

**Identifying and Assessing Bacterial Contamination in Stormwater
Runoff and Receiving Streams**

**A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville**

Songyi Liu

August 2021

Copyright © 2021 by Songyi Liu
All rights reserved

ACKNOWLEDGEMENTS

I owe my deepest gratitude to my supervisor, Dr. Qiang He, for his continuous support, encouragement, inspiration, and guidance in my research work. Without him, I would not have made it through my Ph.D. degree.

I would like to extend my sincerest gratitude to my committee members, Dr. John Schwartz, Dr. Shuai Li, and Dr. Jun Lin, for their excitement and interest in my projects. Their advice and perspective helped me develop an in-depth understanding of my research topic.

I would also like to express my gratitude to my colleagues in the Ph.D. program, Dr. Liu Cao, Dr. Yabing Li, Dr. Si Chen, Dr. Kristen Wyckoff, Dr. Lu Yang, Dr. Yongfeng Wang, Charles Cianciolo, Clifford Swanson, Kailong Ye, and Cole Emmett, for their companionship and assistance during this Ph.D. adventure.

I also thank my cooperators, Dr. Adrian Gonzalez from Water Quality Core Facility; Da Hu and Yifang Xu from Dr. Li's lab; Daniel Williams from the Center for Environmental Biotechnology; Yongchao Xie and Yanchen Sun from Dr. Loffler's lab, for their generous help in my research.

Lastly, I'd like to thank my wife and parents for their love, patience, and continuous support.

ABSTRACT

Stormwater runoff has been recognized as a significant non-point source of pollution responsible for serious ecological impacts. Microbial pollutants represent one of the major contaminants of concern in stormwater runoff and the receiving streams. Traditionally, fecal indicator bacteria (FIB) have been used extensively to assess the microbial quality of water, particularly the extent of fecal contamination in water. However, the application of FIB as indicators of microbial contamination bears known disadvantages of cultivation-dependent techniques, such as the inability to identify the majority of microbial populations that are not amenable to cultivation under laboratory conditions. The effectiveness of a cultivation-independent sequencing technique targeting the 16S rRNA genes was evaluated by comparison with FIB enumeration. Tested with samples from various environmental matrices, including stream water, stormwater runoff, sewage, and stream sediments, the efficacy of the sequencing techniques was validated with the ability to profile the microbial community present in stormwater runoff and stream water. Comparison of microbial contamination in stormwater and receiving streams indicates that roadway traffic characteristics and seasonal effects had considerable impacts on the extent of microbial contamination in stormwater runoff and subsequently the receiving streams.

The survival of *E. coli* strains isolated from pavement runoff, as representatives of microbial contaminants of concern, was investigated under relevant environmental conditions, showing that these microorganisms exhibited limited persistence on asphalt pavement and polystyrene plastic surfaces with the maximum T_{90} averaging 0.25 days. However, the presence of moisture was shown to significantly prolong the viability of these strains, suggesting the potential persistence of these microbial populations in the stormwater runoff and stream water.

Natural variations of the bacterioplankton communities in the streams were evaluated. The results showed while spatial and temporal variations were observed, the stream sizes and characteristics were larger drivers for stream microbial community. Further, the impact of stormwater runoff onto microbial communities in receiving streams was investigated. In stream wetflow, the microbes sourced from the stormwater runoff were determined, as well as the overall contributions that stormwater runoff had on the microbial water quality of streams. Findings in this study could be helpful for developing stormwater runoff control measures.

TABLE OF CONTENTS

Chapter 1	1
Introduction and Literature Review	1
1. Runoff impact	1
2. Microbial contamination	3
3. Persistence and Decay of FIB	4
3.1 Sunlight	5
3.2 Temperature	5
3.3 Alternative Habitats	6
3.4 sources of the FIB	6
4. Microbial community analyses to diagnose pollution conditions	6
References	9
Chapter 2	21
Evaluating the Accuracy of EPA Method 1604 for <i>E. coli</i> and Total Coliform Enumeration by 16S rRNA Gene Sequencing	21
Abstract	21
1. Introduction	21
2. Materials and Methods	23
2.1 Environmental matrices selection	23
2.2 Isolate collection	23
2.3 DNA Extraction and High-Throughput 16S rRNA Amplicon Sequencing	24
2.4 Sequence Analysis	24
2.5 Estimation of method effectiveness	25
3. Results and Discussions	25
3.1 Diversity of Putative TC isolates	25
3.2 Bacterial isolates responsible in <i>E. coli</i>	27
3.3 Bacterial isolates responsible in non- <i>E. coli</i> TC	27
3.4 Effectiveness of the culture-based testing method	29
4. Conclusions	31
References	34
Chapter 3	37
Comparison of the Presence and Behavior of Fecal Indicator Bacteria and Turbidity in Stormwater Runoff	37
Abstract	37
1. Introduction	37
2. Materials and Methods	39
2.1 Sample Collection	39
2.2 Sample Analysis	39
2.3 Statistical analysis	39
3. Results and Discussion	41
3.1 Correlation among water quality parameters	41
3.2 The impact of rainfall characteristics	41

3.3 Seasonal variations.....	43
3.4 The impact of site features	47
4. Conclusions.....	47
References.....	51
Chapter 4	54
Persistence of Environmental <i>E. coli</i> Strains from Stormwater Runoff.....	54
Abstract.....	54
1. Introduction.....	55
2. Materials and Methods.....	56
2.1 Study site.....	56
2.2 Preparation of bacterial strains and inoculum.....	57
2.3 Bacterial Inoculation	57
2.4 Controlled environmental chamber.....	58
2.5 Microbial recovery and enumerating	58
2.6 Statistical analysis and data modeling	59
3. Results and Discussion	60
3.1 The low persistence conditions	60
3.2 Desiccation resistance and persistence of <i>E. coli</i> on surfaces.....	60
3.3 Persistence of <i>E. coli</i> in sufficient water.....	66
4. Conclusions.....	68
References.....	70
Chapter 5	74
Natural Stabilities and Dynamics in Bacterioplankton Communities in the Unimpaired Streams within an Ecoregion	74
Abstract.....	74
1. Introduction.....	74
2. Materials and Methods.....	76
2.1 Sites selection.....	76
2.2 Fecal indicator bacteria test.....	76
2.3 DNA extraction and high-throughput 16S rRNA amplicon sequencing	76
2.4 Bioinformatic and statistical analyses.....	76
3. Results and Discussions.....	78
3.1 Fecal indicator bacteria in the unimpaired streams.....	78
3.2 Species abundance distributions for the bacterioplankton communities	80
3.3 Short period fluctuations in bacterioplankton community structure within a season	80
3.4 Short period fluctuations in the abundances of individual ASVs within a season	82
3.5 Seasonal shifts in bacterioplankton community compositions at each site.	85
3.6 Habitat associations for bacterioplankton	85
4. Conclusions.....	91
References.....	92
Chapter 6	95
Microbial Diversity in Unimpaired Streams across Ecoregions.....	95
Abstract.....	95

1. Introduction.....	95
2. Materials and Methods.....	96
2.1 Sites selection.....	96
2.2 Fecal indicator bacteria test.....	97
2.3 DNA extraction and high-throughput 16S rRNA amplicon sequencing	97
2.4 Bioinformatic and statistical analyses.....	97
3. Results and Discussions.....	99
3.1 Fecal indicator bacteria in the ecoregion streams	99
3.2 Habitat filtering by stream order level, and not ecoregion, is the major driver shaping microbial structures.....	99
3.3 The abundances of individual ASVs across stream sites.....	102
3.4 Putative source environments for the microbial community in the reference stream water	102
3.5 spatial factors related to microbial community composition.....	105
4. Conclusions.....	107
References.....	108
Chapter 7	110
The Impact of Storm Events on Receiving Stream Water Bacterial Community.....	110
Abstract.....	110
1. Introduction.....	110
2. Materials and Methods.....	111
2.1 Sites selection.....	111
2.2 FIB and turbidity test	113
2.3 DNA extraction and high-throughput 16S rRNA amplicon sequencing	113
2.4 Bioinformatic and statistical analyses.....	113
3. Results and Discussions.....	114
3.1 FIB and Turbidity in stormwater runoff and receiving stream water	114
3.2 Bacterial community structure in the receiving stream baseflow, wetflow, and stormwater runoff.....	116
3.3 The abundances of individual ASVs in the receiving stream baseflow, wetflow, and stormwater runoff.....	119
3.4 Sourcetracker analysis	123
4. Conclusions.....	127
References.....	128
Conclusion	131
Appendix.....	133
VITA.....	136

Chapter 1

Introduction and Literature Review

1. Runoff impact

Freshwater ecosystems, such as streams, rivers, lakes, and wetlands, are valuable public resources, provide drinking water, wildlife habitat, recreational and tourism opportunities, water for agriculture and industrial uses, and so on (Cope 1966; Cabelli, Dufour, et al., 1982). Protection of natural water resources will also protect ecosystems, human health, and the economy. US public and private actors spend over \$35 billion total per year on reducing surface water pollution, and the cost is higher than most other US environmental programs (Keiser, Kling, et al., 2019).

Stormwater runoff may be responsible for serious ecological impacts, especially for roadway runoff which contains suspended solids, metal elements, pathogens, organic and inorganic compounds (Stotz 1987; Barrett, Irish Jr, et al., 1998; Han, Lau, et al., 2006; Lau, Han, et al., 2009). As a significant non-point source of pollution, runoff is mainly discharged into the receiving water. For example, streams, lakes, and coastal beaches were frequently reported for contaminants after storm events (Arnold and Gibbons 1996; Gnecco, Berretta, et al., 2005; Flint and Davis 2007; Klimaszewska, Polkowska, et al., 2007; Kayhanian, Stransky, et al., 2008).

The first flush of runoff is a mass flushing of pollutants in the initial period, which may contain higher concentrations of pollutants than those observed during the latter stages of the stormwater runoff (Vorreiter and Hickey 1994; Sansalone and Buchberger 1997; Deletic 1998; Kayhanian and Stenstrom 2005; Mangani, Berloni, et al., 2005; Kayhanian and Stenstrom 2008). Therefore, numerous researches focused on the first flush effects over seasons or based on individual storm events. In general, the first flush impacted by numerous parameters, such as impervious area, watershed area, rainfall intensity, antecedent dry period (Kayhanian, Singh, et al., 2003; Shinya, Tsuruho, et al., 2003; Westerlund, Viklander, et al., 2003; Kayhanian, Stransky, et al., 2008); and first flush effect with high concentrations has been studied for water quality, chemical constituents, solid particle size distribution and toxicity (Kayhanian and Stenstrom 2005; Maniquiz, Lee, et al., 2010).

The Water Pollution Control Act of 1972, which emphasized the goal of fishable and swimmable waters, underpinned necessary regulatory action. And all states should control the discharge of

pollutants of stormwater runoff into receiving water bodies (Athayde, Shelley, et al., 1983). In order to conform to discharge regulations, the total maximum daily load (TMDL) of impaired water bodies and water quality standards were required to compute the pollutant loading (Bickford, Toll, et al., 1999). A decrease in stormwater contamination can be accomplished by source control. In any case, source control alone may not wipe out the discharged contamination loading, and information on water quality of stormwater runoff is required to accomplish cost-effective treatments of contamination decreasing (Lloyd, Wong, et al., 2002; Jiang 2004; Kay, Bartram, et al., 2004; Kayhanian, Regenmorte, et al., 2004; Abrishamchi, Massoudieh, et al., 2010).

Conventional water quality parameters, metal constituents, toxicity, polycyclic aromatic hydrocarbons (PAHs), and nutrients in stormwater runoff were studied frequently (Bourcier and Hindin 1979; Cabelli, Dufour, et al., 1982; Stenstrom, Silverman, et al., 1984; Makepeace, Smith, et al., 1995; Bickford, Toll, et al., 1999; Haile, Witte, et al., 1999; Andrews and Sutherland 2004; Aryal, Furumai, et al., 2005; Gnecco, Berretta, et al., 2005; Lau, Han, et al., 2009; Davis and Birch 2010). Conventional water quality parameters include total suspended solids (TSS), volatile suspended solids (VSS), total organic carbon, chemical oxygen demand (COD), total dissolved solids (TDS), dissolved organic carbon (DOC), biochemical oxygen demand (BOD), oil and grease (O&G), hardness as CaCO₃, temperature, and pH (Campos, Avant, et al.,). Some variations of water quality parameters were observed among different researches, and the inconsistency might be related to many factors, including rain intensity, land use, sample collection methods, and analytical methods (Stotz 1987; Helsel and Cohn 1988; Driscoll, Shelley, et al., 1990; Thomson, McBean, et al., 1997; Bertrand-Krajewski, Chebbo, et al., 1998; Deletic and Maksimovic 1998).

Metal constituents include aluminum (Al), lead (Pb) the metalloid arsenic (As), cadmium (Cd), chromium (Cr), nickel (Ni), iron (Fe), and zinc (Zn), copper (Cu) (Bourcier and Hindin 1979; Roger, Montrejaud-Vignoles, et al., 1998; Sansalone, Koran, et al., 1998; Wu, Allan, et al., 1998; Bäckström, Nilsson, et al., 2003; Kayhanian and Stenstrom 2005; Lau and Stenstrom 2005; Hallberg, Renman, et al., 2007). One crucial observation is the inconstancy of these metal concentrations. For example, the concentrations of Cu, Pb, and Zn in North America were lower than in European countries (Pagotto, Legret, et al., 2000).

Nutrients monitored from different studies include total nitrogen (TN), nitrates (NO₃⁻), nitrites (NO₂⁻), ammonium (NH₄⁺), total Kjeldahl nitrogen (TKN), total phosphorus (TP), and phosphate

(PO_4^{3-}) (Barrett, Kearfott, et al., 2006; Han, Lau, et al., 2006). Nitrogen and phosphorus constituents can be changed within the environment in different forms, with a considerable effect (Mangani, Berloni, et al., 2005; White and Boswell 2006; Berhanu Desta, Bruen, et al., 2007; Eriksson, Baun, et al., 2007). Overall, the sources of N and P in runoff may be from traffic and non-traffic sources, and the non-traffic sources like soil and vegetation from surrounding land uses are the major contributions (Flint and Davis 2007; Yufen, Xlaoke, et al., 2008; Maniquiz, Lee, et al., 2010). Pathogens, herbicides, and pesticides were less frequently tested in stormwater runoff, though they are important parameters and studied regularly in receiving water bodies (An, Kampbell, et al., 2002; Federigi, Verani, et al., 2019).

Finding solid connections among water quality parameters and constituents would decrease the number of monitoring parameters and save analytical expenses. Based on previous studies, there are huge variants for the correlations between the water quality parameters and constituents (Lau, Han, et al., 2009; Davis and Birch 2010). Under some specific conditions and certain management programs, the higher correlations ($R^2 > 0.8$) among parameters can potentially decrease analytical costs. For example, M. Kayhanian showed four parameters (TSS, TDS, Fe, TOC) might be monitored to serve as other thirteen water quality parameters (turbidity, TKN, O&G, TPH, Cu, DOC, Ni, Pb, Zn, EC, Cd, Cl, Cr) (Kayhanian, Suverkropp, et al., 2007).

2. Microbial contamination

Microbial contaminants, such as pathogenic bacteria, viruses, protozoa, and antibiotic resistant genes, have been concerned with the contamination of water bodies (Ashbolt and Bruno 2003; Federigi, Verani, et al., 2019). Studies have reported a high prevalence of microbial contaminants in sewer overflows, defective septic systems, roof-harvested rainwater, and agricultural runoff transported to and Polluted water bodies (Ihssen, Grasselli, et al., 2007; Paule-Mercado, Ventura, et al., 2016; Olds, Corsi, et al., 2018). However, the studies of microbial contamination in roadway stormwater runoff were rarely (Saget, Chebbo, et al., 1996; Jiang, Noble, et al., 2001; Thurston, Gerba, et al., 2001; Ahmed, Hamilton, et al., 2018; Boehm, Graham, et al., 2018; Hamilton, Reyneke, et al., 2019).

It is impossible to monitor all pathogens in water bodies for the determination of microbial contamination because of the great diversity of pathogens. Therefore, fecal indicator bacteria (FIB), like total coliform, fecal coliforms, *Escherichia coli*, and *enterococci*, were selected to assess the

microbial quality due to high levels in feces and likely presences of potential pathogens (Sidhu, Gernjak, et al., 2012).

High concentrations of FIB (>1,000 CFU/100 mL *Escherichia coli*, and >10,000 CFU/100 mL total coliform) are frequently found in stormwater runoff and receiving water bodies (Chong, Sidhu, et al., 2013; Olds, Corsi, et al., 2018; Steele, Blackwood, et al., 2018). And of course, most of the concentrations exceeded the threshold value of FIB for the water quality standards by one or more orders of magnitude (Murphy, Meng, et al., 2017). Moreover, studies have reported that the loadings of FIB over wet weather conditions in receiving waters were higher than base flow over dry weather conditions (Tseng and Jiang 2012; Boehm, Graham, et al., 2018).

