Relationship between toxic cyanobacterial blooms, physico-chemical factors and potential multiple source excreta contamination in an affected watershed.

Valentina Muñoz Ramos
Department of Natural Resource Sciences
McGill University, Montreal
April 2013

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Master of Science

Abstract

Missisquoi Bay (MB) is a temperate eutrophic freshwater ecosystem located in an agricultural watershed and it frequently experiences toxic *Microcystis*-dominated cyanobacterial blooms. Cyanobacterial population dynamics are influenced by a plethora of factors that may differ from system to system, requiring a site-specific assessment of bloom-promoting factors to design more effective bloom prevention or remediation strategies. This is the first biomonitoring study that combined data from high-throughput 16S rRNA gene amplicon sequencing, qPCR and environmental parameters from temporal and spatial samples to identify the main bloom-promoting factors. In addition, high-throughput amplicon sequencing of mitochondrial DNA genes was performed to qualitatively identify potential external sources of nutrients originating from animal excreta. Particular emphasis was placed on 1) determining whether there was a link between nutrients from external sources and cyanobacterial blooms and 2) analyzing in situ the effect of environmental factors (particularly nitrogen (N) and phosphorus (P) concentrations) on the dynamics of cyanobacterial community composition, abundance and toxicity.

The concentrations of total P (TP) and total N (TN) in MB in 2009 correlated significantly with the abundance of total cyanobacterial cells, the *Microcystis* 16S rRNA and *mcyD* genes and intracellular microcystin. The results suggest that external sources of nutrients, such as surface runoff and animal excreta, played a significant role in the load of nutrients into the bay and thus in the proliferation of toxic cyanobacterial blooms. This was indicated by the detection of non-aquatic mitochondrial DNA hosts in the bay and the relationships between the pattern of surface runoff, nutrient concentrations, *E. coli* counts and total cyanobacterial abundance. Potential sources of nutrients from non-aquatic animal excreta in the system comprised rodents, birds, cattle and humans, indicating that efforts are required to control pollution from animal excreta in MB.

During the growing season, the major cyanobacterial taxa were members of the orders *Chroococcales* and *Nostocales*. The genus *Microcystis* was identified as the main *mcyD*-carrier and main microcystin producer, hence the most problematic taxon in the cyanobacterial bloom. The correlations observed with environmental parameters suggest that increasing nutrient concentrations and TN:TP (mass) ratios approaching 11:1,

coupled with an increase in temperature, promoted *Microcystis*-dominated toxic cyanobacterial blooms. Although the importance of nutrient ratios and absolute concentrations on cyanobacterial and *Microcystis* dynamics has been documented, this is the first time that an optimum TN:TP ratio for *Microcystis* dominance has been observed in the field. This observation provides further support to the theory that nutrient supply ratios are an important determinant of species composition in natural phytoplankton assemblages. Although the validity and prediction potential of this optimum ratio for *Microcystis* dominance has yet to be verified through longer-term studies, it may provide practical guidelines for nutrient management strategies to avoid the proliferation of this toxin producing cyanobacterial genus in MB.

Résumé

La baie Missisquoi (BM) est un écosystème d'eau douce tempérée situé dans un bassin versant où les activités agricoles sont importantes. La baie est fréquemment exposée à la prolifération de *Microcystis,* des cyanobactéries qui ont un potentiel toxique. La dynamique des populations de cyanobactéries est influencée par une multitude de facteurs qui peuvent varier d'un site à l'autre. Une caractérisation spécifique des facteurs qui sont responsables des floraisons est donc nécessaire afin de développer des mesures préventives et des stratégies de remédiation efficaces.

C'est la première fois qu'une étude de biosurveillance combine des données de séquençage d'amplicons ARN ribosomal 16S à haut débit, la PCR en temps réel et des paramètres environnementaux d'échantillons qui ont été prélevés de façon spatio-temporelle pour identifier les facteurs qui causent les floraisons. De plus, le séquençage à haut débit d'amplicons ciblant l'ADN mitochondrial a été utilisé pour identifier qualitativement les sources potentielles de nutriments qui proviennent d'excréments d'origine animale. Une insistance a été accordée plus spécifiquement à 1) l'évaluation de l'existence d'un lien entre les nutriments provenant de sources externes et les périodes de floraison de cyanobactéries et 2) l'analyse in situ des impacts des facteurs environnementaux (particulièrement les concentrations d'azote (N) et de phosphore (P)) sur la structure des communautés de cyanobactéries, leur abondance et leur toxicité.

Les concentrations de P et de N à la baie Missisquoi en 2009 étaient fortement corrélées avec l'abondance des cellules de cyanobactéries, des gènes de *Microcystis* (ARNr 16S et *mcyD*) ainsi qu'avec la concentration de microcystine intracellulaire. Les résultats suggèrent que les sources externes de nutriments telles que le ruissellement de surface et les excréments d'origine animale ont contribué de manière importante à la charge d'éléments nutritifs dans la baie et ainsi à la prolifération de cyanobactéries qui produisent des toxines. Ces conclusions ont été corroborées par la présence d'ADN mitochondrial d'hôtes non aquatiques et les liens entre les types de ruissellement de surface, les concentrations de nutriments, et les dénombrements de *E.coli* et de cyanobactéries. Les sources potentielles de nutriments provenant des excréments animaux dans la baie incluaient les rongeurs, les oiseaux, le bétail et les humains. Ces résultats indiquent que des efforts sont requis dans la baie Missisquoi pour contrôler la pollution

provenant des excréments d'origine animale.

Pendant la période de floraison, la communauté de cyanobactéries était principalement composée de *Chroococcales et Nostocales*. Les cyanobactéries du genre *Microcystis* étaient à la fois les principales porteuses du gène *mcyD* et les principales productrices de microcystine. Ce taxon a donc été le plus problématique pendant les fleurs d'eau. Les corrélations observées entre les facteurs environnementaux suggèrent que la dominance des *Microcystis* a été associée à l'augmentation des concentrations en nutriments, à un ratio TN:TP (masse) d'une valeur aux environs de 11:1, combinée à une hausse de température.

Même si l'importance des concentrations et des ratios de nutriments sur la dynamique des cyanobactéries et des *Microcystis* a déjà été documentée, c'est la première fois qu'un ratio optimum pour le genre *Microcystis* est observé sur le terrain. Cette observation est en accord avec la théorie voulant que les ratios de nutriments soient importants pour déterminer la composition des espèces faisant partie de la communauté phytoplanctonique.

La validité et le potentiel de prédiction de ce ratio optimum pour la dominance des *Microcystis* devraient être vérifiés à l'aide d'études à long terme, car il pourrait devenir un critère important pour développer des stratégies de gestion de nutriments qui permettraient d'éviter la prolifération de ce genre de cyanobactéries nocives.

Acknowledgments

I would like to thank my supervisor, Dr. Charles Greer, for opening the doors for me into the world of molecular biology, which I was able to learn and explore in the great environment of your lab. Thank you very much for the opportunity to study cyanobacterial blooms, a topic I soon learnt to be passionate about. I also thank you for all your input, editing suggestions and stipend. Thanks also to my co-supervisor Dr. Lyle Whyte for this opportunity and valuable feedback after my seminar presentations.

I would like to specially thank Nathalie Fortin, I greatly appreciate all the techniques that you gracefully, patiently and thoroughly taught me, including field sampling, preparation of samples for further nutrient, toxin and DNA analyses, *E. coli* counting, DNA extraction, PCR and qPCR. Thank you for collecting the 2009 samples and for collecting the nutrient and toxin data from external contributors. Thank you also for your constant support and advice, for your valuable input and editing during my thesis writing, your help during the translation of my abstract into French, as well as for your smile that so many times lifted my spirits.

Thanks Dr. Étienne Yergeau for always having your door open to provide advice. Thank you very much for listening to my ideas and for motivating me to pursue my scientific curiosity. Thank you for all your advice on statistics, for introducing me into the world of bioinformatics and for all the suggestions during the sequencing data analysis. In addition, thank you for your help during the translation of my abstract into French. I would like to let you know that your constant optimism and passion for the field were inspiring to me.

Thanks Dr. Terrence Bell for all your technical, academic and moral support throughout my masters. I specially appreciate your valuable input prior to every presentation and during my thesis-writing period. Besides providing me with guidance, your comments and suggestions certainly helped me to improve my scientific thinking.

Thanks to all the research officers and technical officers of the Greer lab, your technical expertise, assistance and advice helped me to improve my skills in numerous instances. I would like to specially thank Julie Champagne, Miria Elias, Dr. David Juck, Dr. Louis Jugnia, Dr. Luke Masson, Minh Vuong, Diane Labbé, Suzanne Labelle, Claude Masson, Christine Maynard, and Sylvie Sanschagrin. Suzanne, thank you for helping me

with data compilation and abstract translation. Julie and Christine, thank you for the sequencing analysis. Miria, I would like to say that the month that I had the opportunity to work by your side was a great experience mainly due to your optimism, resourcefulness and charisma. Despite the fact that in the end the Luminex data was not included in my thesis, I was able to profit a lot from your knowledge while optimizing that technique. It was a pleasure to learn from you. Diane, thank you for your input, technical assistance and advice. I will always remember our days in the lab and our Spanish-French interactions with a smile on my face.

Working in this lab also provided me with the opportunity to meet wonderful people from all over the world and make new and lasting friendships. Christine, Lia, Silvia, Paola, Nina, Tuana, Katrina, Terry and Étienne, the ups and downs during these years were spiced up with fun thanks to all of you. I would also like to thank my friend Lina Arcila, our conversations about biology resulted in epiphanies for which I will be forever grateful.

Special thanks to my family for always being there and providing me with encouragement at all times.

Tuan, this experience was more meaningful and enjoyable with you by my side. Thank you so much for you tolerance and support throughout this important period in my life.

I would also like to thank the Natural Sciences and Engineering Research Council of Canada (NSERC), the Max Bell Foundation and McGill University for their financial support.

The numerous examples of endurance, tenacity and perseverance existent in my family have been a source of inspiration to me and have given me the strength to get back on my feet during difficult times. Beloved family, despite of the vast geographical distance that separates us, your constant support and love have accompanied me everyday of my life. You have never left my heart.

This work is dedicated to you.

Contribution of the author and other researchers

The candidate designed novel primers used for high-throughput sequencing of partial 16S rRNA cyanobacterial genes employing the Ion Torrent platform and for quantification of *Microcystis* 16S rRNA genes via qPCR. The candidate extracted the DNA, performed the PCR, most of the qPCR analyses and produced the amplicons for sequencing and conducted all the data analyses. Julie Champagne provided part of the *mcyD* qPCR data. Christine Maynard and Julie Champagne conducted the sequencing for the study.

Part of this research branches out from a larger multidisciplinary epidemiological study. Therefore, this project makes use of data produced by other researchers. The identification and enumeration of cyanobacterial species were performed by Irina Moukhina from the laboratory of Dr. David Bird at UQAM. Dr. David Bird also provided the nutrient and microcystin (ELISA) data. The *E. coli* count data was produced by Nathalie Fortin. Sampling and filtrations were performed by the NRC field team and by students under the supervision of Dr. David Bird and Dr. Benoit Lévesque.

The candidate wrote all the content of this thesis, with valuable suggestions provided by Dr. Charles W. Greer, Nathalie Fortin, Dr. Luke Masson, Minh Vuong and Dr. Lyle G. Whyte. Dr. Étienne Yergeau and Dr. Terrence Bell provided guidance and suggestions during data analysis and writing.

Table of contents

Abstract II
Résumé IV
Acknowledgments VI
Dedication VIII
Contribution of the author and other researchers IX
Table of contents X
List of tables XII
List of figures XIII
List of abbreviations XV
Chapter 1 - Introduction 1
Problematic: Harmful cyanobacterial blooms 1
Study site: Missisquoi Bay of Lake Champlain 2
Research Objectives and Outline 4
Chapter 2 - Literature Review 5
2.1 Cyanobacteria, adaptations and bloom-forming taxa 5
2.2 Factors influencing cyanobacterial growth and dominance 6
2.2.1 Physico-chemical factors 6
Nutrients 6
Phosphorus and nitrogen concentrations 7
Nitrogen to phosphorus ratios (N:P) 8
Temperature 11
Light 12
pH 12
2.2.2 Biotic factors 13
2.3 Cyanobacterial toxicity 13
2.3.1 Microcystin toxicity and water quality guidelines 14
2.3.2 Microcystin producing genera 14
2.3.3 Microcystin chemical structure and stability 15
2.3.4 Microcystin biosynthetic mechanism, evolution and potential functions 17
2.3.5 Factors influencing microcystin production 18
2.4 Recent advances in molecular biological techniques and their relevance to cyanobacterial ecology studies 20
2.4.1 Characterization of cyanobacterial diversity during algal blooms 20
2.4.2 Detection and quantification of genotypes with quantitative PCR 21
2.5 identification of external sources of nutrients 22
2.5.1 Molecular methods for the direct identification of animal sources of excreta contamination 22
2.5.2 *E. coli* as a potential indicator of recent nutrient input from external sources 23
Chapter 3 - Material and Methods 25
3.1 Sampling sites 25
3.2 Sample collection 25
3.3. DNA sample preparation, extraction and quantification 27

3.3.1 Quantitative PCR (qPCR)....27
3.3.1.1 Assessment of PCR inhibition....27
3.3.1.2 Quantitative PCR of Microcystis 16S rRNA gene....28
3.3.1.3 Quantitative PCR of microcystin gene....31
3.3.2 High-throughput sequencing of PCR amplicons with the Ion Torrent platform....31
3.3.2.1 PCR amplification of partial cyanobacterial 16S rRNA and mtDNA genes....32
3.3.2.2 Gene amplicon sequencing....33
3.3.2.3 Bioinformatic analyses....34
3.4 Determination of sample microcystin concentration (ELISA)....35
3.5 Cyanobacterial abundance, *E. coli* counts and environmental parameters....35
3.6 Data and statistical analyses....36

Chapter 4 – Results....38
4.1 General description of environmental conditions in MB/Pike River area during 2009....38
4.2 Nutrient input and its relation with indicators of fecal contamination and cyanobacterial blooms in Missisquoi Bay/Pike River during 2009....43
4.2.1 Spatial analysis of phosphorus and nitrogen concentrations and their relationship with rainfall....43
4.2.2 *E. coli* as a potential indicator of recent nutrient input from external sources....48
4.2.3 Potential sources of nutrients into Missisquoi Bay/Pike River area from animal excreta....50
4.2.4 Relationship between toxic cyanobacterial blooms, nutrient input and *E. coli* in Missisquoi Bay during 2009....51
4.3 Spatio-temporal characterization of major cyanobacterial taxa and their association with environmental parameters in Missisquoi Bay/Pike River during 2009....51
4.3.1 Significant relations between the relative abundance of major cyanobacterial taxa and environmental parameters in Missisquoi Bay....57
4.3.2 Succession of major cyanobacterial taxa and environmental parameters in littoral station BM5 and pelagic station Pel2....59
4.3.3 Relationship between *Microcystis* relative abundance and nutrient concentrations and ratios....64
4.4 Spatio-temporal dynamics of cyanobacterial toxicity in Missisquoi Bay during 2009.68
4.4.1 Significant relationships between microcystin concentrations, total cyanobacterial cells and biomolecular markers for total and toxic cyanobacteria and *Microcystis* abundance....68
WHO water quality guidelines and microcystin concentrations in Missisquoi Bay in 2009....68
4.4.2 Significant relationships between microcystin concentrations, *Microcystis* 16S rRNA gene, estimated total cyanobacteria 16S rRNA gene, total cyanobacterial cells, *mcyD* gene and environmental parameters....71
4.4.3 *Microcystis* 16S rRNA gene, estimated total cyanobacteria 16S rRNA gene, *mcyD* gene, microcystin concentrations and their relationship with nutrient ratios....73

Chapter 5 – Discussion and Conclusions....77
5.1 Nutrient input and its relation with indicators of animal excreta contamination and cyanobacterial blooms in Missisquoi Bay/ Pike River during 2009....77
5.2 Major cyanobacterial taxa in Missisquoi Bay....81
5.2.1 High-throughput sequencing of partial 16S rRNA genes....81
5.2 2 Factors determining the relative abundance of major cyanobacterial taxa in Missisquoi Bay....84
5.3 Cyanobacteria and microcystin dynamics in Missisquoi Bay during 2009....90

5.3.1 Significant relationships between microcystin concentrations and *mcyD* gene, *Microcystis* 16S rRNA gene, estimated total cyanobacteria 16S rRNA gene, and total cyanobacterial cells. 90
5.3.2 Correlations between toxin concentrations, *mcyD*, 16S rRNA genes, total cyanobacterial cells and environmental factors. 92
5.4 Conclusions 97

List of references 101

Appendices 117
Appendix A – Primer design and validation for partial cyanobacterial 16S rRNA gene amplicon sequencing with Ion Torrent 117
Appendix B – Primer design and validation for partial *Microcystis* 16S rRNA gene amplification with qPCR 122
Appendix C – Target animals for mitochondrial DNA sequencing analysis 127

List of tables

Table 2.1 – Members of the following genera have been identified to produce microcystin in natural populations and/or in monocyanobacterial cultures 15
Table 3.1 – Sampling stations in Missisquoi Bay and tributaries 25
Table 3.2 – Primer sequences used in this study 29
Table 4.1 – Ranges of total and dissolved nutrient concentrations recorded in Missisquoi Bay and tributary stations during the 2009 sampling period. 42
Table 4.2 – Mean sampling season nutrient concentrations in the tributaries, littoral and pelagic stations of the Missisquoi Bay/Pike River area 47
Table 4.3 – Relationships between increasing contact of the station type to the shore and nutrient concentrations as determined by Cochran–Mantel–Haenszel (CMH) and Spearman Rank correlations. 47
Table 4.4 – Spearman Rank correlations between Rainfall (rain d-7) and nutrient concentrations at each station type in the Missisquoi Bay/Pike River area in 2009 ..47
Table 4.5 – Relationships between *E. coli* counts and nutrient concentrations as determined by Spearman Rank correlation on the 2009 data from the Missisquoi Bay/Pike River area. 49
Table 4.6 – Significant relationships between the relative abundance of cyanobacterial taxa and physicochemical parameters and ranges of these parameters in which the taxa was observed to dominate (determined independently for each parameter). 58
Table 4.7 – Averages of environmental parameters during the 2009 sampling period in littoral station BM5 and pelagic station Pel2. 63
Table 4.8 – Nutrient concentrations in the MB on July 25/09. 65
Table 4.9 – Significant relationships between *Microcystis* 16S rRNA gene, estimated total cyanobacteria 16S rRNA gene, *mcyD* gene, and microcystin concentrations. ..68
Table 4.10 – Significant relationships of microcystin, *mcyD* gene, *Microcystis* 16S rRNA gene concentrations with environmental parameters and minimum values of these parameters at which the toxin or biomolecular markers were observed to reach the specified concentrations 72

Table 4.11 – Significant relationships of total cyanobacterial cells (microscopy) and estimated total cyanobacterial 16S rRNA gene concentrations with environmental parameters and minimum values of these parameters at which the indicators of cyanobacterial abundance were observed to reach the specified concentrations.....73
Table B.1 – Effect of annealing temperature on the specificity of the primers UcyaF3/R3 for its intended target, *Microcystis*.....125
Table C.1 – Target animals for mitochondrial DNA analysis 127

List of figures

Figure 2.1 – Resource gradient in competition theory.....10
Figure 2.2 – Chemical structure of microcystin-LR. 16
Figure 3.1 – Sampling stations in Missisquoi Bay and tributaries.....26
Figure 4.1 – Total cyanobacterial cells/ml in Missisquoi Bay during 2009. 39
Figure 4.2 – Intracellular and extracellular microcystin in Missisquoi Bay in 2009.....39
Figure 4.3 – *mcyD* and *Microcystis* 16S rRNA gene dynamics in MB stations during 2009. The error bars correspond to the standard deviations (n=3). 40
Figure 4.4 – Nutrient concentrations over time in a) Littoral station BM5 and b) Pelagic station Pel2. Error bars represent standard deviations. 41
Figure 4.5 – TN:TP ratio over time in Missisquoi Bay and tributary stations.....42
Figure 4.6 – Total phosphorus and its relationship with rainfall and station type. 44
Figure 4.7 – Dissolved phosphorus and its relationship with rainfall and station type. ...45
Figure 4.8 – Total nitrogen, and its relationship with rainfall and station type in 2009 samples. The arrow represents the positive trend observed between TN and station contact with the shore (Table 4.3). Asterisks on top of the bars denote whether a significant correlation was found between rain and TN at each station type (Table 4.4). (*) $p<0.05$; (**) $p<0.01$. 46
Figure 4.9 – Dissolved nitrogen, and its relationship with rainfall and station type in 2009 samples. The arrow represents the positive trend observed between DN and station contact with the shore (Table 4.3). Asterisks on top of the bars denote whether a significant correlation was found between rain and DN at each station type (Table 4.4). (*) $p<0.05$; (**) $p<0.01$. 46
Figure 4.10 – *E. coli* counts, and its relationship with rainfall and station type in 2009 samples. The arrow represents the positive trend observed between *E. coli* counts and station contact with the shore. Asterisks on top of the bars denote whether a significant correlation was found between rain and *E. coli* counts at each station type. (*) $p<0.05$; (**) $p<0.01$.....48
Figure 4.11 – Presence of non-aquatic mitochondrial DNA hosts identified to the genus level in Missisquoi Bay/Pike River area. 50
Figure 4.12 – Temporal variation of *Microcystis* in relation to cyanobacterial orders in the littoral station BM5 during 2009.....53
Figure 4.13 – Temporal variation of *Microcystis* in relation to cyanobacterial orders in the Pelagic station Pel2 during 2009.....54
Figure 4.14 – Temporal variation of *Microcystis* in relation to cyanobacterial orders in the tributary station Pike River during 2009.....55

Figure 4.15 – Spatial variation of *Microcystis* in relation to cyanobacterial orders in July 25, 2009 in the different stations in Missisquoi Bay.....56
Figure 4.16 – Temporal profile of the dynamics of environmental and biomolecular parameters in littoral station BM5. 60
Figure 4.17 – Temporal profile of the dynamics of environmental and biomolecular parameters in pelagic station Pel2.....62
Figure 4.18 – TN:TP mass ratio over time in the littoral station BM5 and the pelagic station Pel2 of Missisquoi Bay in 2009. The vertical lines indicate the period from July 15 to August 19. 64
Figure 4.19 – Spatial variation of *Microcystis* in relation to cyanobacterial orders in MB on July 25, 2009 based on increasing TP concentrations. 66
Figure 4.20 – Spatial variation of *Microcystis* in relation to cyanobacterial orders in MB on July 25, 2009 based on increasing nutrient ratios. 67
Figure 4.21 – Relationship between *Microcystis* and nutrient ratios in MB stations at water temperatures equal to or higher than 20°C.....67
Figure 4.22 – Relationship between microcystin and the *mcyD* gene, total cyanobacterial cells and *Microcystis* 16S rRNA gene copies. 70
Figure 4.24 – Estimated total cyanobacteria and *Microcystis* 16S rRNA genes, *mcyD* gene and intracellular microcystin in relation to the TN:TP mass ratio on July 25th, 2009 in Missisquoi Bay. The error bars correspond to the standard deviation (n=3) for the molecular markers (genes). 76
Figure 4.25 – Ratio of *mcyD* copies and estimated *Microcystis* cells on July 25th, 2009 in Missisquoi Bay. It was assumed that *Microcystis* carries two 16S rRNA operons per cell. The error bars represent the relative error of the ratio between *mcyD* (n=3) and *Microcystis* 16S rRNA gene (n=3) qPCR data. 76
Figure A.1 – Comparison of the estimated total cyanobacterial 16S rRNA copies/ml determined based on order and genus. The error bars indicate the standard deviations (n=3).....120
Figure A.2 –Comparison of the percentages of specified genera out of total cyanobacteria determined based on order (a) and genus (b) in sampling station BM5. See text for details. 121

List of abbreviations

BLAST	basic local alignment search tool
CFU	colony forming units
CHAB	cyanobacterial harmful algal blooms
DNA	deoxyribonucleic acid
FIB	fecal indicator bacteria
MB	Missisquoi Bay
MID	multiplex identifier
mtDNA	mitochondrial DNA
NGS	next generation sequencing
PCR	polymerase chain reaction
pLuxA	luxA plasmid DNA
qPCR	quantitative PCR or real time PCR
RDP	Ribosomal Database Project
rRNA	ribosomal ribonucleic acid

Chapter 1 - Introduction

Problematic: Harmful cyanobacterial blooms

Phytoplankton, the autotrophic guild of plankton, provide a wide range of ecological services: they produce and replenish oxygen levels in the waters they inhabit while removing carbon dioxide, contribute substantially to total global primary production (Horner and Schrader 1982; Field et al. 1998), and remove nutrients from water (Fisher et al. 1988). However, excess growth of phytoplankton can lead to undesirable consequences. When proliferation and accumulation of phytoplankton results in a visible discoloration of the water, it is usually referred to as a bloom (Paerl et al. 2001). Blooms can negatively affect the health of the environment (Davis 1964) and in turn impact the economy, since the resulting deterioration in water quality hampers tourism and recreational activities, leading to declines in fisheries productivity (Anderson et al. 2000), contamination of potable water supplies and complications in the water treatment processes (Ghernaout et al. 2010). Cyanobacteria, the prokaryotic members of the phytoplankton, are important components of blooms worldwide, and a reason for concern due to the capability of some taxa to produce toxins that can directly affect human and animal health (Falconer 1999). This further exacerbates the negative impacts of blooms on human, environmental, and economic health.

Cyanobacterial blooms have been observed in freshwater, estuarine, and marine environments, and can develop at the surface (Fortin et al. 2010) or at other depths (Jacquet et al. 2005) in the water column. Blooms can result from an unbalanced situation between bottom-up control factors, such as nutrient and light availability (Geider 1987; Davis et al. 2010), and top-down control factors such as grazing and predation (Gobler et al. 2007). These, among other factors, in combination with the physical hydrological regime of the body of water, such as vertical mixing (Huisman et al. 2004) and retention time (Soares et al. 2012), provide a wide set of chemical, physical and biotic variables that interplay (Paerl 1988), resulting in the differences in incidence, diversity and successional patterns in cyanobacterial blooms that have been observed from temperate to tropical systems worldwide (Rinta-Kanto et al. 2009; Liu et al. 2011; Te and Gin 2011; Al-Tebrineh et al. 2012; Sitoki et al. 2012; Werner et al. 2012). With rising global

temperatures and predicted changes in precipitation patterns that may lead to increased nutrient loading from surface runoff, the dominance of cyanobacteria and the incidence of toxic blooms are expected to increase (Davis et al. 2009; Te and Gin 2011; Paerl and Paul 2012; Schindler 2012).

Study site: Missisquoi Bay of Lake Champlain

Lake Champlain is the sixth largest inland freshwater body in the United States; its basin is a 21,150 km^2 drainage area (Shanley and Denner 1999) that encompasses regions of Vermont (56%), New York (37%) and Quebec (7%) (USEPA 1997). Occupying only 1,136 km^2 (5.4%) of its basin (Shanley and Denner 1999), Lake Champlain flows north from Whitehall, New York towards the Richelieu River in Quebec, where it joins the St. Lawrence River. The lake also flows south through a canal built in 1823, connecting it to the Hudson River, which drains into the Atlantic Ocean at New York. This study was conducted in the Missisquoi Bay (MB) area of Lake Champlain and the Pike River. This bay is located to the north of the Northeast Arm of Lake Champlain and has a surface area of 77.5 km^2 (Simoneau 2007). MB extends from the delta of the Missisquoi River north into the province of Quebec, and it flows south into the Inland Sea (Northeast Arm). The watersheds of the Missisquoi, Pike and Rock Rivers drain directly into this bay. With marsh-like qualities, Missisquoi Bay is shallow (4.3m), relatively warm, well mixed, unstratified and has restricted circulation (USEPA 1997; Facey et al. 2012). MB waters play a key role in local recreational and economic activities. In addition, Lake Champlain represents the major drinking water source for some towns and cities in the province of Quebec and in the states of Vermont and New York (USEPA 1997). This is a concern because MB frequently exhibits toxic cyanobacterial blooms during the summer (Smeltzer et al. 2012).

Over the period of 1964 to 2009, August mean surface temperatures in Lake Champlain increased by 1.6–3.8 °C (Smeltzer et al. 2012). MB remained mesotrophic until roughly the 1970s (Levine et al. 2012). However, from 1979-2009, despite the fact that the total nitrogen (TN) concentrations decreased (~25%), total phosphorus (TP) concentrations increased by 72% (20 µg/l), chlorophyll a concentrations doubled, and cyanobacterial dominance increased (Smeltzer et al. 2012). Toxic *Microcystis* dominated

blooms are now frequent and microcystin concentrations (~ 5 μg/l) that exceed WHO drinking-water quality guidelines have been previously reported in MB (Davis et al. 2009; Fortin et al. 2010). It has been suggested that the shift to cyanobacterial dominance within the phytoplankton community in MB may have been the result of increasing water temperatures, increasing P loads from land-use changes and decreasing TN levels, resulting in a decrease in the TN:TP ratio (Smeltzer et al. 2012).

Because of the large ratio of the MB watershed to lake surface, the land use of this watershed heavily impacts the water quality of the lake (Beck et al. 2012). In the watershed, nutrients find their way into the water from discernible, well defined and discrete sources such as industrial, commercial or sewage treatment plants (i.e. point sources), or from many diffused sources as a result of rainfall or snowmelt that generate land runoff (i.e. non-point sources). Both Quebec and Vermont have taken steps towards reducing P loading into MB from wastewater and non-point sources, as summarized by Beck et al. (2012). In an effort to further control and reduce P loading into MB from point and non-point sources and meet the guidelines of the EPA Clean Water Act, it was determined in 1993 and reaffirmed in 2002, that the annual in-lake TP concentrations for the euphotic zone in MB needed to be reduced to 25 μg/l (NRD 2008; Beck et al. 2012). At the present, point-sources account for only 4-5% of the total P loading to MB from Quebec, while non-point sources are regarded as the main contributors to high P levels in the lake (Beck et al. 2012). Although only ~25% of the MB watershed is devoted to agricultural practices, 65 to 70% of the total annual P load is attributed to agricultural runoff (Troy et al. 2007). In the Quebec portion of the watershed, which represents 42% of the whole MB watershed (Beck et al. 2012), the basin territory is comprised of forest (59%), agriculture (33%) and urban (1%) areas, while the surface water accounts for only 7% (Beck et al. 2012). In terms of agriculture, by 2001, 56% of the planted areas were large-scale crops such as corn (Simoneau 2007), which represents a relatively elevated risk for erosion (Gasser et al. 2010). Livestock units (44,881) were mainly composed of pigs (49%) and cattle (39%), generating large quantities of manure. Wildlife presence in the area is also important: much of the surroundings of the bay are located within the Missisquoi Wildlife Refuge, and the bay is also in the trajectory of the North Atlantic Flyway. Taking into account that the temporal distribution of precipitation intensity and

snowmelt in Quebec is likely to lead to saturation-excess runoff (Michaud 2004), and considering the fact that fecal matter and urine are nutrient rich, the pattern of land use in this watershed suggests that animal excreta could be an important non-point source of nutrients entering the bay via surface runoff.

In view of the heterogeneity of the stressors and responses observed from system to system during algal blooms, a site-specific characterization of the cyanobacterial taxa involved in the bloom, their potential toxicity, as well as an *in situ* analysis of the response of cyanobacterial community composition to environmental variation are essential in order to identify bloom-promoting environmental factors (Huisman and Hulot 2005). Knowledge of the site-specific bloom-promoting factors combined with the identification of external sources of nutrients will contribute to the design of more effective management strategies for the abatement of toxic cyanobacterial blooms in this affected area.

Research Objectives and Outline

The objective of this project was to expand knowledge of the sources of nutrients that entered Missisquoi Bay of Lake Champlain during 2009 and to analyze, *in situ*, the effect of nutrients and other physico-chemical parameters on the dynamics of cyanobacterial community composition, abundance and toxicity in this ecosystem. The underlying hypothesis for this study is that once optimal temperatures are established, cyanobacterial dynamics are mainly determined by nutrient concentrations in MB. The project had 2 main objectives:

1) To analyze whether external sources of nutrients had an effect on the cyanobacterial bloom.
2) To identify major cyanobacterial taxa, potential microcystin producers and elucidate their dynamics in response to the changing environmental conditions, particularly nutrient variation, in MB.

Chapter 2 - Literature Review

2.1 Cyanobacteria, adaptations and bloom-forming taxa

Cyanobacteria comprise a diverse group of oxygenic photosynthetic prokaryotes that originated more than 2.5 billion years ago (Summons et al. 1999; Schopf 2002). An ancient origin (Amard and Bertrand-Sarfati 1997) and their subsequent evolutionary history have provided cyanobacteria with a range of biochemical, physical, biotic and ecological adaptations that allow them to thrive under varied and extreme environmental conditions (Paerl et al. 2001). Some of these adaptations include: cellular nutrient (e.g., phosphorus, nitrogen, carbon) storage (Kromkamp 1987), carbon concentrating mechanisms (Song and Qiu 2007) and nitrogen fixation (Haselkorn 2007) that allow them to thrive under nutrient deficient conditions; siderophores that enable metal sequestration (Itou et al. 2001); production of a mucilaginous sheath and other extracellular polysaccharides (Leak 1967) to tolerate desiccation and freezing (Tamaru et al. 2005); gas vesicle production that allows for buoyancy regulation (Brookes and Ganf 2001), allowing them to take advantage of high levels of irradiance at the surface and nutrient pools at lower depths; optimum growth rates at higher temperatures (generally >25°C) than eukaryotic algae (Robarts and Zohary 1987); photoprotection (Mandal et al. 2011) and/or chromatic adaptation (Kehoe 2010) by virtue of accessory pigments allowing them to tolerate high irradiance at the water surface and/or capture photons at wavelengths not efficiently trapped by chlorophyll in deeper regions of the water column.

In aquatic environments, surface blooms can rapidly form when cells have high growth rates and specially during periods of water column stability in the case of cyanobacterial cells that produce gas vesicles (Paerl et al. 2001). For these species to persist through time, their growth rate, which depends on nutrient supply, light and temperature (Hecky and Kilham 1988), must exceed or at least be equal to losses due to senescence, grazing, dilution or sedimentation (Reynolds 1984). The pattern of cyanobacterial succession in aquatic environments depends on the capability of each taxon to thrive under the changing environmental conditions. Hence, during the progression of a bloom, several strains or species of cyanobacteria may alternately dominate the water body (Paerl et al. 2001). Major bloom-forming cyanobacterial taxa exhibit different combinations of the adaptations previously mentioned, allowing them to

establish dominance under harsh environmental conditions, including high competition for nutrients, light availability and the high pH levels (Paerl 1988; Rantala et al. 2006; Liu et al. 2011) associated with blooms events. Cyanobacteria bloom-forming genera include: *Anabaena, Aphanizomenon, Lyngbya, Microcystis, Nodularia, Gloeotrichia, Cylindrospermopsis, Woronichinia*, and *Planktothrix* (*Oscillatoria*) (Fogg 1969; Reynolds and Walsby 1975; Pitois et al. 1997; Keil et al. 2002; Paul et al. 2005; Rajaniemi-Wacklin et al. 2006; Willame et al. 2006; Carey et al. 2008; Sitoki et al. 2012; Yamamoto and Shiah 2012). Having a global distribution, *Microcystis* spp. are among the bloom-forming cyanobacteria most commonly reported to dominate during bloom events, particularly in freshwater systems (Chorus and Bartram 1999).

2.2 Factors influencing cyanobacterial growth and dominance

Several abiotic and biotic factors, such as the availability of nutrients, light, temperature, the hydrological regime of the reservoir, grazing and physiological adaptations of cyanobacterial species, interplay and have an effect in the development of a bloom, its duration and the dominance by cyanobacterial species (Tilman et al. 1982; Paerl 1988). In this section the main factors suggested to influence cyanobacterial growth and dominance are reviewed.

2.2.1 Physico-chemical factors

Nutrients

Eutrophication, or nutrient enrichment (Smith et al. 1999), is part of the natural evolution of aquatic ecosystems, especially of rivers, ponds lakes and other shallow water bodies (Vasconcelos 2006). However, human activities have accelerated this process, affecting the water quality of aquatic systems and disrupting their balance, leading to the development of harmful algal blooms, hypoxia and fish kills among other undesirable consequences (Davis 1964; Paerl et al. 2001). Eutrophication is expected to be exacerbated by climate change, since changes in patterns of precipitation duration and intensity may further increase nutrient loading to aquatic systems from non-point sources in some areas (Paerl and Paul 2012; Schindler 2012).

Since nutrients (particularly phosphorus and nitrogen) have been identified as variables controlling phytoplankton growth and biomass (Owens and Esaias 1976; Hecky and Kilham 1988), there is general agreement that the development of harmful algal

blooms and their increased incidence have largely resulted from the increase in phosphorus and nitrogen concentrations from point and non-point sources (Smith et al. 1999; Schindler 2012). However, it is not clear how the stoichiometric relationship between these nutrients and/or their absolute concentrations promote cyanobacterial blooms.

Phosphorus and nitrogen concentrations

Once light and temperature are adequate (i.e., non-limiting factors), the growth of phytoplankton can be limited by the availability of nutrients (Hecky and Kilham 1988). In aquatic ecosystems, the availability of P and N are generally lower than those required for plant growth (Paerl et al. 2001), hence an influx of such nutrients is generally associated with a stimulation of phytoplankton growth.

Irrespective of N:P ratios, the absolute concentration of nutrients, particularly P, has been linked to phytoplankton growth in freshwaters. Whole-lake experimentation has demonstrated that P loading is crucial for the development of algal blooms (Schindler 1974; Schindler 1977). Positive associations between P concentrations and cyanobacterial biomass have also been reported (Kotak et al. 2000).

Management strategies to target P loading, coupled with changes in patterns of land use, increased human activities and urbanization in watersheds over the past several decades, have changed the dynamics of nutrient input into aquatic environments, resulting in N loads, sometimes at higher rates than P loads (Paerl and Scott 2010; Paerl and Paul 2012). This increase in nitrogen loadings has sparked interest in investigating the role that N plays in the promotion of cyanobacterial blooms. Gobler and colleagues (2007) suggested that N may play an equally important role to P in bloom promotion, particularly in ones dominated by the non-diazotrophic (i.e. non-nitrogen fixing) genus *Microcystis*. Their field studies, coupled with nutrient amendment analyses, examined the direct effects of N on growth and toxin production. Their results indicated that N loading can promote higher growth rates of the total algal community and *Microcystis* as well as an increase in toxin concentrations (Gobler et al. 2007). However, based on nutrient amendment analysis under different temperature regimes in samples collected from four freshwater lakes in the Northeast US (Davis et al. 2009), it was found that although increases in N concentrations significantly increased the growth rate of toxic *Microcystis*

in some lakes, increases in temperature coupled with higher P concentrations produced the greatest growth rates in toxic *Microcystis* in most of their experiments.

The importance of both macronutrients on the proliferation of cyanobacteria has led to a general disagreement as to which nutrient should be controlled in order to abate cyanobacterial blooms. Management strategies have often been focused on reducing P alone in order to prevent blooms (Paerl et al. 2001). Gobler and colleagues (2007) consider that this strategy is insufficient to prevent toxic cyanobacterial blooms in nitrogen limited systems. They further suggested that management strategies should target both N and P in order to prevent/mitigate the proliferation of all clades of toxic cyanobacterial blooms. In contrast, Schindler (Schindler 1977; Schindler 2012) suggests that management should focus on P reduction since whole-lake enrichment experimentation indicates that increasing P loads result in an increase of algal biomass, and he considers that controlling N inputs may in fact be counterproductive since it may result in low N:P ratios, thus promoting the proliferation of nitrogen-fixing cyanobacteria. This debate will clearly benefit from further studies that analyze, *in situ*, the response of toxic cyanobacterial blooms to the dynamics of these nutrients in natural waters, which was the foundation for the present study.

Nitrogen to phosphorus ratios (N:P)

N:P ratios could be at the base of the co-existence and competitive elimination patterns observed during succession in phytoplankton assemblages (Rhee 1978; Rhee and Gotham 1980; Smith 1982; Tilman et al. 1982; Hecky and Kilham 1988). The rational for this theory is linked to the fact that in order to reproduce, algae require elements (e.g., macronutrients such as N or P) in a relatively fixed proportion (Hecky and Kilham 1988). Since the biochemical function of each element is unique and they cannot be substituted (Droop 1974; Rhee 1978; Smith 1982), stoichiometric requirements need to be fulfilled for algae to grow (Hecky and Kilham 1988). Unialgal laboratory experiments under dual nutrient limitation showed no multiplicative or additive effect of the two nutrient concentrations (Droop 1974; Rhee 1978; Rhee and Gotham 1980), but rather a threshold pattern in which a single nutrient, the one present in the least quantity relative to the species specific nutrient demands, limited growth; the non-limiting nutrient had no effect on growth. Nonetheless, co-limitation by a physical factor (e.g., light, temperature) and a

nutrient is possible and has been observed in cyanobacteria via unialgal chemostats (Healey 1985). The species-specific nutrient demands can then be represented as an optimum nutrient ratio (Rhee 1978), which for N and P has been determined to be equal to the ratio of the subsistence quotas or minimum intracellular concentrations of N and P (Rhee 1978; Rhee and Gotham 1980). At the optimum nutrient ratio the species consumes both nutrients (Tilman et al. 1982), and divergence from this ratio results in nutrient conditions that limit the growth of the species (Rhee 1978), such that, above the optimum ratio the growth is only determined by P, and below the optimum growth is solely determined by N. The optimum N:P atomic ratios have been determined for several phytoplankton species in laboratory experiments, they were demonstrated to be species-specific and hence to vary among phytoplankton species (Rhee and Gotham 1980).

Due to interspecies differences in their optimum nutrient ratio (Rhee and Gotham 1980; Oh and Rhee 1991), it has been suggested that the proportions of the supplied nutrients can have an influence on which species will dominate the phytoplankton community (Rhee 1978; Rhee and Gotham 1980; Hecky and Kilham 1988). In addition to the kind of nutrient limitation, the degree of limitation could also be inferred from optimum nutrient ratios (Rhee and Gotham 1980).Comparing the population sizes of species competing for 2 resources to their species-specific optimum nutrient ratios, may be used as an approach to predict the outcome of resource competition (Tilman et al. 1982). In the schematic representation of this principle (Fig 2.1), for example, supply of N and P to the medium resulting in a cellular N:P ratio closer to the optimum for species A would lead to the dominance of species A and to the competitive exclusion of species B provided their maximum growth rates are similar (Rhee and Gotham 1980).

Attempts have been made to define environmental nitrogen to phosphorus ratios (as dissolved or total nutrients) that could be used in order to predict whether phytoplankton are nutrient limited (Sakamoto 1966; Smith 1982; Smith 1983), and in which cases certain phytoplankton members will dominate in freshwater systems (Smith 1982; Smith 1983). This approach has provided evidence of the influence that nutrient ratios have on the species composition of natural phytoplanktonic communities (Smith 1983; Liu et al. 2011).

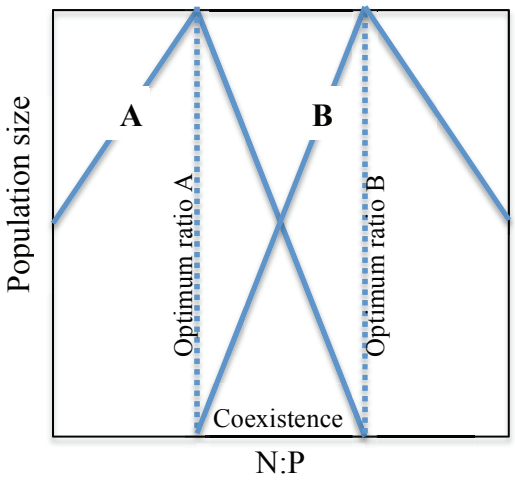

Figure 2.1 – Resource gradient in competition theory.
A and B represent hypothetical species; N:P, nitrogen to phosphorus ratio. Species B has a higher optimum nutrient ratio than species A, hence species A dominates at a lower nutrient ratio than species B and vice versa. At intermediate ratios the two species will be limited by different nutrients (i.e., A by P and B by N); hence they can coexist (Rhee and Gotham 1980). Adapted from Tilman et al. (1982).

