UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

ORIGINE ET DYNAMIQUE D'*ESCHERICHIA COLI* DANS LE FLEUVE SAINT-LAURENT SOUS L'INFLUENCE DES APPORTS ANTHROPIQUES

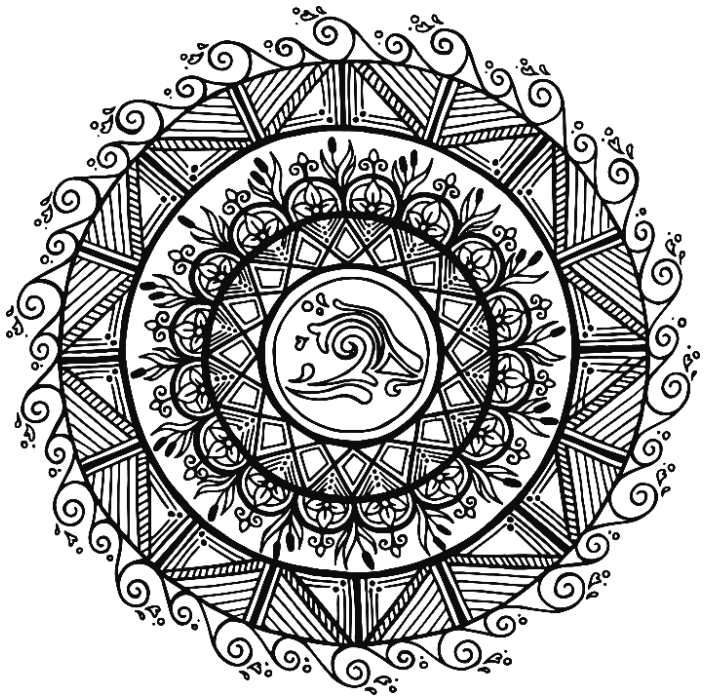
ORIGIN AND DYNAMICS OF ESCHERICHIA COLI IN THE ST. LAWRENCE RIVER UNDER THE INFLUENCE OF ANTHROPOGENIC INPUTS

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

PAR
SILVIA ESTHELA RODRIGUEZ ARGUELLES

AOÛT 2024

Ohén:ton Karihwatéhkwen



Ohneka’shòn:’a

(L’eau)

Nous remercions toute l’eau du monde d’étancher notre soif et de nous donner de la force. L’eau est source de vie. Nous connaissons son pouvoir sous de nombreuses formes: chutes et pluie, brumes et ruisseaux, rivières et océans. D’une seule voix, nous adressons nos salutations et nos remerciements à l’esprit de l’eau.

Désormais, nos esprits ne font plus qu’un.

Le discours de l’Action de grâce est connu sous le nom de Ohén:ton Karihwatéhkwen ou « les mots qui viennent avant tout ». Il s’agit d’une pratique Haudenosaunee favorisant la lucidité qui peut avoir cours dans des contextes formels, tels que des cérémonies et des événements importants, mais ce discours est également prononcé chaque matin lorsque le soleil se lève pour reconnaître tout ce qui a été créé, pour exprimer sa gratitude et pour continuer à vivre avec un bon état d’esprit. Lorsque nous continuons à exprimer notre gratitude à l’égard de tout ce qui a été créé, les éléments naturels gardent leur équilibre naturel. Ainsi, nous respectons nos instructions originelles de prendre soin de toute la création et, en retour, ces éléments continueront à respecter les leurs. Illustration par Victoria Ransom (https://riverrapport.ca/thanksgiving-address/).

Acknowledgments

Throughout my life, I have lived in various locations where, each time, I experienced a different connection to nature. From the coral coasts of my home country (Mexico) to the snowy peaks of far away New Zealand, my work in field research has taught me many ways to understand and think about environmental management issues. Over the years, whether I was aware of it or not, my interest in nature turned into passionate work. In 2021, I was presented with the incredible opportunity to pursue a master's degree in environmental sciences, specializing in water quality. Moreover, this enticing project happened to take place in a country I admire for its cultural diversity and, above all, its breathtaking natural beauty. So here I am today, presenting the results of my studies.

First, I would like to express my gratitude to my supervisor, François Guillemette, for providing me with the opportunity to develop this research in collaboration with his laboratory. His guidance, support, and mentorship have been invaluable throughout my learning and development journey. I also extend my sincere thanks to my co-supervisor, Jérôme Comte, who welcomed me into his laboratory at INRS, providing the essential support and training needed to achieve the objectives of this project.

Secondly, I want to extend my gratitude to the members of my committee. Thomas A. Edge, for his kindness and exceptional dedication, his passion for the subject and his encouragement have inspired me to go deeper into the importance of natural resource management, particularly concerning water. Additionally, I am profoundly thankful to Leigh McGaughey for her strong support and guidance throughout these years. Her willingness to offer unconditional assistance, belief in my abilities, and dedication to the evolution of the project have not only facilitated its progress but also contributed significantly to my personal and professional growth.

I am deeply grateful to the River Institute, Jeff Ridal, Christina Collard, Matthew Windle, and Leigh McGaughey for welcoming me since the summer of 2019, initially as a volunteer and now as I pursue my master's degree. Their kindness, support and

encouragement have been essential in my academic journey, and I appreciate their continuous guidance and dedication to my progress. I extend my sincere gratitude to Brian Hickey, Stephany Hildebrand, Zach Zwanenburg, Philip Ling and all dedicated staff and volunteers of the River Institute for their invaluable support and dedication during the fieldwork conducted over two summers for this project. Special thanks to Stephany Hildebrand for her incredible photography, which captured the essence of the St. Lawrence River and improved my presentations. I am also deeply grateful to Cristina Charette and Erin Smith for their invaluable support and guidance in analyzing my data; their expertise was a fundamental part in my professional development. And to the rest of the staff, although I don't mention it specifically, they were always supporting me and motivating me to achieve my success.

I want to extend my thanks to all the crew members of the *Lampsilis* 2022 expedition, whose dedication made it an unforgettable experience during that summer of sampling. Additionally, I am grateful to my laboratory MicrEau at UQTR for their support and guidance throughout this process.

I am grateful to Charis Wrong from INRS for her invaluable instruction throughout this journey, introducing me to new methods and consistently demonstrating the best willingness to guide me through this process.

I am thankful to the MITACS scholarship awarded to me through the MITACS Accelerate program, in collaboration with the River Institute. This scholarship provided me with the opportunity to fulfill my dream of studying abroad, and its contribution was essential in achieving this goal.

I want to express my gratitude to my friends, especially Claudia Galindo, Ana Farias, and Daniela Almaraz, whose support and encouragement have been a constant source of strength for me. I consider myself fortunate to have them in my life. Additionally, I am deeply thankful to my partner, Laurence Lambin-Gagnon, for his constant support, love and daily motivation that have pushed me forward. Lastly, I extend my appreciation to Ti-Mi, for his warm presence on my lap and his comforting purrs.

Above all, my family's unconditional support has been my greatest strength. Despite the physical distance, my parents, Silvia Arguelles and Javier Rodriguez, have been my pillars of strength, offering constant communication, invaluable advice, and good wishes. I am immensely grateful for the love and support of my brothers, Javier and Daniel Rodriguez, as well as my sisters-in-law, Julieta and Irene. Although being away from my nephew, Andrés, was difficult, he served as a powerful motivator for me to persevere and aspire to be a positive influence in his life. I hope to do the same for my new niece Fran. Additionally, I extend my gratitude to my aunts, Cocaya, Lupita, Boby, Marcha, Abuela Estela and Ceci, for their continuous encouragement and support throughout my journey.

To everyone who has been a part of my journey, whether mentioned here or not, your influence has been deeply meaningful. From the bottom of my heart, Thank you.

Preface

This project has been made possible by a MITACS scholarship awarded to Silvia Rodriguez as part of the MITACS Accelerate program, with funding provided in collaboration with the River Institute in Cornwall, Ontario. The scientific cruises on the St. Lawrence River, conducted onboard the Lampsilis research vessel, were made possible through the financial assistance of the Réseau Québec Maritime and the Government of Quebec via the Odyssée Saint-Laurent research program. Shallow waters sampling was supported by ZIP les Deux-Rives river keeper organization and the River Institute, and associated volunteers. This thesis represents a significant contribution to the research program of the Université du Québec à Trois-Rivières (UQTR) Chair on the Ecology of the St. Lawrence River, under the leadership of François Guillemette. The support and resources provided by these organizations and individuals, as their contributions, have been instrumental in the successful completion of this study.

Résumé

Ce mémoire présente une enquête approfondie sur la santé du fleuve Saint-Laurent, explorant ses dynamiques écologiques, ses défis et l'impact de ses activités anthropiques, avec un accent particulier sur la contamination microbienne par les coliformes fécaux tels qu'*Escherichia coli*. Le fleuve Saint-Laurent, comme d'autres grands systèmes fluviaux, fournit de nombreux services écosystémiques et joue un rôle crucial en reliant la terre et les lacs à l'océan. Il assure notamment un approvisionnement en eau potable, des opportunités récréatives, un soutien à l'agriculture par l'irrigation et la production d'électricité. Cependant, les activités humaines ont considérablement modifié ces services et, plus largement, les processus naturels du fleuve, entraînant une contamination de nature physique, chimique et bactériologique.

Malgré la surveillance historique de la contamination fécale par les autorités gouvernementales, une compréhension exhaustive de l'origine et des dynamiques de cette contamination fécale au sein du paysage fluvial complexe fait encore défaut. Cette recherche vise à expliquer comment certaines caractéristiques du paysage (par exemple, les zones urbaines, les masses d'eau, les lacs fluviaux, les systèmes insulaires) peuvent moduler le transport et la persistance d'*Escherichia coli* en aval. Basée sur un échantillonnage étendu du fleuve Saint-Laurent et des avancées récentes en traçage des sources microbiennes, ce mémoire vise également à identifier les points chauds de contamination microbienne à l'aide d'*Escherichia coli* dans le fleuve Saint-Laurent et à déterminer la persistance en aval et les sources potentielles de contamination bactérienne à travers le système.

Au cours de cinq expéditions de 2017 à 2022 (sauf en 2019), 320 échantillons d'eau ont été collectés à bord du navire de recherche *Lampsilis*, complétés par 198 échantillons d'eau peu profonde (<1 m) collectés à partir de petites embarcations. Des technologies moléculaires avancées, y compris la PCR numérique, ont été employées pour analyser les échantillons et déterminer la concentration d'*Escherichia coli* ainsi que les marqueurs ADN ciblant les sources humaines, de ruminants, de goélands et de porcs. L'étude a

constamment détecté *Escherichia coli* dans la plupart des échantillons analysés, avec 48% des sites dépassant le seuil de sécurité pour le contact humain direct et 19% pour le seuil de contact indirect. Des variations notables de la concentration bactérienne ont été observées entre les différentes masses d'eau, avec des concentrations particulièrement élevées près des zones urbaines, notamment autour du rejet des eaux usées de Montréal. La persistance des concentrations d'*Escherichia coli* sur des centaines de kilomètres en aval suggère un impact important et soutenu de cette source de contamination dans le fleuve. L'analyse temporelle a révélé une influence significative sur les niveaux d'*Escherichia coli* en 2021 et 2022, bien que des relations avec des facteurs spécifiques, tels que les niveaux d'eau, les précipitations et d'autres facteurs environnementaux, n'aient pas été clairement identifiées, suggérant que des variables non prises en compte pourraient avoir influencées les données pendant ces années. Les dynamiques dans les zones proches du rivage suggèrent une plus grande influence des sources de contamination locales comparativement aux niveaux plus stables observés dans les masses d'eau principales. Le signal humain a été identifié comme la principale source de contamination fécale, bien que quelques instances localisées de contamination par les oiseaux aient également été détectées. Malgré la bonne relation entre la concentration d'*Escherichia coli* et les copies d'ADN humain (HF 183), des écarts dans de nombreux sites suggèrent la présence de sources supplémentaires non identifiées de contamination fécale, probablement d'animaux sauvages. Ces résultats soulignent l'impact significatif des activités humaines sur la qualité de l'eau et les écosystèmes en aval, mettant en évidence que des méthodes spécialisées comme le traçage des sources microbiennes sont essentielles pour comprendre les dynamiques de contamination dans les paysages fluviaux complexes.

Le dernier chapitre synthétise les résultats de la recherche en connaissances pratiques pour la gestion et la restauration des écosystèmes, l'amélioration des mesures de qualité de l'eau et la poursuite de la recherche scientifique. En examinant l'écosystème du fleuve Saint-Laurent, cette mémoire apporte des connaissances dans le domaine de la gestion des écosystèmes aquatiques, en se concentrant sur les défis de la contamination microbienne dans les grands fleuves. En plus d'augmenter notre compréhension du rôle des microbes dans le fonctionnement et la santé des systèmes fluviaux, cette mémoire

contribuera au développement d'indicateurs écologiques microbiens de la qualité de l'eau qui seront inclus dans un rapport sur la santé de l'écosystème pour le Haut-Saint-Laurent en collaboration avec l'Institut du Fleuve Saint-Laurent à Cornwall, Ontario.

Mots-clés: Activités anthropiques, écosystèmes aquatiques, *Escherichia coli* (*E. coli*), contamination fécale, sources humaines et animales, grands fleuves, contamination microbienne, traçage des sources microbiennes, écosystèmes fluviaux, fleuve Saint-Laurent, stations de traitement des eaux usées, qualité de l'eau.

Summary

This thesis presents a comprehensive investigation of the St. Lawrence River, exploring into its ecological dynamics, challenges, and the impact of anthropogenic activities, with a particular focus on microbial contamination by fecal coliforms such as *Escherichia coli*. The St. Lawrence River, like other large river systems, plays a crucial role in connecting land and lakes to the ocean, providing multiple ecosystem services. This includes supplying drinking water, offering recreational opportunities, supporting agriculture through irrigation, and generating electricity. However, human activities have significantly altered these services and, more broadly, natural river processes, leading to contamination from physical, chemical, and bacteriological sources.

Despite historical monitoring of fecal contamination by governmental authorities, a comprehensive understanding of the origin and dynamics of fecal contamination within this complex river landscape is still lacking. This research seeks to explain how specific features of the landscape (e.g., urban areas, water masses, fluvial lakes, islands systems) may modulate the transport and persistence of *Escherichia coli* downstream. Based on extensive sampling of the St. Lawrence River and recent advances in microbial source tracking, this thesis also aims to identify hotspots of microbial contamination using *Escherichia coli* in the St. Lawrence River and determine downstream persistence and potential sources of bacterial contamination through the system.

Over five expeditions from 2017 to 2022 (excluding 2019), 320 water samples were collected onboard the research vessel *Lampsilis*, complemented by 198 shallow water samples (<1 m) collected from small boats. Advanced molecular technologies, including digital PCR, were employed to analyze the samples for *Escherichia coli* concentration and DNA markers targeting human, ruminant, gull, and pig sources. The study consistently detected *Escherichia coli* in most of the analyzed samples, with 48% of sites exceeding the safe threshold for direct human contact, and 19% for indirect contact. Notable variations in bacterial concentration were observed across different water masses, with particularly high concentrations near urban areas, especially around Montreal's

wastewater outflow. The persistence of *Escherichia coli* concentrations for hundreds of kilometers downstream suggests a potentially significant and sustained impact of this contamination source on the SLR water quality. Temporal analysis revealed a significant influence on *Escherichia coli* levels in 2021 and 2022, although the specific drivers, such as water levels, precipitation, and other environmental factors, clear relationships were not identified, suggesting that unaccounted variables could have impacted the data during these years. The dynamics in nearshore areas suggest a greater influence of local contamination sources, with *Escherichia coli* concentrations varying more significantly year to year, reflecting localized impacts from nearby sources compared to the more consistent levels observed in the water masses. Human signal was identified as the primary sources of fecal contamination, although a few localized instances of bird contaminations were also detected. Although there was a good relationship between *Escherichia coli* concentration and human DNA copies (HF 183), deviations in many sites suggesting the presence of additional, unidentified sources of fecal contamination, presumably from wild animals. These findings underscore the significant impact of human activities on water quality and downstream ecosystems, emphasizing that specialized methods such as microbial source tracking are essential for understanding contamination dynamics in complex river landscapes.

The final chapter synthesize research findings into practical knowledge for ecosystems management and restoration, improved water quality measures and continued scientific research. By examining the St. Lawrence River's ecosystem, this thesis contributes knowledge to the field of aquatic ecosystem management, focusing on the challenges of microbial contamination in large rivers. In addition to increasing our understanding of the role of microbes in the functioning and health of river systems, this thesis will contribute to the development of microbial ecological indicators of water quality that will be included in an ecosystem health report for the Upper St. Lawrence River in collaboration with the St. Lawrence River Institute in Cornwall, Ontario.

Keywords: Anthropogenic activities, aquatic ecosystems, *Escherichia coli* (*E. coli*), fecal contamination, human and animal sources, large rivers, microbial contamination,

microbial source tracking, river ecosystems, St. Lawrence River, wastewater treatment plants, water quality.

Table of Contents

List of Figures and Tables

Figure **Page**

Table **Page**

List of Abbreviations and Acronyms

CFU	*Colony forming unit*
CSO	*Combined sewer overflow*
DNA	*Deoxyribonucleic acid*
dPCR	*Digital polymerase chain reaction*
E. coli	*Escherichia coli*
FC	*Fecal contamination*
FIB	*Fecal indicator bacteria*
LOD	*Limit of detection*
LOQ	*Limit of quantification*
LSLR	*Lower St. Lawrence River*
MST	*Microbial source tracking*
PCR	*Polymerase chain reaction*
RBT	*Risk-based threshold*
RI	*River Institute*
RQM	*Réseau Québec maritime*
SLR	*St. Lawrence River*
UNDP	*United Nations Development Program*
UQTR	*Université du Québec à Trois-Rivières*
UPSLR	*Upper St. Lawrence River*
WHO	*World Health Organization*
WWTP	*Wastewater Treatment Plant*

Chapter 1

Research problem

1. 1 Fecal contamination in large river systems

The St. Lawrence River – Great Lakes system is the second largest waterway in North America, with a length of 3700 km (Marcogliese et al., 2015) and draining over 25% of the world's fresh water supply, ranking thirteenth among the world's largest drainage basins (Barth & Veizer, 2004; Dang et al., 2022). Stretching 1,600 km from Lake Ontario to the Atlantic Ocean (Dang et al., 2022), the St. Lawrence River (SLR) passes through various Indigenous territories of Mohawk, Oneida, Onondaga, Abenaki, Dawnland Confederacy, Atikamekw, Innu, Wendake-Nionwentsio, Nanrantsouak, Maliseet, among others (Native Land Digital, 2024), and shares borders with both the USA and Canada. Its watershed covers an area of over 1 million km^2, with a mean annual discharge of 12,600 m^3 s^{-1} at Quebec City (Barth & Veizer, 2004). The SLR is characterized by three distinct, heterogeneous, water masses flowing side by side downstream (Hudon & Carignan, 2008), along with many complexities including fluvial lakes, islands, and inputs from tributaries such as the Ottawa River (ECCC, 2021).

Despite its societal and ecological importance, the SLR faces growing anthropogenic pressures from urbanization, industrialization, and agriculture. Millions of people in Canada and the United States live near the SLR, with the majority residing in Canada, downstream from Lake Ontario. Population densities decline, especially in United States, where agricultural land becomes more prevalent. While farming is still widespread in Canada, it is declining in the United States as one moving south, allowing for reforestation and an increase in forested land over the last half-century. Quebec has also experienced a decline in agriculture areas, but the remaining lands have seen intensified agricultural activities (Thorp et al., 2005). The construction of the St. Lawrence Seaway System (Vincent & Dodson, 1999), has altered the fluvial section of the river for navigation and hydropower, resulting in changes to hydrological dynamics, water levels,

and currents (Hudon et al., 2017). Approximately forty million inhabitants benefit from its watershed (Marcogliese et al., 2015), engaging in activities such as shipping, power generation, tourism, and recreation, which collectively pose threats to the river's ecological integrity and functioning. In the las four centuries, water pollution has increased and changed in its nature (Thorp et al., 2005). Wastewater from urban and agricultural runoffs represents potential sources of pollution, characterized by increased concentrations of nutrients and pathogens (Ashbolt, 2004; Goswami et al., 2018). This affects downstream communities, leading to environmental degradation, economic losses, and health risks (Staley et al., 2018; Wen et al., 2017a).