Sources of stormwater FIB are believed nonhuman unless sewer or septic system of human origin contaminating the stormwater. Domestic and wildlife animals deposit feces that wash from soils or paved surfaces (Hörman, Rimhanen-Finne, et al., 2004; Noble, Griffith, et al., 2006; Brownell, Harwood, et al., 2007; Field and Samadpour 2007; Rajal, McSwain, et al., 2007; Stoeckel and Harwood 2007; Ahmed, Huygens, et al., 2008; Ahmed, Stewart, et al., 2008; Davies, Mitchell, et al., 2008; McQuaig, Scott, et al., 2009; Parker, McIntyre, et al., 2010). Urban or high-density residential areas are found that FIB levels in receiving waters over wet weather conditions are higher compared to rural or sparsely residential areas due to high concentrations of FIB in runoff from sewer and septic overflows, commercial markets, and impervious cover (Haile, Witte, et al., 1999; Young and Thackston 1999; Thurston, Gerba, et al., 2001; Lloyd, Wong, et al., 2002; McCarthy, Mitchell, et al., 2006; Struck, Selvakumar, et al., 2008; Page, Dillon, et al., 2010; Soller, Schoen, et al., 2010; Stumpf, Piehler, et al., 2010). For example, Young and Thackston showed that *E. coli* and fecal coliform in streams were related to the densities of development, housing, population, impervious area, and domestic animals (Young and Thackston 1999).

3. Persistence and Decay of FIB

FIB has been used to present potential fecal contamination and pathogens of enteric origins (Grant 1997; Haile, Witte, et al., 1999). However, high concentrations of FIB were found in some habitats where recent fecal materials were rarely reported, which suggests specific FIB populations survived for extended periods outside of feces (Hoff and Akin 1986; England, Lee, et al., 1993; Sjogren 1994). Understanding the extent to FIB which can survive in natural environments is essential to estimating the impacts of fecal contamination events.

The characteristics of natural habitats can contribute to rapid decay or extended survival or even growth of FIB, including sunlight, nutrients, temperature, and environmental types (e.g., liquid types and solid surfaces) (Buchanan and Edelson 1999; Himathongkham, Bahari, et al., 1999; Tao, Bausch, et al., 1999; Rozen and Belkin 2001; Ihssen and Egli 2004; Franz, Van Diepeningen, et al., 2005; Large, Walk, et al., 2005; Bergholz and Whittam 2007; Ihssen, Grasselli, et al., 2007; Semenov, Van Bruggen, et al., 2007; Tabe, Oloya, et al., 2008; Vital, Hammes, et al., 2008; Sidhu, Toze, et al., 2015). Moreover, different types of FIB are shown variant survival abilities under the same environmental conditions (Girvan, Campbell, et al., 2005; Large, Walk, et al., 2005; Fisher and Meakins 2006; Taylor 2008; Sidhu, Toze, et al., 2015).

3.1 Sunlight

FIB that in natural environments are faced with numerous environmental pressures, including sunlight. Multiple studies found the impact of sunlight is significant for FIB (Zhou, Wang, et al., 2015).

The mechanisms of bacterial inactivation by sunlight, including UVA (315 to 400 nm) -induced photo-oxidative damage, and UVB (280 to 315 nm) -induced mutations of pyrimidine dimers (Boyle, Sichel, et al., 2008). Photo-oxidative damage for bacterial inactivation can be either exogenous or endogenous scenarios. In the exogenous mechanism, the sources of reactive oxygen species and free radicals are present in the natural habitats; on the opposite, in the endogenous scenario, the radiations are absorbed by internal cellular components, which is able to make damaging compounds inside the cell (Ahmed, Richardson, et al., 2014; Forte Giacobone and Oppezzo 2015; Korajkic, Wanjugi, et al., 2019).

3.2 Temperature

FIB are commonly mesophiles, and the optimal growth temperatures are from 20°C to 40°C, which consistent with the temperature characteristic of gastrointestinal tracts in warm-blood animals (Savageau 1983). In many natural systems, the temperature fluctuates on a daily and seasonal basis. In nutrient-rich environments, higher temperatures result in elevated metabolic activities, but the survival of FIB tends to be prolonged at lower temperatures when nutrients are deficient (Klein and Alexander 1986; Buchanan and Edelson 1999; Tao, Bausch, et al., 1999; Semenov, Van Bruggen, et al., 2007).

3.3 Alternative Habitats

The types of microbial habitats differ widely and affect the persistence of FIB. In solid surfaces, moisture content or water activity can be one of the critical parameters for the survival of FIB. Extreme drought might result in cell death (Gagliardi and Karns 2000; Sang Ho, Baumler, et al., 2000; Rozen and Belkin 2001; Jiang, Morgan, et al., 2002; Natvig, Ingham, et al., 2002; Topp, Welsh, et al., 2003; Bach, Stanford, et al., 2005).

In liquid habitats, nutrients are generally detrimental to the survival of FIB (Klein and Alexander 1986; Rozen and Belkin 2001). For example, in eutrophic aquatic habitats, FIB can survive longer or even facilitate proliferation (Ihssen and Egli 2004). However, nutrient scarcity in environments showed reduced chances of survival of FIB or pathogens (Klein and Alexander 1986).

3.4 sources of the FIB

The diversity and variability of FIB showed different capabilities to respond to the effects of a given stressor (Ihssen and Egli 2004; Schubert, Dufke, et al., 2004). In previous studies, FIB isolated from animal feces (dog, bovine, deer, pig, or goose feces) or sewage and human feces typically decayed more rapidly than those from natural environments, like surface water, soil, and sediments (Large, Walk, et al., 2005; Bergholz and Whittam 2007; Oliver, Clegg, et al., 2007; Habteselassie, Bischoff, et al., 2008; Saleh-Lakha, Shannon, et al., 2008; Tabe, Oloya, et al., 2008; Harwood, Staley, et al., 2014; Oladeinde, Bohrmann, et al., 2014).

4. Microbial community analyses to diagnose pollution conditions

Although FIB tests were commonly used in stormwater management programs, one of the limitations of FIB is that they rarely have a strong correlation with the presence of pathogens in the same samples (Sidhu, Ahmed, et al., 2013). This may be expected, as FIB in environmental samples are sourced from feces, including humans and other animals, but human pathogens mainly derived from human feces (Bichai and Ashbolt 2017; Steele, Blackwood, et al., 2018). Moreover, it may be significantly different in the decay rates of FIB and other pathogens. Even selective media showed somewhat incorrect rates for presenting the culturable FIB (Federigi, Verani, et al., 2019). Hence, monitoring of FIB in terms of pathogenic pollutions may not yield meaningful outcomes.

The rapid improvement of nucleic acid sequencing technologies has shown the capacity to characterize the microorganisms in complex environmental matrices (Loman, Misra, et al., 2012; Kumar, Menon, et al., 2013; Salipante, Sengupta, et al., 2013; Sunagawa, Mende, et al., 2013; Sinclair, Osman, et al., 2015). Advances in next-generation sequencing revealed new opportunities for pathogenic contamination assessment through analysis of microbial communities (Shendure, Porreca, et al., 2005; Wang, Garrity, et al., 2007; Lazarevic, Whiteson, et al., 2009; Claesson, Wang, et al., 2010; McLellan, Huse, et al., 2010; Kim and Bae 2011; Ashbolt 2015; Eren, Morrison, et al., 2015; Madoui, Engelen, et al., 2015; Nshimiyimana, Freedman, et al., 2015; Federigi, Verani, et al., 2019).

The high-throughput 16S rRNA amplicon sequencing is one of the culture-independent methods used for alternative microbiome analysis (Liu, Lozupone, et al., 2007; Kennedy, Hall, et al., 2014; Logares, Sunagawa, et al., 2014; Schloss, Westcott, et al., 2015; Sinclair, Osman, et al., 2015). Most microorganisms contain the 16S rRNA gene, which is in conjunction with nine hypervariable regions flanked by conserved genomics (Peer, Neefs, et al., 1993). It provides an opportunity to identify the microbial taxonomy of the species associated with different environmental matrices based on designing generic PCR primers to amplify and sequence the hypervariable regions (Chakravorty, Helb, et al., 2007; Liu, Desantis, et al., 2008; Humblot and Guyot 2009).

One of the advantages of utilizing the high-throughput 16S rRNA amplicon sequencing approach is that the numerous bioinformatics tools can be available for free and easy operating, including MOTHUR and QIIME, which are the two most used software packages. Based on the nucleotide sequence similarity, the sequences are clustered into Operational Taxonomic Units (OTU) by bioinformatic tools; or based on exact amplicon sequence variant was observed in each sample, the end result can be an amplicon sequence variant (ASV) table, which is a higher-resolution than the OTU table (Schloss, Westcott, et al., 2009; Caporaso, Kuczynski, et al., 2010; Unno, Jang, et al., 2010). ASVs or OTUs are then identified as the microorganisms compared against databases (Quast, Pruesse, et al., 2013).

For microbial quality assessment, microbial composition analyses have been used to yield insights into potential biological risk. A recent study of bacterial contamination environment samples applied the high-throughput 16S rRNA amplicon sequencing approach is to track sources and proportions of contamination in marker-gene based on a Bayesian approach, which is called

SourceTracker (Knights, Kuczynski, et al., 2011). The method takes contamination sources as entire source communities into the contaminated communities, which are the sink communities, where the mixing proportions are trying to be determined.

Although NGS-based approaches have the potential to be an alternative microbial quality assessment, current frameworks heavily rely upon FIB measurements (Sales-Ortells and Medema 2015; Harmel, Hathaway, et al., 2016; Federigi, Verani, et al., 2019). More further studies will be needed to assess NGS-based methods for pollution detection.

References

- Abrishamchi, A., A. Massoudieh and M. Kayhanian (2010). "Probabilistic modeling of detention basins for highway stormwater runoff pollutant removal efficiency." Urban Water Journal **7**(6): 357-366.
- Ahmed, W., K. A. Hamilton, A. Lobos, B. Hughes, C. Staley, M. J. Sadowsky and V. J. Harwood (2018). "Quantitative microbial risk assessment of microbial source tracking markers in recreational water contaminated with fresh untreated and secondary treated sewage." Environment International **117**: 243-249.
- Ahmed, W., F. Huygens, A. Goonetilleke and T. Gardner (2008). "Real-time PCR detection of pathogenic microorganisms in roof-harvested rainwater in Southeast Queensland, Australia." Applied and Environmental Microbiology **74**(17): 5490-5496.
- Ahmed, W., K. Richardson, J. P. S. Sidhu, P. Jagals and S. Toze (2014). "Inactivation of faecal indicator bacteria in a roof-captured rainwater system under ambient meteorological conditions." Journal of Applied Microbiology **116**(1): 199-207.
- Ahmed, W., J. Stewart, D. Powell and T. Gardner (2008). "Evaluation of Bacteroides markers for the detection of human faecal pollution." Letters in Applied Microbiology **46**(2): 237-242.
- An, Y. J., D. H. Kampbell and G. Peter Breidenbach (2002). "Escherichia coli and total coliforms in water and sediments at lake marinas." Environmental Pollution **120**(3): 771-778.
- Andrews, S. and R. A. Sutherland (2004). "Cu, Pb and Zn contamination in Nuuanu watershed, Oahu, Hawaii." Science of the Total Environment **324**(1-3): 173-182.
- Arnold, C. L. and C. J. Gibbons (1996). "Impervious Surface Coverage: The Emergence of a Key Environmental Indicator." Journal of the American Planning Association **62**(2): 243-258.
- Aryal, R. K., H. Furumai, F. Nakajima and M. Boller (2005). "Dynamic behavior of fractional suspended solids and particle-bound polycyclic aromatic hydrocarbons in highway runoff." Water Research **39**(20): 5126-5134.
- Ashbolt, N. J. (2015). "Microbial Contamination of Drinking Water and Human Health from Community Water Systems." Current environmental health reports **2**(1): 95-106.
- Ashbolt, N. J. and M. Bruno (2003). "Application and refinement of the WHO risk framework for recreational waters in Sydney, Australia." Journal of Water and Health **1**(3): 125-131.
- Athayde, D. N., P. E. Shelley, E. D. Driscoll, D. Gaboury and G. Boyd (1983). "Results of the nationwide urban runoff program" Results of the Nationwide Urban Runoff Program.
- Bäckström, M., U. Nilsson, K. Håkansson, B. Allard and S. Karlsson (2003). "Speciation of heavy metals in road runoff and roadside total deposition." Water, Air, and Soil Pollution **147**(1-4): 343-366.
- Bach, S. J., K. Stanford and T. A. McAllister (2005). "Survival of Escherichia coli O157:H7 in feces from corn- and barley-fed steers." FEMS Microbiology Letters **252**(1): 25-33.
- Barrett, M. E., L. B. Irish Jr, J. F. Malina Jr and R. J. Charbeneau (1998). "Characterization of highway runoff in Austin, Texas, area." Journal of Environmental Engineering **124**(2): 131-137.

Barrett, M. E., P. Kearfott and J. F. Malina Jr (2006). "Stormwater quality benefits of a porous friction course and its effect on pollutant removal by roadside shoulders." Water Environment Research **78**(11): 2177-2185.

Bergholz, T. M. and T. S. Whittam (2007). "Variation in acid resistance among enterohaemorrhagic *Escherichia coli* in a simulated gastric environment." Journal of Applied Microbiology **102**(2): 352-362.

Berhanu Desta, M., M. Bruen, N. Higgins and P. Johnston (2007). "Highway runoff quality in Ireland." Journal of Environmental Monitoring **9**(4): 366-371.

Bertrand-Krajewski, J. L., G. Chebbo and A. Saget (1998). "Distribution of pollutant mass vs volume in stormwater discharges and the first flush phenomenon." Water Research **32**(8): 2341-2356.

Bichai, F. and N. Ashbolt (2017). "Public health and water quality management in low-exposure stormwater schemes: A critical review of regulatory frameworks and path forward." Sustainable Cities and Society **28**: 453-465.

Bickford, G., J. Toll, J. Hansen, E. Baker and R. Keessen (1999). "Aquatic ecological and human health risk assessment of chemicals in wet weather discharges in the Sydney region, New South Wales, Australia." Marine Pollution Bulletin **39**(1-12): 335-345.

Boehm, A. B., K. E. Graham and W. C. Jennings (2018). "Can We Swim Yet? Systematic Review, Meta-Analysis, and Risk Assessment of Aging Sewage in Surface Waters." Environmental Science and Technology **52**(17): 9634-9645.

Bourcier, D. R. and E. Hindin (1979). "Lead, iron, chromium and zinc in road runoff at Pullman, Washington." Science of the Total Environment, The **12**(3): 205-215.

Boyle, M., C. Sichel, P. Fernández-Ibáñez, G. B. Arias-Quiroz, M. Iriarte-Puña, A. Mercado, E. Ubomba-Jaswa and K. G. McGuigan (2008). "Bactericidal effect of solar water disinfection under real sunlight conditions." Applied and Environmental Microbiology **74**(10): 2997-3001.

Brownell, M. J., V. J. Harwood, R. C. Kurz, S. M. McQuaig, J. Lukasik and T. M. Scott (2007). "Confirmation of putative stormwater impact on water quality at a Florida beach by microbial source tracking methods and structure of indicator organism populations." Water Research **41**(16): 3747-3757.

Buchanan, R. L. and S. G. Edelson (1999). "Effect of pH-dependent, stationary phase acid resistance on the thermal tolerance of *Escherichia coli* O157:H7." Food Microbiology **16**(5): 447-458.

Cabelli, V. J., A. P. Dufour, L. J. McCabe and M. A. Levin (1982). "Swimming-associated gastroenteritis and water quality." American Journal of Epidemiology **115**(4): 606-616.

Campos, C. J. A., J. Avant, N. Gustar, J. Lowther, A. Powell, L. Stockley and D. N. Lees (2015). "Fate of Human Noroviruses in Shellfish and Water Impacted by Frequent Sewage Pollution Events." Environmental Science and Technology **49**(14): 8377-8385.

Caporaso, J. G., J. Kuczynski, J. Stombaugh, K. Bittinger, F. D. Bushman, E. K. Costello, N. Fierer, A. G. Peña, J. K. Goodrich, J. I. Gordon, G. A. Huttley, S. T. Kelley, D. Knights, J. E. Koenig, R. E. Ley, C. A. Lozupone, D. McDonald, B. D. Muegge, M. Pirrung, J. Reeder, J. R. Sevinsky, P. J.