From observations in Japanese lakes (Sakamoto 1966) it was suggested that chlorophyll yield was dependant only on TN when the TN:TP (reported as by weight) was below 10:1 and only on TP when the ratio was above 17:1. In regards to cyanobacteria, from analysis including data from 17 lakes worldwide and using the TN:TP in the water column as a surrogate for the N:P supply ratio, Smith (1983) reported a tendency for cyanobacterial algal blooms to occur at epilimnetic TN:TP ratios (reported as by weight) below 29:1. The observation of cyanobacterial dominance at low TN:TP ratios led Smith (1983) to suggest that, in general, cyanobacteria are better competitors for nitrogen than for phosphorus when compared to other groups of algae. Within the TN:TP ratio threshold in which cyanobacteria have been observed to dominate in natural waters (<29:1 reported as by weight (Smith 1983)), different cyanobacterial genera/species dominate depending on their physiological attributes. A boundary of TN:TP of 22:1 (reported as by weight) was identified to provide a distinct delimitation between lakes dominated by nitrogen-fixing cyanobacterial species (<22:1), from the ones where these species were rarely observed (>22:1) (Smith et al. 1995). The cosmopolitan incidence of toxic *Microcystis* blooms has led researchers to analyze the factors that may promote dominance by this non-diazotrophic genus in natural assemblages. From a study encompassing 11 years of environmental data of Lake Taihu,

China, a negative correlation was observed between *Microcystis* cell abundance and the TN:TP ratio; *Microcystis* tended to dominate the phytoplankton community at TN:TP ratios below 30:1 (reported as by mass) (Liu et al. 2011).

On the other hand, Reynolds (1998) argues that changes in nutrient ratios are not the promoters of the algal response. In limnetic aquatic enclosures, increases in the absolute concentration of the added nutrients, led to seasonal variation in dominance of nitrogen fixing species (decrease) in spite of the low nutrient ratio (Reynolds 1986). Indeed, the ability to predict cyanobacterial blooms based on environmental N:P ratios alone is impaired in some ecosystems. Particularly in hypereutrophic systems that have nutrient concentrations that exceed phytoplankton growth demands (Paerl 1988), where recycling of nutrients by biological factors maintains them at high concentrations and where favourable climatic conditions are not constant (e.g., temperate systems) (Reynolds 1998). In these cases, regardless of N:P ratios, the nitrogen and phosphorus concentrations *per se* may not be limiting growth; but cyanobacterial growth may be periodically limited by other factors, such as light or temperature (Paerl 1988; Reynolds 1998).

Temperature

Higher temperatures have often been associated with dominance of cyanobacteria in algal blooms. Hence, in temperate freshwater systems cyanobacterial blooms are frequently observed during the warmest periods of the year in nutrient rich environments (Paerl 1988; Davis et al. 2009). There is a general consensus that increasing global temperatures will lengthen the periods suitable for bloom formation and enhance environmental conditions that favour the dominance of cyanobacterial species (Davis et al. 2009; Te and Gin 2011; Paerl and Paul 2012). This is partly due to the fact that cyanobacteria reach optimum growth and photosynthetic rates at higher temperatures (generally >25°C) than do eukaryotic algae (Robarts and Zohary 1987). The high temperature optimum growth rates in cyanobacteria becomes a competitive advantage in nutrient rich environments allowing them to outcompete other members of the phytoplankton during warm periods (Paerl and Paul 2012). *Microcystis* has been observed to grow optimally at temperatures equal or higher than 25°C, however the optima can vary between cyanobacterial genera and even species (Robarts and Zohary 1987).

Differences in optimum temperatures for growth within cyanobacteria species (Robarts and Zohary 1987) may further affect the composition and community structure of a cyanobacterial bloom at high environmental temperatures. This is supported by laboratory competition experiments where *Microcystis aeruginosa* outcompeted the filamentous species *Phormidium tenue* at temperatures higher than 25°C under nitrogen limitation (Fujimoto et al. 1997).

There are other indirect effects of temperature that can result in cyanobacterial dominance. In shallow reservoirs, warming of the sediment-water interface can indirectly promote cyanobacterial dominance by accelerating P release from the sediments (Jensen and Andersen 1992; Smeltzer et al. 2012).

Light

Cyanobacteria exhibit photoprotective mechanisms (Mandal et al. 2011) and produce accessory pigments that allow for chromatic adaptation and the capture of photons at wavelengths not efficiently used by chlorophyll in deeper regions of the water column (Kehoe 2010). These adaptations enable cyanobacteria, particularly the ones able to control vertical migration using gas vesicles (Brookes and Ganf 2001) during periods of low or no turbulence mixing in the water (Huisman et al. 2004), to exploit habitats with high and low irradiances along the water column. In addition, buoyant cyanobacterial cells in the surface water, especially once dense blooms have formed, reduce light availability for other autotrophs while increasing their own access to light (Paerl 1988), thus enhancing their competitive dominance.

pH

Waters that support extensive blooms are usually characterized by high pH levels (Rantala et al. 2006; Liu et al. 2011) due to the increased photosynthetic activity by phytoplankton and the concomitant inorganic carbon decrease (Talling 1976; Paerl 1988; López-Archilla et al. 2004). At high pH levels not all phytoplankton are equally efficient in obtaining carbon (Talling 1976). However, *Microcystis* and other cyanobacteria are able to incorporate bicarbonate (HCO^{-3}) directly from the medium (Talling 1976; Qiu and Gao 2002; Chen et al. 2009), which is the dominant form of CO_2 at high pH (pH9) (Paerl 1988), or to transform bicarbonate to CO_2 via enzymatic reactions (Paerl 1988; Takahashi et al. 1990). In addition, cyanobacteria possess carbon-concentrating mechanisms (Song

and Qiu 2007). Hence, at high pH levels, these characteristics provide cyanobacteria with a competitive advantage over phytoplankton that do not possess some or any of these adaptations.

2.2.2 Biotic factors

Zooplankton-cyanobacteria interactions can affect cyanobacterial dynamics and the dominance of certain cyanobacteria taxa. Herbivorous zooplankton can negatively impact cyanobacterial dominance by grazing (Gobler et al. 2007). However, the low nutritional value, toxicity and ability to form large colonies, characteristic of some *Microcystis* species and other cyanobacteria, makes these species a non optimal food choice for some herbivorous zooplankton (Nizan et al. 1986; Fulton and Paerl 1987; Vanderploeg et al. 2001). While cyanobacterial toxicity lowers the ingestion and assimilation rates in certain zooplankton populations (Nizan et al. 1986), the low nutritional value of cyanobacteria can affect their reproduction and survival (Fulton and Paerl 1987). On the other hand, selective grazing and filtering by zooplankton have the potential to increase the incidence of toxic cyanobacterial blooms. It has been observed that zebra mussels do not digest *Microcystis aeruginosa* occurring as large colonies, but rather reject them in pseudofeces (Vanderploeg et al. 2001). Zooplankton can also promote cyanobacterial dominance during the summer by increasing nutrient supply via nutrient regeneration in the epilimnion (Tilman et al. 1982).

2.3 Cyanobacterial toxicity

Toxins produced by cyanobacterial genera can have numerous adverse health effects on humans and other animals (Falconer 1999). According to their toxic effects on animal cell lines, cyanotoxins are usually classified as neurotoxins, such as anatoxin, saxitoxin and neosaxitoxin; dermatotoxins, such as aplysiatoxins, lyngbyatoxin-a and debromoaplysiatoxins; irritant toxins, such as lipopolysaccharide endotoxins; hepatotoxins, such as microcystins, nodularins and cylindrospermopsin; the latter also have cytotoxic properties (Chorus and Bartram 1999). Among these toxins, microcystins and nodularins are most frequently detected in cyanobacterial blooms in fresh and brackish waters worldwide (Sivonen and Jones 1999).

2.3.1 Microcystin toxicity and water quality guidelines

In mammals, hepatotoxic microcystin is transported through the gastrointestinal tract and concentrated in liver cells, where it has the potential to inhibit protein phosphatases 1 and 2A (Falconer and Yeung 1992), leading to the hyperphosphorylation of structural proteins and resulting in disaggregation of internal filaments and cell deformation (Chou et al. 1990). Acute exposure to microcystin could lead to severe liver damage and death (Teixeira et al. 1993; Carmichael et al. 2001; Ding et al. 2002). Even chronic exposure to low concentrations of microcystins are of special concern due to the potential role of microcystins in carcinogenesis (Li et al. 2011). Studies have indicated that microcystins are liver tumor promoters (Nishiwaki-Matsushima et al. 1992). Proteomic analysis indicated that microcystins alter the expression of a number of microRNAs and proteins involved in several pathways related to tumorigenesis and signaling pathways in mice (Zhao et al. 2011). In addition, the increased incidence of liver cancer in endemic areas in southeast China has been associated with drinking water sources potentially contaminated by microcystin (Ueno et al. 1996). An association between microcystin concentration in the water and the incidence of colorectal cancer has also been suggested (Zhou 2002).

The Canadian drinking-water guideline established by Health Canada for microcystin-LR is 1.5 µg/l while the provisional drinking-water guideline established by the WHO is 1 µg/l (WHO 1998). The recreational water quality guidelines for low and moderate probability of adverse health effects due to cyanobacterial cells are 20,000 and 100,000 cells/ml which in a bloom that consists of *Microcystis* cells would represent approximately 4 µg/l and 20 µg/l of microcystin, respectively (WHO 2003).

2.3.2 Microcystin producing genera

Toxic (microcystin) blooms in freshwater systems worldwide have been most frequently associated with the dominance of *Microcystis, Anabaena* and *Planktothrix* (formerly *Oscillatoria*) (Sivonen and Jones 1999). Microcystin production has been observed not only in isolates from water but also from terrestrial environments (Sivonen 2008). In addition to *Microcystis, Anabaena* and *Planktothrix,* microcystins have also been reported to be produced by strains within the genera *Nostoc, Anabaenopsis, Hapalosiphon, Synechocystis, Geitlerinema, Leptolyngbya* and *Gloeotrichia* (Carmichael

1992; Domingos et al. 1999; Sivonen and Jones 1999; Oudra et al. 2002; Codd et al. 2005; Carey et al. 2007; Gantar et al. 2009; Stanić et al. 2011). Microcystin producing genera identified in natural populations and in monocyanobacterial cultures are summarized in Table 2.1. Nonetheless, production of microcystin by several of these genera (Table 2.1.) have yet to be confirmed via detection of microcystins in axenic monocyanobacterial cultures (Codd et al. 2005). Production of microcystin by freshwater pico- and nanoplanktonic cyanobacteria strains from the genera *Aphanocapsa* and *Synechococcus* has also been suggested (Domingos et al. 1999); nonetheless the evidence to date is not conclusive. Microcystin production was also identified in freshwater benthic cyanobacterial isolates most closely related to genera within the order *Oscillatoriales,* such as *Phormidium* or *Lyngbya* (Izaguirre et al. 2007); however in this study there was no certainty in regards to the identity of the microcystin producer.

Table 2.1 – Members of the following genera have been identified to produce microcystin in natural populations and/or in monocyanobacterial cultures

Natural population, dominant genera	**Monocyanobacterial culture. Non-axenic (not bacteria free) or not specified**	**Axenic monocyanobacterial culture (bacteria free)**
Microcystis *Anabaena* *Nostoc* *Planktothrix* (formerly *Oscillatoria*) *Anabaenopsis* *Hapalosiphon* *Gloeotrichia**	*Microcystis* *Anabaena* *Nostoc* *Planktothrix* (formerly *Oscillatoria*) *Hapalosiphon* *Synechocystis** *Geitlerinema** *Leptolyngbya**	*Microcystis* *Anabaena* *Planktothrix* (formerly *Oscillatoria*)

Table adapted from Codd et al. (2005).

(*) Information from strains/colonies isolated from natural populations (Oudra et al. 2002; Carey et al. 2007; Gantar et al. 2009; Stanić et al. 2011) added to the table originally compiled by Codd et al. (2005).

2.3.3 Microcystin chemical structure and stability

Microcystins comprise a family of monocyclic heptapeptides with the general structure cyclo(-d-Ala-l-X-erythro-b-methyl-d-isoAsp- l-Y-Adda-d-isoGlu-N-methyldehydro-Ala) (Fig 2.2). They are characterized by a β-amino acid, Adda (3-amino-9-methoxy-2,6,8-trimethyl-10-phenyldeca-4E, 6E-dienoic acid), which has been identified to be essential for toxic activity (Dawson 1998). A single cyanobacterial strain

can produce several microcystin variants (Nishizawa et al. 2000). Over 89 microcystin structural variants have been reported (Welker and Von Döhren 2006) which primarily differ in the amino acid composition at 2 non-conserved positions and in the methylation/demethylation patterns in the molecule (Sivonen and Jones 1999). Microcystin–LR (fig 2.2), the most common of the variants (Sivonen and Jones 1999), has leucine (L) and arginine (R) in the variable positions. It has been reported that microcystin shows stability in deionized and sterilized water (Cousins et al. 1996), and under extreme temperatures or pH (WHO 1998).

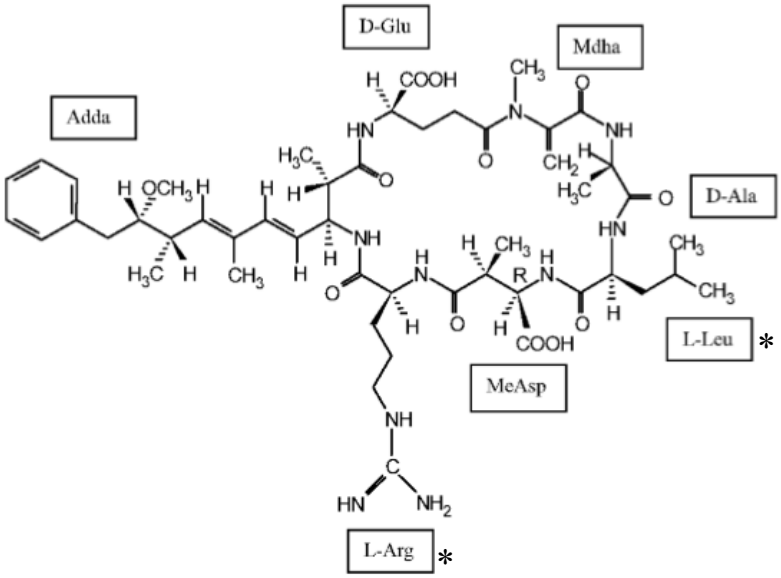

Figure 2.2 – Chemical structure of microcystin-LR.
Amino acids are indicated in boxes: D-Ala – D-alanine; L-Leu – L-leucine; MeAsp – D-erythro-β-methyloaspartic acid; L-Arg – L-arginine; Adda – 3-amino-9-methoxy-2,6,8-trimethyl-10-phenyl-deca-4,6-dienoic acid; D-Glu – D-glutamate; Mdha – N-methylodehydroalanine. (*) Non-conserved amino acids among microcystin variants. Adapted from Mankiewicz et al. (2003).

The purified toxin is stable and decomposes poorly under light irradiation (Tsuji et al. 1994). However, in the presence of cyanobacterial pigments, sunlight can cause significant isomerization of the double bonds and decomposition of microcystin (Tsuji et al. 1994). Rapid decomposition of microcystin has been observed after exposure to UV light at wavelengths around the absorption maxima of the molecule (238–254 nm) (Tsuji et al. 1995). In natural waters, microcystins can also be biodegraded (Cousins et al. 1996; Ho et al. 2012; Shimizu et al. 2012). While the majority of the bacteria reported to degrade microcystins belong to the family *Sphingomonadaceae* (Bourne et al. 1996; Bourne et al. 2001), degradation by other bacteria has also been reported (Ho et al. 2012).

2.3.4 Microcystin biosynthetic mechanism, evolution and potential functions

Microcystins are synthesized nonribosomally (Arment and Carmichael 1996). In *Microcystis aeruginosa,* the 55 kb microcystin biosynthetic gene cluster known as microcystin synthetase, has been characterized (Nishizawa et al. 2000; Tillett et al. 2000; Rhee et al. 2012). The microcystin synthetase gene cluster encodes the enzyme complex responsible for microcystin production, which is an integrated polyketide synthase and peptide synthetase system (Tillett et al. 2000). The microcystin synthetase gene cluster consists of 10 open reading frames arranged as 2 polycistronic operons (*mcyABC* and *mcyDEFGHIJ*), which are transcribed from a central bidirectional promoter localized between *mcyA* and *mcyD* (Kaebernick et al. 2002). While the genes *mcyA* to *mcyE* and *mcyG* encode the catalytic domains of the multifunctional integrated polyketide-peptide enzyme that carry out most of the catalytic reactions involved in the production of microcystin, the other 4 genes (*mcyF* and *mcyH* to *mcyJ*) encode monofunctional proteins presumed to be involved in precursor and tailoring functions (Tillett et al. 2000). It is now widely accepted that the ability of a cyanobacterial strain to produce microcystin depends on the possession and functional expression of the microcystin synthetase genes (Nishizawa et al. 2000; Tillett et al. 2000; Christiansen et al. 2003; Pearson and Neilan 2008).

The composition and gene arrangement in the microcystin synthetase gene cluster is highly similar between *Microcystis* strains (Nishizawa et al. 2000; Tillett et al. 2000; Rhee et al. 2012). The *mcy* genes and their arrangement differ between the genera *Microcystis*, *Anabaena* and *Planktothrix* (Christiansen et al. 2003; Rouhiainen et al. 2004). These 3 genera possess the *mcyABCDEGJ* genes: while the *mcyABC* are arranged in the same order, the order of the other genes differs between these genera. In addition, *Planktothrix* possess the gene *mcyT*, which is absent in the other 2 genera, but does not have the genes *mcyF* and *mcyI* identified in both *Microcystis* and *Anabaena* (Pearson and Neilan 2008).

The distribution of the *mcy* gene cluster among modern cyanobacteria is sporadic (Sivonen and Jones 1999; Rantala et al. 2004). Toxin production is observed in several genera, especially *Microcystis, Anabaena* and *Planktothrix* in freshwaters, which are distantly related to one another (Sivonen and Jones 1999). Moreover, within most genera, there are toxic and non-toxic strains (Vézie et al. 2002; Pearson and Neilan 2008).

Phylogenetic analyses have provided evidence that microcystin biosynthetic genes coevolved with house keeping genes (Rantala et al. 2004). Rantala and colleagues (2004) suggest that the common ancestor of a large number of cyanobacteria used to possess microcystin synthetase genes and that the pattern of *mcy* genes distribution in cyanobacteria today may be due to the fact that these genes have been lost by certain strains during evolution. The variation of *mcy* gene clusters between cyanobacteria have been attributed to recent recombination, mutation and horizontal gene transfer events (Rantala et al. 2004; Tooming-Klunderud et al. 2008).

While the harmful effects that microcystins have on animal and human health are well documented (Falconer 1999), the role that microcystin plays in toxin producing cells remains unclear. The toxin has been proposed to be involved in allelopathic interactions with other cyanobacterial species (Li and Li 2012) and other organisms (Fulton and Paerl 1987; Gobler et al. 2007) and to serve as a feeding deterrent and anti-grazing strategy (Nizan et al. 1986). However the role of microcystin as a defense mechanism against predation has been questioned (Schatz et al. 2007; Zilliges et al. 2011) and may be fortuitous since phylogenetic evidence suggests that the evolutionary origin of microcystin preceded the origin of potential grazers (Rantala et al. 2004). Microcystins have also been suggested to act as infochemicals and signaling molecules released into the medium upon cell lysis (Schatz et al. 2007). Recent research based on proteomic analysis indicates that microcystin inside toxic cells, via covalent binding to other proteins, is involved in protein modulation as well as in protein protection against oxidative damage (Zilliges et al. 2011).

2.3.5 Factors influencing microcystin production

From observations in N-limited (Orr and Jones 1998), P-limited (Oh et al. 2000) and N and P replete (Lyck 2004) *Microcystis aeruginosa* cultures, it has been suggested that the rate of production of microcystin is positively correlated to the rate of cell division, while the cellular microcystin content is negatively correlated to the cell division rate (Orr and Jones 1998; Lyck 2004). Parallel patterns of change in cell quotas of microcystin, chlorophyll a, carbohydrate and protein, throughout cell growth, have further suggested that the production of microcystin is constitutive (Lyck 2004).

That microcystin production is related to cell growth and cell division (Orr and

Jones 1998; Lyck 2004) is supported by transcriptional analyses of the *mcy* genes (Sevilla et al. 2010), but these analyses further suggest that the amount of microcystin produced could be regulated (induced) in response to certain environmental factors. It has been observed that high light intensities and red light increases *mcy* gene transcription (Kaebernick et al. 2000). The transcriptional response of the *mcy* operon is also induced under iron starvation (Sevilla et al. 2008). Despite the fact that microcystin content of the cells has been reported to change depending on the N:P ratio of the culture medium (Lee et al. 2000), transcriptional analysis suggests that the effect of macronutrients on microcystin production may be due to the stimulation of growth of the toxin producer that results in a concomitant increase in toxin concentration, rather than to an induction of the *mcy* operon *per se*. For example, under excess nitrogen conditions, the growth of *Microcystis aeruginosa* was stimulated and microcystin concentrations increased accordingly, however, the microcystin content per cell did not increase, and *mcyD* expression was found to be independent of nitrogen availability (Sevilla et al. 2010).

In freshwater, microcystin concentration dynamics have been associated with changes in the concentrations and ratios of N and P (Kotak et al. 2000). Based on analyses of microcystin concentrations from 246 freshwater Canadian lakes from 2001 to 2011, it was found that microcystin concentrations were high only at low N:P ratios (<23:1 reported as by mass) in nutrient rich waters (Orihel et al. 2012). A negative relation between microcystin concentrations and N:P ratios was also reported in several Alberta lakes (Kotak et al. 2000) and Lake Erie (Rinta-Kanto et al. 2009).

The presence of toxic and nontoxic subpopulations in cyanobacterial blooms (Davis et al. 2009; Rinta-Kanto et al. 2009) contributes to the wide variation in microcystin concentrations during bloom development. Recently, numerous field studies from a variety of environments worldwide have monitored the *mcy* genes in algal blooms and have correlated the abundance of these genes to environmental parameters in order to determine which conditions promote toxic cyanobacterial species (Gobler et al. 2007; Davis et al. 2009; Rinta-Kanto et al. 2009; Davis et al. 2010; Joung et al. 2011). The results are in many instances contradictory. This suggests that the significance of the effect that an environmental factor may have on the promotion of toxic blooms depends on the site-specific characteristics and hence differs between aquatic systems (Rinta-

Kanto et al. 2009; Martins and Vasconcelos 2011).

2.4 Recent advances in molecular biological techniques and their relevance to cyanobacterial ecology studies

2.4.1 Characterization of cyanobacterial diversity during algal blooms

To date, the identification of the cyanobacterial community or major taxa has widely relied on microscopy analysis (Davis et al. 2009; Joung et al. 2011; Sitoki et al. 2012). Sequencing of clone libraries of 16S rRNA and the intergenic spacer region of phycocyanin subunit genes has also been employed (Kim et al. 2006; Wilhelm et al. 2006; Kormas et al. 2011). However these methods are labour-intensive and time-consuming, especially in the analysis of large numbers of samples. In addition morphological analysis of cyanobacteria requires a high level of expertise. As an alternative, high-throughput community fingerprinting methods such as automated ribosomal intergenic spacer analysis (ARISA) (Wood et al. 2008; Rigonato et al. 2012) have been employed to distinguish populations based on length heterogeneity of the PCR amplicon generated from the 16S–23S rRNA intergenic spacer region and in this way analyze community structure. However, unrelated species that exhibit identical lengths of the spacer region can lead to an underestimation of diversity in the community analyzed via ARISA (Fisher and Triplett 1999); making necessary a further sequencing step in order to obtain higher resolution. The emergence of next generation sequencing (NGS) offers a cost-effective approach to simultaneously analyze large numbers of samples from complex environments. Within sequencing platforms, multiplex analyses with the Ion Torrent platform is less expensive than with platforms such as 454/GS-FLX (Whiteley et al. 2012). The Ion Torrent platform determines the DNA sequence by measuring pH changes, and in this way identifies base incorporation during strand synthesis (Rothberg et al. 2011). Hence, the need for expensive light detection systems is avoided and costs are reduced. The high-throughput, cost-effective and semi-quantitative nature of Ion Torrent sequencing exhibits great potential for the study of cyanobacterial diversity and succession pattern during algal blooms. In this project we assessed total cyanobacteria dynamics and diversity using partial 16S rRNA gene high-throughput sequencing via the Ion Torrent platform.

2.4.2 Detection and quantification of genotypes with quantitative PCR

The sequencing of complete genomes of toxic *Microcystis* strains (Kaneko et al. 2007; Frangeul et al. 2008) and the elucidation of the microcystin synthetase gene cluster from cyanobacterial genera including *Microcystis*, *Anabaena* and *Planktothrix* (Nishizawa et al. 2000; Tillett et al. 2000; Christiansen et al. 2003; Rouhiainen et al. 2004; Rhee et al. 2012), have enabled the development of molecular biological techniques to detect and quantify toxic cyanobacterial strains from environmental samples.

Quantitative PCR (qPCR) is an advancement of the qualitative polymerase chain reaction technique (PCR). This sensitive method allows accurate determination of the initial amount of target DNA in a sample (Smith and Osborn 2009). Its quantitative nature further allows establishing correlations between target gene copy numbers and biotic and abiotic environmental parameters. This is a powerful attribute that has been exploited recently in the field of cyanobacterial ecology to better understand cyanobacteria population dynamics in natural assemblages (Martins and Vasconcelos 2011).

This technique has been used to quantify overall cyanobacterial numbers by targeting the 16S rRNA or RNA polymerase genes with cyanobacteria-specific primers (Davis et al. 2009; Rinta-Kanto et al. 2009) or by targeting genes of the phycocyanin operon (Joung et al. 2011), which is ubiquitous among cyanobacteria. The heterogeneity in the DNA sequences of these genetic markers has allowed the design of primers that target specific species. However, complications have been reported when targeting the16S rRNA gene to quantify cyanobacterial abundance, particularly due to the variable number of 16S rRNA operons in cyanobacteria species and the sequence heterogeneity in strains from natural environments (Rinta-Kanto et al. 2009; Engene and Gerwick 2011; Martins and Vasconcelos 2011). Quantitative PCR has also made it possible to investigate the factors that regulate the dynamics of toxin-producing cyanobacteria as well as non-toxin-producing strains of the same species, which are indistinguishable morphologically (Martins and Vasconcelos 2011). By targeting the genes of the microcystin synthetase gene cluster, qPCR was used to quantify total toxic cyanobacteria (Fortin et al. 2010) as well as genus-specific toxic genotypes (Davis et al. 2009; Rinta-Kanto et al. 2009; Joung et al. 2011; Al-Tebrineh et al. 2012) in complex environmental samples. In addition, the low detection limit of qPCR combined with the high correlations observed between toxic genotypes and microcystin concentrations in environmental samples suggests that toxic

genotype monitoring via qPCR may be a better predictor (Fortin et al. 2010) and indicator (Davis et al. 2009; Rinta-Kanto et al. 2009) of microcystin concentrations in aquatic environments than the currently recommended indicators such as total cyanobacterial cell counts and Chl a concentrations (WHO 1998). In this project we quantified *Microcystis* and assessed total cyanobacteria *mcyD* dynamics via qPCR.

2.5 identification of external sources of nutrients

The accelerated eutrophication and resulting proliferation of cyanobacterial blooms in MB have been partly attributed to land use changes in this watershed that have resulted in increasing nutrient loadings into the bay (Smeltzer et al. 2012), especially from non-point sources (Troy et al. 2007). Animals shed large amounts of nutrients (including P and N) in feces and urine (Howery and Pfister 1990; Kirchmann and Pettersson 1994; Sheldrick et al. 2003). The pattern of land use in this watershed (Simoneau 2007), suggests that sources of nutrients into the bay, such as animal excreta, could enter the bay from agricultural sites, manure based crop fertilization or manure disposal, public campgrounds, septic systems and drain fields, wastewater treatment plants or native animals. Hence, sources of excreta pollution should be identified in order to design effective management strategies to control these sources and in this way reduce nutrient loading into aquatic systems such as Lake Champlain.

Since in addition to nutrients, animal shed pathogens and host cells in excreta (Iyengar et al. 1991; Valiere and Taberlet 2000; Padia et al. 2012), in this project we explored the potential of using fecal indicator bacteria (FIB) and genetic markers as indicators of nutrient input into the bay and as a means to identify the animal excreta sources.

2.5.1 Molecular methods for the direct identification of animal sources of excreta contamination

The standard method to measure fecal contamination is to enumerate FIB, such as *E. coli or* enterococci (Field and Samadpour 2007; USEPA 2007). However FIB do not identify the source of the contamination. In the field of fecal source tracking, different indicators and markers have been employed to identify the sources of fecal contamination (Field and Samadpour 2007). Recently, host mitochondrial DNA (mtDNA) sequences have been proposed as a marker for fecal source tracking (Martellini et al. 2005). The

underlying principle of this marker is that animal excreta contains blood, epithelial and intestinal cells (Iyengar et al. 1991; Valiere and Taberlet 2000), hence, in environmental samples mtDNA may be used to track fecal contamination to their specific host (Martellini et al. 2005).

Contrary to FIB, mtDNA is not spread among species, except temporarily after meat consumption (Caldwell et al. 2007). In addition, there is heterogeneity of the mtDNA genome between animal genera (Vuong et al. 2013). These characteristics make mtDNA an ideal marker to be used in order to identify directly the animal source of excreta contamination in complex environmental samples (Martellini et al. 2005; Caldwell et al. 2007; Vuong et al. 2013). Recently, different molecular biological methods have been developed to target mtDNA of several animal species in water samples, such as conventional, nested and multiplex PCR (Martellini et al. 2005; Caldwell et al. 2007), qPCR (Caldwell et al. 2007), and microarrays (Vuong et al. 2013). Mitochondrial sequence information is now widely available for many animals in DNA databases. This allowed us to employ cost-effective, high-throughput methods such as NGS of mtDNA amplicons in order to directly identify potential animal excreta sources of pollution in large numbers of samples from MB.

2.5.2 *E. coli* as a potential indicator of recent nutrient input from external sources

In addition to feces (Padia et al. 2012), soil has been identified as an important source of *E. coli* into aquatic environments (Hardina and Fujioka 1991). It has been demonstrated that *E. coli* can survive, multiply (Byappanahalli 2004; Ishii et al. 2006; Padia et al. 2012) and naturalize (Ishii et al. 2006) in soils. Despite the fact that *E. coli* has also been observed to be able to reproduce in water (Hardina and Fujioka 1991; Padia et al. 2012), aquatic environments are not ideal for the growth of *E. coli* (Whitman et al. 2004). In addition, fish are not sources, but rather vectors of *E. coli* in environmental waters (Hansen et al. 2008). Independent of their source (e.g., animal or soils), once *E. coli* cells arrive in water systems they can sink and/or be inactivated by sunlight via photobiological DNA damage, photooxidation of cellular components or photochemical damage (Whitman et al. 2004). Thus, it has been observed in waters that *E. coli* concentrations decline exponentially as a function of time of day (Whitman et al. 2004). Based on these facts, and considering that both sources of *E. coli* (i.e. feces and soils) are

rich in nutrients, *E coli* counts in surface water samples may be a potential indicator of recent nutrient input from external sources, such as surface runoff or sewage overflows. In the case of recent nutrient input from surface runoff, one would expect to find higher *E. coli* counts in the stations with greater contact with the shore and to find a correlation between *E. coli* counts and rainfall at these stations.

A constraint in employing *E. coli* abundance in surface water samples as an indicator of nutrient input from recent surface runoff, would come from the effects of water mixing events. These mixing events can lead to the resuspension of *E. coli* from the sediments (Ishii et al. 2007; Kiefer et al. 2012) or biofilms (Shikuma and Hadfield 2009). Considering the fact that as the station proximity to the shore increases, the depth of the stations decreases, a confounding effect may arise. Indeed water mixing events alone would suggest a surface runoff pattern for *E. coli,* even in the case of no runoff input. For this reason, the association between the presence of mtDNA from non-aquatic animal hosts and *E. coli* will be analyzed in order to corroborate whether there was *E. coli* input from external sources into the bay. For the purpose of this project, aquatic species were defined as species whose entire life cycle takes place in water.

Chapter 3 - Material and Methods

3.1 Sampling sites

The samples for this study were collected from 11 stations in the Missisquoi Bay /Pike River during 2009 (Table 3.1). There were 5 littoral (BM1, BM2, BM3, BM4 and BM5) and 3 pelagic stations (Pelagic 2, Pelagic Venise and Domaine Florent) in Missisquoi Bay, 1 station located in the Pike River delta (Pike River Mouth) and 2 stations located in tributaries (Pike River and Ewing Creek) (Fig 3.1). When conditions allowed, the stations were sampled from April to December, except for the samples collected from the stations BM1 to BM4 and PelVen, which made part of an epidemiological study that was performed from the end of June until the end of August.

Table 3.1 – Sampling stations in Missisquoi Bay and tributaries

Station Identifier	Station name	Station type	Approximate GPS coordinate (Degrees, minutes, seconds)
BM1	Littoral St. 1	Littoral	N45° 1' 37.0806", W73° 10'53.4606"
BM2	Littoral St. 2	Littoral	N45° 5' 4.4406", W73° 8' 21.0012"
BM3	Littoral St. 3	Littoral	N45° 4' 5.5194", W73° 7' 48.72"
BM4	Littoral St. 4	Littoral	N45° 3' 28.1988", W73° 4' 38.7582"
BM5	Littoral St. 5	Littoral	N45° 2' 23.64", W73° 4' 41.8182"
Pel2	Pelagic St. 2	Pelagic	N45° 2' 9.6612", W73° 6' 23.5794"
PelVen	Pelagic Venise	Pelagic	N45° 4' 4.5012", W73° 8' 37.2582"
DF	Domaine Florent	Pelagic	N45° 4' 13.0188", W73° 7' 0.6594"
PRM	Pike River mouth	river mouth	N45° 4' 14.1594", W73° 5' 48.48"
PR	Pike River	Tributary	N45° 7' 23.6094", W73° 4' 6.8376"
EC	Ewing Creek	Tributary	N45° 7' 31.0758", W73° 4' 39.2442"

3.2 Sample collection

For DNA analyses, the spring and fall samples from BM5, Pike River and Ewing were collected with a sterile glass bottle that was immersed beneath the surface of the water (shallow sampling sites). For the pelagic stations, the photic zone was collected using a peristaltic pump. From the end of June to the end of August, all of the littoral and pelagic stations were sampled as follows: the, surface water was collected with an acrylic tube (93 cm long by 10 cm in diameter) that was immersed horizontally beneath the surface of the water. Three sub samples were combined to form a composite sample.

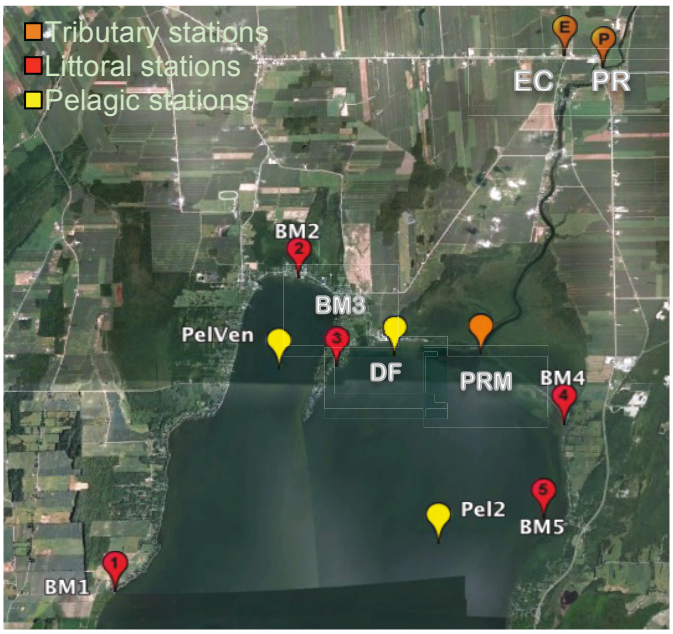

Figure 3.1 – Sampling stations in Missisquoi Bay and tributaries.
Google Earth v. 7.0.1.8250. (May 19, 2012). Missisquoi Bay, Lake Champlain. 45°02'44.86"N, 73°07'57.60"W, elevation 29 m. Eye altitude 29.50 km. Cnes/Spot Image 2012, DigitalGlobe 2012, GeoEye 2012. Available from: earth.google.com. [Accessed November 9, 2012].

For *E.coli* analyses, all samples were collected directly with a sterile bottle immersed beneath the surface of the water during the spring and the fall. From the end of June to the end of August, to ensure that representative samples were collected for the epidemiological study, all of the littoral stations were sampled as follows: two samples were taken at the surface with a sterile bottle along two axis lines generally separated by about 10 metres. The two samples were combined in the lab to form a composite sample. The water samples were stored at 4°C until further laboratory analysis, which took place within 48 h of sample collection.

There were variations on sampling dates for DNA and chemical analyses, as some data were collected as part of other research projects as previously mentioned. For DNA analysis (n=161), samples from both pelagic stations and the littoral station BM5 were collected on a weekly basis from April to June (the sampling of the pelagic stations started in May), every 3 to 5 days from July to August and weekly from September to November 10. From the littoral stations BM1-BM4, samples were collected on a weekly basis during the end of June to the end of August. DF was sampled monthly from May to November, except October. PRM was sampled monthly from May to June, biweekly during July and August and monthly in November. EC was sampled on a weekly basis

during April, on a biweekly basis from May to June and monthly in August. PR was sampled on a weekly basis during April, on a biweekly basis from May to September and monthly in November and December. Due to inaccessibility, only stations BM5 and PR were sampled in December.

3.3. DNA sample preparation, extraction and quantification.

Water samples were filtered (130-400 ml, depending on the amount of suspended solids and/or biomass) onto 0.2-µm hydrophilic polyethersulfone membranes (PAL Corporation, Ann Arbor, MI). The volume filtered per sample was recorded and taken into consideration for standardization during quantitative PCR analysis. The filters were kept frozen at -20°C until processed. A chemical and enzymatic lysis method followed by hot phenol extractions was used to extract DNA as described by Fortin et al. (2010). For each sample, the DNA pellet was resuspended in a standard volume (250 µl) of TE (Tris-Cl, 10 mM; EDTA, 1 mM; pH 8). DNA quantification was performed using the PicoGreen® dsDNA quantitation assay (Invitrogen, Burlington, ON, Canada) and a Safire microplate detection system (Tecan, Männedorf, Switzerland).

3.3.1 Quantitative PCR (qPCR)

3.3.1.1 Assessment of PCR inhibition.

qPCR was used to assess PCR inhibition in the DNA samples by spiking the qPCR master mixture with a known concentration of linearized plasmid DNA containing a 291-bp fragment of the bacterial luciferase gene amplified from *Pseudomonas cepacia* BRI6001L (Masson et al. 1993). In each run, 3 positive controls were included, consisting of spiked master mix with RNAse free water instead of sample DNA, in order to determine the number of gene copies that could be detected in the absence of PCR inhibitors, and thus calculate the percent inhibition for each environmental sample. Two negative non-template controls were included in each run. The qPCR reactions were performed using the QuantiTect SYBR Green PCR kit (Qiagen, Mississauga, Canada) on a Rotor-Gene 3000 apparatus (Corbett Life Science, Sydney, NSW, Australia) in 20 µl volumes containing 4 mM $MgCl_2$, 0.8 µM of each primer (*Lux*AqF2/R primer set, Table 3.2), 10 µl of SYBR green from a QuantiTect PCR kit (Qiagen), 5 µl of diluted DNA sample, and 3 µl of pLuxA (spiking material). The concentrations of the DNA sample dilutions used in qPCR analysis were verified by PicoGreen® dsDNA quantitation assay

(Invitrogen, Burlington, ON, Canada). The qPCR amplification conditions involved a preheating step at 95°C for 15 min and then 45 cycles of 95°C for 10 sec, 60°C for 15 sec, and 72°C for 20 sec. Fluorescence was measured at the end of each cycle at 72°C. The water samples were analyzed in duplicate. Standards were generated using triplicate tenfold serial dilutions of plasmid DNA containing the *LuxA* fragment. The PCR efficiency and the correlation coefficient of the standard curve were 0.95 and 0.99 respectively. In each run, triplicates of the 10^{-3} and 10^{-6} *LuxA* plasmid DNA standards were included. The water sample *CT* values were determined by importing the standard curve and adjusting it to the *CT* values of the 10^{-3} or 10^{-6} *LuxA* plasmid DNA standards.

3.3.1.2 Quantitative PCR of Microcystis 16S rRNA gene

Microcystis 16S rRNA qPCR standard curve

Primers that targeted a specific region of the *Microcystis* 16S rRNA gene were designed (Table 3.2 and Appendix B). The PCR reaction for cloning was performed in a 50 µl volume containing 0.5 µM of each primer (UcyaF3, UcyaR3), 3 units of *Taq* DNA polymerase (GE Healthcare, Baie d'Urfe, Canada), 5 µl of 10X *Taq* polymerase buffer (100 mM Tris-Cl, pH 9.0, 500 mM KCl, 15 mM $MgCl_2$), 1 mM $MgCl_2$, 0.2 mM of each deoxynucleoside triphosphate and 10 ng of *M. aeruginosa* genomic DNA from a pure culture isolated from Lake Champlain and provided by Dr. David Bird from l'Université du Québec à Montréal (UQAM). The PCR cycling conditions were as follows: prior to the addition of *Taq* DNA polymerase, the samples were denatured for 5 min at 96°C. This was followed by 24 cycles of denaturing at 94°C for 1 min, annealing at 60°C for 1 min and extension at 72°C for 1 min. The final extension was at 72°C for 10 min. The PCR product was purified using the GFX ™ PCR DNA and Gel Band Purification Kit (GE Healthcare) and quantified by SYBR® Safe staining (Qiagen) and spot densitometry using a ChemiImager (Alpha Innotech Corporation, San Leandro, CA). Ligation, transformation, and screening of transformants was performed as previously described (Fortin et al. 2010).

Standards were made from tenfold serial dilutions of linearized plasmid DNA containing the UcyaF3/R3 fragment of the *Microcystis* 16S rRNA gene. The qPCR reactions were performed in a Rotor-Gene 3000 apparatus (Corbett Life Science) in 20 µl volumes containing 4 mM $MgCl_2$, 0.8 µM of both primers, 10 µl of SYBR green from a QuantiTect

PCR kit (Qiagen), and 5 µl of template DNA or DNase/RNase-free water (Sigma Aldrich Canada, Oakville, ON) in the case of non-template controls. The qPCR cycling conditions involved an initial preheating at 95°C s for 15 min, followed by 45 cycles of 95°C for 10 s, 60°C for 15 s and 72°C for 20 sec. Fluorescence was measured at the end of each cycle at 72°C. The PCR efficiency and the correlation coefficient of the standard curve were 1.04 and 0.99 respectively.

Table 3.2 – Primer sequences used in this study
*Amplicon size without Ion Torrent adapters

Target (amplicon size and details)	Primer name	Template primer sequence (5'-3') +	Application	Reference and notes
Partial bacterial luciferase alpha subunit (154 bp in *Pseudomonas cepacia* BRI6001L (Masson et al. 1993))	*Lux*AqF2	GCAGAAGACGTA AACCTACTGG	qPCR inhibition assay	Primers amplify the region 965-1116 bp relative to the *luxA* gene sequence of *Vibrio harveyi* (Cohn et al. 1985)
	*Lux*AqR	TTCTTTCATCAAG TCATACCAACAG		
Partial cyanobacterial 16S rRNA gene (161 bp)*	UcyaF4	GCAAGCGTTATCC GGAAT	Ion Torrent sequencing	This study (Appendix A)
	UcyaR6	TCTACGCATTTCA CCGCT		
Partial *Microcystis* 16S rRNA gene (203 bp)*	UcyaF3	AATACGGGGGAG GCAAGCGTTATC	qPCR and Ion Torrent sequencing	This study (Appendix B)
	UcyaR3	ACGCTTTCGCCAC CGATGTTCTT		
Partial cyanobacterial *mcyD* gene of microcystin synthetase enzyme complex (107 bp)	*mcy*D(KS) F	TGGGGATGGACT CTCTCACTTC	qPCR	(Fortin et al. 2010)
	*mcy*D(KS) R	GGCTTCAACATTC GGAAAACG		
Mitochondrial DNA region between 12S and 16S rRNA genes (2 kb)	MI-50F	ACTGGGATTAGA TACCCCACTATG	Nested PCR and Ion Torrent sequencing	(Vuong et al. 2013)
	MI-50R	CGGTCTGAACTCA GATCACGTA		
16S rRNA gene fragments within the 2kb mitochondrial DNA amplicon. Fragment used for sequencing:134 bp* (region 2175-2309 bp in Homo sapiens mitochondrion. GenBank: AF346973.1)	MI50F_Io n	GGCYYWHRAGCA GCCAYCAA		Primers designed by Minh Vuong (Personal communication)
	PyrR1	ATIYTARCAKWRI YICWYCTAT		

+ Degenerate oligonucleotides: Y=C,T; W=A,T; H=A,C,T; R=A,G; K=G,T.
Non-canon base: I= base-pairs with A,T,G,C.