Microbial contamination, and fecal contamination (FC) in particular, introduces pathogens into river ecosystems, presenting a crucial challenge to public health and environmental integrity (Corsi et al., 2014; Ichor et al., 2014; Zandaryaa & Mateo-Sagasta, 2018). However, a comprehensive understanding of the origin, dynamics, and downstream impact of this FC within this complex river landscape is still lacking. Some reasons for this lack include limited spatial coverage in monitoring efforts, which do not directly assess the natural dynamics of FC within and across water masses in the SLR. Water management falls under different jurisdictions in Quebec and Ontario, Canada, and New York State, United States, making it challenging to provide a comprehensive picture of FC from Lake Ontario to the Estuary. These efforts overlook riverine units such as fluvial lakes and islands, potentially influencing the transport and retention of *E. coli*. As the water flows through these rivers unites, they function as natural filters. Within these archipelagos, the interactions between hydrology, ecology, and biogeochemical processes are closely linked, effectively retaining a significant fraction of the transported nutrients and contaminants (Bouwman et al., 2013). Shallow water areas, where FC may not adequately disperse, are also not covered. Additionally, there is a lack of an additional step to identify between human and non-human (e.g., pigs, ruminants, birds) sources of FC using microbial source tracking (MST), which has been historically challenging due to past technology limitations. These deficiencies underscored the importance of building a more comprehensive understanding of human impact on the health of river ecosystems.

Addressing water pollution and microbial contamination in the SLR is crucial for safeguarding environmental integrity and public health. By recognizing and mitigating the anthropogenic pressures on the river ecosystem, we can work towards sustainable resource management and ensure the long-term health and viability of this vital natural source.

1. 2 **Research question**

What are the different anthropogenic sources of *E. coli* in the St. Lawrence River, and how do these sources vary in terms of abundance, distribution, and transportation leading to downstream consequences?

1. 3 **Objectives**

1) Identify hotspots of fecal contamination using *E. coli* in the SLR and determine downstream persistence and dynamics across water masses and features of the riverscape.
2) Identify potential sources of fecal contamination through the system using microbial source tracking.

1. 4 **Hypotheses**

Based on existing research on fecal contamination in the SLR, particularly related to anthropogenic sources of *E. coli*, we hypothesize that:

- The concentration of *E. coli* in the mixed water mass decreases as the water flows downstream from the Montreal effluent through riverine units such as islands and fluvial lakes.
- The SLR exhibits significantly higher concentrations of microbial contamination near sewage discharges from human sources compared to inputs from livestock animals at tributaries, despite the typically perceived higher impact of agricultural activities on water quality.

- Nearshore areas of the SLR exhibit higher concentration of *E. coli* and greater variability in water quality parameters due to closer proximity to anthropogenic inputs, compared to central water masses. Inversely, the central river water masses, being further from these sources, are expected to show lower levels of *E. coli* and reduced variability, primarily due to dilution and dispersal processes.

References

Ashbolt, N. J. (2004). Microbial contamination of drinking water and disease outcomes in developing regions. Toxicology, 198(1-3), 229-238.

Barth, J. A. C., & Veizer, J. (2004). Water mixing in a St. Lawrence River embayment to outline potential sources of pollution. Applied Geochemistry, 19(10), 1637–1641.

Bouwman, A. F., Bierkens, M. F. P., Griffioen, J., Hefting, M. M., Middelburg, J. J., Middelkoop, H., & Slomp, C. P. (2013). Nutrient dynamics, transfer and retention along the aquatic continuum from land to ocean: towards integration of ecological and biogeochemical models. Biogeosciences, 10(1), 1-22.

Corsi, S. R., Borchardt, M. A., Spencer, S. K., Hughes, P. E., & Baldwin, A. K. (2014). Human and bovine viruses in the Milwaukee River watershed: Hydrologically relevant representation and relations with environmental variables. Science of the Total Environment, 490, 849–860.

Dang, D. H., Ma, L., Ha, Q. K., & Wang, W. (2022). A multi-tracer approach to disentangle anthropogenic emissions from natural processes in the St. Lawrence River and Estuary. Water Research, 219, 118588.

Environment and Climate Change Canada. (2021). Canadian environmental sustainability indicators : nutrients in the St. Lawrence River.

Goswami, M., Khan, F. A., Ibrisevic, A., Olsson, P. E., & Jass, J. (2018). Development of Escherichia coli-based gene expression profiling of sewage sludge leachates. Journal of applied microbiology, 125(5), 1502-1517.

Hudon, C., & Carignan, R. (2008). Cumulative impacts of hydrology and human activities on water quality in the St. Lawrence River (Lake Saint-Pierre, Quebec, Canada). Canadian Journal of Fisheries and Aquatic Sciences, 65(6), 1165–1180.

Hudon, C., Gagnon, P., Rondeau, M., Hébert, S., Gilbert, D., Hill, B., Patoine, M., & Starr, M. (2017). Hydrological and biological processes modulate carbon, nitrogen and phosphorus flux from the St. Lawrence River to its estuary (Quebec, Canada). Biogeochemistry, 135(3), 251–276.

Ichor, T., Umeh, E. U., & Duru, E. E. (2014). Microbial Contamination of Surface Water Sources in Rural Areas of Guma Local Government Area of Benue State, Nigeria. Journal of Medical Sciences and Public Health, 2(2), 43–51.

Native Land Digital. (2024). Native Land. Retrieved July 11, 2024, from https://native-land.ca/

Marcogliese, D. J., Blaise, C., Cyr, D., de Lafontaine, Y., Fournier, M., Gagné, F., Gagnon, C., & Hudon, C. (2015). Effects of a major municipal effluent on the St. Lawrence River: A case study. Ambio, 44, 257–274.

Staley, C., Pongsilp, N., Nimnoi, P., Kaiser, T., & Sadowsky, M. J. (2018). Influence of physicochemical factors on bacterial communities along the lower Mekong River assessed by Illumina next-generation sequencing. Water, Air, & Soil Pollution, 229, 1-13.

Thorp, J. H; Lamberti, G. A.; & Casper, A. F. (2005). St. Lawrence River Basin. (pp 983-996) Elsevier Inc.

Vincent, W. F., & Dodson, J. J. (1999). The St. Lawrence River, Canada-USA: the need for an ecosystem-level understanding of large rivers. Japanese Journal of Limnology (Rikusuigaku Zasshi), 60(1), 29-50.

Wen, Y., Schoups, G., & Van De Giesen, N. (2017). Organic pollution of rivers: Combined threats of urbanization, livestock farming and global climate change. Scientific reports, 7(1), 43289.

Zandaryaa, S., & Mateo-Sagasta, J. (2018). Organic matter, pathogens and emerging pollutants.

Chapter 2

Literature review

2. 1 **The Complexity of Large Rivers**

Globally, river basins occupy 69% of the land area, transporting around 19 billion tons of material annually (Gupta, 2007). Rivers are products of the dynamics of the Earth, such as geomorphology, plate tectonics, climate, and volcanism, filling of sedimentary basins, and geologic history that determine their location, size, form, orientation, and evolution among others. When defining a large river, a few characteristics need to be considered, including the drainage basin, which collects rain, ice, or snow that falls in the catchment, the length of the channel, and the volume of discharge that commonly carries large amounts of sediments. Potter (1978) considers four properties for defining large rivers: 1. size of the drainage basin, 2. length of the river, 3. volume of sediment transported, and 4. water discharge. Meade (1996) considers water volume and average suspended sediment discharges to the coastal zone, when defining large rivers. Considering the factors listed above, we can say that large rivers are characterized by large basin areas, long main channels, and high discharge of water and sediments. Figure 2.1 and Table 2.1 show the location of large rivers of the world as defined by these characteristics. The sediment load shows the history of their erosion and the dynamics of the system. Each large river system is unique, its development follows different paths, and its history is complex and depends on geological and climatic factors. What it is evident in present-day rivers is only a representation of a limited part of its long evolutionary history (Gupta, 2007).

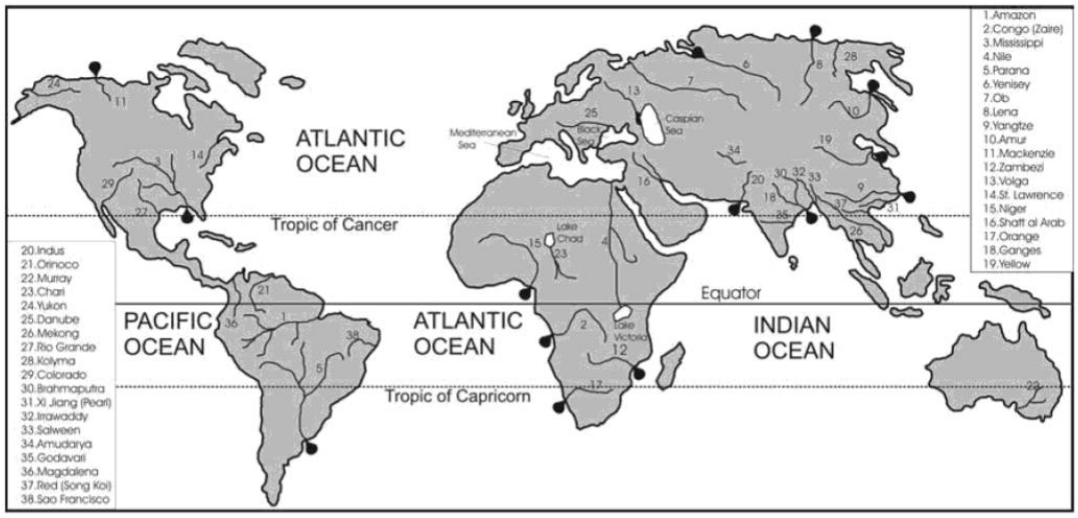


Figure 2.1 The world's primary rivers (From Gupta, 2007).

The headwaters are areas where the water originates, characterized by interactions among hydrologic, geomorphic, and biological processes. These regions are critical for nutrient dynamics and serve as habitats for macroinvertebrates, fish, and amphibians within watersheds. Understanding the spatial and temporal variation in headwaters is essential to comprehend the diversity and heterogeneity of riparian and riverine ecosystem (Gomi et al., 2002). The headwaters of many large rivers originate in mountain ranges and end in the ocean, influenced by changes in sea level and climate change (Gupta, 2007). One unique large river is the St. Lawrence River, in Canada, which does not originate from mountains. Instead, approximately half of its discharge come from the Great Lakes and their many tributaries. The Great Lakes impart specific characteristics to the St. Lawrence River, such as clearer waters and more stable water level compared to other large rivers. This results in the transportation of the smallest amount of suspended sediment (Table 2.1) and hence make the St. Lawrence River less turbid than other large rivers (Thorp et al., 2005).

Table 2.1 List of 15 large rivers and their characteristics (Modified from Gupta, 2007).

River	Annual average water discharges (10^9 m^3)	Length (km)	Drainage basin area (10^6 km^2)	Current average annual suspended sediment discharge (10^6 t)
1) Amazon	6300	6000	5.9	1000-1300
2) Congo	1250	4370	3.75	43
3) Orinoco	1200	770	1.1	150
4) Ganga-Brahmaputra	970	B-2900 G-2525	1.06 (B-0.63)	900-1200
5) Changjiang	900	6300	1.9	480
6) Yenisey	630	5940	2.62	5
7) Mississippi	530	6000	3.22	210
8) Lena	510	4300	2.49	11
9) Mekong	470	4880	0.79	150-170
10) Parana-Uruguay	470	3965	2.6	100
11) St. Lawrence	**450**	**3100**	**1.02**	**3**
12) Irrawaddy	430	2010	0.41	260
13) Ob	400	>5570	2.77	16
14) Amur	325	4060	2.05	52
15) Mackenzie	310	4200	2.00	100

Climate and hydrology are very particular to each large river system, because they drain such a large area with diverse atmospheric patterns, geology, topography, vegetation, land use, and the discharge reflects this mechanism. Precipitation and temperature affect the erosion processes which in turn control the vegetation cover. While large rivers are often located in areas with a large annual rainfall such as in the humid tropics (Figure 2.2), high latitudes rivers that drain North America or Eurasia present perennially baroclinic climates, and the dominant runoff is by snowmelt, resulting in low to moderate annual discharge variability but high monthly discharge variability (Gupta, 2007; Wickert et al., 2016; Wohl et al., 2022).

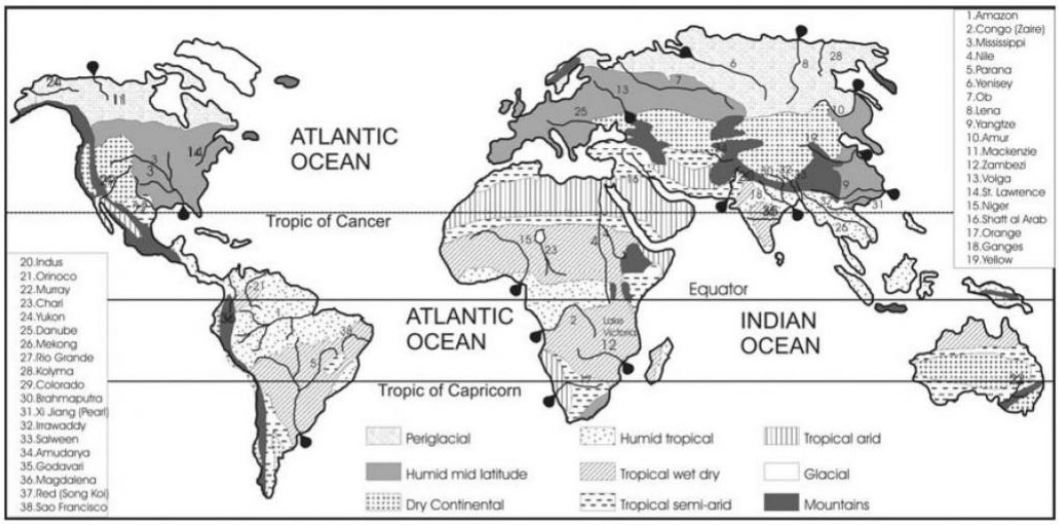


Figure 2.2 The climate settings and variability of large rivers (From Gupta, 2007).

Examples include the Mackenzie River, which originates from the Canadian Rockies and Canada's northern lakes, flowing into the Arctic Ocean, and the Yukon River, with the headwater in the Canadian and Alaskan Cordillera which drains into the Bering Sea, as illustrated in Figure 2.1. The mean annual precipitation across Mackenzie and Yukon basin ranges from 1000 mm in the mountains to less than 250 mm in the lower lands (Gupta, 2007). The maximum discharge is in late spring or early summer for both rivers when the snow is melting. These rivers may also experience important discharge events during summer precipitation. The discharge in winter is generally low due their frozen surface and little surface runoff inputs (Gupta, 2007). These factors may change if it is considered that Canada, like other high latitude regions, is experiencing accelerated and intense warming due to global climate change, altering the distribution of surface water (Stadnyk & Déry, 2021).

Rivers are open lotic systems that have important roles in ecosystems: they are the link between land, lakes, and ocean by the transfer of water and sediments providing a pathway for organic material and nutrients (Savio et al., 2015). As a freshwater system, rivers play a role in animal and plant geography (Potter et al., 2006; Savio et al., 2015), creating habitat for thousands of species, so changes in its geomorphology directly impact

the biota (Gupta, 2007). They also provide essential ecosystem services for humans as a fresh water supply, shipping and transportation, migration corridors for fauna, flooding control, agriculture irrigation, source water for industry, hydropower, inspiration for artists, and recreation among other things (Potter et al., 2006; Savio et al., 2015). As a part of the global biogeochemical cycles, they transform and store terrestrial organic matter (Benstead & Leigh, 2012; Savio et al., 2015) and therefore understanding how the river biota responds to their flow or to changes in the basin can be an important contributor to understanding their cycles (Gupta, 2007).

In general, the global water system has been anthropogenically modified and highly impacted from their natural conditions for different reasons. According to Gupta (2007) migration to, and settlement in, drier regions has led to the building of dams, diverting of water, and the development of irrigation projects, causing more than half of large rivers of the world to be affected. Other modifications to rivers include the creation of channels for flood control and for improving navigation, causing water and land use changes, which have great influence on the dynamics of the channel. The importance of these variables changes over time and evolves in the process. Precipitation and discharge for example, are being affected by climate change, resulting in some instances in lower flow volumes for large rivers, and it was estimated in 1997 that 30% of the global sediment flux is trapped in reservoirs (Gupta, 2007; Walling et al., 2012).

Aquatic ecologists have predominantly directed their attention towards the study of lakes, since large rivers present significant variability and challenges for research, typically, large rivers are approached from an engineering perspective, considering the rivers as water and waste conduits rather than living systems (Vincent & Dodson, 1999). Fortunately, over the past three decades, there has been a shift in the approach to managing rivers and their basins, a growing recognition of rivers as ecological systems, and a desire for environmental management and public participation. However, these shifts present both technical and political challenges. Taking into account future scenarios of climate change and population development growth, the availability and quality of water has become a major concern. Erosion has increased due to deforestation and changes in land

use for crops, and the release of sediment has decreased due to the construction of dams. Increase in human populations along rivers challenges water sanitation, as we have seen that for centuries humans unknowingly contaminated sources of drinking water by dumping raw sewage into rivers or other water bodies, resulting in the appearance of certain diseases and epidemics, affecting large communities (Okello et al., 2019). It is thus necessary that environmental impacts and public health threats in large rivers and basins are incorporated into future monitoring (Friberg et al., 2011).

2.1.1 *Contamination of riverine ecosystems*

Over the course of human civilization, the major role of rivers has been transportation, water supply and waste disposal, but the pressure has been intensified in the last century (Vincent & Dodson, 1999). Surface water is constantly threatened by environmental pollution which is a global problem that affects both industrialized and developing countries in different ways (Ichor et al., 2014). As the human population continues to grow and urbanization and industrialization expand, the sustainable use of ecological resources become an essential matter for the long-term health of the environment (Goswami et al., 2018). Unfortunately, human activities have significantly impacted the water quality of rivers with physical, chemical, and bacteriological variables which alter the environmental conditions for biological communities (Ashbolt, 2004; Yang et al., 1996). In both developed and developing countries, two categories of water pollution correspond to sediments and human and animal waste (Ichor et al., 2014). In large rivers, the main contaminants of water sources are the effluents from wastewater treatment plants (WWTP), agricultural activities, and industrial waste (Čelić et al., 2021). Crop production is the main agricultural activity that pollutes water with nutrients, pesticides, salts, and sediments, while livestock pollutes with organic matter, pathogens, hormones, and antibiotics (Zandaryaa et al., 2018).

Sources of contamination can vary in time and space and can either come from: “point sources” from a single place, easily identified and trackable (i.e., WWTP) or ‘nonpoint sources” from many places, all at once, harder to track (i.e., agricultural or urban runoff) (Corsi et al., 2014; Rock et al., 2015). With respect to the water quality coming

from point sources, certain criteria and standards can be used to manage the concentration of nutrients, oxygen, pathogens, etc., which include limits for discharges. However, nonpoint sources are more complicated since they do not originate from a single discrete source, being the accumulation of small amounts of contaminants collected from a large area. Therefore, diffuse pollution management throughout the basin becomes necessary to limit pollution (Gupta, 2007). Some contamination is due to the lack of protection and natural filtration of the soil, the short distances between where the contamination occurs and the extraction of water and when the rains are intense, the load of pollutants can increase and easily reach the water bodies (Ichor et al., 2014). Human health problems in developing regions are related to unsafe water, poor sanitation, and poor hygiene, where wastewater can be the source of microbial pathogens (Ashbolt, 2004).