- Turnbaugh, W. A. Walters, J. Widmann, T. Yatsunenko, J. Zaneveld and R. Knight (2010). "QIIME allows analysis of high-throughput community sequencing data." Nature Methods **7**(5): 335-336.
- Chakravorty, S., D. Helb, M. Burday, N. Connell and D. Alland (2007). "A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria." Journal of Microbiological Methods **69**(2): 330-339.
- Chong, M. N., J. Sidhu, R. Aryal, J. Tang, W. Gernjak, B. Escher and S. Toze (2013). "Urban stormwater harvesting and reuse: A probe into the chemical, toxicology and microbiological contaminants in water quality." Environmental Monitoring and Assessment **185**(8): 6645-6652.
- Claesson, M. J., Q. Wang, O. O'Sullivan, R. Greene-Diniz, J. R. Cole, R. P. Ross and P. W. O'Toole (2010). "Comparison of two next-generation sequencing technologies for resolving highly complex microbiota composition using tandem variable 16S rRNA gene regions." Nucleic Acids Research **38**(22).
- Cope, O. B. (1966). "Contamination of the freshwater ecosystem by pesticides." Journal of Applied Ecology: 33-44.
- Davies, C. M., V. G. Mitchell, S. M. Petterson, G. D. Taylor, J. Lewis, C. Kaucner and N. J. Ashbolt (2008). Microbial challenge-testing of treatment processes for quantifying stormwater recycling risks and management. Water Science and Technology. **57**: 843-847.
- Davis, B. and G. Birch (2010). "Comparison of heavy metal loads in stormwater runoff from major and minor urban roads using pollutant yield rating curves." Environmental Pollution **158**(8): 2541-2545.
- Deletic, A. (1998). "The first flush load of urban surface runoff." Water Research **32**(8): 2462-2470.
- Deletic, A. B. and Č. T. Maksimovic (1998). "Evaluation of water quality factors in storm runoff from paved areas." Journal of Environmental Engineering **124**(9): 869-879.
- Driscoll, E. D., P. E. Shelley and E. W. Strecker (1990). "Pollutant Loadings and Impacts From Highway Stormwater Runoff Volume III: Analytical Investigation and Research Report." Pollutant Loadings and Impacts from Highway Stormwater Runoff. Vol. III: Analytical Investigation and Research Report.
- England, L. S., H. Lee and J. T. Trevors (1993). "Bacterial survival in soil: Effect of clays and protozoa." Soil Biology and Biochemistry **25**(5): 525-531.
- Eren, A. M., H. G. Morrison, P. J. Lescault, J. Reveillaud, J. H. Vineis and M. L. Sogin (2015). "Minimum entropy decomposition: Unsupervised oligotyping for sensitive partitioning of high-throughput marker gene sequences." ISME Journal **9**: 968-979.
- Eriksson, E., A. Baun, L. Scholes, A. Ledin, S. Ahlman, M. Revitt, C. Noutsopoulos and P. S. Mikkelsen (2007). "Selected stormwater priority pollutants - a European perspective." Science of the Total Environment **383**(1-3): 41-51.
- Federigi, I., M. Verani, G. Donzelli, L. Cioni and A. Carducci (2019). "The application of quantitative microbial risk assessment to natural recreational waters: A review." Marine Pollution Bulletin **144**: 334-350.

Field, K. G. and M. Samadpour (2007). "Fecal source tracking, the indicator paradigm, and managing water quality." Water Research **41**(16): 3517-3538.

Fisher, I. S. and S. Meakins (2006). "Surveillance of enteric pathogens in Europe and beyond: Enter-net annual report for 2004." Euro surveillance : bulletin européen sur les maladies transmissibles = European communicable disease bulletin **11**(8).

Flint, K. R. and A. P. Davis (2007). "Pollutant mass flushing characterization of highway stormwater runoff from an ultra-urban area." Journal of Environmental Engineering **133**(6): 616-626.

Forte Giacobone, A. F. and O. J. Oppezzo (2015). "Survival of *Pseudomonas aeruginosa* exposed to sunlight resembles the phenom of persistence." Journal of Photochemistry and Photobiology B: Biology **142**: 232-236.

Franz, E., A. D. Van Diepeningen, O. J. De Vos and A. H. C. Van Bruggen (2005). "Effects of cattle feeding regimen and soil management type on the fate of *Escherichia coli* O157:H7 and *Salmonella enterica* serovar typhimurium in manure, manure-amended soil, and lettuce." Applied and Environmental Microbiology **71**(10): 6165-6174.

Gagliardi, J. V. and J. S. Karns (2000). "Leaching of *Escherichia coli* O157:H7 in diverse soils under various agricultural management practices." Applied and Environmental Microbiology **66**(3): 877-883.

Girvan, M. S., C. D. Campbell, K. Killham, J. I. Prosser and L. A. Glover (2005). "Bacterial diversity promotes community stability and functional resilience after perturbation." Environmental Microbiology **7**(3): 301-313.

Gnecco, I., C. Berretta, L. G. Lanza and P. La Barbera (2005). "Storm water pollution in the urban environment of Genoa, Italy." Atmospheric Research **77**(1-4 SPEC. ISS.): 60-73.

Grant, M. (1997). "A new membrane filtration medium for simultaneous detection and enumeration of *Escherichia coli* and total coliforms." Appl. Environ. Microbiol. **63**(9): 3526-3530.

Hörman, A., R. Rimhanen-Finne, L. Maunula, C. H. Von Bonsdorff, N. Torvela, A. Heikinheimo and M. L. Hänninen (2004). "Campylobacter spp., Giardia spp., Cryptosporidium spp., Noroviruses, and Indicator Organisms in Surface Water in Southwestern Finland, 2000-2001." Applied and Environmental Microbiology **70**(1): 87-95.

Habteselassie, M., M. Bischoff, E. Blume, B. Applegate, B. Reuhs, S. Brouder and R. F. Turco (2008). "Environmental controls on the fate of *Escherichia coli* in soil." Water, Air, and Soil Pollution **190**(1-4): 143-155.

Haile, R. W., J. S. Witte, M. Gold, R. Cressey, C. McGee, R. C. Millikan, A. Glasser, N. Harawa, C. Ervin, P. Harmon, J. Harper, J. Dermand, J. Alamillo, K. Barrett, M. Nides and G. Y. Wang (1999). "The health effects of swimming in ocean water contaminated by storm drain runoff." Epidemiology **10**(4): 355-363.

Hallberg, M., G. Renman and T. Lundbom (2007). "Seasonal variations of ten metals in highway runoff and their partition between dissolved and particulate matter." Water, Air, and Soil Pollution **181**(1-4): 183-191.

- Hamilton, K. A., B. Reyneke, M. Waso, T. Clements, T. Ndlovu, W. Khan, K. GiGiovanni, E. Rakestraw, F. Montalto, C. N. Haas and W. Ahmed (2019). "A global review of the microbiological quality and potential health risks associated with roof-harvested rainwater tanks." NPJ Clean Water **2**(7).
- Han, Y., S. L. Lau, M. Kayhanian and M. K. Stenstrom (2006). "Characteristics of highway stormwater runoff." Water Environment Research **78**(12): 2377-2388.
- Harmel, R. D., J. M. Hathaway, K. L. Wagner, J. E. Wolfe, R. Karthikeyan, W. Francesconi and D. T. McCarthy (2016). "Uncertainty in monitoring E. coli concentrations in streams and stormwater runoff." Journal of Hydrology **534**: 524-533.
- Harwood, V. J., C. Staley, B. D. Badgley, K. Borges and A. Korajkic (2014). "Microbial source tracking markers for detection of fecal contamination in environmental waters: Relationships between pathogens and human health outcomes." FEMS Microbiology Reviews **38**(1): 1-40.
- Helsel, D. R. and T. A. Cohn (1988). "Estimation of descriptive statistics for multiply censored water quality data." Water Resources Research **24**(12): 1997-2004.
- Himathongkham, S., S. Bahari, H. Riemann and D. Cliver (1999). "Survival of Escherichia coli O157:H7 and Salmonella typhimurium in cow manure and cow manure slurry." FEMS Microbiology Letters **178**(2): 251-257.
- Hoff, J. C. and E. W. Akin (1986). "Microbial resistance to disinfectants: Mechanisms and significance." Environmental Health Perspectives **Vol. 69**: 7-13.
- Humblot, C. and J. P. Guyot (2009). "Pyrosequencing of tagged 16S rRNA gene amplicons for rapid deciphering of the microbiomes of fermented foods such as pearl millet slurries." Applied and Environmental Microbiology **75**(13): 4354-4361.
- Ihssen, J. and T. Egli (2004). "Specific growth rate and not cell density controls the general stress response in Escherichia coli." Microbiology **150**(6): 1637-1648.
- Ihssen, J., E. Grasselli, C. Bassin, P. François, J. C. Piffaretti, W. Köster, J. Schrenzel and T. Egli (2007). "Comparative genomic hybridization and physiological characterization of environment isolates indicate that significant (eco-)physiological properties are highly conserved in the species Escherichia coli." Microbiology **153**(7): 2052-2066.
- Jiang, S. (2004). "Is Urban Runoff a Source of Human Pathogenic Viruses to Recreational Beach Waters? Technical Completion Report, Project No. W-943."
- Jiang, S., R. Noble and W. Chu (2001). "Human adenoviruses and coliphages in urban runoff-impacted coastal waters of Southern California." Applied and Environmental Microbiology **67**(1): 179-184.
- Jiang, X., J. Morgan and M. P. Doyle (2002). "Fate of Escherichia coli O157:H7 in manure-amended soil." Applied and Environmental Microbiology **68**(5): 2605-2609.
- Kay, D., J. Bartram, A. Prüss, N. Ashbolt, M. D. Wyer, J. M. Fleisher, L. Fewtrell, A. Rogers and G. Rees (2004). "Derivation of numerical values for the World Health Organization guidelines for recreational waters." Water Research **38**(5): 1296-1304.

- Kayhanian, M., L. C. Regenmorte and K. Tsay (2004). Characteristics of snowmelt runoff from highways in the Tahoe basin and treatment investigations for improving runoff quality. Transportation Research Record, National Research Council: 112-122.
- Kayhanian, M., A. Singh, C. Suverkropp and S. Borroum (2003). "Impact of annual average daily traffic on highway runoff pollutant concentrations." Journal of Environmental Engineering **129**(11): 975-990.
- Kayhanian, M. and M. K. Stenstrom (2005). Mass loading of first flush pollutants with treatment strategy simulations. Transportation Research Record: 133-143.
- Kayhanian, M. and M. K. Stenstrom (2008). "First flush highway runoff characterization: for stormwater runoff treatment." Stormwater **9**(2): 32-45.
- Kayhanian, M., C. Stransky, S. Bay, S. L. Lau and M. K. Stenstrom (2008). "Toxicity of urban highway runoff with respect to storm duration." Science of the Total Environment **389**(2-3): 386-406.
- Kayhanian, M., C. Suverkropp, A. Ruby and K. Tsay (2007). "Characterization and prediction of highway runoff constituent event mean concentration." Journal of Environmental Management **85**(2): 279-295.
- Keiser, D. A., C. L. Kling and J. S. Shapiro (2019). "The low but uncertain measured benefits of US water quality policy." Proceedings of the National Academy of Sciences **116**(12): 5262-5269.
- Kennedy, K., M. W. Hall, M. D. J. Lynch, G. Moreno-Hagelsieb and J. D. Neufeld (2014). "Evaluating bias of Illumina-based bacterial 16S rRNA gene profiles." Applied and Environmental Microbiology **80**(18): 5717-5722.
- Kim, K. H. and J. W. Bae (2011). "Amplification methods bias metagenomic libraries of uncultured single-stranded and double-stranded DNA viruses." Applied and Environmental Microbiology **77**(21): 7663-7668.
- Klein, T. M. and M. Alexander (1986). "Bacterial inhibitors in lake water." Appl. Environ. Microbiol. **52**(1): 114-118.
- Klimaszewska, K., Z. Polkowska and J. Namieśnik (2007). "Influence of mobile sources on pollution of runoff waters from roads with high traffic intensity." Polish Journal of Environmental Studies **16**(6): 889-897.
- Knights, D., J. Kuczynski, E. S. Charlson, J. Zaneveld, M. C. Mozer, R. G. Collman, F. D. Bushman, R. Knight and S. T. Kelley (2011). "Bayesian community-wide culture-independent microbial source tracking." Nature Methods **8**(9): 761-763.
- Korajkic, A., P. Wanjugi, L. Brooks, Y. Cao and V. J. Harwood (2019). "Persistence and Decay of Fecal Microbiota in Aquatic Habitats." Microbiology and Molecular Biology Reviews **83**(4): e00005-00019.
- Kumar, N. V., T. Menon, P. Pathipati and K. M. Cherian (2013). "16S rRNA sequencing as a diagnostic tool in the identification of culture-negative endocarditis in surgically treated patients." The Journal of heart valve disease **22**(6): 846-849.

- Large, T. M., S. T. Walk and T. S. Whittam (2005). "Variation in acid resistance among shiga toxin-producing clones of pathogenic *Escherichia coli*." Applied and Environmental Microbiology **71**(5): 2493-2500.
- Lau, S. L., Y. Han, J. H. Kang, M. Kayhanian and M. K. Stenstrom (2009). "Characteristics of highway stormwater runoff in Los Angeles: Metals and polycyclic aromatic hydrocarbons." Water Environment Research **81**(3): 308-318.
- Lau, S. L. and M. K. Stenstrom (2005). "Metals and PAHs adsorbed to street particles." Water Research **39**(17): 4083-4092.
- Lazarevic, V., K. Whiteson, S. Huse, D. Hernandez, L. Farinelli, M. Østerås, J. Schrenzel and P. François (2009). "Metagenomic study of the oral microbiota by Illumina high-throughput sequencing." Journal of Microbiological Methods **79**(3): 266-271.
- Liu, Z., T. Z. Desantis, G. L. Andersen and R. Knight (2008). "Accurate taxonomy assignments from 16S rRNA sequences produced by highly parallel pyrosequencers." Nucleic Acids Research **36**(18).
- Liu, Z., C. Lozupone, M. Hamady, F. D. Bushman and R. Knight (2007). "Short pyrosequencing reads suffice for accurate microbial community analysis." Nucleic Acids Research **35**(18).
- Lloyd, S. D., T. H. F. Wong and B. Porter (2002). The planning and construction of an urban stormwater management scheme. Water Science and Technology. **45**: 1-10.
- Logares, R., S. Sunagawa, G. Salazar, F. M. Cornejo-Castillo, I. Ferrera, H. Sarmiento, P. Hingamp, H. Ogata, C. de Vargas, G. Lima-Mendez, J. Raes, J. Poulain, O. Jaillon, P. Wincker, S. Kandels-Lewis, E. Karsenti, P. Bork and S. G. Acinas (2014). "Metagenomic 16S rDNA Illumina tags are a powerful alternative to amplicon sequencing to explore diversity and structure of microbial communities." Environmental Microbiology **16**(9): 2659-2671.
- Loman, N. J., R. V. Misra, T. J. Dallman, C. Constantinidou, S. E. Gharbia, J. Wain and M. J. Pallen (2012). "Performance comparison of benchtop high-throughput sequencing platforms." Nature Biotechnology **30**(5): 434-439.
- Madoui, M. A., S. Engelen, C. Cruaud, C. Belser, L. Bertrand, A. Alberti, A. Lemainque, P. Wincker and J. M. Aury (2015). "Genome assembly using Nanopore-guided long and error-free DNA reads." BMC Genomics **16**(1).
- Makepeace, D. K., D. W. Smith and S. J. Stanley (1995). "Urban Stormwater Quality: Summary of Contaminant Data." Critical Reviews in Environmental Science and Technology **25**(2): 93-139.
- Mangani, G., A. Berloni, F. Bellucci, F. Tatàno and M. Maione (2005). "Evaluation of the pollutant content in road runoff first flush waters." Water, Air, and Soil Pollution **160**(1-4): 213-228.
- Maniquiz, M. C., S. Lee and L. H. Kim (2010). "Multiple linear regression models of urban runoff pollutant load and event mean concentration considering rainfall variables." Journal of Environmental Sciences **22**(6): 946-952.
- McCarthy, D., V. G. Mitchell and A. Deletic (2006). "Escherichia coli levels in urban stormwater." Paper Presented to the Urban Drainage Modelling and International Water Sensitive Urban Design Conference, Melbourne, 4-6 April 2006.

