Integrative microbial contamination assessment for water quality monitoring in the Great Lakes

INTEGRATIVE MICROBIAL CONTAMINATION ASSESSMENT FOR WATER QUALITY MONITORING IN THE GREAT LAKES

By

WENJIE ZHENG, B.SC (HONOURS)

A Thesis
Submitted to the School of Graduate Studies
in Partial Fulfilment of the Requirements
for the Degree
Master of Science

McMaster University

McMaster University MASTER OF SCIENCE (2012) Hamilton, Ontario (Biology)

TITLE: Integrative microbial contamination assessment for water quality monitoring in Great Lakes
AUTHOR: Wenjie Zheng, B. Sc (Honours) (Dalhousie University, Agriculture campus)
SUPERVISOR: Professor H.E. Schellhorn
NUMBER OF
PAGES: x, 94

Abstract

Recreational beaches are important local resources for attracting tourists. It is critical to keep tracking recreational water quality to prevent public health issues. Waterborne pathogens are one of the main elements that could cause recreational water related diseases. Fecal pollution is the primary source of waterborne pathogens. Therefore, it is important to quantify the amount of fecal pollution indicators that are present in the water, particular the human fecal indicator. The primary objective of this thesis is to develop an integrative microbial quality monitoring system to better understand water quality. The first part of this thesis examined the presence of a general fecal pollution indicator (*E. coli*) and a human fecal pollution indicator (human-specific *Bacteroidales*). The correlations between pollution sources and beach water quality were also studied to identify the impact of pollution sources. The results revealed the highly localized correlations at individual beaches depended on the impact from pollution sources. The weak correlations suggested some previous assumed pollution sources may only weakly impacted beach water quality.

Because *E. coli* strains differ enormously in pathogenic potential, it is possible that environmental *E. coli* have different genetic compositions and differential gene expression in genes such as the global stress regulator *rpoD* and *rpoS*. Thus, the second part of this thesis examined genetic composition and gene expression in *E. coli* environmental strains to study how global gene expression is altered in the natural environment. The results revealed differential RpoS expression levels in environmental *E. coli* strains, suggesting that genes regulated by *rpoD* and *rpoS* may have differential

expression levels in environmental strains, compared to commonly studied laboratory strains.

Acknowledgement:

I would like to thank my supervisor, Dr. Herb E. Schellhorn for this great opportunity to work in this wonderful lab and for his supervision and help. I also thank my co-supervisor Dr. Tom. A Edge for his supervision during sampling seasons at Canada Centre for Inland Water.

Special thanks to my wonderful colleagues, Steve Botts, Sarah Chiang, Tao Dong, Hamoody Howard-Azzeh, Lingzi Kong, Andrew O'Brien, Sharmila Sathiasothy, Mohiuddin Shohag, Myra Sohail, Vito Vijayakumaran, Jas Wasniewski, Shirley Wong and Rosemary Yu. Thank you all for those great suggestions and the funny moments. I enjoyed coffee and lunch breaks with every single one of you.

Table of Contents

List of Figures

List of Tables

Part I: Water quality monitoring in Niagara Region

Chapter 1. Introduction

1.1 Microbial source tracking

McMaster University, Environment Canada and the Region of Niagara collaborated to monitor water quality at selected beaches in Niagara Region to prevent public health issue. The Region of Niagara includes a total of 38 beaches, most of which function as recreational beaches for the public. Since the number of summer tourists increases dramatically in a short period of time, it is important to identify and qualify microbial contamination in the water, especially human health related microbes. Compared to other water sources (e.g., drinking water), recreational water has a relatively higher level of microbial contamination (Safe Drinking Water Act, 2002). The water quality objective in Ontario is 100 Colony Forming Units (CFU) *E. coli* per 100mL recreational water (Ministry of Environment and Energy, 1994). Canada has a guideline of 200 CFU/100mL (Health Canada, 2012). When *E. coli* counts exceed the threshold, the beach is considered unsafe and it is closed (posted) until *E. coli* counts drop below the regulatory limit.

One of the new water quality monitoring protocols is microbial source tracking methodology. Microbial source tracking approaches compare the differences of indicator organisms in pollution source and targeted water bodies to determine the impact from the pollution source to the water body (NWRI Scientific Assessment Report, 2006). It

includes two types of tracking approaches: library-dependent methods and library-independent methods.

Library-dependent methodology requires the selection of one or more microbial water quality indicator(s) for different types of water bodies, such as *E. coli* for fresh water and *Enterococcus* for marine water (Health Canada, 2012). Once the microbial indicator is selected, a large collection of water and fecal samples from interested sampling bodies and nearby fecal pollution sources will be required to create a collection of microbial water quality indicator organism. After all the samples are processed, the similarity of the microbial water quality indicator organism from interested sampling sites and nearby pollution sources can be determined by DNA fingerprinting and other techniques (Carlos et al., 2012). Therefore, it was chosen as one methodology of this water quality monitoring project.

However, *E. coli* enumeration protocol takes more than 24 hours. This results in the delay between water sample collection and the actual decision acts (opening or closure) on a beach. For this reason, it is realized that it may not be a perfect approach to protect public health. It is also important to identify the specific pathogen which may lead to a future disease outbreak, especially if there are human fecal pollution sources nearby. In addition, it can take many weeks to complete library-dependent microbial source tracking similarity comparison. Therefore, another microbial source tracking approach, library-independent microbial source tracking method was also employed to determine the specific pollution source at Niagara beaches.

Compared to library-dependent methods, library-independent microbial source tracking methods search for the presence of a host specific microorganism to identify the pollution source, rather than build up a reference library of the microbial water quality indicator. The host specific microorganisms are a group of microorganisms that live in gastro-intestinal tracts naturally. They also have a very important common characteristic, as they should only be present in the specific gastro-intestinal tracts of their host. Therefore, once the presence of the particular host specific microorganism is confirmed in a fecal sample, it represents the presence of the fecal pollution from its host as well. It is usually achieved by detecting the presence of a unique DNA fragment from the host specific microorganism. On that account, library-independent methods require less laboratory work but one the other hand they require the unique selection of host specific microorganism and the development of unique probes. For example, the approach of human-specific *Bacteroidales* detection is used to detect the presence of human fecal contaminant (Ahmed et al., 2012).

Consequently, *E. coli* enumeration and human-specific *Bacteroidales* detection were selected as two steps in this water quality monitoring project. A total of 15 beaches from the Niagara Region were monitored during the summer of 2010, including Long Beach Conservation West (LCW), Long Beach Conservation East (LCE), Long Beach (LB), Nickel Beach (NB), Larraine Road beach (LR), Humberstone Centennial Beach (HC), Bay Crystal beach (BC) and Bernard beach (BE) at Lake Erie; and Queen's Royal beach (QR), Garden City beach (GC), Lakeside beach (LS), Charles Daley beach East (CDE), Charles Daley beach West (CDW), Nelles Park beach (NP) and Fifty Point beach

(FP) at Lake Ontario (Table 1). After sample processing and statistical analysis, beaches were sorted into four groups based on their potential risk levels (Table 4). In 2011, focus was moved to those beaches that showed a higher level of public health risk. Therefore, five out of six beaches from the highest risk class (including QR, GC, LCE, LS and LR) were studied in 2011. Another three beaches (FP, BC and LB) from each other risk classes were also studied as reference beaches.

1.2 *E. coli* enumeration

Recreational water illnesses (RWIs) are caused by waterborne pathogens and chemical toxins (Hlavsa et al., 2011). Waterborne pathogens are the main elements that can cause recreational water related diseases. Fecal pollution is the primary source of waterborne pathogens (Dwight et al., 2005). Therefore, it is very important to quantify the amount of fecal pollution indicators that are present in the water. *E. coli* are been used as one of microbial indicators for fecal contamination in recreational waters because *E. coli* concentration is positively associated with fecal contamination (National Academy of Sciences, 1977). *E. coli* are gram-negative bacteria that are normally found in warm-blooded organisms. They also exist in fecal polluted waters with a large population (Ferguson et al., 2012). *E. coli* is related with swimming associated gastro-intestinal disease (Masters et al., 2011), thus the Canadian federal recreational water safety guideline is set to 200 *E. coli* per 100 mL of recreational water to minimize public health issues (Health Canada, 2012). In the meantime, Ontario has a guideline of 100 *E. coli* per 100 mL, as mentioned above.

Currently, many *E. coli* enumeration methods use a β-glucuronidase assay to detect the presence of *E. coli*. β-glucuronidase is specifically produced by *E. coli* and presents in 97 per cent of *E. coli* strains (Rice et al., 1990). Substrate in the agar is cleaved by *E. coli* and returns a purple color on *E. coli* colony.

1.3 Human-specific *Bacteroidales* detection

Although *E. coli* have been used as microbial water quality indicator for over two decades, there are a few limitations with this *E. coli* enumeration approach. One is the delayed result and the other is that it is not a host specific approach. So it will be difficult to predict the water quality properly and difficult to trace back to the pollution source, especially when human health related pathogens exist. Therefore, an alternative indicator with host specificity will be helpful to monitor water quality that corresponding to human health concern.

The order of *Bacteroidales* includes four families of environmental bacteria as *Bacteroidaceae*, *Rikenellaceae*, *Porphyromonadaceae* and *Prevotellaceae* (Boone and Castenholz, 2001). The use of host-specific *Bacteroidales* to detect presence of human fecal pollution was developed a decade ago based on a *Bacteroidales* 16S rRNA marker (Bernhard and Field, 2000; Bernhard and Field, 2000). Then, this technology was expanded to other fecal source hosts, including dog, pig and cow, etc (Dick et al., 2005; Kildare et al., 2007).

For the purpose of this thesis, we are particular interested in human-specific *Bacteroidales* because it directly links to human health related disease. For human-

specific *Bacteroidales* detection, both conventional PCR and real time PCR (qPCR) assays are currently available. Conventional PCR assay was developed by Bernhard's group in 2000 (Bernhard and Field, 2000; Bernhard and Field, 2000), and it is used by different groups in the world (Ahmed et al., 2012). Real time PCR assays were developed recently and they are compatible with different florescent dyes as SYBR Green and TaqMan. The SYBR Green assay is also known as HF183 SYBR since it adopted HF183 forward primer with newly developed reverse primer (Seurinck et al., 2005). TaqMan based real time PCR assay amplifies a similar human-specific *Bacteroidales* fragment (BacHum) (Kildare et al., 2007). A recent paper reported that SYBR Green assay may have less sensitivity in human-specific *Bacteroidales* detection (5/8) than TaqMan assay (8/8), but TaqMan may have more false-positive results in domestic animal specific *Bacteroidales* detections (Van De Werfhorst et al., 2011). However, due to the small sample size (approximately 10 samples per host), more studies should be done to compare the sensitivity between these two methods.

Chapter 2. Methodology

2.1 Water sample collection

Water samples (500mL) were collected by Niagara Public Health at three sampling locations along the beach weekly and twice a week at knee (~60cm) depth and at nearby point source(s) of pollution. Water bottles were previously autoclaved and were not opened until they were in the water body to prevent contamination. Sampling students started sampling at 8:00 AM, and samples were delivered to Niagara regional Environmental Centre at Thorold, Ontario by 1:00 PM. A duplicate set of sample from transects 1, 3, 5 and point source(s) of pollution was delivered to CCIW at Burlington, Ontario by 2:00 PM.

After water samples arrived, samples were processed under sterilized laboratorial conditions. Differential Coliform (Oxoid, Catalog # 1045092) plates were prepared previously and labeled for the following filtration process. Water samples were then mixed well and serial diluted (normally 1, 10 and 100 mL) before filtered through 0.45 μm membranes (Fisherbrand, Catalog # F1BA15591). The Filter membranes were then placed on Differential Coliform Agar and incubated at 44.5 °C for 18 hours. Generally, plates containing 30-100 *E. coli* colonies were enumerated to increase accuracy (Bertke, 2007).

2.2 DNA extraction

A total of 300 mL of water from each sample was filtered through another 0.45 μm membrane. DNA was extracted from the filtered membrane using PowerSoil® DNA

Isolation Kit (Mo-Bio, Catalog # 12888-100). At the end, a total of 100 µL DNA sample was prepared and stored in Matrix* 2D Barcoded Storage Tubes (Thermo Scientific, Catalog # 3714MTX) at -80° C. By measuring the 260/280 ratio, DNA samples were assessed for protein contamination.

2.3 Human-specific *Bacteroidales* detection

2.3.1 *Conventional PCR*

Extracted DNA was used to detect the presence of human Human-specific *Bacteroidales* marker in recreational water by using conventional PCR based on microbial source tracking methodology (Edge et al., 2010; Sidhu et al., 2012). HF183F (5'-ATC ATG AGT TCA CAT GTC CG-3') was used as forward primer while BAC32R (5'-AAC GCT AGC TAC AGG CTT-3') was the forward primer to detect the presence of human-specific *Bacteroidales* (Ahmed et al., 2008). Each mixture contained 10X PCR buffer, 40 pM of each primer, 0.06% BSA, 0.8 mM of dNTP mixture, 1.25 units of HotMaster Taq, 1 µL of template and PCR grade ddH_2O to a total volume of 25 µL. The reaction was initiated at 94 °C for 2 minutes. The reaction was continued for 35 cycles of 94 °C for 20 sec, 63 °C for 10 sec, 65 °C for 50 sec and final extension at 66 °C for 7 min (Edge and Hill, unpublished data).

For the detection of universal *Bacteroidales*, BAC32R and the modified forward primer BAC708 (5'-CAA TCG GAG TTC TTC GTG-3') were used (Ahmed et al., 2008). The reaction mixture of universal *Bacteroidales* detection was the same as human specific *Bacteroidales* detection except for the different primer set. The reaction was initialized at

94 °C for 2 minutes. The reaction was continued for 35 cycles of 94 °C for 20 sec, 53 °C for 10 sec, 65 °C for 50 sec and final extension at 66 °C for 7 min (Edge and Hill, unpublished data).

After amplification, a mixture of 5 µL of amplified products and 1 µL of 6X DNA dye were mixed well before ran through a 1.25% agarose gel at 170 V for 60 minutes, along with positive and negative controls. Finished agarose gel was stained with ethidium bromide for 20 minutes and destained with ddH_2O before visualized under UV light. The visualized band was considered as a positive result for human-specific *Bacteroidales* marker detection, which indicated the presence of human fecal pollution. This conventional PCR method was validated by amplifying fecal DNA samples (Edge et al., 2010).

2.3.2 Real time PCR (qPCR) assay

Relative quantification of human-specific *Bacteroidales* (Lee et al., 2006) with SYBR Green was performed to examine the correlation between *E. coli* counts and the presence of human-specific *Bacteroidales* at Queen's Royal Beach. HF183F and newly designed reverse primer (5'-TAC CCC GCC TAC TAT CTA ATG-3') were used as the primer set (Seurinck et al., 2005). The reactions were performed in Bio-Rad CFX96 Touch™ Real-Time PCR Detection System. Each mixture contained 5.5 µL of 2X qPCR master mix, 0.25 µM of each primer, 0.22 µL of reference dye (50X), 1 µL of template and qPCR grade ddH_2O to a total volume of 11 µL. The reaction was initiated at 50 °C

for 2 minutes and 95 °C for 10 minutes. The reaction was continued by 40 cycles of 95 °C for 30 s, 53 °C for 1 min, and 60 °C for 1 min (Seurinck et al., 2005)..

2.4 Statistical analysis

After samples were processed, data were analyzed based on five questions as following:

1. Which is the most contaminated beach?

2. The exceed rate (above the provincial guideline) of each beach;

3. Are *E. coli* concentrations along transects correlate with a point source of pollution?

4. Are *E. coli* counts correlate with the presence of human sewage marker?

5. Is there any statistical difference between *E. coli* counts in 2010 and 2011?

Geometric means were calculated to determine the average *E. coli* concentration. Geometric means *n*th (n = the number of the population) root the product of the population (Microsoft Office Excel 2010) rather than simply averaging the sum of the population in arithmetic mean. In this way, the mean will be less affected by factors which have larger numeric ranges. This is especially useful for environmental samples when multiple factors with different numeric ranges are often involved in (Frank, 2009).

Spearman's rank correlation coefficient (r_s) test was employed as a tool to examine the correlation between variables with a significance level of 0.05 (JMP9, SAS). Spearman's test is a non-parametric test for testing the statistical dependence between

variables (Corder and Foreman, 2009). r_s ranges from -1 to 1. The r_s values of 1 and -1 represent the strongly positive and strongly negative correlation, respectively. A value of 0 indicates that there is no correlation between two variables. Compared to other correlation coefficient tests such as the Pearson product-moment correlation coefficient, Spearman's test is less sensitive to "outliers" since it measures correlation by numeric ranks rather than the raw numbers. This is more useful when analyzing environmental samples, when an "outlier" may be the unusual, but correct *E. coli* concentration due to a special environmental event such as rainfall.

Chapter 3. Results and discussion

3.1 *E. coli* contamination averages among beaches

3.1.1 Year 2010

E. coli enumeration

A summary of *E. coli* counts over the summer in 2010 is presented in Table 2. Geometric means in all 15 beaches did not exceed the Ontario provincial guideline, which is 100 CFU/100mL of recreational water. The highest *E. coli* concentration was found at Queen's Royal beach, which reached 67 CFU/100mL (geometric mean). The lowest *E. coli* concentration was found at Bernard Beach, which was 12 CFU/100mL (geometric mean). A cutline of 20% exceedance (>100CFU/100mL) was used to classify beaches to different water quality risk categories. A total of 6 beaches (Long Beach Conservation East, Nickel Beach, Charles Daly East, Bay Crystal Beach, Bernard Beach and Long Beach) were grouped into the low *E. coli* contamination risk group, while other 9 beaches were in the high *E. coli* contamination risk group (Table 4).