2.1.2 *Microbial contamination of riverine ecosystems*

Municipal and industrial waste, commercial activities and run-off from agricultural areas are the major causes of water contamination releasing pathogenic microorganisms into the environment (Goswami et al., 2018). It is difficult to understand and quantify the loads, transport, and fate of pathogens in the environment, due to their spatial and temporal variability (Zandaryaa et al., 2018). As human activity and climate change continue to escalate, the vulnerability of freshwater to pollutants will increase from different sources, such as untreated sewage overflows, runoff of animal excreta from farms, and alga blooms triggered by excessive nutrient loads (WHO, 2021).

Water contamination by pathogens has ecological and toxicological impacts on aquatic organisms, causing health risks, environmental degradation, and economic losses (Zandaryaa & Mateo-Sagasta, 2018). These pathogens are recognized as a potential hazard to human health because of their presence in water recreational areas, in drinking water systems, or in crops contaminated by irrigation (Corsi et al., 2014; Korajkic et al., 2018). While foodborne and airborne transmission play crucial roles in disease outbreaks, waterborne transmission remains a significant and widespread mode of pathogen dissemination, particularly for enteric diseases, although the estimates may not be entirely

well documented, since some of the illnesses go unnoticed, being mild and lasting a couple of days, and not requiring medical treatment (Ichor et al., 2014).

Some of the infectious diseases that can be acquired by contaminated water are viral hepatitis, polio, typhoid and paratyphoid fever, amoebic and bacillary dysentery, botulism, cholera, schistosomiasis, salmonellosis, primary amoebic meningoencephalitis, and giardiasis. They usually come from the feces and urine of infected people or livestock but can also be present in the environment from natural sources such as wildlife (Ichor et al., 2014). There are, however, some efforts to try to remedy the entry of these contaminants into water sources. According to the United Nations Development Program (UNDP), 80% of the global human wastewater goes into waterways without adequate treatment, and that is why they included "Clean water and sanitation", as a sustainable goal, to create strategies for better management (UNDP, 2021). Like the UNDP, there are other international and national organisations that have taken this parameter as a guideline on fresh water for recreational water quality, which aims to protect the public health; the United States Environmental Protection Agency (EPA) based in the Clean Water Act, offers support in creating and executing programs for water pollution control (EPA, 2021); and health Canada create a the "Guidelines for Canadian Recreational Water Quality" to assist authorities responsible for managing recreational waters (Health Canada, 2012). Table 2.2 presents the recommended water quality parameters for *Escherichia coli* (*E. coli*) concentration, measured in colony forming units (CFU) per 100 mL, with the limits for both direct/primary contact and indirect/secondary contact in recreational freshwater setting, across two distinct organizations.

Table 2.2 Guidelines for *E. coli* concentration, recreational water quality: summary table (Modified from EPA, 2021 and Health Canada, 2012).

Organizations	**Consideration**	**Direct/primary contact**	**Indirect/secondary contact**	**Reference**
U.S.EPA	Geometric mean concentration	126 CFU/100 mL		EPA, 2021
	Statistical threshold value	410 CFU/100 mL		
Health Canada	Geometric mean concentration (minimum 5 samples)	200 CFU/100 mL	1000 CFU/100 mL	Health Canada, 2012
	Single sample maximum concentration	400 CFU/100 mL		

Disease-causing pathogens from fecal origin transmitted by drinking water, are known as enteric pathogens (Ashbolt, 2004). Fecal contamination (FC) is the primary contributor to the introduction of these enteric pathogens into the water, along with the introduction of organic matter and nutrients, such as nitrogen, and phosphorus, and other contaminants (Rock et al., 2015). The fecal indicator bacteria (FIB) are used to determine the presence of FC in the water and its occurrence over time ("If" and "when") (Rock et al., 2015). This monitoring approach, used for more than a century in water quality monitoring programs, has been selected for its low pathogenic potential, high abundance in sewage and feces, and its correlation with other pathogens (bacteria and viruses) (Harwood et al., 2013). The most used FIB are fecal coliforms, *E. coli*, and enterococci (Harwood et al., 2013), however, once is introduced into the aquatic environment, the association between this FIB and the pathogens may alter due to various factors, such as dilution, water flow, and pathogen survival in the environment (Vijayan et al., 2023).

Escherichia coli and enterococci are indicators of potential human health risk, due to their high concentration in mammals' feces (Vijayan et al., 2023).Additionally waterflow can contributes to the spread of zoonotic diseases, such as avian influenza virus, with outbreaks often linked to wild birds (McDuie et al., 2022). These factors play a crucial role in assessing the effectiveness of drinking water treatment. The elimination of pathogenic bacteria responsible for cholera and typhoid fevers, can be evaluated using the common FIB, *E. coli* (Ashbolt, 2004). *E. coli* is a bacterium that grows naturally in the lower intestines of mammals and birds, it is an indicator for detecting FC and is used to find hotspots of organic pollution in aquatic ecosystems and their persistence downstream. While many strains of *E. coli* are harmless, some strains can cause infections and illness in humans and animals, leading to symptoms like vomiting and diarrhea (Ashbolt, 2004; EPA, 2021; Tallon et al., 2005). There is a diverse range of distinct *E. coli* ecotypes, each one occupying specific ecological niche by colonizing different host species (Yu et al., 2021).

Agriculture poses a significant challenge to river systems by introducing *E. coli* strains from different hosts, each of which carry ecotypes adapted to their specific gut

environments and highlight the risk of animal waste contaminating water bodies which can cause disease outbreaks (Weller et al., 2022; Yu et al., 2021). A particularly important pathogenic strain for human is *E. coli* O157:H7, which causes outbreaks worldwide (Paula et al., 2014). This strain is regularly found in cattle feces and can be transmitted through contaminated food, water, and direct contact with infected people or animal (Mead et al., 1998). Moreover, agricultural runoff can carry antibiotic-resistant strains of *E. coli*, a concerning development stemming from the use of antibiotics in livestock. These strains can transfer their resistance genes to other bacteria within aquatic ecosystems, increasing public health risks (Yu et al., 2021). Additionally, the high levels of nutrients, including nitrogen and phosphorous from fertilizers found in agricultural runoff, contribute to eutrophication (Browning et al., 2023). These conditions favour the evolution of *E. coli* strains capable of thriving in nutrient-rich environments. Such strains can become "naturalized" to soil, sand, sediments, and algae, potentially integrating into the microbial landscape and altering the natural dynamics of these ecosystems (Ishii et al., 2008).

Urban areas introduce a complex mix of *E. coli* strains into river systems, predominantly sourced from human activities (Walker et al., 2015). These strains often exhibit unique virulence characteristics not commonly found in those associated with agricultural runoff, reflecting the diverse microbial inputs from sewage overflows, improperly managed waste, urban runoff, and urban wildlife (Petersen et al., 2020). Alongside biological contaminants, urban runoff carries a variety of chemical pollutants, such as suspended sediments, metals, polycyclic aromatic hydrocarbons (PAHs), phthalates, alkylphenols (Ps), bisphenol-A (BPA) and pesticides, which can selectively influence the survival and proliferation of *E. coli* ecotypes with higher resilience to toxic environments (Müiller et al, 2020; Ranjan et al., 2022). Additionally, the phenomenon of thermal pollution, primarily driven by the heat-absorbing infrastructure of densely populated areas, can further complicate the microbial dynamics in water bodies. This leads to warmer waters that may favour *E. coli* variants with increased tolerance to temperature shifts, potentially exacerbating the risk of FC (Petersen et al., 2020). Collectively, these factors underscore the critical need for integrated urban water management strategies that

address the multifaceted challenges posed by urban runoff, aiming to safeguard public health and preserve aquatic ecosystems.

Understanding the diverse ecotypes of *E. coli* present in contaminated water bodies and their sources (agricultural vs. urban) is critical for designing effective water management and treatment strategies (Qi et al., 2021). Such strategies may include targeted measures to reduce agricultural runoff through improved waste management practices on farms or to mitigate urban runoff through green infrastructure and better sewage treatment. Additionally, tracking the diversity and abundance of *E. coli* ecotypes can also provide insights into the effectiveness of these strategies and help in assessing the health risks associated with waterborne pathogens in different environmental contexts.

Fecal contamination in waterways can originate from various sources, including treated and partially treated wastewater effluents, combined sewer overflows (CSO), sanitary sewer overflows (SSO), leaks in sanitary and sewer lines, inadequately connected sanitary systems, septic systems, and agricultural runoff from farms and livestock. The most common domestic wastewater treatment systems in urban areas with dense populations are centralized WWTP (Drury et al., 2013). These plants are crucial as untreated municipal wastewater often has high concentration of E. *coli* (Elahi et al., 2017) Normally, treated wastewater, along with discharges from CSO and SSO, are released into surface waters. Agricultural contributions to FC occur through the dispersal of manure in water bodies, pastures, or its use as a fertilizer, from which pathogens can be transported across land and into surface waters (Corsi et al., 2014; Rock et al., 2015). Given the incalculable human and animal sources of FC (Drury et al., 2013), the concentration of fecal indicators such as *E. coli* not only serves as a tracer for wastewater effluents but can also be associated with nutrient levels, water clarity, and changes in ammonium (NH_4), total phosphorus (TP), silicon dioxide (SiO_2), and suspended solids concentrations (Vis et al., 1998).

In complex rivers systems, it becomes necessary to employ genetic-based microbial source tracking (MST) techniques to analyze inputs and loads from different sources, since each source contributes different contaminants, including unique strains of

bacteria (Harwood et al., 2013). These techniques act as an integrative tool that helps understand the changes within hydrographic basin dynamics. By analyzing the sources, this technique provides a comprehensive picture of how contaminants move and change within the system, and facilitates the development of water quality criteria, as well as assisting in management and remediation planning (Behrendt et al., 2002). Despite the use of FIB to assess human health risk, there are some assumptions, as numerous studies have shown weak correlations between FIB concentrations and the presence of pathogens, especially when contamination originates from unknown sources (Harwood et al., 2013). This emphasizes the importance of employing MST methods in understanding and managing potential health risk associated with FC in river ecosystems.

Microbial Source tracking

Microbial source tracking (MST), which emerged at the end of the 20th century, is a toolbox of techniques that identify the sources of FC ("Who" is contributing?) by associating specific fecal microorganisms with particular hosts in environmental waters, these host-associated microorganism can be used as a signature molecule (marker) such as DNA sequences, with the main intention to differentiate between human and non-human sources of FC, while also being able to identify different species of animals (Harwood et al., 2013; Rock et al., 2015).

There are generally two types of methods for MST: 'library-dependent' and 'library-independent'. The library-dependent method is based on isolate-by-isolate identification of cultured bacteria (e.g., *E. coli*) from different fecal sources and water samples, comparing water isolates to a "library" of bacterial strains from known fecal sources. This requires the creation of biochemical (phenotypic) or molecular (genotypic) fingerprints for bacterial strains to compare the similarity of strains isolated from water samples and suspected fecal sources. This method is more expensive and time consuming, and the libraries are temporally and geographically specific, so they are not applied at larger spatial (large watershed) and temporal scales. The library-independent methods are based on the detection of a specific host-associated genetic markers or gene targets identified in the water sample (e.g., 16S rRNA gene of *Bacteroidales*); these methods can

identify fecal sources based on a specific host (genetic marker) of the bacteria without needing a "library" (Rock et al., 2015). *Bacteroides* have several advantageous characteristics that make them widely used in MST. Their high abundance in human and animal feces, minimal potential for environmental growth, and high degree of host specificity make *Bacteroides* a reliable alternative fecal indicator to *E. coli*, as they constitute a significant portion of the fecal bacterial population (Rock et al., 2015). To trace them, some library-independent methods utilize polymerase chain reaction (PCR). This approach reveals source information by targeting and amplifying specific genetic markers unique to microorganisms from different hosts, such as humans or animals. By detecting these markers in the DNA or RNA (Ribonucleic acid) extracted from a water sample, PCR facilitates the identification of FC sources, distinguishing between human and animal origins (Rock et al., 2015). The most common DNA markers used to identify the fecal sources are for humans, pigs, dogs, cows, and gulls (Harwood et al., 2014). An example of this technique is a study conducted over 3 years in the Dargle River, which flows into the Irish Sea in Bray, Dublin. There, 354 samples were collected at 10 sampling sites to track 12 MST markers to distinguish between sources of fecal pollution, where human and ruminant where the most abundant (Ballesté et al., 2020).

The risk to human health from animal FC is often underestimated and under-addressed in the literature, although past events have raised awareness in many localities (Ali, 2004). It is generally perceived as less severe than contamination from human sources, posing challenges for authorities responsible for maintaining water quality for recreational activities and drinking water purposes (Dufour et al., 2012). Yet, zoonotic waterborne infections from domestic or agricultural animal feces still pose a definite health risk, and as such MST markers can provide useful information regarding the links between land use and health risks (Rock et al., 2015). Wildlife plays an important role in the transmission of pathogens, with potentially higher risk, particularly in areas shared between humans and wildlife. A notable example of this phenomenon is evident in urban environments, where gulls emerge as potential sources of pollution. This is attributed to their opportunistic feeding behaviour, as they are often attracted to easily accessible food sources associated with human activities, such as landfills, water treatment plants,

agricultural areas, and livestock areas. In these environments, gulls can ingest microorganisms of human fecal origin. Consequently, they serve as transport vectors for human pathogens, influencing surface waters that are used for purposes such as irrigation, drinking, and recreational activities (Alm et al., 2018; Martín-Vélez et al., 2023). Recognizing the sources and the origin of FC is essential for a consistent evaluation of human health risks (microbial diseases), ensuring safe water access, developing targeted remediation plans (Rock et al., 2015). It may also be a useful tool to improve water management strategies within river catchments (Ballesté et al., 2020).

The limit of detection (LOD) represents the minimum amount reliably detectable during the analytical step, often quantified in gene copies by PCR. On the other hand, the limit of quantification (LOQ) denotes the minimum number of gene copies or the smallest amount of fecal matter that can be measured, results can be expressed as gene copies per ng of DNA, gene copies in 100 mL of water (common units in environmental microbiology) (Harwood et al., 2013). Derived from recent research by A.B. Boehm and J.A. Soller, a risk-based threshold (RBT) for a human-associated marker (HF183, associated to the bacterial genus *Bacteroides*) is set at 525 DNA Copies per 100 mL, considered representative of conditions in line with the Recreational Water Quality Criteria, which corresponds to a 32 illnesses per 100 mL (Boehm et al., 2020; EPA et al., 2004).

References

Ali, S. H. (2004). A socio-ecological autopsy of the *E. coli* O157: H7 outbreak in Walkerton, Ontario, Canada. Social science & medicine, 58(12), 2601-2612.

Alm, E. W., Daniels-Witt, Q. R., Learman, D. R., Ryu, H., Jordan, D. W., Gehring, T. M., & Santo Domingo, J. (2018). Potential for gulls to transport bacteria from human waste sites to beaches. Science of the Total Environment, 615, 123–130.

Ashbolt, N. J. (2004). Microbial contamination of drinking water and disease outcomes in developing regions. Toxicology, 198(1-3), 229-238.

Ballesté, E., Demeter, K., Masterson, B., Timoneda, N., Sala-Comorera, L., & Meijer, W. G. (2020). Implementation and integration of microbial source tracking in a river watershed monitoring plan. Science of the Total Environment, 736, 139573.

Behrendt, H., Kornmilch, M., Opitz, D., Schmoll, O., & Scholz, G. (2002). Estimation of the nutrient inputs into river systems–experiences from German rivers. Regional Environmental Change, 3, 107-117.

Benstead, J. P., & Leigh, D. S. (2012). An expanded role for river networks. Nature Geoscience, 5(10), 678-679.

Boehm, A. B., & Soller, J. A. (2020). Refined ambient water quality thresholds for human-associated fecal indicator HF183 for recreational waters with and without co-occurring gull fecal contamination. Microbial Risk Analysis, 16, 100139.

Browning, D. A., Mausbach, W. E., Stookey, C., Nikolai, S. J., Barrow, J., & Townsend, D. E. (2023). Validating microbial source tracking markers and assessing the efficacy of culturable *E. coli* and enterococcus assays in Ozark streams, USA. Water, Air, & Soil Pollution, 234(6), 348.

Čelić, M., Jaén-Gil, A., Briceño-Guevara, S., Rodríguez-Mozaz, S., Gros, M., & Petrović, M. (2021). Extended suspect screening to identify contaminants of emerging concern in riverine and coastal ecosystems and assessment of environmental risks. Journal of Hazardous Materials, 404, 124102.

Corsi, S. R., Borchardt, M. A., Spencer, S. K., Hughes, P. E., & Baldwin, A. K. (2014). Human and bovine viruses in the Milwaukee River watershed: Hydrologically relevant representation and relations with environmental variables. Science of the Total Environment, 490, 849–860.

Drury, B., Rosi-Marshall, E., & Kelly, J. J. (2013). Wastewater treatment effluent reduces the abundance and diversity of benthic bacterial communities in

urban and suburban rivers. Applied and Environmental Microbiology, 79(6), 1897–1905.

Dufour, A., & Bartram, J. (Eds.). (2012). Animal waste, water quality and human health. IWA publishing.

Elahi, E., Abid, M., Zhang, L., & Alugongo, G. M. (2017). The use of wastewater in livestock production and its socioeconomic and welfare implications. Environmental Science and Pollution Research, 24, 17255-17266.

EPA, U. S. Environmental Protection Agency, (2021). *E. coli* (*Escherichia coli*), Factsheet on water quality parameters, EPA 841F21007F.

Friberg, N., Bonada, N., Bradley, D. C., Dunbar, M. J., Edwards, F. K., Grey, J., ... & Woodward, G. U. Y. (2011). Biomonitoring of human impacts in freshwater ecosystems: the good, the bad and the ugly. In Advances in ecological research . Academic press, 44, 1-68.

Goswami, M., Khan, F. A., Ibrisevic, A., Olsson, P. E., & Jass, J. (2018). Development of Escherichia coli-based gene expression profiling of sewage sludge leachates. Journal of applied microbiology, 125(5), 1502-1517.

Gupta, A. (Ed.). (2007). Large rivers: geomorphology and management. John Wiley & Sons.

Gomi, T., Sidle, R. C., & Richardson, J. S. (2002). Understanding processes and downstream linkages of headwater systems: headwaters differ from downstream reaches by their close coupling to hillslope processes, more temporal and spatial variation, and their need for different means of protection from land use. BioScience, 52(10), 905-916.

Harwood, V. J., Staley, C., Badgley, B. D., Borges, K., & Korajkic, A. (2014). Microbial source tracking markers for detection of fecal contamination in environmental waters: Relationships between pathogens and human health outcomes. In FEMS Microbiology Reviews, 38 (1), 1–40.

Health Canada. (2012). Guidelines for Canadian recreational water quality, Third Edition. Water, air and Climate Change Bureau, Healthy Environments and Consumer Safety Branch, Health Canada, Ottawa, Ontario.

Ichor, T., Umeh, E. U., & Duru, E. E. (2014). Microbial Contamination of Surface Water Sources in Rural Areas of Guma Local Government Area of Benue State, Nigeria. Journal of Medical Sciences and Public Health, 2(2), 43–51.

Ishii, S., & Sadowsky, M. J. (2008). *Escherichia coli* in the environment: implications for water quality and human health. Microbes and environments, 23(2), 101-108.

Korajkic, A., McMinn, B. R., & Harwood, V. J. (2018). Relationships between microbial indicators and pathogens in recreational water settings.

International Journal of Environmental Research and Public Health, 15(12), 2842.

Martín-Vélez, V., Navarro, J., Figuerola, J., Aymí, R., Sabaté, S., Planell, R., Vila, J., & Montalvo, T. (2023). A spatial analysis of urban gulls contribution to the potential spread of zoonotic and antibiotic-resistant bacteria. Science of the Total Environment, 912.

McDuie, F., Matchett, E. L., Prosser, D. J., Takekawa, J. Y., Pitesky, M. E., Lorenz, A. A., ... & Casazza, M. L. (2022). Pathways for avian influenza virus spread: GPS reveals wild waterfowl in commercial livestock facilities and connectivity with the natural wetland landscape. Transboundary and emerging diseases, 69(5), 2898-2912.

Müller, A., Österlund, H., Marsalek, J., & Viklander, M. (2020). The pollution conveyed by urban runoff: A review of sources. Science of the Total Environment, 709, 136125.