- McLellan, S. L., S. M. Huse, S. R. Mueller-Spitz, E. N. Andreishcheva and M. L. Sogin (2010). "Diversity and population structure of sewage-derived microorganisms in wastewater treatment plant influent." Environmental Microbiology **12**(2): 378-392.
- McQuaig, S. M., T. M. Scott, J. O. Lukasik, J. H. Paul and V. J. Harwood (2009). "Quantification of human polyomaviruses JC virus and BK Virus by TaqMan quantitative PCR and comparison to other water quality indicators in water and fecal samples." Applied and Environmental Microbiology **75**(11): 3379-3388.
- Murphy, H. M., Z. Meng, R. Henry, A. Deletic and D. T. McCarthy (2017). "Current Stormwater Harvesting Guidelines Are Inadequate for Mitigating Risk from Campylobacter during Nonpotable Reuse Activities." Environmental Science and Technology **51**(21): 12498-12507.
- Natvig, E. E., S. C. Ingham, B. H. Ingham, L. R. Cooperband and T. R. Roper (2002). "Salmonella enterica serovar typhimurium and Escherichia coli contamination of root and leaf vegetables grown in soils with incorporated bovine manure." Applied and Environmental Microbiology **68**(6): 2737-2744.
- Noble, R. T., J. F. Griffith, A. D. Blackwood, J. A. Fuhrman, J. B. Gregory, X. Hernandez, X. Liang, A. A. Bera and K. Schiff (2006). "Multitiered approach using quantitative PCR to track sources of fecal pollution affecting Santa Monica Bay, California." Applied and Environmental Microbiology **72**(2): 1604-1612.
- Nshimiyimana, J. P., A. J. E. Freedman, P. Shanahan, L. C. H. Chua and J. R. Thompson (2015). ""Variation of Bacterial Communities and Pathogen Taxa as a Function of Land Use and Water Quality in an Urban Tropical Catchment of Singapore". Proceedings of the 115th General Meeting of American Society for Microbiology.
- Oladeinde, A., T. Bohrmann, K. Wong, Purucker, K. Bradshaw, R. Brown, B. Snyder and M. Molina (2014). "Decay of fecal indicator bacterial populations and bovine-associated source-tracking markers in freshly deposited cow pats." Applied and Environmental Microbiology **80**(1): 110-118.
- Olds, H. T., S. R. Corsi, D. K. Dila, K. M. Halmo, M. J. Bootsma and S. L. McLellan (2018). "High levels of sewage contamination released from urban areas after storm events: A quantitative survey with sewage specific bacterial indicators." PLoS Medicine **15**(7).
- Oliver, D. M., C. D. Clegg, A. L. Heathwaite and P. M. Haygarth (2007). "Preferential attachment of Escherichia coli to different particle size fractions of an agricultural grassland soil." Water, Air, and Soil Pollution **185**(1-4): 369-375.
- Page, D., P. Dillon, J. Vanderzalm, E. Bekele, K. Barry, K. Miotlinski and K. Levett (2010). "Managed aquifer recharge case study risk assessments." CSIRO Water for A Healthy Country Flagship Report.
- Pagotto, C., M. Legret and P. Le Cloirec (2000). "Comparison of the hydraulic behaviour and the quality of highway runoff water according to the type of pavement." Water Research **34**(18): 4446-4454.
- Parker, J. K., D. McIntyre and R. T. Noble (2010). "Characterizing fecal contamination in stormwater runoff in coastal North Carolina, USA." Water Research **44**(14): 4186-4194.

- Paule-Mercado, M. A., J. S. Ventura, S. A. Memon, D. Jahng, J. H. Kang and C. H. Lee (2016). "Monitoring and predicting the fecal indicator bacteria concentrations from agricultural, mixed land use and urban stormwater runoff." Science of the Total Environment **550**: 1171-1181.
- Quast, C., E. Pruesse, P. Yilmaz, J. Gerken, T. Schweer, P. Yarza, J. Peplies and F. O. Glöckner (2013). "The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools." Nucleic Acids Research **41**(D1): D590-D596.
- Rajal, V. B., B. S. McSwain, D. E. Thompson, C. M. Leutenegger and S. Wuertz (2007). "Molecular quantitative analysis of human viruses in California stormwater." Water Research **41**(19): 4287-4298.
- Roger, S., M. Montrejaud-Vignoles, M. C. Andral, L. Herremans and J. P. Fortune (1998). "Mineral, physical and chemical analysis of the solid matter carried by motorway runoff water." Water Research **32**(4): 1119-1125.
- Rozen, Y. and S. Belkin (2001). "Survival of enteric bacteria in seawater." FEMS Microbiology Reviews **25**(5): 513-529.
- Saget, A., G. Chebbo and J. L. Bertrand-Krajewski (1996). "The first flush in sewer systems." Water Science and Technology **33**(9): 101-108.
- Saleh-Lakha, S., K. E. Shannon, C. Goyer, J. T. Trevors, B. J. Zebarth and D. L. Burton (2008). "Nitric oxide reductase gene expression and nitrous oxide production in nitrate-grown *Pseudomonas mandelii*." Applied and Environmental Microbiology **74**(22): 6876-6879.
- Sales-Ortells, H. and G. Medema (2015). "Microbial health risks associated with exposure to stormwater in a water plaza." Water Research **74**: 34-46.
- Salipante, S. J., D. J. Sengupta, C. Rosenthal, G. Costa, J. Spangler, E. H. Sims, M. A. Jacobs, S. I. Miller, D. R. Hoogestraat, B. T. Cookson, C. McCoy, F. A. Matsen, J. Shendure, C. C. Lee, T. T. Harkins and N. G. Hoffman (2013). "Rapid 16S rRNA Next-Generation Sequencing of Polymicrobial Clinical Samples for Diagnosis of Complex Bacterial Infections." PLoS ONE **8**(5).
- Sang Ho, C., D. J. Baumler and C. W. Kaspar (2000). "Contribution of dps to acid stress tolerance and oxidative stress tolerance in *Escherichia coli* O157:H7." Applied and Environmental Microbiology **66**(9): 3911-3916.
- Sansalone, J. J. and S. G. Buchberger (1997). "Partitioning and first flush of metals in urban roadway storm water." Journal of Environmental Engineering **123**(2): 134-143.
- Sansalone, J. J., J. M. Koran, J. A. Smithson and S. G. Buchberger (1998). "Physical characteristics of Urban Roadway solids transported during rain events." Journal of Environmental Engineering **124**(5): 427-440.
- Savageau, M. A. (1983). "*Escherichia coli* habitats, cell types, and molecular mechanisms of gene control." American Naturalist **122**(6): 732-744.
- Schloss, P. D., S. L. Westcott, M. L. Jenior and S. K. Highlander (2015). "Sequencing 16S rRNA gene fragments using the PacBio SMRT DNA sequencing system." PeerJ PrePrints **3**.
- Schloss, P. D., S. L. Westcott, T. Ryabin, J. R. Hall, M. Hartmann, E. B. Hollister, R. A. Lesniewski, B. B. Oakley, D. H. Parks, C. J. Robinson, J. W. Sahl, B. Stres, G. G. Thallinger, D. J. Van Horn and C. F. Weber (2009). "Introducing mothur: Open-source, platform-independent,

community-supported software for describing and comparing microbial communities." Applied and Environmental Microbiology **75**(23): 7537-7541.

Schubert, S., S. Dufke, J. Sorsa and J. Heesemann (2004). "A novel integrative and conjugative element (ICE) of *Escherichia coli*: The putative progenitor of the *Yersinia* high-pathogenicity island." Molecular Microbiology **51**(3): 837-848.

Semenov, A. V., A. H. C. Van Bruggen, L. Van Overbeek, A. J. Termorshuizen and A. M. Semenov (2007). "Influence of temperature fluctuations on *Escherichia coli* O157:H7 and *Salmonella enterica* serovar Typhimurium in cow manure." FEMS Microbiology Ecology **60**(3): 419-428.

Shendure, J., G. J. Porreca, N. B. Reppas, X. Lin, J. P. McCutcheon, A. M. Rosenbaum, M. D. Wang, K. Zhang, R. D. Mitra and G. M. Church (2005). "Molecular biology: Accurate multiplex polony sequencing of an evolved bacterial genome." Science **309**(5741): 1728-1732.

Shinya, M., K. Tsuruho, T. Konishi and M. Ishikawa (2003). Evaluation of factors influencing diffusion of pollutant loads in urban highway runoff. Water Science and Technology. **47**: 227-232.

Sidhu, J., W. Gernjak and S. Toze (2012). "Health risk assessment of urban stormwater." Urban Water Security Research Alliance Technical Report No. 102.

Sidhu, J. P. S., W. Ahmed, W. Gernjak, R. Aryal, D. McCarthy, A. Palmer, P. Kolotelo and S. Toze (2013). "Sewage pollution in urban stormwater runoff as evident from the widespread presence of multiple microbial and chemical source tracking markers." Science of the Total Environment **463-464**: 488-496.

Sidhu, J. P. S., S. Toze, L. Hodgers, K. Barry, D. Page, Y. Li and P. Dillon (2015). "Pathogen decay during managed aquifer recharge at four sites with different geochemical characteristics and recharge water sources." Journal of Environmental Quality **44**(5): 1402-1412.

Sinclair, L., O. A. Osman, S. Bertilsson and A. Eiler (2015). "Microbial community composition and diversity via 16S rRNA gene amplicons: Evaluating the illumina platform." PLoS ONE **10**(2).

Sjogren, R. E. (1994). "Prolonged survival of an environmental *Escherichia coli* in laboratory soil microcosms." Water, Air, & Soil Pollution **75**(3-4): 389-403.

Soller, J. A., M. E. Schoen, T. Bartrand, J. E. Ravenscroft and N. J. Ashbolt (2010). "Estimated human health risks from exposure to recreational waters impacted by human and non-human sources of faecal contamination." Water Research **44**(16): 4674-4691.

Steele, J. A., A. D. Blackwood, J. F. Griffith, R. T. Noble and K. C. Schiff (2018). "Quantification of pathogens and markers of fecal contamination during storm events along popular surfing beaches in San Diego, California." Water Research **136**: 137-149.

Stenstrom, M. K., G. S. Silverman and T. A. Bursztynsky (1984). "Oil and grease in urban stormwaters." Journal of Environmental Engineering (United States) **110**(1): 58-72.

Stoeckel, D. M. and V. J. Harwood (2007). "Performance, design, and analysis in microbial source tracking studies." Applied and Environmental Microbiology **73**(8): 2405-2415.

Stotz, G. (1987). "Investigations of the properties of the surface water run-off from federal highways in the FRG." Science of the Total Environment, The **59**(C): 329-337.

- Struck, S. D., A. Selvakumar and M. Borst (2008). "Prediction of effluent quality from retention ponds and constructed wetlands for managing bacterial stressors in storm-water runoff." Journal of Irrigation and Drainage Engineering **134**(5): 567-578.
- Stumpf, C. H., M. F. Piehler, S. Thompson and R. T. Noble (2010). "Loading of fecal indicator bacteria in North Carolina tidal creek headwaters: Hydrographic patterns and terrestrial runoff relationships." Water Research **44**(16): 4704-4715.
- Sunagawa, S., D. R. Mende, G. Zeller, F. Izquierdo-Carrasco, S. A. Berger, J. R. Kultima, L. P. Coelho, M. Arumugam, J. Tap, H. B. Nielsen, S. Rasmussen, S. Brunak, O. Pedersen, F. Guarner, W. M. De Vos, J. Wang, J. Li, J. Doré, S. Dusko Ehrlich, A. Stamatakis and P. Bork (2013). "Metagenomic species profiling using universal phylogenetic marker genes." Nature Methods **10**(12): 1196-1199.
- Tabe, E. S., J. Oloya, D. K. Doetkott, M. L. Bauer, P. S. Gibbs and M. L. Khaita (2008). "Comparative effect of direct-fed microbials on fecal shedding of *Escherichia coli* O157:H7 and *Salmonella* in naturally infected feedlot cattle." Journal of Food Protection **71**(3): 539-544.
- Tao, H., C. Bausch, C. Richmond, F. R. Blattner and T. Conway (1999). "Functional genomics: Expression analysis of *Escherichia coli* growing on minimal and rich media." Journal of Bacteriology **181**(20): 6425-6440.
- Taylor, C. M. (2008). "Enterohaemorrhagic *Escherichia coli* and *Shigella dysenteriae* type 1-induced haemolytic uraemic syndrome." Pediatric Nephrology **23**(9): 1425-1431.
- Thomson, N. R., E. A. McBean, W. Snodgrass and I. B. Monstrenko (1997). "Highway stormwater runoff quality: Development of surrogate parameter relationships." Water, Air, and Soil Pollution **94**(3-4): 307-347.
- Thurston, J. A., C. P. Gerba, K. E. Foster and M. M. Karpiscak (2001). "Fate of indicator microorganisms, giardia and cryptosporidium in subsurface flow constructed wetlands." Water Research **35**(6): 1547-1551.
- Topp, E., M. Welsh, Y. C. Tien, A. Dang, G. Lazarovits, K. Conn and H. Zhu (2003). "Strain-dependent variability in growth and survival of *Escherichia coli* in agricultural soil." FEMS Microbiology Ecology **44**(3): 303-308.
- Tseng, L. Y. and S. C. Jiang (2012). "Comparison of recreational health risks associated with surfing and swimming in dry weather and post-storm conditions at Southern California beaches using quantitative microbial risk assessment (QMRA)." Marine Pollution Bulletin **64**(5): 912-918.
- Unno, T., J. Jang, D. Han, J. H. Kim, M. J. Sadowsky, O. S. Kim, J. Chun and H. G. Hur (2010). "Use of barcoded pyrosequencing and shared OTUs to determine sources of fecal bacteria in watersheds." Environmental Science and Technology **44**(20): 7777-7782.
- Vital, M., F. Hammes and T. Egli (2008). "*Escherichia coli* O157 can grow in natural freshwater at low carbon concentrations." Environmental Microbiology **10**(9): 2387-2396.
- Vorreiter, L. and C. Hickey (1994). Incidence of the first flush phenomenon in catchments of the Sydney region. National Conference Publication - Institution of Engineers, Australia.

- Wang, Q., G. M. Garrity, J. M. Tiedje and J. R. Cole (2007). "Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy." Applied and Environmental Microbiology **73**(16): 5261-5267.
- Westerlund, C., M. Viklander and M. Bäckström (2003). Seasonal variations in road runoff quality in Luleå, Sweden. Water Science and Technology. **48**: 93-101.
- White, S. S. and M. Boswell (2006). "Planning for water quality: Implementation of the NPDES phase II stormwater program in California and Kansas." Journal of Environmental Planning and Management **49**(1): 141-160.
- Wu, J. S., C. J. Allan, W. L. Saunders and J. B. Evett (1998). "Characterization and pollutant loading estimation for highway runoff." Journal of Environmental Engineering **124**(7): 584-592.
- Young, K. D. and E. L. Thackston (1999). "Housing density and bacterial loading in urban streams." Journal of Environmental Engineering **125**(12): 1177-1180.
- Yufen, R., W. Xlaoke, O. Zhiyun, Z. Hua, D. Xiaonan and M. Hong (2008). "Stormwater runoff quality from different surfaces in an urban catchment in Beijing, China." Water Environment Research **80**(8): 719-724.
- Zhou, J., X. C. Wang, Z. Ji, L. Xu and Z. Yu (2015). "Source identification of bacterial and viral pathogens and their survival/fading in the process of wastewater treatment, reclamation, and environmental reuse." World journal of microbiology & biotechnology **31**(1): 109-120.

Chapter 2

Evaluating the Accuracy of EPA Method 1604 for *E. coli* and Total Coliform Enumeration by 16S rRNA Gene Sequencing

Abstract

This study utilized high-throughput 16S rRNA amplicon sequencing to explore the positive identification of the U.S. Environmental Protection Agency-approved culture-based coliform testing method based on phylogenetic affiliation. The enumeration method was tested on the four types of environmental matrices, including stream water, stormwater runoff, sewage, and stream sediments. The results showed detecting true-positive (TP) of non-*E. coli* coliforms was more challenging than detecting TP of *E. coli* in this study. The missed *E. coli* caused false-positive (FP) results in stream water and sewage non-*E. coli* TC isolates, but the major FP groups of stormwater runoff and sediment non-*E. coli* TC isolates were the *Aeromonas* and *Bacillus* non-coliform groups. The sensitivity of the *E. coli* test was the lowest for the stream water samples (75.0%), but was respectable for other mediums, exhibiting values of 93.7%, 98.9%, and 100% for sewage, sediment, and stormwater runoff, respectively. Furthermore, the specificity of the *E. coli* identification was above 95.6% for the four environmental matrices, and the average rate of the overall agreement was 81.2%, 96.5%, 98.5%, and 97.7% for stream water, sewage, sediment, and stormwater runoff, respectively. The results suggested that this culture-based method was accurate at correctly identifying *E. coli* on the four different matrices. Overall, the high-throughput 16S rRNA Amplicon Sequencing was shown to be a useful approach to confirm the positive colonies of the coliform-testing method. Assessment of performance of other traditional enzyme-based bacterial testing methods will be conducted in future studies.