Human-specific *Bacteroidales* maker detection

A summary of percent of sampling days when the human-specific *Bacteroidales* DNA marker was detected at 15 beaches in 2010 is presented in Table 3. Queen's Royal had the highest percent of sampling days showed the positive result for human-specific *Bacteroidales* detection, which was 43%. No human-specific *Bacteroidales* marker was detected at Long Beach Conservation West.

Because there is no standard cutline to distinguish human fecal pollution risk at beach, 10% cutline was established during analysis in this study. Therefore, Bay Crystal Beach (10%), Charles Daly East Beach (13%), Garden City Beach (31%), Long Beach Conservation East (13%), Lorraine Road Beach (10%), Lakeside Beach (14%), Nelles Park Beach (13%) and Queen's Royal Beach were considered as beaches with high human fecal contamination risk sites. Another 7 beaches were graded as low human fecal contamination risk beaches.

Consequently, combined two grouping methods together, 15 beaches were ranked into 4 different categories, including high human fecal contamination with high *E. coli* concentration, low human fecal contamination with high *E. coli* concentration, high human fecal contamination with low *E. coli* concentration and low human fecal contamination with low *E. coli* concentration (Table 4). Queen's Royal Beach, Garden City Beach, Long Beach Conservation East, Lakeside Beach, Lorraine Road Beach and Nelles Park Beach were ranked in the highest risk class due to their high *E. coli* concentration and high human fecal contamination. Long Beach Conservation West, Nickel Beach, Bernard Beach and Long Beach were ranked as the cleanest beaches.

3.1.2 Year 2011

E. coli enumeration

A summary of *E. coli* counts over the summer in 2011 is presented in Table 5. Based on the classification system created on 2010, a beach is considered as a yearly potential health risk site if it was posted (>100 CFU/100mL) for more than 20% of the sampling time over the bathing season (Edge et al., 2010). The results from 2011 indicated that most beaches except LB had exceed the 20% cutoff, suggesting that the potential public health issue still remains at most targeted beaches. In 2010, BC did not reach the 20% exceed cutoff line, but the exceedance reached 21.6% at BC in 2011.

In terms of overall *E. coli* average in 2011, LCE and LR had the overall averages exceed the provincial guideline, while the others were within the provincial guideline. The data suggests high levels of *E. coli* contamination are still remaining in both beaches. LCE was the most contaminated beach in 2011 according to the provincial guideline. The average of *E. coli* concentration at Long Beach Conservation East reached 195.3±634.1 CFU/100mL. The large standard error of mean also suggests the big difference of *E. coli* concentration between sampling days at Long Beach Conservation East over the summer. In the meantime, Water quality at LCE also decreased in 2011 since the overall average was 43.0±52.3 CFU/100mL in 2010. Since no human fecal marker was detected at Long Beach Conservation East, it is possible that other fecal pollution sources such as bird droppings may affected beach water quality at Long Beach Conservation East (Edge and Hill, 2007). On the other hand, *E. coli* contamination was lowest at FP in 2011, with an

overall average of 24.1±57.3 CFU/100mL. Therefore, FP was ranked as the cleanest beach in 2011 by geometric mean.

Human-specific *Bacteroidales* maker detection

A summary of percent of sampling days when the human *Bacteroidales* DNA marker was detected at 7 beaches in 2011 is presented in Table 6. Queen's Royal still had the highest percent of sampling days showed a positive result for human *Bacteroidales* detection in 2011, which was 50%. Garden City Beach also remained a high level of human fecal contamination, which was 27%. No human *Bacteroidales* marker was detected at Long Beach Conservation East.

When applying the 10% human *Bacteroidales* detection cutline into consideration, Queen's Royal Beach and Garden City Beach were classified into the high human fecal contamination risk class, while Bay Crystal Beach (3%), Fifty Point Beach (3%), Long Beach (3%), Long Beach Conservation East (0%), Lorraine Road Beach (7%), Lakeside Beach (7%) grouped into the low human fecal contamination risk class.

Therefore, 8 beaches in 2011 were categorized into 4 different risk classes as 2010 (Table 7). Queen's Royal Beach, Garden City Beach, Lakeside beach, Lorraine Road Beach, Fifty Point Beach and Bay Crystal Beach were ranked into the highest risk group due to the high human fecal contamination risk and high *E. coli* contamination risk. Long Beach Conservation East stayed in low human fecal contamination but high *E. coli* concentration class. Long Beach was divided into the group of high human fecal

contamination and low *E. coli* concentration. No beach was in the low human fecal contamination and low *E. coli* concentration class.

Compared to 2010, a change of beach classifications was observed. Fifty Point Beach, Bay Crystal Beach and Long beach were shifted to the high *E. coli* contamination risk group in 2011, suggesting the increase of *E. coli* contamination risk in these beaches. Bay Crystal Beach was shifted to the high human fecal contamination class, indicating the increased impact from human activities to the beach water quality. Long Beach Conservation East was the only beach showing improvement in this water quality risk classification. It moved from the high human fecal contamination risk class to the low human fecal contamination class.

3.2 Correlations of *E. coli* concentration between beach sampling locations and pollution sources in 2011

E. coli concentrations along transects at each beach were ranked against *E. coli* concentrations at their point sources of pollution individually by Spearman's coefficient correlation test. This study examined how point source(s) of pollution correlated to water quality along transects at each beach. Correlations at one-day post rain event were calculated separately because more *E. coli* may be washed into beach water and lead to the increase of *E. coli* concentration (Edge et al., 2010).

At Fifty Point beach, significant strong correlations were observed between beach sampling locations throughout 2011, r_s= 0.81 between FP1-FP3, r_s= 0.80 between FP1-FP5 and r_s= 0.85 between FP3-FP5 (Table 1, Appendix B). The weak correlations

between FPS and beach sampling locations after rain event suggest that FPS may not specifically affect water quality at Fifty Point Beach (Figure 8, Appendix B).

At Bay Crystal beach, BC3 and BC5 were strongly correlated with BC1, with the r_s values of 0.91 and 0.82, respectively. BC5 was strongly correlated to BC3 with an r_s value of 0.83. The correlations between BCS and all transects were weak for all-summer data (Table 2, Appendix B). This suggests that BCS has little impact on beach water quality at Bay Crystal. Instead, the strong correlations in BC1-BC5 and BC3-BC5 suggesting that there may be a pollution source near BC5 which led to the high *E. coli* concentration at BC5. Interestingly, a strong correlation between BC3 and BCS was identified for rain-day data (Table 9, Appendix B). However, due to the small sample size, more studies should be done to study the impact from BCS to beach water quality after rain event.

At Garden City, GC1 was positively correlated to GC5 and GCS, and GC3 was positively correlated to GC5. This suggests that GCS is one of the targeted pollution sources, since it strongly associated with water quality at GC1 and GC3 (Table 3, Appendix B). In addition to GCS, the strong correlations between GC1 (r_s=0.67) and GC3 (r_s=0.73) to GC5 suggests that there may be another potential pollution source near GC5 or *E. coli* accumulation from water current at GC5. No post rain-event samples were collected in 2011 at Garden City. Therefore, no post rain event correlation study was performed.

Two potential sources of contamination are located on each side of the beach at Long Beach. The correlations between sampling locations were weak except for LB3-LB1. An r_s value of 0.85 indicated LB3 is strongly positively correlated to LB1 (Table 4, Appendix B). No other strong correlations were identified. This suggests that overall speaking, these two point sources of pollution may have limited impact on beach water quality. Post rain correlation study suggested the strong correlations between LBS1 and LB1, suggesting LBS1 may significantly impacted water quality at LB1 (Table 10, Appendix B). The strong correlation between LBS1 and LBS2 suggesting they may link to the same pollution source at rain days (Table 10, Appendix B).

Similar to Long Beach, transects at Long Beach Conservation East did not show strong correlations to the potential pollution source throughout summer 2011, and the only strong correlation identified was between LCE3 and LCE5 (Table 5, Appendix B). However, these two transects are covered with rocks, and no nearby pollution source was identified. The strong correlation between LCE3 and LCE5 may be due to *E. coli* accumulation at the bay-structured beach. However, the strong correlation between LCES and LCE5 post rain event suggests LCES may significantly impacted water quality at LCE5 (Table 11, Appendix B).

Lakeside Beach, located at Port Dalhousie in St. Catherines, is one of the most popular beaches in Niagara Region. Interesting correlations between transects and the pollution sources were determined throughout summer 2011. LS3 was correlated to LS1. The point source of pollution LSS was correlated to LS5 (Table 6, Appendix B). It suggests the impact limitation of LSS to beach water. It may strongly impacted water

quality at LS5, but weakly correlated to water quality at the other two transects. Due to only one post rain event data set was collected at Lake Side beach, the post rain correlation was not performed.

Queen's Royal Park is located at the downstream of Niagara River. *E. coli* contaminations at transects were correlated with each other throughout summer 2011, r_s=0.69 at QR1-QR3, and r_s=0.73 at QR3-QR5 (Table 7, Appendix B). In the meantime, there was not a strong correlation between any transect and the pollution source, suggesting that QRS only has limited impact on beach water quality at Queen's Royal Park throughout the summer. However, the strong correlations between QRS and QR1-QR3 post rain event suggests QRS may significantly impacted water quality at Queen's Royal Beach after the rain events (Table 12, Appendix B).

Lorraine Road beach was covered by algal mat for the entire summer in 2011, and no sample from the pollution source was collected due to safety concerns. Therefore, no correlation coefficient test was performed for the LR data.

3.3 Correlation between *E. coli* concentration and human fecal pollution concentration

At Queen's Royal Beach and Garden City Beach, human sewage markers were detected for 43% in 2010 and 50% in 2011 of the sampling days. Due to the highest rate of human sewage marker presence and the high *E. coli* concentrations, conventional PCR results and *E. coli* counts from Queen's Royal Beach were analyzed to examine the

correlation between *E. coli* concentration and the presence of human sewage marker to determine the impact from human fecal pollution to *E. coli* contamination.

At QR, *E. coli* concentration geometric average in the absence of human sewage markers was 62.9±45.0 CFU/100mL and 65.4±65.3 CFU/100mL in the presence of human sewage markers. A P-value of 0.89 indicated that there is no significant difference in *E. coli* concentration between the presence and absence of human sewage marker. It indicates that *E. coli* concentration in the presence of human sewage marker is not significantly higher than *E. coli* concentration in the absence of human sewage marker. Therefore, the increased presence of human fecal pollution may cause the increase of *E. coli* concentration at Queen's Royal Beach.

Due to the limitations of conventional PCR, real time PCR was performed to quantify the amount of human-specific *Bacteroidales* presented in the water. The real time PCR results revealed that *E. coli* contamination may be weakly correlated with human fecal pollution at Queen's Royal. The regressions between *E. coli* concentration and human fecal concentration were identified at QR1, QR3 and QRS (Figure 1, 2 and 4). However, the small slope and R square values with a significance level of 0.05 indicated that the correlations were weak. No correlation between *E. coli* concentration and human fecal concentration were found at QR5. Melt curve was also analyzed to make sure the specificity of the amplified product (Ririe et al., 1997; Smith et al., 2009) from 60 °C to 95 °C at 0.4 °C per minute (Figure 5) (Seurinck et al., 2005). Like conventional PCR result, real time PCR result suggests that while human fecal pollution may correlate to water quality at Queen's Royal Beach, it may not be the main cause of high *E. coli*. A

recent study indicates that bird droppings may be the major fecal pollution source at Hamilton beaches (Edge and Hill, 2007). A future study can be done to design animal-specific qPCR probes and determine the correlation between animal fecal concentration and *E. coli* contamination.

3.4 Difference between *E. coli* concentrations in 2010 and 2011

The average *E. coli* concentrations from all sampling days (3rd Tuesday of May until last Thursday of August) in 2010 and 2011 were also compared, in order to determine the difference between the two years. Due the different sampling beaches in the two years, *E. coli* concentrations from beach water samples at Bay Crystal Beach, Fifty Point Beach, Garden City Beach, Long beach, Long Beach Conservation East, Lorraine Road Beach, Lakeside Beach, and Queen's Royal Beach in 2010 and 2011 were used to process this study. Overall, *E. coli* concentration average was 30.3±13.4 CFU/100mL (geometric mean) in 2010 and 50.3±84.9 CFU/100mL (geometric mean) in 2011. A P value of 0.0003 (Paired-T test) suggests *E. coli* concentration in 2011 was significant higher than same beaches in 2010. However, it is difficult to determine if the decrease of water quality is due to heavier human activities or environmental events (Passerat et al., 2011). Rain event or animal impact could be the reasons for the change in water quality as well (Vinten et al., 2004; Whitman et al., 2008). Future work can be done to determine the correlation between environmental events and *E. coli* concentration.

Table 1. Sampling Sites in Niagara Region in summer 2010.

Lake Erie	Lake Ontario	Point Sources of Pollution	Drains	Sewage plant	Rivers
Long Beach Conservation West (LCW)	Queen's Royal (QR)	Charles Daly Beach East Source (CDEs)	Low Banks Drain (LB001)	BC-STP Raw (RAW)	Eighteen Mile Creek (ET001)
Long Beach Conservation East (LCE)	Garden City (GC)	Charles Daly Beach West (CDWS)	Casey Drain (CD001)	BC-STP Final Effluent (FE)	Fifteen Mile Creek (FF001)
Long Beach (LB)	Lakeside Beach (LS)	Centennial Park Beach Source (HCS)	Eagle Marsh Drain (EM001)		Forty Mile Creek (FM001)
Nickel Beach (NB)	Charles Daly Beach East (CDE)	Lorraine Road Beach Source (LRS)	Wignell Drain (WD001)		Sixteen Mile Creek (SX001)
Lorraine Road Beach (LR)	Charles Daly Beach West (CDW)	Nickel Beach Source (NBS)	Beaver Dam Creek (BD001)		Twenty Mile Creek (TN006)
Centennial Park Beach (HC)	Nelles Park Beach (NP)	Nelles Park Beach Source (NPS)	Point Abino Drain (PA001)		
Crystal Beach (BC)	Fifty Point Conservation Area (FP)	Queen's Royal Source (QRS)	Six Mile Creek (SM001)		
Bernard Beach (BE)			Krafts Drain (KD001)		

Table 2. Summary for *E. coli* concentration (CFU/100mL) and exceed rate (% > 100 CFU/100mL) at sampling beaches in Niagara Region in 2010.

Beach	Mean	Geo Mean	Std Err Mean	Exceed Rate (%), by year
BC	69	23	12.8	16
BE	37	12	8.0	10
CDE	73	20	14.6	17
CDW	155	27	47.0	26
FP	89	27	17.3	23
GC	286	41	73.2	38
HC	213	42	43.9	33
LB	34	18	4.7	13
LCE	175	43	35.4	39
LCW	97	26	23.1	19
LR	70	28	10.3	23
LS	144	42	32.5	29
NB	77	18	20.5	16
NP	223	34	62.9	35
QR	257	67	69.0	48

Table 3. Percent of sampling days with positively human-specific *Bacteroidales* DNA marker detection at 15 beaches in Niagara Region in 2010.

Beach	Days positive (%)
BC	10
BE	3
CDE	13
CDW	7
FP	7
GC	31
HC	8
LB	7
LCE	13
LCW	0
LR	10
LS	14
NB	3
NP	13
QR	43

Table 4. Water quality monitoring results for beaches at the Region of Niagara in 2010 and their potential risk classification.

	High human fecal contamination (≥ 10% of days showed HF positive)	Low human fecal contamination (< 10% of days showed HF positive)
High *E. coli* concentration (≥20% posting)	QR, GC, LCE, LS, LR, NP	HC, CDE, FP
Low *E. coli* concentration (<20% posting)	CDE, BC	LCW, NB, BE, LB

Table 5. Summary for *E. coli* concentration (CFU/100mL) and exceed rate (% > 100 CFU/100mL) at sampling beaches in Niagara Region in 2011.

Beach	Mean	Geo Mean	Std Err Mean	Exceed Rate (%), by year
BC	123.2	44.5	20.2	21.6
FP	173.1	24.1	57.3	21.1
GC	546.4	78.5	265.5	30.2
LB	158.7	49.3	46.2	17.9
LCE	1892.5	195.3	634.1	43.5
LR	405.4	105.8	91.5	27.6
LS	154.9	58.1	47.3	24.9
QR	142.1	63.0	25.9	37.8

Table 6. Percent of sampling days with positively human-specific *Bacteroidales* DNA marker detection at 7 beaches in Niagara Region in 2011.

Beach	Days positive (%)
BC	3
FP	3
GC	27
LB	3
LCE	0
LR	7
LS	7
QR	50

Table 7. Water quality monitoring results for beaches at the Region of Niagara in 2011 and their potential risk classification.