Okello, P. E., Bulage, L., Riolexus, A. A., Kadobera, D., Kwesiga, B., Kajumbula, H., Mulongo, M., Namboozo, E. J., Pimundu, G., Ssewanyana, I., Kiyaga, C., Aisu, S., & Zhu, B. P. (2019). A cholera outbreak caused by drinking contaminated river water, Bulambuli District, Eastern Uganda, March 2016. BMC Infectious Diseases, 19(1), 1-8.

Petersen, F., & Hubbart, J. A. (2020). Physical factors impacting the survival and occurrence of *Escherichia coli* in secondary habitats. In Water, 12, (6), 1796.

Potter, P. E., Hamblin, K., & Hamblin, W. K. (2006) Big Rivers Worldwide Part 1 Origins. (Vol. 48). Department of Geology, Brigham Young University.

Qi, R., Xue, N., Zhou, X., Zhao, L., & Yang, Y. (2021). Distinct Composition and Assembly Processes of Bacterial Communities in a River From The Arid Area: Ecotypes or Habitat Types?. Microbial ecology, 1-11.

Ranjan, N., Singh, P. K., & Maurya, N. S. (2022). Pharmaceuticals in water as emerging pollutants for river health: A critical review under Indian conditions. In Ecotoxicology and Environmental Safety, 247 114220.

Rock, C., Rivera, B., & Gerba, C. P. (2015). Microbial Source Tracking. In Environmental Microbiology: Third Edition (pp. 309–317). Elsevier Inc. Academic Press.

Savio, D., Sinclair, L., Ijaz, U. Z., Parajka, J., Reischer, G. H., Stadler, P., Blaschke, A. P., Blöschl, G., Mach, R. L., Kirschner, A. K. T., Farnleitner, A. H., & Eiler, A. (2015). Bacterial diversity along a 2600km river continuum. Environmental Microbiology, 17(12), 4994–5007.

Stadnyk, T. A., & Déry, S. J. (2021). Canadian continental-scale hydrology under a changing climate: A review. Water (Switzerland), 13(7), 906.

Tallon, P., Magajna, B., Lofranco, C., & Kam, T. L. (2005). Microbial indicators of faecal contamination in water: A current perspective. In Water, Air, and Soil Pollution, 166 (1–4), 139–166.

Thorp, J. H; Lamberti, G. A.; & Casper, A. F. (2005). St. Lawrence River Basin. Elsevier Inc. Pag 983-996.

United Nations Department of Economic. (2021). World social report 2021: reconsidering rural development. United Nations.

Vijayan, J., Nathan, V. K., Ammini, P., & Ammanamveetil, A. M. H. (2023). Bacterial diversity in the aquatic system in India based on metagenome analysis—a critical review. In Environmental Science and Pollution Research, 30(11), 28383–28406. Springer Science and Business Media Deutschland GmbH.

Vincent, W. F., & Dodson, J. J. (1999). The St. Lawrence River, Canada-USA: the need for an ecosystem-level understanding of large rivers. Japanese Journal of Limnology (Rikusuigaku Zasshi), 60(1), 29-50.

Vis, C., Hudon, C., Cattaneo, A., & Pinel-Alloul, B. (1998). Periphyton as an indicator of water quality in the St Lawrence River (Quebec, Canada). Environmental pollution, 101(1), 13-24.

Walker, J. W., Van Duivenboden, R., & Neale, M. W. (2015). A tiered approach for the identification of faecal pollution sources on an Auckland urban beach. New Zealand Journal of Marine and Freshwater Research, 49(3), 333–345.

Walling, D. E. (2012). The role of dams in the global sediment budget. IAHS-AISH publication, 356, 3-11.

Weller, D. L., Murphy, C. M., Johnson, S., Green, H., Michalenko, E. M., Love, T. M. T., & Strawn, L. K. (2022). Land Use, Weather, and Water Quality Factors Associated With Fecal Contamination of Northeastern Streams That Span an Urban-Rural Gradient. Frontiers in Water, 3, 741676.

Wickert, A. D. (2016). Reconstruction of North American drainage basins and river discharge since the Last Glacial Maximum. Earth Surface Dynamics, 4(4), 831-869.

Wohl, E., & Lininger, K. B. (2022). Hydrology and discharge. Large Rivers: Geomorphology and Management, Second Edition, 42-75.

World Health Organization. (2021). Guidelines on recreational water quality. Volume 1, Coastal and fresh waters.

Yang, C., Telmer, K., & Veizer, J. (1996). Chemical dynamics of the “St. Lawrence” riverine system: δDH2O, δ18OH2O, δ13CDIC, δ34Ssulfate, and dissolved 87Sr/86Sr. Geochimica et Cosmochimica Acta, 60(5), 851-866.

Yu, D., Banting, G., & Neumann, N. F. (2021). A review of the taxonomy, genetics, and biology of the genus Escherichia and the type species Escherichia coli. Canadian Journal of Microbiology, 67(8), 553–571.

Zandaryaa, S., & Mateo-Sagasta, J. (2018). Chapter 8. Organic matter, pathogens and emerging pollutants.

Chapter 3

ASSESING ESCHERICHIA COLI CONCENTRATION AND MICROBIAL SOURCE TRACKING IN A LARGE RIVER SYSTEM: ST. LAWRENCE RIVER

Manuscript in preparation for Journal of Water Research.

Silvia Rodriguez[a, b, c, *], Leigh McGaughey[c], Jérôme Comte[d], Thomas A. Edge[e], François Guillemette[a, b]

[a] Département des sciences de l'environnement, UQTR, Université du Québec à Trois-Rivières, Trois-Rivières, Quebec, H2X 3X8, Canada
[b] Centre for Research on Watershed-Aquatic Ecosystem Interactions (RIVE), Université du Québec à Trois-Rivières, Trois-Rivières, Quebec, Canada
[c] St. Lawrence River Institute, Cornwall, ON
[d] Centre-Eau Terre Environnement, INRS
[e] Department of Biology, McMaster University, 1280 Main Street W., Hamilton, ON L8S 4K1, Canada

*Corresponding author: Silvia Rodriguez

E-mail: Silvia.Rodriguez@uqtr.ca

3. 1 Author's Contributions

Silvia Rodriguez: Contributed to study design, carried on fieldwork and laboratory analysis, conducted statistical analysis of the data, generated figures, and wrote the first draft and edited the article.

Leigh McGaughey: Provide support in project development, secured funding, and edited the paper.

Jérôme Comte: Provide laboratory support, guided the development of the document, and participated in its editing.

Thomas A. Edge: Provide valuable expertise and knowledge, facilitated access to the laboratory for source tracking analyses, guided the development of the research topic, and contributed to editing the paper.

François Guillemette: Designed the study, secured funding, provided guidance and direction throughout the development of the research theme, and contributed to editing the document.

3. 2 Full article in English: Assessing *Escherichia coli* concentration and microbial source tracking in a large river system: St. Lawrence River

Graphical Abstract

Abstract

Despite historical monitoring of fecal contamination by governmental authorities, a comprehensive understanding of the origin and dynamics of fecal contamination within complex river landscapes are still lacking. Here, we investigate the spatial and temporal dynamics of *Escherichia coli* (*E. coli)* contamination in the St. Lawrence River, a significant waterway extending from Lake Ontario to the Atlantic Ocean. Over five summers, water samples were collected from different water masses onboard the research vessel *Lampsilis*, complemented by nearshore sample collection. Through microbial source tracking, we identified the primary sources of *E. coli* using specific markers for human (HF 183), gull (Gull4), pig (Pig2Bac), and ruminant (Rum2Bac) origins. Our findings reveal significant spatial variability in *E. coli* levels, with the highest concentrations near wastewater outflow, underscoring substantial human influence. Surprisingly, concentrations were higher at the river's center compared to the nearshore, suggesting that river inputs from large cities play a crucial role in dispersing pollutants often > 100 km downstream. Areas such as fluvial lakes and islands, showed lower *E. coli* concentrations, indicating natural mitigation effects. Human sources were the predominant contributors to the observed contamination, although occasional detection of gull markers and discrepancies in *E. coli* levels and DNA copies suggest potential untracked wild animal sources.

This study underscores the critical need for identifying specific pollution sources and their dispersion within complex river system to develop effective management and restoration strategies in large rivers.

Key Words: Anthropogenic activities, aquatic ecosystems, *Escherichia coli* (*E. coli*), fecal contamination, human and animal sources, large rivers, microbial contamination, microbial source tracking, river ecosystems, St. Lawrence River, wastewater treatment plants, water quality.

Introduction

Water pollution, encompassing physical, chemical, and biological contaminants, represents a widespread issue affecting both developed and developing countries, with animal and human waste being the primary contributors (Ashbolt, 2004; Ichor et al., 2014). Such contamination poses a significant risk to aquatic life and human health, necessitating rigorous identification and management strategies (Corsi et al., 2014; Zandaryaa et al., 2018). The sources of water contaminations, which have temporal and spatial variations, can be classified in two categories: “point sources”, originating from a single location, easily identified and trackable, and “nonpoint sources”, originating from multiple locations and it is harder to track (Corsi et al., 2014; Rock et al., 2015). Enteric pathogens, including bacteria and virus among others, often originating from fecal contamination (FC) and can cause infections and diseases when they enter the body through the gastrointestinal tract. This represents a major public health concern, highlighting the urgency for precise identification and effective mitigation of such pollutants to ensure water safety for human use (Rock et al., 2015).

Escherichia coli (E. coli), a commonly used fecal indicator bacterium (FIB), naturally found in the lower intestines of mammals and birds, serves as essential tool in assessing water quality due to its association with FC, its relationship with other pathogens and potential health risk. It is also utilized to identify hotspots of organic pollution and trace its persistence downstream, which may be altered due to various factors such as dilution, water flow, and pathogen survival in the environment (Ashbolt, 2004; Elahi et al., 2017; Vijayan et al., 2023). While many strains of *E. coli* are harmless, some strains can cause infections and illness in humans and animals, leading to symptoms like vomiting and diarrhea (Ashbolt, 2004; EPA, 2021; Tallon et al., 2005). Monitoring *E. coli* concentration in large river systems is important for public health across various aspects such as drinking water, recreational activities, and shellfish harvesting (Rock et al., 2015).

Addressing the challenges of FC identification in large river systems necessitates advanced techniques such as microbial source tracking (MST), which offer insight into pollution sources (Harwood et al., 2013; Rock et al., 2015). Traditionally MST studies

have been conducted by quantitative Polymerase Chain Reaction (qPCR) (Harwood et al., 2013). More recently, digital Polymerase Chain Reaction (dPCR), a library independent method, emerges as a promising tool for identifying specific host-associated genetic marker associated with FC (Rock et al., 2015). An example of this technique is a three-year study in the Dargle River, flowing into the Irish Sea in Bray, Dublin, where 354 samples were collected at 10 sites to track 12 markers distinguishing between sources of fecal pollution, with human and ruminant sources being the most abundant (Ballesté et al., 2020). Another example took place in Georgia, USA, where 114 samples were analyzed during a two-year study with three markers, revealing that dogs significantly contributed to the contamination, being the main source of FC (McKee et. al., 2020). Previous MST studies have aimed to identify the sources of diverse FC, but it has rarely been used in the context of large river systems with multiple potential sources and complex flow dynamics. Understanding the sources of FC is beneficial for assessing effective water quality management and remediation strategies (Behrendt et al., 2002; Rock et al., 2015). While zoonotic waterborne infections pose risks to human health, their severity is generally lower compared to those originating from human sources (Harwood et al., 2014). Due to their opportunistic feeding behavior, gulls serve as a vector for human pathogens, highlighting the influence of surface water quality between human activities and wildlife (Alm et al., 2018; Martín-Vélez et al., 2023).

Rivers constitute a vital component of global ecosystems, covering 69% of the Earth's land area, and have diverse characteristics which have been shaped by geological and climatic factors (Gupta, 2007). These open lotic systems not only serve as pathways for organic matter and nutrients but also provide habitat for many species and deliver essential ecosystem services for many regions, which unfortunately this also introduces various stressors on the aquatic environment (Gupta, 2007; Potter et al., 2006; Savio et al., 2015). Given the importance of these systems, the escalating challenges of climate change and population growth have intensified concerns about water availability and quality (Okello et al., 2019).

The St. Lawrence River (SLR), like other large rivers, faces ecological pressures from human activities including shipping, power generation, tourism, recreation, and the provision of drinking water and sanitation, alongside urban development (Hudon et al., 2017). Its watershed faces significant challenges, notably water pollution from both urban and agricultural sources. Urban inputs include domestic sewage, industrial discharges, and stormwater runoff, while agricultural lands contribute through wastewater, runoff, and flood zones such as Lake St. Pierre. Approximately 5 million residents discharge wastewater into the SLR, with varying treatment levels. The Montreal Metropolitan Community makes the largest contribution, releasing about 3.5 million cubic meters daily (41 m^3s^{-1}), (Barth et al., 2004; Marcogliese et al., 2015). However, research on microbial communities and their responses to pollutants within the SLR remains limited, preventing comprehensive understanding and effective management strategies (Maranger et al., 2005).

This study investigates the introduction, persistence, and origins of FC in a large and complex river, the SLR. Over five expeditions from 2017 to 2022 (excluding 2019), 320 water samples were collected onboard the research vessel *Lampsilis*, complemented by 198 shallow water samples (<1 m) collected from small boats to identify hotspots of microbial contamination, focusing on *E. coli* across the riverscape, and its persistence downstream. We then used advanced molecular technologies, including digital PCR, to identify potential sources of bacterial contamination in the SLR and its tributaries. Our analysis considered the variability of nearby land use, such as urban areas and agricultural regions. We utilized markers to trace contamination back to specific sources, including human, pig, ruminant, and gull origins. Our findings offer insights into water contamination sources and how specific features of complex river landscape may affect their persistence downstream in large river system.

Methods

Study site

The study was conducted in the SLR, a major waterway that stretches for 1,600 km along U.S.-Canada border (Dang et al., 2022), connecting Lake Ontario to the Atlantic Ocean, going through different First Nation territories (Figure 3.1). The SLR spanning 550 km from Kingston, Ontario to Quebec City, Quebec, is geographically divided into two main sections: the Upper St. Lawrence River (UPSLR) stretch from Kingston, Ontario to Valleyfield, Quebec, sharing its southern shoreline with the United States of America and Akwesasne. The lower section, known as the Lower St. Lawrence River (LSLR) extends from Montreal to Quebec City, Quebec. Prior to reaching Quebec City, the LSLR presents longitudinal stratification, resulting in three distinct water masses: the main channel originating from Lake Ontario (contributing 60% of total discharge), the northern water mass fed by its major tributary (Ottawa River contributing 16%), and the mixed water mass or Montreal effluent, near Île Aux-Vaches (contributing 0.3%) (Yang et al., 1996). Due to their different physicochemical characteristics, density, and water velocity, these water masses remain on their own course without significant mixing for more than > 100 km. Along its course, the SLR flows through three large fluvial lakes (Lake St. Francis, Lake St. Louis, and Lake St. Pierre) and around various islands.

Sample collection

Sampling from the water masses was conducted aboard the research vessel *Lampsilis* from UQTR, through five scientific expeditions between 2017 and 2022, during the summer (late July and early August). These expeditions covered different segments of the river from Lake Ontario, ON to Cacouna, QC as follows: i) 45 sites over a 250 km stretch from Montreal to Trois-Rivières in 2017, ii) 60 sites over 400 km from Lake Ontario to Trois-Rivières in 2018, iii) 66 sites over 450 km from Montreal to Cacouna in 2020, iv) 80 sites over 700 km from Lake Ontario to Cacouna in 2021, and v) 69 sites over 600 km from Lake Ontario to Quebec City in 2022 (Figure 3.1, Table 3.1). Orthogonal

transects were conducted during sampling, specifically targeting the three distinct water masses.

Table 3.1 Details of sampling conducted aboard the research vessel, 2017-2022, summer expeditions (July-August):

Mission year	Date	Sites	Distance (km)	Start	End
2017	09/07-15/07	45	250	Montreal	Trois-Rivières
2018	16/07-23/07	60	400	Lake Ontario	Trois-Rivières
2020	21/07-30/07	66	450	Montreal	Cacouna
2021	22/07-03/08	80	700	Lake Ontario	Cacouna
2022	28/07-07/08	69	600	Lake Ontario	Quebec City

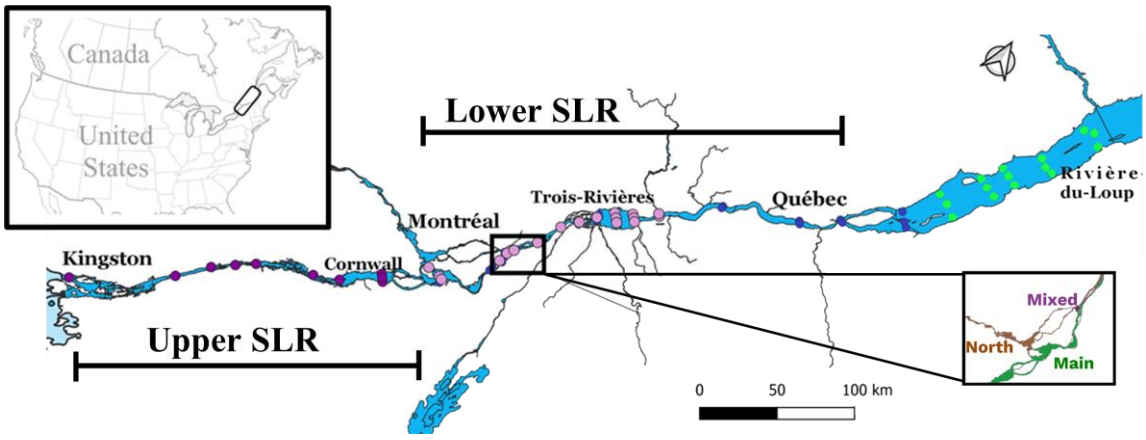


Figure 3.1 Map of sampling locations along the St. Lawrence River, Canada, from Lampsilis Annual Missions (2017, 2018, 2020, 2021 and 2022). Circles indicate sampling sites color-coded to represent distinct stretches along the river: Purple denotes the UPSLR stretch from Lake Ontario to Salaberry-de-Valleyfield, Quebec. Pink indicates the stretch from Montreal to Trois-Rivières, while Blue covers Trois-Rivières to Quebec City, including l'Île-d'Orléans. Green marks the estuary from l'Île-d'Orléans to Cacouna. The Pink and Blue color, denotes the LSLR stretch from Montreal to Quebec City. The figure illustrates three water masses: north (Ottawa River) mixed (Montreal Effluent), and green (main channel, water from Lake Ontario).

Shallow water sampling received support from the River Institute, a non-profit research institute, in 2021 and 2022 to sample 25 sites over 230 km stretch from Lake Ontario to Lake St. Francis in UPSLR, and from the ZIP les Deux-Rives river keeper organization to sample an additional 148 sites over 110 km stretch from Trois-Rivières to

Quebec City in LSLR between 2020 and 2022 (Figure S1). For the MST analysis, we targeted samples that showed >200 CFU, resulting in: i) 118 samples focusing on the three water masses over three years (2017,2018 and 2022), ii) 6 samples for shallow water in UPSLR for one year (2022), and iii) 10 samples for shallow water in LSLR for two years (2020 and 2021).

Water samples were collected from the research vessel *Lampsilis* using a GoFlo water sampler at one meter depth. Each sampler was sterilized with the local water at each station and conditioned three times prior to use. The sampling process involved filling 20 L sterilized polycarbonate bottles, following a three-wash protocol with the same water source before introducing the sample water for further analyses. Shallow water samples obtained from the UPSLR and LSLR, were collected approximately 50 m from the shore at 1 m depth. Shallow water samples were processed following the same protocol employed for water collected aboard the *Lampsilis*. Simultaneously with the collection of water samples, physicochemical parameters, including dissolved oxygen (mg mL^{-1}), pH, specific conductance (µS cm^{-1}), barometric pressure (mmHg), and water temperature (ºC), were measured using a ProDSS multiprobe YSI.