1. Introduction

Escherichia coli (*E. coli*) and Total coliforms (TC), as fecal indicators, are commonly used to assess the potential presence of pathogenic microorganisms in environments. EPA Method 1604 for enumerating TC and *E. coli* is a universally accepted method for quantifying the presence of these organisms in surface water and treated wastewaters (EPA 2002). The method utilizes selective dyes and the specific enzyme activities of TC and *E. coli* to isolate and differentiate *E. coli* and non-*E. coli* TC. After plating, *E. coli* are distinguishable as blue colored colonies through

the action of b-glucuronidase(GUD) enzyme on 5-Bromo-4-Chloro-3-Indolyl-Beta-D-glucuronide (BCIG), and non- *E. coli* TC are recognized as red colored colonies through the action of 2,3,5-Triphenoltetrazolium Chloride (TTC) (Grant 1997). Due to the relatively low cost of implementation, the use of this method has expanded to include a variety of different ecosystems, such as stream water (Hamilton, Kim, et al., 2005), stormwater runoff (Simon and Makarewicz 2009), sewage(Crane, Jackson, et al., 2006), and stream sediments (Boutilier, Jamieson, et al., 2009). However, the microbial communities in these ecosystems are highly diverse, and non-coliform bacterial strains may cause false-positive (FP) results (Bernasconi, Volponi, et al., 2006; Olstadt, Schauer, et al., 2007; Maheux, Huppé, et al., 2008), which execute an unnecessary burden for panic and remediation. And FP results may vary in different environmental matrices (Dunling and FIESSEL 2008; McLain and Williams 2008) due to the distinct bacteria in each ecosystem. Furthermore, while some non-coliform bacteria may result in FP results, some strains of TC and *E. coli* may not respond to the testing method, subsequently causing false-negative (FN) results, which could make ecosystems less impaired than it actually is and thus introducing a public health risk (Cahoon and Thompson 1987; Feng, Lum, et al., 1991; Feng 1995).

Various methods of true identification for colonies resulting from EPA method 1604 have been reported, including API 20E, VITEK2, and Sanger sequencing systems (Aldridge and Hodges 1981; Sanders, Peyret, et al., 2000; Zhang, Hong, et al., 2015). These methods require selecting individual colonies for verification, which can be extremely time-consuming if the accuracy of testing requires numerous colonies to be verified.

High-throughput 16S rRNA amplicon sequencing is revolutionizing the phylogenetic identification of microorganisms. The amplification of 16S rRNA gene regions is a powerful strategy for identifying bacterial populations, and this tool can be used to garner information about the genetic composition of a microbial community as a whole as opposed to bacterial monocultures. This study proposes the use of 16S rRNA sequencing as a faster, more cost-effective method of identifying the positive results in EPA Method 1604 by analyzing the genetic makeup of all colonies present.

In order to test the positive results of the EPA method, *E. coli* and non- *E. coli* TC colonies were isolated from stream water, stormwater runoff, sewage, and stream sediments, and then characterized by the high-throughput 16S rRNA amplicon sequencing method. Results from the

sequencing analysis indicated that, for each sample type, the vast majority of positive detections were actually *E. coli* and non-*E. coli* TC. Moreover, a limited number of non-coliform groups, such as *Bacillus*, *Enterococcus*, *Brevibacillus*, *Stenotrophomonas*, *Lactococcus*, *Streptococcus*, *Flavobacterium*, *Aeromonas*, and *Pseudomonas*, were also detected and caused false-positive (FP) results. The overall results suggested that the EPA method was accurate at correctly identifying *E. coli* and total coliforms on the four different matrices in this study.

2. Materials and Methods

2.1 Environmental matrices selection

A total of 12 environmental samples from sewage, stream water, stream sediment, and stormwater runoff in East Tennessee, USA, were collected between 2017 and 2018. Untreated sewage samples were collected from three different municipal wastewater treatment plants at the inlet under dry weather conditions without any impact of recent storm events.

Stream water samples were collected by submerging a sterile 1 L bottle below the stream surface according to standard methods (Lurry and Kolbe 2000) from three sampling locations in a local stream. All samples represented typical base flow for the stream, meaning at least three days without precipitation were required prior to sampling. Stream sediment samples were collected at each stream water sampling location. A sterile acrylic tube (115mm length) was pressed approximately 15 cm deep into the sediment, and the core was pulled with a gloved hand.

Stormwater runoff samples were collected from three different pavement sites. Sampling was conducted using Nalgene™ Storm Water Samplers, which were installed immediately prior to the anticipated storms. Each sampler collected a full one-liter grab sample representing first-flush stormwater runoff.

After collection, all samples were transported using a cooler with ice and delivered to the laboratory within 6 hours, thus satisfying holding time requirements.

2.2 Isolate collection

The TC and *E. coli* colonies were obtained following the protocol provided by the manufacturer (Hach, Loveland, CO). Briefly, to extract TC and *E. coli* from sediment samples, 10 g of wet sediment was mixed with 100 mL of sterile Phosphate-buffered saline (PBS) and shaken for 5 min

to make a sediment slurry. Water and sediment slurry samples were diluted and filtered to trap TC and *E. coli* on the membrane surface (MilliporeSigma™ HAWG047S6), and then incubated at $36 \pm 2^\circ\text{C}$ for 22 ± 2 hours. The appropriate number of colonies in one membrane is 50-100.

Approximately 200 individual putative *E. coli* colonies from the same sample were picked and transferred to a single membrane using sterile needles. Similarly, 200 individual putative non-*E. coli* TC colonies from the same sample were transferred to a single membrane. Each membrane was ideally composed of a monoculture of each bacterial group: *E. coli*, and non-*E. coli* TC. The membranes were stored at -80°C until DNA extraction.

2.3 DNA Extraction and High-Throughput 16S rRNA Amplicon Sequencing

DNA from the colony isolates was extracted by the FastDNA Spin Kit for Soil (MP Biomedicals) following the manufacturer's protocol.

High-Throughput 16S rRNA Amplicon Sequencing was performed as previously described (Wyckoff, Chen, et al., 2017). The PCR amplification used primers 515-F and 806-R to target the V4 region of the 16S rRNA gene (Caporaso, Lauber, et al., 2012). PCR products were purified using the Wizard SV Gel and PCR Clean-Up System (Promega), and paired-end sequencing was completed using the Illumina MiSeq platform.

2.4 Sequence Analysis

Analysis of sequence reads was conducted using QIIME 2 (Bolyen, Rideout, et al., 2019). Amplicon sequence variants (ASVs) enumeration was performed using the DADA2 pipeline (Callahan, McMurdie, et al., 2016). In contrast to the traditional cluster-based methods based on 97% sequence identity, DADA2 is an error model that can differentiate ASVs based on as little as a single nucleotide. This method is preferred for this study because the 16S V4 region amplicon is short, and some TC species differ at only one position. If traditional 97% identification methods were used, these taxa may be mistakenly grouped together. The obtained ASVs were then matched to the Greengenes 13.8 database.

Sequence alignment and construction of phylogenetic trees were conducted using the EMBL-EBI (Madeira, Lee, et al., 2019) and iTOL online tools (Letunic and Bork 2016).

2.5 Estimation of method effectiveness

The effectiveness of the EPA 1604 method was quantified by calculating specificity, sensitivity, , and overall agreement (equations 1-3) for each sample (Snedecor and Cochran 1989).

$$\text{Specificity} = \left[\frac{\text{TN}}{\text{TN} + \text{FP}} \right] \times 100\% \quad (1)$$

$$\text{Sensitivity} = \left[\frac{\text{TP}}{\text{TP} + \text{FN}} \right] \times 100\% \quad (2)$$

$$\text{Overall agreement} = \left[\frac{\text{TP} + \text{TN}}{\text{Overall}} \right] \times 100\% \quad (3)$$

In these equations, TN stands for true negative, TP stands for true positive, FN stands for false negative, FP stands for false positive, and Overall is the total of TN, TP, FN, and FP.

3. Results and Discussions

3.1 Diversity of Putative TC isolates

A phylogenetic tree was constructed with the high-throughput 16S rRNA amplicon sequencing data. ASVs making up less than 0.2% relative abundance in each sample were filtered out because there were approximately 200 colonies in each sequencing sample, so any ASVs with extremely low proportions were most likely due to contamination during sample processing, not from the colonies themselves. In total, 69 ASVs were obtained in this study, including 49 *Enterobacteriaceae* clusters and 20 non- *Enterobacteriaceae* clusters (Figure 2.1). Although the definitions of coliform bacteria vary between different detection methods, the lactose-fermenting *Enterobacteriaceae* is currently widely used to define TC (Zhang, Hong, et al., 2015).

Obviously, the genera within the family *Enterobacteriaceae*, including *Citrobacter*, *Cronobacter*, *Enterobacter*, *Erwinia*, *Escherichia*, *Klebsiella*, *Kosakonia*, *Pantoea*, *Pseudescherichia*, *Raoultella*, *Salmonella*, *Serratia*, and *Trabulsiella*, are genetically distinct organisms. However, the similarity between the two clusters with the largest phylogenetic variance in *Enterobacteriaceae* is still 92.07%. In the 11 different *Escherichia* genera, the lowest similarity is 99.12%.

Among the 20 non-TC ASVs, *Streptococcus*, *Flavobacterium*, *Aeromonas*, and *Pseudomonas* have been observed as common causes of FP results (Frampton and Restaino 1993; Grant 1997). However, *Bacillus*, *Enterococcus*, *Brevibacillus*, *Stenotrophomonas*, and *Lactococcus* were the

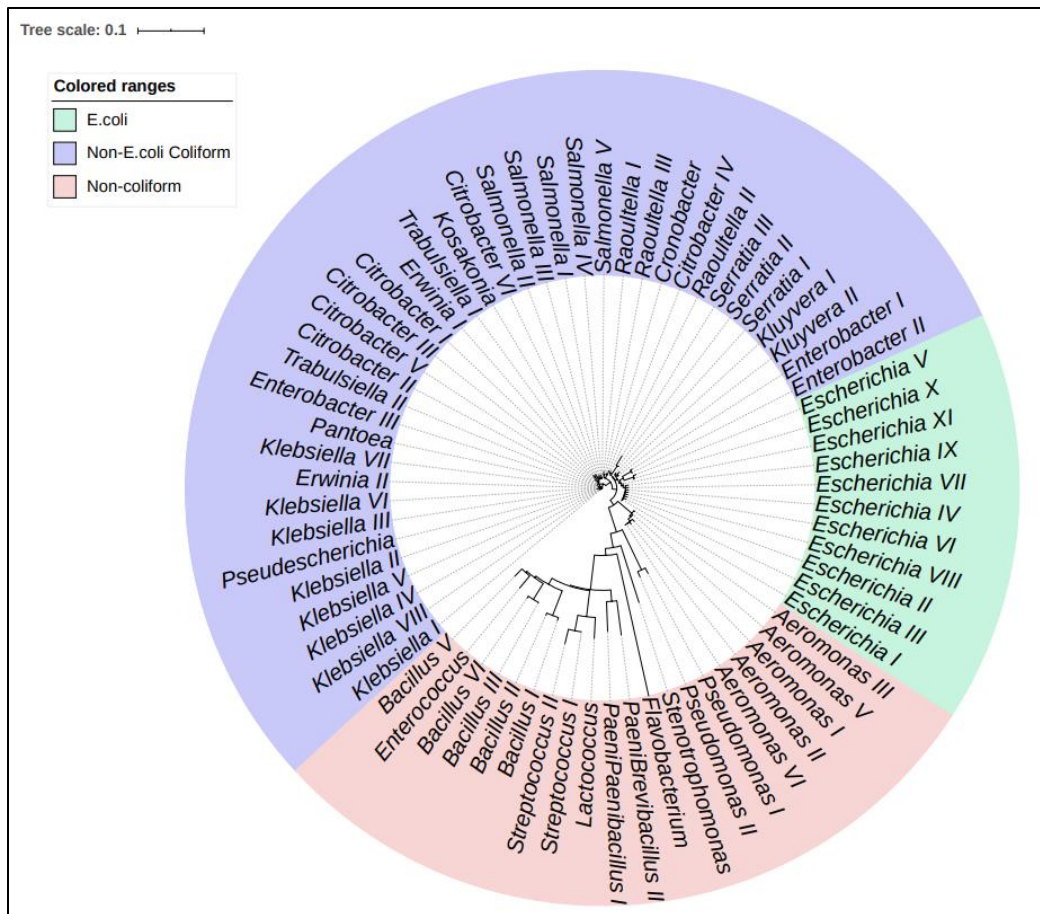


Figure 2.1. Phylogenetic tree of ASVs using 16S rRNA sequencing following DADA2 Pipeline obtained from the putative *E. coli* and non-*E. coli* TC colonies inferred by Neighbor-Joining. The total 69 ASVs included 20 non-coliform, 37 non-*E. coli* coliform, and 12 *E. coli* ASVs.

first reported in this study.

3.2 Bacterial isolates responsible in *E. coli*

Among the putative *E. coli* isolates in the four different environmental matrices, a total of 15 ASVs were identified (Figure 2.1). Four of these ASVs belonged to the *Escherichia* genera, and these ASVs made up greater than 95.3% of the community in all samples (Figure 2.2).

Although the stormwater runoff showed the highest variability of the isolates, including 12 different ASVs, other sample types contained only 3 or 4 ASVs. The *Escherichia_I* was the most abundant ASV across all samples. Moreover, in sediment and stream water samples, it was the only *Escherichia* ASV present.

There are 11 non-*Escherichia* ASVs that caused false positives, and the relative abundances of these ASVs ranged from 0.12% to 1.35% (Figure 2.2). Of ASVs that resulted in FP for *E. coli*, 3 belonged to the *Enterobacteriaceae* group (non-*E. coli* TC), including *Salmonella_V*, *Raoultella_II*, and *Klebsiella_VIII*. Non-TC groups that caused FP were *Aeromonas*, *Bacillus*, *Enterococcus*, *Flavobacterium*, *Brevibacillus*, *Pseudomonas*, and *Stenotrophomonas*.

Although FP results were found to be present using EPA Method 1604, the impacts were minor. Additionally, numerous previous studies have confirmed this method for *E. coli* quantification in the upper 90% range (Pearez, Berrocal, et al., 1986; Sarhan and Foster 1991; Grant 1997; Hamilton, Kim, et al., 2005; Olstadt, Schauer, et al., 2007; Zhang, Hong, et al., 2015).

3.3 Bacterial isolates responsible in non- *E. coli* TC

Non-TC groups caused the FP results for both *E. coli* and non-*E. coli* TC tests. *E. coli* that was not registered by the EPA method caused both FP of non-*E. coli* TC, as well as FN of *E. coli*. The primary true positive (TP) genera found in all samples were *Klebsiella*, *Citrobacter*, *Enterobacter*, *Raoultella*, *Salmonella*, *Serratia*, and *Trabulsiella* (Figure 2.3). The overall relative abundances of true non-*E. coli* TC results in sediment, stormwater runoff, sewage, and stream water were 78.4%, 91.9%, 91.8%, and 61.7%, respectively.

No missed *E. coli* was found in stormwater runoff non-*E. coli* TC isolates, and only 1.1% missed *E. coli* was found in sediment non-*E. coli* TC isolates. The FP results present in stormwater runoff and sediment non-*E. coli* TC isolates were primarily caused by non-TC groups.

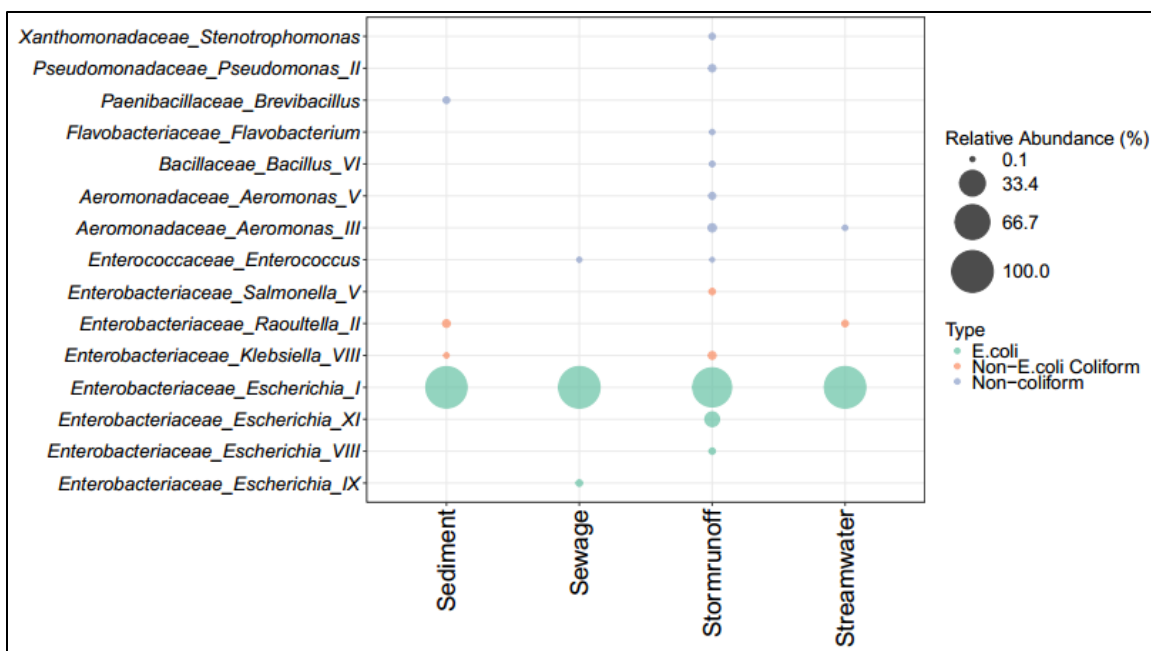


Figure 2.2. Relative abundances of ASVs got from putative *E. coli* colonies. The percentages of true *E. coli* were 98.0%, 99.9%, 95.3%, and 99.4% in sediment, sewage, stormwater runoff, and stream water, respectively.

