversity in environmental studies. *Molecular Ecology Resources*, 15(3), 526-542. <https://doi.org/10.1111/1755-0998.12336>

ANNEXE A

Sites et points d'échantillonnage

Cours d'eau	Code	Coordonnées	Date d'échantillonnage (JJ/MM/AAAA)
St-André	ANDR	46.21257241 -73.23926845	03/08/2022
Bonaventure	AVEN	46.09477651 -73.2318137	26/07/2022
Ruisseau des Bélanger	BELA	46.2836318 -73.0384573	27/07/2022
Ruisseau Bibeau	BIBE	46.11365526 -73.29410089	26/07/2022
Bois Blanc	BOIB	46.18865485 -73.05740056	31/08/2021
Brielle	BRIE	46.15668649 -72.6914982	30/08/2021
Cachée	CACH	46.19861659 -73.10214841	31/08/2021
Ste-Catherine	CATH	46.12019565 -73.21318894	26/07/2022
St-Charles	CHAR	46.2962721 -72.6929727	25/07/2022
Chicot	CHIC	46.23026719 -73.21445059	03/08/2022
Colbert	COLB	46.11666184 -72.76301175	01/09/2021
David Houle	DAVH	46.18151257 -72.66043999	30/08/2021
Des Frères	FRER	46.14183244 -72.71603191	30/08/2021
La Grande Décharge	GDEC	46.17139498 -73.0324313	01/08/2022
Aux Glaises	GLAI	46.30798337 -72.75028209	31/08/2021
Grande Ligne	GLIG	46.16022619 -73.05212083	01/08/2022
Landroche	LAND	46.12871305 -72.73940831	30/08/2021
Langevin	LANG	46.2876809 -72.8284767	27/07/2022
Lavoie (au rang du petit bois)	LAVC	46.11730904 -72.84137809	01/09/2021

Lavoie (à la sortie d'Odanak)	LAVO	46.08517029 -72.81115791	28/07/2022
Lévesque	LEVQ	46.11123 -72.776763	02/08/2022
Landry Henault	LHEN	46.15561 -73.142659	26/07/2022
Lemir	LMIR	46.164022 -72.681239	04/08/2022
Petite rivière du loup	LOUP	46.29516052 -73.03874208	27/07/2022
De l'ormière	ORMI	46.2290746 -73.0689978	01/08/2022
Ruisseau Pelletier	PELL	46.2986124 -72.9427016	27/07/2022
Aux sables	SABL	46.288496 -72.669514	25/07/2022
Sarrazin	SARR	46.18187813 -73.10589474	31/08/2021
Ruisseau Sévère René	SEVR	46.11462819 -72.58449106	04/08/2022
Bois Tobanek (Concession St-François, décharge de la 2 ^e concession)	TOBA	46.03528506 -72.81962028	01/09/2021
Petite Yamachiche	YAMA	46.30037792 -72.81809082	27/07/2022
Rivière St-Zéphirin	ZEPH	46.0826452 -72.61797953	04/08/2022