E. coli enumeration and DNA extraction

For *E. coli* concentration, 100 mL of water sample were filtered onto 0.45 µm nitrocellulose filter (MF). Depending on the level of FC expected, determined by previous samplings or specific location characteristics, dilution of 1:10, 1:100 or 1:1000 were performed with sterilized water. Once filtered, the filter was placed onto a Chromo Select agar plate enriched with X-glucuronide to support selective growth of *E. coli*, ensuring that no air was trapped. While X-glucuronide agar, generally considered very specific for *E. coli* enumeration, competition with other bacteria may still occur, potentially leading to a small percentage of false positive (Vergine, et al., 2017). Agar plates were then placed in an incubator at 44.5 ℃ for 24 hours. Results were reported as a colony forming units (CFU) per 100 mL. For DNA extraction, 300 mL was filtered (0.22 µm PES, Polyethersulfone filters) and stored at -80 ℃ following the protocol described in Edge et al. (2021). Filters were then thawed to conduct DNA extraction, using the Dneasy Power

Soil Pro Kit (Qiagen Inc., 2021) per the guidelines stipulated by the manufacturer, and stored at -80℃ for further analysis.

Microbial source tracking using digital PCR

Specific primers and probes sets were used, for the human HF183 marker (Green et al., 2014, AEM), ruminant Rum2Bac marker (Mieszkin et al., 2010, JAM), gull Gull4 marker (Ryu et al., 2012, AEM), and swine Pig2Bac marker (Mieszkin et al., 2009, AEM), due to their specific relationship with human environments and anthropogenic activities. dPCR reactions were conducted as a duplex for HF183 and Gull4 markers, without interference, and single Rum2Bac, Pig2Bac and mitochondrial DNA markers. For each dPCR reaction, the following components were combined:1 µL nuclease-free water, 0.75 µL of each 900 nM forward and reverse primer, 0.75 µL of each 250 nM probe, and 7.5 µL QuantStudio™ 3D Digital PCR Master Mix v.2 (ThermoFisher), along with 2µL of extracted DNA template. These reaction mixtures were then loaded into a 20,000 micro-well chip (QuantStudio™ 3D Digital PCR 20 Chip v2), where each partition had a volume of 755 pL. The loading process utilize the QuantStudio™ 3D Chip Loader (ThermoFisher). Subsequent PCR amplification was conducted using the ProFlex PCR System (ThermoFisher), following a thermal cycling (96℃ for 10 min, 40 cycles of 60℃ for 2 min and 98℃ for 30 s, and 2 min a 60℃). After PCR amplification, the 20 K Chip was subjected to analysis using the QuantStudio™ 3D Digital PCR Instrument (ThermoFisher). Positive and negative controls were processed alongside each set. (Edge et al., 2021).

The acquired results were processed by the QuantStudio™ AnalysisSuite™ software. This software performed multiple tasks, including the determination of threshold fluorescence values for the ROX reference dye, which helped identify qualified PCR well partitions. Additionally, the software analyzed the FAM and VIC dye signals to detect positive reactions related to the DNA markers. For further quantification, the AnalysisSuite™ software employed a Poisson Plus modeling technique. This approach allows the calculation of target concentrations within the sample. The final outcomes were reported as DNA copy numbers per 100 mL. No-template PCR controls used in the study

yielded negative results, demonstrating the absence of contamination that could affect dPCR assays. The average number of wells analyzed on duplex PCR chips for the HF183 and Gull4 markers was 17,499 (+/− 635), while for the Rum2Bac it was 17,252 (+/− 1,726), and for the Pig2Bac it was 17,892 (+/− 484). (Edge et al., 2021).

Thresholds and Guidelines for Human Health Protection

Evaluating water quality concerning human health risks linked to recreational water exposure, the established Health Canada guidelines from 2012 specify *E. coli* concentrations of 200 CFU per 100 mL for direct contact and 1,000 CFU per 100 mL for indirect contact (Health Canada, 2012). Traditionally, the standard for direct contact refers to the geometric mean concentration (based on a minimum of five samples). However, in our methodology, where we analyzed only one sample per station, we have adapted this guideline to apply a threshold of 200 CFU per 100 mL for each individual sample, ensuring rigorous assessment and alignment with safety standards for direct contact. We recognized that several significant changes have been included in the current fecal indicator guidelines due to the most recent epidemiological studies (Health Canada, 2023). Despite these updates, we have chosen to maintain continuity with Health Canda guidelines from 2012 to align with the initial project. For the HF183 DNA marker, a risk-based threshold (RBT) of 525 DNA copies per 100 mL is applied, derived from a risk assessment conducted by Boehm et al. (2020).

A detection threshold for DNA markers in the dPCR assay for this study was established by requiring a minimum of three PCR positive wells on a chip. This criterion was defined based on clear clusters distinctly separated from the fluorescence levels measured in PCR negative wells. The established threshold corresponds to a detection limit of 10 DNA copies per 100 mL. Although, there is still work to be done regarding defining definitive threshold using MST, current studies suggest a minimum of five PCR positive wells which is equivalent to a detection limit of 17 DNA copies per 100 mL (Edge et al., 2021).

Statistical analysis

All statistics analyses and data visualization were performed using R (v.4.3.0, R Core Team 2023) and R-studio (v.2023.09.0+463, R Core Team 2023), with graphical representation created using *ggplot2* package (Wickham et al., 2016), unless specified otherwise. Generalized Additive Models (GAMs), as statistical models, were implemented using the *mgcv* package v4.3.0 (Wood, 2017).

To identify hotspots of microbial contamination, GAMs were employed to investigate the spatial and temporal variation in *E. coli* concentrations from headwaters and across water masses. GAMs are an extension of the Generalized Linear Model and are characterized by their incorporation of smooth functions, allowing for linear and non-linear relationships between predictor variables and the response variable (Gomez-Rubio, 2018). Two GAMs were built with E. *coli* concentrations as the response variable. *E. coli* concentration was log-transformed to ensure model assumptions (Figure S4). The first GAM included a smooth by factor interaction between distance from headwater and water masses, a parametric effect of water masses, and a random intercept of year. The second GAM included latitude and longitude smooth interactions using a Duchon spline, and a smooth function of year. GAMs diagnostics and the maximum basis function (k) were verified using the *gam. Check* function and the function *appraise* from the *gratia* R package v0.8.1 (Simpson, 2023).

To identify potential bacterial contamination sources, correlation and regression analyses for MST were performed to explore the relationship between DNA markers (HF183, Gull4, Rum2bac, and Pig2bac) and *E. coli* concentration. Multiple linear regression models assessed the association between human DNA copies levels and the distance from headwaters across different water masses. All tests were conducted to assess statistical significance for deviations in either direction from the null hypothesis, and results with a p-value <0.05 were considered statistically significant. This threshold was applied across all analytical methods to maintain consistency in the interpretation of results. Additionally, a GAM was built to assess the temporal and spatial trends in Human DNA copies (HF183). The GAM model included log-transformed Human DNA copies

(HF183) as the response variable, latitude, and longitude smooth interactions using a Duchon spline, and a smooth function of year. Assumptions were checked as described previously for (Figure S4).

Results

Spatial and temporal variation in E. coli concentration

Over the five years sampled, a total of 341 water samples were collected along the SLR, from Kingston, Ontario to Cacouna, Quebec. *E. coli* was detected in 97% (330/341) of these samples. Examining threshold levels, significant variation was observed across different water masses and regions. When considering the total percentage exceedances of the three water masses combined, 23% of the samples exceeded the threshold for indirect contact, and 50% for direct contact (Figure S2A). When examining the individual water masses, we observed that the mixed water mass exhibits the highest contamination levels, with 60% exceeding the indirect contact limit, and 40% the direct contact limit. In the north water mass, 35% of samples exceeded the indirect contact limit and 53% the direct contact limit. Finally, in the main channel, the concern was related to direct contact, with 53% of samples exceeding this threshold (Figure S2A). When comparing shallow water samples from the upper and lower sections of the river (Figure S2B), a significant difference was observed ($p < 0.05$). In the UPSLR, 16% exceeded the threshold for indirect contact, and 40% for direct contact. In LSLR shallow water samples, levels were slightly higher, with 17% exceeding the indirect contact limit and 55% the direct contact limit.

The GAM showed a significant effect of distance from headwaters on *E. coli* concentration across water masses along the SLR. Spatial analysis along the St. Lawrence River, spanning from Kingston Ontario to Cacouna Quebec, revealed significant variations in *E. coli* concentrations (Figure 3.2, Table 3.2). Areas near Kingston and Cacouna presented comparatively low *E. coli* concentrations compared to other urban areas (Figure 3.2). We detected hotspots of microbial contamination near urban areas such

as Cornwall, Montreal, Trois-Rivières, and Quebec City with *E. coli* concentrations being the highest near the Montreal area (Figure 3.2).

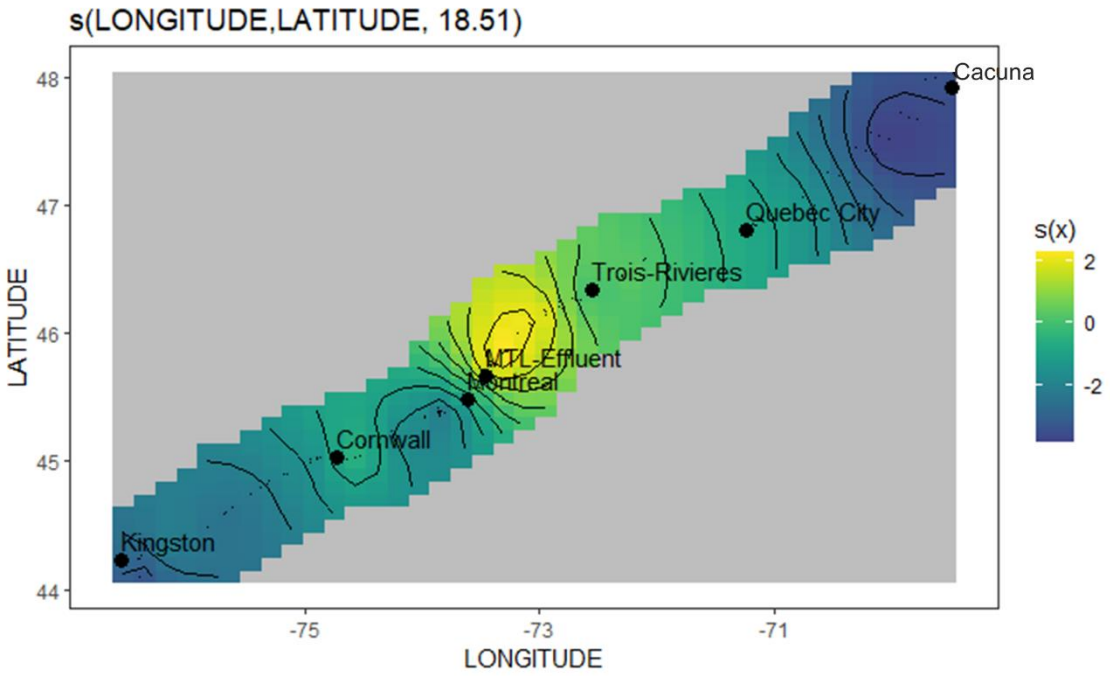


Figure 3.2 Generalized Additive Model (GAM) describing the spatial effect on *E. coli* concentration in the St. Lawrence River, from 2017 to 2022, from Kingston, Ontario to Cacouna, Quebec.

Table 3.2 The Generalized Additive Model (GAM) summary of the spatial and temporal effects of *E. coli* concentration (CFU per 100 mL) as a response variable in the St. Lawrence River, depicted as a function of year and longitude and latitude. Edf indicates the estimated degrees of freedom and Ref.df represents the residual degrees of freedom. Significant parametric (A) and partial effects (B) are indicated by $p < 0.05$.

A. Parametric coefficients	**Estimate**	**Standard Error**	**t-value**	***p*-value**
(Intercept)	5.68383	0.09827	57.84	<2e-16
B. Smooth terms	**edf**	**Ref.df**	**F-value**	***p*-value**
s(Year)	3.756	3.963	26.646	<2e-16
s(Longitude, Latitude)	18.506	34	9.189	<2e-16
R^2- adjusted	**0.534**			
Deviance explained	**56.40%**			
N	**341**			

The *E. coli* concentrations in UPSLR initially had low levels in the main channel, but with the incorporation of the north and mixed water masses, the LSLR levels varied and, after reaching maximum concentration (Montreal area), they declined with distance downstream (Figure S3). The temporal analysis was focused on the LSLR due to its consistent sampling throughout the five-year study (Table 3.3, Figure 3.3). Our results revealed a significant decrease in *E. coli* concentrations with distance from the headwaters in the main and mixed water masses (Figure 3.3A-B). In contrast, a notable increase in *E. coli* concentration was detected only within the north water mass near Montreal followed by a decline in *E. coli* concentration (Figure 3.3C). In the mixed water mass, elevated *E. coli* concentrations were evident, reaching millions of colonies. The average concentration stands at 7,836,750 CFU per 100 mL, with the maximum count of 27,700,000 CFU per 100 mL recorded during the summer of 2022.

Table 3.3 The Generalized Additive Model (GAM) summary of the partial effect of *E. coli* concentration (CFU per 100 mL) as response variable across different water masses in the Lower St. Lawrence River, depicted as a function of distance and the random effect of year. edf indicates the estimated degrees of freedom and Ref.df represents the residual degrees of freedom. Significant parametric (A) and partial effects (B) are indicated by $p < 0.05$.

A. Parametric coefficients	**Estimate**	**Standard Error**	**t-value**	***p*-value**
(Intercept)	5.2747	0.4489	11.751	<2e-16
Mixed Water Mass	2.7270	0.1951	13.976	<2e-16
North Water Mass	1.1726	0.1983	5.913	1.57e-08
B. Smooth terms	**edf**	**Ref.df**	**F-value**	***p*-value**
s(Distance): Main	1.001	1.001	12.17	0.000601
s(Distance): Mixed	3.271	4.043	72.98	<2e-16
s(Distance): North	5.923	7.215	10.69	<2e-16
s(Year)	3.860	4.000	28.29	<2e-16
R^2- adjusted	**0.78**			
Deviance explained	**79.7%**			
N	**204**			

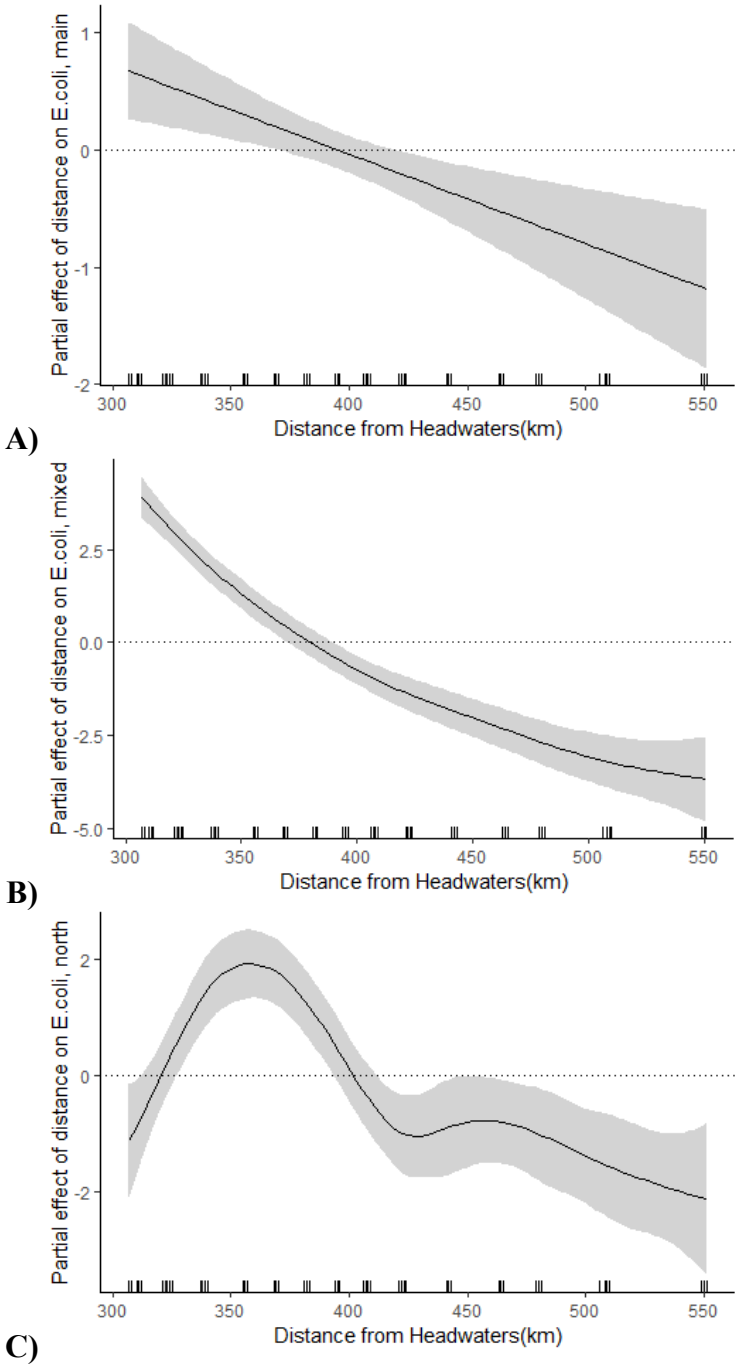


Figure 3.3 Generalized Additive Model (GAM) describing the partial effect of distance from headwaters (km) on *E. coli* concentration (CFU per 100 mL) across different water masses in the Lower St. Lawrence River while accounting for the random effect of year. A) Main water mass. B) Mixed water mass. C) North water mass. The shaded area indicates the 95% confidence interval, dashed line indicates the null effect, and the rug plot (tick line on the x axis) are the observations of the predictor variables.

Unlike the different water masses that exhibited similar trends of *E. coli* concentration through the years, shallow waters exhibited more dynamic influences due to localized inputs or runoff. During 2021 and 2022, the north shore sampling sites of the UPSLR had lower levels than the south shore. On the contrary, in the LSLR the north shore sampling sites exhibited higher concentration than the south shore (Figure 3.4).

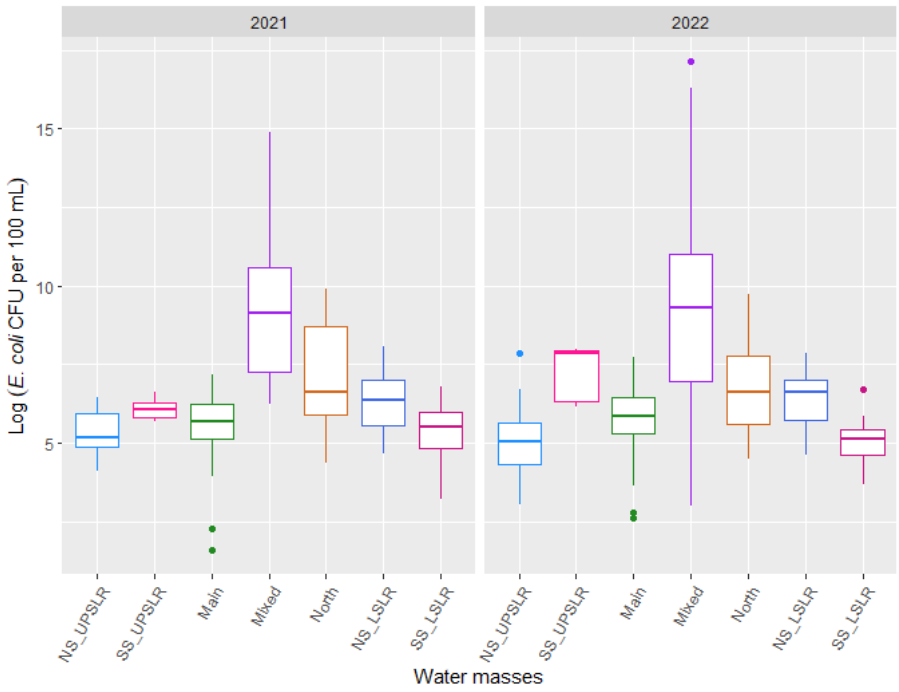


Figure 3.4 Boxplots of *E. coli* concentration for each water mass (main, north, and mixed) and shallow waters from north shore (NS) and south shore (SS) from UPSLR and LSLR, during 2021 and 2022 sampling period, from Kingston, Ontario to Quebec City, Quebec.