The *Aeromonas* genus was a common instigator of FP results, and it was found in all samples. *Aeromonas* abundance was especially high in sediment non-*E. coli* TC isolates (14.2%). *Bacillus* was another common genus resulting in FP results in non-*E. coli* TC isolates, and it was found in high abundance in sediment and stormwater runoff samples (5.2% and 4.7%, respectively). Although *Enterococcus*, *Citrobacter*, *Lactococcus*, *Pseudomonas*, and *Stenotrophomonas* also occurred in non-*E. coli* TC isolates, the abundances of these genera were all lower than 0.4%.

An overall rate of FP results was calculated by combining missed *E. coli* rates with the frequency of FP results in non-*E. coli* TC isolates due to non-TC groups. The overall frequency of FP in sediment, stormwater runoff, sewage, and stream water was 7.6%, 8.1%, 21.2%, and 38.4%, respectively. The FP frequency in non-*E. coli* TC isolates was much higher than solely *E. coli* isolates. Therefore, the method for quantifying non-*E. coli* TC is not as accurate as the method for quantifying only *E. coli*.

Some ASVs were shown in both *E. coli* TC isolates and non-*E. coli* TC isolates, including *Aeromonas*_III, *Aeromonas*_V, *Bacillus*_VI, *Escherichia*_I, *Klebsiella*_VIII, *Raoultella*_II, *Salmonella*_V, *Enterococcus*, and *Stenotrophomonas*. The reason for this discrepancy may be that the 16S rRNA Amplicon Sequencing only sequences a fraction of the microbial genes, and the short reads may not have the resolution necessary to be classified as a separate ASV (Franzosa, Hsu, et al., 2015). Therefore, there can be multiple different strains in the same ASV, and the different strains may have different responses for the culture-based TC testing. For example, 31 GUD-negative *E. coli* strains have been reported to avoid identification in the culture-based testing method (Feng, Lum, et al., 1991).

3.4 Effectiveness of the culture-based testing method

The performance of the culture-based *E. coli* testing method was evaluated for specificity, sensitivity, and overall agreement for various environmental mediums. Specificities were above 95.6% for all mediums, and the sewage medium reported 99.9% specificity, indicating that nearly all positive detections are the correct coliform classification (Figure 2.4). Numerous previous studies have evaluated the specificities of this method for *E. coli* quantification based on different methods, and the ranges were above 85% (Pearez, Berrocal, et al., 1986; Sarhan and Foster 1991; Grant 1997; Hamilton, Kim, et al., 2005; Olstadt, Schauer, et al., 2007; Zhang, Hong, et al., 2015).

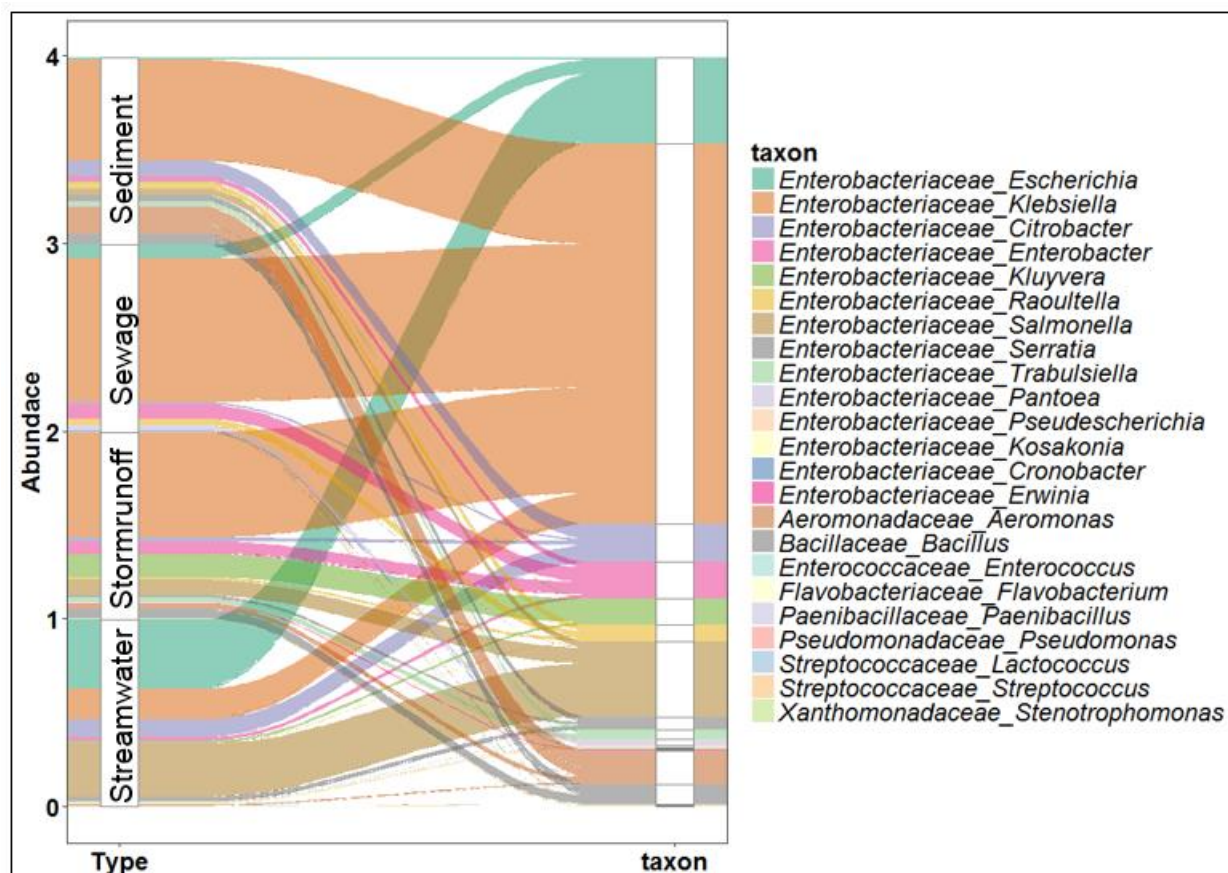


Figure 2.3. Relative abundance of ASVs from putative non-*E. coli* TC. 78.7%, 92.5%, 91.9% and 61.6% true non-*E. coli* TC, 1.1%, 7.1%, 0% and 37.4% missed *E. coli*, and 20.2%, 0.4%, 8.1% and 1.0% non-TC in sediment, sewage, stormwater runoff, and stream water respectively.

Sensitivity was high for stormwater runoff, sediment, and sewage (Figure 2.5). The sensitivity of the method for the stormwater runoff was especially high (100%) due to a complete lack of FN results. The sensitivity of the method for sewage and sediment mediums was 93.7% and 98.9%, respectively. However, the culture-based method exhibited the lowest sensitivity for the stream water medium at 75.0% due to numerous undetected *E. coli* in non-*E. coli* TC isolates. Additionally, the standard deviation for the stream water samples was 12.7%, indicating that the triplicates were quite variable.

In general, the overall agreement exhibited by the culture-based *E. coli* testing for sewage, sediment, and stormwater runoff samples was higher than 96.5% for each, suggesting that this method is highly accurate at correctly identifying *E. coli* in these mediums based on the 16S rRNA gene sequencing conformation (Figure 2.6). Although high specificity for stream water was shown in this study, the low sensitivity caused the overall agreement to decrease to a final value of 81.2%.

4. Conclusions

The EPA 1604 culture-based method for quantifying the presence of *E. coli* and total coliforms in aquatic ecosystems has become widely used due to its ease of implementation and relatively low cost. It is important to confirm the positive results of the method for various environmental matrices, and advances in the high-throughput 16S rRNA amplicon sequencing allow this tool to be used to explore the results of this culture-based test. In this study, the high specificity exhibited by the enumeration method indicates that the method is accurate in identifying *E. coli* colonies on the four types of ecosystems. However, elevated levels of FP results were found for the non-*E. coli* TC test, indicating that the method may decrease in accuracy when attempting to identify a taxonomically larger subset of bacteria. Overall, the high-throughput 16S rRNA Amplicon Sequencing was found to be an effective method of verifying the accuracy of The EPA coliform testing method. It should also be mentioned that, due to the limited samples in this study, observations need to be validated with more samples from different environmental matrices. Moreover, further efforts will extend the use of 16S rRNA sequencing to assessing the performance of other traditional enzyme-based bacterial testing methods.

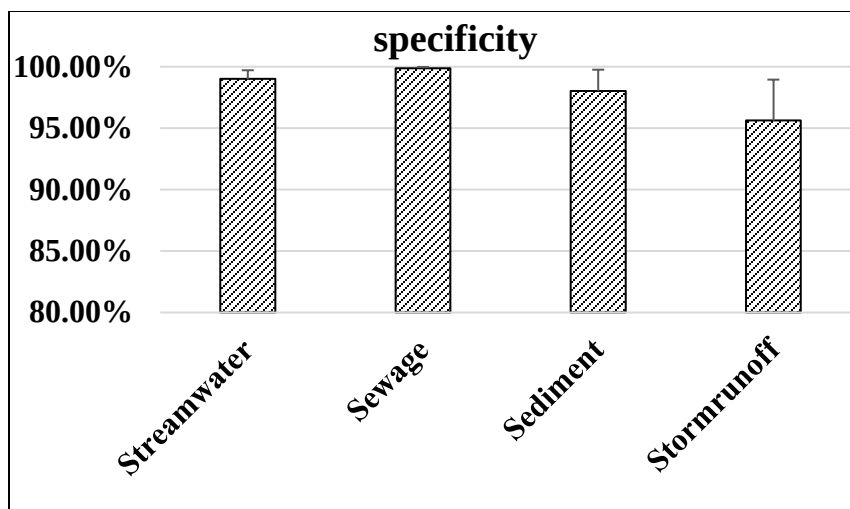


Figure 2.4. Specificity of putative *E. coli*. The average rates of 99.0%, 99.9%, 98.0%, and 95.6% for stream water, sewage, sediment, and stormwater runoff, respectively.

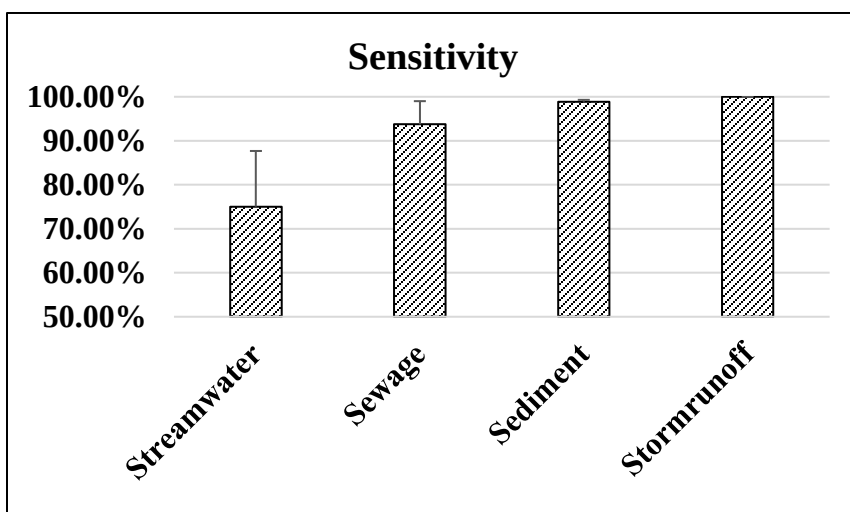


Figure 2.5. Sensitivity of putative *E. coli*. The average rates of 75.0%, 93.7%, 98.9%, and 100% for stream water, sewage, sediment, and stormwater runoff, respectively.

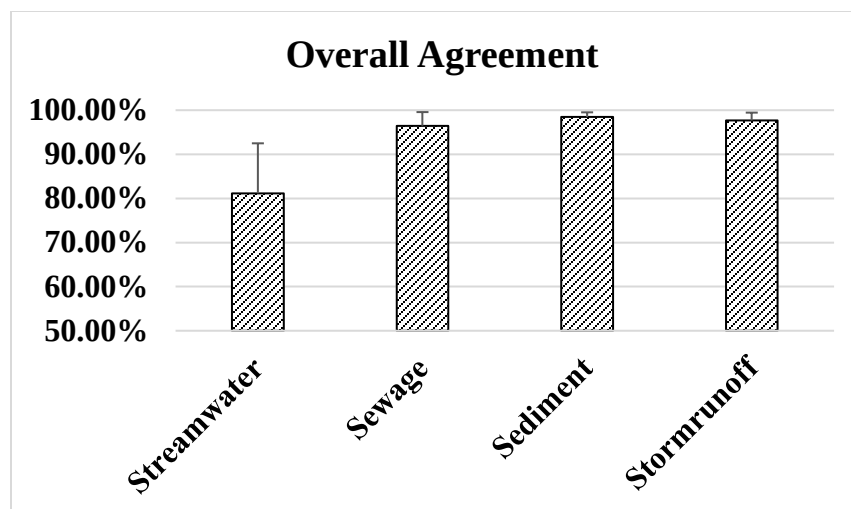


Figure 2.6. The overall agreement of putative *E. coli*. The average rates of 81.2%, 96.5%, 98.5%, and 97.7% for stream water, sewage, sediment, and stormwater runoff, respectively.

References

- Aldridge, K. E. and R. L. Hodges (1981). "Correlation studies of entero-set 20, API 20E and conventional media systems for Enterobacteriaceae identification." Journal of clinical microbiology **13**(1): 120-125.
- Bernasconi, C., G. Volponi and L. Bonadonna (2006). "Comparison of three different media for the detection of *E. coli* and coliforms in water." Water science and technology **54**(3): 141-145.
- Bolyen, E., J. R. Rideout, M. R. Dillon, N. A. Bokulich, C. C. Abnet, G. A. Al-Ghalith, H. Alexander, E. J. Alm, M. Arumugam and F. Asnicar (2019). "Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2." Nature biotechnology **37**(8): 852-857.
- Boutilier, L., R. Jamieson, R. Gordon, C. Lake and W. Hart (2009). "Adsorption, sedimentation, and inactivation of *E. coli* within wastewater treatment wetlands." Water research **43**(17): 4370-4380.
- Cahoon, F. E. and J. S. Thompson (1987). "Frequency of *Escherichia coli* 0157: H7 isolation from stool specimens." Canadian journal of microbiology **33**(10): 914-915.
- Callahan, B. J., P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson and S. P. Holmes (2016). "DADA2: high-resolution sample inference from Illumina amplicon data." Nature methods **13**(7): 581.
- Caporaso, J. G., C. L. Lauber, W. A. Walters, D. Berg-Lyons, J. Huntley, N. Fierer, S. M. Owens, J. Betley, L. Fraser and M. Bauer (2012). "Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms." The ISME journal **6**(8): 1621.
- Crane, R., C. B. Jackson, R. Sedlacek and D. Walker (2006). "EPA approves new test procedures for the analysis of microbiological pollutants in wastewater and sludge." Loveland, CO: Hach Company. Available online at <http://www.hach.com/fmmimghach>.
- Dunling, W. and W. FIESSEL (2008). "Evaluation of media for simultaneous enumeration of total coliform and *Escherichia coli* in drinking water supplies by membrane filtration techniques." Journal of Environmental Sciences **20**(3): 273-277.
- EPA, U. (2002). Method 1604: Total Coliforms and *Escherichia coli* in Water by Membrane Filtration Using a Simultaneous Detection Technique, EPA-821-R-02-023 US Environmental Protection Agency.
- Feng, P. (1995). "*Escherichia coli* serotype O157: H7: novel vehicles of infection and emergence of phenotypic variants." Emerging infectious diseases **1**(2): 47.
- Feng, P., R. Lum and G. W. Chang (1991). "Identification of uidA gene sequences in beta-D-glucuronidase-negative *Escherichia coli*." Appl. Environ. Microbiol. **57**(1): 320-323.
- Frampton, E. and L. Restaino (1993). "Methods for *Escherichia coli* identification in food, water and clinical samples based on beta-glucuronidase detection." Journal of Applied Bacteriology **74**(3): 223-233.
- Franzosa, E. A., T. Hsu, A. Sirota-Madi, A. Shafquat, G. Abu-Ali, X. C. Morgan and C. Huttenhower (2015). "Sequencing and beyond: integrating molecular'omics' for microbial community profiling." Nature Reviews Microbiology **13**(6): 360.