Limit of detection (LOD) for *Microcystis* 16S rRNA gene qPCR

The LOD of the *Microcystis* 16S rRNA gene qPCR, employing the UcyaF3/R3 primer set, was determined using qPCR and the LOD tool from GenEx Pro software (version 4.4.2; MultiD Analyses AB, Göteborg, Sweden) as described by Fortin et al. (2010). The amplification conditions were the same as the ones used for the qPCR standard curve. At a cutoff value of 33 cycles and 95% confidence level, the LOD was determined to be 31.15 copies per reaction (6.23 copies/µl).

Quantitative PCR analysis of *Microcystis 16SrRNA* gene from Missisquoi Bay samples

For the determination of *Microcystis* 16S rRNA gene copy numbers in Missisquoi Bay and tributary water samples, qPCR was performed on three fivefold dilutions of each DNA sample. The concentration of the 10^0 dilutions was verified by fluorescence using the PicoGreen® dsDNA quantitation assay (Invitrogen). In each run, triplicates of the 10^{-4} and 10^{-8} plasmid DNA standards were used as positive controls, and DNase/RNase-free water (Sigma Aldrich Canada) was used as the no-template control. The amplification conditions were the same as the ones used for the qPCR standard curve. The water sample *CT* values were determined by importing the standard curve and adjusting it to the *CT* values of the 10^{-4} or 10^{-8} plasmid DNA standards used as the positive control.

Each value below the *Microcystis* 16S rRNA gene qPCR LOD was recorded as zero copies/µl, unless inhibition had been demonstrated, in which case they were removed from further analysis. The final qPCR results for each sample were standardized based on the total amount of DNA extracted for each sample and the volume of water filtered for the sample. The values were expressed as *Microcystis* 16S rRNA gene copies in total DNA extract/ml of filtered water sample. In order to remove from the analysis the amount of copies amplified from non-intended targets, the *Microcystis* qPCR data was corrected based on the percentage of *Microcystis* sequences amplified per sample via UcyaF3/R3 Ion Torrent sequencing results (Appendix B). These data are referred to as *Microcystis* 16S rRNA copies/ml in further analysis. Assuming that *Microcystis* has 2 rRNA operons per genome (Rinta-Kanto et al. 2009; Engene and Gerwick 2011), an estimate of *Microcystis* cell abundance was made based on the *Microcystis* F3/R3 qPCR output. These data are referred to as estimated *Microcystis* cells/ml in further analyses. The *Microcystis* 16S rRNA copies/ml values were also analyzed based on the cyanobacterial

(UcyaF4/R6) Ion Torrent sequencing results (section 3.3.2) in order to determine relative numbers of 16S rRNA gene copies of total *Cyanobacteria* in relation to *Microcystis* 16S rRNA gene copy numbers. This is referred to as estimated total cyanobacterial 16S rRNA copies/ml in further analyses.

3.3.1.3 Quantitative PCR of microcystin gene

Amplification of the β-ketoacyl synthase subunit of the *mcyD* gene was done via qPCR in order to detect and monitor the dynamics of microcystin-producing cyanobacteria and to determine the correlation between toxigenic cyanobacteria and toxin concentration in the Missisquoi Bay/Pike River area.

Microcystin gene qPCR was performed on all of the samples for which there was DNA available. qPCR amplification was performed with the *mcyD*(KS) primers and conditions described in Fortin et al. (2010), with the exception that the annealing temperature was at 54°C instead of 58°C. The PCR efficiency and the correlation coefficient of the standard curve were 1.00 and 0.99 respectively. Each value below the *mcyD* qPCR LOD (Fortin et al. 2010) was recorded as zero copies/µl unless inhibition was demonstrated, in which case they were removed from further analysis.

The output was standardized based on the total amount of DNA extracted for each sample and the volume of water filtered for the sample and is referred to as *mcyD* copies/ml in further analysis.

3.3.2 High-throughput sequencing of PCR amplicons with the Ion Torrent platform

High-throughput sequencing of partial cyanobacterial 16S rRNA gene amplicons was employed in order to monitor the diversity and community composition of cyanobacteria in the Missisquoi Bay/Pike River area during 2009 and establish the relationship between the relative abundance of major cyanobacterial taxa and environmental parameters. A total of 57 DNA samples were sequenced with primers designed to target the cyanobacterial 16S rRNA gene (Appendix A).

For the purpose of this project, aquatic species were defined as species whose entire life cycle takes place in water. Considering the presence of mitochondrial DNA (mtDNA) from non-aquatic species as a prospective indicator of nutrient input from animal excreta, high-throughput sequencing using the Ion Torrent platform was also performed in order to identify non-aquatic animal genera potentially contributing to

nutrient input into the bay/Pike River area. A 2kb region in the 12S rRNA – Val 7 tRNA – 16S rRNA area of the mitochondrial genome allows for the identification of at least 28 commercial and domestic animals (Vuong et al. 2013). In our methodology, a nested PCR, within this 2 kb region, was required in order to amplify a shorter PCR product (205 bp) suitable for sequencing with the Ion Torrent platform (MI50F_Ion/PyrR1 primer set, table 3.2). In a validation step for this methodology, sequencing analysis of the 205 bp fragment amplified from a genomic DNA mixture of the 28 target animals revealed that the region amplified allowed the identification of 26 of these animals (Appendix C). A total of 27 DNA samples from MB/Pike River were sequenced.

3.3.2.1 PCR amplification of partial cyanobacterial 16S rRNA and mtDNA genes

Cyanobacterial 16S rRNA gene amplification

Cyanobacterial 16S rRNA gene amplification was carried out with the primer set UcyaF4/R6 (table 3.2), which had been modified as described by Yergeau et al. (2012) in order to contain the adapters required by the Ion Torrent sequencer as well as a unique (10 nt) multiplex identifier (MID) tag.

The PCR reactions were performed in 30 µl volumes containing 0.5 µM of each forward and reverse primer, 0.2 mM of each deoxynucleoside triphosphate, 2.5 mM $MgCl_2$, 0.75 U of *Taq* DNA polymerase (GE Healthcare), in 3 µl (1 X) of the 10X *Taq* DNA polymerase buffer provided (100 mM Tris-Cl, pH 9.0, 500 mM KCl, 15 mM $MgCl_2$) and 20 ng of template DNA. For samples in which more than 60% PCR inhibition was observed, two PCR reactions using 10 ng of DNA were performed and pooled upon completion. RNAse free water served as the no-template controls while 10 ng of *M. aeruginosa* genomic DNA was added for the positive controls. PCR cycling conditions involved an initial denaturation at 95°C for 5 min, followed by 25 cycles of 95°C for 30 s, 53°C for 30s, 72°C for 45s and ending with a final extension step at 72°C for 10 min. The reaction produced a single 224 bp fragment, which was prepared for sequencing as described in the next section.

mtDNA amplification

The amplicons to be sequenced were obtained via nested PCR targeting a fragment within the 2kb mtDNA region between the 12S and 16S rRNA genes. In order

to reduce the bias due to the use of degenerate primers in the second PCR, 3 replicates of each sample were included in the first PCR and its products pooled to be used in the second PCR. In the first PCR the reactions were performed in 25 µl volumes containing 2 µM of each forward and reverse primer (MI50F_Ion/PyrR1, Table 3.2), 1.3 mg/ml BSA, 1X KAPA2G Robust DNA polymerase mix (Kapa Biosystems Inc., Boston, MA, USA). The amount of template DNA added to the first PCR reaction was determined based on the results from the qPCR test for inhibition (section 3.3.1.1) and was used as a standard for all samples (5 ng). The first PCR cycling conditions involved an initial denaturation step at 95°C for 5 min, followed by 35 cycles of 95°C for 30 s, 50°C for 30s, 72°C for 1 min and ending with a final extension step at 72°C for 7 min. The second PCR generates the amplicons to be sent for sequencing, therefore we utilized the MI50_Ion primer set that had been modified as previously described, but using the template primers MI50F_Ion and PyrR1 (Table 3.2). The second PCR reactions were performed in 50 µl volumes containing 1 µM of each primer, 0.2 mM of each deoxynucleoside triphosphate, 4 mM $MgCl_2$, 1.25 U of *Taq* DNA polymerase (GE Healthcare), in 5 µl (1 X) of the 10X *Taq* DNA polymerase buffer provided (100 mM Tris-Cl, pH 9.0, 500 mM KCl, 15 mM $MgCl_2$) and 2 µl of pooled product from first PCR as template DNA. The second PCR cycling conditions involved an initial denaturation at 94°C for 5 min, followed by 25 cycles of 94°C for 30 s, 50°C for 30s, 72°C for 45 s and ending with a final extension step at 72°C for 7 min. The 205 bp fragment was prepared for sequencing as described in the next section.

3.3.2.2 Gene amplicon sequencing

PCR products were purified from agarose gels using a QIAquick Gel Extraction kit (Qiagen), quantified using PicoGreen® dsDNA quantitation assay (Invitrogen), diluted to a concentration of 5 x 10^8 molecules and pooled in an equimolar ratio resulting in two pools (one for the cyanobacterial 16S rRNA gene amplicon and one for the mtDNA amplicon). Sequencing was conducted on an Ion Torrent Personal Genome Machine™ using the Ion 314™ chip following the manufacturer's protocols as described in Yergeau et al. (2012).

3.3.2.3 Bioinformatic analyses

Sequence data were analyzed using the RDP pyrosequencing pipeline (http://pyro.cme.msu.edu/) (Cole et al. 2009). The sequences were binned according to their MID tags (only accepting perfect matches) and the MID and forward primer were trimmed. For cyanobacterial 16S rRNA gene amplicons, all sequences that had unidentified bases (Ns), an average expected quality score lower than 17, or were shorter than 150 bp, were removed from further analysis. The remaining sequences from the 57 distinct datasets, were then submitted to the RDP Classifier tool, using a 0.5 bootstrap cutoff, in order to identify cyanobacterial sequences. The FASTA Sequence Selection tool from RDP was used to remove from further analysis any sequence that was not assigned to the phylum *cyanobacteria* in each dataset. The remaining sequences from each dataset were aligned, using the program BLASTN, against a cyanobacteria/chloroplast database (total of 3,338 sequences of which 2,852 were cyanobacterial sequences) downloaded from the RDP database as of February 2012 (Databases described in detail in appendix A). The 100 best hits from the BLASTN search were kept for each read in every dataset. The BLAST files were imported into MEGAN4 to assign the reads of each dataset to appropriate taxa in the NCBI taxonomy using the lowest common ancestor algorithm (LCA) (Huson et al. 2011) with default parameters except: maximum number of matches per read, 100; Minimum Support, 1; Minimum Score, 44 (this score threshold removed hits with an E value higher than 10^{-5}); Top Percentage, 0.5 (this parameter allowed to only analyze matches with the best score). For mtDNA amplicons, all sequences that had unidentified bases (Ns), an average expected quality score lower than 15 or were shorter than 50 bp, were removed from further analysis. The sequences from the 27 resulting datasets were aligned with BLASTN against a mitochondrial DNA database downloaded from the NCBI site as of October 2011. The 100 best hits were kept for each read in every dataset. The BLAST files were imported into MEGAN4 to assign the reads of each dataset to appropriate taxa in the NCBI taxonomy using the lowest common ancestor algorithm (LCA) (Huson et al. 2011) with the parameters described before, except: Minimum Support, 5 (i.e. only genera that were represented by at least 5 sequences were considered to be present in the samples); Minimum score, 48 (in order to discard hits with an E value higher than 10^{-5}).

3.4 Determination of sample microcystin concentration (ELISA)

To estimate microcystin concentrations in the water samples, 20 to 500 ml of sample (depending on the density of the planktonic particles) were filtered through a Whatman GF/F 0.7 μm glass-fiber filter (Whatman, Inc., Florham Park, NJ). Filtrand and filtrate were used to detect intracellular (particulate) and extracellular (dissolved) toxins respectively.

The concentration of microcystins was determined at UQAM in the laboratory of David Bird using enzyme-linked immunoabsorbent assays (ELISA) in 96-well plates, according to manufacturer's instructions (Abraxis LLC, Warminster, PA USA). This test is sensitive to the ADDA functional group of these hepatotoxins, and the standard curve is based on microcystin-LR.

For intracellular toxins, a quarter of the filter was suspended in 1 ml of distilled water and sonicated on ice at full power for 1 minute followed by centrifugation at 13,000 x g for 10 min. Following centrifugation, the supernatant was collected for measurement. The dissolved microcystin concentration was measured in the GF/F 0.7 μm filtrate of filtered water samples. In general, the toxin extracts were diluted 20 times for dissolved toxins, 40 times for particulates, to avoid interference present in raw water. If necessary, extracts were diluted to 50 times to enter the linear region of the standard curve estimated by nonlinear regression, based on a logistic model with 4 parameters. The detection limit was set at three times the standard deviation of the blank. Values below this limit were considered to be not distinguishable from zero.

3.5 Cyanobacterial abundance, *E. coli* counts and environmental parameters

Phytoplankton subsamples (20 ml) were preserved with Lugol's iodine. The identification and enumeration of cyanobacterial species were performed by Irina Moukhina from the laboratory of Dr. David Bird at UQAM. Counts were performed using an inverted microscope by the Utermöhl method (Utermöhl 1958).

For *E. coli* analyses, the water was filtered in duplicates within 24-48 hours, on 0.45 μm cellulose ester membranes (Millipore Corporation, Bedford MA, É.-U.). The filters were processed in duplicates, on MI-Agar plates and incubated as recommended by the protocol MA.700-ECCtmi 1.0 from the 'Centre d'expertise en analyses environnementales du Québec (CEAEQ)'. Fluorescent and blue colonies, equal to 0.5 mm

or larger, were counted after a 24 hour incubation period at 35^0C ±0.5 °C (CEAEQ, 2006).

Nutrient analyses were conducted in the GRIL nutrient laboratory at UQAM. Duplicate samples were taken for measuring nutrient concentrations. Samples for dissolved nutrients were taken from the filtrate of the GF/F 0.7 μm glass-fiber filter (Whatman, Inc., Florham Park, NJ) used to filter water samples for intracellular toxins. Total and dissolved phosphorus concentrations were evaluated by the molybdenum-blue method after persulfate digestion and autoclaving. The digested samples were analyzed by spectrophotometry at 890nm. Total and dissolved nitrogen concentrations were measured as nitrates after alkaline persulfate digestion. Colorimetric analyses were carried out on an Alpkem-Flow Solution IV 0-1-Analytical RA autoanalyzer as described by the manufacturer (O-I-Analytical, College Station, TX U.S.A.)

Environmental parameters such as pH and water temperature were obtained with a submersible YSI 6600 v2 water quality multi-probe (YSI, Yellow Springs, Ohio, USA). Air temperature was recorded on site with a Kestrel 3500 weather meter (Nielsen-Kellerman, Bothwyn, PA, USA). Daily precipitation data for Philipsburg, QC (45°02'00.000" N, 73°05'00.000" W) was obtained from the National Climate Data and Information Archive of Environment Canada. The data, recorded in mm, were summed for the preceding 7 days up to, but excluding, the day of sampling; it is referred to as 7-day precipitation in further analysis.

3.6 Data and statistical analyses

For the spatial analysis of the relationship between nutrients (section 4.2.1) or *E.coli* counts (section 4.2.2) with station type and rainfall in the study area in 2009, the data was analyzed based on the station type. A rank was assigned to each station type depending on its degree of contact with the shore: Pelagic stations (PelVen, Pel2, DF), being furthest from the shore, were given a rank of 1, littoral stations (BM1-BM5) were given a rank of 2 and the tributary stations (PR, EC) were given a rank of 3. The PRM station was excluded from this analysis because it was the only station sampled at the delta of PR, and it did not fit in the rank criteria. In order to analyze the trend between station contact with the shore and the concentration of nutrients or *E. coli,* Cochran–Mantel–Haenszel (CSMH) statistics were performed and further verified with Spearman's

Rank correlations (rs). For CSMH, the continuous quantitative data for nutrient concentrations and the categorical data for station type contact with the shore were used, controlling for 3 rain levels (1=0-20 mm, 2=21-40 mm, 3=41-60 mm). For rs, the continuous quantitative data for nutrient concentrations and the categorical data for station type contact with the shore (1-3) were used. In order to analyze whether there was a relationship with nutrient concentrations/*E. coli* counts and rainfall, Spearman's Rank correlations were performed using the continuous quantitative data for nutrient concentrations and rainfall (rain d-7) at each of the station types (i.e. tributary, pelagic and littoral). A nonparametric analysis of variance, the Kruskal-Wallis test, was employed in order to verify whether the average annual nutrient concentrations or *E. coli* counts differed between station types. In addition, correlations were tested between dissolved and total phosphorus (μg/l), dissolved and total nitrogen (mg/l), and *E. coli* counts (CFU/100 ml). The Fisher exact test was employed to analyze the association between the detection of *E. coli* and non-aquatic mtDNA hosts. The statistical analyses were performed using SAS (v.9.3, SAS Institute Inc., Cary, NC). Nonparametric scores were specified for all tests. Outputs with a p value<0.05 were considered significant and their specific null hypothesis were rejected.

All samples tested from Ewing station had high levels of PCR inhibition and were removed from molecular biological analyses. The *mcyD* gene was not found in the PR station, and only 3 *Microcystis* sequences were found in this station (1 in July 12/09 and 2 in August 8/09) based on Ion Torrent sequencing analysis of the cyanobacterial 16S rRNA gene amplicon (UcyaF4/R6 primers). Therefore, the analysis of major cyanobacterial taxa found over time and space and of the dynamics of *Microcystis,* the *mcyD* gene, and microcystin within the cyanobacterial community profile in relation to environmental factors during 2009, was done on the littoral, pelagic and Pike River Mouth stations of Missisquoi Bay, excluding the data from the tributary stations. Since Tw, pH and DOC available data values did not differ significantly among the sampling stations, their daily average value was used for all the stations in the correlation analyses. Spearman Rank Correlations (rs) were calculated between the variables of interest employing R (v.2.15.0, The R foundation for Statistical Computing, Vienna, Austria). Two-tailed $p < 0.05$ was considered significant.

Chapter 4 – Results

4.1 General description of environmental conditions in MB/Pike River area during 2009

In this section, Missisquoi Bay (MB) stations include data from all of the stations located in the bay, i.e. littoral, pelagic and PRM stations. Missisquoi Bay/Pike River area refers to the data from all sampling stations, including the tributary stations PR and EC.

Cyanobacterial cells in the bay in 2009, 49.0% of the samples (n=343) were below the WHO recreational water guidelines for relatively low probability of adverse health effects (20,000 cyanobacterial cells/ml), 32.7% of the samples had between 20,000 and 100,000 cyanobacterial cells/ml, and 18.3% of the samples had cyanobacterial counts above 100,000, which is the threshold for moderate probability of adverse health effects (WHO 2003). The highest cyanobacterial cell count was observed on 21/7/09 at the littoral station BM4 with 14,493,570 cells/ml (Fig 4. 1).

From the samples collected for intracellular toxin analysis (n=349), 65.9% were below the WHO provisional drinking-water guideline value (Fig. 4.2), 13.8% of the samples had concentrations between 1-4 µg/l, 13.2% of the samples had concentrations between 4-20 µg/l and 7.1% had intracellular concentrations higher than 20 µg/l. The highest concentration of intracellular microcystin recorded in the bay was 773.75 µg/l on 16/7/09 at BM5.

Analyses of extracellular toxins (n=308) revealed that 94.5% of the samples were below the drinking-water guideline for microcystin-LR and only 3 samples had extracellular toxin concentrations higher than 4 µg/l. The maximum extracellular toxin concentration in MB samples was 5.12 µg/l at BM3 on 22/6/09.

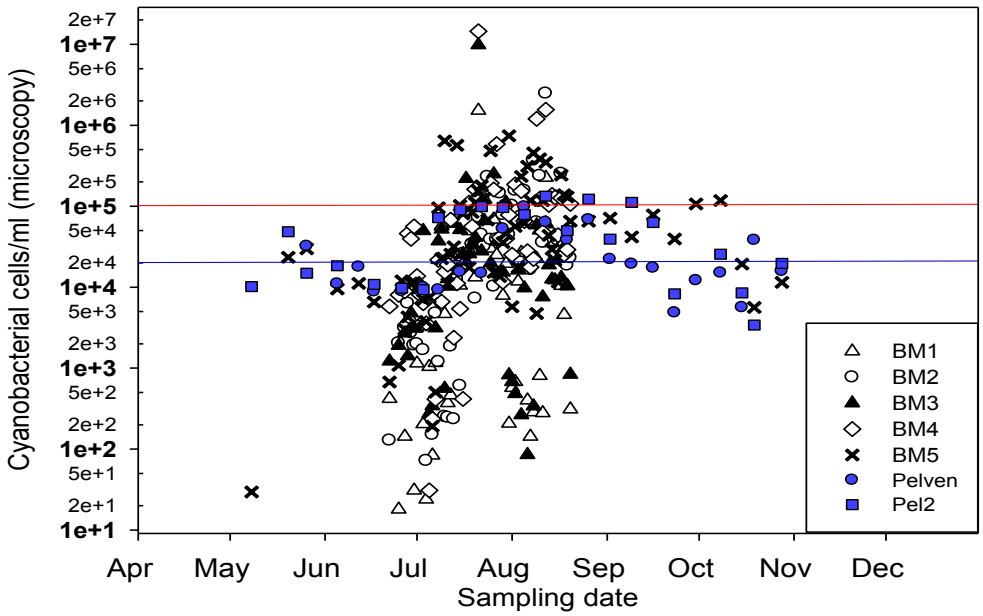

Figure 4.1 – Total cyanobacterial cells/ml in Missisquoi Bay during 2009.
Samples for cyanobacterial cell counting were collected from the littoral stations BM1-BM4 from late June to late August. For the littoral station BM5 and both pelagic stations, samples were collected from early May to late October. The horizontal lines indicate the WHO recreational water guidelines for relatively low probability of adverse health effects (blue) and the moderate probability of adverse health effects (red) due to cyanobacterial density.

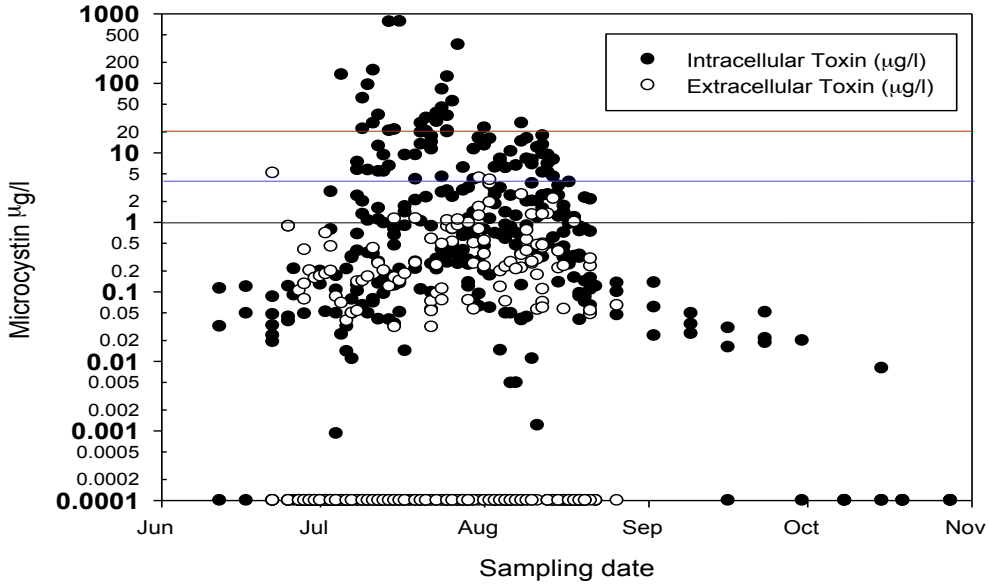

Figure 4.2 – Intracellular and extracellular microcystin in Missisquoi Bay in 2009
Samples for toxin analysis were collected from all MB stations. The horizontal black line represents the provisional drinking-water guideline value for microcystin-LR; the other two lines indicate the microcystin concentrations expected at a cell density of 20,000 cells/ml (blue) and 100,000 cells/ml (red, potential health impact in recreational waters due to toxin) of microcystin producing genera (WHO 2003). A value of 0.0001 µg/l for intracellular microcystin indicates an observation below the detection limit; the detection limit for extracellular microcystin was 0.02 µg/l.

The *mcyD* gene was detected in the bay stations but not in the tributary station, PR. The maximum number of *mcyD* copies in the bay was 6,708,091 copies/ml recorded on 25/7/09 at Pel2. *Microcystis* 16S rRNA gene copies followed a similar pattern as the *mcyD* gene (Fig 4.3). The highest number of *Microcystis* 16S rRNA gene copies was observed at PRM on August 8 with 14,423,840 copies/ml while very low copy numbers were observed in PR, with a maximum of 20 copies/ml on August 8.

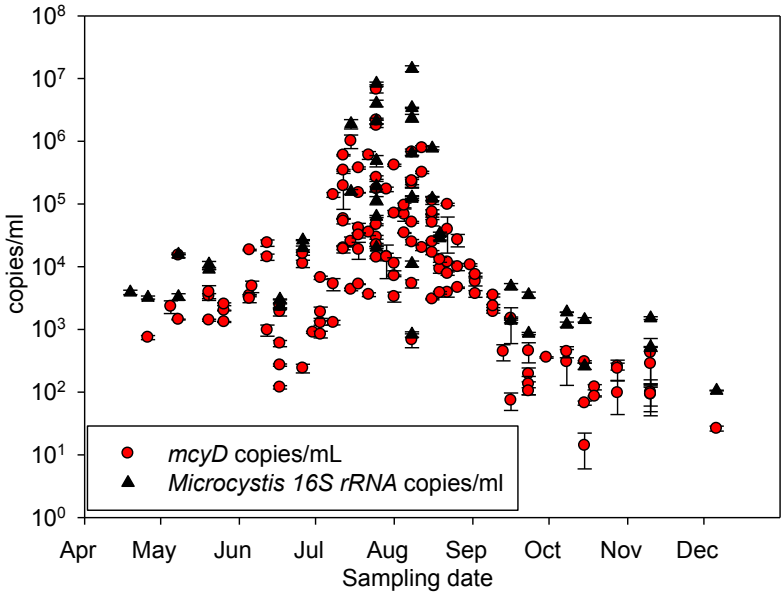

Figure 4.3 – *mcyD* and *Microcystis* 16S rRNA gene dynamics in MB stations during 2009. The error bars correspond to the standard deviations (n=3).

While temperature, DOC and pH values fluctuated with time in MB, their average annual values did not differ significantly among sampling stations (Kruskal Wallis test, Tw, p= 0.94; Ta, p= 0.73; pH= 0.96, DOC, p= 0.82). pH and DOC values were not recorded from tributary stations. The sampling campaign covered almost the full range of above-zero temperatures in the bay. The air temperature was 2.8 °C at the beginning of the sampling season (5/4/09) and -1 °C at the end (6/12/09). The maximum and minimum values recorded during the sampling season were 26.5 °C (18/8/09) and -1 °C (6/12/09). The average pH value in MB was 8.18; it ranged from 7.44 (23/9/09) to 9.31 (22/7/09). The average DOC value in MB was 5.15 mg/l, ranging from 3.89 (8/5/09) to 6.71 mg/l (29/7/09).

Nutrient concentrations fluctuated highly over time, and their pattern differed between stations (Fig 4.4). The spatial pattern of nutrient concentrations in the system was studied further in the next section. In general, MB experienced eutrophic conditions (CCME 2004) throughout 2009, with an average TP concentration of 70.22 µg/l, ranging from mesotrophic to hyper-eutrophic concentrations (Table 4.1). These concentrations far exceeded the TP annual mean concentrations recommended for this lake segment (25 µg/l) (NRD 2008).

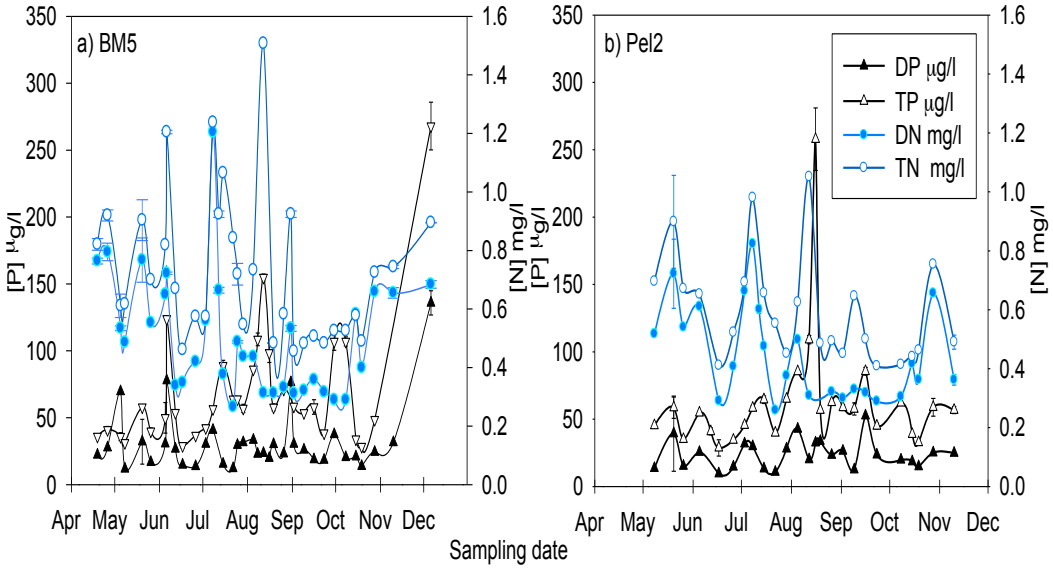

Figure 4.4 – Nutrient concentrations over time in a) Littoral station BM5 and b) Pelagic station Pel2. Error bars represent standard deviations.

The mean DP concentration observed in MB was 28.97 µg/l. The average TP (157.73 µg/l) and DP (64.59 µg/l) concentrations for the tributary stations in 2009 were higher than those observed in bay. The average TN and DN concentrations for the bay were 0.77 mg/l and 0.54 mg/l, respectively. As in the case of P, the average TN (2.64 mg/l) and DN (2.53 mg/l) concentrations for the tributary stations were higher than the ones observed in the bay.

Table 4.1 – Ranges of total and dissolved nutrient concentrations recorded in Missisquoi Bay and tributary stations during the 2009 sampling period.

Nutrients (units)	MB stations		Tributary stations	
	Minimum	**Maximum**	**Minimum**	**Maximum**
TP (µg/l)	28.49 BM5 19/10/09	267.98 BM5 6/12/09	25.95 EC 6/6/09	642.67 PR 17/6/09
DP (µg/l)	3.95 PelVen 12/7/09	135.82 BM5 6/12/09	19.17 EC 19/4/09	181.79 EC 17/6/09
TN (mg/l)	0.30 PelVen 28/10/09	2.28 PRM 20/5/09	0.84 PR 31/8/09	7.98 EC 17/6/09
DN (mg/l)	0.26 Pel2 22/7/09	2.19 PRM 20/5/09	0.81 EC 31/8/09	7.27 EC 17/6/09

Over time, nutrient concentrations differed between stations (Fig 4.4), and hence, so did the TN:TP ratios. As a result, periods of P or N limitation or both, were experienced at different points in time in the different stations (Fig 4.5). The relationships between TN:TP ratios and cyanobacterial growth and species composition were analyzed in subsequent sections.

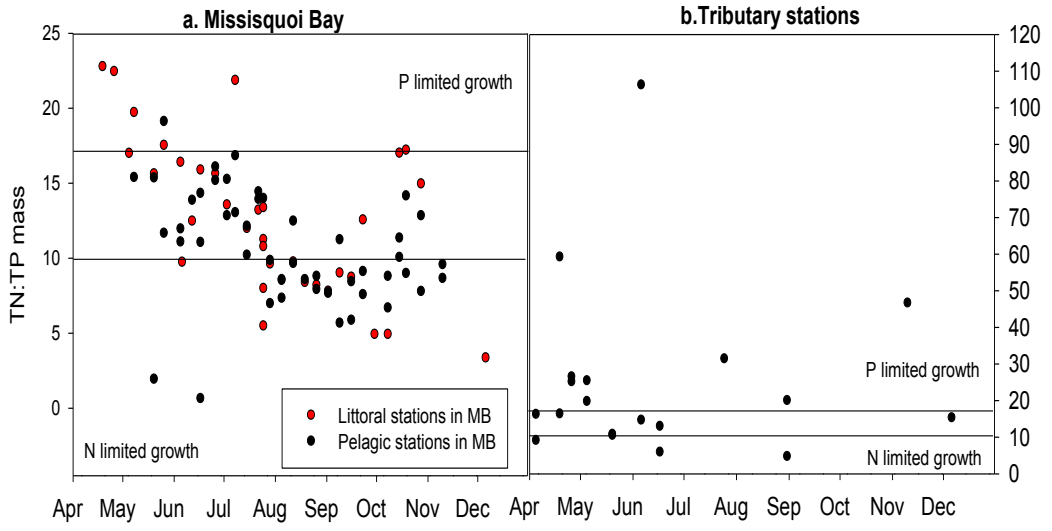

Figure 4.5 – TN:TP ratio over time in Missisquoi Bay and tributary stations.
The horizontal bars represent the thresholds for N and P limited growth in phytoplankton as determined by Sakamoto (1966): N-limited growth TN:TP <10:1 (mass); P-limited growth TN:TP>17:1 (mass). At intermediate TN:TP ratios either nutrient could be limiting growth.

4.2 Nutrient input and its relation with indicators of fecal contamination and cyanobacterial blooms in Missisquoi Bay/Pike River during 2009

In this section, the spatial distribution of nutrient concentrations and their relation with rainfall was determined. Ion Torrent sequencing of mtDNA amplicons was employed in order to identify potential sources of animal excreta contamination in the Bay. In addition, the suitability of *E. coli* as an indicator of nutrient input from external sources was explored. The aim was to increase knowledge about the nutrient sources that found their way into the bay and to analyze their relationship with the incidence of cyanobacterial blooms in this watershed during 2009.

4.2.1 Spatial analysis of phosphorus and nitrogen concentrations and their relationship with rainfall

In order to verify whether there was a trend of nutrient input from external sources, such as surface runoff, it was expected to see that in the Missisquoi Bay/PR area there would be higher nutrient concentrations in the stations with greater contact with the shore, and a relationship between rainfall and nutrient levels specially in the stations closest to the shore. For this analysis, the data was analyzed based on the station type and a rank was assigned to each station type depending on the degree of contact with the shore as described in section 3.6. To visualize in a simple manner the relationships studied, the data was graphed utilizing a categorical classification of rain and ranges of the nutrient concentrations (Fig 4.6-4.9).

In summary, spatially, all nutrient concentrations tended to significantly increase as station contact with the shore increased. For all nutrients, their mean sampling season values differed among station types; however, they were not statistically different between all station types for some nutrients, and only P (dissolved and total) was found to correlate with rain in at least one of the station types.

The sampling season mean TP concentration of each type of station was significantly different only between the pelagic and tributary stations (Table 4.2). The concentration of TP tended to significantly increase with increased contact with the shore; however, this correlation was not very strong (Table 4.3). The correlation between rain and TP was strong and significant but only detected in the tributary stations (Table 4.4) (Fig 4.6).

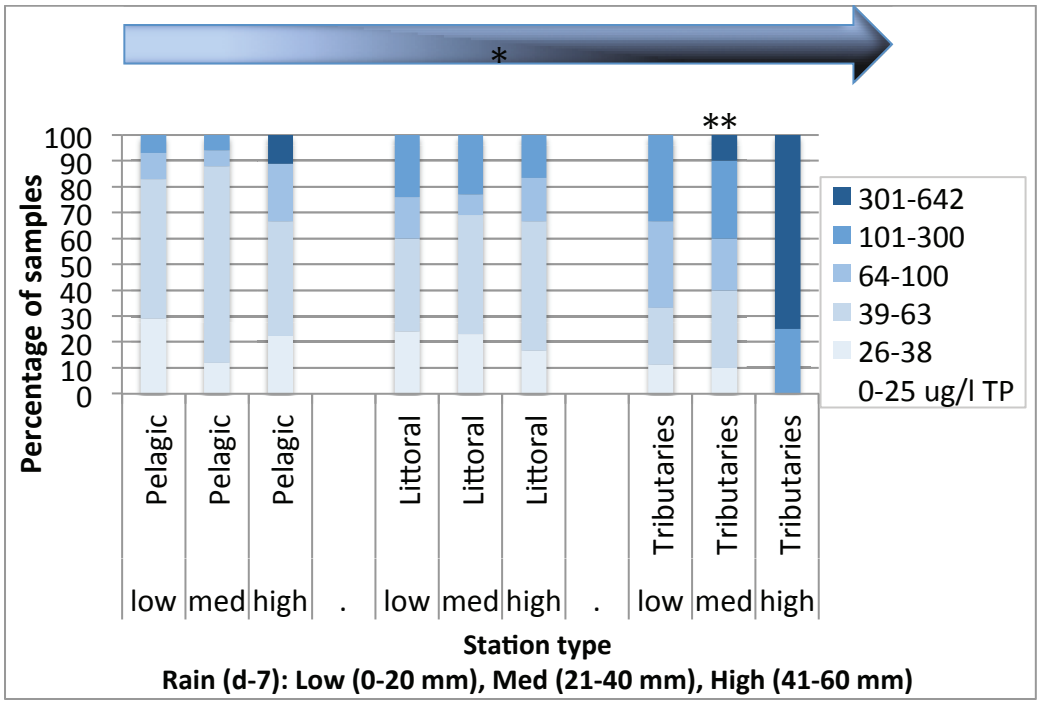

Figure 4.6 – Total phosphorus and its relationship with rainfall and station type.
The arrow represents the positive trend observed between TP and station contact with the shore (Table 4.3). Asterisks on top of the bars denote whether a significant correlation was found between rain and TP at each station type (Table 4.4). (*) $p<0.05$; (**) $p<0.01$.

The sampling season mean DP concentration of each type of station was significantly different between all stations except between the littoral and pelagic stations (Table 4.2). The concentration of DP tended to significantly increase as contact with the shore increased; this correlation was stronger than the one observed with TP (Table 4.3). The correlation between rain and DP was strong and significant; it was detected in the tributary stations as well as in the littoral stations (Table 4.4) (Fig 4.7).

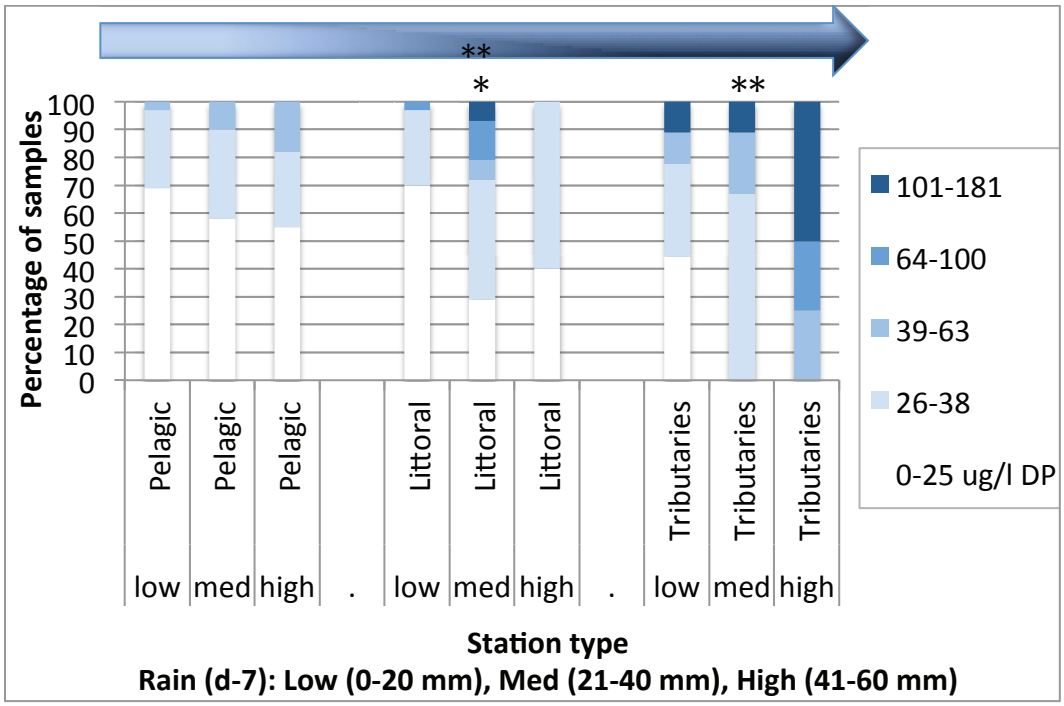

Figure 4.7 – Dissolved phosphorus and its relationship with rainfall and station type. The arrow represents the positive trend observed between DP and station contact with the shore (Table 4.3). Asterisks on top of the bars denote whether a significant correlation was found between rain and DP at each station type (Table 4.4). (*) $p<0.05$; (**) $p<0.01$.

Both TN and DN showed the same pattern. Their sampling season mean concentration for each station type was significantly different between the three station types (Table 4.2). A strong and significant correlation was found between the concentrations of both types of nitrogen and contact of the station with the shore (Table 4.3). There was no significant correlation between rain and DN or TN in any of the station types (Table 4.4) (Figs 4.8 and 4.9).

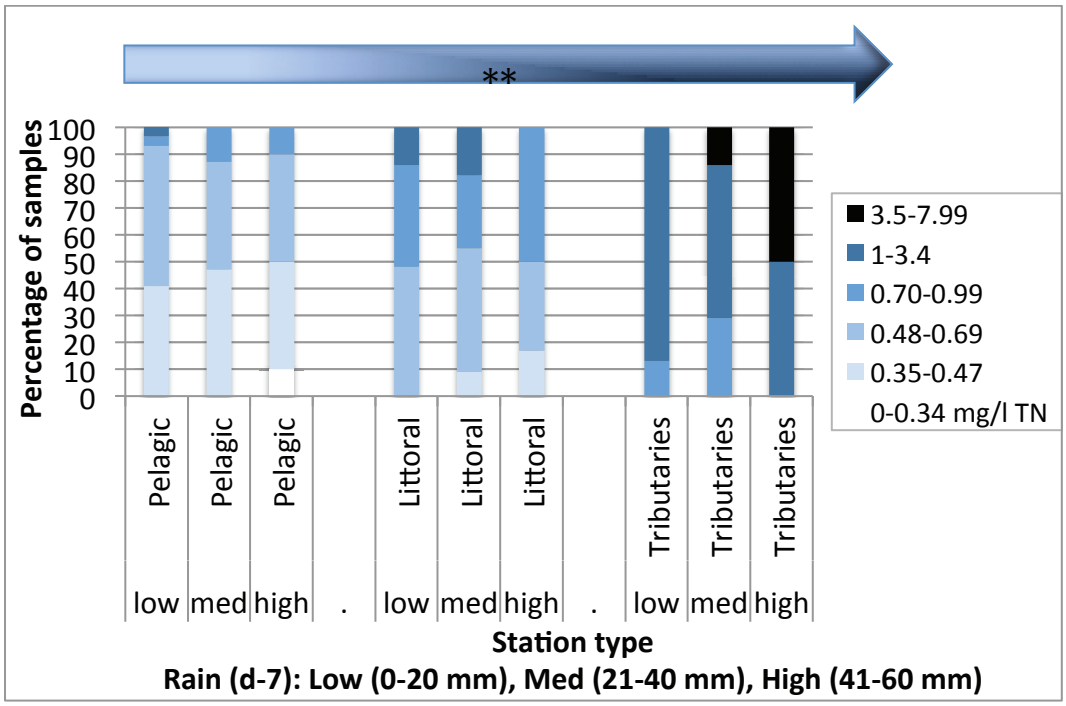

Figure 4.8 – Total nitrogen, and its relationship with rainfall and station type in 2009 samples. The arrow represents the positive trend observed between TN and station contact with the shore (Table 4.3). Asterisks on top of the bars denote whether a significant correlation was found between rain and TN at each station type (Table 4.4). (*) $p<0.05$; (**) $p<0.01$.