A temporal effect during the five-year study, revealed a significant influence of year on *E. coli* concentrations in the SLR (Table 3.2, Figure 3.5). *E. coli* concentrations remained stable on average from 2017 to 2019, followed by an increase from 2020 to 2021, which seemed to level off in 2022 (Figure 3.5).

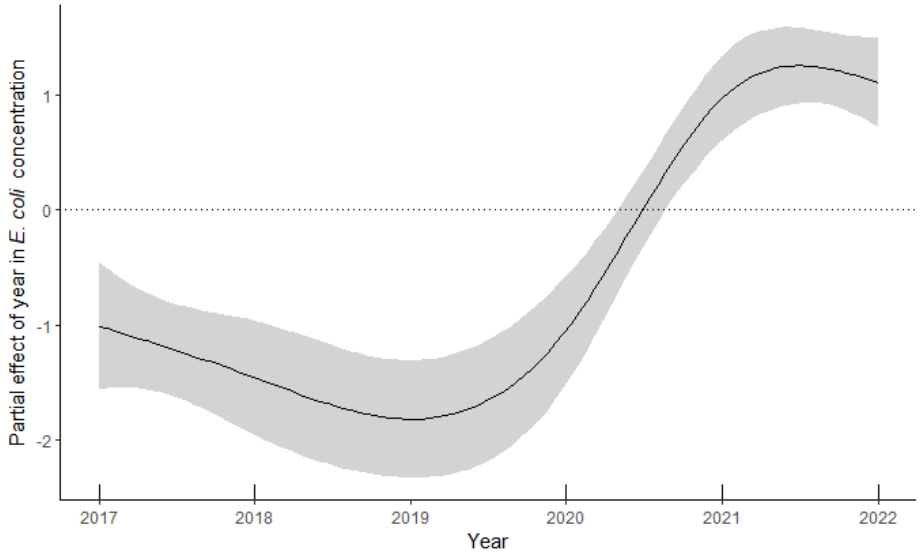


Figure 3.5 Generalized Additive Model (GAM) describing the temporal partial effect on *E. coli* concentration in the St. Lawrence River, from 2017 to 2022. Shaded area indicates the 95% confidence interval, dashed line indicates the null effect, and the rug plot (tick line on the x axis) are the observations of the predictor variables, with an extrapolation of the year 2019.

Microbial source tracking

The human marker (HF183) was detected in 71.9% (95/132) of samples, with a varying detection rate observed across different water masses: 94% in the mixed water mass, 81% in the north water mass, and 52% in the main water mass. We found 29% of observations exceeded the 525 DNA copies per 100 mL risk-based threshold (RBT). Notably, this contamination is not evenly distributed across different water masses. Of the samples exceeding the RBT, 15% originate from the mixed water mass and 11% from the north water mass. Shallow water areas exhibited distinct detection rates, with 33% for UPSLR, and 100% for LSLR. The gull marker (Gull4) was detected in 56% of samples, while the pig marker (Pig2bac) had a lower detection rate of 3.5%. The ruminant marked (Rum2bac) was not detected in this study.

Human DNA copies significantly varied along the SLR, extending from Kingston, Ontario to Trois-Rivières, Quebec (Table 3.4, Figure 3.6). A similar pattern to that observed for *E. coli* was noted, with concentrations increasing near urban areas and

particularly high levels in the Montreal area. The highest concentrations recorded were 411,067 DNA copies per 100 mL for HF183, observed at the Montreal effluent site. In contrast, the UPSLR exhibited the lowest concentrations for both *E. coli* concentration and DNA copies per 100 mL for HF183.

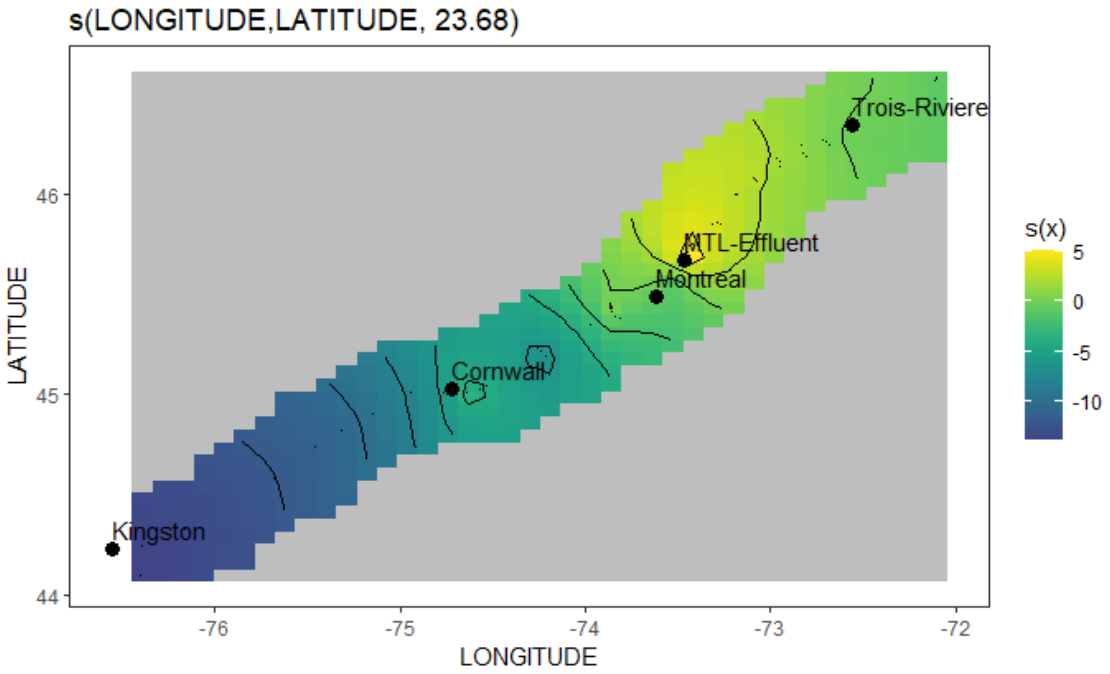


Figure 3.6 Generalized Additive Model (GAM) describing the spatial effect of human DNA copes (HF183) in the St. Lawrence River, for the years 2017, 2018 and 2022, from Kingston, Ontario to Trois-Rivières, Quebec.

Table 3.4 The Generalized Additive Model (GAM) summary of the spatial effect of Human DNA copies (HF183) as response variable in the St. Lawrence River, depicted as a function of year and longitude and latitude. Edf indicates the estimated degrees of freedom and Ref.df represents the residual degrees of freedom. Significant parametric (A) and partial effects (B) are indicated by $p < 0.05$.

A. Parametric coefficients	**Estimate**	**Standard Error**	**t-value**	***p*-value**
(Intercept)	4.8038	0.4586	10.47	<2e-16
B. Smooth terms	**edf**	**Ref.df**	**F-value**	***p*-value**
s(Longitude, Latitude)	23.679	39	8.858	<2e-16
s(Year)	1.638	2	4.214	0.00697
R^2- adjusted	**0.538**			
Deviance explained	**81%**			
N	**118**			

The scatter plot correlation strength indicates the positive linear relationship between *E. coli* concentration and human DNA copies (Figure 3.7, Correlation coefficient = 0.89). Human DNA copies were predominant, having a positive significant correlation with *E. coli* levels, followed by gulls. The pig signal was detected at a single site, showing notably low levels (Figure 3.7A). We then assessed how the human marker would vary across water masses and found positive correlations with the mixed water mass presenting the highest positive correlation. In contrast, shallow waters displayed a negative correlation, possibly influenced by dynamic factors and wildlife inputs, suggesting that *E. coli* presence is not always related to human sources (Figure 3.7B).

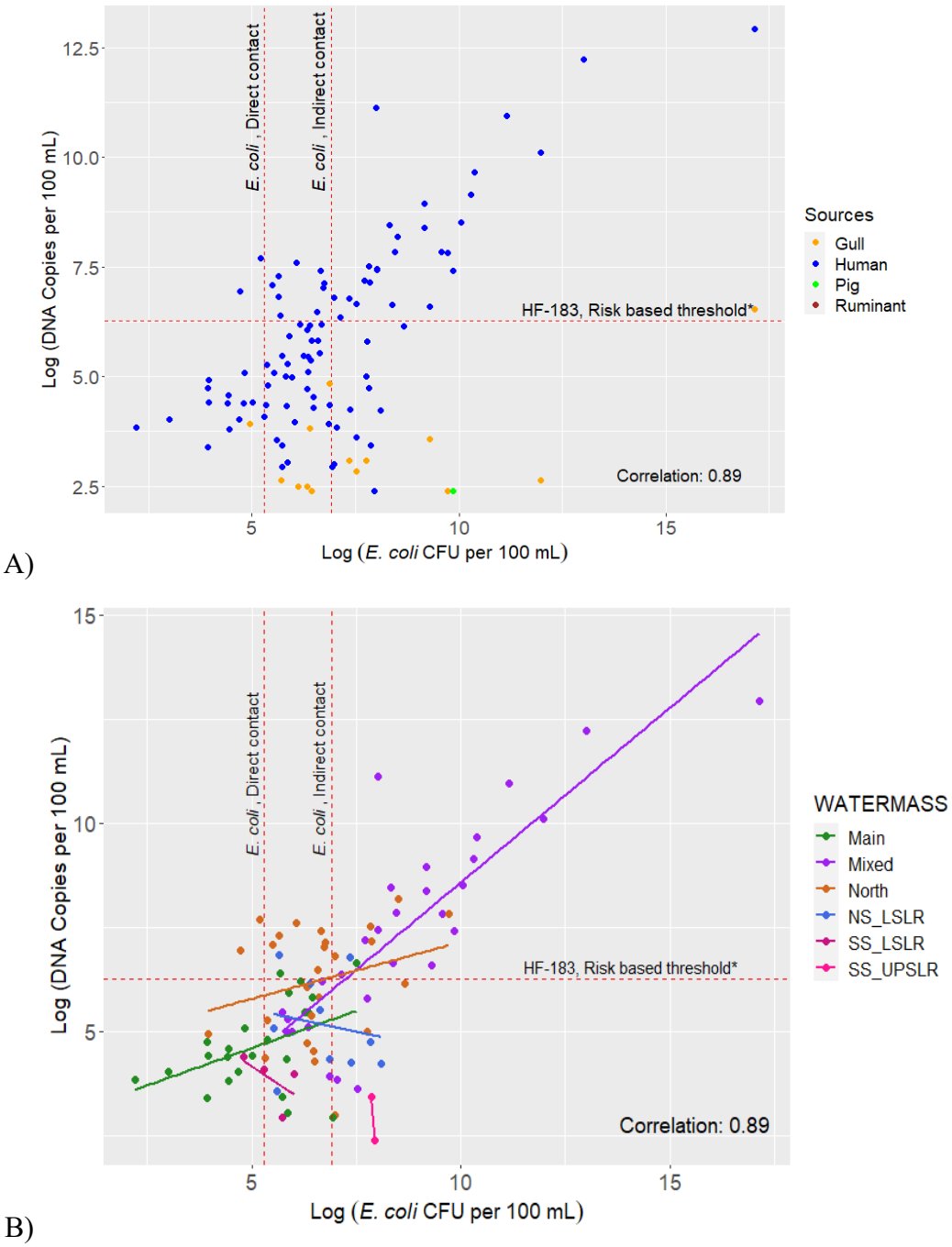


Figure 3.7 Scatter plot illustrating the relationship between DNA copies of specific markers (log-transformed) and *E. coli* concentration (CFU per 100 mL) (log-transformed) in the St. Lawrence River. **A)** Markers include human (HF183), gull (Gull4), pig (Pig2bac), and ruminant (Rum2bac). **B)** Human marker (HF183), from 2017 to 2022, excluding 2019.

Discussion

In this study, we employed *E. coli* as a FIB to assess the presence and persistence of FC across a large river system. Our findings reveal a significant spatial heterogeneity in *E. coli* levels across the SLR, highlighting the challenge in generalizing its concentration throughout the river system. This variability was influenced by anthropogenic sources and modulated by physical and environmental factors, including the unique geographic characteristics of the river, such as urban proximity, islands, fluvial lakes, and tributaries. This suggests that anthropogenic pollution, which is similar to the dynamic nature of river ecosystems, varies across the river both in terms of quantity and origin. This leads to diverse concentrations and persistence of *E. coli*, with possible ecological and human health consequences in the SLR (Ashbolt, 2004; Elahi et al., 2017).

The presence of FC was found in most samples, with varying *E. coli* concentration across different water masses and shallow waters. The results indicate that the SLR generally exceeds the safety limits for both direct and indirect contact, but not uniformly across all water masses. Water from Lake Ontario (main) had the highest percentage of samples within safety limits (47%), while water from Ottawa River (north) had only a small percentage within safety limits (12%). Water from Montreal effluent (mixed) exhibited the highest percentage of samples exceeding safety limits (100%), which is in line with the fact that waste waters only undergo primary treatment at Montreal's WWTP. These results also highlight the significant footprint of large cities on water quality in large river systems. In shallow waters, the LSLR had a higher percentage of samples exceeding safety limits compared to UPSLR, reflecting the higher density of human population and riverine cities located along the LSLR. Moreover, the data show that *E. coli* levels near the shore are more variable than central water masses, as shown in Figures 3.4 and 3.7. This variability is likely influenced by a changing combination of local factors, inputs from tributaries, wildlife, agricultural activities, and urban aeras, making each sampling site unique.

The spatial analysis further suggests that effluent discharges and specific locations along the SLR could be the contributing factor of FC, as shown by hotspots identified in

the SLR (Figure 3.2). Notably, *E. coli* concentration in the mixed water mass followed a dilution pattern, well captured by our GAM model (Figure 3.3B). This dilution to acceptable levels (below safety limits of 1,000 CFU and 200 CFU) does not occur until more than 100 km downstream, towards Quebec City and the estuarine environment. In the northern water mass, which flows extensively through the island system, there was a sharp drop in *E. coli* counts, shown in Figure 3.3C, suggesting that natural processes and/or cell sedimentation and burial may play a filtering role in such system at the river landscape level. These dynamics also suggest a complex interaction of anthropogenic and environmental factors shaping the distribution of *E. coli* along the SLR.

The temporal analysis showed a significant variation in our data with a sharp increase in *E. coli* concentration detected in 2021 and 2022. This increase may be due to natural events, population growth, and landscape alterations. We explored potential links to changes in water level, precipitation, and populations along the river by analyzing historical climate data, water level records, and demographic information using linear regression analysis and a significance level of $p < 0.05$. However, despite these efforts, we could not determine a specific cause or driver, highlighting the importance of long-term monitoring to better understand water quality dynamics in large river systems like the SLR. Further investigation should also consider the timing of agricultural inputs, migratory bird patterns, and precipitation events to better improve our understanding of how inputs into the river fluctuate, as these factors are influenced by basin runoff.

Animal and human waste are the primary contributors to surface water pollution (Ashbolt, 2004; Ichor et al., 2014). In this study, human origin of contamination was the most frequent, with some samples exceeding health risk limits for both *E. coli* and HF183 (Boehm et al. 2020; EPA, 2021; Health Canada, 2012), especially in the mixed water mass. The highest *E. coli* concentrations were recorded near the Montreal outflow site, while the UPSLR showed the lowest levels. These findings align with the case study on the largest volume of wastewater discharge in the SLR from the Montreal WWTP (Marcogliese et al., 2015), and with other studies in the Great Lakes areas (Edge et al., 2021). This domination of human sources in the SLR system calls for a better management

of wastewaters such as a reduction in overflows during extreme precipitation events and improvements in wastewater treatment with the implementation of additional disinfection steps.

Interestingly, we found that gull could cause significant increases in *E. coli* concentration, but in the case of the SLR, these increases seemed to be more localized to a few sites, particularly in shallow waters and areas influenced by urban stormwaters. Gulls are highly adaptable to various habitat, including human-altered environments, their presence in shallow water is likely due to available food sources and breeding preferences, while in urban areas, they are attracted to food sources such as garbage (Anderson, 2013). On the other hand, it should be noted that while gull contamination may present some threat to human health, zoonotic waterborne infections generally present a lower risk compared to pathogens of human origin, due to a mismatch in host-infection capabilities (Harwood et al., 2014), but have nonetheless been shown to be vectors causing severe outbreaks depending on the pathogen involved. However, the presence of different contamination sources in the same area can create more harmful conditions than a single source alone (Alm et al., 2018; Martín-Vélez et al., 2023; Rock et al., 2015).

This study has limitations in the sampling process, as samples were collected annually during the summer without replication at each site. This single-season approach prevents understanding the river's year-round dynamics, including the impacts of seasonal events like rainfall, ice melt, water level fluctuations, agricultural runoff (including fertilizer), and bird migration, among others (Li et al., 2021). Finally, some sites corresponding to high *E. coli* levels and low population density showed minimal signal from our four markers, suggesting the presence of wildlife contamination or other sources. For instance, extensive wetlands are presented through the river which are known habitats for waterflows and other wildlife (McDuie et al., 2022). Expanding the research to watersheds and other markers could help our understanding of infectious diseases, that can come from wildlife (Ichor et al., 2014). Establishing a clear detection threshold criterion for DNA markers based on the number of positive wells could improve accuracy. This approach could also help refine the RBT for HF183 by examining the relationship

between *E. coli* concentration and human DNA copies, as we observed that in areas where *E. coli* concentrations exceeded the health risk limit (Health Canada, 2012), the RBT of 525 copies found by Boehm et al. (2020) being much lower.

Conclusions

In this study we identified hotspots of microbial contamination within the SLR. The presence of human DNA dominated these hotspots, indicating extensive inputs of non-disinfected wastewaters from urban areas. This influence is higher in regions that present both high human population density and the presence of inadequate or minimally treated water treatment plants, demonstrated by the significant impact observed near the Montreal effluent.

The noticeable association between high concentration of *E. coli* and wastewater discharges at the central channel (mixed water mass) highlights the fundamental role of urban areas in influencing water quality. The extent of this influence is evident in its downstream persistence, as we measured the presence of FC way beyond 100 km downstream of the discharge point.

Additionally, nearshore areas were found to be more dynamic, which highlights the need for specific studies and monitoring programs using approaches to MST similar to our own to comprehensively assess water quality and reduce potential harmful exposition to contaminated waters. Recognizing the complexities of large rivers and microbial contamination in these dynamic zones is crucial to developing specific management strategies.

Finally, our study highlights the importance of considering both localized and general influences to formulate strategies that mitigate the impact of anthropogenic activities on water quality, eventually to contribute to the preservation of aquatic ecosystems.

Acknowledgments

We thank the St. Lawrence River Institute, Jeff Ridal, Christina Collard, Matthew Windle, Brian Hickey, Stephany Hildebrand, Zach Zwanenburg, Philip Ling and all dedicated staff and volunteers for supporting in fieldwork and data obtaining; and Cristina Charette and Erin Smith for their contribution in the statistical analysis. We are also grateful to Charis Wrong at INRS for assistance with lab analyses.

Funding

This project has been made possible by a MITACS scholarship awarded to Silvia Rodriguez as part of the MITACS Accelerate program, with funding provided in collaboration with the River Institute in Cornwall, Ontario. The scientific cruises on the St. Lawrence River, conducted onboard the *Lampsilis* research vessel, were made possible through the financial assistance of the Réseau Québec Maritime and the Government of Quebec via the Odyssée Saint-Laurent research program. Shallow water sampling was supported by ZIP les Deux-Rives river keeper organization and the River Institute, and associated volunteers. This thesis represents a significant contribution to the research program of the Université du Québec à Trois-Rivières (UQTR) Chair on the Ecology of the St. Lawrence River, under the leadership of François Guillemette. The support and resources provided by these organizations and individuals, as their contributions, have been instrumental in the successful completion of this study.