Grant, M. (1997). "A new membrane filtration medium for simultaneous detection and enumeration of *Escherichia coli* and total coliforms." Appl. Environ. Microbiol. **63**(9): 3526-3530.

Hamilton, W. P., M. Kim and E. L. Thackston (2005). "Comparison of commercially available *Escherichia coli* enumeration tests: Implications for attaining water quality standards." Water Research **39**(20): 4869-4878.

Letunic, I. and P. Bork (2016). "Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees." Nucleic acids research **44**(W1): W242-W245.

Lurry, D. L. and C. M. Kolbe (2000). Interagency field manual for the collection of water-quality data, US Department of the Interior, US Geological Survey.

Madeira, F., J. Lee, N. Buso, T. Gur, N. Madhusoodanan, P. Basutkar, A. Tivey, S. C. Potter, R. D. Finn and R. Lopez (2019). "The EMBL-EBI search and sequence analysis tools APIs in 2019." Nucleic acids research.

Maheux, A. F., V. Huppé, M. Boissinot, F. J. Picard, L. Bissonnette, J.-L. T. Bernier and M. G. Bergeron (2008). "Analytical limits of four β -glucuronidase and β -galactosidase-based commercial culture methods used to detect *Escherichia coli* and total coliforms." Journal of microbiological methods **75**(3): 506-514.

McLain, J. E. and C. F. Williams (2008). "Seasonal variation in accurate identification of *Escherichia coli* within a constructed wetland receiving tertiary-treated municipal effluent." Water research **42**(15): 4041-4048.

Olstadt, J., J. J. Schauer, J. Standridge and S. Kluender (2007). "A comparison of ten USEPA approved total coliform/*E. coli* tests." Journal of water and health **5**(2): 267-282.

Pearez, J., C. I. Berrocal and L. Berrocal (1986). "Evaluation of a commercial β -glucuronidase test for the rapid and economical identification of *Escherichia coli*." Journal of applied bacteriology **61**(6): 541-545.

Sanders, C. C., M. Peyret, E. S. Moland, C. Shubert, K. S. Thomson, J.-M. Boeufgras and W. E. Sanders (2000). "Ability of the VITEK 2 Advanced Expert System To Identify β -Lactam Phenotypes in Isolates of Enterobacteriaceae and *Pseudomonas aeruginosa*." Journal of Clinical Microbiology **38**(2): 570-574.

Sarhan, H. and H. Foster (1991). "A rapid fluorogenic method for the detection of *Escherichia coli* by the production of β -glucuronidase." Journal of applied bacteriology **70**(5): 394-400.

Simon, R. D. and J. C. Makarewicz (2009). "Storm water events in a small agricultural watershed: Characterization and evaluation of improvements in stream water microbiology following implementation of Best Management Practices." Journal of Great Lakes Research **35**: 76-82.

Snedecor, G. W. and W. G. Cochran (1989). "Statistical methods, 8thEdn." Ames: Iowa State Univ. Press Iowa.

Wyckoff, K. N., S. Chen, A. J. Steinman and Q. He (2017). "Impact of Roadway Stormwater Runoff on Microbial Contamination in the Receiving Stream." Journal of environmental quality **46**(5): 1065-1071.

Zhang, Y., P.-Y. Hong, M. W. LeChevallier and W.-T. Liu (2015). "Phenotypic and phylogenetic identification of coliform bacteria obtained using 12 coliform methods approved by the US Environmental Protection Agency." Appl. Environ. Microbiol. **81**(17): 6012-6023.

Chapter 3

Comparison of the Presence and Behavior of Fecal Indicator Bacteria and Turbidity in Stormwater Runoff

Abstract

Stormwater runoff from urban impervious surfaces presents a high potential risk of non-point source contaminations. In this study, *Escherichia coli* (*E. coli*), total coliform (TC), and turbidity in stormwater runoff from roadway pavement and culvert were evaluated over 23 storm events in east Tennessee. First flush grab samples were collected from 6 sites, including a rural minor collector road, a rural minor arterial road, an urban minor collector road, an urban minor arterial road, a rural interstate divided highway, and an urban interstate divided highway in 2016-2017. Turbidity and fecal indicator bacteria (FIB) in runoff showed great variances across different sites and different storm events, but the elevated pollutant levels present a potential public health hazard. Moreover, Turbidity and FIB were lack of correlation and revealed different distribution patterns in this study. Analyses suggested that the turbidities were similar in pavement runoff and culvert runoff on the same site from the same storm event, but concentrations of FIB were not. In addition, results showed turbidities in runoff were highly correlated with precipitation intensity (PI) and antecedent dry period (ADP), but poorly correlated between FIB and these storm weather conditions. However, seasonal differences for FIB were noted, and FIB concentrations in winter were lower than in other seasons. Further, the impact of site features was analyzed in this study. The FIB concentrations are higher at the urban minor collector than the rural minor collector, and higher at urban minor arterial than the rural minor arterial. Higher annual average daily traffic (AADT) road sites, higher FIB concentrations in this study. Overall, the two different indicator bacteria showed a strong correlation, but differential behavior of FIB and turbidity was revealed.

1. Introduction

Urbanization development has increased the presence of impervious surfaces (e.g., roadways and parking lots) so that stormwater runoff passes overland, carrying pollutants to receiving water bodies (Stotz 1987). Compared with other runoff, the roadway runoff presents a high potential risk of non-point source contaminations (Athayde, Shelley, et al., 1983; Driscoll, Shelley, et al., 1990). And the contaminants in the runoff depend on numerous factors, such as seasonality, precipitation

intensity (PI), antecedent dry period (ADP), traffic characteristic, and watershed features (Thomson, McBean, et al., 1997; Kayhanian, Singh, et al., 2003; Westerlund, Viklander, et al., 2003; Han, Lau, et al., 2006; Han, Lau, et al., 2006; Hallberg, Renman, et al., 2007; Li and Barrett 2008). Thus, remarkable spatial and temporal variations were shown in both compositions and concentrations. Conventional water quality parameters (turbidity, pH, temperature, and total suspended solids), metal constituents, toxicity, polycyclic aromatic hydrocarbons (PAHs), fecal indicator bacteria (FIB), and nutrients in stormwater runoff were studied frequently (Thomson, McBean, et al., 1997; Kayhanian, Suverkropp, et al., 2007; Kayhanian and Paytan 2011). Researchers have been trying to find correlations between different quality parameters in order to reduce test costs (Kayhanian, Suverkropp, et al., 2007; Thomson, McBean, et al., 1997). Moreover, some studies believed that FIB attached to particles and were transported together, and indeed, FIB were correlated with turbidity in some water bodies (Coulliette and Noble 2008; Kayhanian, Fruchtmann, et al., 2012; Paule-Mercado, Ventura, et al., 2016). In this study, turbidity and fecal indicator bacteria (FIB), including *E. coli* and total coliforms (TC), in stormwater runoff were evaluated.

Pollutants in roadway runoff may exhibit a “first flush” effect, which is a mass flushing of contaminants in the initial period and may contain higher concentrations of pollutants than those observed during the latter stages of the stormwater runoff (Sansalone and Buchberger 1997; Bertrand-Krajewski, Chebbo, et al., 1998; Deletic 1998; Kayhanian and Stenstrom 2005; Kayhanian and Stenstrom 2008). Therefore, first flush grab samples were collected in the initial rain period in order to eliminate intra-event variability in the study.

Typically, there are two types of runoff from transportation facilities. One is pavement runoff which is exclusively from the pavement, and another is culvert runoff which is including pollutants generated from the pavement and surrounding land uses (Wemple and Jones 2003). Few studies compared water qualities of the two types of runoff. Our main objectives were: (1) evaluating the presence of first flush runoff from pavement and culvert for FIB and turbidity; (2) examining the effect factors for the contaminants in the runoff, including PI, ADP, traffic characteristic, watershed features, and seasonal effects; (3) comparing results for spatial and temporal variability FIB and turbidity.

2. Materials and Methods

2.1 Sample Collection

Stormwater runoff samples were collected from 6 roadway sites in east Tennessee, USA, including a rural minor collector road, a rural minor arterial road, an urban minor collector road, an urban minor arterial road, a rural interstate divided highway, and an urban interstate divided highway (Figure 3.1 and Table 3.1).

At each site, culvert runoff and pavement runoff were collected by Nalgene™ Storm Water Samplers, which were installed on sampling sites one day before the rain. The bottles were washed, rinsed with deionized water, and autoclaved; EPDM gasket, polypropylene ball valve, and closure were washed by 70% ethanol for sterilization. The sampler collects a full liter of first flush stormwater runoff into a Nalgene™ sample bottle, and a floating ball valve automatically closes the bottle when it is full, preventing sample dilution. After sample collection, bottles were transported using a cooler with ice to the laboratory within 6 hours. 23 storm events were monitored from November 2016 to October 2017. ADP and PI at the first hour were recorded through weather history. (<https://www.wunderground.com/history/daily/us/tn/knoxville/KTYS>)

2.2 Sample Analysis

Sample analyses were generally performed within 8 hours of sample collection. Turbidity was measured by the turbidity meter following the manufacturer's instruction (Thermo Scientific™ AQ3010). FIB, including *E. coli* and total coliforms, were determined by growth on the selective and differential membrane filtration m-Colibblue24® broth (Hach Co., Loveland, CO) following the protocol provided by the manufacturer. Briefly, one hundred milliliters of each stormwater runoff sample with appropriate dilutions were filtered, and the then filter was attached on a pad which was soaked with 2 ml of the broth. The filters were incubated at 37 °C for 24 ± 2 hours.

2.3 Statistical analysis

All statistical analyses were performed using R 3.50 (Team 2017). A Kruskal–Wallis test was performed to determine statistical differences between spatial and temporal different samples (Breslow 1970). Spearman Correlation was utilized to explore correlations between water quality parameters and climate factors (Wissler 1905).

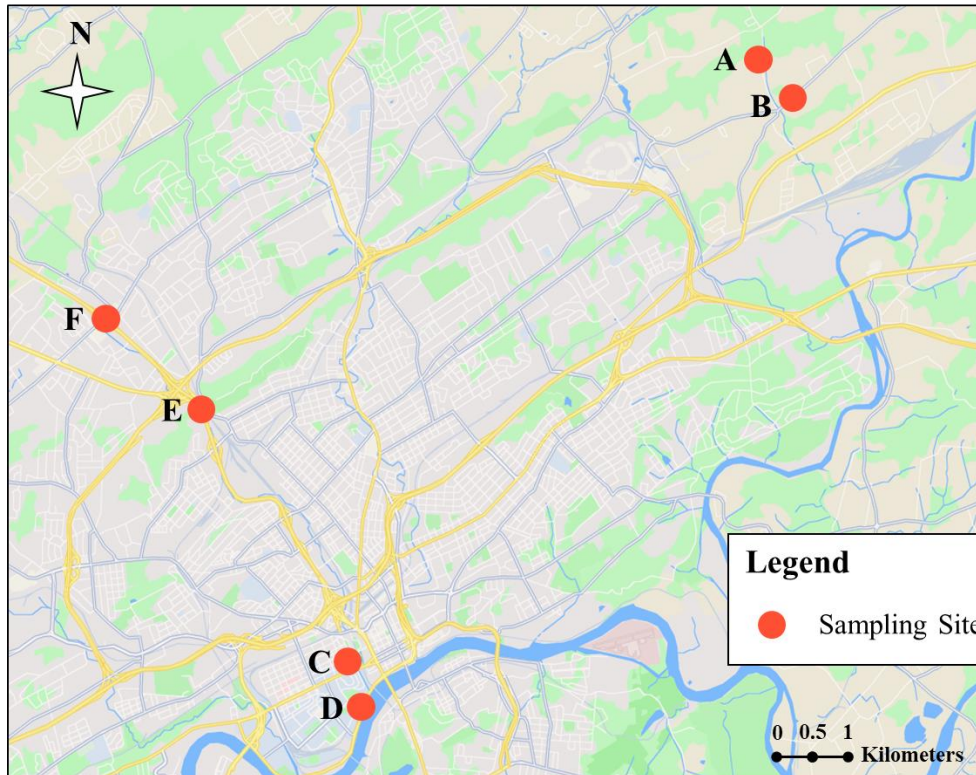


Figure 3.1. Sampling sites. Each sampling site includes both pavement runoff and culvert runoff.

Table 3.1. Summary of sampling locations

Site ID	Road Class	annual average daily traffic	Surrounding land use description
A	Minor Collector	1409	Forest and rural low density agricultural and residential areas
B	Minor Arterial	6721	Forest and rural low density agricultural and residential areas
C	Minor Collector	5473	Developed area of school and residential housing
D	Minor Arterial	16209	Developed area of school and residential housing
E	Interstate divided highway	60852	Forest and low density commercial and residential areas
F	Interstate divided highway	87092	busy commercial area

3. Results and Discussion

Overall, 23 storm events were monitored FIB and turbidity between November 2016 and October 2017, at 6 roadway sites in Knoxville, TN, USA, including A -- a rural minor collector road, B -- a rural minor arterial road, C -- an urban minor collector road, D -- an urban minor arterial road, E -- a rural interstate divided highway, and F -- an urban interstate divided highway (Figure 3.1 and Table 3.1). The two rural local roads, two urban local roads, and two divided highways are closed and within 3 kilometers, respectively. Moreover, the geographical distances between each two of the sites are less than 15 kilometers.

The variations of water quality parameters were huge. The *E. coli* concentrations were from below the limit of detection (10 CFU/100mL) to more than 10^5 CFU/100 mL, the concentrations of TC are from 50 CFU/100 mL to more than 10^6 CFU/100 mL, and the turbidities were from 5 NTU to 120 NTU.

3.1 Correlation among water quality parameters

Correlations between FIB and turbidity were shown in figure 3.2. *E. coli* was not correlated with turbidity (Figure 3.2A). However, *E. coli* and TC were not only correlated highly, but also followed linear regression (R-squared is 0.92) (Figure 3.2B). Of course, TC could not have a correlation with turbidity.

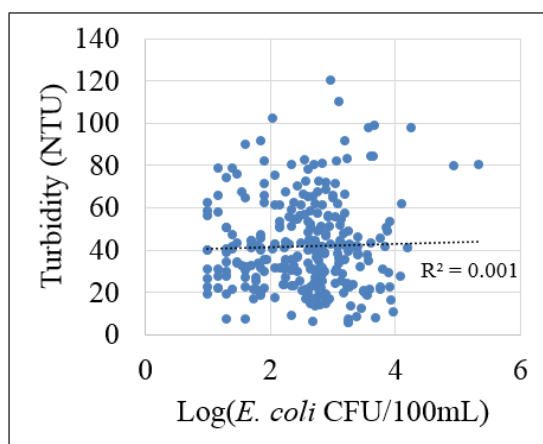
Previous studies showed correlations between FIB and turbidity in water bodies were variable. Some studies found strong correlations between FIB and turbidity, but some showed poor correlations, even though it was widely known that FIB attached to particles and they were transported together (Coulliette and Noble 2008; Kayhanian, Fruchtmann, et al., 2012; Paule-Mercado, Ventura, et al., 2016).

In order to find out the factors that may cause the differential behavior of FIB and turbidity in the stormwater runoff, several potential influences, including PI, ADP, traffic characteristics, watershed features, and seasonal effects, were evaluated in this study.

3.2 The impact of rainfall characteristics

In this study, Pearson correlations were used to figure out the relationship between rainfall characteristics and water quality parameters. Rainfall characteristics were including PI and ADP.

A.



B.

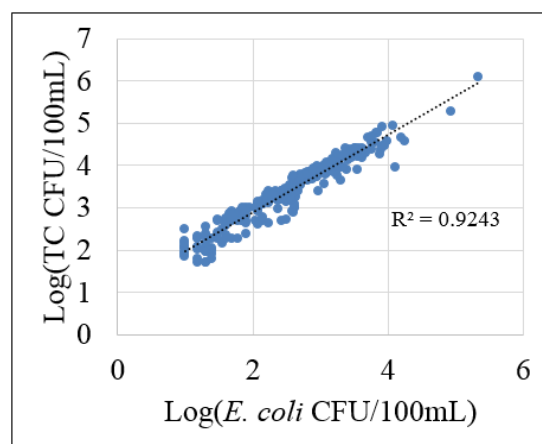


Figure 3.2. Correlations among water quality parameters. (A) The correlation between *E. coli* and turbidity; (B) The correlation between *E. coli* and total coliform.