	High human fecal contamination (≥ 10% of days showed HF positive)	Low human fecal contamination (< 10% of days showed HF positive)
High *E. coli* concentration (≥20% posting)	QR, GC, LS, LR, FP and BC	LCE
Low *E. coli* concentration (<20% posting)	LB	Not Applicable

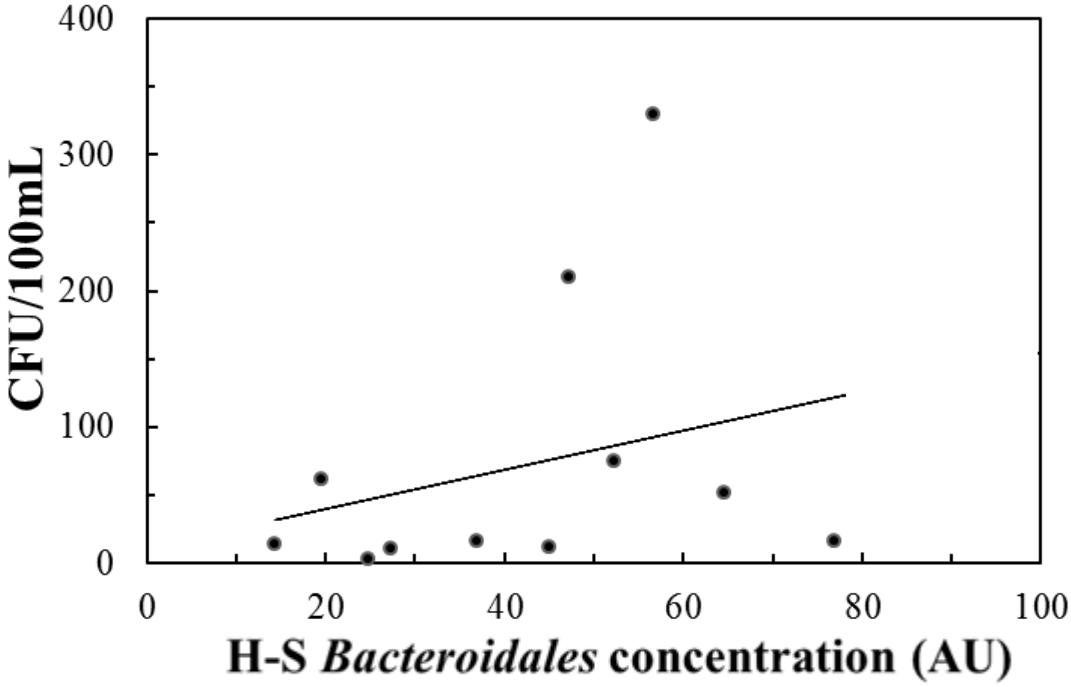


Figure 1. Correlation between *E. coli* concentration and human-specific *Bacteroidales* concentration at Queen's Royal Sampling Location #1 in Arbitrary Units.

$y = 1.44x + 10.90$, $R^2 = 0.22$

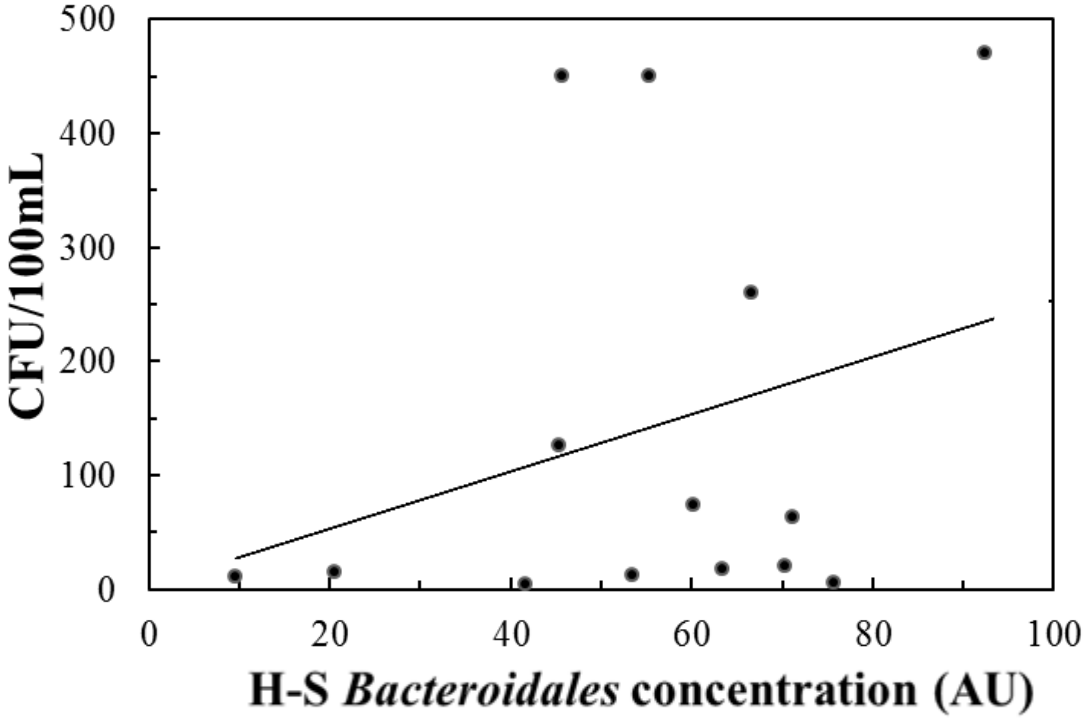


Figure 2. Correlation between *E. coli* concentration and human-specific *Bacteroidales* concentration at Queen's Royal Sampling Location #3 in Arbitrary Units.

$y = 2.51x + 2.65$, $R^2 = 0.14$

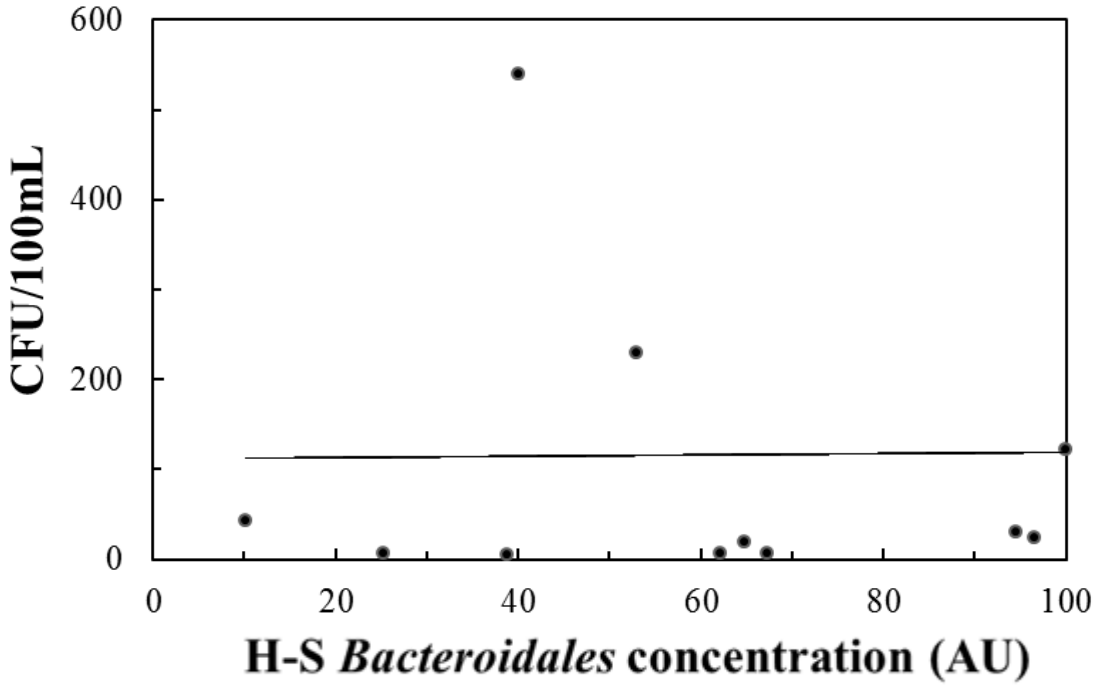


Figure 3. Correlation between *E. coli* concentration and human-specific *Bacteroidales* concentration at Queen's Royal Sampling Location #5 in Arbitrary Units.

$y = 0.076x + 112.01$, $R^2 = 0.0004$

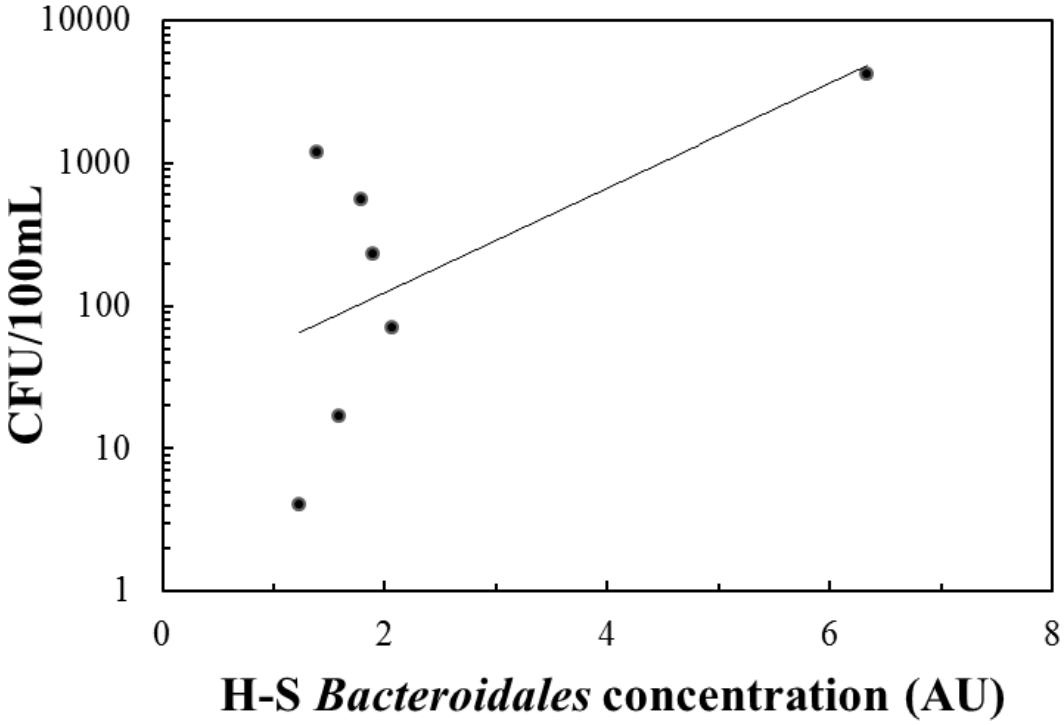


Figure 4. Correlation between *E. coli* concentration and human-specific *Bacteroidales* concentration at Queen's Royal Pollution Source in Arbitrary Units.

$y = 793.88x - 950.75$, $R^2 = 0.88$

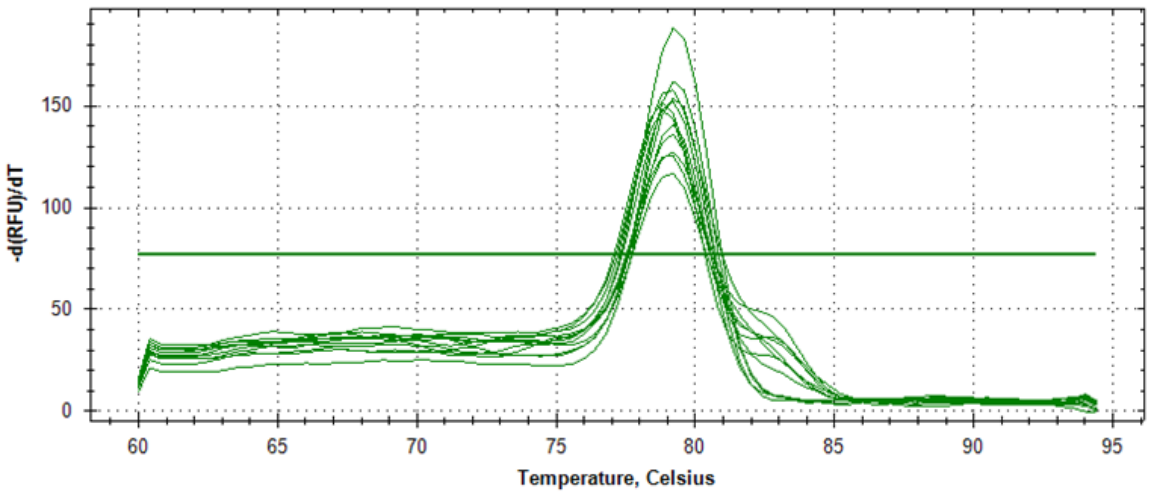


Figure 5. Melt curve of the human-specific *Bacteroidales* amplification (the change in fluorescence with temperature (dRFU/dT) versus temperature).

Chapter 4. Conclusion

The risk classification of beaches in Niagara Region was successful established. All beaches were divided into different categories depending on their *E. coli* concentration and the frequency of human fecal detection. Overall, water quality at investigated beaches in 2011 was decreased, compared to the same beaches investigated in 2010.

Water quality at most investigated beaches was relatively high. *E. coli* concentration at all beaches in 2010 was below the provincial guideline. In 2011, water quality at most beaches decreased but the average concentrations still remained within the provincial guideline. Two beaches (Long Beach Conservation East and Lorraine Road Beach) exceeded 100 CFU/100mL in 2011, suggesting an increased *E. coli* contamination issue at these two beaches. The human-specific *Bacteroidales* marker was not commonly detected at most beaches. However, at Queen's Royal Beach (43% in 2010 and 50% in 2011) and Garden City Beach (31% in 2010 and 27% in 2011), the detection rates were relatively high, compared to other beaches. It suggests that these two beaches may need remediation to ensure the water is suitable for public bathing and swimming.

Since it has been determined that human sewage may not be the major determinant for *E. coli* contamination at certain beaches, it would be interesting to develop more animal-specific *Bacteroidales* probes for real time PCR assays to determine the correlations between *E. coli* concentration and different animal hosts at individual beach.

Part II. Sigma factor expression in *E. coli* environmental isolates

Chapter 1. Introduction

1.1 Sigma factors

Sigma factors (σ factors) are a group of proteins that are essential for the initiation of RNA synthesis. Sigma factors are part of the RNA polymerase holoenzyme complex, which enables specific binding of RNA polymerases to promoters (Osawa and Yura, 1981; Jishage et al., 1996). Sigma factors can be divided into two families, the σ^{70} family and the σ^{54} family. Within the group of σ^{70} family, sigma factors are classified into four different subgroups depends on their structures and functions. Group 1 sigma factors are the primary sigma factors which responsible for cell growth, survival and the most RNA synthesis in exponential phase. They also have a structure that is close related to σ^{70}. Group 2 sigma factors have a similar structure to σ^{70} but not essential for cell growth. Group 3 includes the alternative sigma factors that enable the specific binding of RNA polymerase. Group 4 sigma factors mainly responsible for extracytoplasmic singles (Lonetto et al., 1992).

E. coli contains one primary sigma factor σ^{70} and six alternative sigma factors as, σ^{54}, σ^{38}, σ^{32}, σ^{28}, σ^{24} and σ^{19} (Jishage et al., 1996). RpoD (σ^{70}), the primary *E. coli* sigma factor or the 'housekeeping' sigma factor, has a molecular weight of 70 kDa. As a RpoD

group 1 sigma factor, RpoD regulates most genes that are essential for survival and growth in exponential phase. Its expression is also reported throughout the growth of the cells (Shiozawa et al., 1996). Mutation and deletion of *rpoD* could lead to an altered promoter specificity or even cell death (Gardella et al., 1989; Keener and Nomura, 1993).

σ^{54} (RpoN), encoded by *rpoN*, a member of the σ^{54} family, is an alternative sigma factor that responsible for nitrogen limitation. RpoN controls the transformation of more than 30 promoters in *E. coli*, but not all these promoters are response to nitrogen utilization (Zimmer et al., 2000). Few of them act as nitrogen assimilation supporter (Reitzer and Schneider, 2001).

σ^{38} (RpoS), encoded by *rpoS*, is another alternative sigma factor in *E. coli*. RpoS controls the expression of global stress response genes which are important in stationary phase. In *E. coli*, RpoS is responses to general stresses, such as heat, osmotic, oxidative and UV, etc. (Cheville et al., 1996; Hengge-Aronis, 2002). RpoS is known to control approximately 10% of the *E. coli* K-12 genome during stress conditions and stationary phase (Dong and Schellhorn, 2009). By using the general stress sigma factor RpoS as a model, a higher portion of *rpoS* mutants have been reported in the natural environment, compared to laboratory environment (Chiang et al., 2011). This suggests that RpoS may be an important factor for *E. coli* survival under environmental selection.

σ^{32} (RpoH), encoded by *rpoH*, is another alternative sigma factor that responses to heat stress. The expression of RpoH increases during the temperature changes from 30 ℃ to 42 ℃. RpoH degrades if the temperature stabilized at 30 ℃ or 42 ℃. It has a very short

half-life of one minute (Straus et al., 1987; Tilly et al., 1989). Due the instability of RpoH, continued transcription of *rpoH* is required for cell survival under such conditions (Erickson and Gross, 1989). Interestingly, the structure of *rpoH* mRNA is more important than amino acid sequence to the function of RpoH (Nagai et al., 1991). Once the cells enter a higher temperature (> 42 ℃), the expression of RpoE will be activated.

σ^{28} (RpoF), encoded by *fliA*, is an alternative sigma factor that associates with the formation of flagellar (Makinoshima et al., 2003). The expression of RpoF is reported to be negatively regulated by RpoS. (Makinoshima et al., 2003; Dong et al., 2011).

σ^{24} (RpoE), encoded by *rpoE*, is an alternative sigma factor that responsible for the extracytoplasmic/extreme heat stress. RpoE acts as the main regulator of the heat stress response, under conditions of extreme heat (>42 ℃). RpoE is also responsible for the transcription of *rpoH*. Expression of Rpo and RpoE has been studied in K12 strains but to a lesser extent in environmental strains (Dong and Schellhorn, 2009).

σ^{19} (FecI), encoded by *fecI*, is a sigma factor that responses to ferric citrate transport regulation. In *E. coli* K-12 strain, the ferric citrate transport system has three core proteins as FecA, FecR and FecI. FecI stays in the cytoplasm and guides RNAP to the promoters of *fecABCDE* to start transcription (Lonetto et al., 1994).

1.2 Phenotypic diversity in *E. coli*

Adaptation is critical for bacterial survival in the natural environment (Oulkheir et al., 2007), it usually appears when bacteria are facing stressed environment. It helps

bacteria survive and grow better in the environment where they are living in. It is usually achieved by the selection of non-synonymous mutations (Galhardo et al., 2007) or modified gene expressions (Velliou et al., 2011) when bacteria are experiencing stress. For example, increased salt resistant caused by the increased expression of membrane porins was observed when in *E. coli* mutants when they are experiencing increased osmotic stress in sewage treatment plant (Puranik et al., 2012).