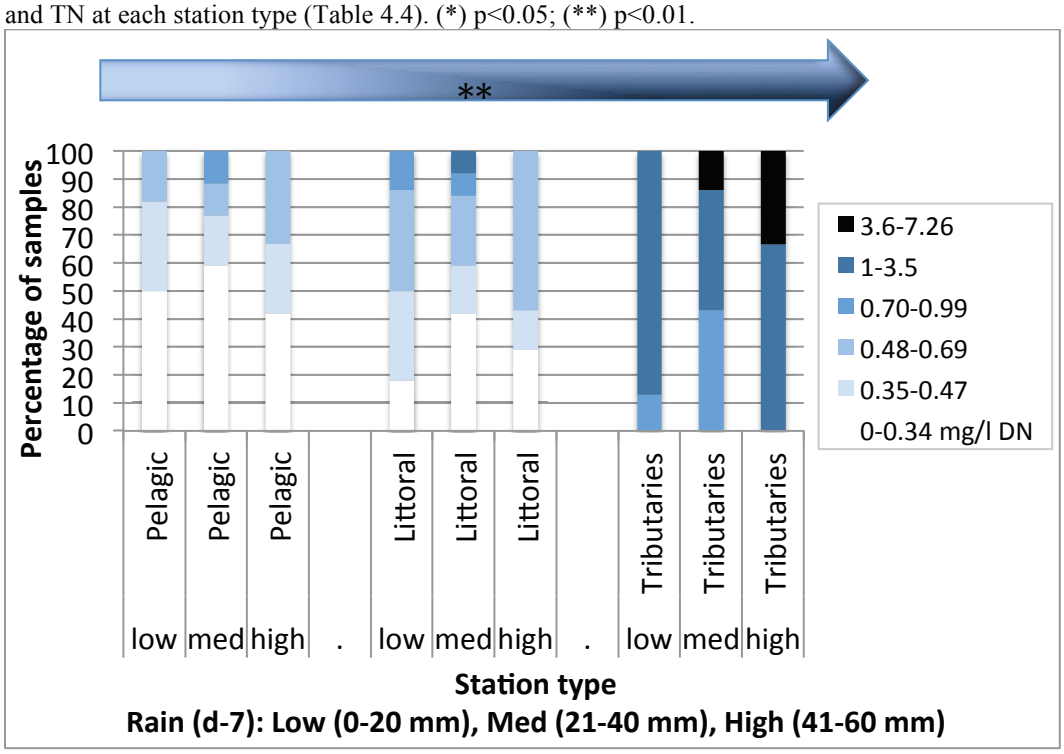

Figure 4.9 – Dissolved nitrogen, and its relationship with rainfall and station type in 2009 samples. The arrow represents the positive trend observed between DN and station contact with the shore (Table 4.3). Asterisks on top of the bars denote whether a significant correlation was found between rain and DN at each station type (Table 4.4). (*) $p<0.05$; (**) $p<0.01$.

Table 4.2 – Mean sampling season nutrient concentrations in the tributaries, littoral and pelagic stations of the Missisquoi Bay/Pike River area

Parameter	Pelagic	Littoral	Tributaries
DP (µg /l)	23.12[a]	30.24[a]	61.00[b]
TP (µg /l)	69.21[a]	70.84[ab]	157.73[b]
DN (mg/l)	0.39[a]	0.50[b]	2.40[c]
TN (mg/l)	0.53[a]	0.74[b]	2.50[c]

DP (dissolved phosphorus), TP (total phosphorus), DN (dissolved nitrogen), TN (total nitrogen).
The superscripts (a,b,c) identify which concentrations were significantly different between station types as determined by pairwise Kruskal-Wallis tests.

Table 4.3 – Relationships between increasing contact of the station type to the shore and nutrient concentrations as determined by Cochran–Mantel–Haenszel (CMH) and Spearman Rank correlations.

Parameter	**Cochran–Mantel–Haenszel** (CSMH correlation/p)	**Spearman Rank correlations** (Rho estimates/n/p)
DP	18.72(*)/p<0.0001	0.40(*)/126/ p<0.01
TP	4.37(*)/p=0.0366	0.21(*)/121/ p<0.05
DN	48.46(*)/p<0.0001	0.64(*)/119/ p<0.01
TN	55.63(*)/p<0.0001	0.72(*)/111/ p<0.01

Abbreviations: DP (dissolved phosphorus), TP (total phosphorus), DN (dissolved nitrogen), TN (total nitrogen), n (number of data points in the correlation tested), p (p-value).
(*) indicates a significant correlation.

Table 4.4 – Spearman Rank correlations between Rainfall (rain d-7) and nutrient concentrations at each station type in the Missisquoi Bay/Pike River area in 2009

Parameter	Tributaries	Littoral	Pelagic
DP	0.55(*)/22/ 0.008	0.37(*)/45/ /0.01	0.07/59/ 0.6
TP	0.65(*)/23/ 0.0009	0.05/44 /0.8	0.09/54/ 0.5
DN	0.42/21/ 0.059	-0.055/41 /0.7	0.06/57 /0.7
TN	0.40/21/ 0.07	0.06/38/ 0.7	-0.07/52/ 0.6

The data is displayed in the following order: Spearman rho estimates (rs)/number of data points in the correlation tested (n)/ p value. (*) indicates a significant correlation.

4.2.2 *E. coli* as a potential indicator of recent nutrient input from external sources

Nutrient concentrations were found to be higher in the stations with greater contact with the shore, and a correlation between rainfall and P concentrations were found in these stations, suggesting a pattern of nutrient input from external sources. To verify whether nutrient (N and P) external loads contribute significantly to the nutrient concentrations observed in the bay, *E. coli* counts were used as an indicator of nutrient input from external sources, such as nutrient rich surface runoff or sewage overflows. It was first determined whether *E. coli* was a suitable indicator for recent nutrient input from external sources. In order for *E. coli* to be a suitable indicator it was expected that 1) *E. coli* counts increased as the contact of the station with the shore increased, 2) a positive relationship between *E. coli* and rainfall would be observed at the station closest to the shore (Fig 4.10), and 3) a positive association between *E. coli* and the presence of mitochondrial DNA of non-aquatic species in the bay would be observed.

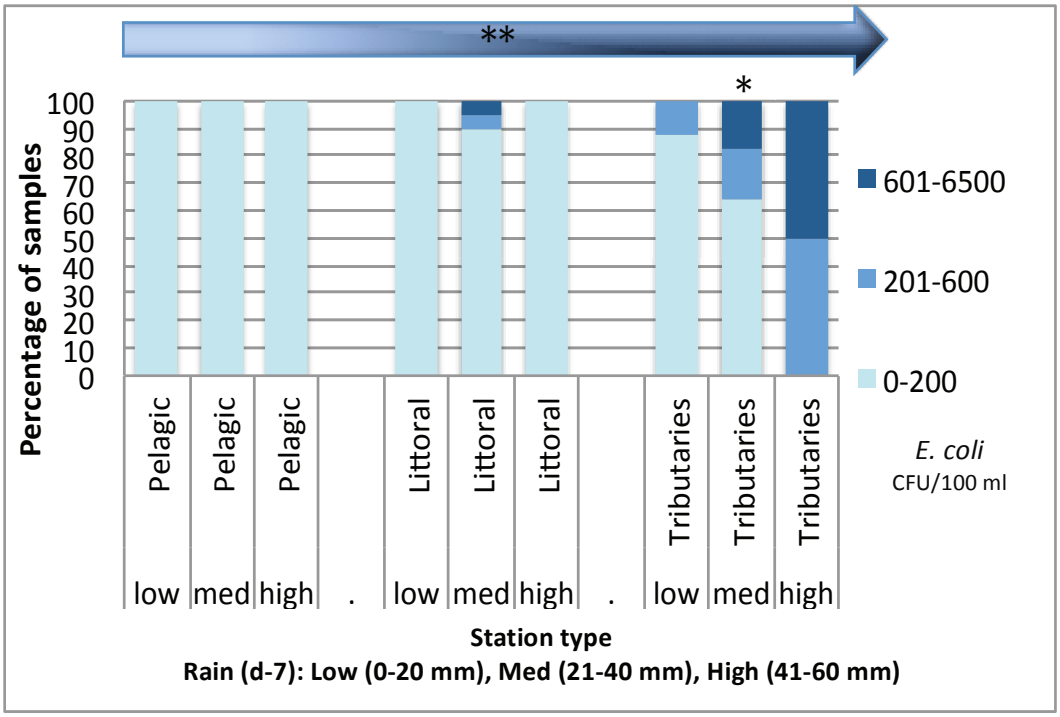

Figure 4.10 – *E. coli* counts, and its relationship with rainfall and station type in 2009 samples. The arrow represents the positive trend observed between *E. coli* counts and station contact with the shore. Asterisks on top of the bars denote whether a significant correlation was found between rain and *E. coli* counts at each station type. (*) $p<0.05$; (**) $p<0.01$.

The geometric mean *E. coli* counts were 1.30, 3.83 and 79.56 CFU/100 ml for the pelagic, littoral and tributary stations, respectively. The sampling season average *E. coli*

counts were 3.65 CFU/100 ml in the pelagic stations; 92.82 CFU/100 ml in the littoral stations; and 528.57 CFU/100 ml in the tributary stations. The sampling season mean *E. coli* counts of each type of station was significantly different between the littoral and tributary stations (Kruskal Wallis test Chi sq=16.99, p<0.0001) and between the pelagic and tributary stations (Kruskal Wallis test Chi sq=23.35, p<0.0001). The *E. coli* counts tended to increase as the station contact with the shore increased (CSMH correlation=23.45, p<0.0001; rs=0.57, n=79, p<0.0001). The correlation between rain and *E. coli* counts was only significant in the tributary stations (rs=0.48, n=34, p<0.03) (Fig 4.10). From the samples analyzed for mitochondrial DNA (n=27), in all of the samples in which *E. coli* was detected (n=17) non-aquatic mitochondrial DNA hosts were also detected. On the other hand, no *E. coli* was detected in samples in which there was no detection of mtDNA from non-aquatic hosts. Hence, the likelihood of detecting *E. coli* was evidently higher in samples in which mtDNA from non-aquatic hosts was detected (Fisher exact test one-tailed p=0.01). This positive association between the presence of non-aquatic mtDNA hosts and *E. coli* provides evidence for the presence of *E. coli* from external sources and also suggests that these external sources include animal excreta.

Table 4.5 – Relationships between *E. coli* counts and nutrient concentrations as determined by Spearman Rank correlation on the 2009 data from the Missisquoi Bay/Pike River area.

	TP	**DN**	**TN**	***E. coli***
DP	0.48/122**	0.50/117**	0.40/111**	0.51/61**
TP		0.23/109**	0.44/106**	0.36/56**
DN			0.82/114**	0.65/52**
TN				0.63/46**

The data is displayed in the following order: Spearman rho estimates (rs)/number of data points in the correlation tested (n)/ p value. (NS) indicates a correlation with a p value>0.05; (*) p<0.05; (**) p<0.01. DP (dissolved phosphorus), TP (total phosphorus), DN (dissolved nitrogen), and TN (total nitrogen). For the correlations, the quantitative data was used for all parameters.

It is worth mentioning that there was no correlation between *E. coli* counts and air or water temperature. In summary, these analyses suggests that *E. coli* counts in surface water samples may be used as an indicator of nutrient input from external sources. The

positive correlation observed between *E. coli* counts and all of the nutrients (Table 4.5), suggests that nutrients from external sources contribute significantly to the nutrient concentrations observed in the system.

4.2.3 Potential sources of nutrients into Missisquoi Bay/Pike River area from animal excreta

Since *E. coli* counts were significantly higher in the tributary stations when compared to the other station types, a similar relationship was expected with mtDNA from non-aquatic species. All the PR samples were therefore sent for mtDNA sequencing in order to identify potential sources of nutrients into the system from non-aquatic animal excreta. In addition, 2 samples collected on the same dates were sequenced from each MB station in order to compare whether the non-aquatic mtDNA hosts differed within stations (Fig 4.11).

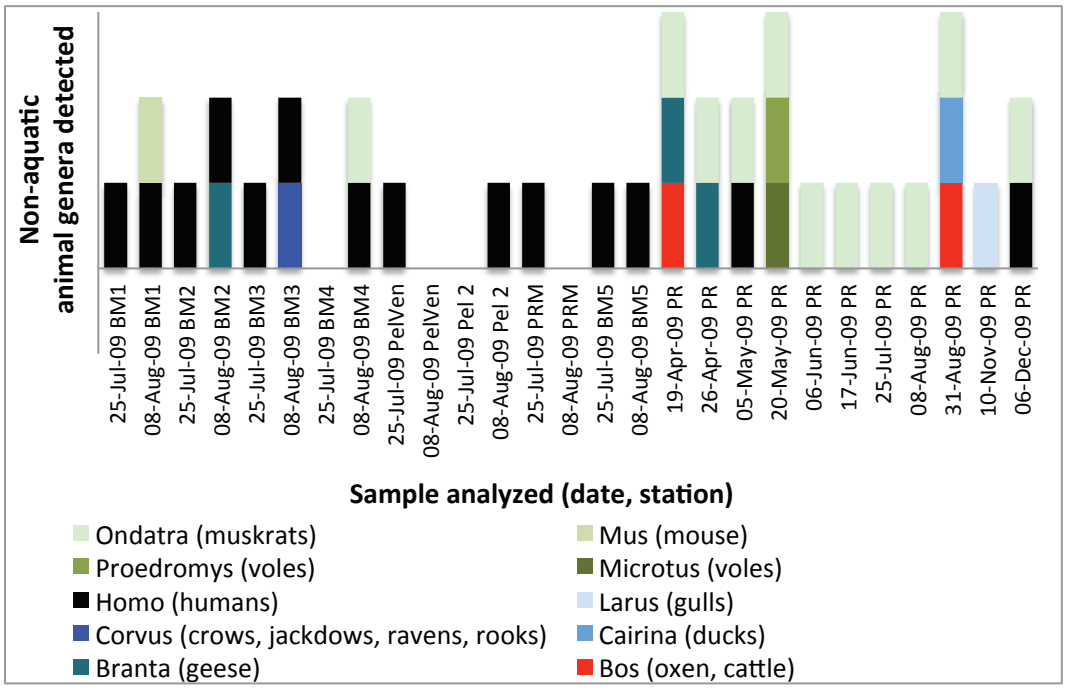

Figure 4.11 – Presence of non-aquatic mitochondrial DNA hosts identified to the genus level in Missisquoi Bay/Pike River area.
For the purpose of this project, aquatic species were defined as species whose entire life cycle takes place in water. The X-axis shows the sampling date and station of the samples analyzed. The Y-axis denotes the animal genera detected in the sample. A genus was considered present/detected in the sample if represented by at least 5 mtDNA sequences.

From the 27 samples sequenced 175,115 reads were recovered. A total of 82,723 sequences met the quality criteria established (Section 3.3.2.3). The average depth of

coverage was 3,063 sequences per sample; however the variation was large (min: 65 sequences per sample; max: 10,553 sequences per sample). For the purpose of this project, aquatic species were defined as species whose entire life cycle takes place in water. Due to the fact that a nested PCR was required in order to produce the mtDNA amplicons sent for sequencing, the output could not be treated as quantitative; hence a qualitative taxon-based analysis was conducted. A genus was considered detected/present in the sample if represented by at least 5 mtDNA sequences.

The non-aquatic mtDNA hosts observed in the bay included rodents, birds, cattle and humans. Contrary to what was expected in an agricultural watershed, the presence of farm animals was very low in the samples tested. Indeed the *Bos* genus was only found in 2 PR samples. When comparing the samples from Jul 25th and August 8th across the system, it could be observed that the human presence was stronger in the bay stations compared to the tributary station PR where wildlife, particularly rodents, was predominant (Fig. 4.11).

It is worth noting that mtDNA from non-aquatic hosts was not detected in 4 of the samples and that fish genera were only detected in PR (data not shown).

4.2.4 Relationship between toxic cyanobacterial blooms, nutrient input and *E. coli* in Missisquoi Bay during 2009

When analyzing the relationships between *E. coli* counts from surface water samples - our indicator of nutrient input from external sources - and biological parameters of cyanobacterial abundance and toxicity (i.e. cyanobacterial abundance (cell/ml), microcystin concentrations (µg/l)), *E. coli* counts correlated significantly with cyanobacterial cell abundance (r_s=0.41, n=34, p=0.0172). This suggests that there is a link between nutrient input from external sources and the incidence of cyanobacterial blooms in the bay.

4.3 Spatio-temporal characterization of major cyanobacterial taxa and their association with environmental parameters in Missisquoi Bay/Pike River during 2009

High-throughput amplicon sequencing of the partial cyanobacterial 16S rRNA gene was performed on 57 water samples (see Appendix A.2 for sequencing output details), in order to verify which environmental parameters were significantly related to the cyanobacterial dynamics in the bay (section 4.3.1). Based on this information, the

cyanobacterial species composition through time and space was characterized in detail in relation to the fluctuating environmental parameters. Major cyanobacterial taxa and potential microcystin producers were also identified (section 4.3.2 and 4.3.3).

To follow the succession in each of the stations, the littoral station BM5 (Fig 4.12), the pelagic station Pel2 (Fig 4.13) and the tributary station PR (Fig 4.14) were analyzed. In order to characterize the cyanobacterial species composition at one point in time across the bay, all the MB samples from July 25th were sequenced (Fig 4.15). There was a particular interest to elucidate the cyanobacterial composition on this date since the highest *mcyD* concentrations in the bay were registered on July 25th at Pel2.

From the 57 samples sequenced in this study, a total of 306,380 sequences met the quality criteria established (Section 3.3.2.3). There was an average depth of coverage of 5,375 sequences per sample (min: 1,375 sequences per sample; max: 8,012 sequences per sample). The average number of cyanobacterial sequences per sample was 1,208. However, as expected, the number of cyanobacterial sequences per sample varied depending on the station and the time of the year. All the cyanobacterial orders, except for the *Gloeobacterales,* were observed in all of the stations at some point in time. The major orders were *Chroococcales, Nostocales* and *Oscillatoriales.* In the order *Chroococcales*, the predominant genera were *Microcystis* and *Synechococcus,* but the relative abundance of *Chroococcales* other than *Microcystis* was generally lower than the relative abundance of *Microcystis* alone. In the order *Nostocales* the predominant genus was *Anabaena;* and in the order *Oscillatoriales* the predominant genus was *Leptolyngbya.*

In BM5, the community appeared to be dominated by members of the *Chroococcales*, *Nostocales* and *Oscillatoriales* (Fig 4.12). From the *Chroococcales*, the genus *Microcystis* was present in most of the samples analyzed throughout the sampling season. The orders *Stignometales*, *Prochlorales* and *Pleurocapsales* had very low relative abundances (<1%) throughout the samples with the *Pleurocapsales* reaching a maximum of 13.9% on Dec 6/09. The main bloom at BM5 took place during July and August when the cyanobacterial community was mainly dominated by members of the order *Chroococcales*, particularly from the genus *Microcystis*, and by members of the order *Nostocales*. A smaller bloom took place in September, when *Nostocales* and *Chroococcales* other than *Microcystis,* mainly dominated the community.

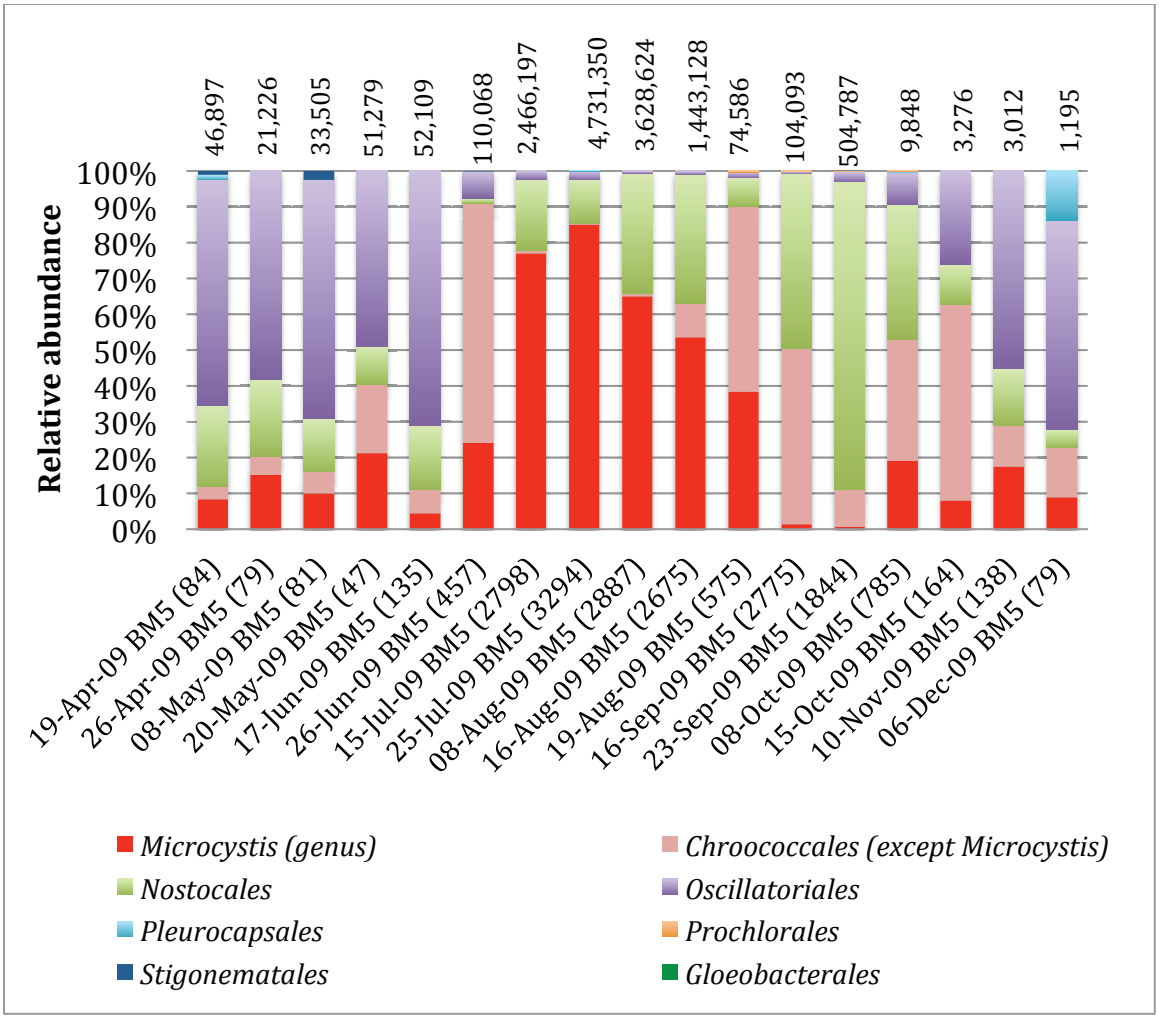

Figure 4.12 – Temporal variation of *Microcystis* in relation to cyanobacterial orders in the littoral station BM5 during 2009.

Two samples per month were sequenced in order to provide insight into the temporal cyanobacterial succession at the sampling station. The X-axis shows the sampling date and the number of total cyanobacterial sequences identified to the order level in parenthesis. Only *Microcystis* is shown from the total genera identified. The percentages on the Y-axis reflect the relative abundance of the orders/genera in the sample. The numbers on top of the bars represent the estimated total cyanobacterial 16S rRNA gene copies/ml.

In Pel2, a similar pattern to the one in BM5 was observed (Fig 4.13). The community was again dominated by members of the *Chroococcales*, *Nostocales* and *Oscillatoriales*. From the *Chroococcales*, the genus *Microcystis* was present in most of the samples analyzed throughout the sampling season. The orders *Stignometales*, *Prochlorales* and *Pleurocapsales* had low relative abundances (<1%) throughout the samples with the *Stignometales* reaching a maximum of 1.8% on May 8/09. At this station, the main bloom also took place during July and August when the cyanobacterial

community was mainly dominated by members of the order *Chroococcales*, particularly from the genus *Microcystis*, and by members of the order *Nostocales*. A smaller bloom took place during September, when *Nostocales* and *Chroococcales,* other than *Microcystis,* mainly dominated the community. In both BM5 and Pel2, there seemed to be an inverse relationship between *Microcystis* and the relative abundance of *Chroococcales* other than *Microcystis* and *Oscillatoriales*.

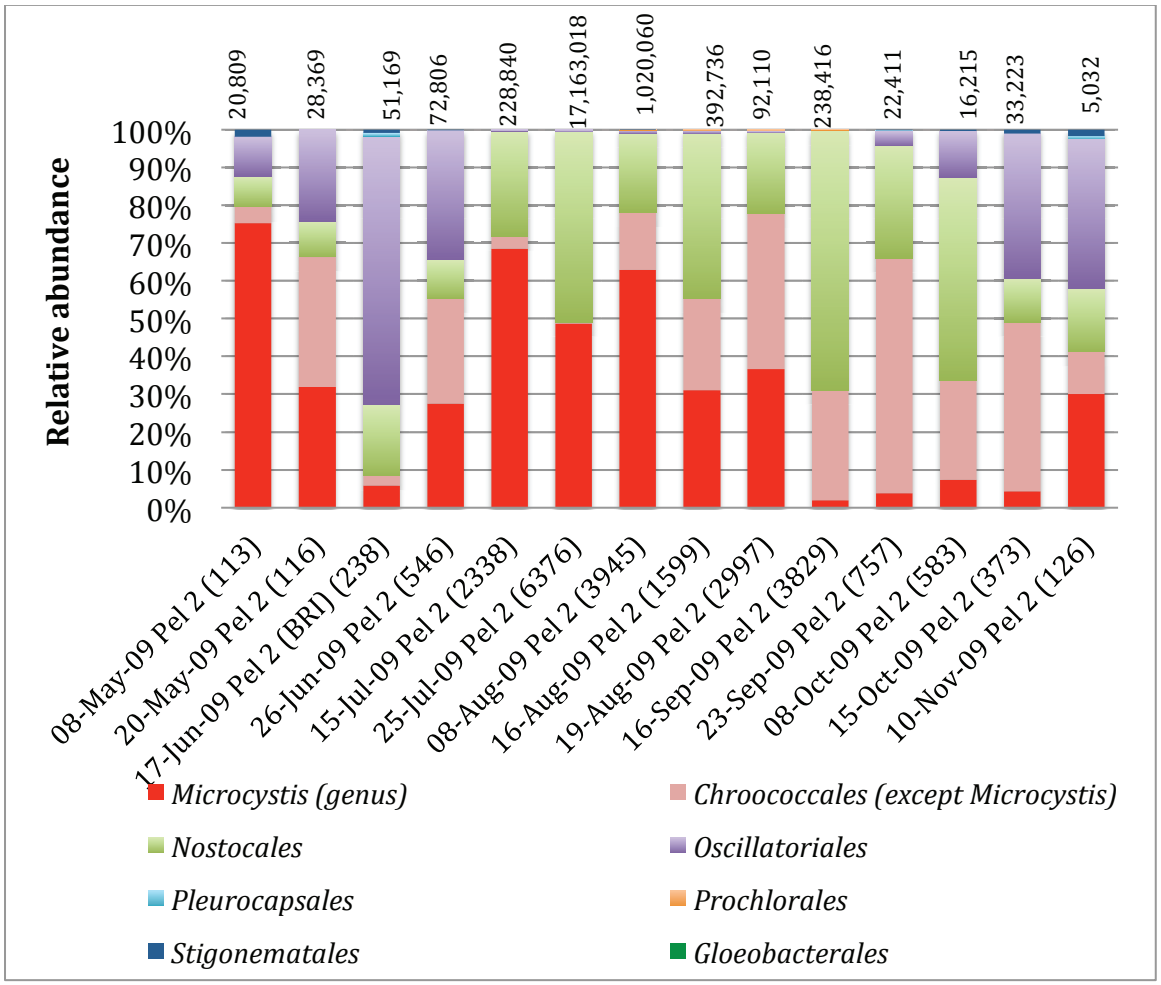

Figure 4.13 – Temporal variation of *Microcystis* in relation to cyanobacterial orders in the Pelagic station Pel2 during 2009.

Two samples per month were sequenced in order to provide insight into the temporal cyanobacterial succession at the sampling station. The X-axis shows the sampling date and the number of total cyanobacterial sequences identified to the order level in parenthesis. Only *Microcystis* is shown from the total genera identified. The percentages on the Y-axis reflect the relative abundance of the orders/genera in the sample. The numbers on top of the bars represent the estimated total cyanobacterial 16S rRNA gene copies/ ml.

A completely different pattern was observed in PR in terms of major orders and their relative abundances (Fig 4.14). The orders dominating the samples were *Pleurocapsales*, *Oscillatoriales* and *Chroococcales*. *Microcystis* was found at low relative abundances and only in two samples (1% and 11% on July 12 and August 8, respectively). It is worth noting that the total number of cyanobacterial sequences was very low in each PR sample, probably due to an unfavourable hydrodynamic regime for cyanobacteria. Due to the low number of sequences, the data from this station was not analyzed further and was not included in the statistical analysis.

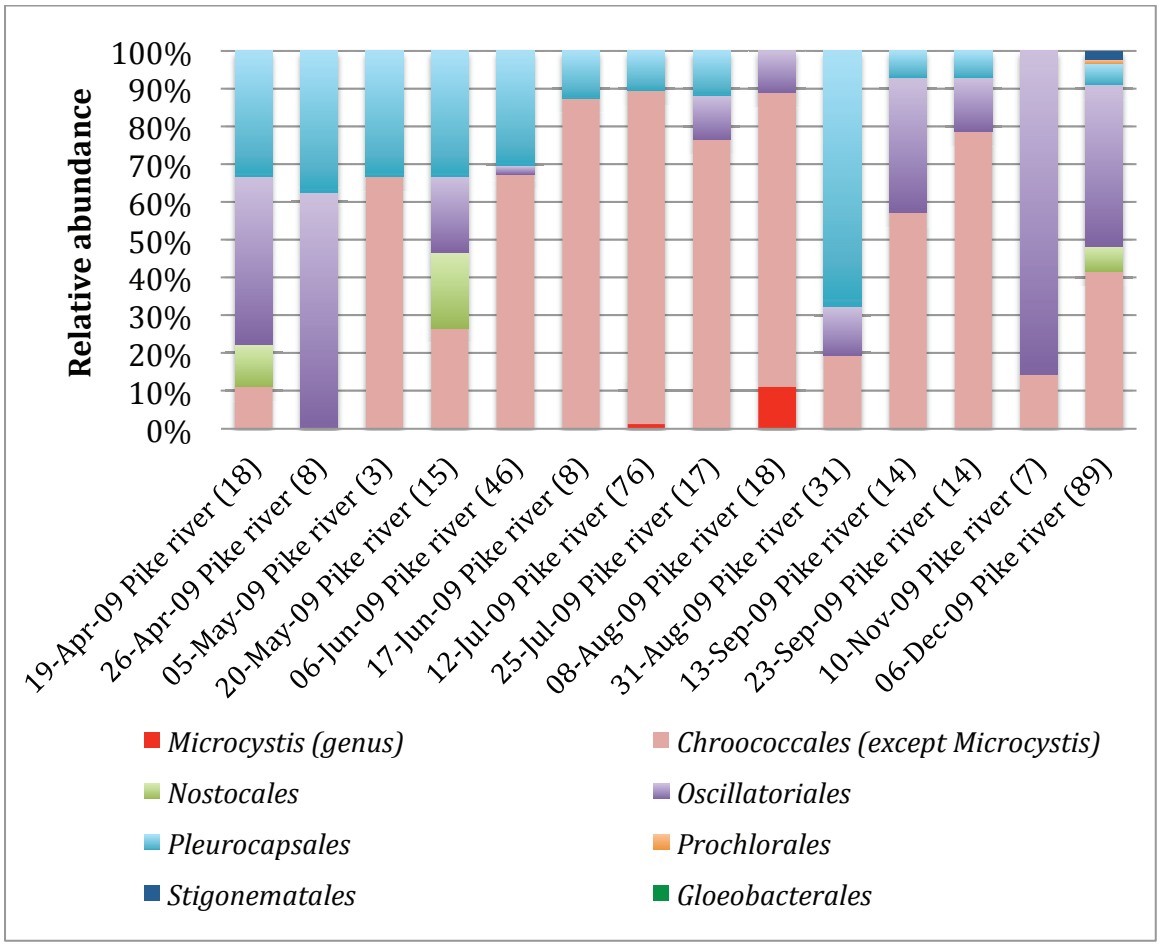

Figure 4.14 – Temporal variation of *Microcystis* in relation to cyanobacterial orders in the tributary station Pike River during 2009.
Two samples per month were sequenced in order to provide insight into the temporal cyanobacterial succession at the sampling station. The X-axis shows the sampling date and the number of total cyanobacterial sequences identified to the order level in parenthesis. Only *Microcystis* is shown from the total genera identified. The percentages on the Y-axis reflect the relative abundance of the orders/genera in the sample.

On July 25th, the cyanobacterial community throughout the bay was composed of potentially toxic taxa (Fig 4.15), especially from the orders *Nostocales* and *Chroococcales*. The genus *Microcystis* was strongly present throughout the bay, but it was not the dominating taxon in all the sampling stations. This observation is further analyzed in section 4.3.3.

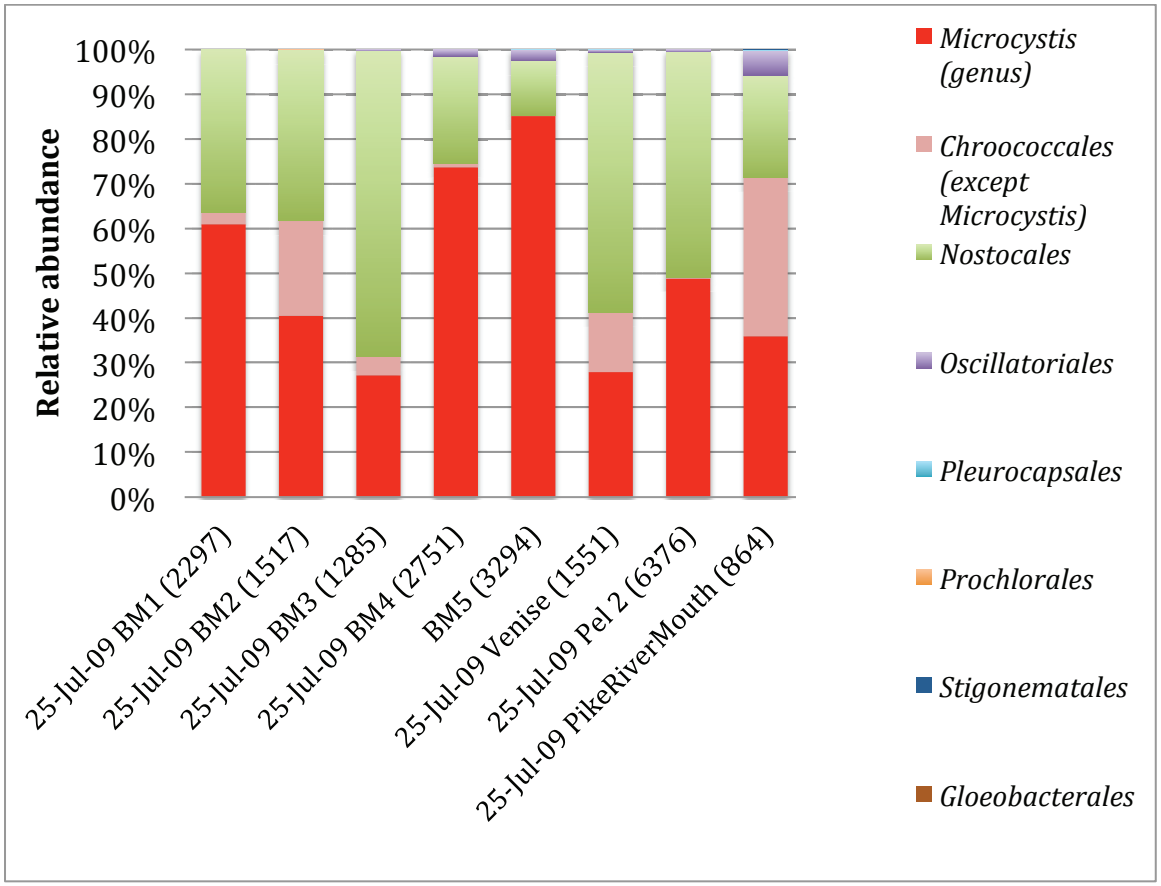

Figure 4.15 – Spatial variation of *Microcystis* in relation to cyanobacterial orders in July 25, 2009 in the different stations in Missisquoi Bay.
The X-axis shows the sampling date and the number of total cyanobacterial sequences identified to the order level in parenthesis. The percentages on the Y-axis reflect the relative abundance of the orders/genus in the sample.

4.3.1 Significant relations between the relative abundance of major cyanobacterial taxa and environmental parameters in Missisquoi Bay

Spearman rank correlations (rs) were performed between the relative abundances of major cyanobacterial taxa in MB (i.e *Microcystis*, *Chroococcales* other than *Microcystis*, *Nostocales* and *Oscillatoriales*) and the different environmental parameters in order to verify which parameters were significantly related to the cyanobacterial dynamics observed in Missisquoi Bay. The environmental parameters tested were air temperature (Ta), water temperature (Tw), pH, dissolved organic carbon (DOC), chlorophyll a (Chl a), DP, TP, TN, DN, TN:TP (mass), microcystin concentrations and *mcyD* gene abundance. For this purpose all the data from the sequenced MB samples (i.e. all stations excluding tributaries) were included in the analysis (n= 43).

In addition, thresholds of physicochemical parameters in which each taxa was observed to dominate (i.e. relative abundance higher than 50%), were established independently for each parameter, based on the data from the samples sequenced that were included in the statistical analysis (Table 4.6).

It was not always possible to identify a threshold in a specific parameter when the taxon dominated, because in some instances, the taxon dominated at wide ranges of the parameter. For example, *Microcystis* dominance was mostly observed at pH levels higher than 8.66, with only one exception where *Microcystis* reached 75% of the sample at a pH of 7.5.

In general, *Oscillatoriales* dominance was mainly observed at TP concentrations lower than 42 µg/l and TN:TP mass ratios higher than 14:1, with one exception on December 6th, in which the TP concentration was 268 µg/l and the TN:TP ratio was 3.33:1. On that date, the air temperature was -1 °C. The relative abundance of *Oscillatoriales* did not exceed 10% at water temperatures higher than 15 °C, except for one occasion.

Table 4.6 – Significant relationships between the relative abundance of cyanobacterial taxa and physicochemical parameters and ranges of these parameters in which the taxa was observed to dominate (determined independently for each parameter).

Cyanobacterial taxa	Parameter	Spearman Rank correlation (rho/n/p)	Parameter threshold (taxa >50%)
Microcystis	Tw (°C)	0.45/ 43/ <0.01	≥14.4
	pH	0.57/ 38/ <0.01	Not distinct*
	TP (µg/l)	0.44/ 37/ <0.01	≥39.23
	TN (mg/l)	0.52/ 32/ <0.01	≥0.52
	mcyD gene (copies/ml)	0.81/ 42/ 0	≥15,218
	Extracellular microcystin (µg/l)	0.47 /20/ 0.04	Not distinct*
	Intracellular microcystin (µg/l)	0.79/ 33/ <0.01	≥0.37
Chroococcales* other than *Microcystis	*mcyD* gene (copies/ml)	- 0.51/ 42/ <0.01	≤51,091
	Intracellular microcystin (µg/l)	- 0.68/ 33/ <0.01	≤0.36
Nostocales	DOC (mg/l)	0.44/ 20/ 0.049	≥4.91
	TP (µg/l)	0.38/ 37/ 0.02	≥38.51
	DN (mg/l)	- 0.35/ 33/ 0.049	≤0.39
	TN:TP (mass)	- 0.40/ 30/ 0.03	≤12.54
Oscillatoriales	Tw (°C)	- 0.69/ 36/ <0.01	Not distinct*
	pH	- 0.66/ 38/ <0.01	≤7.81
	DOC (mg/l)	- 0.74/ 20/ <0.01	≤4.49
	TP (µg/l)	- 0.45/ 37/ <0.01	Not distinct*
	TN:TP (mass)	0.55/ 30/ <0.01	Not distinct*
	mcyD gene (copies/ml)	- 0.55/ 42/ <0.01	≤2,419
	Intracellular microcystin (µg/l)	- 0.39/ 33/ 0.03	≤0.12

*Not distinct refers to the fact that the taxon was observed to dominate at concentrations of the parameter encompassing its whole range. Refer to the text for details.

-**Note**: The interpretation of these thresholds should be done cautiously. For example, percentages of *Microcystis* per sample higher than 50% were only observed at temperatures higher than 14°C; nonetheless this does not mean that higher temperatures always coincided with dominance of *Microcystis*. This only means that in the samples where *Microcystis* dominated, the temperatures were higher than 14°C.

Relative abundance of *Microcystis* and other taxa

During the visual analysis of the temporal profiles (Fig 4.12- 4.13) it was noted that in terms of relative abundance, some cyanobacterial taxa tended to have an inverse relationship with the dynamics of *Microcystis* in MB. This relationship was significant for *Chroococcales* other than *Microcystis* ($rs= -0.52$, $p<0.01$), and for *Oscillatoriales* ($rs= -0.41$, $p<0.01$).

Potential toxin producing taxa

Only the relative abundance of *Microcystis* significantly and positively correlated with microcystin and *mcyD* gene concentrations (Table 4.6) despite the fact that other potential toxin producing genera were present, such as *Leptolyngbya* and *Anabaena* (correlation data not shown for these genera). This suggests that *Microcystis* was the main microcystin producer in the MB.

4.3.2 Succession of major cyanobacterial taxa and environmental parameters in littoral station BM5 and pelagic station Pel2

Based on previous results, only the orders *Chroococcales, Nostocales and Oscillatoriales* were analyzed in detail. The significant relationships found between the relative abundance of cyanobacterial taxa and the physicochemical and biomolecular parameters in MB (Table 4.6), describe for the most part, the succession of cyanobacterial taxa observed in the stations BM5 (Fig 4.12) and Pel2 (Fig 4.13). When there were exceptions, or when some of the environmental conditions were outside the ranges in which dominance by each taxa was generally observed, nutrient concentrations and temperature seemed to determine the final outcome.

In BM5 (Fig 4.12), *Oscillatoriales* dominated in samples from April to mid-June; when they reached a peak. From April to May 8th, the environmental conditions in the station complied with the general concentrations in which this taxon usually dominated in the bay (Fig 4.16). On May 20, an increase in TP (57.8 µg/l) was accompanied by a relative decrease in the dominance of *Oscillatoriales* and an increase in *Chroococcales*. However, *Microcystis* did not

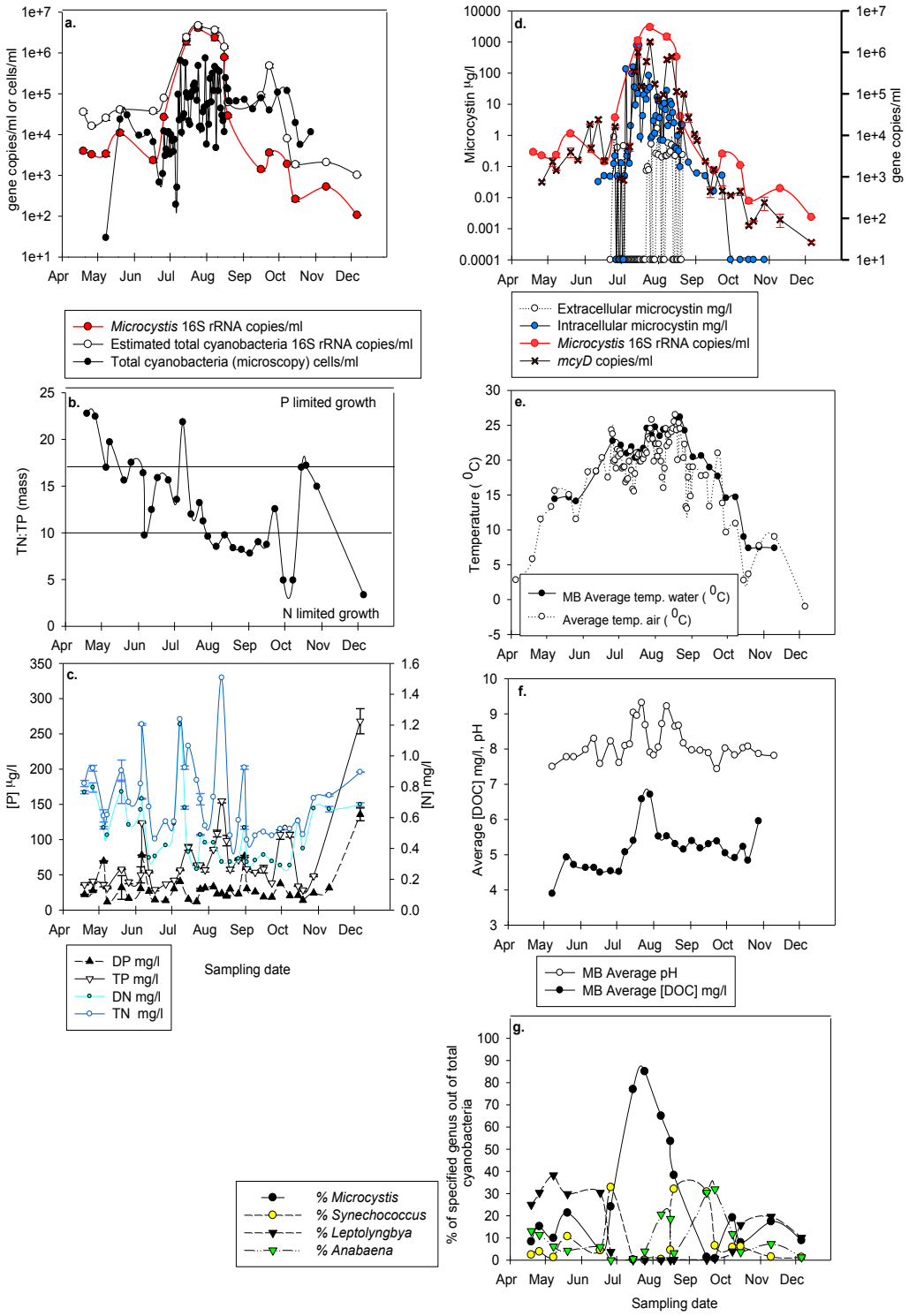

Figure 4.16 – Temporal profile of the dynamics of environmental and biomolecular parameters in littoral station BM5.

a) Total cyanobacteria abundance (microscopy), *Microcystis* and total cyanobacteria 16S rRNA gene abundance; b) TN:TP ratio; c) total and dissolved N and P concentrations; d) microcystin concentration, *mcyD* gene and *Microcystis* 16S rRNA gene abundance; e) temperature; f) pH and dissolved organic carbon (DOC) concentrations; g) relative abundance of main cyanobacteria genera at BM5. Figures e and f reflect MB averages as opposed to BM5 specific data.

dominate despite the presence of nutrient concentrations at which it dominated on other dates, perhaps due to the relatively low water temperature (14.7°C). On June 17, the phosphorus concentrations were relatively low (TP=29.05 µg/l) and *Oscillatoriales* peaked. In late June, with an increase in temperature and phosphorus concentrations (TP=36.76 µg/l; 23.72°C), we observed a shift of dominance towards the *Chroococcales*, where genera other than *Microcystis* prevailed. The *Chroococcales* and the *Nostocales* dominated during the summer. However, it was interesting to note that the relative abundance of *Microcystis* alone was significantly higher than the relative abundance of the *Nostocales* from July 15 to August 19 (inclusive) (Wilcoxon-Mann-Whitney $p<0.01$). *Microcystis* reached its dominance peak in late July when the water temperature was 20.54°C, TP=64.00 µg/l, and TN=0.72 mg/l and the pH=8.68, and subsided towards the end of August, where the concentrations of TN were below 0.52 and the TN:TP mass was below 9:1. The *Nostocales* reached its dominance peak towards the end of September when the water temperature was 17.7°C, DOC=5.38 mg/l, TN= 0.48 mg/l, DN= 0.32 mg/l, TP=38.51 µg/l and the TN:TP mass was about 12.5:1. As the temperature started to decrease in October, the relative abundance of *Oscillatoriales* started to increase and dominate while the presence of other orders decreased in the samples despite the fact that the nutrient concentrations were particularly high at the end of the sampling campaign, generating an apparent nitrogen limiting environment (Ta= -1°C; TP= 267.98; TN=0.90 mg/l; TN:TP mass=3.33:1).