Declaration of interest statement

No potential conflict of interest was reported by the authors.

References

Alm, E. W., Daniels-Witt, Q. R., Learman, D. R., Ryu, H., Jordan, D. W., Gehring, T. M., & Santo Domingo, J. (2018). Potential for gulls to transport bacteria from human waste sites to beaches. Science of the Total Environment, 615, 123–130.

Anderson, C. M. (2017). Annual spatial ecology of herring gulls (Larus argentatus) in eastern North America (Doctoral dissertation, Acadia University).

Ashbolt, N. J. (2004). Microbial contamination of drinking water and disease outcomes in developing regions. Toxicology, 198(1-3), 229-238.

Ballesté, E., Demeter, K., Masterson, B., Timoneda, N., Sala-Comorera, L., & Meijer, W. G. (2020). Implementation and integration of microbial source tracking in a river watershed monitoring plan. Science of the Total Environment, 736, 139573.

Barth, J. A. C., & Veizer, J. (2004). Water mixing in a St. Lawrence River embayment to outline potential sources of pollution. Applied Geochemistry, 19(10), 1637–1641.

Behrendt, H., Kornmilch, M., Opitz, D., Schmoll, O., & Scholz, G. (2002). Estimation of the nutrient inputs into river systems - experiences from German rivers. In Journal of Material Cycles and Waste Management, 3 (1–3), 107–117.

Boehm, A. B., & Soller, J. A. (2020). Refined ambient water quality thresholds for human-associated fecal indicator HF183 for recreational waters with and without co-occurring gull fecal contamination. Microbial Risk Analysis, 16, 100139.

Corsi, S. R., Borchardt, M. A., Spencer, S. K., Hughes, P. E., & Baldwin, A. K. (2014). Human and bovine viruses in the Milwaukee River watershed: Hydrologically relevant representation and relations with environmental variables. Science of the Total Environment, 490, 849–860.

Dang, D. H., Ma, L., Ha, Q. K., & Wang, W. (2022). A multi-tracer approach to disentangle anthropogenic emissions from natural processes in the St. Lawrence River and Estuary. Water Research, 219, 118588.

Edge, T. A., Boyd, R. J., Shum, P., & Thomas, J. L. (2021). Microbial source tracking to identify fecal sources contaminating the Toronto Harbour and Don River watershed in wet and dry weather. Journal of Great Lakes Research, 47(2), 366–377.

Elahi, E., Abid, M., Zhang, L., & Alugongo, G. M. (2017). The use of wastewater in livestock production and its socioeconomic and welfare implications. Environmental Science and Pollution Research, 24(21), 17255–17266.

EPA, U. S. Environmental Protection Agency, (2021). *E. coli* (*Escherichia coli*), Factsheet on water quality parameters, EPA 841F21007F.

Gomez-Rubio, V. (2018). Generalized Additive Models: An Introduction with R (2nd Edition). Journal of Statistical Software, 86,1-5.

Green, H. C., Haugland, R. A., Varma, M., Millen, H. T., Borchardt, M. A., Field, K. G., Walters, W. A., Knight, R., Sivaganesan, M., Kelty, C. A., & Shanks, O. C. (2014). Improved HF183 quantitative real-time PCR assay for characterization of human fecal pollution in ambient surface water samples. Applied and Environmental Microbiology, 80(10), 3086–3094.

Gupta, A. (Ed.). (2007). Large rivers: geomorphology and management. John Wiley & Sons.

Harwood, V. J., Staley, C., Badgley, B. D., Borges, K., & Korajkic, A. (2013). Microbial source tracking markers for detection of fecal contamination in environmental waters: Relationships between pathogens and human health outcomes 38(1), 1-40.

Health Canada. (2012). Guidelines for Canadian recreational water quality, Third Edition. Water, air and Climate Change Bureau, Healthy Environments and Consumer Safety Branch, Health Canada, Ottawa, Ontario.

Health Canada. (2023). Guidelines for Canadian recreational water quality, Indicators of Fecal Contamination, Health Canada, Ottawa, Ontario.

Ichor, T., Umeh, E. U., & Duru, E. E. (2014). Microbial Contamination of Surface Water Sources in Rural Areas of Guma Local Government Area of Benue State, Nigeria. Journal of Medical Sciences and Public Health, 2(2), 43–51.

Li, E., Saleem, F., Edge, T. A., & Schellhorn, H. E. (2021). Biological indicators for fecal pollution detection and source tracking: A review. In Processes, 9(11), 2058.

Maranger, R. J., Pace, M. L., del Giorgio, P. A., Caraco, N. F., & Cole, J. J. (2005). Longitudinal spatial patterns of bacterial production and respiration in a large river-estuary: Implications for ecosystem carbon consumption. Ecosystems, 8(3), 318–330.

Marcogliese, D. J., Blaise, C., Cyr, D., de Lafontaine, Y., Fournier, M., Gagné, F., Gagnon, C., & Hudon, C. (2015). Effects of a major municipal effluent on the St. Lawrence River: A case study. In Ambio, 44(4), 257–274.

Martín-Vélez, V., Navarro, J., Figuerola, J., Aymí, R., Sabaté, S., Planell, R., Vila, J., & Montalvo, T. (2023). A spatial analysis of urban gulls contribution to the potential spread of zoonotic and antibiotic-resistant bacteria. Science of the Total Environment, 912, 168762.

McKee, A., Molina, M., Cyterski, M., & Couch, A. (2020). Microbial source tracking (MST) in Chattahoochee River National Recreation Area: seasonal and

precipitation trends in MST marker concentrations, and associations *with E. coli* levels, pathogenic marker presence, and land use. Water research, 171, 115435.

Mieszkin, S., Furet, J. P., Corthier, G., & Gourmelon, M. (2009). Estimation of pig fecal contamination in a river catchment by real-time PCR using two Pig-Specific Bacteroidales 16S rRNA genetic markers. Applied and Environmental Microbiology, 75(10), 3045–3054.

Mieszkin, S., Yala, J. F., Joubrel, R., & Gourmelon, M. (2010). Phylogenetic analysis of Bacteroidales 16S rRNA gene sequences from human and animal effluents and assessment of ruminant faecal pollution by real-time PCR. Journal of Applied Microbiology, 108(3), 974–984.

Okello, P. E., Bulage, L., Riolexus, A. A., Kadobera, D., Kwesiga, B., Kajumbula, H., Mulongo, M., Namboozo, E. J., Pimundu, G., Ssewanyana, I., Kiyaga, C., Aisu, S., & Zhu, B. P. (2019). A cholera outbreak caused by drinking contaminated river water, Bulambuli District, Eastern Uganda, March 2016. BMC Infectious Diseases, 19(1), 1-8.

Potter, P. E., Hamblin, K., & Hamblin, W. K. (2006). Big Rivers Worldwide Part 1 Origins. (Vol. 48). Department of Geology, Brigham Young University.

R Core Team (2023) R: A language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rock, C., Rivera, B., & Gerba, C. P. (2015). Microbial Source Tracking. In Environmental Microbiology: Third Edition (pp. 309–317). Elsevier Inc.

Ryu, H., Griffith, J. F., Khan, I. U. H., Hill, S., Edge, T. A., Toledo-Hernandez, C., Gonzalez-Nieves, J., & Domingoa, J. S. (2012). Comparison of gull feces-specific assays targeting the 16S rRNA Genes of Catellicoccus marimammalium and Streptococcus spp. Applied and Environmental Microbiology, 78(6), 1909–1916.

Savio, D., Sinclair, L., Ijaz, U. Z., Parajka, J., Reischer, G. H., Stadler, P., Blaschke, A. P., Blöschl, G., Mach, R. L., Kirschner, A. K. T., Farnleitner, A. H., & Eiler, A. (2015). Bacterial diversity along a 2600km river continuum. Environmental Microbiology, 17(12), 4994–5007.

Simpson G. (2023). gratia: Graceful ggplot-Based Graphics and Other Functions for GAMs Fitted using mgcv. R package version 0.8.1.

Tallon, P., Magajna, B., Lofranco, C., & Kam, T. L. (2005). Microbial indicators of faecal contamination in water: A current perspective. In Water, Air, and Soil Pollution, 166(1–4), 139–166.

Vergine, P., Salerno, C., Barca, E., Berardi, G., & Pollice, A. (2017). Identification of the faecal indicator Escherichia coli in wastewater through the β-D-glucuronidase activity: comparison between two enumeration methods,

membrane filtration with TBX agar, and Colilert-18. Journal of Water and Health, 15(2), 209-217.

Vijayan, J., Nathan, V. K., Ammini, P., & Ammanamveetil, A. M. H. (2023). Bacterial diversity in the aquatic system in India based on metagenome analysis—a critical review. In Environmental Science and Pollution Research, 30(11), 28383–28406. Springer Science and Business Media Deutschland GmbH.

Wickham, H., Chang, W., & Wickham, M. H. (2016). Package 'ggplot2'. Create elegant data visualisations using the grammar of graphics, 2(1), 1-189.

Wood, S. N. (2017). Generalized additive models: an introduction with R. chapman and hall/CRC.

Yang, C., Telmer, K., & Veizer, J. (1996). Chemical dynamics of the "St. Lawrence" riverine system: δDH2O, δ18OH2O, δ13CDIC, δ34Ssulfate, and dissolved 87Sr/86Sr. Geochimica et Cosmochimica Acta, 60(5), 851-866.

Zandaryaa, S., & Mateo-Sagasta, J. (2018). Chapter 8. Organic matter, pathogens and emerging pollutants.

Table captions

Table **Page**

Figure captions

Figure **Page**

Supplementary Figures

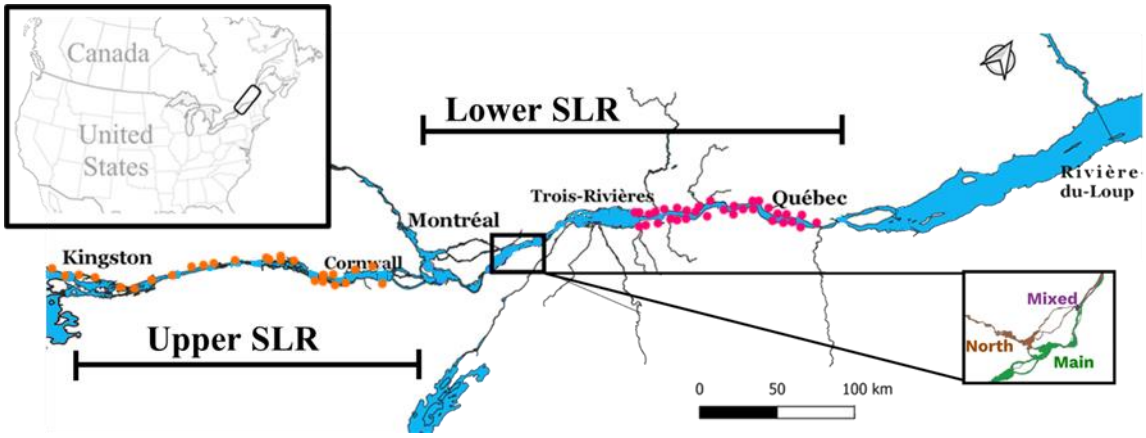


Figure S1 Map of shallow water sampling locations along the St. Lawrence River, Canada. Circles indicate sampling sites color-coded to represent distinct stretches along the river: Orange denotes the UPSLR stretch from Lake Ontario to Salaberry-de-Valleyfield, Quebec. Pink indicates the LSLR stretch from Trois-Rivières to Quebec City. The figure illustrates three water masses: north (Ottawa River) mixed (Montreal Effluent), and green (main channel, water from Lake Ontario).

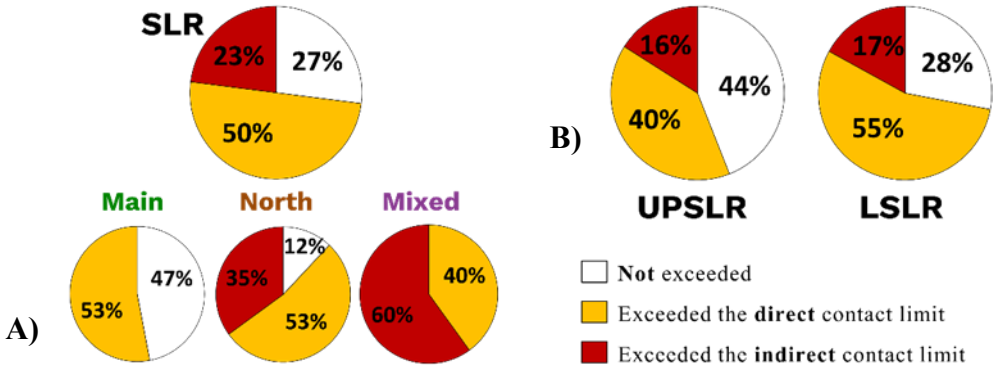


Figure S2 Proportion of sites exceeding the limits according to established guidelines for *E. coli* concentrations set at 200 CFU per 100 mL for direct contact and 1,000 CFU per 100 mL for indirect contact (Health Canada, 2012). **A)** Samples from SLR and the distinct water masses (main, north and mixed) from 2017 to 2022, excluding 2019.**B)** Shallow water samples from UPSLR and LSLR collected in 2021 and 2022.

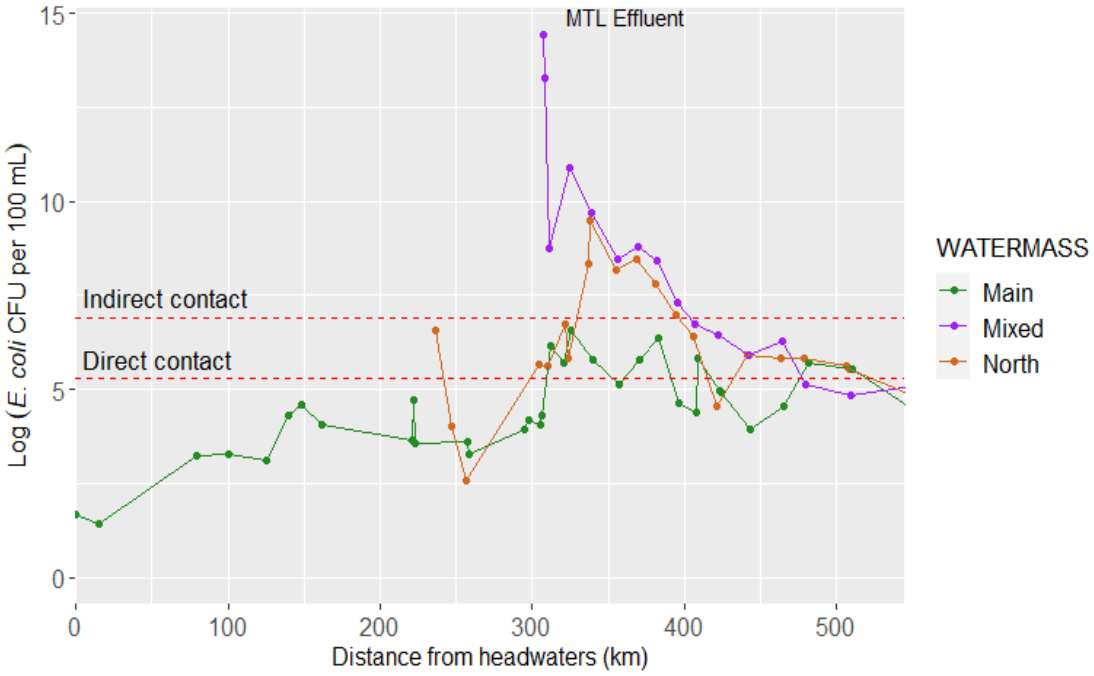
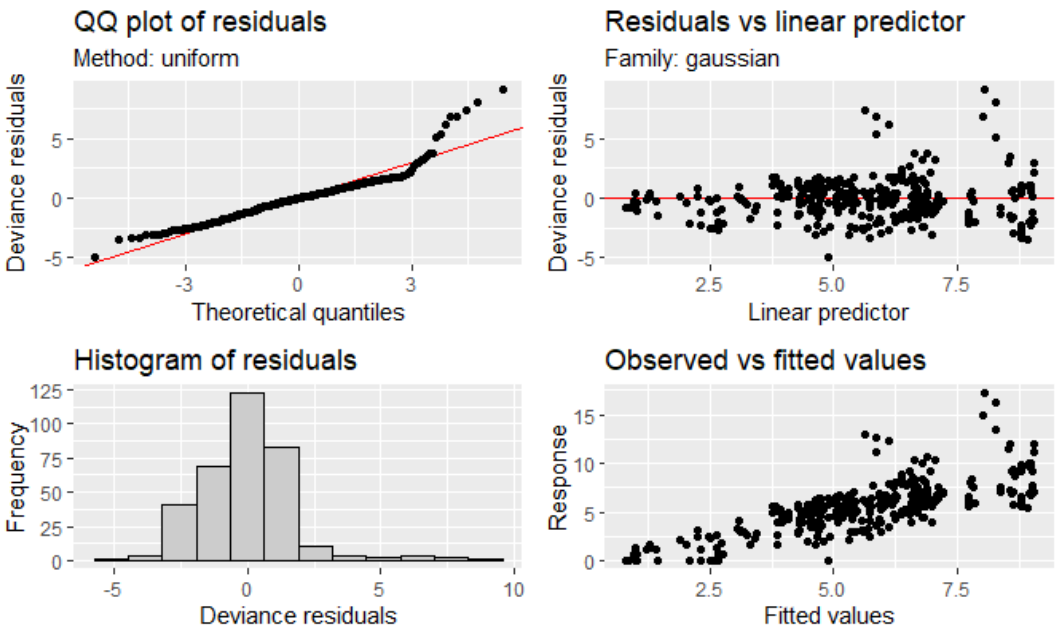
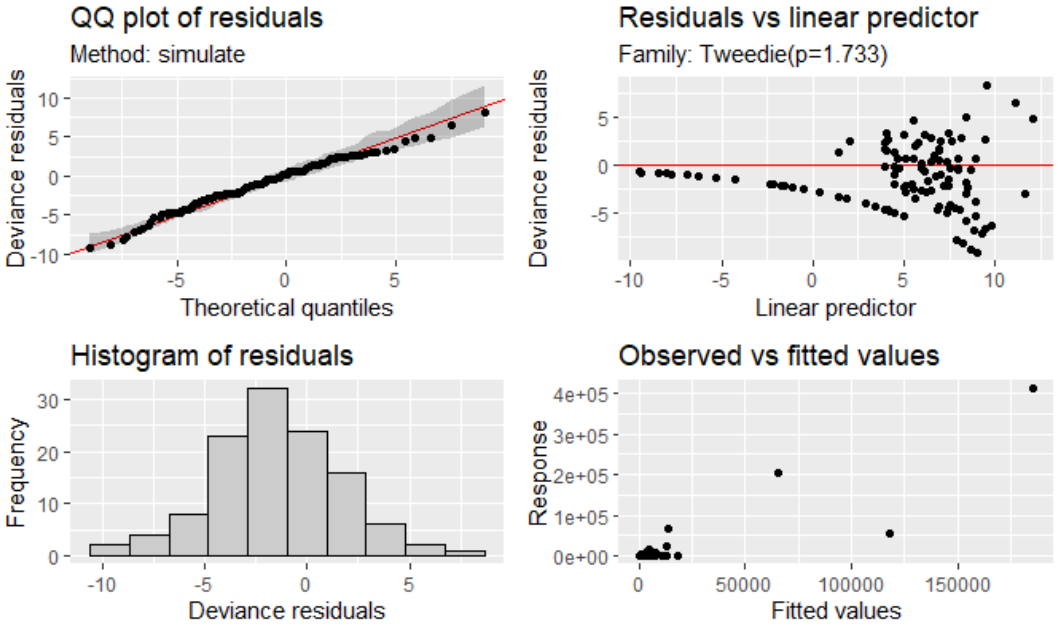


Figure S3 Variation in *E. coli* concentration (CFU per 100 mL) (log-transformed) across different water masses of the St. Lawrence River, categorized into main (green), mixed (purple), and north (orange) sections, plotted against distance from headwaters (km). Dashed line indicates the limits for recreational water exposure, the established guidelines for *E. coli* concentrations are 200 CFU per 100 mL for direct contact and 1,000 CFU per 100 mL for indirect contact (Health Canada, 2012).