Non-synonymous mutation sometimes leads to differential gene expressions. A recent study observed differential expressions introduced by different mutations in the same *E. coli* strain when *E. coli* were exposed to antibiotics (Ahmetagic and Pemberton, 2011).Therefore, different bacterial phenotypes may exist in the natural environment due to different gene composition and gene expression. This suggests that gene expressions in *E. coli* environmental strains may be different from commonly used *E. coli* laboratory strains.

Current water quality monitoring approach was established based on the assumption that environmental *E. coli* are functioning identically as a species. However, with the developed understanding of *E. coli*, the limitation of this water quality approach was observed in recent years. *E. coli* strains differ enormously in pathogenic potential. Most *E. coli* strains are harmless to human health while only a few strains carrying disease-causing genetic determinants (e.g., strain O157:H7). The importance of *E. coli* to human health issues is thus affected by genetic information (DNA) and gene expression in *E. coli* strains in addition to microbial numbers. This illustrates that the number of *E.*

coli in the water may not direct link to chance of getting sick. Differential gene expressions were observed in environmental strains at stationary phase (Chiang and Schellhorn, 2010). Previous work also observed differential gene expression in *E. coli* laboratory strains throughout growth (Dong and Schellhorn, 2009). This reveals the possibility of differential gene expression *E. coli* environment strains throughout growth.

The natural environment does not always provide the optimal growth condition as primary habitat. *E. coli* thus often experience the stressed environment which limits their growth including temperature, pH, nutrient and osmotic stresses, etc (Bissonnette et al., 1975; Armalyte et al., 2008). Therefore, environmental *E. coli* need to have the ability to adapt to their secondary habitat, such as the natural environment. It is usually achieved by genetic mutation. This process may be associated with non-synonymous mutations in stress-related genes, such as *rpoS* and *rpoE*. Previous studies have showed that RpoE and RpoS may control the expression of same gene under stressed conditions (Du et al., 2011). So, they may both play important roles in *E. coli* adaptation. Therefore, it is important to compare the change of expression levels of RpoS and RpoE between *rpoS* non-synonymous mutations and wild types. By doing that, we could reveal if RpoS and RpoE are the important factors for *E. coli* adaptation.

Sigma factors are the important global regulators in *E. coli*. Most essential genes for survival and growth are regulated by sigma factors. Therefore, this study is focussed on *E. coli* how gene expression is altered in *E. coli* environmental strains by analyzing differential expression levels in selected sigma factors.

In this study, the expression of RpoD, RpoS and RpoE in environmental *rpoS* mutants obtained from a range of natural water sources, including raw sewage and beach water, were examined along with the environmental *rpoS* wild type strains, laboratory and pathogenic strains. The objective is to understand how gene expression is altered in *E. coli* laboratory strain and environmental strains (wild type and *rpoS* mutants) by analyzing differences in RpoD, RpoS and RpoE expression levels.

Chapter 2. Methodology

2.1 *E. coli* Confirmation

After Differential Coliform Agar incubation (Section 2.1, Part I), single blue colonies were re-streaked onto MacConkey plates at 37.5 ℃ overnight for *E. coli* candidate purification (MacConney et al., 1905). After incubation, pink colored candidates were then suspended in 200mL 1X saline solution in 96-well microliter plate for further confirmation tests.

E. coli confirmation test includes two parts, EC-MUG test, and Tryptone test. EC-MUG is a media contains 4-methylumbelliferyl-β-D-glucuronide (MUG). *E. coli* contain the enzyme glucuronidase, after overnight incubation, glucuronidase will cleave MUG into blue fluorescent that can be visualized under UV light (Macconkey, 1905). For Tryptone test, *E. coli* contain the enzyme tryptophanase, after overnight incubation, tryptophanase will cleave tryptophan into indolic metabolites. By adding a small drop of "Kovacs" (aldehyde group of ρ-dimethylaminobenzaldehyde), a red ring will be formed if the candidate is *E. coli* (Hall, 1985; Bonadonna et al., 2007).

Therefore, during confirmation test, four sets of different media were prepared: *E. coli* candidate in saline solution, EC-MUG media, Tryptone media and the storage media TSB (Tryptic Soy Broth). Candidates were replicated into 96-well microtiter plate with a steel replicator. Between different sets of *E. coli* plates, the replicator was washed through bleach and dH_20, and then flamed with 95% ethanol to avoid cross-contamination.

Candidates with positive blue fluorescent in EC-MUG test and red ring formation in Tryptone test were confirmed as *E. coli* and were stored in TSB with 20% glycerol at -80 ℃ (Hassan and Frank, 2004).

2.2 Identification of *rpoS* mutants

A total of 1008 *E. coli* natural isolates were isolated at the end of the collection period (Table 8). Isolates were screened for two highly RpoS-dependent phenotypes: succinate growth and catalase activity (Chen et al., 2004; Chiang et al., 2011). A total of 10 isolates showed positive growth in succinate media and negative response in catalase activity were identified as the *rpoS* mutant candidates. The *rpoS* region of these candidates was amplified with PCR. An *E. coli* single colony was suspended in 100 µL ddH_2O and used as PCR template. ML08-4514 (5'-CTT-GCA-TTT-TGA-AAT-TCG-TTA-CA-3') was used as the forward primer, and ML08-4515 (5'-TTA-ACG-CCA-TTC-TCG-GTT-TTA-C-3'), ML08-5873 (5'-GGT-GCA-ATC-TCC-AGC-CG-3') AND ML08-5874 (5'-GGA-GAA-TCG-TGG-CTT-AGT-CAG-3') were used as three different sets of reverse primers (Chen et al., 2004). Reaction mix includes 2 µL of 10X PCR buffer, 0.8 mM of dNTP mixture, 1.6 µM of each primer, 8 Units of homemade Taq, 1 µL of DNA template and PCR grade ddH_2O to a total volume of 20 µL. The reaction was initialized at 95 °C for 2 minutes. The reaction was continued by 30 cycles of 95 °C for 30 sec, 58 °C for 30 sec, 72 °C for 1 min and final extension at 72 °C for 10 min (Chiang et al., 2011).

Amplified PCR products were purified (NucleoSpin® Gel and PCR Clean-up) and were then sequenced to confirm changes in *rpoS*. Identified *rpoS* mutants were used as experimental strains to process the following experiment with wild type controls.

Table 8. 1008 *E. coli* environmental isolates collection obtained from August/30 2010.

Type	No. of isolates
Beach water	562
Stream water runoff	150
Drains	97
Sewage treatment plant raw sample	30
Sewage treatment plant final effluent	45
Algae Bloom	13
Rivers	111

2.3 Sample collection

Isolated strains were grown in LB, and samples were collected throughout growth at desired points. OD (Optical density) 600nm was measured with a spectrophotometer as an indicator of cell growth. Samples were collected at designated OD_{600} points as follows: 0.1, 0.3, 0.6, 0.8, 1.0, 1.2 and 1.5. Samples from overnight culture and 72 hour growth were also taken as long term cultures. Protein samples from K-12 MG1655 and HS2210 (MG1655 Δ*rpoS*) overnight cultures were also collected as two sets of controls.

2.4 SDS-PAGE

Pelleted cells were washed twice to minimize salt impact. Protein samples were prepared by suspending the remaining pellets in 1X SDS loading buffer to the same concentration (equivalent OD_{600} of 1.0). Protein samples were then been boiled to lyse cells and denature proteins. Protein samples were stocked at -20 ℃. A total of 10 μg

protein from each sample was subjected to electrophoresis on 10% SDS gels with stacking gels (Dong et al., 2009). Protein staining was also performed to another gel as loading controls.

2.5 Immunoblot analysis

Antibodies specific for RpoD, RpoS and RpoE were used to determine expression levels. Polyvinylidene fluoride (PVDF) membrane was soaked in methanol for exact 1 minute with shaking before protein transfer. Proteins were transferred to PVDF membrane by semi-dry transfer electrophoresis at 25 V for 20 minutes. Membranes were then blocked with 5% milk in TBST for one hour with shaking to prevent unspecific binding of the primary antibody. Then, membranes were incubated in the appropriate primary antibodies (Neoclone RNAP RpoD, RpoS and RpoE, catalog# W0004, SP002-W0002, and WP007, respectively; all made at 1:10,000 in 5% skimmed milk) overnight at 4℃ with shaking. In the second morning, membranes were washed using TBST three times before incubation with the secondary antibody (BioRad Goat anti-mouse lgG (H+L)-HRP Conjugate, catalog# 170-6516, made at 1:3,000 in 5% skimmed milk) for one hour at room temperature with shaking. Membranes were washed three times with TBST at ten minutes each. Drained membranes were placed in enhanced chemiluminescence (ECL) solution for 1 minute. Finally, membranes were wrapped in Saran wrap and placed in a film cassette carefully, avoiding air bubble formation. Films were exposed in the dark room with proper exposure times depending on the observed expression levels.

Chapter 3. Results

3.1 *rpoS* mutant screening

After catalase bubble test and succinate growth test, 10 out of 1008 natural isolates showed negative responses in catalase test and positive responses in succinate growth. These isolates were determined as potential candidates with *rpoS* mutations. Nucleotide sequencing returned one isolate, R4D8, as a non-synonymous mutation at its 43rd amino acid site with an A to T transversion, which led to a Glutamine to Glutamic Acid switch. The mutant frequency was thus 0.098%.

Besides R4D8, another three *rpoS* environmental mutants and two wild type isolates that have been identified previously (Chiang et al., 2011) were added into the following experiment along with laboratory strain MG1655 and pathogenic strain EDL933 (Table 9).

Table 9. Environmental *E. coli* isolates used in this study.

Isolate	Location	Type	Type of mutation (*rpoS*)	Source
MG1655	N/A	Laboratory	WT	Lab
AZB07	Hamilton	Sediment Core	G--> A transition	(Chiang et al., 2011)
AZB10	Hamilton	Sediment Core	WT	(Chiang et al., 2011)
BNB03	Toronto	Sediment Core	1 bp deletion	(Chiang et al., 2011)
ECE12	Main and King	Combined Sewage Flow	5 bp deletion	(Chiang et al., 2011)
ECH01	Hamilton Harbour	Beach	WT	(Chiang et al., 2011)
EDL933	N/A	Pathogenic	WT	Lab
R4D8	Charles Daley West	Storm water outfall	A--> T transversion	CCIW, sequenced by W. Zheng

3.2 Immunoblot

3.2.1 RpoD Expression

As the primary sigma factor, consistent RpoD expression in MG1655 was observed. The expression of RpoD in MG1655 remained at high level across all protein samples (Figure 6). Within the four environmental *rpoS* mutants, AZB07 and ECE12 had the same levels of RpoD expression as MG1655, while the expressions of RpoD in the other mutants BNB03 and R4D8 decreased throughout growth. RpoD expressions in two wild type strains (AZB10, ECH01) and pathogenic strain (EDL933) also decreased throughout growth, compared to the laboratory strain MG1655.

3.2.2 RpoS Expression

The expression of RpoS in MG1655 was consistent with previously published works (Dong et al., 2008; Dong and Schellhorn, 2009; Dong et al., 2011). RpoS was first detected when OD_{600} reached 0.3, peaked at OD_{600} 1.0 and remained at the same level until the end of the experiment period. Within the *rpoS* mutants, the expression of RpoS heavily decreased in AZB07, AZB10 and ECE12, and lightly reduced in R4D8, compared to the expression level in MG1655. In R4D8, RpoS expression was observed at OD_{600} of 0.6, and peaked at OD_{600} of 1.0. Interestingly, after that point, RpoS expression then declined and remained at a relative low level. For the two wild type strains, RpoS in AZB10 and ECH01 were both observed at OD_{600} of 0.3. Then, the expression increased until it reached its maximum level when cells entered stationary phase (OD_{600} of 1.5). At the end, RpoS expression slightly decreased in the long term culture (overnight and 72h

sample). In EDL933, the pathogenic strain, RpoS reached the maximum expression level at OD_{600} of 0.8, and the expression remained until cells entered stationary phase. In overnight and long term culture, the expression level was lower, compared to stationary phase (Figure 7).

3.2.3 RpoE Expression

A consistent high level of RpoE expression was observed in MG1655 from the beginning of experiment until the end. The same levels of RpoE expression were observed in all other strains. The size of in ECE12 strain was observed to be 5kDA larger than MG1655, while other environmental strains have the same size (Figure 8).

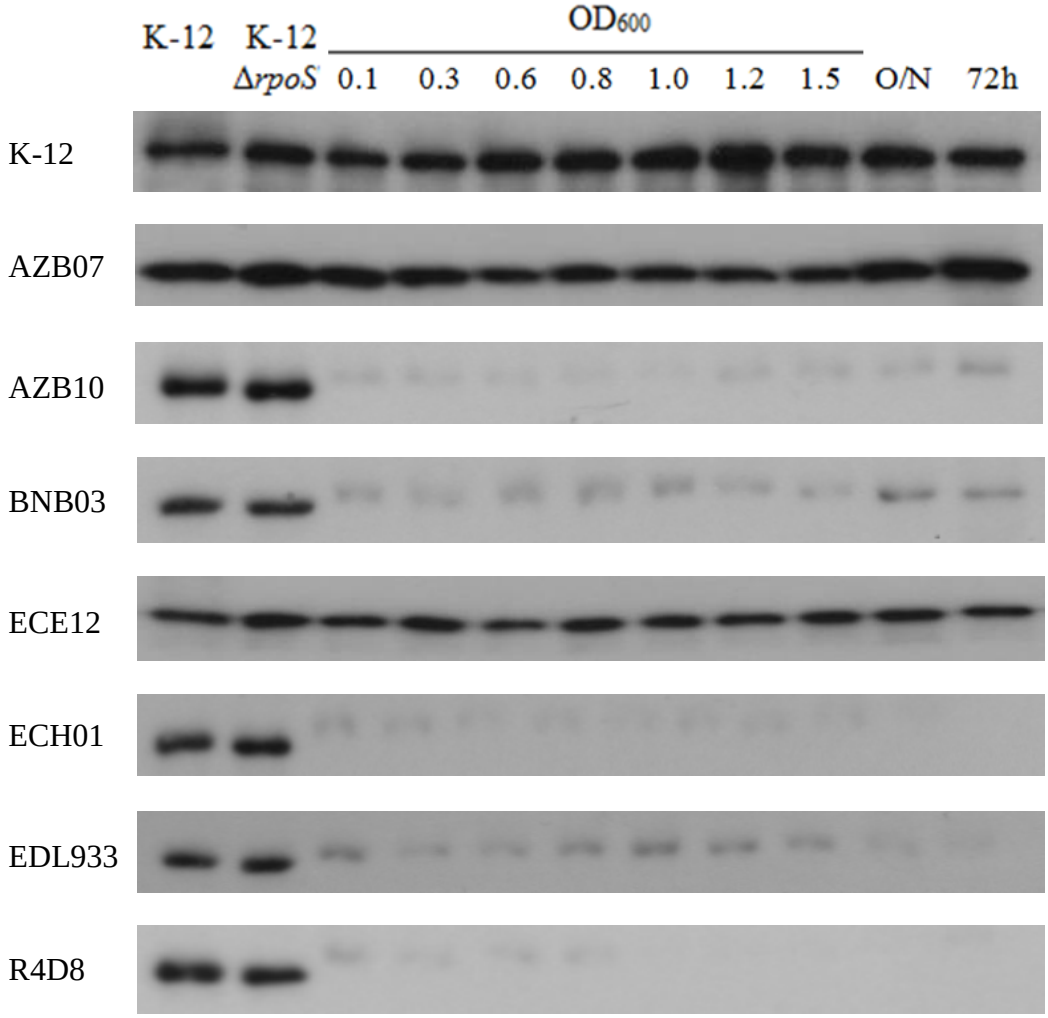


Figure 6. RpoD expression levels in *E. coli* laboratory, environmental and pathogenic strains. Control 1: K- 12 MG1655. Control 2: HS2210 (MG1655 *ΔrpoS*).