In Pel2 a slightly different pattern of dominance was observed (Fig 4.13). This variation in cyanobacterial dominance between the two stations might stem from the different dynamics observed in terms of nutrient concentrations. While there were no significant differences between the annual averages of the environmental parameters analyzed in the two sampling stations, the TN concentrations were significantly higher at BM5 (Table 4.7). In addition, in each of the stations the nutrients seemed to peak at different dates over time, as did the pattern of nutrient limitation (Fig 4.16 b, c; Fig 4.17 b, c).

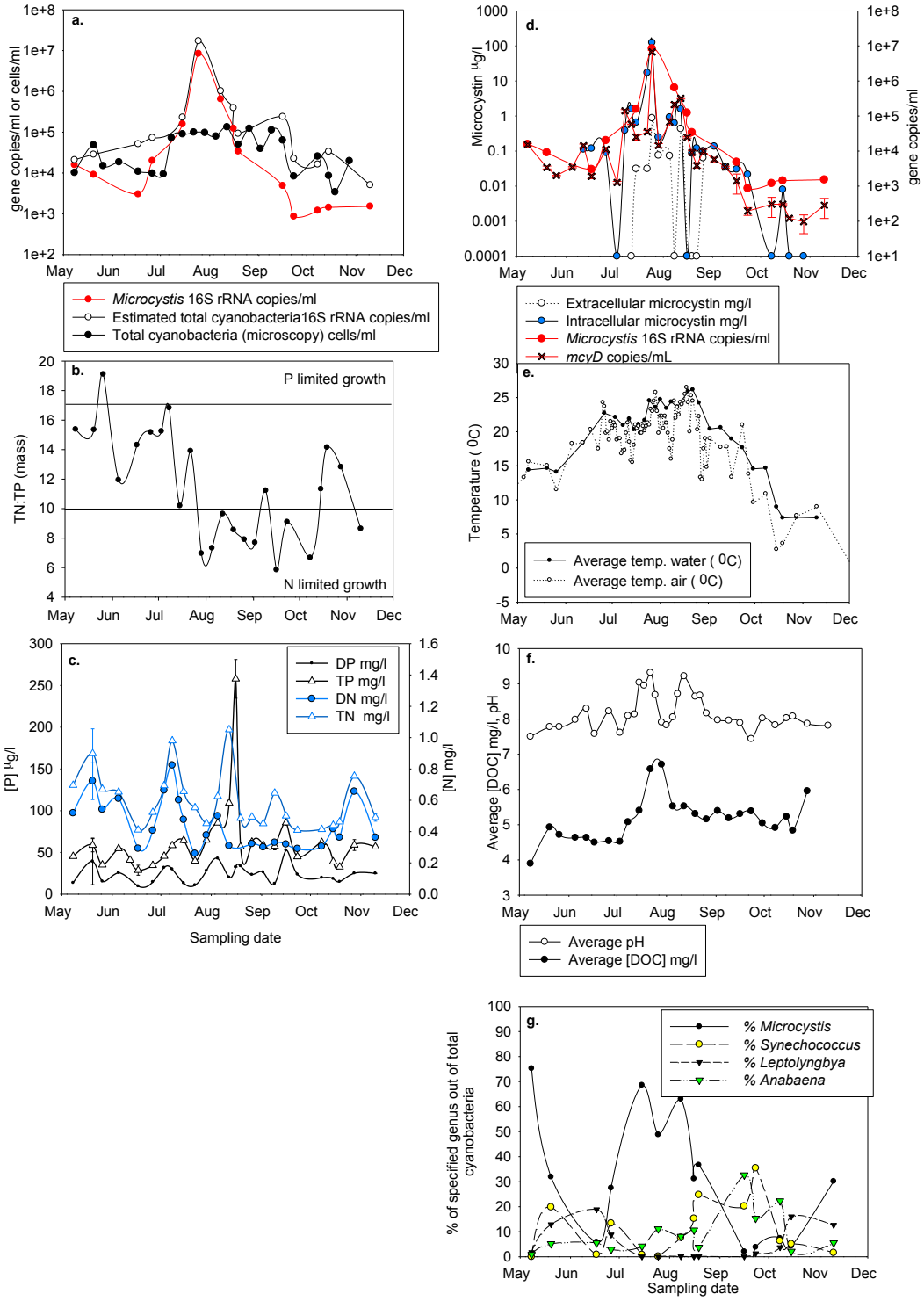

Figure 4.17 – Temporal profile of the dynamics of environmental and biomolecular parameters in pelagic station Pel2

a) Total cyanobacteria abundance (microscopy), *Microcystis* and total cyanobacteria 16S rRNA gene abundance; b) TN:TP ratio; c) total and dissolved N and P concentrations; d) microcystin concentration, *mcyD* gene and *Microcystis* 16S rRNA gene abundance; e) temperature; f) pH and dissolved organic carbon (DOC) concentrations; g) relative abundance of main cyanobacteria genera at BM5. Figures e and f reflect MB averages as opposed to Pel2 specific data.

Table 4.7 – Averages of environmental parameters during the 2009 sampling period in littoral station BM5 and pelagic station Pel2.

Parameter	BM5 (Average ± sd)	Pel2 (Average ± sd)	Wilcoxon-Mann-Whitney test*
[TP] µg/l	69.31 ± 46.92	63.00 ± 43.63	P>0.05
[DP] µg/l	31.72 ± 24.20	24.08 ± 10.69	P>0.05
[TN] mg/l	**0.73 ± 0.25**	**0.60 ± 0.18**	**P<0.05**
[DN] mg/l	0.51 ± 0.20	0.45 ± 0.16	P>0.05
[Chl a] µg/l	17.25 ± 19.13	13.20 ± 12.61	P>0.05
[DOC] mg/l	5.14 ± 0.52	5.14 ± 0.50	P>0.05
[pH]	8.18 ± 0.57	8.23 ± 0.58	P>0.05
Ta °C	17.54 ± 6.65	16.64 ± 5.94	P>0.05
Tw °C	19.11 ± 6.10	18.71 ± 6.25	P>0.05

*One-sided p value.
sd: standard deviation

As opposed to what was observed in BM5, the *Oscillatoriales* was not the dominant taxon in Pel2 during May: this may be attributed to the higher nutrient concentrations (Fig 4.17 c) resulting in a shift towards the *Chroococcales*. On June 17, the TP decreased to 28.7 µg/l and the *Oscillatoriales* reached their highest dominance peak. As the temperature, pH and nutrient concentrations increased in the summer months, the cyanobacterial community was dominated by the *Chroococcales* and *Nostocales* as observed in BM5. However, in Pel2, the relative abundance of *Microcystis* alone was not significantly higher than the relative abundance of *Nostocales* from July 15 to August 19 (inclusive) (Wilcoxon-Mann-Whitney p>0.05). This scenario may be due to the different nutrient dynamics between stations resulting in relatively lower TN:TP mass ratios during this period in Pel2 (Fig 4.18).

The *Nostocales* reached its dominance peak a little earlier at Pel2 than at BM5 during September, but as observed in BM5, it did so when the *Microcystis* bloom had subsided. At that time, the water temperature was 18.95°C, DOC=5.30 mg/l, DN= 0.32 mg/l, TP=85.56 µg/l, TN=0.50 and the TN:TP mass= 5.8:1. As in BM5, when the temperature started to decrease in October, the relative abundance of *Oscillatoriales* started to increase, and this taxon dominated towards the end of the sampling campaign in Pel2. Finally. the genus *Microcystis* showed an increase in relative abundance consistent with an increase in nutrient concentrations in November. Nonetheless the *Microcystis* 16S rRNA gene copy abundance was lower than 1,600 copies/ml by the end of the sampling season (Fig 4.17 c).

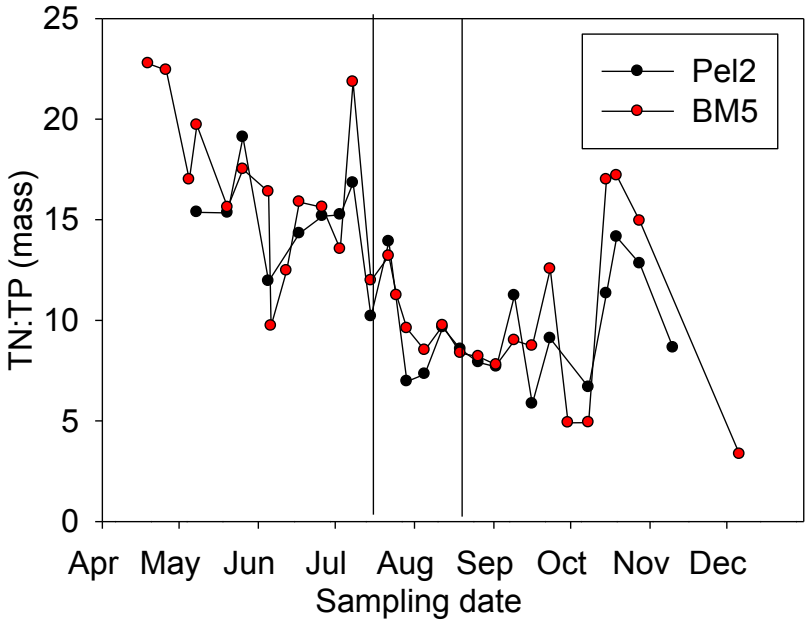

Figure 4.18 – TN:TP mass ratio over time in the littoral station BM5 and the pelagic station Pel2 of Missisquoi Bay in 2009. The vertical lines indicate the period from July 15 to August 19.

In both stations, the dynamics of the *Microcystis* 16S rRNA copies/ml closely followed the pattern of estimated total cyanobacteria 16S rRNA copies/ml, *mcyD* gene and intracellular microcystin, especially during the summer when the *Microcystis* bloom took place (Fig 4.16 a, d; Fig 4.17 a, d). In general, the temporal profiles at both stations, showed that *Microcystis* and *Nostocales* were the predominant taxa when microcystin was detected, while the relative abundances of other orders were below 3%. Correlation analysis suggested that *Microcystis* was the main microcystin producer (Table 4.6).

4.3.3 Relationship between *Microcystis* relative abundance and nutrient concentrations and ratios

On July 25th, the abundance of *Microcystis* in the bay was high, but it was not the dominant taxon in all of the sampling stations (Fig 4.19). While environmental factors, such as pH, temperature and DOC, were relatively constant among stations, nutrient concentrations showed high variability. In this section, the relation between the relative abundance (dominance) of *Microcystis* and nutrient concentrations and ratios is investigated once temperature is not limiting growth any longer. For this analysis, it was

assumed that at water temperatures ≥20°C, temperature was not limiting growth of *Microcystis* since *Microcystis* 16S rRNA gene concentrations of 40,000 copies/ml and even higher than 200,000 copies/ml were observed at water temperatures ≥20°C (table 4.10).

On July 25th, the average water temperature of the bay was 24.5°C; the average pH was 8.68; the DOC was not available on this date and Chl a concentrations ranged from 2.5 to 5.5 µg/l. All the stations had eutrophic P concentrations; the nutrient concentrations are shown in table 4.8.

Table 4.8 – Nutrient concentrations in the MB on July 25/09.

	DP (ug/l)	**sd**	**TP** (ug/l)	**sd**	**DN** (mg/l)	**sd**	**TN** (mg/l)	**sd**
BM1	22.72	0.40	39.23	1.80	0.44	0.01	0.52	0.01
BM2	35.35	1.10	162.89	22.56	0.57	0.29	0.89	0.06
BM3	24.27	0.60	88.36	7.79	0.39	0.01	0.70	0.03
BM4	38.10	0.40	107.13	5.19	0.59	0.00	1.15	0.14
BM5	29.56	1.50	64.01	0.50	0.49	0.00	0.72	0.04
PELVEN	5.43	NA	NA	NA	0.32	0.05	0.48	0.03
PEL2	0.90	NA	NA	NA	0.78	NA	NA	NA
PRM	28.86	0.30	68.74	2.99	0.87	0.03	0.99	0.01

sd= standard deviation; NA= data not available.

Based on the correlation analyses and nutrient concentrations alone, *Microcystis* should dominate at high TN and TP concentrations. However, the pattern observed across the stations does not reflect this (Fig 4.19), suggesting that there may be another factor controlling the relative abundance of *Microcystis*.

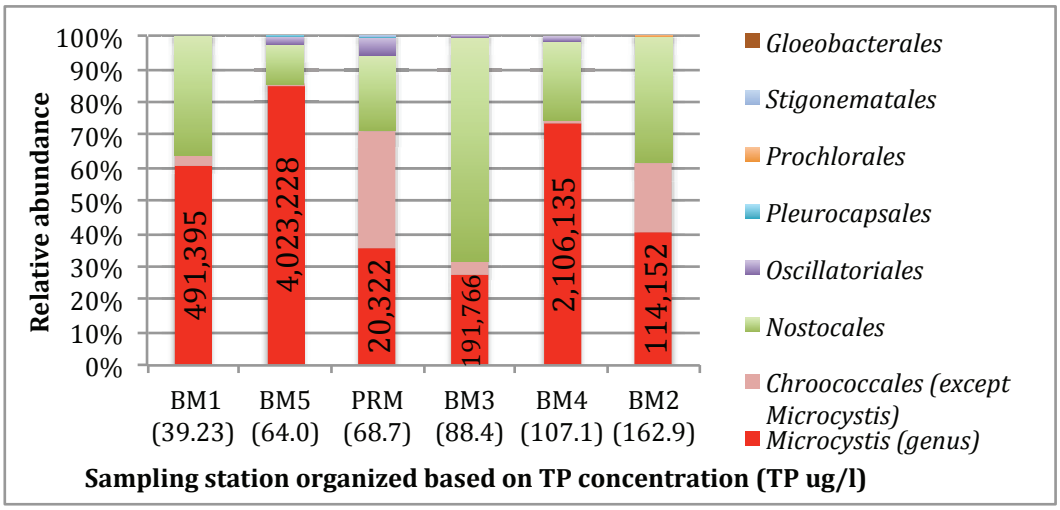

Figure 4.19 – Spatial variation of *Microcystis* in relation to cyanobacterial orders in MB on July 25, 2009 based on increasing TP concentrations.
The X-axis shows the sampling station and the concentration of TP in parenthesis. The percentages on the Y-axis reflect the relative abundance of the orders/genus in the sample. The numbers within the bars represent the abundance of *Microcystis* 16S rRNA gene copies (copies/ml). The stations Pel2 and PelVen were not included due to the fact that they did not have complete nutrient data. Note that the average Tw on this date was 24.5°C and that all stations had nutrient concentrations ≥39 µg/l of TP and ≥0.52 mg/l of TN (Table 4.8).

The pattern of *Microcystis* dominance across stations on July 25th seem to be better explained by a non-linear, non-monotonic relation with the nutrient ratios. The relative abundance of *Microcystis* seemed to increase as the TN:TP (mass) ratio approached 11:1 (Fig 4.20). The abundance of *Microcystis* also seemed to increase as the TN:TP (mass) ratio approached 11:1 (Fig 4.20, numbers within the bars). In order to verify whether this relationship between *Microcystis* relative abundance and the nutrient ratios held in the rest of the dataset, all the samples that had sequencing data, complete nutrient and that were collected at water temperatures higher than 20°C, were analyzed in detail (Fig 4.21). These samples had nutrient concentration ranges of 34.48-162.89 µg/l for TP and 0.48-1.15 mg/l for TN. Interestingly it was observed that regardless of the sampling date or station, the relative abundance of *Microcystis* still increased as the TN:TP (mass) ratio approached 11:1. However, the abundance of *Microcystis* 16S rRNA gene copies did not always followed such pattern, probably reflecting the effect of absolute nutrient concentrations on biomass production (Fig 4.21).

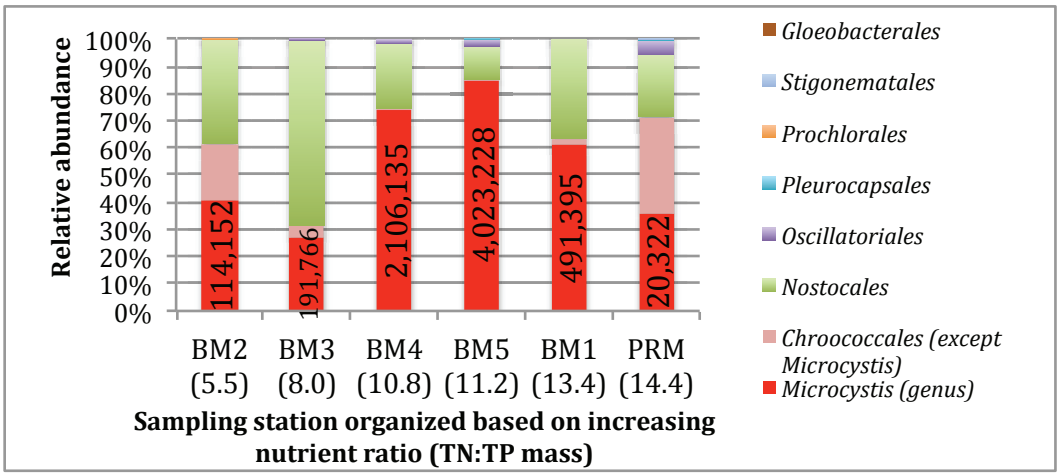

Figure 4.20 – Spatial variation of *Microcystis* in relation to cyanobacterial orders in MB on July 25, 2009 based on increasing nutrient ratios.

The X-axis shows the sampling station and the TN:TP (mass) in parenthesis. The percentages on the Y-axis reflect the relative abundance of the orders/genus in the sample. The numbers within the bars represent the abundance of *Microcystis* 16S rRNA gene copies (copies/ml). The stations Pel2 and PelVen were not included due the fact that they did no have complete nutrient data. Note that the average Tw on this date was 24.5°C and that all stations had nutrient concentrations ≥39 µg/l of TP and ≥0.52 mg/l of TN (Table 4.8).

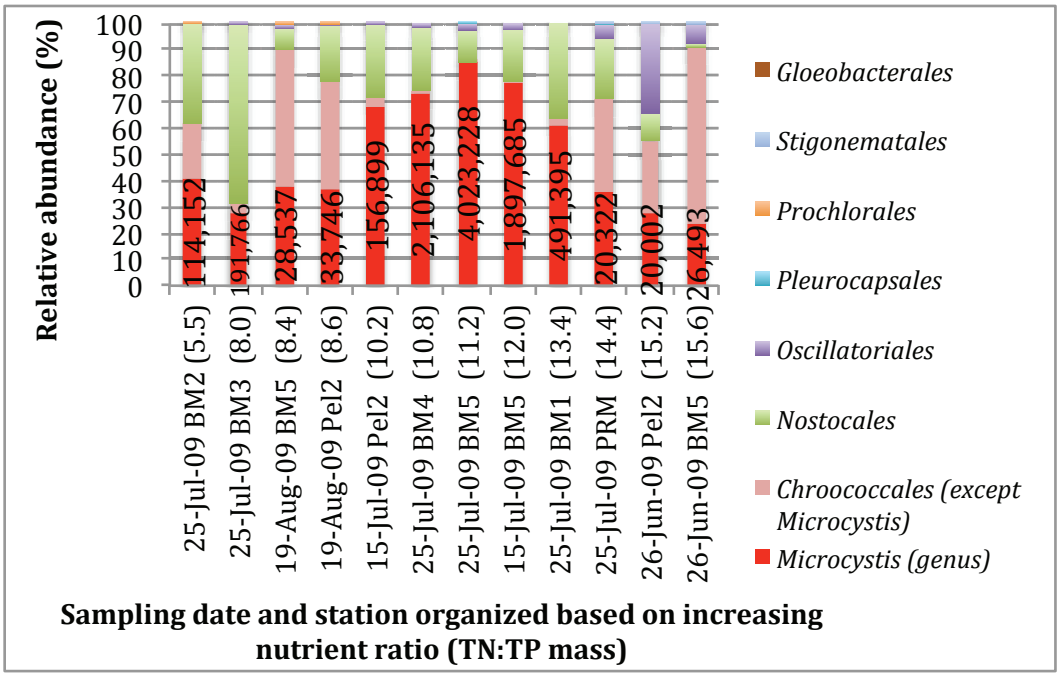

Figure 4.21 – Relationship between *Microcystis* and nutrient ratios in MB stations at water temperatures equal to or higher than 20°C.

The X-axis shows the sampling station and the TN:TP (mass) in parenthesis. The percentages on the Y-axis reflect the relative abundance of the orders/genus in the sample. The numbers within the bars represent the abundance of *Microcystis* 16S rRNA gene copies (copies/ml). Note that these samples had nutrient concentration ranges of 34.48-162.89 µg/l for TP and 0.48-1.15 mg/l for TN.

4.4 Spatio-temporal dynamics of cyanobacterial toxicity in Missisquoi Bay during 2009

4.4.1 Significant relationships between microcystin concentrations, total cyanobacterial cells and biomolecular markers for total and toxic cyanobacteria and *Microcystis* abundance.

Spearman rank correlations (rs) were performed between the abundances of the *Microcystis* 16S rRNA gene, estimated total cyanobacteria 16S rRNA gene, total cyanobacterial cells, *mcyD* gene and microcystin. For this purpose all of the data from the MB samples were included in the analysis. Data from the tributary stations were excluded from this analysis because no *mcyD* gene was detected in the PR station and EC had high levels of PCR inhibitors hampering the molecular biological analysis (Table 4.9).

Table 4.9 – Significant relationships between *Microcystis* 16S rRNA gene, estimated total cyanobacteria 16S rRNA gene, *mcyD* gene, and microcystin concentrations.

	Intracellular microcystin	**Extracellular microcystin**	***mcyD* gene**	***Microcystis* 16S rRNA gene**	**Estimated total cyanobacteria 16S rRNA gene**
Intracellular microcystin					
Extracellular microcystin	0.29/307/ p<0.01				
***mcyD* gene**	0.86/100/ p<0.01	0.29/59/ p<0.05			
***Microcystis* 16S rRNA gene**	0.89/33/ p<0.01	0.54/20/ p<0.05	0.96/42/ p<0.01		
Estimated total cyanobacteria 16S rRNA gene	0.82/33/ p<0.01	0.47/20/ p<0.05	0.88/42/ p<0.01	0.91/43/ p<0.01	
Total cyanobacteria cell abundance (microscopy)	0.59/326/ p<0.01	0.28/288/ p<0.01	0.47/101/ p<0.01	0.55/31/ p<0.01	0.52/31/ p<0.01

The data is displayed in the following order: Spearman rho estimates (rs)/number of data points in the correlation tested (n)/ p value.

WHO water quality guidelines and microcystin concentrations in Missisquoi Bay in 2009

WHO environmental guidelines of low and moderate probability of adverse health effects due to cyanobacterial cells are 20,000 and 100,000 cells/ml, respectively, which in a bloom consisting mainly of *Microcystis* cells would represent approximately 4 µg/l and

20 µg/l of microcystin, respectively (WHO 2003). The underlying assumption behind the WHO freshwater quality guidelines for moderate probability of adverse health effects is that *Microcystis* has an average toxin content per cell of 200 fg, varying from 100 to 500 fg, depending on the degree of toxicity of the bloom and the regional differences in microcystin content of the cells (Falconer et al. 1999). The average and median toxin content per toxic cell (based on the *mcyD* gene) observed in the bay were below the average toxin content per cell that was used to develop WHO guidelines. Nevertheless, the range of toxin content per *mcyD* gene was very wide and exceeded 500 fg per *mcyD* gene copy. Based on our estimate of intracellular microcystin per *mcyD* gene copy, from 75 samples of MB that had intracellular microcystin concentrations higher than the detection limit (0.02 ug/l) and were screened for the *mcyD* gene, the observed microcystin content per *mcyD* copy (estimated toxic cell) in the bay ranged from 1.33 to 1,455 fg/copy. The average microcystin content per *mcyD* copy (estimated toxic cell) was 61.28 fg/copy, while the median value was 18.40 fg of microcystin per *mcyD* gene copy (estimated toxic cell).

In our data, the total cyanobacterial cell abundance surpassed the freshwater quality guidelines on several occasions in which the toxin levels were below the guidelines (Fig 4.22 b). On the other hand, the concentrations of microcystin in the bay, related to the *mcyD* gene and estimated *Microcystis* cells complied fairly well with WHO environmental guidelines (Fig 4.22 a and b). This is assuming that cyanobacterial cells carry one *mcyD* gene copy per cell (Kaebernick et al. 2002), and that *Microcystis* has two rRNA operons per genome (Rinta-Kanto et al. 2009; Engene and Gerwick 2011).
In terms of *mcyD* copies in the bay in relation to the WHO recreational water guidelines, only one sample out of 100 had a microcystin concentration higher than 20 µg/l (20.7 µg/l) and less than 100,000 *mcyD* gene copies/ml (46,782 ± 1,613 copies/ml) (Fig 4.22 a). From 33 samples screened for the *Microcystis* 16S rRNA gene, all of the samples with microcystin concentrations higher than 20 µg/l had estimated *Microcystis* cells/ml higher than 100,000 cells/ml (Fig 4.22 c).

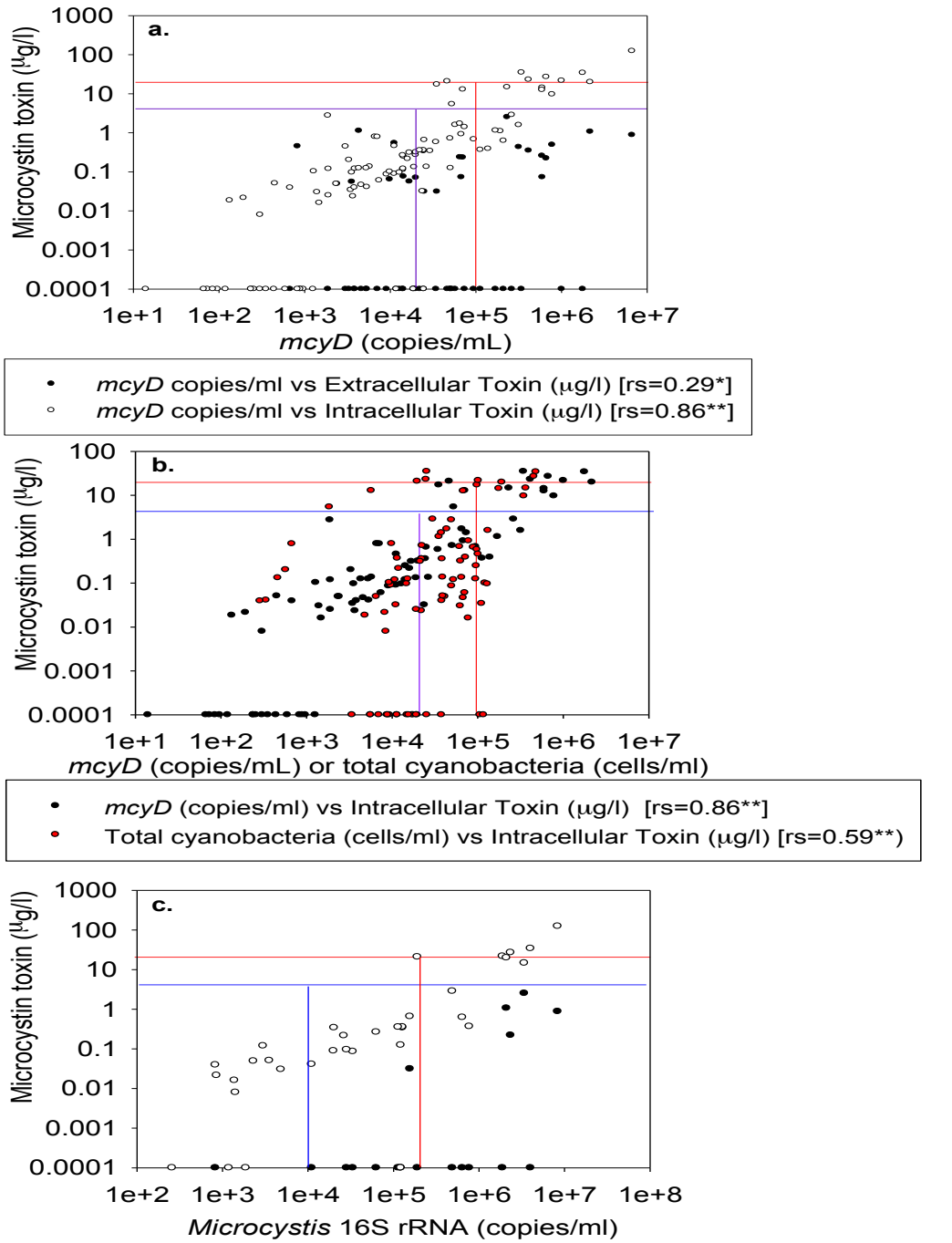

Figure 4.22 – Relationship between microcystin and the *mcyD* gene, total cyanobacterial cells and *Microcystis* 16S rRNA gene copies.

The lines represent the WHO freshwater quality guidelines of relatively low probability of adverse health effects (blue) and the moderate probability of adverse health effects (red) in terms of cyanobacterial cells (vertical). The horizontal lines represent the expected microcystin concentrations at such cell densities (WHO 2003). A value of 0.0001 µg/l for intracellular microcystin indicates an observation below the detection limit, the detection limit for extracellular microcystin was 0.02 µg/l. (rs) Spearman's rho; (*) p value<0.05; (**) p value<0. 01.

4.4.2 Significant relationships between microcystin concentrations, *Microcystis* 16S rRNA gene, estimated total cyanobacteria 16S rRNA gene, total cyanobacterial cells, *mcyD* gene and environmental parameters

Spearman rank correlations (rs) were performed between the abundances of *Microcystis* 16S rRNA gene, estimated total cyanobacteria 16S rRNA gene, total cyanobacterial cells, *mcyD* gene, microcystin concentrations and the different environmental parameters analyzed (T air, T water, pH, DOC, Chl a, nutrient concentrations and TN:TP mass ratio) in order to investigate which parameters may have been significantly related to the incidence of *Microcystis* and toxic cyanobacterial blooms observed in Missisquoi Bay during 2009 (Table 4.10 and 4.11). Data from the tributary stations were excluded for the reason previously mentioned.

In order to create a frame of reference, the parameters that were significantly associated were compared, independently, to a framework developed based on the WHO provisional drinking water quality guideline for microcystin and the freshwater quality guidelines for relatively low and moderate probability of adverse health effects due to cyanobacterial cells. In this study, the WHO provisional drinking-water quality guideline of 1 ug/l was used instead of Quebec's guideline of 1.5 ug/l because its criterion is more stringent. The framework in terms of microcystin, was established at 1, 4 and 20 µg/l, which would represent approximately 5,000, 20,000 and 100,000 cells/ml of microcystin producing genera, or 5,000, 20,000 and 100,000 copies/ml of the *mcyD* gene assuming that the average toxin content per cell is 200 fg (Falconer et al. 1999) and that cyanobacteria only carry 1 copy per cell (Kaebernick et al. 2002). For the *Microcystis* 16S rRNA gene this framework would represent approximately 10,000, 40,000 and 200,000 copies/ml assuming that *Microcystis* carries 2 operons of this gene per cell (Rinta-Kanto et al. 2009; Engene and Gerwick 2011).

Table 4.10 – Significant relationships of microcystin, *mcyD* gene, *Microcystis* 16S rRNA gene concentrations with environmental parameters and minimum values of these parameters at which the toxin or biomolecular markers were observed to reach the specified concentrations

Intracellular microcystin (µg/l)	**Spearman Rank correlation** (rho/n/p)	**Minimum values of the parameters at which these or higher microcystin concentrations were observed**		
		1 µg/l	**4 µg/l**	**20 µg/l**
T water (°C)	0.47/122/p<0.01	20.32	20.32	20.32
pH	0.43/122/p<0.01	7.61	7.83	7.83
DOC (mg/l)	0.49/94/ p<0.01	4.52	5.07	5.4
Chl a (µg/l)	0.35/69/ p<0.01	2.2	2.2	3.4
TP (µg/l)	0.37/77/ p<0.01	39.23	39.76	64.00
DN (mg/l)	0.27/78/ p<0.05	0.26	0.26	0.38
TN (mg/l)	0.56/70/ p<0.01	0.52	0.55	0.70

***mcyD* gene (copies/ml)**	**Spearman Rank correlation** (rho/n/p)	**Minimum values of the parameters at which these or higher gene concentrations were observed**		
		5,000 copies/ml	**20,000 copies/ml**	**100,000 copies/ml**
T water (°C)	0.64/113/p<0.01	14.4	18.3	20.3
pH	0.63/121/p<0.01	7.5	7.8	7.8
DOC (mg/l)	0.30/76/p<0.01	3.89	4.63	5.07
Chl a (µg/l)	0.33/89/p<0.01	1.6	2.2	3.4
TP (µg/l)	0.38/98/p<0.01	34.48	34.53	39.23
TN (mg/l)	0.52/91/p<0.01	0.45	0.49	0.52

***Microcystis* 16S rRNA gene (copies/ml)**	**Spearman Rank correlation** (rho/n/p)	**Minimum values of the parameters at which these or higher gene concentrations were observed**		
		10,000 copies/ml	**40,000 copies/ml**	**200,000 copies/ml**
T water (°C)	0.69/36/ p<0.01	14.4	20.3	20.3
pH	0.68/38/ p<0.01	7.5	8.68	8.68
TP (µg/l)	0.39/37/ p<0.05	34.48	36.08	39.23
TN (mg/l)	0.36/32/ p<0.05	0.48	0.48	0.52

Extracellular microcystin

No significant associations were found with any of the environmental parameters tested.

Table 4.11 – Significant relationships of total cyanobacterial cells (microscopy) and estimated total cyanobacterial 16S rRNA gene concentrations with environmental parameters and minimum values of these parameters at which the indicators of cyanobacterial abundance were observed to reach the specified concentrations

Total cyanobacteria (microscopy) (cell/ml)	Spearman Rank correlation (rho/n/p)	Minimum values of the parameters at which these or higher cell concentrations were observed		
		5,000 cells/ml	20,000 cells/ml	100,000 cells/ml
T water (°C)	0.21/121/ p<0.05	7.36	7.36	14.56
pH	0.35/127/ p<0.01	7.44	7.44	7.83
DOC (mg/l)	0.43/102/ p<0.01	3.89	4.52	5.40
Chl a (µg/l)	0.40/68/ p<0.01	Not distinct*	Not distinct*	Not distinct*
DP (µg/l)	0.22/84/ p<0.05	9.56	10.77	12.11
TP (µg/l)	0.67/82/ p<0.01	28.49	34.53	44.51
TN (mg/l)	0.34/77/ p<0.01	0.31	0.36	0.48
TN:TP (mass)	-0.41/74/p<0.01	Not distinct*	Not distinct*	Not distinct*

Total cyanobacteria 16S rRNA gene (copies/ml)	Spearman Rank correlation (rho/n/p)	Minimum values of the parameters at which these or higher gene concentrations were observed
		200,000 copies/ml
T water (°C)	0.65/36/ p<0.01	17.68
pH	0.66/38/ p<0.01	Not distinct*
DOC (mg/l)	0.49/20/ p<0.05	5.30

*Not distinct: the specified gene or cell concentrations were observed at environmental parameter concentrations that encompassed the whole range of data.

4.4.3 *Microcystis* 16S rRNA gene, estimated total cyanobacteria 16S rRNA gene, *mcyD* gene, microcystin concentrations and their relationship with nutrient ratios

In section 4.3 it was noted that the relative abundance of total *Microcystis* would increase as the nutrient ratio approached 11:1 at water temperatures equal to or higher than 20°C, and that in general, the abundance of *Microcystis* 16S rRNA gene copies also followed this trend. In addition, correlation analysis suggested that this genus was the main *mcyD*-carrier and microcystin producer. We therefore expected to see that the dynamics of microcystin concentrations and the abundance of the *mcyD* gene would follow a similar pattern. All the data points available from MB in terms of the *mcyD* gene, *Microcystis* 16S rRNA gene and microcystin concentrations from samples collected at water temperatures equal to or higher than 20°C were plotted against the TN:TP (mass) ratio (Fig 4.23). This analysis showed that *mcyD* gene copy numbers higher than 100,000

copies/ml and total microcystin concentrations higher than 1 µg/l are found in a TN:TP mass ratio range from 5:1 to 15:1, and that they increase as the TN:TP mass ratio approaches 11:1, closely following the dynamics of dominance and abundance of total *Microcystis*. However, low microcystin and *mcyD* concentrations were also recorded at a nutrient ratio of 11:1 (Fig 4.23).

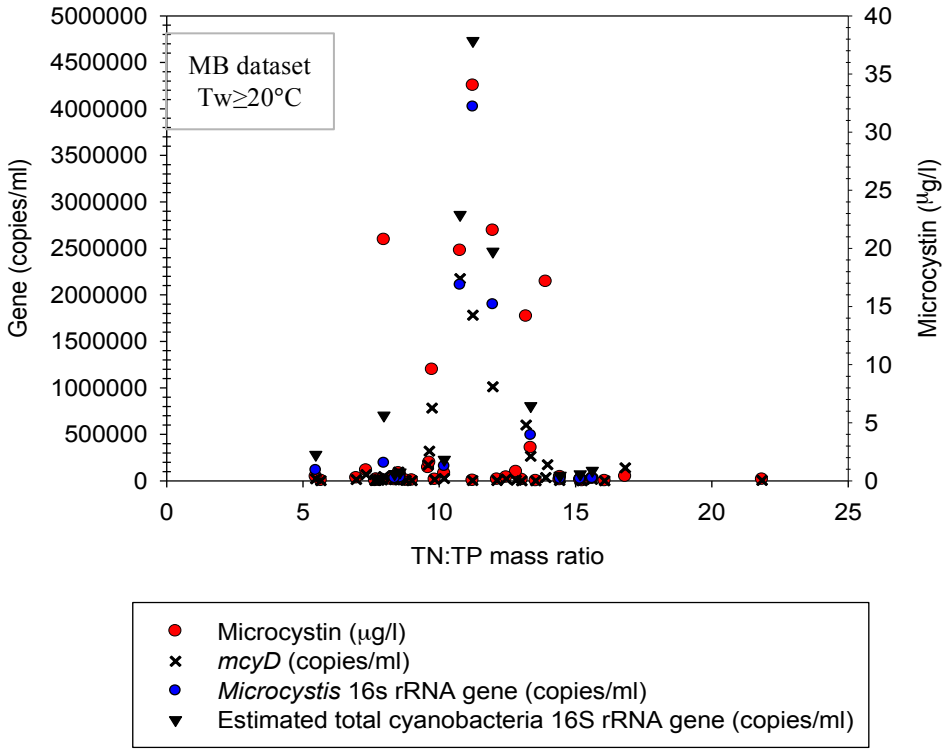

Figure 4.23 – Estimated total cyanobacteria and *Microcystis* 16S rRNA genes, *mcyD* gene and intracellular microcystin in relation to the TN:TP mass ratio during 2009 in Missisquoi Bay (all samples collected at Tw≥20°C).

To further study the relationship observed between microcystin and *mcyD* concentrations with the nutrient ratio, the dynamics of these parameters were analyzed at one time point. On July 25th, the average temperatures of the bay were Ta= 21°C and Tw=24.5°C; the average pH was 8.68 and Chl a concentrations ranged from 2.5 to 5.5 µg/l. The DOC values were not available on this date. All the stations had eutrophic P concentrations; the nutrient status is shown in table 4.8. The highest concentration of the *mcyD* gene and microcystin concentrations found in the samples were 6.71 million copies/ml and 124.87 µg/l, respectively; both were recorded in Pel2 on July 25th/09.

Unfortunately this station did not have a complete nutrient dataset; therefore it was excluded from this analysis.

Again, there was no linear relationship between the nutrient concentrations in the stations and the microcystin concentrations or the abundances of the genes analyzed that could explain the dynamics observed in the bay on this day (Table 4.8). Nonetheless, despite the fact that none of these parameters had a monotonic relationship with the TN:TP mass ratio (Spearman rank correlation $p>0.05$ for all of the parameters), it was observed that the dynamics observed in the bay on July 25th were slightly better explained by the relationship with the TN:TP ratio, although not linear, nor monotonic (Fig 4.24). The concentrations of the biomolecular parameters in question seemed to increase as the TN:TP mass ratio approached 11:1, which is congruent with the TN:TP mass ratio at which the relative abundance of *Microcystis* peaked.

Nevertheless, this apparent trend with the TN:TP mass ratio still does not account for some of the dynamics observed on July 25th. For example, the intracellular microcystin per *mcyD* gene copy (fg/copy), ranged from 9 to 19 fg per *mcyD* copy at all stations, except at BM3, which had 443 fg of intracellular toxin per *mcyD* copy. For this reason, BM3 had a microcystin concentration comparable to the one found in BM4, despite of the fact that BM3 had roughly 46 times fewer *mcyD* copies than BM4 (Fig 4.24). When looking at the relative abundance of *Microcystis* in the samples, related to the relative abundance of cyanobacterial orders, it was observed that there is a strong presence of other cyanobacterial orders (e.g., *Nostocales*, *Oscillatoriales* and other *Chroococcales*) that comprise potential microcystin-producing genera, especially in BM4 (Fig 4.20).

It was also interesting to note that in some of the stations such as PRM, pel2 and BM4 there were more *mcyD* gene copies than estimated *Microcystis* cells (Fig 4.25), corroborating that *Microcystis* was not the only *mcyD* gene carrier in the bay.