A)



B)

Figure S4 Diagnostic plots: **A)** *E. coli* concentration GAM fit with a gaussian distribution. **B)** Human DNA copes (HF183) GAM fit with a Tweedie distribution.

Chapter 4

Conclusions and Recommendations

4. 1 **Findings Overview**

Human and animal waste are some of the main contributors to water pollution, representing a widespread issue (Ashbolt, 2004; Ichor et al., 2014). This contamination poses a significant risk to aquatic ecosystems and human health, necessitating rigorous identification and management strategies (Corsi et al., 2014; Zandaryaa et al., 2018). Monitoring *E. coli* concentration in large river systems is crucial for public health, affecting drinking water, recreational activities, and food consumption (Rock et al., 2015). Our study takes an important step towards understanding the influence of anthropogenic sources of bacterial contamination on *Escherichia Coli (E. coli)* concentrations in the St. Lawrence River (SLR) and identifying their sources.

Initially, we identified hotspots of microbial contamination by analyzing *E. coli* levels and its persistence downstream. These hotspots were determined based on safety limits for recreational water exposure: 200 CFU (Colony Forming Units) per 100 mL for direct contact and 1,000 CFU per 100 mL for indirect contact. These thresholds, based on environmental and health guidelines for water quality (Health Canada, 2012), allowed us to quantitatively assess areas at high risk of microbial contamination. We then employed microbial source tracking (MST) techniques, utilizing human HF183 marker (Green et al., 2014, AEM), ruminant Rum2Bac marker (Mieszkin et al., 2010, JAM), gull Gull4 marker (Ryu et al., 2012, AEM), and swine Pig2Bac marker (Mieszkin et al., 2009, AEM), to identify potential sources of contamination. For the human marker, a risk-based threshold (RBT) of 525 DNA copies per 100 mL was applied, derived from a risk assessment conducted by Boehm et al. (2020).

The SLR, considered as a large river (Gupta, 2007), is characterized by three distinct and heterogeneous water masses flowing side by side downstream (Hudon &

Carignan, 2008). It also includes fluvial lakes, islands, and tributary inputs (CECC, 2021). We found that these complexities and dynamics significantly influence the distribution of bacterial contamination, with diverse *E. coli* concentrations in the different water masses decreasing after passing through various riverine units. In shallow water, the distribution plays a different role due to limited dispersion and greater influence from direct inputs and runoff.

During our five-year scientific expeditions, we detected *E. coli* present in 97% of the 341 samples. The *E. coli* concentrations were not uniformly distributed within the SLR. The Figure 3.2 from Chapter 3, highlights hotspots near urban areas, such as Cornwall, Montreal, Trois-Rivières, and Quebec City. The highest concentrations were found in the Montreal area, specifically in the mixed water mass originating from Montreal's effluent, with 60% of the samples exceeding the indirect contact limit, and 40% exceeding the direct contact limit (Chapter 3, Figure S2A).

Identifying the sources and pathways of fecal contamination (FC) in large river systems requires advanced techniques like MST to differentiate between human and non-human sources (Harwood et al., 2013; Rock et al., 2015). Using specific primers and probes for the human, ruminant, gull, and swine, we found that the human marker was detected in 71.9% (95/132) of samples. The detection rate varied across different water masses, with the highest presence in the mixed water mass at 94% of samples, exceeding the RBT by 15%, confirming that the *E. coli* concentration in this water mass originates from human sources. Other markers were rarely detected, except for the gull marker, which was precent in 56% of samples (Chapter 3, Figure 3.7). In shallow waters, the human marker was found in 33% of the samples in the Upper St. Lawrence River (UPSLR), and 100% for Lower St. Lawrence River (LSLR). However, some sites showed high *E. coli* concentrations with low or no presence of the tested markers, suggesting that nearshore waters might be influenced by other sources.

The impact of wastewater discharge into a large river depends on the city's size, the wastewater treatment plant (WWTP) methods, and the river's size and flow rate (García-Armisen et al., 2014). Drury et al. (2013) found in their study of effluent influence

in freshwater ecosystems that WWTP effluent, in heavily urbanized regions, significantly alters the chemical and biological characteristics of the receiving ecosystem. Our findings have a significant implication for public health, environmental management, and policymaking. Persistent high levels of *E. coli* downstream from urban areas underscore the urgent need for improved wastewater treatment processes and stricter regulations to protect freshwater ecosystems and human health.

The correlation observed between human DNA copies and *E. coli* levels around Montreal serves as a clear indicator of the direct impact of human activities on water quality. While *E. coli* concentration decreased downstream, they did not reach acceptable safety limits (below 1,000 CFU and 200 CFU) until over 100 km downstream, towards Quebec City and encountering the estuarine environment. Understanding the distance required for *E. coli* levels to drop below safety threshold emphasizes the importance of continuous monitoring and management of water quality over large distances to ensure safer water for downstream communities. This not only protects human health but also supports recreational activities such as swimming and beach use, reducing health risk associated with waterborne pathogens and improve ecosystem resilience, supporting biodiversity and ecosystem services that benefit both human population and wildlife.

Addressing these challenges, our study highlights the urgent need to improve wastewater management strategies to mitigate FC in SLR and others large rivers. Additionally, our findings emphasize the importance of using MST to identify the origins of contamination.

4. 2 **Research recommendations**

In this study, we encounter several limitations, that need to be addressed. Starting with our sampling process, samples were collected annually during the summer, without replication at each site. Due to this single season sampling approach, we were unable to understand the river's dynamics throughout the year, which is crucial for assessing the impacts of seasonal events such as rainfall, ice melt, water level fluctuations, agricultural runoff (including fertilizer), bird migration, among others (Li et al., 2021). Future

investigations of a similar nature should improve methodological approaches by incorporating duplicate sampling for both *E. coli* and MST, alongside the incorporation of negative controls during water filtration (Edge et al., 2021). These measures would improve the reliability and consistency of the outcomes obtained.

The geographical extent of sampling of the SLR varied from year to year, with some years covering the entire SLR while other years focused on shorter sections. The extent of sampling creates data gaps for shallow waters along the UPSLR on the south shore, specifically in the United States, and the LSLR was not sampled concurrently onboard the research vessel *Lampsilis*. These variations could lead to potential discrepancies in the data or models developed. For instance, discrepancies may alter our estimate of the spatial distribution of FC, potentially leading to an underestimation or overestimation of contamination in areas not consistently sampled. Generalizing near shore areas to a certain water quality is challenging due to their more dynamic nature. A more homogeneous sampling approach could offer a clearer and more detailed picture of water quality in these areas. Additionally, timing could be a critical factor for accuracy of the temporal variations, as events such as rainfall and water level fluctuations can influence *E. coli* concentrations through runoff and dilution effects. However, this variability does not significantly impact our study's conclusions, as our data provides a comprehensive overview of *E. coli* contamination distribution across the different water masses.

Including additional environmental factors in the analysis, such as rain events, water levels, water flow, and turbidity would also be beneficial. Seasonal rainfalls often lead to a notable increase in fecal pollution, indicating a seasonal pattern in waterborne disease outbreaks (Li et al., 2021). Marking our samples as wet or dry for further analysis, would provide information about the impact of rain and runoff. These factors are essential for a comprehensive understanding of the river's dynamics and its influence on the survival or concentration of bacteria.

Additional funding to facilitate expanding the scope of this research is needed. With increased funding, we could invest in more markers to expand the research and

explore different sources, improve our analytical capabilities, and allow implementation of seasonal sampling to monitor changes through time. For instance, our study aimed to identify anthropogenic sources of FC, yet the observation of elevated *E. coli* concentration in areas lacking human influence or other markers used in our study suggests a need to incorporate wildlife markers. Identifying non-human sources of contamination will further contribute to understanding potential health risks to humans and inform management strategies. Also, by improving our data collection and analysis, we can fill existing knowledge gaps and contribute to the development and implementation of effective river management practices. This would support our goal of mitigating the impact of FC on river ecosystems, contributing to the health and safety of the communities that depend on these vital water resources.

4. 3 **Future Direction**

The rapid evolution of technology in DNA, digital PCR, software, markers primers and probes, among others, offers promising ways for advancing our research methodologies. Looking for continues updates is beneficial for enhancing data analysis and ensuring the adequacy of our protocols. In microbial source tracking, verifying the relevance and effectiveness of markers specific to our study area is essential (Harwood et al., 2013; Rock et al., 2015).

The role of wildlife in contributing to FC, as previously mentioned, is particularly important, especially in wetlands, understanding this aspect can give us important information on the health and biodiversity of riverine environments. Distinguishing between human and animal FC is critical, as human contamination poses unique risks, including the introduction of harmful bacteria and pollutants that affect both human health and ecological integrity. Literature suggests that the combination of various sources might necessitate adjusting the RBT for the human marker (Boehm et al. 2020). Reviewing the RBT in this region could provide insights into its accuracy, especially in scenarios involving multiple contamination sources.

Water-related diseases are significant sources of infection worldwide, as mentioned before, affecting human health and their link to climate changes can impact water availability and quality due the increase in human movement, population growth and the associated risk of pathogens, vectors, and infections. While scientist understand how climate may affect weather patterns, the consequences and influences on microbiological water quality are less studied. Waterborne diseases are sensitive to the environmental change, and climate change will alter rainfall events, storms, and droughts (Nichols et al., 2018). The rate of most ecological changes in ecosystems is slow, often over decades. Therefore, long term environmental monitoring is essential to identify ecological variables (Smeltzer et al., 2012). Such monitoring helps to detect any changes over time, as monitoring systems need to be maintained for long periods. When analysing results, is crucial to consider them in a historical context. Maintaining long term records is vital for assessing present and future policy decisions and identifying the consequences of drivers such as climate change, which may harm ecosystems, and long term data can help validate simulation models and test hypothesis (Burt et al., 2008). As we recommend, long term monitoring is critical to capturing the dynamic of the SLR, considering annual cycles and the impacts of climate change, population growth, and landscape alterations. This monitoring can provide valuable information on temporal trends and the effectiveness of management strategies.

Finally, there are promising opportunities for collaboration with researchers and institutions to further enrich this study. Partnering with experts in bacterial diversity can improve our understanding of microbial communities within the river ecosystem. This collaboration could help to understand how microbial contamination from different anthropogenic sources influences the diversity of natural bacterial communities in the SLR. Then, collaborating with the St. Lawrence River Institute offers an opportunity to explore ecological indicators through the River Rapport project. This project specializes in identifying different ecological indicators to assess the health of the UPSLR and sharing these findings with the community. These collaborations could expand our study and contribute to interdisciplinary exchanges, leading to more comprehensive results.

4. 4 River Management Implications

Water management is particularly susceptible to conflict due to the extensive use of resources by diverse and dispersed groups of people with different degrees of power and influence. The significance of water resources varies among these groups, often leading to shortages and unfair distribution (Furber et al., 2016). In recent years, researchers, communities, and public agencies have collaborated to protect and manage the SLR, facing new challenges annually (Farrell et al.,2007). This collaboration has led to international cooperation to establish regulations and standards protocols, such as the International Joint Commission (IJC), the St. Lawrence Action Plan (2011 - 2026), and the Great Lakes Water Quality Agreement (GLWQA), all having in common goal to restore and protect the waters. Considering that downstream communities often are the most affected by upstream actions or negligence, our study emphasizes the importance of equitable responsibility to ensure effective pollution management of FC.

The spatial variation in *E. coli* concentrations across large rivers, such as the SLR, is significantly influenced by different inputs along the river through human activities, as illustrated in Chapter 3 (Figure 3.2 and Figure 3.6). This anthropogenic impact highlights the challenges of managing a river that not only covers large geographical areas, where water masses and river units may change the way that microbial contamination migrates down the river, but also crosses different jurisdictional boundaries. Effective management of such resources requires a multijurisdictional approach, emphasizing collaboration and clear communication among stakeholders (Léonard et al., 2004). A holistic approach, where stakeholders prioritize the common good over individual interests, is also essential for addressing these complexities and ensuring responsible water quality management (Gupta, 2007). Based on this, it is important that FC is considered a priority in the SLR management strategy. This involves establishing joint monitoring programs and creating a centralized data-sharing platform. Additionally, engaging with different agencies, such as the Canada Water Agency, IJC, and St. Lawrence Action Plan, through agreements that commit to work to specific objectives related to improving water quality in the SLR, is crucial.

The concentration of *E. coli* as an indicator of FC and potential threats to human health and the ecosystem (Corsi et al., 2014), highlights the need to improve WWTP infrastructure to mitigate human-sourced contamination, managing agricultural inputs, with the recommendation of best practices and buffer zones, can reduce runoff from crops and livestock. Understanding the role of wildlife and wetlands in microbial contamination and remediation near these areas provides insights for more effective pollution control measures. Accurate identification of the sources not only helps pinpoint host specific pathogens but also determines control point and informs remediation strategies (Li et al., 2021). This can significantly reduce harmful bacteria levels entering the SLR, facilitating more accurate water quality monitoring. In our study, human activities along the river, such as urban runoff and WWTP, contribute differently to contamination at various points. Consequently, we identified that water masses and water units, such as island and lakes, significantly influence how microbial contamination, in this case, *E. coli* concentration, is transported downstream, affecting water quality. This distribution of contamination is not uniform across the river, resulting in areas with varying concentration of pollutants. This variability is important to consider when planning recreational activities, decision should be informed by localized contamination data, considering the traces present from upstream inputs. Areas close to WWTP or those affected by downstream trace pollution require more rigorous management strategies, informed by long term monitoring to ensure public health.

Looking toward the future, we need to consider the impacts of population growth and climate change on the health of the SLR and other rivers (ECCC, 2021). The responsibility for river conservation rest not only with environmental and governmental agencies but also with individuals and communities. Each of us plays a crucial role in preserving clean water and healthy ecosystems (De Rosa et al., 1999). Through this study we address some challenges that the SLR like others large rivers are facing, hoping to transform our findings into actionable strategies that improve the health and resilience of these vital waterways.

References

Ashbolt, N. J. (2004). Microbial contamination of drinking water and disease outcomes in developing regions. Toxicology, 198(1-3), 229-238.

Boehm, A. B., & Soller, J. A. (2020). Refined ambient water quality thresholds for human-associated fecal indicator HF183 for recreational waters with and without co-occurring gull fecal contamination. Microbial Risk Analysis, 16, 100139.

Burt, T. P., Howden, N. J. K., Worrall, F., & Whelan, M. J. (2008). Importance of long-term monitoring for detecting environmental change: lessons from a lowland river in south east England. Biogeosciences, 5(6), 1529-1535.

Corsi, S. R., Borchardt, M. A., Spencer, S. K., Hughes, P. E., & Baldwin, A. K. (2014). Human and bovine viruses in the Milwaukee River watershed: Hydrologically relevant representation and relations with environmental variables. Science of the Total Environment, 490, 849–860.

De Rosa, C. T., Rosemond, Z. A., Cibulas, W., & Gilman, A. P. (1999). Research management in the Great Lakes and St. Lawrence River basins: challenges and opportunities. Environmental research, 80(3), 274-279.

Drury, B., Rosi-Marshall, E., & Kelly, J. J. (2013). Wastewater treatment effluent reduces the abundance and diversity of benthic bacterial communities in urban and suburban rivers. Applied and Environmental Microbiology, 79(6), 1897–1905.

Edge, T. A., Boyd, R. J., Shum, P., & Thomas, J. L. (2021). Microbial source tracking to identify fecal sources contaminating the Toronto Harbour and Don River watershed in wet and dry weather. Journal of Great Lakes Research, 47(2), 366–377.

Environment and Climate Change Canada. (2021). Canadian environmental sustainability indicators : nutrients in the St. Lawrence River.

Farrell, J. M., Klindt, R. M., Casselman, J. M., LaPan, S. R., Werner, R. G., & Schiavone, A. (2007). Development, implementation, and evaluation of an

international muskellunge management strategy for the upper St Lawrence River. Environmental Biology of Fishes, 79, 111-123.

Furber, A., Medema, W., Adamowski, J., Clamen, M., & Vijay, M. (2016). Conflict management in participatory approaches to water management: A case study of Lake Ontario and the St. Lawrence River Regulation. Water, 8(7), 280.

García-Armisen, T., Inceoğlu, Ö., Ouattara, N. K., Anzil, A., Verbanck, M. A., Brion, N., & Servais, P. (2014). Seasonal variations and resilience of bacterial communities in a sewage polluted urban river. Plos one, 9(3), e92579.

Green, H. C., Haugland, R. A., Varma, M., Millen, H. T., Borchardt, M. A., Field, K. G., Walters, W. A., Knight, R., Sivaganesan, M., Kelty, C. A., & Shanks, O. C. (2014). Improved HF183 quantitative real-time PCR assay for characterization of human fecal pollution in ambient surface water samples. Applied and Environmental Microbiology, 80(10), 3086–3094.

Gupta, A. (Ed.). (2007). Large rivers: geomorphology and management. John Wiley & Sons.

Harwood, V. J., Staley, C., Badgley, B. D., Borges, K., & Korajkic, A. (2013). Microbial source tracking markers for detection of fecal contamination in environmental waters: Relationships between pathogens and human health outcomes. FEMS microbiology reviews, 38(1), 1-40.

Health Canada. (2012). Guidelines for Canadian recreational water quality, Third Edition. Water, air and Climate Change Bureau, Healthy Environments and Consumer Safety Branch, Health Canada, Ottawa, Ontario.

Hudon, C., & Carignan, R. (2008). Cumulative impacts of hydrology and human activities on water quality in the St. Lawrence River (Lake Saint-Pierre, Quebec, Canada). Canadian Journal of Fisheries and Aquatic Sciences, 65(6), 1165–1180.

Ichor, T., Umeh, E. U., & Duru, E. E. (2014). Microbial Contamination of Surface Water Sources in Rural Areas of Guma Local Government Area of Benue State, Nigeria. Journal of Medical Sciences and Public Health, 2(2), 43–51.

Léonard, N. J., Taylor, W. W., & Goddard, C. (2004). Multijurisdictional management of lake sturgeon in the Great Lakes and St. Lawrence River. In Sturgeons and paddlefish of North America (pp. 231-251). Dordrecht: Springer Netherlands.

Li, E., Saleem, F., Edge, T. A., & Schellhorn, H. E. (2021). Biological indicators for fecal pollution detection and source tracking: A review. In Processes, 9(11), 2058.

Mieszkin, S., Furet, J. P., Corthier, G., & Gourmelon, M. (2009). Estimation of pig fecal contamination in a river catchment by real-time PCR using two pig-specific Bacteroidales 16S rRNA genetic markers. Applied and environmental microbiology, 75(10), 3045–3054.

Mieszkin, S., Yala, J. F., Joubrel, R., & Gourmelon, M. (2010). Phylogenetic analysis of Bacteroidales 16S rRNA gene sequences from human and animal effluents and assessment of ruminant faecal pollution by real-time PCR. Journal of Applied Microbiology, 108(3), 974–984.

Nichols, G., Lake, I., & Heaviside, C. (2018). Climate change and water-related infectious diseases. Atmosphere, 9(10), 385.

Rock, C., Rivera, B., & Gerba, C. P. (2015). Microbial Source Tracking. In Environmental Microbiology: Third Edition (pp. 309–317). Elsevier Inc.

Ryu, H., Griffith, J. F., Khan, I. U. H., Hill, S., Edge, T. A., Toledo-Hernandez, C., Gonzalez-Nieves, J., & Domingoa, J. S. (2012). Comparison of gull feces-specific assays targeting the 16S rRNA Genes of Catellicoccus marimammalium and Streptococcus spp. Applied and Environmental Microbiology, 78(6), 1909–1916.

Smeltzer, E., d Shambaugh, A., & Stangel, P. (2012). Environmental change in Lake Champlain revealed by long-term monitoring. Journal of Great Lakes Research, 38, 6-18.

Zandaryaa, S., & Mateo-Sagasta, J. (2018). Chapter 8. Organic matter, pathogens and emerging pollutants.