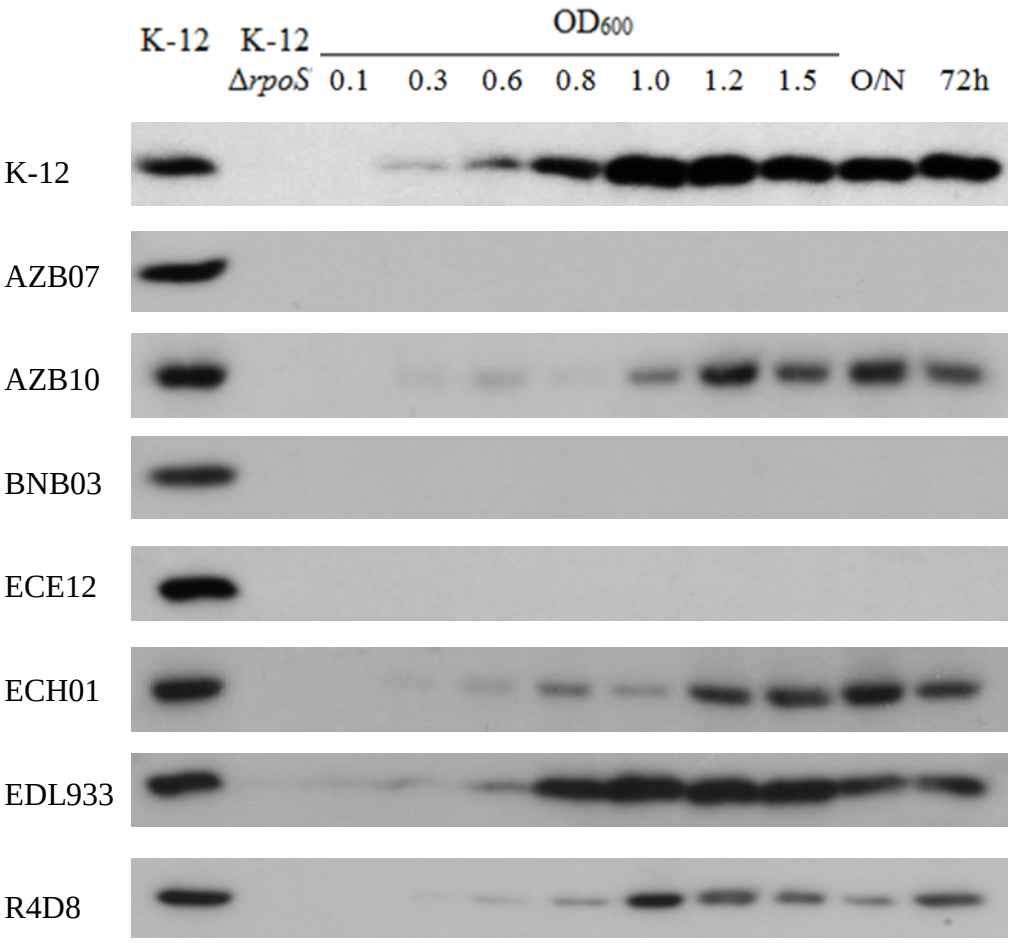


Figure 7. RpoS expression levels in *E. coli* laboratory, environmental and pathogenic strains. Control 1: K- 12 MG1655. Control 2: HS2210 (MG1655 *ΔrpoS*).

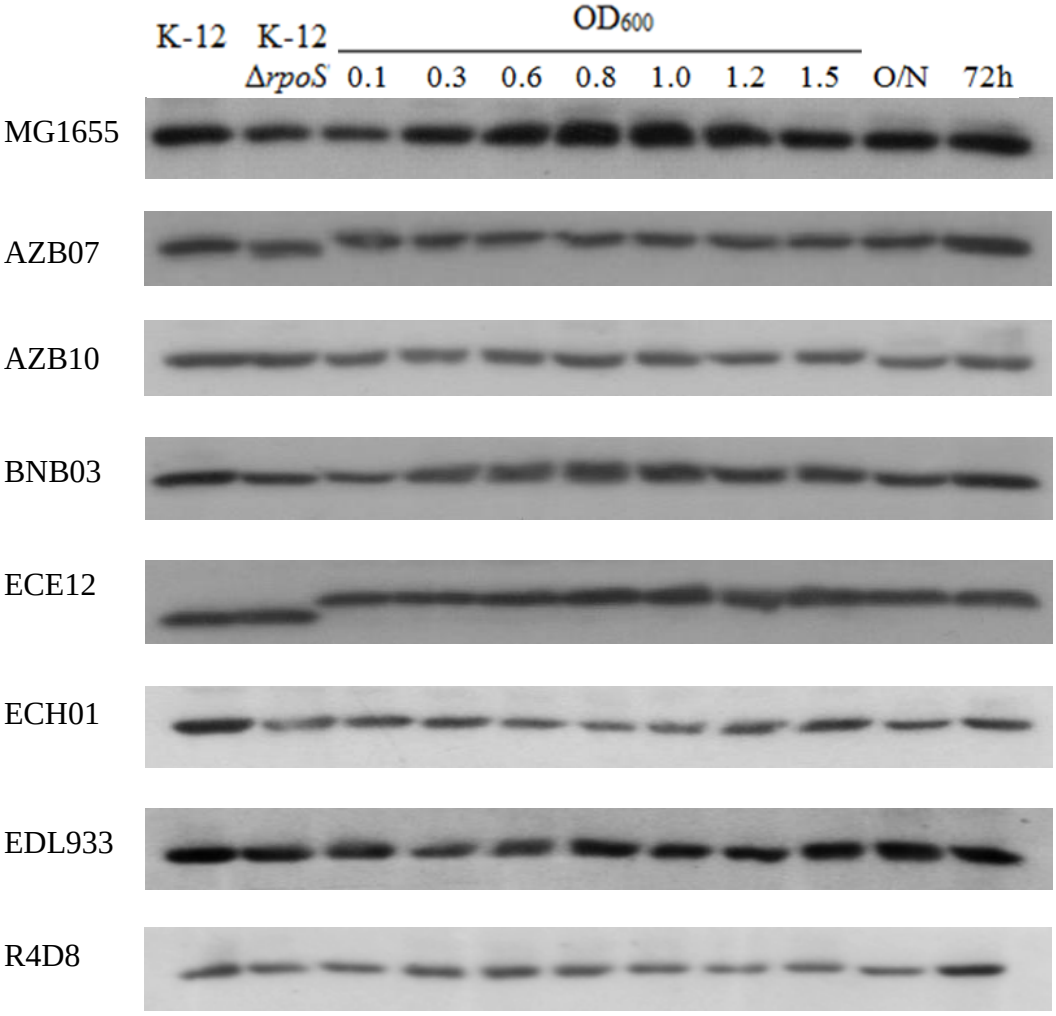


Figure 8. RpoE expression levels in *E. coli* laboratory, environmental and pathogenic strains. Control 1: K- 12 MG1655. Control 2: HS2210 (MG1655 *ΔrpoS*).

Chapter 4. Discussion

The expressions of RpoS in three *rpoS* mutants (AZB07, BNB03 and ECE12) heavily reduced. In the other *rpoS* mutant R4D8, although RpoS was expressed at the same stage as MG1655, but the expression level was relatively lower, compared to MG1655. RpoS is usually detected either during starvation condition or when cells enter stationary phase (Dong and Schellhorn, 2009) to slow down *E. coli*'s growth but to maximize survival rate. About 200 genes that essential for survival are regulated by RpoS. The change of RpoS expression levels will results in the change of RpoS regulated downstream essential genes expression levels. This also will introduce the difference in global gene expression in bacteria (Alvarez-Ordonez et al., 2012).

Under laboratory investigation, the spontaneous mutation rate has been reported around 10^{-10} (Drake et al., 1998; Wielgoss et al., 2011). Previous study has reported a relatively high *rpoS* mutation frequency (0.3%) in the natural environment (Chiang et al., 2011). This study also confirmed a higher environmental *rpoS* mutation frequency of 0.098%. Mutations in *rpoS* suggest an environmental selection of *rpoS* mutant and its importance to *E. coli* adaptation in the natural environment (Chiang and Schellhorn, 2010). It also suggests that RpoS might be an important regulator during environmental adaptation. Interestingly, it has been reported that some *E. coli* strains contain an insertion named IS10 in their *ots-BA* operons. This operon is likely to control some regulations that were controlled by *rpoS* in *rpoS* mutants (Stoebel et al., 2009). This also suggested the possibility of gene expression diversity in *E. coli* environmental strains.

Continuous high levels of RpoE expression were observed in all strains throughout this experiment, suggesting that RpoE may not play an important role in environmental adaption of *E. coli* from Niagara Region. It is possible that in Niagara Region, water temperature is less likely to exceed the optimal growth temperature *E. coli*'s, which is 37.5 ℃. Surprisingly, the size of RpoE in ECE12 was appeared to be 5kDA larger than in MG1655. Nucleotide sequencing has indicated that *rpoE* in ECE12 is identical to MG1655. The change of RpoE size may due to changes at translational or post-translational level.

RpoD was tested because it is the house-keeping factor and it transcribes most essential genes. Interestingly, unexpected expression levels were observed in the control sigma factor, RpoD. In two out of four *rpoS* mutated strains (BNB03 and R4D8), two *rpoS* wild type strains (AZB10 and ECH01) and pathogenic strain (EDL933), RpoD was expressed at relatively low levels, compared to the laboratory stain K-12 MG1655. From theory, RpoD expression is expected at a high level in most conditions. It is possible that the relatively higher level of RpoD expression in K-12 MG1655 is due to years' selection of fast-grow *E. coli* stain in rich media under laboratory conditions. However, this optimal growth condition does not often appear in the natural environment. This suggests the differential RpoD expressions among environmental strains to laboratory strain. The difference in RpoD expression levels across environmental isolates suggests that RpoD may play an important role during environmental selection.

When comparing expression levels of RpoD and RpoS between experimental isolates and laboratory strain MG1655, differential expressions were observed. No experimental strain expressed the same levels of RpoD and RpoS as MG1655 (Figure 7). The reduced levels of RpoS expression were observed in AZB10, ECH01, EDL933 and R4D8. The regular levels of RpoD expression were observed in three strains (including two *rpoS* mutants AZB07 and ECE12), compared to MG1655. On the other hand, RpoD expression reduced group includes two *rpoS* mutants (BNB03 and R4D8), suggesting that there may not have a correlation between RpoD expression and *rpoS* mutants. The changes of phenotypes are likely caused by adaption to stressed conditions in the natural environment, such as carbon source limitation (Ferenci and Spira, 2007; Olvera et al., 2009). Recent studies also revealed different responses of RpoS when pathogenic *E. coli* were subjecting to heat and cold stress conditions (Vidovic et al., 2011; Vidovic et al., 2012). This is a good example of phenotypic difference in environmental *E. coli* caused by differential gene expression. In this study, all isolates were grown at the same laboratory conditions and all protein samples were harvested at the same growth points. So the phenotypic differences observed were less likely caused by laboratory growth conditions but differential gene expression caused by different genetic composition. The differences in genetic composition are probably maybe by mutagenesis during environmental selection, which helps environmental *E. coli* to survive and grow in the new environment.

In conclusion, adaptation is critical for *E. coli* to survive and change their lifestyles from the primary habitat to the secondary habitat (White et al., 2011). The

change of expression levels in RpoD and RpoS suggests that they may the important regulators during environmental adaptation, while RpoE is likely to have little impact on natural adaptation in these experimental strains. Because RpoD and RpoS regulate most essential genes for survival and growth in *E. coli*, the different expression levels of RpoD and RpoS suggested the possibility of differential global gene expressions in environmental *E. coli* strains.

Bibliographies:

Ahmed, W., N. Masters, et al. (2012). "Consistency in the host specificity and host sensitivity of the Bacteroides HF183 marker for sewage pollution tracking." *Lett Appl Microbiol* **55**(4): 283-289.

Ahmed, W., N. Masters, et al. (2012). "Consistency in the host specificity and host sensitivity of the Bacteroides HF183 marker for sewage pollution tracking." *Lett Appl Microbiol*.

Ahmed, W., J. Stewart, et al. (2008). "Evaluation of Bacteroides markers for the detection of human faecal pollution." *Lett Appl Microbiol* **46**(2): 237-242.

Ahmetagic, A. and J. M. Pemberton (2011). "Antibiotic resistant mutants of Escherichia coli K12 show increases in heterologous gene expression." *Plasmid* **65**(1): 51-57.

Alvarez-Ordonez, A., M. Begley, et al. (2012). "Polymorphisms in rpoS and heterogeneity in stress tolerance in natural isolates of Cronobacter sakazakii." *Appl Environ Microbiol*.

Armalyte, J., V. Seputiene, et al. (2008). "An Escherichia coli asr mutant has decreased fitness during colonization in a mouse model." *Res Microbiol* **159**(6): 486-493.

Bernhard, A. E. and K. G. Field (2000). "Identification of nonpoint sources of fecal pollution in coastal waters by using host-specific 16S ribosomal DNA genetic markers from fecal anaerobes." *Appl Environ Microbiol* **66**(4): 1587-1594.

Bernhard, A. E. and K. G. Field (2000). "A PCR assay To discriminate human and ruminant feces on the basis of host differences in Bacteroides-Prevotella genes encoding 16S rRNA." *Appl Environ Microbiol* **66**(10): 4571-4574.

Bertke, E. E. (2007). "Composite Analysis for Escherichia coli at Coastal Beaches." *Journal of Great Lakes Research* **33**(2): 335-341.

Bissonnette, G. K., J. J. Jezeski, et al. (1975). "Influence of environmental stress on enumeration of indicator bacteria from natural waters." *Appl Microbiol* **29**(2): 186-194.

Bonadonna, L., C. Cataldo, et al. (2007). "Comparison of methods and confirmation tests for the recovery *Escherichia coli* in water." *Desalination* **213**: 18-23.

Boone, D. R. and R. W. Castenholz (2001). *Bergey's Manual of Systematic Bacteriology*. New York, Springer-Verlag.

Carlos, C., F. Alexandrino, et al. (2012). "Use of Escherichia coli BOX-PCR fingerprints to identify sources of fecal contamination of water bodies in the State of São Paulo, Brazil." *Journal of Environmental Management* **93**: 38-43.

Chen, G., C. L. Patten, et al. (2004). "Positive selection for loss of RpoS function in Escherichia coli." *Mutat Res* **554**(1-2): 193-203.

Cheville, A. M., K. W. Arnold, et al. (1996). "rpoS regulation of acid, heat, and salt tolerance in Escherichia coli O157:H7." *Appl Environ Microbiol* **62**(5): 1822-1824.

Chiang, S. M., T. Dong, et al. (2011). "Phenotypic diversity caused by differential RpoS activity among environmental Escherichia coli isolates." *Appl Environ Microbiol* **77**(22): 7915-7923.

Chiang, S. M. and H. E. Schellhorn (2010). "Evolution of the RpoS regulon: origin of RpoS and the conservation of RpoS-dependent regulation in bacteria." *J Mol Evol* **70**(6): 557-571.

Corder, G. W. and D. I. Foreman (2009). *Nonparametric Statistics for Non-Statisticians: A Step-by-Step Approach*.

Dick, L. K., A. E. Bernhard, et al. (2005). "Host distributions of uncultivated fecal Bacteroidales bacteria reveal genetic markers for fecal source identification." *Appl Environ Microbiol* **71**(6): 3184-3191.

Dong, T., B. K. Coombes, et al. (2009). "Role of RpoS in the virulence of Citrobacter rodentium." *Infect Immun* **77**(1): 501-507.

Dong, T., M. G. Kirchhof, et al. (2008). "RpoS regulation of gene expression during exponential growth of Escherichia coli K12." *Mol Genet Genomics* **279**(3): 267-277.

Dong, T. and H. E. Schellhorn (2009). "Control of RpoS in global gene expression of Escherichia coli in minimal media." *Mol Genet Genomics* **281**(1): 19-33.

Dong, T., R. Yu, et al. (2011). "Antagonistic regulation of motility and transcriptome expression by RpoN and RpoS in Escherichia coli." *Mol Microbiol* **79**(2): 375-386.

Drake, J. W., B. Charlesworth, et al. (1998). "Rates of spontaneous mutation." *Genetics* **148**(4): 1667-1686.

Du, H., M. Wang, et al. (2011). "Coregulation of gene expression by sigma factors RpoE and RpoS in Salmonella enterica serovar Typhi during hyperosmotic stress." *Curr Microbiol* **62**(5): 1483-1489.

Dwight, R. H., L. M. Fernandez, et al. (2005). "Estimating the economic burden from illnesses associated with recreational coastal water pollution--a case study in Orange County, California." *J Environ Manage* **76**(2): 95-103.

Edge, T. A. and S. Hill (2007). "Multiple lines of evidence to identify the sources of fecal pollution at a freshwater beach in Hamilton Harbour, Lake Ontario." *Water Res* **41**(16): 3585-3594.

Edge, T. A., S. Hill, et al. (2010). "Microbial Source Tracking Studies at Niagara Region Beaches, Progress Report for 2010."

Edge, T. A., S. Hill, et al. (2010). "Library-dependent and library-independent microbial source tracking to identify spatial variation in faecal contamination sources along a Lake Ontario beach (Ontario, Canada)." *Water Sci Technol* **62**(3): 719-727.

Erickson, J. W. and C. A. Gross (1989). "Identification of the sigma E subunit of Escherichia coli RNA polymerase: a second alternate sigma factor involved in high-temperature gene expression." *Genes Dev* **3**(9): 1462-1471.

Ferenci, T. and B. Spira (2007). "Variation in stress responses within a bacterial species and the indirect costs of stress resistance." *Ann N Y Acad Sci* **1113**: 105-113.

Ferguson, A. S., A. C. Layton, et al. (2012). "Comparison of fecal indicators with pathogenic bacteria and rotavirus in groundwater." *Sci Total Environ* **431**: 314-322.

Frank, S. A. (2009). "The common patterns of nature." *J Evol Biol* **22**(8): 1563-1585.

Galhardo, R. S., P. J. Hastings, et al. (2007). "Mutation as a stress response and the regulation of evolvability." *Crit Rev Biochem Mol Biol* **42**(5): 399-435.

Gardella, T., H. Moyle, et al. (1989). "A mutant Escherichia coli sigma 70 subunit of RNA polymerase with altered promoter specificity." *J Mol Biol* **206**(4): 579-590.

Hall, L. P. (1985). "A direct plate count method for detecting E. coli in frozen food by detecting indole in colonies." *Food Microbiology* **1985**(2): 31-37.

Hassan, A. N. and J. F. Frank (2004). "Attachment of Escherichia coli O157:H7 grown in tryptic soy broth and nutrient broth to apple and lettuce surfaces as related to cell hydrophobicity, surface charge, and capsule production." *International Journal of Food Microbiology* **96**: 103-109.

Health Canada (2012). "Guidance on Waterborne Bacterial Pathogens."

Hengge-Aronis, R. (2002). "Signal transduction and regulatory mechanisms involved in control of the sigma(S) (RpoS) subunit of RNA polymerase." *Microbiol Mol Biol Rev* **66**(3): 373-395, table of contents.

Hlavsa, M. C., V. A. Roberts, et al. (2011). "Surveillance for waterborne disease outbreaks and other health events associated with recreational water --- United States, 2007--2008." *MMWR Surveill Summ* **60**(12): 1-32.