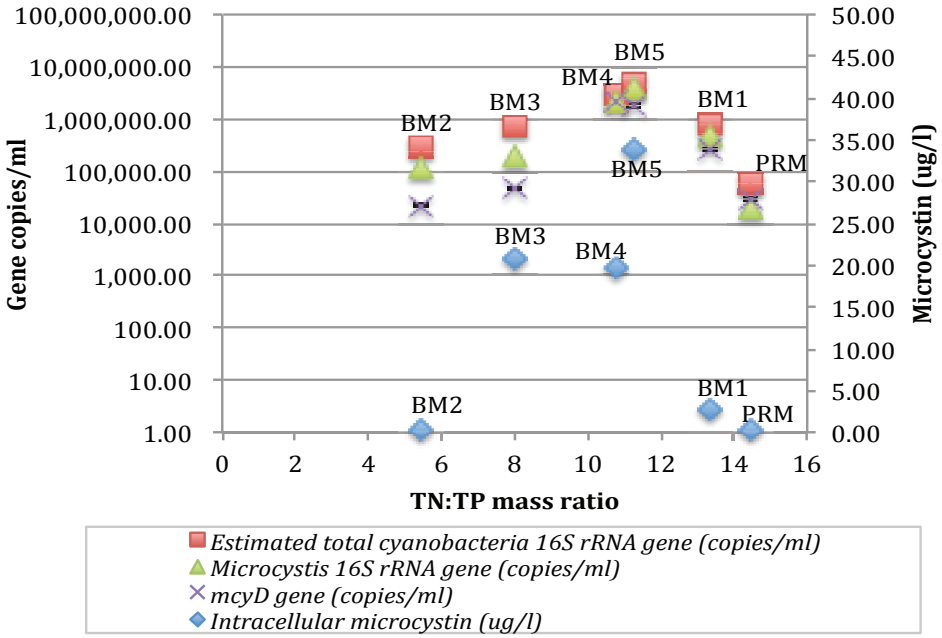

Figure 4.24 – Estimated total cyanobacteria and *Microcystis* 16S rRNA genes, *mcyD* gene and intracellular microcystin in relation to the TN:TP mass ratio on July 25th, 2009 in Missisquoi Bay. The error bars correspond to the standard deviation (n=3) for the molecular markers (genes).

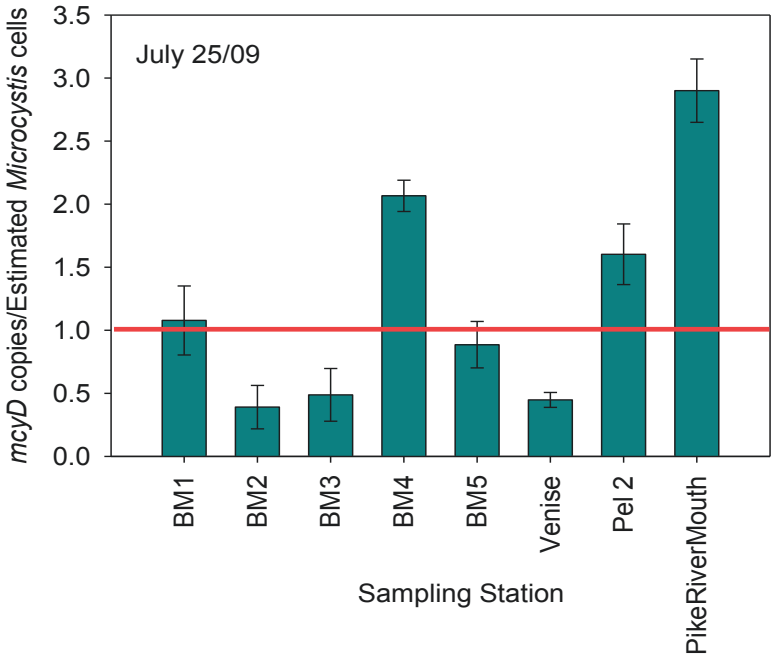

Figure 4.25 – Ratio of *mcyD* copies and estimated *Microcystis* cells on July 25th, 2009 in Missisquoi Bay. It was assumed that *Microcystis* carries two 16S rRNA operons per cell. The error bars represent the relative error of the ratio between *mcyD* (n=3) and *Microcystis* 16S rRNA gene (n=3) qPCR data.

Chapter 5 – Discussion and Conclusions

5.1 Nutrient input and its relation with indicators of animal excreta contamination and cyanobacterial blooms in Missisquoi Bay/ Pike River during 2009

While an overall decreasing trend of nutrient input has been observed in Lake Champlain, the concentration of P has increased by 72% in Missisquoi Bay during the past 5 decades, accompanied by a doubling of Chl a concentrations and an increasing dominance of cyanobacteria (Levine et al. 2012; Smeltzer et al. 2012).

In this highly agricultural watershed, where the land devoted to agriculture encompasses approximately 33% of the area (Beck et al. 2012), unincorporated manure during crop fertilization can account for a considerable amount of P input in surface runoff (Michaud 2004). External nutrient sources (e.g. from surface runoff, defective septic systems and drain fields or sewage overflows from plants located upstream of PR) seemed to substantially contribute to the nutrient concentrations in MB and Pike River during 2009, especially of P. Phosphorus concentrations (both total and dissolved) increased in the stations as contact with the shore increased, and a significant correlation with rainfall was detected with P concentrations in the stations closest to the shore, i.e. tributary stations. Nonetheless, despite that the average phosphorus concentration was higher in the littoral than in the pelagic station, the difference was not statistically significant. This probably reflects the effect of the internal P influx and resuspension from the sediments that can take place during hypoxic periods (Testa and Kemp 2012) and thus may contribute to P concentrations in the water column of shallow well mixed systems (Gobler et al. 2007).

The concentrations of N also increased in the stations as contact with the shore increased. However, contrary to what was observed with P, no significant correlations were detected between N concentrations and rainfall at the stations closest to the shore. This suggests that the higher nitrogen concentrations observed at the stations closest to the shore may also be attributed to N that reaches the lake some time after rainfall events have occurred (e.g. due to leaching and transportation to tile drainage) or to internal sources of N. While phosphates in general are strongly adsorbed to soil particles, nitrates are highly soluble, hence higher leaching of N than P has been observed in tile drain effluents (Johnston et al. 1965). In terms of potential internal sources of N, it was

interesting to note that all the samples analyzed for mtDNA from the tributary station PR (the station closest to the shore) had a pronounced presence of rodents, particularly muskrats, throughout most of the year. Muskrat lifestyle is highly dependent on water (Errington 1951; Ruys et al. 2011), and their diet changes seasonally following the availability of forage species during the year. It has been hypothesized that muskrats meet their nutrient requirements during the summer by consuming high-protein, low-fiber aquatic plants (Campbell and MacArthur 1996). The habitat and dietary requirements of rodents, such as muskrats, in our area of study may represent a constant and direct source of nutrient input into the water (i.e. not carried in via surface runoff) from animal excreta. If muskrat's excreta is rich in nitrogen, especially urine, which is richer in nitrogen than in phosphorus in small mammals (Clark et al. 2005), this observation may also partly contribute to the lack of correlation between nitrogen and rain in our tributary stations. Although this remains highly speculative, future studies of animal excreta effects on nutrient input in the bay could be performed using qPCR and mtDNA gene targets combined with sequencing and nutrient analyses.

Runoff from fields where manure has been applied, in addition to nutrients, sediment, and organic solids, can transport pathogens, such as *E. coli*, to surface waters. In addition to feces (Padia et al. 2012), soil has been identified as an important source of *E. coli* into aquatic environments (Hardina and Fujioka 1991). It has been demonstrated that *E. coli* can survive, multiply (Byappanahalli 2004; Ishii et al. 2006; Padia et al. 2012) and naturalize (Ishii et al. 2006) in soils. Other sources of *E. coli* in the water include defective septic systems and drain fields as well as sewage overflows. In 2009 alone, there were 155 overflow events: 137 associated with rain, 13 associated with snow melt and 5 associated with damages and/or repair work from the Bedford wastewater treatment plant (MAMROT 2010). In the MB/Pike River area, *E. coli* counts demonstrated a pattern of surface runoff likely enhanced by sewage overflow events after intense rainfall during 2009. The *E. coli* counts in the studied stations significantly increased as the station contact with the shore, and a correlation between *E. coli* counts and rainfall was observed in the tributary stations. This, in combination with the significant correlations detected between *E. coli* counts and the concentrations of nutrients, suggest that *E. coli* counts from surface water samples is a good indicator of recent nutrient input into the water

system from animal excreta and/or surface runoff. A constraint to employing *E. coli* abundance in surface water samples as an indicator of recent nutrient input comes from the confounding effect derived from water circulation and/or water mixing events that may lead to the resuspension of *E. coli* from the sediments (Ishii et al. 2007; Kiefer et al. 2012) and/or biofilms (Shikuma and Hadfield 2009). Nonetheless, the significant association observed between the presence of *E. coli* and the detection of mtDNA from non-aquatic hosts provides evidence that *E. coli* from external sources did find their way into the bay.

Different point and non-point sources can contribute to the animal excreta contamination of a waterbody, including sewage overflows from wastewater treatment plants, defective septic systems and drain fields, livestock, birds, and wildlife. In agreement with the pattern of land use of the MB watershed (Beck et al. 2012), the mtDNA hosts observed in the bay comprised rodents, birds, cattle and humans. The detection of human mtDNA was prevalent in the bay, and wildlife, particularly rodents, were mainly detected in the tributary station, PR. The detection of farm animals was low in all the samples tested. This is contrary to what was expected in this agricultural watershed, that is used for the growth of cattle and pigs (Simoneau 2007) and heavily relies on crop fertilization practices (manure and chemical fertilizers). It is important to emphasize that while our characterization of non-aquatic mtDNA hosts in the system through sequencing allowed the detection and identification of non-aquatic mtDNA hosts, this identification was not exhaustive for reasons discussed below; hence the lack of detection of hosts does not necessarily mean that they were absent.

The low detection of animals of agricultural interest, combined with the fact that fish genera were only detected in PR (the shallowest sampling station) suggest that mtDNA might be short lived in the environment. Recent work in our laboratory indicates that the half-life of mtDNA in lake water samples can vary from 2 to 10 days depending on the water temperature (L. Masson, personal communication). It is also possible that mtDNA from non-aquatic species was present in the environment but at quantities below the limit of detection of the PCR reaction, or that the method is subject to primer bias and only detected sequences that were highly abundant, or that the animals present in the studied area may not be well represented in the DNA databases. Lack of amplification

due to mtDNA damage could have also been possible since it has been suggested that mtDNA is more susceptible to damage by reactive oxygen species than nuclear DNA inside the cells (Richter et al. 1988; Salazar and Van Houten 1997; Yakes and Van Houten 1997). The caveats of the approach employed may have also hindered the ability to perform a more exhaustive and reliable identification of all the potential target animals that could have been otherwise detected had not primer bias occurred (refer to Appendix C): A degenerate set of primers was used for the second reaction of the nested PCR amplification of the mtDNA in order to produce a shorter-length amplicon suitable for Ion Torrent sequencing; unfortunately, multiple bands of different sizes were produced during the amplification of mtDNA from environmental samples. Since the Ion Torrent platform can handle fragments shorter than 300 bp at the moment, only the expected 205 bp band was sequenced for all of the samples. Thus, it may be possible that the four samples where mtDNA from non-aquatic species was not detected (and all the other samples) had mtDNA from non-aquatic species but with higher molecular weight amplicons. In addition, the high number of PCR cycles associated with nested PCR may have also exacerbated the PCR bias and production of PCR artifacts (Qiu et al. 2001; Lee et al. 2012), especially in samples with low DNA template (mtDNA) concentrations. Since it is not possible to know *a priori* which amount of total DNA will have enough target DNA (mtDNA), a standard of 5 ng of total DNA was employed for the amplification reactions of all samples. In this regard, PCR biases and artifacts due to a low mtDNA template concentration could explain the large variation in sequencing reads recovered per sample once low-quality reads had been removed.

In spite of all the discussed drawbacks, our approach still allowed a preliminary qualitative identification of potential sources of animal excreta from non-aquatic species and their comparison across samples since all samples were treated equally (i.e. all the mtDNA amplicons were produced from a standard initial amount of total DNA (5 ng) for each sample and the amplicons from the same size molecular weight band of all samples were pooled in an equimolar ratio and sent for sequencing). In the future, the characterization of non-aquatic mtDNA hosts in the MB/Pike river samples from 2009 will be validated with a primer set that delivers a single length PCR product suitable for sequencing with the Ion Torrent platform (M. Vuong, personal communication) and the

PCR protocol will be optimized to avoid multiple size amplicons in environmental samples. In this way the PCR biases will be minimized, better quality reads will be produced and a deeper depth of coverage during sequencing will be attained, thus facilitating more reliable and conclusive analyses.

In MB during the 2009 sampling campaign, the concentrations of phosphorus and nitrogen correlated significantly with the abundance of total cyanobacterial cells, *Microcystis* 16S rRNA gene, the *mcyD* gene and intracellular microcystin. We suggest that external sources of nutrients, such as surface runoff and animal excreta, played a significant role in the load of nutrients into MB and thus in the proliferation of toxic cyanobacterial blooms. This is supported by i) the pattern of surface runoff observed in nutrient concentrations and *E. coli* counts, ii) the significant correlations between *E. coli* counts and nutrient concentrations, iii) the detection of non-aquatic mitochondrial DNA hosts in the system, and iv) the significant correlation between the concentrations of *E. coli* counts and total cyanobacterial cells in MB.

5.2 Major cyanobacterial taxa in Missisquoi Bay

5.2.1 High-throughput sequencing of partial 16S rRNA genes

We developed primers that allowed us to analyze the dynamics of the relative abundance of major cyanobacterial taxa and their succession in MB during 2009 using high-throughput sequencing. Since high-throughput sequencing is semi-quantitative, we could establish whether there were significant correlations between the dynamics of the environmental parameters and the relative abundance of major cyanobacterial taxa, and thus determine which parameters may have played a role in the observed succession patterns. Careful attention was placed on primer design, the 16S rRNA gene region amplified and the amplification conditions employed (See Appendix A) to minimize potential PCR biases due to the negative effect they can have on the determination of community structure from 16S rRNA gene amplicons (Hamady and Knight 2009; Lee et al. 2012).

In the future, our approach to the determination of cyanobacterial community structure will greatly benefit from a revised sequence database in which the identification of cyanobacterial strains is based on a polyphasic approach (Hoffmann et al. 2005; Willame et al. 2006; Thomazeau et al. 2010), as well as from further sequencing of

cyanobacterial lineages that are currently absent or poorly represented in the databases. At present, our approach was subject to the caveats associated with using the 16S rRNA gene and the sequence information that is currently available in the databases. To date, cyanobacterial identification has widely depended on either phenotypic or genotypic markers (Willame et al. 2006). Within each of these bases for cyanobacterial identification there are complications that obstruct the development of a consistent cyanobacterial classification. The morphology and morphometry of cyanobacteria can change depending on growth conditions, or can be altered in culture. In addition, some cyanobacteria with the same genetic makeup present different phenotypes depending on physiological conditions, while some strains with similar morphotypes differ in their 16S rRNA gene sequence (Anagnostidis and Komárek 1985; Castenholz and Norris 2005; Willame et al. 2006). As a result, genetic relationships do not always correspond to classifications that were obtained from morphological data (Lyra et al. 2001; Willame et al. 2006; Thomazeau et al. 2010). Consequently, sequence databases currently contain cyanobacterial sequences that have been incorrectly named since their identities, based on the sequence annotation, do not correspond to their closest related organisms on a genetic basis (Willame et al. 2006). Nonetheless, the 16S rRNA gene remains a good marker for establishing relationships between cyanobacteria since it is the molecular marker with the highest representation in the GenBank database (Thomazeau et al. 2010). In addition, its use in phylogenetic studies has been considered efficient for the identification of cyanobacteria to the genus level, while it has sometimes challenged traditional classifications that are based on morphological characters alone (Lyra et al. 2001; Willame et al. 2006; Thomazeau et al. 2010).

In our approach, a taxon was assigned to each sequence using the LCA algorithm from the MEGAN4 software (Huson et al. 2011). This helped to account for incongruities between sequences in the database that have been named based on genetic or morphological characters. The taxon assigned to each sequence depended on whether the sequences from the database that more closely related to our sequence had the same common ancestor, independent of the method by which those database sequences had been identified. From the cyanobacterial sequences, 98.92% were classified to the order level and 89.55% to the genus level by MEGAN4 analysis. As expected, not all of the

sequences that were classified to the order level could successfully be assigned at the genus level, especially those from the orders *Nostocales* and *Oscillatoriales.* This was partly because of inconsistencies in the cyanobacterial classification in the database due to the lack of congruence between morphotypes and genetic makeup. For example the 16S rRNA gene does not support the differentiation between *Anabaena* and *Aphanizomenon*, despite the fact that they are morphologically distinct (Lyra et al. 2001; Willame et al. 2006; Thomazeau et al. 2010). For this reason, we decided that the total number of cyanobacterial sequences should be determined from the total amount of sequences identified to the order level, and that the relative abundances of the main genera would be standardized based on this. This ensured that the base pool of total cyanobacteria sequences did not neglect members of the cyanobacterial population, and that it only varied based on the amount of cyanobacterial sequences, as opposed to the presence of certain unidentifiable cyanobacteria at the genus level (Appendix A). This was essential since our purpose was to analyze total cyanobacterial dynamics. The homogeneous and monophyletic nature of the *Microcystis* genus (Lyra et al. 2001; Komárková et al. 2005; Willame et al. 2006) allowed for differentiation of this genus within its order and between this genus and cyanobacterial genera of the other orders, although not within *Microcystis* species. This enabled us to analyze the dynamics of this globally important genus in relation to other major taxa during a harmful algal bloom. Another advantage of our approach is that it allowed the detection of pico- and nonoplanktonic cyanobacteria, such as *Synechococcus*, that could be missed by microscopic analysis (Wilhelm et al. 2006). This was partly reflected in the imperfect correlation between our estimated total cyanobacterial 16S rRNA gene copies/ml and total cyanobacterial cells/ml (microscopy data) ($sr=0.52$, $n=31$, $p<0.01$).

The appropriate coverage depth for sequencing analysis depends on the purpose of the analysis, the target gene and region sequenced, the diversity of the targets in the sample and the read length (Hamady and Knight 2009). It has been suggested that a coverage depth of 1,000 sequences per sample provides a good trade-off between number of samples and depth of sampling (Hamady and Knight 2009; Kuczynski et al. 2010). But the depth of coverage required may be lower if the goal is to estimate the major bacterial phyla, let alone major cyanobacterial orders. Previous studies have been able to detect

major patterns of variation among real and simulated microbial communities with as few as 100 sequences per sample (Ley et al. 2008; Kuczynski et al. 2010). In our case, rather than a complete characterization of all sequences in each of the samples, our purpose was to perform a quantitative (i.e. determination of relative abundances), taxon-based characterization of the major cyanobacterial taxa within the cyanobacterial community through time and space in order to analyze community composition dynamics in relation to the changing environmental conditions in a bloom affected area such as MB in 2009. Thus, taking into consideration the goal of the project and the resources available, a large number of samples (n=43) and a relatively shallow depth of coverage (approximately 1,000 sequences per sample) seemed more appropriate for our study as opposed to sequencing fewer samples at deeper depth. Our sequencing analysis involved a sampling coverage depth of at least 1,375 good-quality sequences per sample, with an average of 5,000 good-quality sequences per sample. This depth of coverage allowed sampling an average of 1,208 cyanobacterial sequences per sample. However, as expected, the number of cyanobacterial sequences per sample varied depending on the station and the time of the year. In general, all the stations gave a substantial number of sequences; the majority of the samples had at least 100 cyanobacterial sequences while some samples had as many as 6,376 cyanobacterial sequences during the summer. Hence a sampling depth of coverage of at least 1,000 sequences per sample was appropriate, especially during bloom periods. Refer to Appendix A.4 for PCR optimization that may lead to even greater recovery of cyanobacterial sequences during future sequencing analyses employing the primer set UcyaF4/R6.

5.2 2 Factors determining the relative abundance of major cyanobacterial taxa in Missisquoi Bay

The cyanobacterial community of the photic zone in MB was mainly composed of the orders *Chroococcales*, *Nostocales* and *Oscillatoriales* during the sampling season in 2009. While it is widely recognized that no individual environmental factor promotes bloom formation, nutrients and temperature are often identified as key factors in the development of cyanobacterial blooms (McQueen and Lean 1987; Paerl 1988; Elliott et al. 2006; Paerl and Huisman 2008; Davis et al. 2009). From our observations, when certain environmental conditions were outside the ranges in which dominance by each

taxa was generally observed, nutrient concentrations and temperature, seemed to play an important role in determining the pattern of succession of the cyanobacterial community. As expected, the highest estimated cyanobacterial 16S rRNA gene copy numbers and cyanobacterial cell counts were found during the warmest period of the sampling season, especially during the summer, where the temperature reached conditions that are considered optimal for cyanobacterial growth (Robarts and Zohary 1987).

Under favourable conditions, especially relatively high temperature and nutrient concentrations, the genus *Microcystis* dominated. *Microcystis* 16S rRNA gene abundance reached peak concentrations in the summer along the bay, with more than $8x10^6$ copies/ml and $4x10^6$ copies/ml on July 25th in the Pel2 and BM5 stations, respectively. The maximum number of *Microcystis* 16S rRNA gene copies was recorded on Aug 8th, at PRM with more than 14 $x10^6$ copies/ ml. Previous studies on natural systems have found that phosphorus and nitrogen concentrations, (Gobler et al. 2007; Davis et al. 2009; Maske et al. 2010) as well as temperature (Davis et al. 2009; Liu et al. 2011) are factors that have a significant and positive effect on the biomass of *Microcystis*. In addition, positive correlations between *Microcystis* and pH have also been reported (Liu et al. 2011). In line with these studies, we found that the relative abundance of *Microcystis* in the bay was significantly correlated with temperature, TP, TN, and pH. In our samples, the genus *Microcystis* was generally dominant (relative abundances in sample higher than 50%) at temperatures higher than 14°C, TP concentrations higher than 39 µg/l, TN levels equal to or higher than 0.52 mg/l and pH levels equal to or higher than 7.5, independently. It is worth emphasizing that the interpretation of these thresholds should be done cautiously, since they were determined independently for each parameter. For example, percentages of *Microcystis* per sample higher than 50% were only observed at temperatures higher than 14°C, however this does not mean that temperatures higher than 14°C always coincided with dominance of *Microcystis*. The minimum pH level at which *Microcystis* relative abundances higher than 50% were observed in at least one of the samples was 7.5 (average pH range in MB=7.44-9.31), suggesting that this genus can dominate over a wide pH range. This observation contributes to the idea that pH is not the principal cause for *Microcystis* dominance among cyanobacterial taxa. During bloom events the pH tends to increase due to the increased photosynthetic activity by

phytoplankton and the concomitant inorganic carbon decrease (Talling 1976; Paerl 1988; López-Archilla et al. 2004). The fact that *Microcystis* was the only taxa that positively correlated with pH may have been circumstantial and simply be due to the fact that this genus was the dominating taxon during the bloom.

Previous studies have suggested that low nitrogen to phosphorus ratios favour cyanobacterial dominance in freshwater lakes (Smith 1983) and that *Microcystis* tend to dominate the phytoplankton community at TN:TP ratios below 30:1 (reported as by mass) (Liu et al. 2011). In line with these studies, the dynamics of *Microcystis* in MB were better explained by a non-linear non-monotonic relationship with nutrient ratios. At water temperatures equal to or higher than 20°C, the dominance of *Microcystis* tended to increase as the TN:TP (mass) ratio approached 11:1, irrespective of the sampling date or station. Our observations are in agreement with the idea that stoichiometric requirements need to be fulfilled in order for algae to grow, since the biochemical function of each nutrient is unique and cannot be substituted (Hecky and Kilham 1988). Optimum atomic N:P ratios have been previously identified in unialgal laboratory experiments; they were demonstrated to be species-specific and hence to vary among phytoplankton species (Rhee and Gotham 1980). Based on laboratory analyses, the optimum N:P atomic ratio for *Microcystis* sp. was suggested to be 9 (Rhee and Gotham 1980) and 11 for a *Microcystis aeruginosa* strain isolated from flooded rice paddles (Oh and Rhee 1991). In addition, laboratory analysis on *Microcystis* strain PCC7820 indicated that a decreasing N:P ratio in the growth medium stimulated its growth (Kotak et al. 2000).

It has been suggested that in a phytoplanktonic assemblage the nutrient ratios determine the species composition and their relative biomass while the absolute nutrient concentrations determine the total biomass (Bulgakov and Levich 1999). In line with this suggestion, *Microcystis* dominance (relative abundance) and abundance (*Microcystis* 16S rRNA gene copies/ml) in MB was generally higher at nutrient ratios approaching 11:1, but the effect of absolute nutrient concentrations on *Microcystis* abundance was also evident: *Microcystis* abundance correlated positively and significantly with TN and TP. In some instances, nutrient ratios further from 11:1, that had higher absolute nutrient concentrations, resulted in a higher copy number of *Microcystis* 16S rRNA genes than nutrient ratios closer to 11:1 but with lower absolute nutrient concentrations.

We suggest that the dynamics of *Microcystis* in MB during 2009 were determined by both absolute nutrient concentrations and their ratio, in such a way that once temperatures were favourable for *Microcystis* to thrive, the dominance of *Microcystis* depended on the stoichiometric relation between N and P while the absolute nutrient concentrations influenced the amount of biomass that could be supported. *Microcystis* seemed to have an optimal TN:TP (mass) ratio of 11:1, hence departure from this ratio at certain points in time in MB would limit the growth of *Microcystis* either by N or P. For example, in the stations that were analyzed in detail, the pelagic Pel2 and littoral BM5, on June 26th the water temperature was 23.7°C and TN concentrations were higher than 0.52 mg/l. However the TP concentrations were 36.76 and 34.48 µg/l at BM5 and Pel2, respectively, resulting in a TN:TP (mass) ratio>15:1. The fact that *Microcystis* did not dominate on this date may be due to P-limiting conditions for members of this genus. On the other hand, the demise of the main *Microcystis* bloom during the summer (to <5,000 *Microcystis* 16S rRNA gene copies during mid to late September) coincided with a decrease in TN concentrations (<0.52 mg/l). At this point, during mid-September, the water temperature had decreased to 18.9°C but TP concentrations were still well above 39 µg/l and the TN:TP mass ratios were 5.84 and 8.72 for the pelagic and littoral station, respectively, suggesting N-limitation. In line with our observations, a study involving nutrient enrichment experiments on samples from a eutrophic system, found that *Microcystis* demise was linked with N-limitation that led to reduced growth rates (Gobler et al. 2007). In addition, nitrogen limitation can also lead to a reduction in buoyancy and sinking of *Microcystis* cells (Brookes and Ganf 2001; Chu et al. 2007), since nitrogen is necessary for the production of proteins that are required for gas vesicle formation.

In MB, the relative abundance of *Oscillatoriales* and *Chroococcales* other than *Microcystis* tended to decrease significantly when *Microcystis* increased. However, while *Chroococcales* other than *Microcystis* were frequently found when *Microcystis* was present, the *Oscillatoriales* generally thrived under conditions where *Microcystis* did not. The relative abundance of *Oscillatoriales* in the bay significantly and negatively correlated with temperature, pH, TP, DOC, while it positively correlated with the TN:TP mass ratio. In our samples, the order *Oscillatoriales* was generally dominant at temperatures lower than 14°C, pH levels lower than 7.82, DOC levels lower than 4.5

mg/l, TP concentrations lower than 42 μg/l and TN:TP mass ratios higher than 14:1, independently. Nonetheless, when the temperatures were low for other taxa to thrive, the *Oscillatoriales* could also dominate at eutrophic and even hypereutrophic P concentrations and concomitant N-limiting conditions. Moreover, *Oscillatoriales* also dominated on a date in which the air temperature was as high as 20°C while the TP concentration was low (28.7 μg/l). This suggests that members of the *Oscillatoriales* have the potential to thrive under a wide range of temperature and nutrient conditions. Indeed, *Leptolyngbya,* the predominant genus within this order in our samples, has shown robust growth rates under culture conditions, growing well through a temperature range of 22°C to 40°C and tolerating pH levels up to 11 and high solar irradiances (Araujo et al. 2011; Taton et al. 2012). However, members from this order in our samples were mostly confined to periods where the conditions were unfavourable for the other main taxa to bloom, probably due to less efficient competition mechanisms in the *Oscillatoriales* against the *Nostocales* and *Chroococcales*, such as *Microcystis,* during the summer. Nevertheless, *Oscillatoriales*, have the potential to also grow at low light levels (Reynolds et al. 2002; Li and Brand 2007; Araujo et al. 2011; Pumas et al. 2011), and lower temperatures than other cyanobacterial taxa (Robarts and Zohary 1987). These attributes likely enabled them to exploit growing periods earlier and later during the year in MB. In addition, some members of the order *Oscillatoriales*, including members from the *Leptolyngbya* genus (Misra and Tuli 2000; Li et al. 2010), have been reported to carry out nitrogen fixation, allowing them to cope with nitrogen limitation conditions.

The relative abundance of *Nostocales* in the bay significantly correlated with DOC, TP and negatively correlated with DN and the TN:TP mass ratio. The *Nostocales* tended to co-exist with *Microcystis* during spring and summer, and eventually peaked during September once the main *Microcystis* bloom had subsided (to <5,000 *Microcystis* 16S rRNA gene copies during mid to late September), coinciding with a decrease in temperature and nitrogen concentrations; at this point, the TN:TP mass ratios were 5.84:1 and 8.72:1 for the pelagic and littoral station, respectively, suggesting N-limitation. This may be attributed to the nitrogen fixation capabilities of *Nostocales*, such as *Anabaena* (Haselkorn 2007), and their ability to tolerate lower temperatures than *Microcystis* (Robarts and Zohary 1987).

Microcystis and *the Nostocales* seemed to dominate at similar TP concentrations. Nonetheless, it was interesting to note that in both stations analyzed, *Microcystis* bloomed first despite the fact that *Microcystis* does not have the competitive advantage of fixing nitrogen. This could partly be attributed to the favourable TN concentrations and TN:TP ratios in both stations during the periods when *Microcystis* dominated. When *Microcystis* dominated during July in both stations, and also in early May in Pel2, the TP and temperature conditions were favourable for either taxon to dominate, but the nitrogen concentrations were high (TN>0.52 mg/l) and the nutrient (mass) ratios were close to 11:1. These conditions probably favoured *Microcystis* to establish dominance over *Nostocales* and the rest of the cyanobacterial community. It was interesting to note that while at BM5 the relative abundance of *Microcystis* alone was significantly higher than the *Nostocales* during the main *Microcystis* bloom, this was not the case in Pel2. This could also be partly explained by the nutrient ratio, since during this period BM5 exhibited nutrient ratios closer to 11:1 than did Pel2. Another explanation for *Microcystis* outcompeting the *Nostocales* and maintaining its dominance during optimum environmental conditions may have been an allelopathic effect of microcystin on the growth of *Anabaena*. In the bay, intracellular and extracellular microcystin, as well as the *mcyD* gene concentrations, were positively and significantly correlated only with the genus *Microcystis*; suggesting that the main microcystin producers in the system were members from this genus. An allelopathic effect of microcystin on the growth of *Anabaena* has been previously suggested based on observations from competition experiments between toxic *Microcystis aeruginosa* and *Anabaena* PCC7120 (Li and Li 2012). However, further research is needed in this area. More solid evidence may be provided from future competition experiments co-culturing a toxic *Microcystis* strain with *Anabaena* and a *mcy* gene mutant of the same *Microcystis* strain with *Anabaena*. This strategy would reduce the bias that may arise from using different toxic and non-toxic *Microcystis* strains (Briand et al. 2012). In addition these experiments should take into account the effects of *Microcystis* cell-free filtrates and pure microcystin on the growth of *Anabaena.* At present, the allelopathic effects of microcystin on cyanobacterial population dynamics remain highly controversial (Kardinaal et al. 2007; Schatz et al. 2007; Briand et al. 2012; Li and Li 2012).

Although the importance of nutrient ratios and absolute concentrations on cyanobacterial dynamics had been previously and widely recognized (Rhee 1978; Rhee and Gotham 1980; Smith 1982; Tilman et al. 1982; Hecky and Kilham 1988), ranges of TN:TP ratios in which cyanobacteria tends to dominate had been previously observed in the field (Smith 1983; Smith et al. 1995), optimum N:P ratios for *Microcystis* had been determined in laboratory experiments (Rhee and Gotham 1980; Oh and Rhee 1991) and ranges of TN:TP ratios in which *Microcystis* tends to dominate had been previously observed in the field (Liu et al. 2011), to the best of our knowledge, this is the first time that an apparent optimum TN:TP ratio for *Microcystis* dominance has been observed in the field. This observation would provide further support to the theory that N:P ratios are an important determinant of species composition in phytoplankton assemblages (Rhee 1974; Rhee 1978; Rhee and Gotham 1980; Tilman et al. 1982; Smith 1983). In the future, the relationship between *Microcystis* and nutrient ratios should be further studied in other years in MB and in different aquatic systems in order to verify whether this relationship holds from year to year in MB and to analyze how this relationship changes depending on varying environmental conditions and cyanobacterial assemblages among aquatic systems.

5.3 Cyanobacteria and microcystin dynamics in Missisquoi Bay during 2009

5.3.1 Significant relationships between microcystin concentrations and *mcyD* gene, *Microcystis* 16S rRNA gene, estimated total cyanobacteria 16S rRNA gene, and total cyanobacterial cells.

Microcystin concentrations, the abundances of the genes (*mcyD*, *Microcystis* 16S rRNA, estimated total cyanobacteria 16S rRNA) and total cyanobacterial cells were mutually, positively and significantly correlated in MB. This is in agreement with previous correlation analysis done on Lake Erie samples (Rinta-Kanto et al. 2009), except that we also found our estimated total cyanobacterial 16S rRNA gene to be highly correlated with intracellular microcystin, while this was not the case in the study by Rinta-Kanto et al. (2009) due to the fact that the cyanobacterial community in Lake Erie was dominated by members of the non-toxic genus *Synechococcus*. MB samples were dominated by *Microcystis* and despite of the presence of other potentially toxic genera detected via sequencing, such as *Anabaena* (Sivonen 1996) and *Leptolyngbya*

(Richardson et al. 2007; Frazão et al. 2010), only *Microcystis* positively correlated with microcystin and the *mcyD* gene. Since the microcystin production rate is highly correlated with growth rate and cell division (Orr and Jones 1998; Lyck 2004; Briand et al. 2012) and the dynamics of estimated total cyanobacterial 16S rRNA gene abundance in MB were highly dependent on *Microcystis* 16S rRNA gene abundance, it is not surprising to observe strong correlations between the abundances of estimated total cyanobacteria and *Microcystis* with microcystin concentrations in the bay.

Microcystis was not the only *mcyD* gene carrier in the bay, since on occasions the *mcyD* gene concentration was higher than the concentrations of estimated *Microcystis* cells, and other orders comprising potentially toxic genera were identified in the bay. However, the genus *Microcystis* was identified as the main *mcyD* gene carrier and microcystin producer in the bay during 2009 and this was supported by: i) the fact that only the relative abundance of *Microcystis* correlated significantly with the concentrations of microcystin and the *mcyD* gene, ii) high and significant correlations between the abundance of the 16S rRNA gene of *Microcystis* and the concentrations of microcystin and the *mcyD* gene, iii) *mcyD* gene copies higher than 100,000 copies/ml and total microcystin concentrations higher than 1 µg/l were found at a TN:TP mass ratio range from 5 to 15, and they tended to increase as the TN:TP mass ratio approached 11, following the conditions observed to favour *Microcystis* abundance and relative abundance once optimum temperature conditions were established in MB.

Microcystis was also the main microcystin producer in MB during 2006 (Fortin et al. 2010). It is interesting to note that in MB, the abundance of the *mcyD* gene and the toxicity of the blooms have increased. In 2006, the *mcyD* gene did not surpass $1x10^5$ copies/ml (Fortin et al. 2010) and the maximum microcystin concentrations reported were lower than 5 µg/l (Davis et al. 2009; Fortin et al. 2010). The highest concentration of *mcyD* gene found in 2009 samples screened for this gene was $6.7x10^6$ *mcyD* gene copies/ml; this sample had a microcystin concentration of 124.87 µg/l. However, toxin concentrations as high as 774 µg/l were observed in MB (*mcyD* data not available).

5.3.2 Correlations between toxin concentrations, *mcyD*, 16S rRNA genes, total cyanobacterial cells and environmental factors.

No significant associations were found between extracellular microcystin and any of the environmental parameters tested. On the other hand, intracellular microcystin was positively associated with water temperature, pH, DOC, Chl a, TP, TN and DN. It was observed that the *mcyD* gene correlated with the same environmental parameters as did intracellular toxin, except for DN, while the *Microcystis* 16S rRNA gene correlated with the same environmental parameters as the *mcyD* gene, except for DOC and Chl a. Total cyanobacterial cell abundance positively correlated with water temperature, pH, DOC, Chl a, TP, DP and TN, and correlated negatively with the TN:TP mass ratio. Our estimate of total cyanobacterial 16S rRNA gene abundance correlated with the same parameters as total cyanobacterial cell abundance, with the exception of Chl a, nutrient concentrations and ratios.

In agreement with our observations, cyanobacterial growth and blooms are usually associated with higher Chl a concentrations and temperatures, as well as high pH and DOC levels due in part to the increased photosynthetic activity by phytoplankton and the concomitant inorganic carbon decrease (Talling 1976; Paerl 1988; López-Archilla et al. 2004), the excretion of DOC during photosynthesis (Fallowfield and Daft 1988) and the fact that optimum temperatures for maximum growth for cyanobacteria are usually higher than those observed for other members of the phytoplankton (Robarts and Zohary 1987). Based on cyanobacterial abundance, only the *Microcystis* 16S rRNA gene did not correlate with DOC or Chl a levels while the estimated total cyanobacterial 16S rRNA gene did not correlate with Chl a.

While total cyanobacteria cell abundance, derived from microscopy, correlated with nutrient concentrations, our estimate of total cyanobacterial 16S rRNA gene abundance did not correlate with any of the nutrient concentrations. This lack of correspondence may stem from the fact that a wider range of cyanobacteria, that may have a broader range of metabolic needs, could be detected via the amplification of the 16S rRNA gene, which is not the case for microscopic analysis. During microscopy, species may be frequently missed due to their small size (Wilhelm et al. 2006) or the fact that it is difficult to distinguish individual cells in filamentous cyanobacteria (Ernst et al. 2006). Nevertheless, the major caveats of using the 16S RNA gene as a proxy for

cyanobacterial cell abundance could have also impacted the correlation analysis. Different cyanobacterial genomes possess different numbers of rRNA operons. With an average of 1.8 copies per genome, the number of rRNA operons in cyanobacteria can vary from 1 to 4 copies (Engene and Gerwick 2011). This may have led to an overestimation of the cyanobacterial population, depending on which strains were present. In addition, the sequence heterogeneity in the 16S rRNA gene of cyanobacterial cells from natural environments for which there is still no knowledge in sequence databases, complicates the development of primers that that are able to amplify the entire cyanobacterial population from environmental samples (Rinta-Kanto et al. 2009).

Associations between the abundances of the biomolecular/cell abundance markers and environmental factors have been previously reported from a diverse range of aquatic systems worldwide. However, the degree and significance of the associations reported in the literature are varied and sometimes contradictory. This may be due to the fact that environmental parameters may differ from system to system in terms of the degree of relevance and the effect that they have on cyanobacterial dynamics (Rinta-Kanto et al. 2009). In agreement with our findings, correlations between TP and the abundance of *Microcystis,* the *mcyD* gene and microcystin have been observed in Lake Erie (Rinta-Kanto et al. 2009), and in the Daechung Reservoir, Korea, where significant correlations were also found between temperature and microcystin and both total and toxic *Microcystis* (Joung et al. 2011). The abundance of total cyanobacteria, total *Microcystis* and the *mcyB* gene were also found to correlate with TP and with water temperature (except the *mcyB* gene) in a subtropical shallow eutrophic lake in China (Xu et al. 2010). Both TN and TP were positively associated with total and toxigenic *Microcystis* populations, although only TN was positively associated with microcystin, in a tropical reservoir in Singapore (Te and Gin 2011).

The positive correlations observed between the abundance of the biomolecular/cell abundance markers analyzed and the growth-limiting environmental parameters, such as nutrients and temperature, in combination with the positive correlation observed between the abundance of these markers and microcystin concentrations, are in agreement with the hypothesis that there is a linear relationship between cell division rates and microcystin production (Orr and Jones 1998; Lyck 2004).

Hence the factors that promote cell growth also indirectly promote the production of microcystin (Orr and Jones 1998). Since *Microcystis* was identified as the main *mcyD*-carrier and microcystin producer, it was understandable to find the highest *mcyD* gene copies and microcystin concentrations in the range of the environmental parameters that promoted the growth of *Microcystis*. Total microcystin concentrations higher than 1 µg/l and *mcyD* gene copies higher than 100,000 copies/ml were found at a TN:TP mass ratio range from 5:1 to 15:1, and they increased as the TN:TP mass ratio approached 11:1, closely following the dynamics of dominance and abundance of total *Microcystis.* A relationship between microcystin concentrations and nutrient ratios has been previously reported. A recent study of 246 fresh water bodies across Canada (Orihel et al. 2012) found that microcystin concentrations were elevated at low TN:TP (<23:1 reported as by mass) under high nutrient concentrations. A negative relationship between microcystin concentrations and N:P ratios was also reported in several Alberta lakes (Kotak et al. 2000) and Lake Erie (Rinta-Kanto et al. 2009). Kotak and colleagues (2000) suggested that microcystin concentrations are regulated by factors that have an influence on the growth of the microcystin producer and on the cellular production of microcystin, and that the conditions that may be optimal for growth may not necessarily be optimal for microcystin production. They further suggested that the effect of the nutrient ratio on microcystin production is due to the influence that the nutrient ratios have on the growth of *Microcystis* rather than on the production of microcystin. Our data support this idea. However, despite the fact that the highest microcystin and *mcyD* concentrations were observed at nutrient ratios favouring the growth and dominance of *Microcystis* (i.e. ratios approaching 11:1 at water temperatures higher than 20°C), we also observed low microcystin and *mcyD* concentrations under these conditions. In addition, a large variation was observed in microcystin content per *mcyD* copy (fg/copy) that could not be explained by the trend observed with the nutrient ratio. Perhaps the inconsistencies found between the abundances of the *mcyD* gene and microcystin concentrations and this trend with the TN:TP mass ratio found with total *Microcystis* abundance*,* could be partly attributed to: i) during a bloom not all *Microcystis* strains are toxic, and toxic and non-toxic genotypes can co-exist (Vezie et al. 1998; Rinta-Kanto et al. 2009); ii) under certain conditions, the cost of microcystin production may outweigh its benefits leading to

dominance by non-microcystin producing genotypes (Briand et al. 2012); iii) *Microcystis* was not the only *mcyD* gene-carrier in MB; iv) microcystin content can differ between genera and even between isolates within the same genus (Vezie et al. 1998); v) microcystin cell quota can differ at different growth phases because it has been suggested to depend on the cell division rate (Lyck 2004); vi) microcystin cell content in toxic cyanobacteria has been observed to increase in the presence of other cyanobacteria (Briand et al. 2012; Li and Li 2012); and vii) the transcriptional response of the *mcy* operon has been observed to be induced under certain conditions, such as high light intensity and red light (Kaebernick et al. 2000) and iron starvation (Sevilla et al. 2008). Hence, at nutrient ratios approaching 11:1 under favourable temperatures, the *mcyD* gene and toxin concentrations were not always observed to increase because the dynamics of the *mcyD* gene and microcystin production in the bay are not solely dictated by the abundance of total *Microcystis.* In the future, a similar analysis should be done targeting *Microcystis*-specific *mcy* genes, combined with transcriptional analyses in the laboratory, in order to elucidate the factors that promote the toxic *Microcystis* genotypes in MB and microcystin production.

The correlations for TN were higher for the *mcyD* gene and microcystin concentrations than for the *Microcystis* 16S rRNA gene. This may be related to the high nitrogen cost involved in producing the toxin, since nitrogen accounts for over 14% of the molecular weight of microcystin-LR (Botes et al. 1985), and it is also required by the cells for the biosynthesis of this metabolite. Hence, the production of microcystin may be one of the reasons for toxic cells to have higher nitrogen requirements than non-toxic cells in order to grow (Vézie et al. 2002; Briand et al. 2012). However, previous studies have found negative correlations between nitrogen concentrations and toxic *Microcystis* genotypes and microcystin concentrations (Rinta-Kanto et al. 2009). We also observed a positive, although weak, correlation between DN and microcystin concentrations. However, this correlation does not necessarily mean that high nitrogen availability induces the production of microcystin. Transcriptional analysis of the *mcyD* gene under various concentrations of nitrogen revealed no considerable changes in the expression of this gene or microcystin cell content in a toxic *Microcystis* strain exposed to excess or limited nitrate conditions (Sevilla et al. 2010).

New intracellular roles for microcystin have been suggested recently, such as protein modulation and protection by covalent binding of microcystin to proteins (Zilliges et al. 2011) as well as free radical scavenging (Dziallas and Grossart 2011), which may provide a competitive advantage to microcystin-producing species under high light and other conditions that induce oxidative stress. This competitive advantage may partly contribute to the dominance of *Microcystis* and the peaks observed for the *mcyD* gene and microcystin toxin in MB during the warmest months, which are often associated with high solar irradiance. The weak correlation observed between intracellular and extracellular microcystin may further support an intracellular role for microcystin. The highest extracellular microcystin reported in MB in our period of study was 5.12 µg/l while intracellular concentrations reached levels as high as 774 µg/l. It is worth mentioning that if microcystin binds covalently to proteins (Zilliges et al. 2011), it calls for an evaluation on the repercussions that this may have on the detection of microcystin via the methods employed to date.