Jishage, M., A. Iwata, et al. (1996). "Regulation of RNA polymerase sigma subunit synthesis in Escherichia coli: intracellular levels of four species of sigma subunit under various growth conditions." *J Bacteriol* **178**(18): 5447-5451.

Keener, J. and M. Nomura (1993). "Dominant lethal phenotype of a mutation in the -35 recognition region of Escherichia coli sigma 70." *Proc Natl Acad Sci U S A* **90**(5): 1751-1755.

Kildare, B. J., C. M. Leutenegger, et al. (2007). "16S rRNA-based assays for quantitative detection of universal, human-, cow-, and dog-specific fecal Bacteroidales: a Bayesian approach." *Water Res* **41**(16): 3701-3715.

Lee, C., J. Kim, et al. (2006). "Absolute and relative QPCR quantification of plasmid copy number in Escherichia coli." *J Biotechnol* **123**(3): 273-280.

Lonetto, M., M. Gribskov, et al. (1992). "The sigma 70 family: sequence conservation and evolutionary relationships." *J Bacteriol* **174**(12): 3843-3849.

Lonetto, M. A., K. L. Brown, et al. (1994). "Analysis of the Streptomyces coelicolor sigE gene reveals the existence of a subfamily of eubacterial RNA polymerase sigma factors involved in the regulation of extracytoplasmic functions." *Proc Natl Acad Sci U S A* **91**(16): 7573-7577.

Macconkey, A. (1905). "Lactose-Fermenting Bacteria in Faeces." *J Hyg (Lond)* **5**(3): 333-379.

MacConney, A., M. B., et al. (1905). "Lactose-fermenting bacteria in faeces." *The Journal of Hygiene* **5**(3): 333-379.

Makinoshima, H., S. Aizawa, et al. (2003). "Growth phase-coupled alterations in cell structure and function of Escherichia coli." *J Bacteriol* **185**(4): 1338-1345.

Masters, N., A. Wiegand, et al. (2011). "Escherichia coli virulence genes profile of surface waters as an indicator of water quality." *Water Res* **45**(19): 6321-6333.

Ministry of Environment and Energy (1994). *Water Management*.

Nagai, H., H. Yuzawa, et al. (1991). "Interplay of two cis-acting mRNA regions in translational control of sigma 32 synthesis during the heat shock response of Escherichia coli." *Proc Natl Acad Sci U S A* **88**(23): 10515-10519.

National Academy of Sciences (1977). *Drinking Water and Health, Part I*. Washington, D.C.

NWRI Scientific Assessment Report (2006). *Microbial source tracking in aquatic ecosystems: the state of the science and an assessment of needs*. Burlington, Ontario, National Water Research Institute, Environment Canada.

Olvera, L., A. Mendoza-Vargas, et al. (2009). "Transcription analysis of central metabolism genes in Escherichia coli. Possible roles of sigma38 in their expression, as a response to carbon limitation." *PLoS One* **4**(10): e7466.

Osawa, T. and T. Yura (1981). "Effects of reduced amount of RNA polymerase sigma factor on gene expression and growth of Escherichia coli: studies of the rpoD450 (amber) mutation." *Mol Gen Genet* **184**(2): 166-173.

Oulkheir, S., K. Ounine, et al. (2007). "The effect of salt concentration and pH on the heat resistance of Escherichia coli in tryptic soy broth." *Acta Microbiol Immunol Hung* **54**(4): 399-412.

Passerat, J., N. K. Ouattara, et al. (2011). "Impact of an intense combined sewer overflow event on the microbiological water quality of the Seine River." *Water Res* **45**(2): 893-903.

Puranik, S., S. Shaligram, et al. (2012). "Demonstration of sequential adaptation strategy for developing salt tolerance in bacteria for wastewater treatment: A study using Escherichia coli as model." *Bioresour Technol* **121C**: 282-289.

Reitzer, L. and B. L. Schneider (2001). "Metabolic context and possible physiological themes of sigma(54)-dependent genes in Escherichia coli." *Microbiol Mol Biol Rev* **65**(3): 422-444, table of contents.

Rice, E. W., M. J. Allen, et al. (1990). "Efficacy of beta-glucuronidase assay for identification of Escherichia coli by the defined-substrate technology." *Appl Environ Microbiol* **56**(5): 1203-1205.

Ririe, K. M., R. P. Rasmussen, et al. (1997). "Product differentiation by analysis of DNA melting curves during the polymerase chain reaction." *Anal Biochem* **245**(2): 154-160.

Safe Drinking Water Act (2002). Ontario Drinking Water Quality Standards.

Seurinck, S., T. Defoirdt, et al. (2005). "Detection and quantification of the human-specific HF183 Bacteroides 16S rRNA genetic marker with real-time PCR for assessment of human faecal pollution in freshwater." *Environ Microbiol* **7**(2): 249-259.

Shiozawa, T., C. Ueguchi, et al. (1996). "The rpoD gene functions as a multicopy suppressor for mutations in the chaperones, CbpA, DnaJ and DnaK, in Escherichia coli." *FEMS Microbiol Lett* **138**(2-3): 245-250.

Sidhu, J. P., L. Hodgers, et al. (2012). "Prevalence of human pathogens and indicators in stormwater runoff in Brisbane, Australia." *Water Res*.

Smith, E., M. E. Jones, et al. (2009). "Quantitation of DNA methylation by melt curve analysis." *BMC Cancer* **9**: 123.

Stoebel, D. M., K. Hokamp, et al. (2009). "Compensatory evolution of gene regulation in response to stress by Escherichia coli lacking RpoS." *PLoS Genet* **5**(10): e1000671.

Straus, D. B., W. A. Walter, et al. (1987). "The heat shock response of E. coli is regulated by changes in the concentration of sigma 32." *Nature* **329**(6137): 348-351.

Tilly, K., J. Spence, et al. (1989). "Modulation of stability of the Escherichia coli heat shock regulatory factor sigma." *J Bacteriol* **171**(3): 1585-1589.

Van De Werfhorst, L. C., B. Sercu, et al. (2011). "Comparison of the host specificities of two bacteroidales quantitative PCR assays used for tracking human fecal contamination." *Appl Environ Microbiol* **77**(17): 6258-6260.

Velliou, E. G., E. V. Derlinden, et al. (2011). "Heat adaptation of Escherichia coli K12: effect of acid and glucose." *Procedia Food Science* **1**: 987-993.

Vidovic, S., A. K. Mangalappalli-Illathu, et al. (2011). "Prolonged cold stress response of Escherichia coli O157 and the role of rpoS." *Int J Food Microbiol* **146**(2): 163-169.

Vidovic, S., A. K. Mangalappalli-Illathu, et al. (2012). "Heat acclimation and the role of RpoS in prolonged heat shock of Escherichia coli O157." *Food Microbiol* **30**(2): 457-464.

Vinten, A. J., D. R. Lewis, et al. (2004). "Predicting the effect of livestock inputs of E. coli on microbiological compliance of bathing waters." *Water Res* **38**(14-15): 3215-3224.

White, A. P., K. A. Sibley, et al. (2011). "Intergenic sequence comparison of Escherichia coli isolates reveals lifestyle adaptations but not host specificity." *Appl Environ Microbiol* **77**(21): 7620-7632.

Whitman, R. L., K. Przybyla-Kelly, et al. (2008). "Sunlight, season, snowmelt, storm, and source affect E. coli populations in an artificially ponded stream." *Sci Total Environ* **390**(2-3): 448-455.

Wielgoss, S., J. E. Barrick, et al. (2011). "Mutation Rate Inferred From Synonymous Substitutions in a Long-Term Evolution Experiment With Escherichia coli." *G3 (Bethesda)* **1**(3): 183-186.

Zimmer, D. P., E. Soupene, et al. (2000). "Nitrogen regulatory protein C-controlled genes of Escherichia coli: scavenging as a defense against nitrogen limitation." *Proc Natl Acad Sci U S A* **97**(26): 14674-14679.

Appendix A: Maps for sampling beaches

Fig 1. Sampling locations at Bay Crystal Beach

Fig 2. Sampling locations at Bernard Avenue Beach

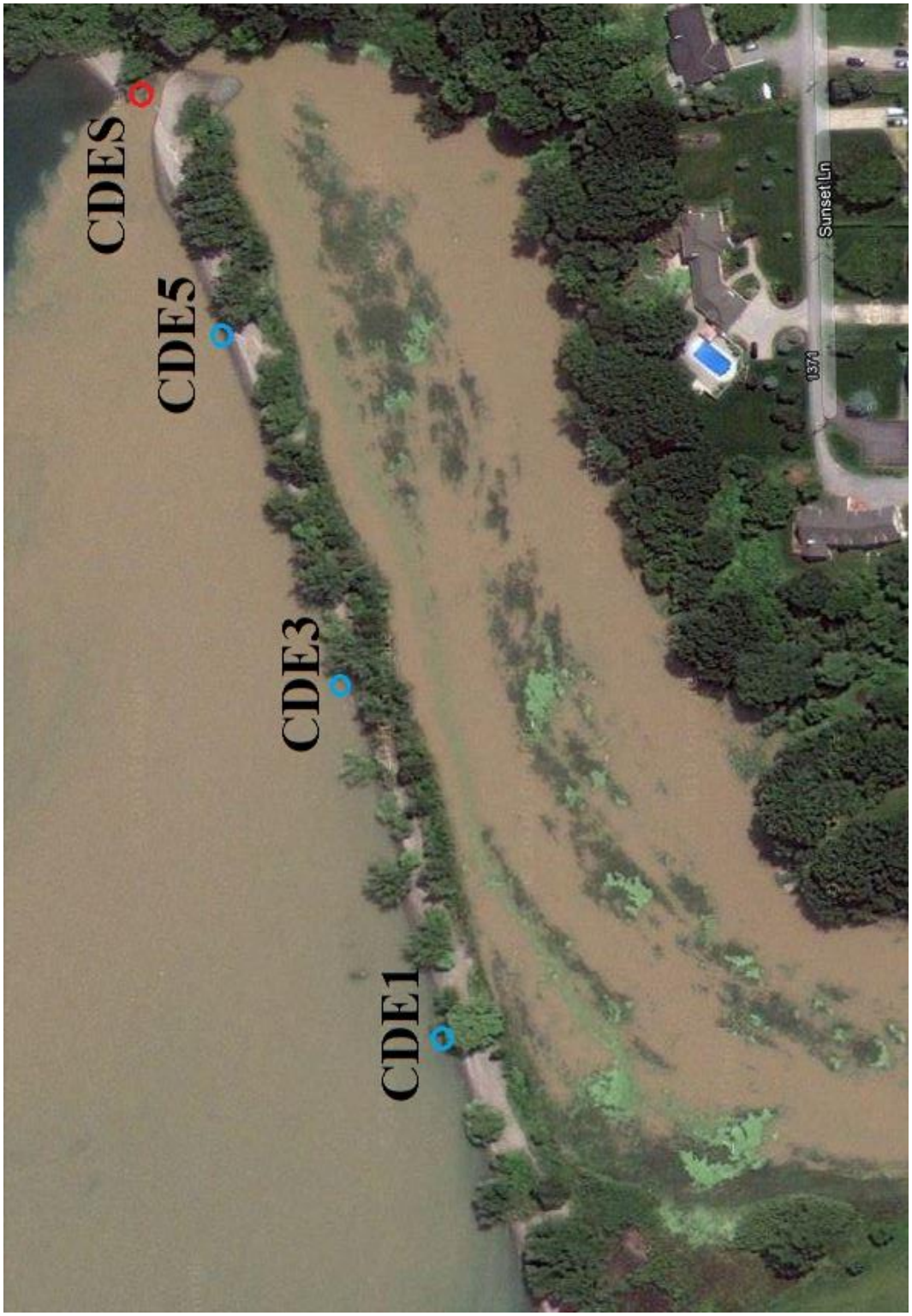


Fig 3. Sampling locations at Charles Daly Beach East Beach

Fig 4. Sampling locations at Charles Daly Beach West Beach

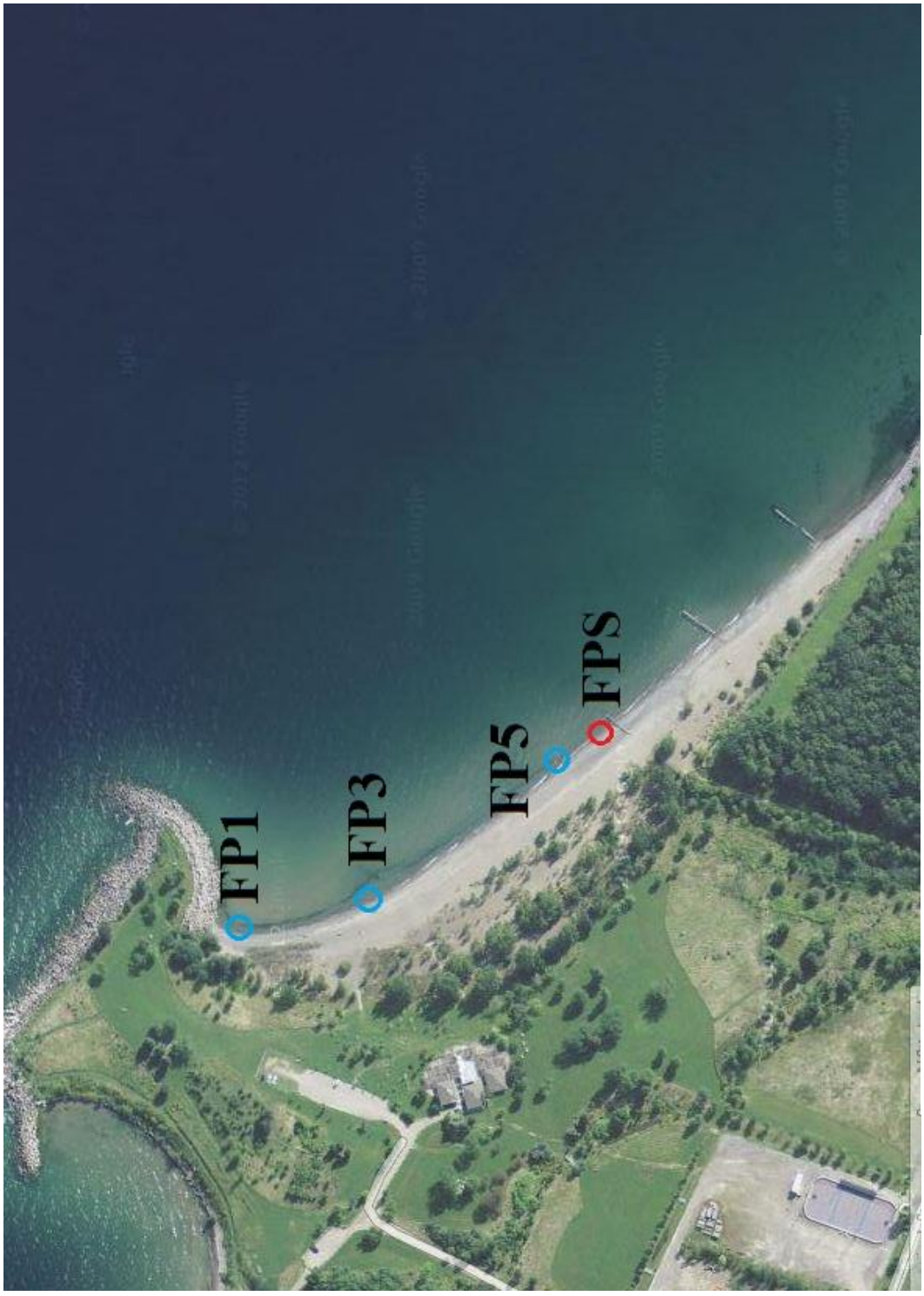


Fig 5. Sampling locations at Fifty Point Beach

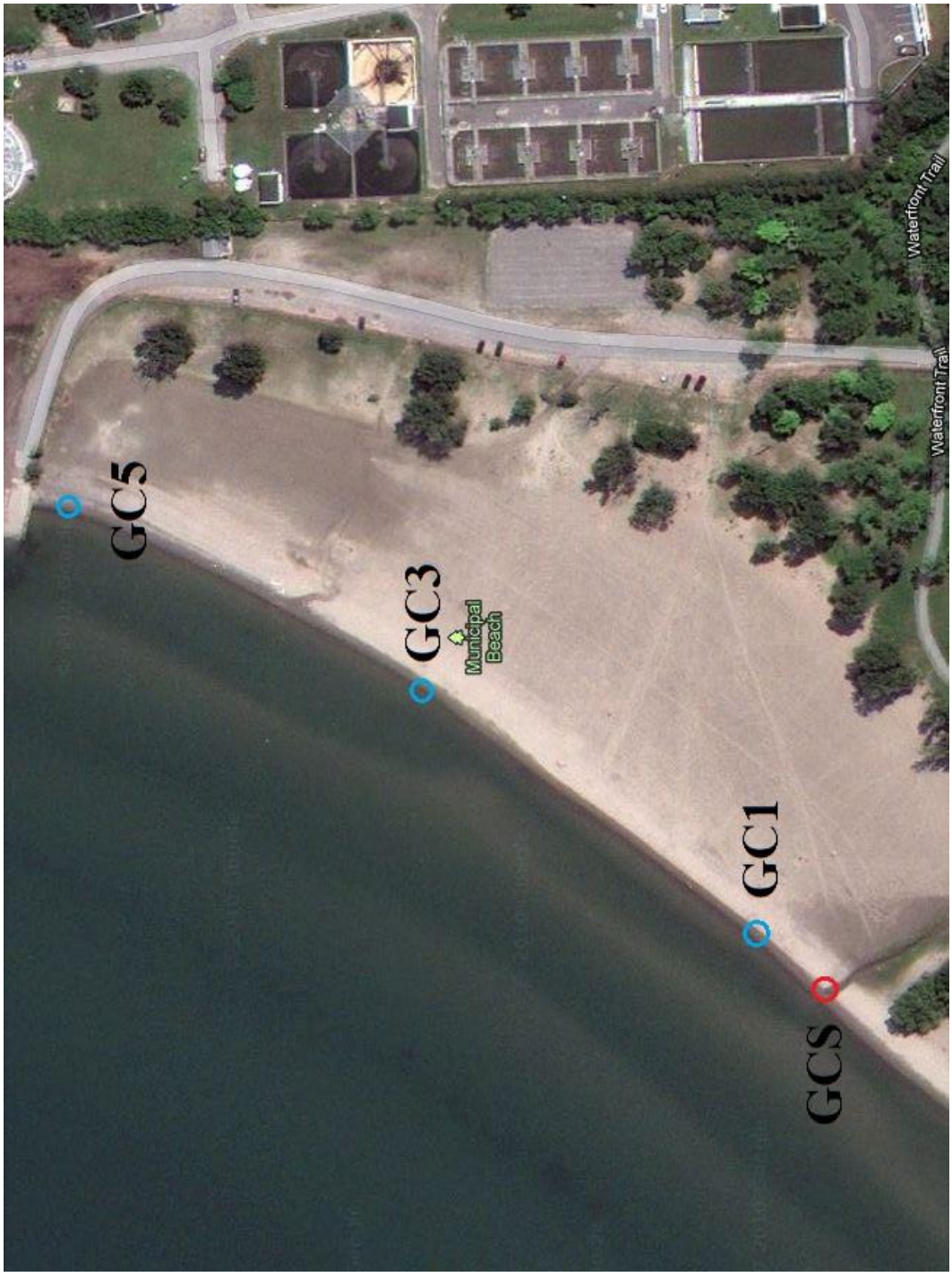


Fig 6. Sampling locations at Garden City Beach

Fig 7. Sampling locations at Centennial Park Beach

Fig 8. Sampling locations at Long Beach

Fig 9. Sampling locations at Long Beach Conservation East

Fig 10. Sampling locations at Long Beach Conservation West

Fig 11. Sampling locations at Larraine Road Beach

Fig 12. Sampling locations at Lakeside Beach

Fig 13. Sampling locations at Nickel Beach

Fig 14. Sampling locations at Nelles Park Beach

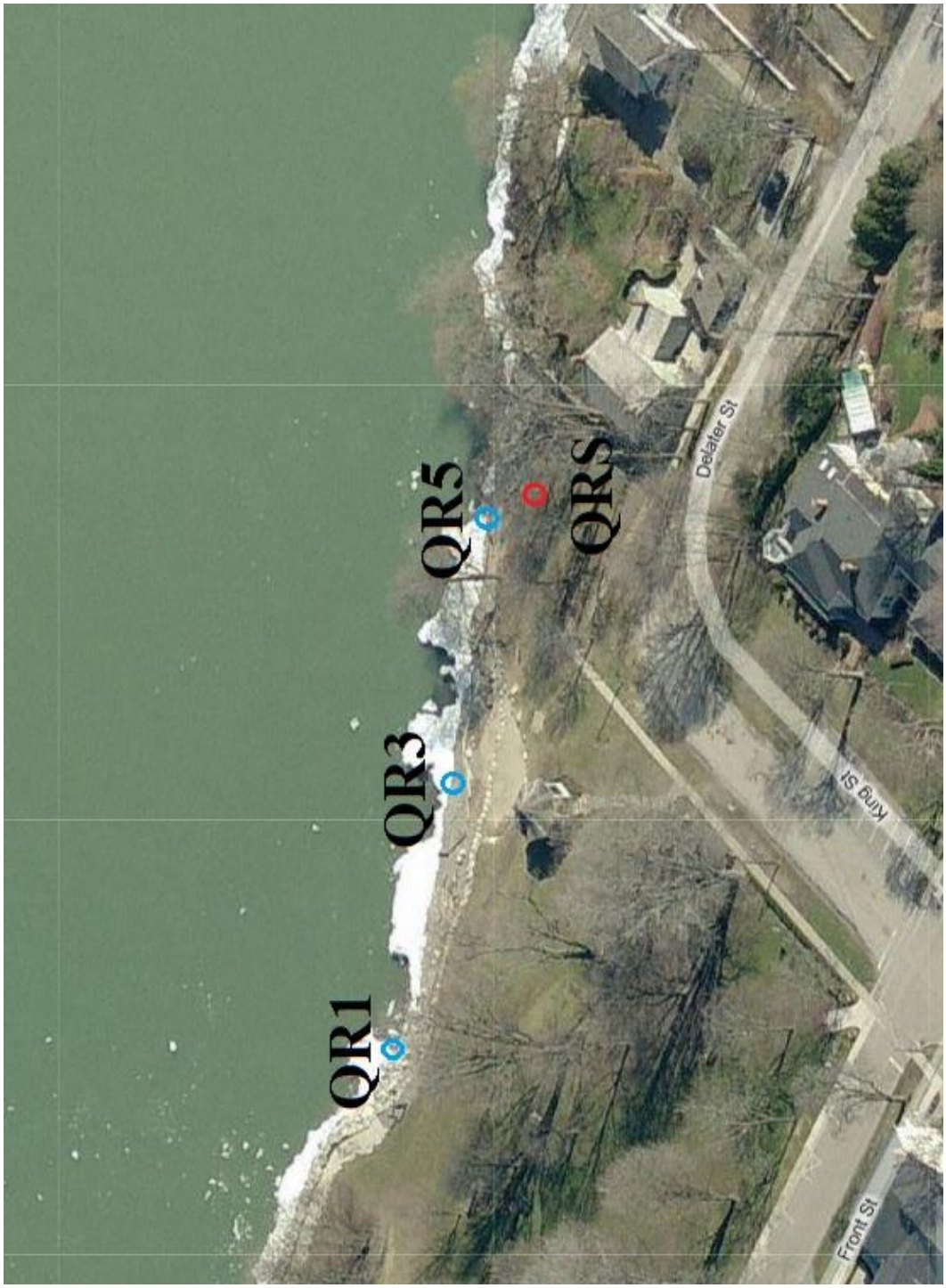


Fig 15. Sampling locations at Queen's Royal Beach

Appendix B. Statistical analysis results for Niagara water quality monitoring project

Table 1. Spearman correlation coefficient table for Fifty Point all-summer data in 2011.

Variable 1	Variable 2	r_s	P value
FP3	FP1	0.81	0.01
FP5	FP1	0.80	0.01
FP5	FP3	0.83	0.01
FPS	FP1	0.33	0.08
FPS	FP5	0.25	0.18
FPS	FP3	0.14	0.45

Table 2. Spearman correlation coefficient table for Bay Crystal all-summer data in 2011.

Variable 1	Variable 2	r_s	P value
BC3	BC1	0.91	0.01
BC5	BC1	0.82	0.01
BC5	BC3	0.83	0.01
BCS	BC1	0.10	0.75
BCS	BC5	0.03	0.92
BCS	BC3	0.02	0.94

Table 3. Spearman correlation coefficient table for GC all-summer data in 2011.

Variable 1	Variable 2	r_s	P value
GC3	GC1	0.77	0.01
GC5	GC1	0.67	0.01
GC5	GC3	0.73	0.01
GCS	GC1	0.80	0.80
GCS	GC3	0.20	0.20
GCS	GC5	0.20	0.20

Table 4. Spearman correlation coefficient table for LB all-summer data in 2011

Variable 1	Variable 2	r_s	P value
LB3	LB1	0.85	0.01
LBS1	LB3	0.50	0.17
LBS1	LB1	0.43	0.24
LB5	LB3	0.40	0.03
LBS1	LB5	0.37	0.33
LB5	LB1	0.25	0.18
LBS2	LBS1	0.22	0.57
LBS2	LB5	0.12	0.55
LBS2	LB3	0.09	0.63
LBS2	LB1	0.02	0.92

Table 5. Spearman correlation coefficient table for LCE all-summer data in 2011

Variable 1	Variable 2	r_s	P value
LCE5	LCE3	0.86	0.01
LCE5	LCE1	0.40	0.03
LCE3	LCE1	0.35	0.06
LCES	LCE1	0.26	0.36
LCES	LCE3	0.24	0.40
LCES	LCE5	0.23	0.43

Table 6. Spearman correlation coefficient table for LS all-summer data in 2011

Variable 1	Variable 2	r_s	P value
LS3	LS1	0.73	0.01
LSS	LS5	0.70	0.19
LS5	LS3	0.12	0.50
LSS	LS3	0.10	0.87
LSS	LS1	0.10	0.87
LS5	LS1	0.08	0.64

Table 7. Spearman correlation coefficient table for QR all-summer data in 2011.

Variable 1	Variable 2	r_s	P value
QR5	QR3	0.73	0.01
QR3	QR1	0.69	0.01
QR5	QR1	0.48	0.01
QRS	QR3	0.18	0.33
QRS	QR1	0.11	0.54
QRS	QR5	0.03	0.86

Table 8. Spearman correlation coefficient table for FP rain-day data in 2011 (n=5).

Variable 1	Variable 2	r_s	P value
FP5	FP3	1.00	0.01
FP5	FP1	0.80	0.10
FP3	FP1	0.80	0.10
FPS	FP1	0.70	0.19
FPS	FP5	0.30	0.62
FPS	FP3	0.30	0.62

Table 9. Spearman correlation coefficient table for BC rain-day data in 2011 (n=3).

Variable 1	Variable 2	r_s	P value
BC5	BC1	1.00	0.01
BCS	BC3	1.00	0.01
BC3	BC1	0.50	0.67
BC5	BC3	0.50	0.67
BCS	BC1	0.50	0.67
BCS	BC5	0.50	0.67

Table 10. Spearman correlation coefficient table for LB rain-day data in 2011 (n=3).

Variable 1	Variable 2	r_s	P value
LBS1	LB1	1.00	0.01
LBS2	LBS1	1.00	0.01
LB3	LB1	0.50	0.67
LB5	LB3	0.50	0.67
LBS1	LB5	0.50	0.67
LB5	LB1	0.50	0.67
LBS1	LB3	0.50	0.67
LBS2	LB5	0.50	0.67
LBS2	LB3	0.50	0.67
LBS2	LB1	0.50	0.67

Table 11. Spearman correlation coefficient table for LCE rain-day data in 2011 (n=4).

Variable 1	Variable 2	r_s	P value
LCES	LCE5	1.00	0.01
LCE3	LCE1	0.80	0.20
LCE5	LCE3	0.80	0.20
LCES	LCE1	0.50	0.67
LCES	LCE3	0.50	0.67
LCE5	LCE1	0.40	0.60

Table 12. Spearman correlation coefficient table for QR rain-day data in 2011 (n=5).

Variable 1	Variable 2	rs	P value
QRS	QR1	1.0	0.01
QRS	QR3	0.90	0.04
QR3	QR1	0.90	0.04
QR5	QR1	0.10	0.87
QRS	QR5	0.10	0.87
QR5	QR3	0.00	1.00

Appendix C: Protein gels

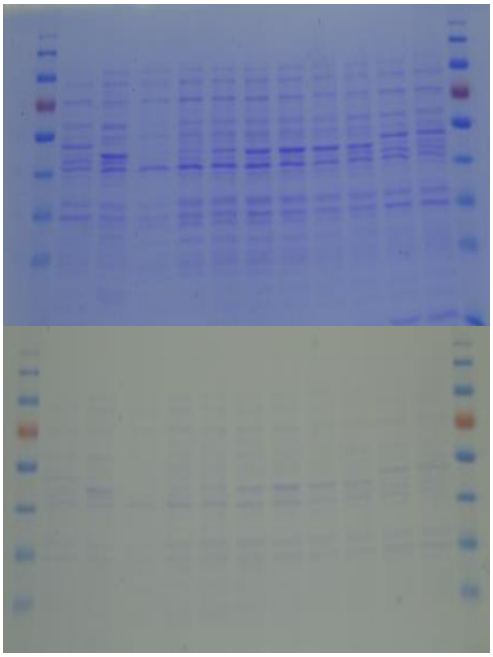


Figure 1. Protein gels for MG1655. Top: 3 hours destain; bottom: overnight destain.

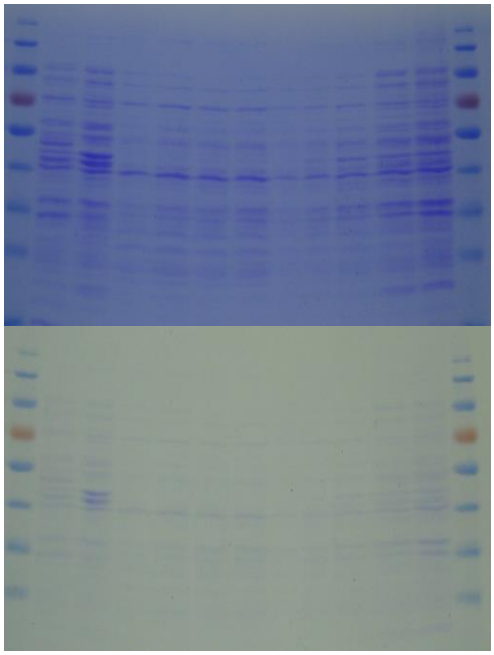


Figure 2. Protein gels for AZB07. Top: 3 hours destain; bottom: overnight destain.

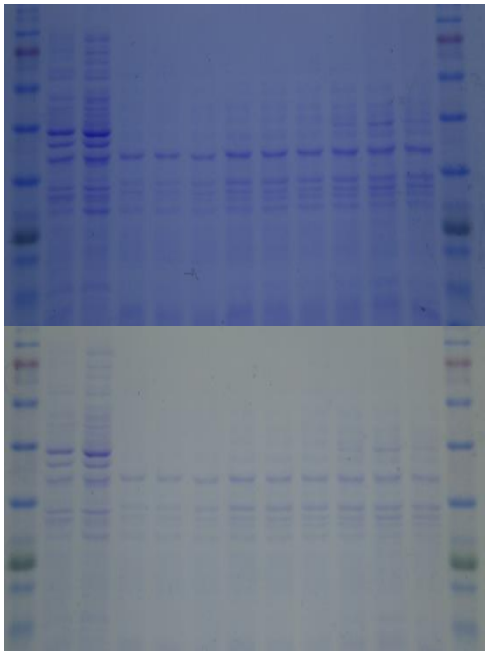


Figure 3. Protein gels for AZB10. Top: 3 hours destain; bottom: overnight destain.

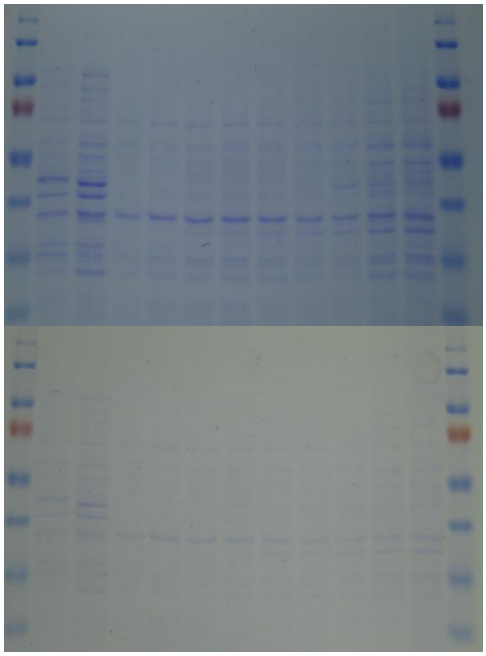


Figure 4. Protein gels for BNB03. Top: 3 hours destain; bottom: overnight destain.

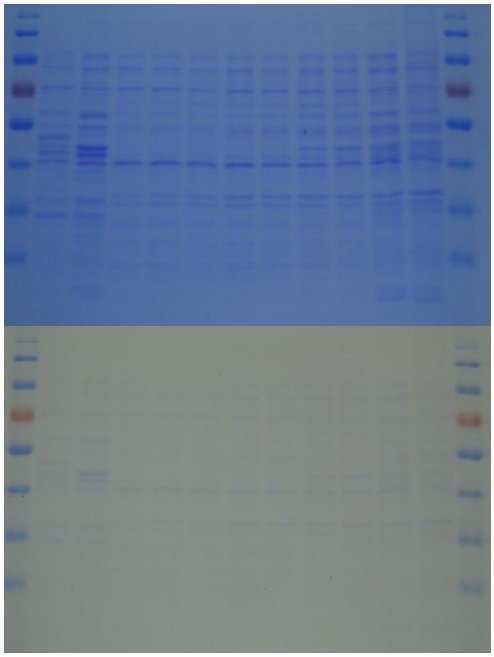


Figure 5. Protein gels for ECE12. Top: 3 hours destain; bottom: overnight destain.

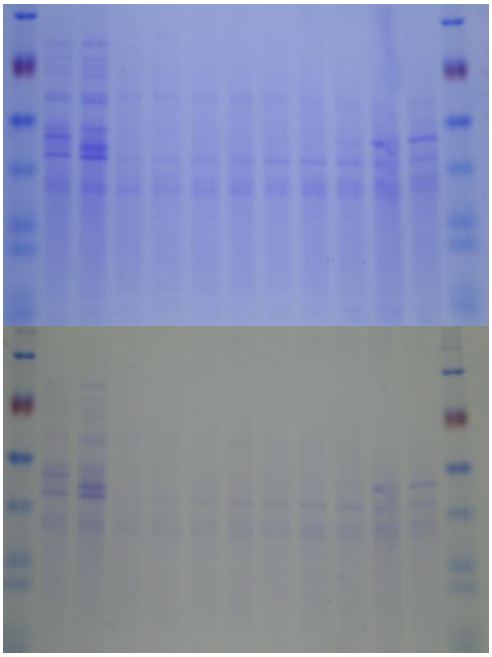


Figure 6. Protein gels for ECH01. Top: 3 hours destain; bottom: overnight destain.