Production of microcystin requires the presence of the microcystin biosynthesis gene cluster (*mcy* operons) (Tillett et al. 2000), hence only cells that possess it have the potential to produce microcystin. Despite the fact that the presence of the *mcy* genes in potentially toxic cells does not guarantee production of microcystin (e.g., due to mutations along the *mcy* gene cluster (Kaebernick et al. 2001; Kurmayer et al. 2004)), in agreement with Davis et al. (2009), we consider that the *mcyD* gene may be a better indicator of microcystin concentrations *in situ* than total cyanobacterial cells or Chl a, which are the parameters that the WHO recreational water quality guidelines are currently based on. From our correlation analysis, microcystin, especially intracellular microcystin, correlated more strongly with the *mcyD* gene abundance ($sr=0.86$, $p<0.01$), than with total cyanobacterial cell abundance ($sr=0.59$, $p<0.01$), or Chl a concentrations ($sr=0.35$, $p<0.01$). Future guidelines based on molecular markers, such as *mcy* genes, should have a broad applicability in different systems. We recommend that *mcy*-carrying cyanobacterial cells be used as an indicator of environmental microcystin concentrations, as was done in this study, rather than *mcy*-carrying *Microcystis* or total *Microcystis* cells, since not all toxic blooms worldwide are dominated by *Microcystis*.

5.4 Conclusions

A link was demonstrated between nutrient inputs from external sources and the incidence of the cyanobacterial bloom that took place in MB in 2009. The concentrations of P and N in MB in 2009 correlated significantly with the abundance of total cyanobacterial cells, the *Microcystis* 16S rRNA and *mcyD* gene numbers and intracellular microcystin concentration. We suggest that external sources of nutrients, such as surface runoff and animal excreta, played a significant role in the loading of nutrients into the bay and thus in the proliferation of toxic cyanobacterial blooms. This was based on i) the pattern of surface runoff observed in nutrient concentrations and *E. coli* counts, ii) the correlation between *E. coli* counts and nutrient concentrations, iii) the detection of non-aquatic mitochondrial DNA in the bay, and iv) the significant correlation between *E. coli* counts and total cyanobacterial abundance. Potential sources of nutrients from non-aquatic animal excreta in the system comprised rodents, birds, cattle and humans.

The cyanobacterial community of the photic zone of MB was mainly composed of the orders *Chroococcales*, *Nostocales* and *Oscillatoriales* in 2009. During the growing season, the major cyanobacterial taxa were members of the orders *Chroococcales* and *Nostocales*.

Potential microcystin producing genera that were present included *Microcystis, Anabaena* and *Leptolyngbya*. Nevertheless, correlation analysis between the relative abundance of the main cyanobacterial taxa, the abundance of *Microcystis* 16S rRNA and *mcyD* genes and microcystin concentrations suggested that *Microcystis* was the main *mcyD* carrier and microcystin producer, and hence the most problematic genus during 2009. Thus, focus was placed on determining the environmental factors that had a significant effect on the proliferation of *Microcystis* and the production of microcystin during the toxic cyanobacterial bloom that took place in 2009 in MB. Interestingly, once optimum temperatures for *Microcystis* growth had been established, the dynamics of *Microcystis* dominance and abundance in MB were better explained by a non-linear, non-monotonic relationship with nutrient ratios than by a linear relationship with nutrient concentrations. At water temperatures higher than 20°C, the dominance of *Microcystis* tended to increase as the TN:TP (mass) approached 11:1, irrespective of the sampling date or station. We suggest that the dynamics of *Microcystis* abundance and dominance

were determined by nutrient concentrations and ratios, in such a way that once temperatures and nutrient concentrations were favourable for *Microcystis* to thrive, the dominance of *Microcystis* relative to other cyanobacteria depended on the stoichiometric relationship between N and P, while the absolute nutrient concentrations determined how much biomass could be supported. *Microcystis* seemed to have an optimal TN:TP (mass) ratio close to 11:1, hence departure from this ratio at certain times in MB would limit the growth of *Microcystis* either by N or P.

Although the importance of nutrient ratios and absolute concentrations of N and P on cyanobacterial dynamics and dominance have been observed both in the laboratory and field (Rhee 1978; Rhee and Gotham 1980; Smith 1982; Tilman et al. 1982; Smith 1983; Hecky and Kilham 1988; Oh and Rhee 1991; Smith et al. 1995; Liu et al. 2011), to the best of our knowledge, this is the first time that an optimum TN:TP ratio for *Microcystis* dominance has been observed in the field. This observation provides further support to the theory that N:P ratios are an important determinant of species composition in the phytoplankton community (Rhee 1974; Rhee 1978; Rhee and Gotham 1980; Tilman et al. 1982; Smith 1983).

Intracellular concentrations of microcystin equal to or higher than 20 μg/l were observed in the bay at Tw > 20°C; pH >7.8; DOC > 5 mg/l; Chl a > 3 μg/l; TP ≥ 64 μg/l; TN ≥ 0.70 mg/l; DN ≥ 0.38 mg/l, independently. This toxin concentration is 20 times the WHO provisional drinking-water guideline concentration for microcystin-LR and represents a moderate probability of adverse health effects. This is a concern since toxin concentrations as high as 774 μg/l were recorded during 2009 in MB. The highest microcystin and *mcyD* concentrations were observed under the temperatures and TN:TP ratios that favoured the growth and dominance of the genus *Microcystis* relative to other cyanobacteria; however, low microcystin and *mcyD* concentrations were also observed under these conditions. Thus, the dynamics of the *mcyD* gene and microcystin production in MB were not solely dictated by the abundance of total *Microcystis*. This was expected since the *Microcystis* genus is known to have toxic and non-toxic species, and other potential toxin-producing taxa were also present in MB. In the future, similar analyses should be done but targeting *Microcystis*-specific *mcy* genes combined with transcriptional analyses in the field and laboratory in order to elucidate the factors that

promote toxic *Microcystis* genotypes and microcystin production in MB. Most of our exploratory data analysis relied on correlations and associations between biotic and abiotic factors. In the future, deeper statistical analysis could be performed in order to move towards a more mechanistic model to explain the toxic-bloom dynamics.

Under a global warming scenario, where the August average surface water temperatures in Lake Champlain have increased by 1.6-3.8°C over the last 5 decades (Smeltzer et al. 2012), combined with increasing nutrient loading, our data suggests that the incidence and toxicity of cyanobacterial blooms will increase in this temperate aquatic system. An increase of cyanobacterial bloom incidents under these stressors has been suggested in the literature (Paerl and Huisman 2008; Davis et al. 2009; Liu et al. 2011; Paerl and Paul 2012; Schindler 2012; Smeltzer et al. 2012). Thus, it is imperative to increase efforts to reduce nutrient input from external sources into the bay. The presence of non-aquatic mtDNA hosts in our sampling sites suggests that more effort can be made to control nutrient pollution from animal excreta.

Although the validity and prediction potential of the optimum TN:TP ratio for *Microcystis* dominance has yet to be verified by including data from multiple years in MB, this optimum ratio would be of practical importance since it could provide a guideline for nutrient management strategies for avoiding the proliferation of the main toxin producer in MB. There is general agreement that nutrient load reduction should be directed at reversing eutrophication. Traditionally, management strategies targeting nutrient removal have mainly been focused on P alone (Schindler 2012). However, others suggest that N should also be controlled (Gobler et al. 2007; Paerl et al. 2011). Based on our results, the dominance and abundance of the main microcystin producer occurred at high nutrient concentrations reaching a peak at TN:TP mass ratios approaching 11:1. Hence, for management strategies that involve nutrient removal techniques (e.g. nutrient removal from waste water or nutrient precipitation within lakes), we consider that the control of P alone is appropriate because it is easier to avoid nutrient ratios that stimulate noxious cyanobacterial species such as *Microcystis* by controlling only one nutrient. Increase of the nutrient ratio could also be achieved by adding N to the system, however, this may not be a sound long-term option since it would require nutrient addition. Controlling both nutrients may not only be costly, but counterproductive, since the final

nutrient ratio may be difficult to predict; if the final nutrient ratio is at the low end of the spectrum, other noxious, potentially toxic cyanobacteria that are able to fix nitrogen may be stimulated (Schindler 1977; Smith et al. 1995; Schindler et al. 2008). It seems that the ultimate solution, as suggested by Schindler (1977; 2012) is to concentrate efforts on reduction of P concentrations. In this way the final nutrient ratio increases, diverging from those that stimulate cyanobacteria while promoting other phytoplanktonic species (Schindler 1977; Schindler 2012). Based on our observations, the target annual in-lake TP concentration recommended for the euphotic zone in MB (25 μg/l) (NRD 2008) seems appropriate since cyanobacterial blooms higher than 20,000 cells/ml and microcystin concentrations higher than 1 μg/l in MB during 2009 were only observed at TP concentrations higher than 25 μg/l.

List of references

Al-Tebrineh, J., C. Merrick, D. Ryan, A. Humpage, L. Bowling and B. A. Neilan 2012. Community composition, toxigenicity, and environmental conditions during a cyanobacterial bloom occurring along 1,100 Kilometers of the Murray River. Applied and Environmental Microbiology 78, 263-272.

Amard, B. and J. Bertrand-Sarfati 1997. Microfossils in 2000 Ma old cherty stromatolites of the Franceville Group, Gabon. Precambrian Research 81, 197-221.

Anagnostidis, K. and J. Komárek 1985. Modern approach to the classification system of cyanophytes. 1 - Introduction. Archiv für Hydrobiologie 39, 291-302.

Anderson, D. M., P. Hoagland, Y. Kaoru and A. W. White 2000. Estimated annual economic impacts from Harmful Algal Blooms (HABs) in the United States. Technical Report WHOI-2000-11(pp 3-8), Woods Hole Oceanographic Institution, Woods Hole.

Araujo, E. D., J. Patel, C. Araujo, S. P. Rogers, S. M. Short, D. A. Campbell and G. S. Espie 2011. Physiological characterization and light response of the CO2-concentrating mechanism in the filamentous cyanobacterium *Leptolyngbya* sp. CPCC 696. Photosynthesis Research 109, 85-101.

Arment, A. R. and W. W. Carmichael 1996. Evidence that microcystin is a thio-template product. Journal of Phycology 32, 591-597.

Ashelford, K. E., A. J. Weightman and J. C. Fry 2002. PRIMROSE: a computer program for generating and estimating the phylogenetic range of 16S rRNA oligonucleotide probes and primers in conjunction with the RDP‐II database. Nucleic Acids Research 30, 3481-3489.

Beck, E., E. Bochove, E. Smeltzer and D. Leblanc 2012. International Missisquoi Bay study board final report to the International Joint Commission. Retrieved from the International Joint Comission website: http://www.ijc.org/missisquoibayreport/wp-content/uploads/2012/04/International-Missisquoi-Bay-Study-Board.pdf.

Botes, D. P., P. L. Wessels, H. Kruger, M. T. C. Runnegar, S. Santikarn, R. J. Smith, J. C. J. Barna and D. H. Williams 1985. Structural studies on cyanoginosins-LR, -YR, -YA, and -YM, peptide toxins from *Microcystis aeruginosa*. Journal of the Chemical Society, Perkin Transactions 1, 2747-2748.

Bourne, D. G., G. J. Jones, R. L. Blakeley, A. Jones, A. P. Negri and P. Riddles 1996. Enzymatic pathway for the bacterial degradation of the cyanobacterial cyclic peptide toxin microcystin LR. Applied and Environmental Microbiology 62, 4086-4094.

Bourne, D. G., P. Riddles, G. J. Jones, W. Smith and R. L. Blakeley 2001. Characterization of a gene cluster involved in bacterial degradation of the cyanobacterial toxin microcystin LR. Environmental Toxicology 16, 523-534.

Briand, E., M. Bormans, C. Quiblier, M.-J. Salençon and J.-F. Humbert 2012. Evidence of the cost of the production of microcystins by *Microcystis aeruginosa* under differing light and nitrate environmental conditions. PLoS ONE 7, e29981.

Brookes, J. D. and G. G. Ganf 2001. Variations in the buoyancy response of *Microcystis aeruginosa* to nitrogen, phosphorus and light. Journal of Plankton Research 23, 1399-1411.

Bulgakov, N. G. and A. P. Levich 1999. The nitrogen : phosphorus ratio as a factor regulating phytoplankton community structure Archiv für Hydrobiologie 146, 3-22.

Byappanahalli, M. 2004. Indigenous soil bacteria and low moisture may limit but allow faecal bacteria to multiply and become a minor population in tropical soils. Water science and technology 50, 27-32.
Caldwell, J. M., M. E. Raley and J. F. Levine 2007. Mitochondrial multiplex real-time PCR as a source tracking method in fecal-contaminated effluents. Environmental Science & Technology 41, 3277-3283.
Campbell, K. L. and R. A. MacArthur 1996. Seasonal changes in gut mass, forage digestibility, and nutrient selection of wild muskrats (*Ondatra zibethicus*). Physiological Zoology 69, 1215-1231.
Carey, C. C., J. F. Haney and K. L. Cottingham 2007. First report of microcystin-LR in the cyanobacterium *Gloeotrichia echinulata*. Environmental Toxicology 22, 337-339.
Carey, C. C., K. C. Weathers and K. L. Cottingham 2008. *Gloeotrichia echinulata* blooms in an oligotrophic lake: Helpful insights from eutrophic lakes. Journal of Plankton Research 30, 893-904.
Carmichael, W. W. 1992. Cyanobacteria secondary metabolites—the cyanotoxins. Journal of Applied Microbiology 72, 445-459.
Carmichael, W. W., S. M. Azevado, J. S. An, R. J. Molica, E. M. Jochimsen, S. Lau, K. L. Rinehart, G. R. Shaw and G. K. Eaglesham 2001. Human fatalities from cyanobacteria: Chemical and biological evidence for cyanotoxins. Environmental Health Perspectives 109, 663-668.
Castenholz, R. W. and T. B. Norris 2005. Revisionary concepts of species in the Cyanobacteria and their applications. Algological Studies/Archiv für Hydrobiologie 117, 53-69.
CCME (2004). Canadian water quality guidelines for the protection of aquatic life. Phosphorus: Canadian Guidance Framework for the Management of Freshwater Systems. In: Canadian environmental quality guidelines, 2004, Canadian Council of Ministers of the Environment, Winnipeg.
Chen, Z., H. Cheng and X. Chen 2009. Effect of Cl- on photosynthetic bicarbonate uptake in two cyanobacteria *Microcystis aeruginosa* and *Synechocystis* PCC6803. Chinese Science Bulletin 54, 1197-1203.
Chorus, I. and J. Bartram 1999. (Eds.). Toxic cyanobacteria in water: A guide to their public health consequences, monitoring and management (pp 41-177). London: E & FN Spon.
Chou, Y.-H., J. R. Bishcoff, D. Beach and R. D. Goldman 1990. Intermediate filament reorganization during mitosis is mediated by p34cdc2 phosphorylation of vimentin. Cell 62, 1063-1071.
Christiansen, G., J. Fastner, M. Erhard, T. Börner and E. Dittmann 2003. Microcystin biosynthesis in *Planktothrix*: Genes, evolution, and manipulation. Journal of Bacteriology 185, 564-572.
Chu, Z., X. Jin, B. Yang and Q. Zeng 2007. Buoyancy regulation of *Microcystis flos-aquae* during phosphorus-limited and nitrogen-limited growth. Journal of Plankton Research 29, 739-745.
Clark, J., E. Hellgren, J. Parsons, E. Jorgensen, D. Engle and D. Leslie 2005. Nitrogen outputs from fecal and urine deposition of small mammals: Implications for nitrogen cycling. Oecologia 144, 447-455.

Codd, G. A., L. F. Morrison and J. S. Metcalf 2005. Cyanobacterial toxins: Risk management for health protection. Toxicology and Applied Pharmacology 203, 264-272.
Cohn, D. H., A. J. Mileham, M. I. Simon, K. H. Nealson, S. K. Rausch, D. Bonam and T. O. Baldwin 1985. Nucleotide sequence of the *luxA* gene of *Vibrio harveyi* and the complete amino acid sequence of the alpha subunit of bacterial luciferase. Journal of Biological Chemistry 260, 6139-6146.
Cole, J. R., Q. Wang, E. Cardenas, J. Fish, B. Chai, R. J. Farris, A. S. Kulam-Syed-Mohideen, D. M. McGarrell, T. Marsh, G. M. Garrity and J. M. Tiedje 2009. The Ribosomal Database Project: Improved alignments and new tools for rRNA analysis. Nucleic Acids Research 37, D141-D145.
Cousins, I. T., D. J. Bealing, H. A. James and A. Sutton 1996. Biodegradation of microcystin-LR by indigenous mixed bacterial populations. Water Research 30, 481-485.
Davis, C. 1964. Evidence for the eutrophication of Lake Erie from phytoplankton records. Limnology and Oceanography 9, 275-283.
Davis, T., M. Harke, M. Marcoval, J. Goleski, C. Orano-Dawson, D. Berry and C. Gobler 2010. Effects of nitrogenous compounds and phosphorus on the growth of toxic and non-toxic strains of *Microcystis* during cyanobacterial blooms. Aquatic Microbial Ecology 61, 149-162.
Davis, T. W., D. L. Berry, G. L. Boyer and C. J. Gobler 2009. The effects of temperature and nutrients on the growth and dynamics of toxic and non-toxic strains of *Microcystis* during cyanobacteria blooms. Harmful Algae 8, 715-725.
Dawson, R. M. 1998. The toxicology of microcystins. Toxicon 36, 953-962.
Ding, W.-X., H.-M. Shen and C.-N. Ong 2002. Calpain activation after mitochondrial permeability transition in microcystin-induced cell death in rat hepatocytes. Biochemical and Biophysical Research Communications 291, 321-331.
Domingos, P., T. K. Rubim, R. J. Molica, S. M. Azevedo and W. W. Carmichael 1999. First report of microcystin production by picoplanktonic cyanobacteria isolated from a northeast Brazilian drinking water supply. Environmental Toxicology 14, 31-35.
Droop, M. R. 1974. The nutrient status of algal cells in continuous culture. Journal of the Marine Biological Association of the United Kingdom 54, 825-855.
Dziallas, C. and H.-P. Grossart 2011. Increasing oxygen radicals and water temperature select for toxic *Microcystis* sp. PLoS ONE 6, e25569.
Elliott, J., I. Jones and S. Thackeray 2006. Testing the sensitivity of phytoplankton communities to changes in water temperature and nutrient load in a temperate lake. Hydrobiologia 559, 401-411.
Engene, N. and W. H. Gerwick 2011. Intra–genomic 16S rRNA gene heterogeneity in cyanobacterial genomes. Fottea 1, 17-24.
Ernst, B., S. Neser, E. O'Brien, S. J. Hoeger and D. R. Dietrich 2006. Determination of the filamentous cyanobacteria *Planktothrix rubescens* in environmental water samples using an image processing system. Harmful Algae 5, 281-289.
Errington, P. L. 1951. Concerning fluctuations in populations of the prolific and widely distributed muskrat. The American Naturalist 85, 273-292.
Facey, D. E., J. E. Marsden, T. B. Mihuc and E. A. Howe 2012. Lake Champlain 2010: A summary of recent research and monitoring initiatives. Journal of Great Lakes Research 38, Supplement 1, 1-5.

Falconer, I., J. Bartram, I. Chorus, T. Kuiper-Goodman, H. Utkilen, M. Burch and G. A. Codd 1999. Safe levels and safe practices. In I. Chorus and J. Bartram, (Eds.), Toxic cyanobacteria in water: A guide to their public health consequences, monitoring and management (pp 155-177). London: E & FN Spon.

Falconer, I. R. 1999. An overview of problems caused by toxic blue–green algae (cyanobacteria) in drinking and recreational water. Environmental Toxicology 14, 5-12.

Falconer, I. R. and D. S. K. Yeung 1992. Cytoskeletal changes in hepatocytes induced by *Microcystis* toxins and their relation to hyperphosphorylation of cell proteins. Chemico-Biological Interactions 81, 181-196.

Fallowfield, H. J. and M. J. Daft 1988. The extracellular release of dissolved organic carbon by freshwater cyanobacteria and algae and the interaction with Lysobacter CP-1. British Phycological Journal 23, 317-326.

Field, C. B., M. J. Behrenfeld, J. T. Randerson and P. Falkowski 1998. Primary production of the biosphere: Integrating terrestrial and oceanic components. Science 281, 237-240.

Field, K. G. and M. Samadpour 2007. Fecal source tracking, the indicator paradigm, and managing water quality. Water Research 41, 3517-3538.

Fisher, M. M. and E. W. Triplett 1999. Automated approach for ribosomal intergenic spacer analysis of microbial diversity and its application to freshwater bacterial communities. Applied and Environmental Microbiology 65, 4630-4636.

Fisher, T. R., L. W. Harding Jr, D. W. Stanley and L. G. Ward 1988. Phytoplankton, nutrients, and turbidity in the Chesapeake, Delaware, and Hudson estuaries. Estuarine, Coastal and Shelf Science 27, 61-93.

Fogg, G. E. 1969. The Leeuwenhoek lecture, 1968: The physiology of an algal nuisance. Proceedings of the Royal Society of London. Series B. Biological Sciences 173, 175-189.

Fortin, N., R. Aranda-Rodriguez, H. Jing, F. Pick, D. Bird and C. W. Greer 2010. Detection of microcystin-producing cyanobacteria in Missisquoi Bay, Quebec, Canada, using quantitative PCR. Applied and Environmental Microbiology 76, 5105-5112.

Frangeul, L., P. Quillardet, A.-M. Castets, J.-F. Humbert, H. Matthijs, D. Cortez, A. Tolonen, C.-C. Zhang, S. Gribaldo, J.-C. Kehr, Y. Zilliges, N. Ziemert, S. Becker, E. Talla, A. Latifi, A. Billault, A. Lepelletier, E. Dittmann, C. Bouchier and N. Tandeau de Marsac 2008. Highly plastic genome of *Microcystis aeruginosa* PCC 7806, a ubiquitous toxic freshwater cyanobacterium. BMC Genomics 9, 274.

Frazão, B., R. Martins and V. Vasconcelos 2010. Are known cyanotoxins involved in the toxicity of picoplanktonic and filamentous north atlantic marine cyanobacteria? Marine Drugs 8, 1908-1919.

Fujimoto, N., R. Sudo, N. Sugiura and Y. Inamori 1997. Nutrient-limited growth of *Microcystis aeruginosa* and *Phormidium tenue* ad competition under various N:P ratios and temperatures. Limnology and Oceanography 42, 250-256.

Fulton, R. S. and H. W. Paerl 1987. Toxic and inhibitory effects of the blue-green algae *Microcystis aeruginosa* on herbivorous zooplankton. Journal of Plankton Research 9, 837-855.

Gantar, M., R. Sekar and L. L. Richardson 2009. Cyanotoxins from black band disease of corals and from other coral reef environments. Microbial Ecology 58, 856-864.

Gasser, M., O. M. Bolinder, S. Martel, D. Poulin, I. Beaudin, A. R. Michaud and A. Drouin 2010. Impacts agroenvironnementaux associés à la culture et au prélèvement de biomasses végétales agricoles pour la production de bioproduits industriels. Rapport final (pp 20-42). Quebec: Institut de recherche et de développement en agroenvironnement (IRDA)

Geider, R. J. 1987. Light and temperature dependence of the carbon to chlorophyll a ratio in microalgae and cyanobacteria: Implications for physiology and growth of phytoplankton. New Phytologist 106, 1-34.

Ghernaout, B., D. Ghernaout and A. Saiba 2010. Algae and cyanotoxins removal by coagulation/flocculation: A review. Desalination and Water Treatment 20, 133-143.

Gobler, C. J., T. W. Davis, K. J. Coyne and G. L. Boyer 2007. Interactive influences of nutrient loading, zooplankton grazing, and microcystin synthetase gene expression on cyanobacterial bloom dynamics in a eutrophic New York lake. Harmful Algae 6, 119-133.

Hamady, M. and R. Knight 2009. Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. Genome Research 19, 1141-1152.

Hansen, D. L., J. J. Clark, S. Ishii, M. J. Sadowsky and R. E. Hicks 2008. Sources and sinks of *Escherichia coli* in benthic and pelagic fish. Journal of Great Lakes Research 34, 228-234.

Hardina, C. M. and R. S. Fujioka 1991. Soil: The environmental source of *Escherichia coli* and *enterococci* in Hawaii's streams. Environmental Toxicology and Water Quality 6, 185-195.

Haselkorn, R. 2007. Heterocyst Differentiation and Nitrogen Fixation in *Anabaena*. In C. Elmerich and W. E. Newton (Eds.), Associative and Endophytic Nitrogen-fixing Bacteria and Cyanobacterial Associations (pp 233-255). Dordrecht: Springer Netherlands. 41, 65-68.

Healey, F. P. 1985. Interacting effects of light and nutrient limitation on the growth rate of *Synechococcus linearis* (*cyanophyceae*). Journal of Phycology 21, 134-146.

Hecky, R. E. and P. Kilham 1988. Nutrient limitation of phytoplankton in freshwater and marine environments: A review of recent evidence on the effects of enrichment. Limnology and Oceanography 33, 796-822.

Ho, L., E. Sawade and G. Newcombe 2012. Biological treatment options for cyanobacteria metabolite removal – A review. Water Research 46, 1536-1548.

Ho, L., T. Tang, D. Hoefel and B. Vigneswaran 2012. Determination of rate constants and half-lives for the simultaneous biodegradation of several cyanobacterial metabolites in Australian source waters. Water Research 46, 5735-5746.

Hoffmann, L., J. Komárek and J. Kaštovský 2005. System of cyanoprokaryotes (cyanobacteria) state in 2004. Archiv für Hydrobiologie 117, 95-115.

Horner, R. and G. C. Schrader 1982. Relative contributions of ice algae, phytoplankton, and benthic microalgae to primary production in nearshore regions of the Beaufort Sea. Arctic 35, 485-503.

Howery, L. D. and J. A. Pfister 1990. Dietary and fecal concentrations of nitrogen and phosphorus in penned white-tailed deer does. The Journal of Wildlife Management 54, 383-389.

Huisman, J. and F. D. Hulot 2005. Population dynamics of harmful cyanobacteria. Harmful Cyanobacteria. J. Huisman, H. P. Matthijs and P. Visser, Springer Netherlands. 3: 143-176.

Huisman, J., J. Sharples, J. M. Stroom, P. M. Visser, W. E. A. Kardinaal, J. M. H. Verspagen and B. Sommeijer 2004. Changes in turbulent mixing shift competition for light between phytoplankton species. Ecology 85, 2960-2970.
Huson, D. H., S. Mitra, H.-J. Ruscheweyh, N. Weber and S. C. Schuster 2011. Integrative analysis of environmental sequences using MEGAN4. Genome Research 21, 1552-1560.
Ishii, S., D. L. Hansen, R. E. Hicks and M. J. Sadowsky 2007. Beach sand and sediments are temporal sinks and sources of *Escherichia coli* in Lake Superior. Environmental Science & Technology 41, 2203-2209.
Ishii, S., W. B. Ksoll, R. E. Hicks and M. J. Sadowsky 2006. Presence and growth of naturalized *Escherichia coli* in temperate soils from Lake Superior watersheds. Applied and Environmental Microbiology 72, 612-621.
Itou, Y., S. Okada and M. Murakami 2001. Two structural isomeric siderophores from the freshwater cyanobacterium *Anabaena cylindrica* (NIES-19). Tetrahedron 57, 9093-9099.
Iyengar, V., G. P. Albaugh, A. Lohani and P. P. Nair 1991. Human stools as a source of viable colonic epithelial cells. The FASEB Journal 5, 2856-2859.
Izaguirre, G., A.-D. Jungblut and B. A. Neilan 2007. Benthic cyanobacteria (*Oscillatoriaceae*) that produce microcystin-LR, isolated from four reservoirs in southern California. Water Research 41, 492-498.
Jacquet, S., J.-F. Briand, C. Leboulanger, C. Avois-Jacquet, L. Oberhaus, B. Tassin, B. Vinçon-Leite, G. Paolini, J.-C. Druart, O. Anneville and J.-F. Humbert 2005. The proliferation of the toxic cyanobacterium *Planktothrix rubescens* following restoration of the largest natural French lake (Lac du Bourget). Harmful Algae 4, 651-672.
Jensen, H. S. and F. O. Andersen 1992. Importance of temperature, nitrate, and pH for phosphate release from aerobic sediments of four shallow, eutrophic lakes. Limnology and Oceanography 37, 577-589.
Johnston, W. R., F. Ittihadieh, R. M. Daum and A. F. Pillsbury 1965. Nitrogen and phosphorus in tile drainage effluent. Soil Science Society of America Journal 29, 287-289.
Joung, S.-H., H.-M. Oh, S.-R. Ko and C.-Y. Ahn 2011. Correlations between environmental factors and toxic and non-toxic *Microcystis* dynamics during bloom in Daechung Reservoir, Korea. Harmful Algae 10, 188-193.
Kaebernick, M., E. Dittmann, T. Börner and B. A. Neilan 2002. Multiple alternate transcripts direct the biosynthesis of microcystin, a cyanobacterial nonribosomal peptide. Applied and Environmental Microbiology 68, 449-455.
Kaebernick, M., B. A. Neilan, T. Börner and E. Dittmann 2000. Light and the transcriptional response of the microcystin biosynthesis gene cluster. Applied and Environmental Microbiology 66, 3387-3392.
Kaebernick, M., T. Rohrlack, K. Christoffersen and B. A. Neilan 2001. A spontaneous mutant of microcystin biosynthesis: Genetic characterization and effect on *Daphnia*. Environmental Microbiology 3, 669-679.
Kaneko, T., N. Nakajima, S. Okamoto, I. Suzuki, Y. Tanabe, M. Tamaoki, Y. Nakamura, F. Kasai, A. Watanabe, K. Kawashima, Y. Kishida, A. Ono, Y. Shimizu, C. Takahashi, C. Minami, T. Fujishiro, M. Kohara, M. Katoh, N. Nakazaki, S. Nakayama, M. Yamada, S. Tabata and M. M. Watanabe 2007. Complete genomic structure of the

bloom-forming toxic cyanobacterium *Microcystis aeruginosa* NIES-843. DNA Research 14, 247-256.
Kardinaal, W., L. Tonk, I. Janse, S. Hol, P. Slot, J. Huisman and P. M. Visser 2007. Competition for light between toxic and nontoxic strains of the harmful cyanobacterium *Microcystis*. Applied and Environmental Microbiology 73, 2939-2946.
Kehoe, D. M. 2010. Chromatic adaptation and the evolution of light color sensing in cyanobacteria. Proceedings of the National Academy of Sciences 107, 9029-9030.
Keil, C., A. Forchert, J. Fastner, U. Szewzyk, W. Rotard, I. Chorus and R. Krätke 2002. Toxicity and microcystin content of extracts from a *Planktothrix bloom* and two laboratory strains. Water Research 36, 2133-2139.
Kiefer, L. A., D. R. Shelton, Y. Pachepsky, R. Blaustein and M. Santin-Duran 2012. Persistence of *Escherichia coli* introduced into streambed sediments with goose, deer and bovine animal waste. Letters in Applied Microbiology 55, 345-353.
Kim, S.-G., S.-K. Rhee, C.-Y. Ahn, S.-R. Ko, G.-G. Choi, J.-W. Bae, Y.-H. Park and H.-M. Oh 2006. Determination of cyanobacterial diversity during algal blooms in Daechung Reservoir, Korea, on the basis of *cpcBA* intergenic spacer region analysis. Applied and Environmental Microbiology 72, 3252-3258.
Kirchmann, H. and S. Pettersson 1994. Human urine - Chemical composition and fertilizer use efficiency. Nutrient Cycling in Agroecosystems 40, 149-154.
Komárková, J., M. A. Mugnai, C. Sili, O. Komárek and S. Turicchia 2005. Stable morphospecies within the 16S rRNA monophyletic genus *Microcystis* (Kutzing) Kutzing. Algological Studies/Archiv für Hydrobiologie 117, 279-295.
Kormas, K. A., S. Gkelis, E. Vardaka and M. Moustaka-Gouni 2011. Morphological and molecular analysis of bloom-forming cyanobacteria in two eutrophic, shallow Mediterranean lakes. Limnologica - Ecology and Management of Inland Waters 41, 167-173.
Kotak, B. G., A. K. Y. Lam, E. E. Prepas and S. E. Hrudey 2000. Role of chemical and physical variables in regulating microcystin-LR concentrations in phytoplankton of eutrophic lakes. Canadian Journal of Fisheries and Aquatic Sciences 57, 1584-1593.
Kromkamp, J. 1987. Formation and functional significance of storage products in cyanobacteria. New Zealand Journal of Marine and Freshwater Research 21, 457-465.
Kuczynski, J., Z. Liu, C. Lozupone, D. McDonald, N. Fierer and R. Knight 2010. Microbial community resemblance methods differ in their ability to detect biologically relevant patterns. Nature Methods, 813–819.
Kurmayer, R., G. Christiansen, J. Fastner and T. Börner 2004. Abundance of active and inactive microcystin genotypes in populations of the toxic cyanobacterium *Planktothrix* spp. Environmental Microbiology 6, 831-841.
Leak, L. V. 1967. Fine structure of the mucilaginous sheath of *Anabaena* sp. Journal of Ultrastructure Research 21, 61-74.
Lee, C. K., C. W. Herbold, S. W. Polson, K. E. Wommack, S. J. Williamson, I. R. McDonald and S. C. Cary 2012. Groundtruthing Next-Gen Sequencing for Microbial Ecology–Biases and Errors in Community Structure Estimates from PCR Amplicon Pyrosequencing. PLoS ONE 7, e44224.
Lee, S. J., M. H. Jang, H. S. Kim, B. D. Yoon and H. M. Oh 2000. Variation of microcystin content of *Microcystis aeruginosa* relative to medium N:P ratio and growth stage. Journal of Applied Microbiology 89, 323-329.

Levine, S. N., A. Lini, M. L. Ostrofsky, L. Bunting, H. Burgess, P. R. Leavitt, D. Reuter, A. Lami, P. Guilizzoni and E. Gilles 2012. The eutrophication of Lake Champlain's northeastern arm: Insights from paleolimnological analyses. Journal of Great Lakes Research 38, Supplement 1, 35-48.

Ley, R. E., M. Hamady, C. Lozupone, P. J. Turnbaugh, R. R. Ramey, J. S. Bircher, M. L. Schlegel, T. A. Tucker, M. D. Schrenzel, R. Knight and J. I. Gordon 2008. Evolution of mammals and their gut microbes. Science 320, 1647-1651.

Li, Y., J. A. Chen, Q. Zhao, C. Pu, Z. Qiu, R. Zhang and S. Weiqun 2011. A cross-sectional investigation of chronic exposure to microcystin in relationship to childhood liver damage in the Three Gorges Reservoir Region, China. Environmental Health Perspectives 119, 1483-1488.

Li, Y. and D. Li 2012. Competition between toxic *Microcystis aeruginosa* and nontoxic *Microcystis wesenbergii* with *Anabaena* PCC7120. Journal of Applied Phycology 24, 69-78.

Li, Z. and J. Brand 2007. *Leptolyngbya nodulosa* sp. nov. (*Oscillatoriaceae*), a subtropical marine cyanobacterium that produces a unique multicellular structure. Phycologia 46, 396-401.

Li, Z., J. Yu, K.-R. Kim and J. Brand 2010. Nitrogen fixation by a marine non-heterocystous cyanobacterium requires a heterotrophic bacterial consort. Environmental Microbiology 12, 1185-1193.

Liu, X., X. Lu and Y. Chen 2011. The effects of temperature and nutrient ratios on *Microcystis* blooms in Lake Taihu, China: An 11-year investigation. Harmful Algae 10, 337-343.

López-Archilla, A., D. Moreira, P. López-García and C. Guerrero 2004. Phytoplankton diversity and cyanobacterial dominance in a hypereutrophic shallow lake with biologically produced alkaline pH. Extremophiles 8, 109-115.

Lyck, S. 2004. Simultaneous changes in cell quotas of microcystin, chlorophyll a, protein and carbohydrate during different growth phases of a batch culture experiment with *Microcystis aeruginosa*. Journal of Plankton Research 26, 727-736.

Lyra, C., S. Suomalainen, M. Gugger, C. Vezie, P. Sundman, L. Paulin and K. Sivonen 2001. Molecular characterization of planktic cyanobacteria of *Anabaena*, *Aphanizomenon*, *Microcystis* and *Planktothrix* genera. International Journal of Systematic and Evolutionary Microbiology 51, 513-526.

MAMROT 2010. Évaluation de performance des ouvrages municipaux d'assainissement des eaux pour l'année 2009 (ISBN 978-2-550-59204-4). Retrieved from the Quebec Ministry of Municipal Affairs, Regions and Land Occupancy website: http://www.mamrot.gouv.qc.ca.

Mandal, S., J. Rath and S. P. Adhikary 2011. Adaptation strategies of the sheathed cyanobacterium *Lyngbya majuscula* to ultraviolet-B. Journal of Photochemistry and Photobiology B: Biology 102, 115-122.

Mankiewicz, J., Z. Walter and M. Zalewski 2003. Natural toxins from cyanobacteria. Acta Biologica Cracoviensia 45, 9-20.

Martellini, A., P. Payment and R. Villemur 2005. Use of eukaryotic mitochondrial DNA to differentiate human, bovine, porcine and ovine sources in fecally contaminated surface water. Water Research 39, 541-548.

Martins, A. and V. Vasconcelos 2011. Use of qPCR for the study of hepatotoxic cyanobacteria population dynamics. Archives of Microbiology 193, 615-627.

Maske, S., L. Sangolkar and T. Chakrabarti 2010. Temporal variation in density and diversity of cyanobacteria and cyanotoxins in lakes at Nagpur (Maharashtra State), India. Environmental Monitoring and Assessment 169, 299-308.

Masson, L., Y. Comeau, R. Brousseau, R. Samson and C. Greer 1993. Construction and application of chromosomally integrated *lac-lux* gene markers to monitor the fate of a 2,4-dichlorophenoxyacetic acid-degrading bacterium in contaminated soils. Microbial releases : viruses, bacteria, fungi 1, 209-216.

McQueen, D. J. and D. R. S. Lean 1987. Influence of water temperature and nitrogen to phosphorus ratios on the dominance of blue-green algae in Lake St. George, Ontario. Canadian Journal of Fisheries and Aquatic Sciences 44, 598-604.

Michaud, A. R. M. R. L. 2004. Cropping, soil type and manure application effects on phosphorus export and bioavailability. Canadian Journal of Soil Science 84, 295‐305.

Misra, H. S. and R. Tuli 2000. Differential expression of photosynthesis and nitrogen fixation genes in the cyanobacterium *Plectonema boryanum*. Plant Physiology 122, 731-736.

Nishiwaki-Matsushima, R., T. Ohta, S. Nishiwaki, M. Suganuma, K. Kohyama, T. Ishikawa, W. W. Carmichael and H. Fujiki 1992. Liver tumor promotion by the cyanobacterial cyclic peptide toxin microcystin-LR. Journal of Cancer Research and Clinical Oncology 118, 420-424.

Nishizawa, T., A. Ueda, M. Asayama, K. Fujii, K.-i. Harada, K. Ochi and M. Shirai 2000. Polyketide synthase gene coupled to the peptide synthetase module involved in the biosynthesis of the cyclic heptapeptide microcystin. Journal of Biochemistry 127, 779-789.

Nizan, S., C. Dimentman and M. Shilo 1986. Acute toxic effects of the cyanobacterium *Microcystis aeruginosa* on *Daphnia magna*. Limnology and Oceanography 31, 497–502.

NRD 2008. Vermont water quality standards. State of Vermont Natural Resources Board. Water Resources Panel. Vt. Code R. 12 004 052 January 2008. http://www.nrb.state.vt.us/wrp/publications/wqs.pdf.

Oh, H.-M., S. J. Lee, M.-H. Jang and B.-D. Yoon 2000. Microcystin production by *Microcystis aeruginosa* in a phosphorus-limited chemostat. Applied and Environmental Microbiology 66, 176-179.

Oh, H.-M. and G. Y. Rhee 1991. A comparative study of microalgae isolated from flooded rice paddies: light-limited growth, C fixation, growth efficiency and relative N and P requirement. Journal of Applied Phycology 3, 211-220.

Orihel, D. M., D. F. Bird, M. Brylinsky, H. Chen, D. B. Donald, D. Y. Huang, A. Giani, D. Kinniburgh, H. Kling, B. G. Kotak, P. R. Leavitt, C. C. Nielsen, S. Reedyk, R. C. Rooney, S. B. Watson, R. W. Zurawell and R. D. Vinebrooke 2012. High microcystin concentrations occur only at low nitrogen-to-phosphorus ratios in nutrient-rich Canadian lakes. Canadian Journal of Fisheries and Aquatic Sciences 69, 1-6.

Orr, P. T. and G. J. Jones 1998. Relationship between microcystin production and cell division rates in nitrogen-limited *Microcystis aeruginosa* cultures. Limnology and Oceanography 43, 1604-1614.

Oudra, B., M. Loudiki, V. Vasconcelos, B. Sabour, B. Sbiyyaa, K. Oufdou and N. Mezrioui 2002. Detection and quantification of microcystins from cyanobacteria

strains isolated from reservoirs and ponds in Morocco. Environmental Toxicology 17, 32-39.
Owens, O. V. H. and W. E. Esaias 1976. Physiological responses of phytoplankton to major environmental factors. Annual Review of Plant Physiology 27, 461-483.
Padia, R., R. Karthikeyan, S. Mukhtar and I. Parker 2012. Occurrence and fate of *E. coli* from various non-point sources in a subtropical watershed. Journal of Natural and Environmental Sciences 3, 9-18.
Paerl, H. W. 1988. Nuisance phytoplankton blooms in coastal, estuarine, and inland waters. Limnology and Oceanography 33, 823-847.
Paerl, H. W., R. S. Fulton, P. H. Moisander and J. Dyble 2001. Harmful freshwater algal blooms, with an emphasis on cyanobacteria. The Scientific World Journal 1, 76-113.
Paerl, H. W., N. S. Hall and E. S. Calandrino 2011. Controlling harmful cyanobacterial blooms in a world experiencing anthropogenic and climatic-induced change. Science of the Total Environment 409, 1739-1745.
Paerl, H. W. and J. Huisman 2008. Blooms like it hot. Science 320, 57-58.
Paerl, H. W. and V. J. Paul 2012. Climate change: Links to global expansion of harmful cyanobacteria. Water Research 46, 1349-1363.
Paerl, H. W. and J. T. Scott 2010. Throwing fuel on the fire: Synergistic effects of excessive nitrogen inputs and global warming on harmful algal blooms. Environmental Science & Technology 44, 7756-7758.
Paul, V., R. Thacker, K. Banks and S. Golubic 2005. Benthic cyanobacterial bloom impacts the reefs of South Florida (Broward County, USA). Coral Reefs 24, 693-697.
Pearson, L. A. and B. A. Neilan 2008. The molecular genetics of cyanobacterial toxicity as a basis for monitoring water quality and public health risk. Current Opinion in Biotechnology 19, 281-288.
Pitois, S. G., M. H. Jackson and B. J. B. Wood 1997. Summer bloom of *Gloeotrichia echinulata* and *Aphanizomenon flos-aquae* and phosphorus levels in Antermony Loch, central Scotland. International Journal of Environmental Health Research 7, 131-140.
Pumas, C., P. Vacharapiyasophon, Y. Peerapornpisal, P. Leelapornpisid, W. Boonchum, M. Ishii and C. Khanongnuch 2011. Thermostablility of phycobiliproteins and antioxidant activity from four thermotolerant cyanobacteria. Phycological Research 59, 166-174.
Qiu, B. and K. Gao 2002. Effects of CO2 enrichment on the bloom-forming cyanobacterium *Microcystis aeruginosa* (*cyanophyceae*): Physiological responses and relationships with the availability of dissolved inorganic carbon. Journal of Phycology 38, 721-729.
Qiu, X., L. Wu, H. Huang, P. E. McDonel, A. V. Palumbo, J. M. Tiedje and J. Zhou 2001. Evaluation of PCR-Generated Chimeras, Mutations, and Heteroduplexes with 16S rRNA Gene-Based Cloning. Applied and Environmental Microbiology 67, 880-887.
Rajaniemi-Wacklin, P., A. Rantala, M. A. Mugnai, S. Turicchia, S. Ventura, J. Komárková, L. Lepistö and K. Sivonen 2006. Correspondence between phylogeny and morphology of *Snowella* spp. and *Woronichinia naegeliana*, cyanobacteria commonly occurring in lakes. Journal of Phycology 42, 226-232.
Rantala, A., D. P. Fewer, M. Hisbergues, L. Rouhiainen, J. Vaitomaa, T. Börner and K. Sivonen 2004. Phylogenetic evidence for the early evolution of microcystin synthesis.