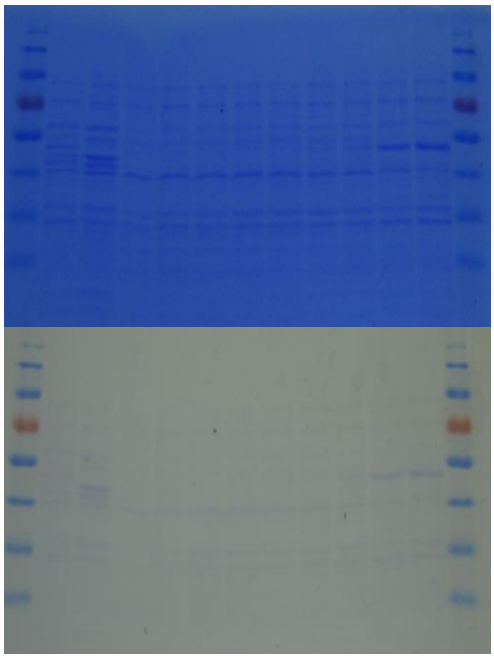


Figure 7. Protein gels for EDL933. Top: 2 hours destain; bottom: overnight destain.

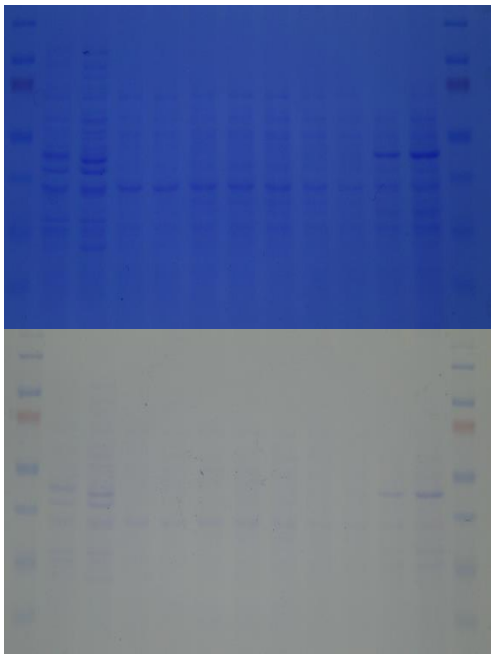


Figure 8. Protein gels for R4D8. Top: 3 hours destain; bottom: overnight destain

Appendix D: Standard Operation Procedures:

Immunoblot (Western blot):

Protein sample preparation

1. Streak out culture from -80 ℃ freezer on a 1X LB plate overnight.
2. Inoculate a single colony to 1X LB liquid culture with antibiotics; incubate at 37.5 ℃ with shaking at 200 rpm for overnight.
3. Subculture strain in a dilution at 1:10,000 into 50mL 1X LB, incubate at 37.5 ℃ with shaking at 200 rpm.
4. Take culture samples at desired OD_{600} points (K-12 MG1655 takes about 3 hours to get to OD_{600} = 0.1).
5. Transfer culture sample to a 1.5 mL Eppendorf tube, centrifuge for 5 min at 13,000g.
6. Slowly discard supernatant with a pipette tip.
7. Suspend sample in 1X SDS loading buffer.
8. Boil SDS culture sample in boiling water to lyse cell and denature proteins.
9. Store protein sample at -20 ℃ for short term storage (up to 3 weeks) and store at -80 ℃ for long term storage

SDS gel preparation

1. Clean glasses with dH_20 and 95% Ethanol, wipe to dry.
2. Assemble glasses in green holders.

3. Slowly pipette 4.7mL of resolving gel mixture into assembled glasses
4. Cover the top of the mixture with dH_20.
5. Let it solidify for 20 min.
6. Slowly remove water from the top with Kimwipe.
7. Slowly pipette stacking gel mixture into assembled glasses.
8. Place a cleaned comb into assembled glasses quickly.
9. Let it solidify for 20 min.

SDS-PAGE

1. After solidified, gently remove assembled glasses out of the green holder
2. Assemble Mini-Protean tetra apparatus.
3. Pour SDS running buffer into the apparatus.
4. Load 3μL of protein ladder and 10μg of protein into well.
5. Run gel at 50V for 40min until sample has reached resolving gel.
6. Run get at 100V for 50min until sample has reached the bottom.

Protein gel staining:

1. Slowly remove gel and place it in 0.1% Coomassie Blue for 1h.
2. Place gel in destain solution for 3h.
3. Take a picture.
4. Place gel back into destain solution overnight.
5. Take a picture.

Semi-dry transfer

1. Cut PVDF membrane into desired size (7*9cm)
2. Cut filter papers into desired size (8*10cm)
3. Soak PVDF membrane in methanol for exact 1min, and place it in transfer buffer immediately.
4. Place 6 filter papers (per gel) and gel in transfer buffer as well.
5. Create a sandwich cassette in semi-dry transfer machine (From Bottom to top: 3 filter papers, PVDF membrane, gel and 3 filter papers)
6. Run gel at 25V for 20min.

Label membrane:

1. Slowly take membrane out of the chamber, and block it in 5% milk (5g skim milk power in 100mL 1X TBST) for 1h.
2. Gently cut membrane down to a size that can fit in the 47mm Petri dish.
3. Block membrane with primary antibody (1:10,000 dilution, made in 5% milk) overnight.
4. Save primary antibody and store it at -20 ℃.
5. Wash membrane 3X with 1xTBST on shaker, 10min each.
6. Block membrane in secondary antibody (1:30,000 dilution, made in 5% milk) for 1h.
7. Save secondary antibody and store it at -20 ℃.
8. Wash membrane 3X with 1xTBST on shaker, 10min each.

9. Mix 5mL of reagent 1 and 5mL of reagent 2 shortly before using it (ECL Western blotting detection reagents)
10. Put membrane in the mix for 1min.
11. With a tweezers, gently drain the solution and seal the membrane in Saran wrap.

Visualization:

1. Place the membrane in a Kodak cassette
2. Expose GE Healthcare Amersham Hyperfilm ECL with membrane.

Reagents:

2X SDS loading buffer:

12.5mL 1M Tris-Cl (pH 6.8)

8.7mL 100% glycerol

2.5mL 2-mercaptoethanol

10mL 10% SDS

1mL 1% bromophenol blue

15.3 mL ddH20

10x TBST (500mL)

43.5g NaCl

50mL 1M Tris-Cl (pH 8), 2.5mL Tween 20

ddH20

10x transfer buffer (1L)

3.025g Tris

14.4g glycine

200mL methanol

ddH_20

10x running buffer (1L)

30g Tris

144g clycine

ddH_20

(add 5mL 10% SDS to 500mL 1x running buffer for SDS gels)

0.1% Coomassie blue (500mL)

0.5g Coomassie

200mL methanol

50mL acetic acid

Filer through a 0.45μm filter

Destain solution (1L)

450mL methanol

100mL acetic acid

ddH20

Resolving gel mix (per 2 gels)

4mL ddH20

2.5mL 1.5M Tris-Cl (pH 8.8)

0.1mL 10% SDS

3.3mL 30% Acry/Bis

0.1mL 10% APS

0.01mL TEMES

Stacking gel mix (per 2 gels)

3.6mL ddH_20

0.63mL 1M Tris-Cl (pH6.8)

0.05mL 10% SDS

0.66mL 30% Acry/Bis

0.05mL 10% APS

0.005mL TEMES

Appendix E: Defence presentation

Integrative microbial contamination assessment for water quality monitoring in the Great Lakes

Wenjie Zheng
M.Sc. Defence

Department of Biology
McMaster University

- Part I: Water quality monitoring in Niagara Region in 2010 and 2011
- Part II: Sigma factor expression in *E. coli* environmental isolates

Part I: Water quality monitoring in Niagara Region

- Recreational beaches - attracting tourists
- Waterborne pathogens - main disease causing element (WHO, 2004)
- Fecal pollution - major source of waterborne pathogens (WHO, 2003)
- It is important to identify and quantify fecal contamination in the water, especially human health related microbes

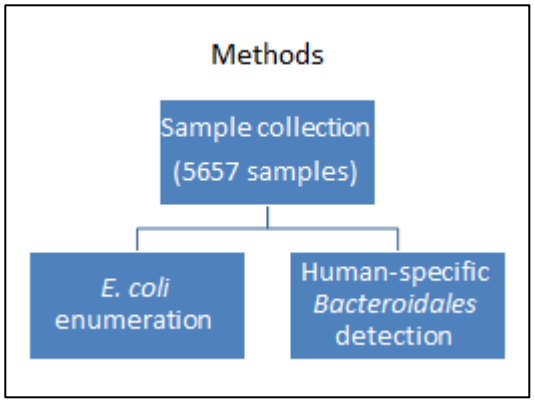
Microbial Source Tracking

- Library-dependent method
- Microbial water quality indicator: e.g., *E. coli*
  - Ontario: 100CFU/100 mL (Ministry of Environment and Energy, 1994)
- Library-independent method
- Host specific microorganism: e.g., human-specific *Bacteroidales* (Bernhard et al., 2000; Seurinck et al., 2005)

Objectives

- 1. Validate a microbial contamination assessment system for water quality monitoring
- 2. Examine the impact of pollution sources on beach water
  - *E. coli* pollution from nearby point sources of pollution
  - Human fecal pollution at individual beaches

Sampling Sites



Beaches – Potential Risk Classification[1] in 2010

	High human fecal contamination (≥ 10% of days showed human fecal positive)	Low human fecal contamination (< 10% of days showed human fecal positive)
High *E. coli* concentration (≥20% beach closure)	Queen's Royal Garden City Long Beach Con. East Lakeside Lorraine Road Nelle's Park	Fifty Point Humberstone Centennial Charles Daly Beach West
Low *E. coli* concentration (<20% beach closure)	Bay Crystal Charles Daly Beach East	Long Beach Long Beach Con. West Nickel Beach Bernard

(1) Edge et al., 2010

Beaches – Potential Risk Classification in 2011

	High human fecal contamination (≥ 10% of days showed human fecal positive)	Low human fecal contamination (< 10% of days showed human fecal positive)
High *E. coli* concentration (≥20% beach closure)	Queen's Royal Garden City Lakeside Lorraine Road Fifty Point Bay Crystal	Long Beach Con. East
Low *E. coli* concentration (<20% beach closure)	Long Beach	

Impact of pollution sources on beach water

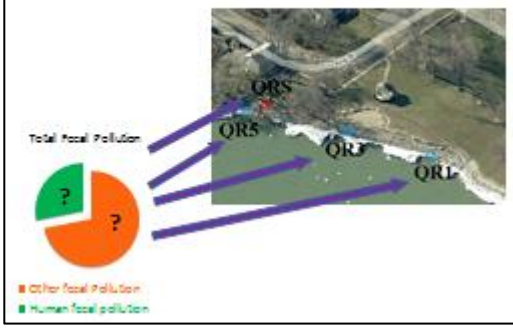
1. Correlations of *E. coli* concentration between transects and pollution sources at Queen's Royal (QR) Beach

Variable 1	Variable 2	r²
QR5	QR1	1.0
QR5	QR3	0.90
QR3	QR1	0.90
QR5	QR1	0.10
QR5	QR5	0.10
QR5	QR3	0.00

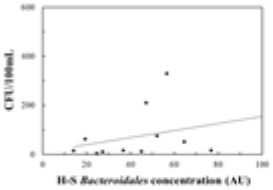
Variable 1	Variable 2	r²
QR5	QR3	0.79
QR3	QR1	0.69
QR5	QR1	0.49
QR5	QR3	0.19
QR5	QR1	0.11
QR5	QR5	0.02

Impact of pollution sources on beach water

2. Correlation between *E. coli* concentration and human fecal concentration



Correlation between *E. coli* concentration and human-specific *Bacteroidales* concentration



Correlations at transects:

QR1 $y = 1.66x + 10.80$ $R^2 = 0.33$

QR3 $y = 3.31x + 3.83$ $R^2 = 0.14$

QR5 $y = 0.07x + 113.01$ $R^2 = 0.0004$

QR3 $y = 33.81x^{0.44}$ $R^2 = 0.88$

Figure 1. Correlation between *E. coli* concentration (CFU/100mL) and human-specific (H-S) *Bacteroidales* concentration at Queen's Royal (QR) Transect 1 in Arbitrary Units (AU).

Part I Conclusion

- Integrative microbial contamination assessment system wasvalidated
- The localized impacts from pollution sources at individual sites were determined

Part II: Sigma factor expression in *E. coli* environmental isolates

- *E. coli* strains differ enormously in pathogenic potential as only few strains carry disease-causing genetic determinants (Vogt and Dippold, 2005)
- The importance of *E. coli* to human health issues is thus affected by genetic information (DNA) and gene expression in *E. coli* strains in addition to microbial numbers (Tauschek et al., 2002)

- This study examined the differences in genetic composition and gene expression between laboratory strains and environmental strains
- RpoS: the general stress response sigma factor (Hengge-Aronis, 2002)
- RpoE: the extracytoplasmic/extreme heat stress sigma factor (Lipinska et al., 1990)
- RpoD: the "housekeeping" sigma factor (Shiozawa et al., 1996)
- In different environments, the expression of sigma factors may be changed (Chiang et al., 2011;Dong and Schellhorn, 2009)

Objective

- Understand how gene expression is altered in *E. coli* environmental strains (wild type and *rpoS* mutants) and laboratory strains by analyzing differences in RpoS, RpoE and RpoD expression levels throughout growth

Table 2. *E. coli* isolates used in this study

Isolate	Location	Type	Type of mutation (*rpoS*)	Source
K-12 MG1655	N/A	Laboratory	WT	Schellhorn Lab
AZ807	Hamilton	Sediment Core	G--> A transition	(Chiang et al., 2011)
AZ810	Hamilton	Sediment Core	WT	(Chiang et al., 2011)
BNB03	Toronto	Sediment Core	1 bp deletion	(Chiang et al., 2011)
ECE12	Main and King	Combined Sewage Flow	5 bp deletion	(Chiang et al., 2011)
ECH01	Hamilton Harbour	Beach	WT	(Chiang et al., 2011)
EDL933	N/A	Pathogenic	WT	Schellhorn Lab
R4D8	Charles Daly West	Storm Water Outfall	A--> T transversion	CDW, sequenced by W. Zheng

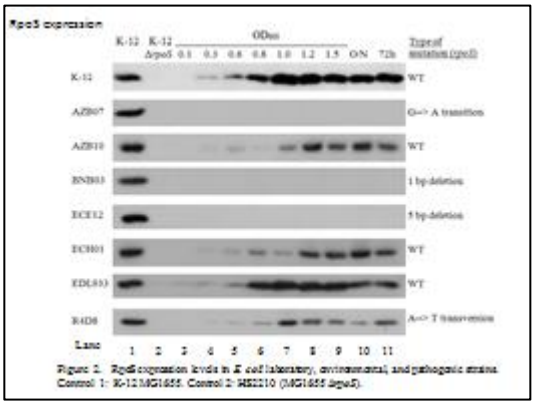


Figure 2. RpoS expression levels in *E. coli* laboratory, environmental, and pathogenic strains. Control 1: K-12 MG1655. Control 2: RH2210 (MG1655 *ΔrpoS*).

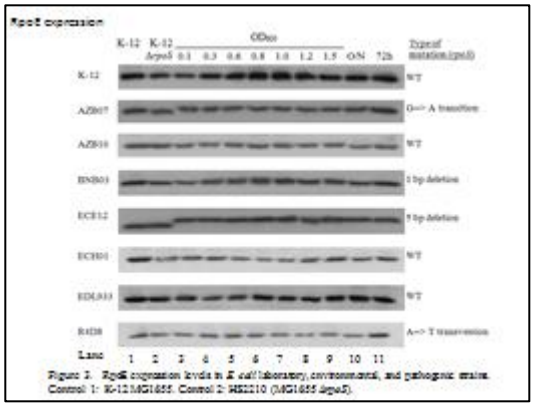


Figure 3. RpoS expression levels in *E. coli* laboratory, environmental, and pathogenic strains. Control 1: K-12 MG1655. Control 2: HS2210 (MG1655 *ΔrpoS*).

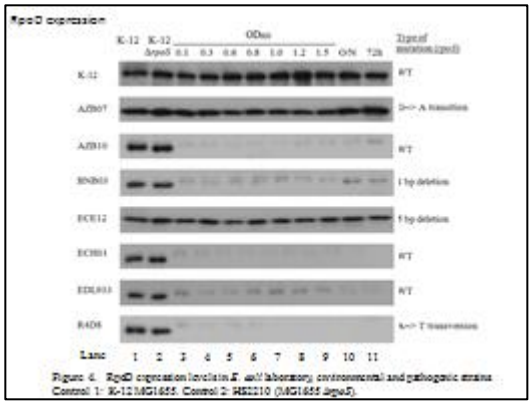


Figure 4. RpoD expression levels in *E. coli* laboratory, environmental and pathogenic strains. Control 1: K-12 MG1655. Control 2: HS2210 (MG1655 *ΔrpoS*).

Conclusion

- Differential gene expression in RpoS and RpoD were observed between *E. coli* environmental strains and commonly used laboratory strains
- Suggests that genes regulated by RpoD and RpoS have differential expression between *E. coli* environmental strains and commonly used laboratory strains

Future work

- Use animal-specific *Bacteroidales* probes to determine the impact of animal feces on beach water quality in the Region of Niagara

Acknowledgement

- Dr. Herb E. Schellhorn
- Dr. Tom A. Edge and Dr. Richard A. Morton
- Lab members
- WaterSmart Niagara & Mitacs for funding & scholarship support

References


Thank you ☺