Proceedings of the National Academy of Sciences of the United States of America 101, 568-573.
Rantala, A., P. Rajaniemi-Wacklin, C. Lyra, L. Lepistö, J. Rintala, J. Mankiewicz-Boczek and K. Sivonen 2006. Detection of microcystin-producing cyanobacteria in Finnish lakes with genus-specific microcystin synthetase gene E (*mcyE*) PCR and associations with environmental factors. Applied and Environmental Microbiology 72, 6101-6110.
Reynolds, C. 1998. What factors influence the species composition of phytoplankton in lakes of different trophic status? Hydrobiologia 369, 11-26.
Reynolds, C. S. 1984. Phytoplankton periodicity: the interactions of form, function and environmental variability. Freshwater Biology 14, 111-142.
Reynolds, C. S. 1986. Experimental manipulations of the phytoplankton periodicity in large limnetic enclosures in Blelham Tarn, English Lake District. Hydrobiologia 138, 43-64.
Reynolds, C. S., V. Huszar, C. Kruk, L. Naselli-Flores and S. Melo 2002. Towards a functional classification of the freshwater phytoplankton. Journal of Plankton Research 24, 417-428.
Reynolds, C. S. and A. E. Walsby 1975. Water blooms. Biological Reviews 50, 437-481.
Rhee, G. Y. 1974. Phosphate uptake under nitrogen limitation by *Scenedesmus* sp. and its ecological implications. . Journal of Phycology, 470-475.
Rhee, G. Y. 1978. Effects of N:P atomic ratios and nitrate limitation on algal growth, cell composition, and nitrate uptake. Limnology and Oceanography 23, 10-25.
Rhee, G. Y. and I. J. Gotham 1980. Optimum N:P ratios and coexistence of planktonic algae. Journal of Phycology 16, 486-489.
Rhee, J. S., H. U. Dahms, B. S. Choi, J. S. Lee and I. Y. Choi 2012. Identification and analysis of whole microcystin synthetase genes from two Korean strains of the cyanobacterium *Microcystis aeruginosa*. Genes and Genomics 34, 435-439.
Richardson, L. L., R. Sekar, J. L. Myers, M. Gantar, J. D. Voss, L. Kaczmarsky, E. R. Remily, G. L. Boyer and P. V. Zimba 2007. The presence of the cyanobacterial toxin microcystin in black band disease of corals. FEMS Microbiology Letters 272, 182-187.
Richter, C., J. W. Park and B. N. Ames 1988. Normal oxidative damage to mitochondrial and nuclear DNA is extensive. Proceedings of the National Academy of Sciences 85, 6465-6467.
Rigonato, J., A. D. Kent, D. O. Alvarenga, F. D. Andreote, R. M. Beirigo, P. Vidal-Torrado and M. F. Fiore 2012. Drivers of cyanobacterial diversity and community composition in mangrove soils in south-east Brazil. Environmental Microbiology, Online ISSN: 1462-2920 (Published online in advance of the print journal).
Rinta-Kanto, J. M., E. A. Konopko, J. M. DeBruyn, R. A. Bourbonniere, G. L. Boyer and S. W. Wilhelm 2009. Lake Erie Microcystis: Relationship between microcystin production, dynamics of genotypes and environmental parameters in a large lake. Harmful Algae 8, 665-673.
Robarts, R. D. and T. Zohary 1987. Temperature effects on photosynthetic capacity, respiration, and growth rates of bloom‐forming cyanobacteria. New Zealand Journal of Marine and Freshwater Research 21, 391-399.
Rothberg, J. M., W. Hinz, T. M. Rearick, J. Schultz, W. Mileski, M. Davey, J. H. Leamon, K. Johnson, M. J. Milgrew, M. Edwards, J. Hoon, J. F. Simons, D. Marran, J. W. Myers, J. F. Davidson, A. Branting, J. R. Nobile, B. P. Puc, D. Light, T. A. Clark,

M. Huber, J. T. Branciforte, I. B. Stoner, S. E. Cawley, M. Lyons, Y. Fu, N. Homer, M. Sedova, X. Miao, B. Reed, J. Sabina, E. Feierstein, M. Schorn, M. Alanjary, E. Dimalanta, D. Dressman, R. Kasinskas, T. Sokolsky, J. A. Fidanza, E. Namsaraev, K. J. McKernan, A. Williams, G. T. Roth and J. Bustillo 2011. An integrated semiconductor device enabling non-optical genome sequencing. Nature 475, 348–352.

Rouhiainen, L., T. Vakkilainen, B. L. Siemer, W. Buikema, R. Haselkorn and K. Sivonen 2004. Genes coding for hepatotoxic heptapeptides (microcystins) in the cyanobacterium *Anabaena* Strain 90. Applied and Environmental Microbiology 70, 686-692.

Ruys, T., O. Lorvelec, A. Marre and I. Bernez 2011. River management and habitat characteristics of three sympatric aquatic rodents: common muskrat, coypu and European beaver. European Journal of Wildlife Research 57, 851-864.

Sakamoto, M. 1966. Primary production by phytoplankton community in some Japanese lakes and its dependence on lake depth. Archiv für Hydrobiologie 62, 1-28.

Salazar, J. J. and B. Van Houten 1997. Preferential mitochondrial DNA injury caused by glucose oxidase as a steady generator of hydrogen peroxide in human fibroblasts. Mutation Research/DNA Repair 385, 139-149.

Schatz, D., Y. Keren, A. Vardi, A. Sukenik, S. Carmeli, T. Börner, E. Dittmann and A. Kaplan 2007. Towards clarification of the biological role of microcystins, a family of cyanobacterial toxins. Environmental Microbiology 9, 965-970.

Schindler, D. W. 1974. Eutrophication and recovery in experimental lakes: Implications for lake management. Science 184, 897-899.

Schindler, D. W. 1977. Evolution of phosphorus limitation in lakes: Natural mechanisms compensate for deficiencies of nitrogen and carbon in eutrophied lakes. Science 195, 260-262.

Schindler, D. W. 2012. The dilemma of controlling cultural eutrophication of lakes. Proceedings of the Royal Society B: Biological Sciences 279, 4322-4333.

Schindler, D. W., R. E. Hecky, D. L. Findlay, M. P. Stainton, B. R. Parker, M. J. Paterson, K. G. Beaty, M. Lyng and S. E. M. Kasian 2008. Eutrophication of lakes cannot be controlled by reducing nitrogen input: Results of a 37-year whole-ecosystem experiment. Proceedings of the National Academy of Sciences 105, 11254-11258.

Schopf, J. W. (2002). The fossil record: Tracing the roots of the cyanobacterial lineage. The Ecology of Cyanobacteria. B. Whitton and M. Potts, Springer Netherlands**:** 13-35.

Sevilla, E., B. Martin-Luna, L. Vela, M. T. Bes, M. F. Fillat and M. L. Peleato 2008. Iron availability affects *mcyD* expression and microcystin-LR synthesis in *Microcystis aeruginosa* PCC7806. Environmental Microbiology 10, 2476-2483.

Sevilla, E., B. Martin-Luna, L. Vela, M. T. Bes, M. L. Peleato and M. Fillat 2010. Microcystin-LR synthesis as response to nitrogen: Transcriptional analysis of the *mcyD* gene in *Microcystis aeruginosa* PCC7806. Ecotoxicology 19, 1167-1173.

Shanley, J. B. and J. C. Denner 1999. The hydrology of the Lake Champlain basin. In Lake Champlain in Transition: from research toward restoration, edited by T.O Manley and P.L. Manley. Washington, DC: Americal Geophysical Union. p 41-67.

Sheldrick, W., J. K. Syers and J. Lingard 2003. Contribution of livestock excreta to nutrient balances. Nutrient Cycling in Agroecosystems 66, 119-131.

Shikuma, N. J. and M. G. Hadfield 2009. Marine biofilms on submerged surfaces are a reservoir for *Escherichia coli* and *Vibrio cholerae*. Biofouling 26, 39-46.

Shimizu, K., H. Maseda, K. Okano, T. Kurashima, Y. Kawauchi, Q. Xue, M. Utsumi, Z. Zhang and N. Sugiura 2012. Enzymatic pathway for biodegrading microcystin LR in *Sphingopyxis* sp. C-1. Journal of Bioscience and Bioengineering 114, 630-634.
Simoneau, M. 2007. État de l'écosystème aquatique du bassin versant de la baie Missisquoi : faits saillants 2001-2004, Québec, ministère du Développement durable, de l'Environnement et des Parcs, Direction du suivi de l'état de l'environnement, ISBN 978-2-550-49625-0
Sitoki, L., R. Kurmayer and E. Rott 2012. Spatial variation of phytoplankton composition, biovolume, and resulting microcystin concentrations in the Nyanza Gulf (Lake Victoria, Kenya). Hydrobiologia 691, 109-122.
Sivonen, K. 1996. Cyanobacterial toxins and toxin production. Phycologia 35, 12-24.
Sivonen, K. 2008. Emerging high throughput analyses of cyanobacterial toxins and toxic cyanobacteria. In H.K Hudnell (Ed.), Cyanobacterial Harmful Algal Blooms: State of the Science and Research Needs (pp 539-557). New York: Springer.
Sivonen, K. and G. Jones 1999. Cyanobacterial toxins. In I. Chorus and J. Bartram, (Eds.), Toxic cyanobacteria in water: A guide to their public health consequences, monitoring and management (pp 41-111). London: E & FN Spon.
Smeltzer, E., A. Shambaugh and P. Stangel 2012. Environmental change in Lake Champlain revealed by long-term monitoring. Journal of Great Lakes Research 38, Supplement 1, 6-18.
Smith, C. J. and A. M. Osborn 2009. Advantages and limitations of quantitative PCR (Q-PCR)-based approaches in microbial ecology. FEMS Microbiology Ecology 67, 6-20.
Smith, V. H. 1982. The nitrogen and phosphorus dependance of algal biomass in lakes: An empirical and theoretical analysis. Limnology and Oceanography 27, 1101-1112.
Smith, V. H. 1983. Low nitrogen to phosphorus ratios favor dominance by blue-green algae in lake phytoplankton. Science 221, 669-671.
Smith, V. H., V. J. Bierman, B. L. Jones and K. E. Havens 1995. Historical trends in the Lake Okeechobee ecosystem IV. Nitrogen:phosphorus ratios, cyanobacterial dominance, and nitrogen fixation potential. Archiv für Hydrobiologie 107, 71-88.
Smith, V. H., G. D. Tilman and J. C. Nekola 1999. Eutrophication: Impacts of excess nutrient inputs on freshwater, marine, and terrestrial ecosystems. Environmental Pollution 100, 179-196.
Soares, M. C. S., M. M. Marinho, S. M. O. F. Azevedo, C. W. C. Branco and V. L. M. Huszar 2012. Eutrophication and retention time affecting spatial heterogeneity in a tropical reservoir. Limnologica - Ecology and Management of Inland Waters 42, 197-203.
Song, Y. and B. Qiu 2007. The CO2-concentrating mechanism in the bloom-forming cyanobacterium *Microcystis aeruginosa* (*Cyanophyceae*) and effects of UVB radiation on its operation. Journal of Phycology 43, 957-964.
Stanić, D., S. Oehrle, M. Gantar and L. L. Richardson 2011. Microcystin production and ecological physiology of Caribbean black band disease cyanobacteria. Environmental Microbiology 13, 900-910.
Summons, R. E., L. L. Jahnke, J. M. Hope and J. A. Logan 1999. 2-Methylhopanoids as biomarkers for cyanobacterial oxygenic photosynthesis. Nature, 554-557.
Takahashi, K., T. Yoshioka, E. Wada and M. Sakamoto 1990. Temporal variations in carbon isotope ratio of phytoplankton in a eutrophic lake. Journal of Plankton Research 12, 799-808.

Talling, J. F. 1976. The depletion of carbon dioxide from lake water by phytoplankton. Journal of Ecology 64, 79-121.
Tamaru, Y., Y. Takani, T. Yoshida and T. Sakamoto 2005. Crucial role of extracellular polysaccharides in desiccation and freezing tolerance in the terrestrial cyanobacterium *Nostoc commune*. Applied and Environmental Microbiology 71, 7327-7333.
Taton, A., E. Lis, D. M. Adin, G. Dong, S. Cookson, S. A. Kay, S. S. Golden and J. W. Golden 2012. Gene transfer in *Leptolyngbya* sp. strain BL0902, a cyanobacterium suitable for production of biomass and bioproducts. PLoS ONE 7, e30901.
Te, S. H. and K. Y.-H. Gin 2011. The dynamics of cyanobacteria and microcystin production in a tropical reservoir of Singapore. Harmful Algae 10, 319-329.
Teixeira, M. G., M. C. Costa, V. de Carvalho, M. S. Pereira and E. Hage 1993. Gastroenteritis epidemic in the area of the Itaparica Dam, Bahia, Brazil. Bulletin of the Pan American Health Organization 27, 244–253.
Testa, J. M. and W. M. Kemp 2012. Hypoxia-induced shifts in nitrogen and phosphorus cycling in Chesapeake Bay. Limnology and Oceanography 57, 835-850.
Thomazeau, S., A. Houdan-Fourmont, A. Couté, C. Duval, A. Couloux, F. Rousseau and C. Bernard 2010. The contribution of sub-Saharan African strains to the phylogeny of cyanobacteria: Focusing on the *Nostocaceae* (*Nostocales*, cyanobacteria). Journal of Phycology 46, 564-579.
Tillett, D., E. Dittmann, M. Erhard, H. von Döhren, T. Börner and B. A. Neilan 2000. Structural organization of microcystin biosynthesis in *Microcystis aeruginosa* PCC7806: An integrated peptide–polyketide synthetase system. Chemistry & Biology 7, 753-764.
Tilman, D., S. S. Kilham and P. Kilham 1982. Phytoplankton community ecology: The role of limiting nutrients. Annual Review of Ecology, Evolution, and Systematics 13, 349-372.
Tooming-Klunderud, A., B. Mikalsen, T. Kristensen and K. S. Jakobsen 2008. The mosaic structure of the *mcyABC* operon in *Microcystis*. Microbiology 154, 1886-1899.
Troy, A., D. Wang and D. Capen 2007. Updating the Lake Champlain basin land use data to improve prediction of phosphorus loading. In D. Wang (Ed.), Report to the Lake Champlain Basin Program (pp 1-53). Grand Isle, Vermont: LCBP.
Tsuji, K., S. Naito, F. Kondo, N. Ishikawa, M. F. Watanabe, M. Suzuki and K.-i. Harada 1994. Stability of microcystins from cyanobacteria: Effect of light on decomposition and isomerization. Environmental Science & Technology 28, 173-177.
Tsuji, K., T. Watanuki, F. Kondo, M. F. Watanabe, S. Suzuki, H. Nakazawa, M. Suzuki, H. Uchida and K.-I. Harada 1995. Stability of microcystins from cyanobacteria—II. Effect of UV light on decomposition and isomerization. Toxicon 33, 1619-1631.
Ueno, Y., S. Nagata, T. Tsutsumi, A. Hasegawa, M. F. Watanabe, H.-D. Park, G.-C. Chen, G. Chen and S.-Z. Yu 1996. Detection of microcystins, a blue-green algal hepatotoxin, in drinking water sampled in Haimen and Fusui, endemic areas of primary liver cancer in China, by highly sensitive immunoassay. Carcinogenesis 17, 1317-1321.
USEPA 1997. Deposition of air pollutants to the Great Waters, second Report to Congress. Prepared for the office of Air Quality Planning and Standards, U.S. EPA. Research Triangle, NC. EPA-453/R-97-011.
USEPA 2007. National water quality inventory: Report to congress, United States Environmental Protection Agency, Washington, DC. EPA-841-R-07-001.

Utermöhl, H. 1958. Zur Vervolkommnung der quantitativen phytoplankton- Methodik. Mitt. Internat. Ver. Für Theoret. Angew. Limnologie 9, 1-38.
Valiere, N. and P. Taberlet 2000. Urine collected in the field as a source of DNA for species and individual identification. Molecular Ecology 9, 2150-2152.
Vanderploeg, H. A., J. R. Liebig, W. W. Carmichael, M. A. Agy, T. H. Johengen, G. L. Fahnenstiel and T. F. Nalepa 2001. Zebra mussel (*Dreissena polymorpha*) selective filtration promoted toxic *Microcystis* blooms in Saginaw Bay (Lake Huron) and Lake Erie. Canadian Journal of Fisheries and Aquatic Sciences 58, 1208-1221.
Vasconcelos, V. 2006. Eutrophication, toxic cyanobacteria and cyanotoxins: when ecosystems cry for help. Limnetica 25, 425-432.
Vezie, C., L. Brient, K. Sivonen, G. Bertru, J. C. Lefeuvre and M. Salkinoja-Salonen 1998. Variation of microcystin content of cyanobacterial blooms and isolated strains in Lake Grand-Lieu (France). Microbial Ecology 35, 126-135.
Vézie, C., J. Rapala, J. Vaitomaa, J. Seitsonen and K. Sivonen 2002. Effect of nitrogen and phosphorus on growth of toxic and nontoxic *Microcystis* strains and on intracellular microcystin concentrations. Microbial Ecology 43, 443-454.
Vuong, N.-M., R. Villemur, P. Payment, R. Brousseau, E. Topp and L. Masson 2013. Fecal source tracking in water using a mitochondrial DNA microarray. Water Research 47, 16-30.
Welker, M. and H. Von Döhren 2006. Cyanobacterial peptides – Nature's own combinatorial biosynthesis. FEMS Microbiology Reviews 30, 530-563.
Werner, V. R., H. D. Laughinghouse, M. F. Fiore, C. L. Sant'Anna, C. Hoff, K. R. de Souza Santos, E. B. Neuhaus, R. J. R. Molica, R. Y. Honda and R. O. Echenique 2012. Morphological and molecular studies of *Sphaerospermopsis torques-reginae* (Cyanobacteria, *Nostocales*) from South American water blooms. Phycologia 51, 228-238.
Whiteley, A. S., S. Jenkins, I. Waite, N. Kresoje, H. Payne, B. Mullan, R. Allcock and A. O'Donnell 2012. Microbial 16S rRNA Ion Tag and community metagenome sequencing using the Ion Torrent (PGM) Platform. Journal of Microbiological Methods 91, 80-88.
Whitman, R. L., M. B. Nevers, G. C. Korinek and M. N. Byappanahalli 2004. Solar and temporal effects on *Escherichia coli* concentration at a Lake Michigan swimming beach. Applied and Environmental Microbiology 70, 4276-4285.
WHO 1998. Cyanobacterial toxins: microcystin-LR. In: Guidelines for drinking water quality. 2nd Edition, Addendum to Vol. 2. Health criteria and other supporting information (pp. 95–110). Geneva: World Health Organization.
WHO 2003. Algae and cyanobacteria in fresh water. In: Guidelines for safe recreational water environments. Vol. 1: Coastal and fresh waters (pp. 136–158). Geneva: World Health Organization.
Wilhelm, S. W., G. S. Bullerjahn, M. L. Eldridge, J. M. Rinta-Kanto, L. Poorvin and R. A. Bourbonniere 2006. Seasonal hypoxia and the genetic diversity of prokaryote populations in the central basin hypolimnion of Lake Erie: Evidence for abundant cyanobacteria and photosynthesis. Journal of Great Lakes Research 32, 657-671.
Willame, R., C. Boutte, S. Grubisic, A. Wilmotte, J. Komárek and L. Hoffmann 2006. Morphological and molecular characterization of planktonic cyanobacteria from Belgium and Luxembourg. Journal of Phycology 42, 1312-1332.

Wood, S. A., D. Mountfort, A. I. Selwood, P. T. Holland, J. Puddick and S. C. Cary 2008. Widespread distribution and identification of eight novel microcystins in antarctic cyanobacterial mats. Applied and Environmental Microbiology 74, 7243-7251.

Xu, Y., G. Wang, W. Yang and R. Li 2010. Dynamics of the water bloom-forming *Microcystis* and its relationship with physicochemical factors in Lake Xuanwu (China). Environmental Science and Pollution Research 17, 1581-1590.

Yakes, F. M. and B. Van Houten 1997. Mitochondrial DNA damage is more extensive and persists longer than nuclear DNA damage in human cells following oxidative stress. Proceedings of the National Academy of Sciences 94, 514-519.

Yamamoto, Y. and F. K. Shiah 2012. Factors related to the dominance of *Cylindrospermopsis raciborskii* (cyanobacteria) in a shallow pond in northern taiwan. Journal of Phycology 48, 984-991.

Yergeau, E., J. R. Lawrence, S. Sanschagrin, M. J. Waiser, D. R. Korber and C. W. Greer 2012. Next-generation sequencing of microbial communities in the Athabasca River and its tributaries in relation to oil sands mining activities. Applied and Environmental Microbiology 78, 7626-7637.

Zhao, Y., P. Xie and H. Fan 2011. Genomic profiling of microRNAs and proteomics reveals an early molecular alteration associated with tumorigenesis induced by MC-LR in mice. Environmental Science & Technology 46, 34-41.

Zhou, L., Yu, H., Chen, K., 2002. Relationship between microcystin in drinking water and colorectal cancer. Biomedical and Environmental Sciences 15, 166–171.

Zilliges, Y., J.-C. Kehr, S. Meissner, K. Ishida, S. Mikkat, M. Hagemann, A. Kaplan, T. Börner and E. Dittmann 2011. The cyanobacterial hepatotoxin microcystin binds to proteins and increases the fitness of *Microcystis* under oxidative stress conditions. PLoS ONE 6, e17615.

Appendices

Appendix A – Primer design and validation for partial cyanobacterial 16S rRNA gene amplicon sequencing with Ion Torrent

A.1 Design and *in silico* analysis.

A primer set compatible with the Ion Torrent sequencing platform was designed that targets the cyanobacterial 16S rRNA gene. We aligned at least one representative sequence of the 139 cyanobacterial genera present on the NCBI database as of February 2012. A basic local alignment was conducted, using BLAST, to identify a conserved region shared by the sequences and the reference sequence *M. aeruginosa* NIES-843 (GenBank AP009552.1). Primer-BLAST (default parameters) was employed to find primers specific to the region identified. The chosen primer set was named UcyaF4/R6 (Table 3.2).

Primer characteristics: the PCR specifications of the primers were verified with MacVector (International Biotechnologies, Inc., New Haven, Conn). Both primers are 18 nt in length, have 50% GC content and the Tm difference of the pair is 1.1°C (UcyaF4 Tm, 50.1°C; UcyaR6 Tm, 48.9°C). None of the primers form self 3' dimers or hairpins, however F4 forms a self-duplex. The PCR product generated with the UcyaF4/R6 primer pair is 161 bp (amplification region in *M. aeruginosa* AP009552.1: 1886296 to1886456 bp). **Primer phylogenetic range:** in order to determine the coverage rates of the primers UcyaF4/R6 *in silico*, we downloaded nearly full-length 16S rRNA gene sequences from the RDP website (version 10.8) as of February 2012 (Cole et al. 2009). The RDP database sequence characteristics chosen were as follows: ≥ 1,200 nt in size, good quality, type and non-type strains, from isolates, NCBI taxonomy. The dataset included 142,154 eubacterial sequences from which 2,852 (2%) were cyanobacterial sequences. The coverage rate of each primer was measured as a percentage of the RDP database sequences in which the primer sequence was found. This analysis was performed using the programs Primrose and OligoCheck (Ashelford et al. 2002) allowing 0, 1 and 2 mismatches, at the most, between the primer and the database sequences. UcyaF4/R6 coverage rate of the eubacteria and cyanobacteria databases was 5.7% and 96.4%, respectively, when no mismatches were allowed. The primer coverage rate increased to 63.9% and 99.3% of the eubacteria and cyanobacteria database, respectively, when 2

mismatches were allowed. This *in silico* analysis revealed that the primer set UcyaF4/R6 provides good coverage of cyanobacteria sequences in the RDP database, but that it is not entirely specific to cyanobacteria. Nevertheless, this primer set is suitable for the sequencing of cyanobacteria employing the Ion Torrent platform, since undesired sequences can be filtered out and removed from subsequent analyses. The phylogenetic range of the primers was also verified *in silico* with a primer-BLAST search. Of the 139 cyanobacterial genera present in the NCBI database as of February 2012, 7 genera were not amplified with this primer set. However, these 7 genera are not well represented in the NCBI database, and the sequences that are present are not full length16S rRNA genes.

PCR fragment identification resolution: In order to verify *in silico* whether the PCR fragment would allow differentiation among cyanobacteria, a pairwise similarity matrix was calculated using MacVector ver. 10.6 (MacVector Inc., Cary, NC, USA). The alignment consisted of the UcyaF4/R6 amplification regions of at least one representative of each of the cyanobacterial genera amplified based on the primer-BLAST analysis mentioned previously. A total of 140 cyanobacterial sequences were used in the alignment. The pairwise similarity matrix revealed that a cutoff equal to or higher than 96.99% identity would allow differentiation between cyanobacterial orders. However, within the following orders, a cutoff of 99% still did not allow differentiation between certain genera. For example, in the order *Nostocales*, it was not possible to distinguish between *Cylindrospermum*/*Rivularia*, *Raphidiopsis/Cylindrospermopsis*, or *Aphamizomenon/Cuspidothrix/Anabaena*, or *Tolypothrix/Coloedesmium*; in the order *Oscillatoriales* between *Leptolyngbya/ Microcoleus* or between *Planktothricoide/ Tychonema*. Finally, it was not possible to distinguish between *Symphionemyosis/ Iphinoe* in the order *Stignometales*.

A.2 Analysis of the performance of the UcyaF4/R6 primer set designed in this study

From the 57 samples sequenced in this study, a total of 306,380 sequences met the quality criteria established (i.e. all sequences that contained undetermined bases (*N*), had an average expected quality score lower than 17, or were shorter than 150 bp were removed from further analysis). There was an average depth of coverage of 5,375 sequences per sample (min: 1,375 sequences per sample; max: 8,012 sequences per sample). From the total amount of good quality sequences, a total of 304,781 were

assigned to a taxon by the RDP classifier. From the classified sequences, 69,637 (22.84%) were assigned to the *Cyanobacteria*, while the remaining sequences were classified into other eubacterial phyla or plastids; no *Archaea* were amplified. From the cyanobacterial sequences, 68,887 (98.92%) were classified to the order level and 62,363 (89.55%) to the genus level by MEGAN4 analysis. The average number of cyanobacterial sequences per sample was 1,208. However, as expected, the number of cyanobacterial sequences per sample varied depending on the station and the season of the year. In general, all the stations, except PR, gave a substantial number of sequences; the lowest per sample was 47, however the majority of the samples had at least 100 sequences, with some samples as high as 6,376 cyanobacterial sequences during the summer. On the other hand, PR, probably due to its hydrological regime, had a low number of cyanobacterial sequences throughout the year; the highest number of cyanobacterial sequences per sample was 89 while the average was 26. In the sequencing output all the cyanobacterial orders were observed (*Chroococcales, Nostocales, Oscillatoriales, Pleurocapsales, Prochlorales, Stigonematales*) except for *Gloeobacterales*. As expected, most of the sequences that were not classified to the genus level were from the orders *Nostocales* and *Oscillatoriales*. The genus *Microcystis* was successfully differentiated from the genus *Synechocystis*.

A.3 Analysis and implications of the parameters derived from sequencing employing UcyaF4/R6 primers

The *Microcystis* 16S rRNA gene copies/ml qPCR values were combined with the *Microcystis* percentage in the sample, as determined by UcyaF4/R6 total *Cyanobacteria* Ion Torrent sequencing results, in order to estimate relative numbers of 16S rRNA gene copies of total cyanobacteria. There was a slight variation in estimates, depending on whether the total number of cyanobacterial sequences was determined from sequences classified to the order or genus level (Fig A.1).

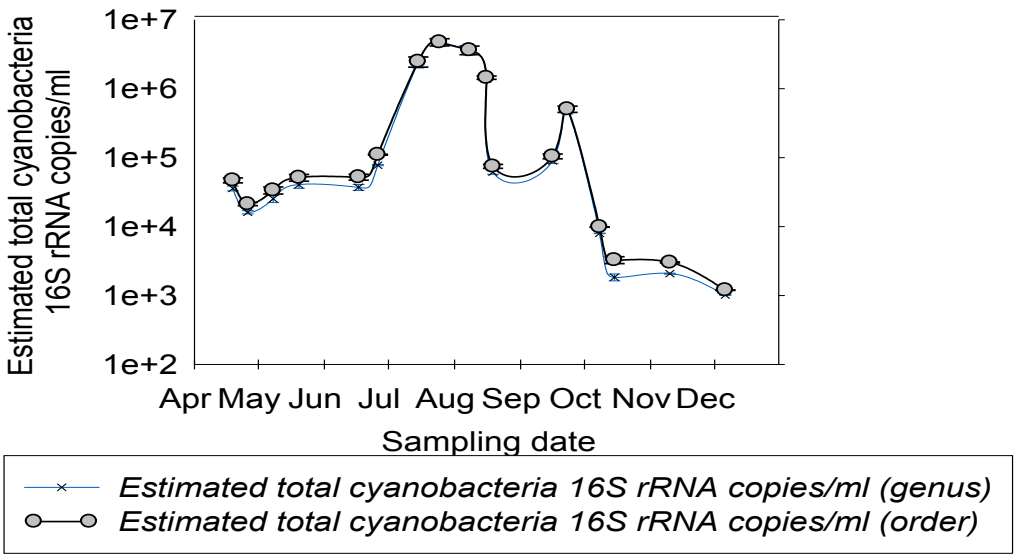

Figure A.1 – Comparison of the estimated total cyanobacterial 16S rRNA copies/ml determined based on order and genus. The error bars indicate the standard deviations (n=3).

Cyanobacteria have been traditionally classified based on morphological traits, however, in some cases, this classification is not supported based on analysis of the 16S rRNA gene. As a result, genetic relationships do not always correspond to classifications that were obtained from morphological data (Lyra et al. 2001; Willame et al. 2006; Thomazeau et al. 2010). Because of this reason, in addition with the fact that the UcyaF4/R6 primers amplify a partial region of the cyanobacterial 16S rRNA gene, not all sequences classified at the order level are successfully assigned at the genus level, especially those from the orders *Nostocales* and *Oscillatoriales* as noted in section A.1. Hence an estimate of the total cyanobacterial sequences based on the sequences classified up to the genus level neglects a portion of the total cyanobacterial community, and in this way may lead to a biased estimate of the percentages of cyanobacterial genera when unidentifiable cyanobacterial species are present (Fig. A.2). Therefore, it was established that a genus relative abundance should be determined out of the total amount of sequences identified to the order level instead of the total amount of sequences identified to the genus level. There was a significant positive correlation between our estimated total cyanobacterial 16S rRNA gene copies/ml and total cyanobacterial cells/ml (microscopy data) as suggested by Spearman correlation ($sr=0.52$, $n=31$, $p<0.01$).

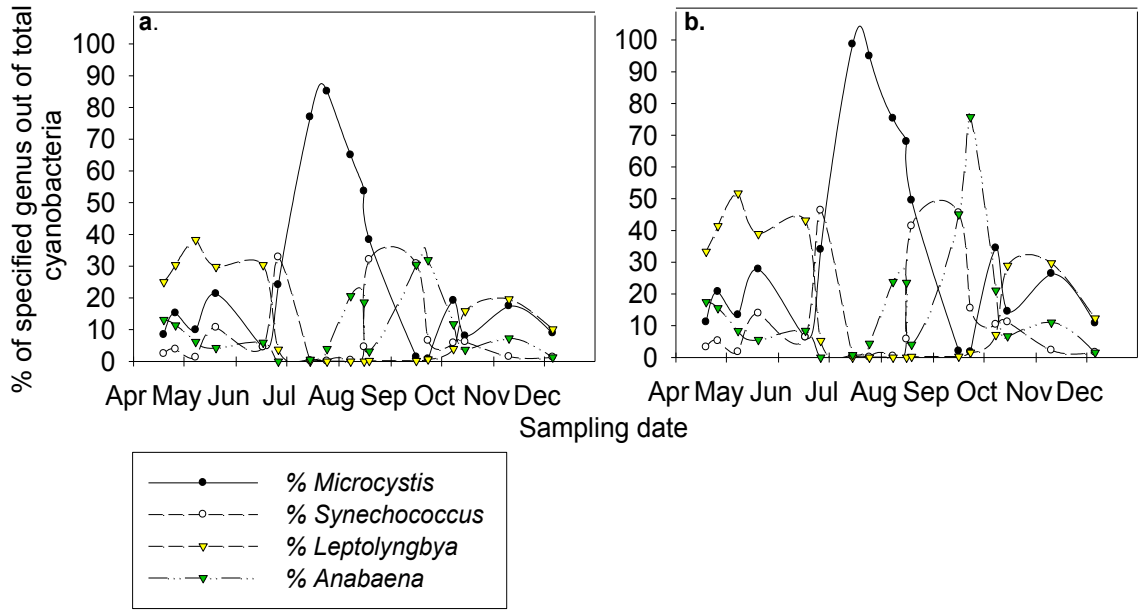

Figure A.2 –Comparison of the percentages of specified genera out of total cyanobacteria determined based on order (a) and genus (b) in sampling station BM5. See text for details.

A.4 Future optimization of sequencing employing UcyaF4/R6 primers

Based on *in silico* analysis, the primer set UcyaF4/R6 amplified 96.4% of the cyanobacterial database with zero mismatches and 99.3% of this database when 2 mismatches were allowed, hence a very low annealing temperature (53°C) was chosen at first in order to minimize PCR amplification bias that would result in preferential amplification of certain cyanobacterial genera, in this way impacting negatively the cyanobacterial profile. Nonetheless, for future experiments employing the primer set UcyaF4/R6, it is recommended to make a moderate increase in the annealing temperature during PCR, and in this way increase the percentage of cyanobacterial sequences out the total number of sequences amplified without compromising the final cyanobacterial profile. In PCR optimization experiments where the annealing temperature varied (appendix B.3), it was noted that even eliminating the annealing step still allows the amplification, though in low numbers, of sequences that have up to 4 mismatches with the primer set UcyaF3/R3. Hence, a moderate increase from 53°C to 56°C in the annealing temperature of PCR amplifications with the primer set UcyaF4/R6 may not substantially bias the cyanobacterial profile, while it may increase the overall number of cyanobacterial sequences amplified.

Appendix B – Primer design and validation for partial *Microcystis* 16S rRNA gene amplification with qPCR

B.1 Design and *in silico* analysis.

The primers were designed following the same procedure mentioned in appendix A. The primer set was named UcyaF3/R3 (Table 3.2).

Primer characteristics: UcyaF3 primer is 24 nt in length, has 54.2% GC content and a Tm of 62.1°C. UcyaR3 is 23 nt in length, has 52.2% GC content and a Tm of 62.6°C. Neither of the primers forms self 3' dimers, hairpins or self-duplex. The size of the UcyaF3 primer pair is 203 bp (amplification region in *M. aeruginosa* AP009552.1: 1886284 to1886486 bp). **Primer phylogenetic range:** The genus *Microcystis,* represented in the databases by 169 sequences, accounted for 0.1% and 5.9% of the sequences found in our eubacterial and cyanobacterial databases, respectively. When no mismatches were allowed, the UcyaF3/R3 primer set coverage rate of the eubacteria and cyanobacteria databases was 0.1% and 5.8%, respectively. Hence, the UcyaF3/R3 primers are highly specific for *Microcystis* when no mismatches are allowed, amplifying 96% of all *Microcystis* sequences in our database and only one non-target sequence in the genus *Synechocystis* (FM177504.1). When 2 mismatches were allowed, the primer set coverage rate increased to 0.8% and 41.8% of the eubacteria and cyanobacteria databases, respectively. However, when mismatches are allowed, these primers can amplify sequences of members from other cyanobacterial genera, other phylogenetic groups of bacteria, and eukaryotic algae. For these reasons, and taking into consideration that we would be using a double-stranded DNA dye for qPCR, hence requiring excellent target specificity, the amplification of *Microcystis* by UcyaF3/R3 qPCR reactions took place under very stringent annealing conditions and further validation of the amplification products of these primers was carried out via Ion Torrent sequencing. **PCR fragment identification resolution:** A pairwise similarity matrix suggested that the primer set amplifies a region of the 16S rRNA gene that allows identification of *Microcystis* up to the genus level. In addition, a cutoff equal to or higher than 97.99% allows distinguishing between *Microcystis* and *Sinechocystis* FM177504.1.

B.2 UcyaF3/R3 primer validation and verification of sequences amplified in this study

All the samples that were amplified via UcyaF3/R3 *Microcystis* 16S rRNA gene qPCR were sent for high-throughput sequencing, as described in section 3.3.2 but using the template specific primers UcyaF3/R3, in order to verify the phylogenetic composition of the amplified products, and correct when needed the UcyaF3/R3 *Microcystis* 16S rRNA gene qPCR output based on the percentage of *Microcystis* sequences amplified out of the total bacterial amplified. For MEGAN analysis, the procedure followed was the one described in section 3.3.2 with the following exceptions: The resulting sequencing dataset from the RDP pyrosequencing pipeline was aligned, using the program BLASTN, against the eubacterial database downloaded from the RDP database (142,154 eubacterial sequences, as of February 2012, with the following characteristics: ≥ 1,200 nt in size, good quality, type and non-type strains, from isolates). The BLAST files were imported into MEGAN4 to assign the reads of each dataset to appropriate taxa in the NCBI taxonomy using the lowest common ancestor algorithm (LCA) (Huson et al. 2011) with the fallowing parameters: maximum number of matches per read, 100; Minimum Support, 1; Minimum Score, 50 (this score threshold removed hits with an E value higher than 10^{-5}); Top Percentage, 0.5 (this parameter allowed to select matches with the best score and in this way help to discriminate between *Synechocystis* and *Microcystis* genera*)*; Minimum complexity filter, 0.3. The percentage *Microcystis* was determined out of the complete set of sequences that were successfully assigned to the genus level.

From the 57 samples sequenced for this study, a total of 628,214 sequences amplified met the quality criteria established in section A.2. Based on the RDP classifier, the sequences amplified were assigned mainly to the *Cyanobacteria* phylum, the rest were assigned to plastids or other bacterial phyla such as *Firmicutes*, *Actinobacteria* and *Proteobacteria*. No *Archaea* were amplified. MEGAN4 analysis, of the blast files resulting from the BLASTN search against the eubacterial database, revealed that from the total number of sequences amplified that met the quality criteria, 512,370 (81.6%) sequences were successfully assigned to the genus level, and of these, 336,448 (66.1%) were assigned to the genus *Microcystis*. In the sequencing analysis, the genus *Microcystis* was successfully differentiated from the genus *Synechocystis,* which allowed correcting the *Microcystis* qPCR output. It was observed that the majority of the samples that had low percentages of *Microcystis* were either spring samples or PR samples, which suggests

that in cases where the target template (*Microcystis*) is in low abundance, the UcyaF3/R3 primers, even at the stringent PCR annealing temperature of 60°C, still amplify considerable percentages of unintended targets. One of the most common unintended targets amplified in spring samples was *Phalacroma mitra* (AB199885.1), whose DNA sequence has 1 and 2 mismatches with the forward and reverse primers, respectively. In view of these results, all the UcyaF3/R3 qPCR data was corrected as mentioned in section 3.3.1.2 using the Ion Torrent sequencing output.

B.3 Further optimization of PCR amplification with UcyaF3/R3 primer set

In order to evaluate whether the primer specificity for *Microcystis* could be increased, ten samples were amplified at even more stringent PCR conditions, via a 2-step PCR, which skips the annealing step. Then the amplification products were sequenced as described in section B.2. Six of these samples were Lake Champlain samples that had previously been amplified with an annealing temperature of 60 °C, allowing comparison between the two protocols. The four other samples came from other lakes. To verify the variation within a sequencing run, the sample Aug 8/09 BM3 was amplified in duplicate via 2-step PCR and included in the same sequencing run. The percentage of *Microcystis* sequences out of the total bacterial amplified was determined as described in section B.2.

Eliminating the annealing step during PCR amplification successfully improved the specificity of UcyaF3/R3 for its intended target (Table B.1).

High specificity for *Microcystis* was achieved even in samples where *Microcystis* was not abundant. For example, with the 2-step PCR protocol, the percentage of *Microcystis* amplified increased by 92% in the sample Apr 19/09 BM5 and by 90% in the sample Apr 26/09 BM5. In addition, despite the fact that in the Lake Vert sample (where no *Microcystis* sequences were found based on the sequencing output) the primers did amplify unintended targets, the total number of sequences amplified was 158, which is low compared with the 6,278 sequences amplified from Apr 19/09 BM5, a sample in which Microcystis abundance was relatively low compared to summer samples. Most of the sequences amplified from L. Vert were identified as *Planktothrix agardhii* (FJ159128), whose DNA sequence has 1 and 3 mismatches with the forward and reverse UcyaF3/R3 primers, respectively.

Table B.1 – Effect of annealing temperature on the specificity of the primers UcyaF3/R3 for its intended target, *Microcystis*.
Comparison of percentages of *Microcystis* amplified per sample.

	Annealing Temperature: 60°C			**2-Step PCR (no annealing temp)**		
Sample amplified	Total # seq.	*Microcystis* # seq.	***Microcystis* (%)**	Total # seq.	*Microcystis* # seq.	***Microcystis* (%)**
Apr 19/09 BM5	3,070	223	**7.3%**	6,278	6,230	**99.2%**
Apr 26/09 BM5	6,225	598	**9.6%**	9,184	9,144	**99.6%**
May 20/09 BM5	1,012	6,664	**15.2%**	9,338	9,257	**99.1%**
Sep 16/09 BM5	5,455	7,650	**71.3%**	6,782	6,775	**99.9%**
Aug 8/09 BM3 (technical replicate)	9,960	11,420	**87.2%**	7,110	7,102	**99.9%**
Aug 8/09 BM3 (technical replicate)	NA			5,230	5,230	**100%**
July 23/08 L. Vert	NA			158	0	**0%**
Aug 4/09 L. Choiniere	NA			8,762	8,301	**94.7%**
Aug 11/09 L. Choiniere	NA			6,108	6,060	**99.2%**
Sep 15/09 L. Choiniere	NA			7,583	7,496	**98.9%**

Seq: sequences.
NA: not available

High percentages of *Microcystis* sequences were also obtained from Lake Choiniere samples. The lowest percentage observed (94.7%) was from the Aug 4/09 sample. In this sample, most of the remaining 5.3% sequences were mainly identified as *Woronichinia naegeliana* (AJ781043), whose DNA sequence has 2 and 1 mismatch with the forward and reverse primers, respectively. Almost no variation in the percentage of *Microcystis* amplified during 2-step PCR (0.1%) was found (compare Aug 8 BM3 technical duplicates, Table B.1).

In conclusion, the primer set UcyaF3/R3 has a high specificity for *Microcystis* when employing 2-step PCR, otherwise, even high annealing temperatures, such as 60°C, during PCR, still allows abundant amplification of unintended templates, which necessitates secondary sequencing analysis using the same primers and PCR conditions in order to correct the qPCR data. To avoid additional analyses in the future, it is recommended to implement a 2-step real-time PCR.

Appendix C – Target animals for mitochondrial DNA sequencing analysis

In order to verify whether the shorter region amplified within the 2kb fragment of mtDNA still allowed the identification of the target animals, a mixture of genomic DNA from the 28 target animals was amplified using the MI50F_Ion/PyrR1 primer set (table 3.2) and the 205 bp fragment was sent for sequencing using the Ion Torrent platform. From the 28 initial target animals, only 2 animals were not identified (Table C.1, in red).

Table C.1 – Target animals for mitochondrial DNA analysis
In red: target animals that were not identified after sequencing analysis of a mixture of genomic DNA comprised of the 28 target animals in the list.

Scientific name (genus species)	Common name
Alces alces	Moose
Anas platyrhynchos	Mallard duck
Bos taurus	Cow
Branta canadensis	Canada goose
Canis latrans	Coyote
Canis lupus familiaris	Dog
Capra hircus	Goat
Castor canadensis	Beaver
Cervus elaphus	Red Deer
Columba livia	Pigeon
Coturnix coturnix	Quail
Dromaius novaehollandiae	Emu
Equus caballus	Horse
Felis catus	Cat
Gallus gallus	Poultry
Homo sapiens	Human
Larus spp.	Gull
Meleagris gallopavo	Turkey
Mus musculus	Mouse
Odocoileus virginianus	White-tailed deer
Ondatra zibethicus	Muskrat
Oryctolagus cuniculus	Rabbit
Ovis aries	Sheep
Phalacrocorax spp.	Cormorant
Procyon lotor	Raccoon
Rangifer tarandus	Caribou
Struthio camelus	Ostrich
Sus scrofa	Pig