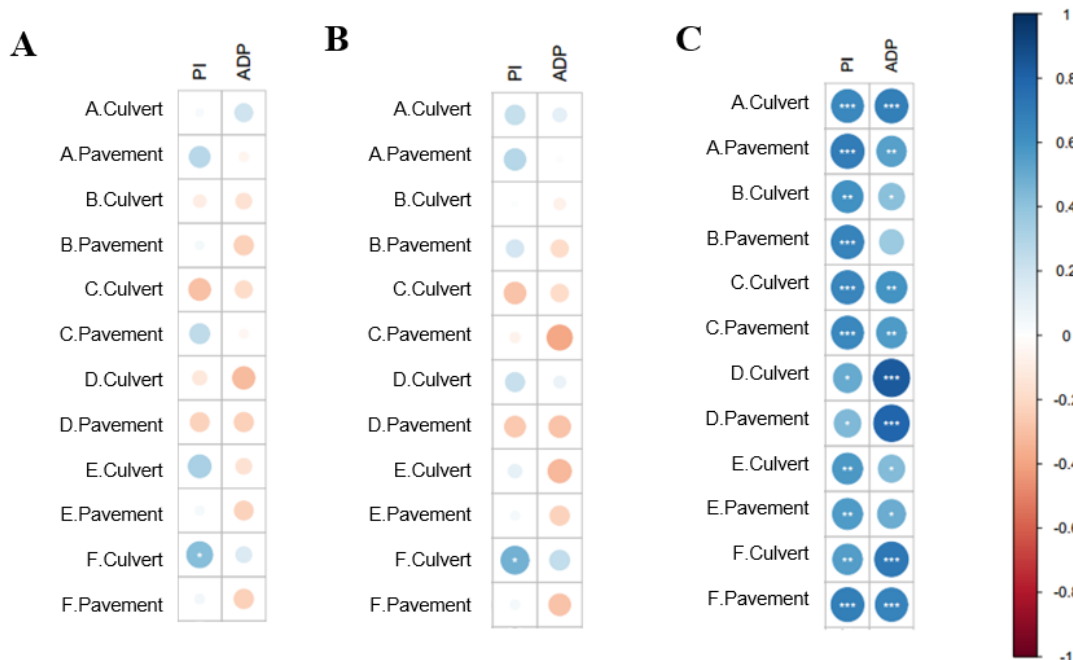


Figure 3.3. Correlation of weather conditions and water quality parameters. (A) *E. coli*, (B) Total coliform, (C) turbidity with PI and ADP at each site. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

Due to high variances of the water quality parameters across sites, even in the same storm events (Figure 3.3 and Figure 3.S1), the Pearson correlation was tested site by site. The results showed both PI and ADP were more positively correlated with turbidities of each site than FIB; lack of correlations was found between FIB and rainfall characteristics. However, the degrees of correlations of rainfall characteristics and turbidities were different at each site. The reason may be that the PI and ADP data were collected from online weather history, and the exact rainfall characteristics varied by sites.

The multiple regression model was also performed to analyze the combined effect of PI and ADP for turbidity (Table 3.2). Although the turbidity of runoff from different sites showed varied significant degrees of PI and ADP, all coefficients of determination (R^2) for this model were higher than 0.38 with a p-value less than 0.01. Moreover, the culvert runoff and pavement runoff from the same site were shown the similar significant levels and coefficients of PI and ADP, and this is further explained that the pavement runoff and culvert runoff are both impacted by rainfall features which were highly similar in the same storm event at the same site.

Although turbidities between sites were shown significant correlations, no significant correlations between FIB levels at any locations, even the pavement runoff and culvert runoff from the same site could not be found the correlations of FIB (Figure 3.S2). This further supports the assertion that the rainfall features may not impact the FIB transport patterns.

3.3 Seasonal variations

Seasonal variations in FIB are shown in figure 3.5 and figure 3.6. Cumulative probability plots showed huge variations of FIB within winter, the range of *E. coli* concentrations was from below the limit of detection (10 CFU/100mL) to more than 10^3 CFU/100 mL, and the range of TC concentrations was from 50 CFU/100mL to more than 10^4 CFU/100 mL, but *E. coli* and TC both had the lowest concentrations during winter compared to other seasons based on the results of the Kruskal-Wallis test.

Although FIB concentrations in winter were significantly lower than in spring, summer, and fall, there is no significant difference among spring, summer, and fall, except the slightly significant difference between spring and summer, which may be because early spring or icy spring showed the similar FIB patterns with winter. Previous studies also showed that FIB concentrations in

Table 3.2. Multiple Regression of Turbidity.

site	Runoff type	Variable 1 -- PI		Variable 2 -- ADP		Overall	
		Coef	P-value	Coef	P-value	R-Squared	P-value
A	Culvert	74.558	0.00631	4.344	0.00228	0.6344	0.00004266
	Pavement	103.007	0.00171	3.034	0.04899	0.5768	0.0001844
B	Culvert	95.484	0.01051	2.029	0.25209	0.4105	0.005068
	Pavement	111.689	0.00266	1.172	0.48781	0.452	0.00244
C	Culvert	95.08	0.00449	3.851	0.01913	0.5733	0.0001999
	Pavement	98.102	0.00688	3.719	0.03624	0.53	0.0005256
D	Culvert	41.648	0.09893	7.349	0.00000718	0.7332	0.000001826
	Pavement	29.562	0.2468	6.466	0.000051	0.6548	0.00002405
E	Culvert	113.482	0.0218	3.141	0.191	0.3825	0.008059
	Pavement	114.094	0.0256	4.16	0.0992	0.4123	0.00492
F	Culvert	57.512	0.045714	5.309	0.000961	0.6013	0.0001014
	Pavement	119.649	0.00214	5.703	0.00356	0.6532	0.00002513

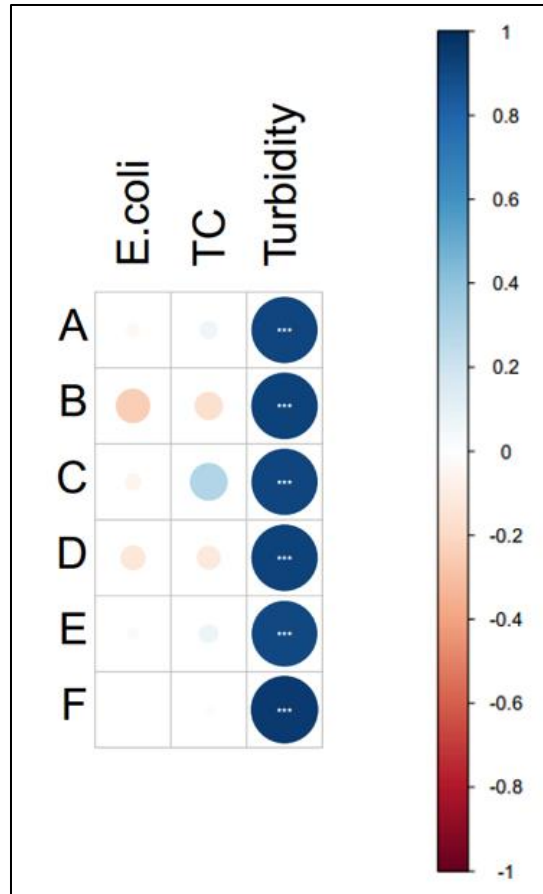


Figure 3.4. Correlation of *E. coli*, Total coliform, and turbidity between pavement runoff and culvert runoff at each site. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

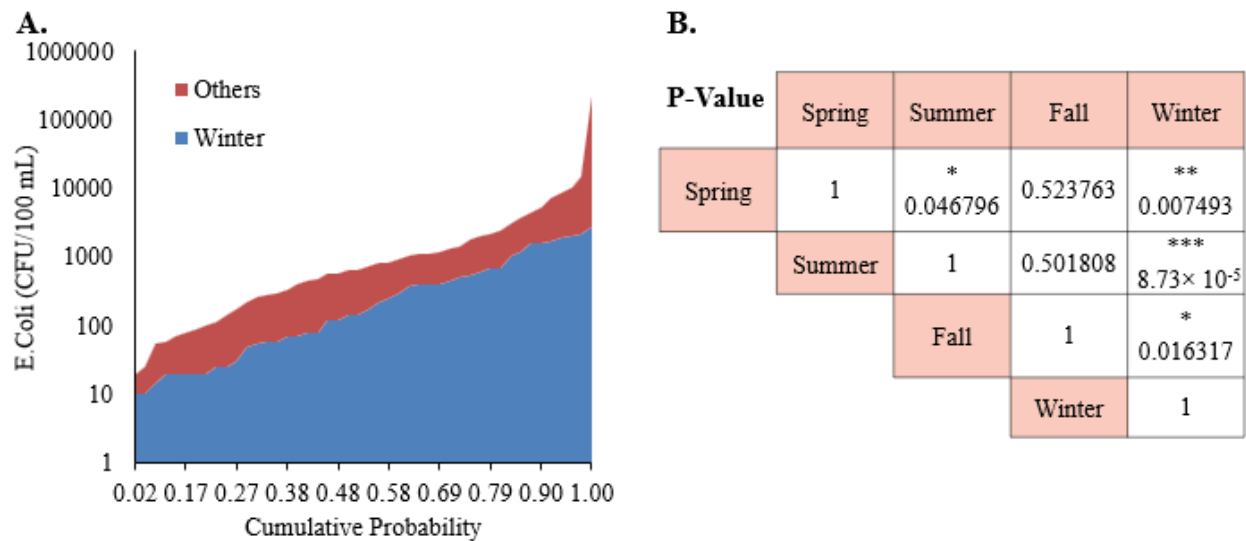


Figure 3.5. Seasonality of *E. coli* concentrations; (A) Cumulative probability plots for Winter and other seasons; (B) P values for two-group comparisons done by T-Test. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

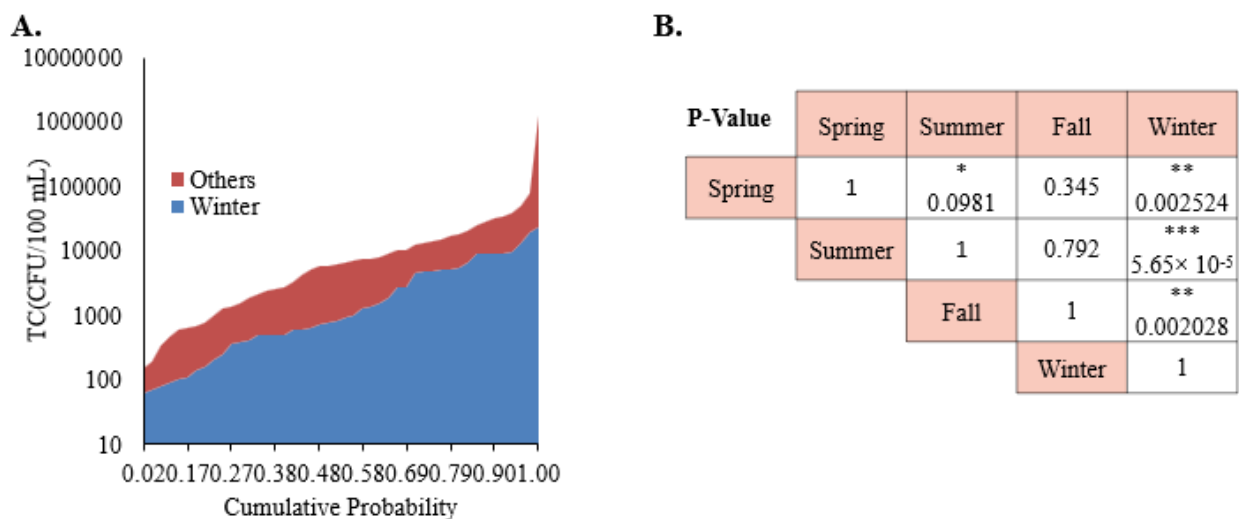


Figure 3.6. Seasonality of total coliform concentrations; (A) Cumulative probability plots for Winter and other seasons; (B) P values for two-group comparisons done by T-Test. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

stormwater runoff were lowest in winter (McCarthy, Mitchell, et al., 2006; Selvakumar and Borst 2006; Krometis, Characklis, et al., 2010). However, there is no seasonal difference in turbidity.

3.4 The impact of site features

To identify the impact of site features on water qualities of runoff, we analyzed the water quality in two potential forces: surrounding land use pattern and annual average daily traffic (AADT).

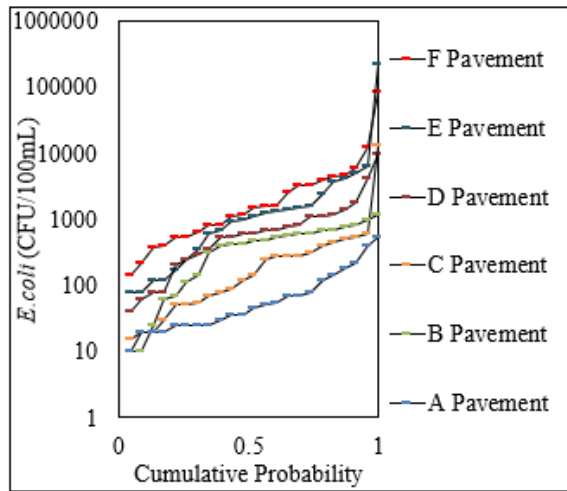
Surrounding land use patterns were including rural and urban. Both surrounding land use pattern and AADT were reported for the impact on water quality: the urban sites that had higher human activities might increase pollutants concentration in stormwater runoff, and higher AADT caused a higher probability to transport contamination into runoff (Sansalone and Buchberger 1997; Deletic and Maksimovic 1998; Shinya, Tsuchinaga, et al., 2000; Kayhanian, Singh, et al., 2003; Aryal, Furumai, et al., 2005; Kayhanian, Fruchtmann, et al., 2012). In this study, the FIB concentrations are higher at the urban minor collector than the rural minor collector, and higher at urban minor arterial than the rural minor arterial, but no significant difference between rural and urban highway (Figure 3.7 – 3.10).

The relationships between AADT and FIB were also analyzed. Table 3.1 was shown AADT in each site, from lowest to highest, the rank is A, C, B, D, E, and F. In this study, the sites with higher AADT appeared higher FIB concentrations in runoff, although there were a few exceptions. There is no significant difference in FIB in the stormwater runoff between D and E sites, and E and F sites, but a significant difference showed between D and F sites. The FIB concentrations seemed increasing limited in the runoff from the higher AADT sites, like D, E, and F. The reason may be the AADT could not affect FIB concentrations after it arrived at a certain high level. More sites need to be studied in the future to determine the hypothesis. Although we found the relationships between site features and FIB, surrounding land use pattern and AADT seemed no impact on the turbidity in stormwater runoff.

4. Conclusions

Stormwater runoff, including pavement runoff and culvert runoff, were monitored over 23 storm events at 6 sites for *E. coli*, TC, and turbidity. All first flush grab samples were collected in the initial rain period for all storm events in order to eliminate intra-event variability. Results showed great variations of these parameters among different sites and different storm events. But there

A.

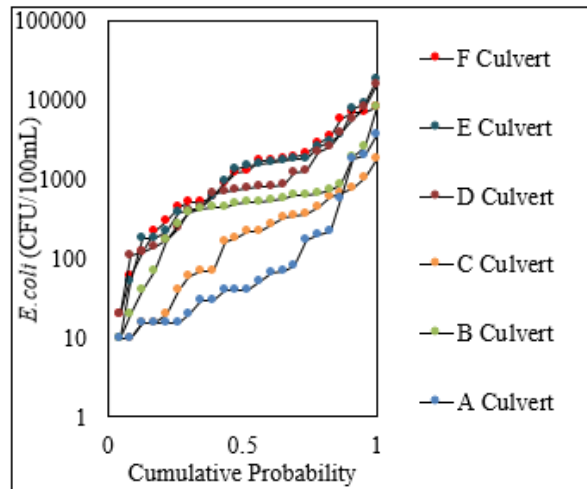


B.

P-Value	A	C	B	D	E	F
A	1	** 7.41×10^{-3}	*** 5.67×10^{-7}	*** 1.37×10^{-11}	*** 1.24×10^{-12}	*** 6.44×10^{-16}
C		1	* 1.44×10^{-2}	*** 7.64×10^{-6}	*** 2.57×10^{-7}	*** 9.19×10^{-10}
B			1	* 2.43×10^{-2}	** 1.16×10^{-3}	*** 2.20×10^{-5}
D				1	0.200	* 2.26×10^{-2}
E					1	0.438
F						1

Figure 3.7. The relationship between *E. coli* concentrations in pavement runoff and road features. (A) Cumulative probability plots for each site; (B) P values for two-group comparisons done by T-Test. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

A.

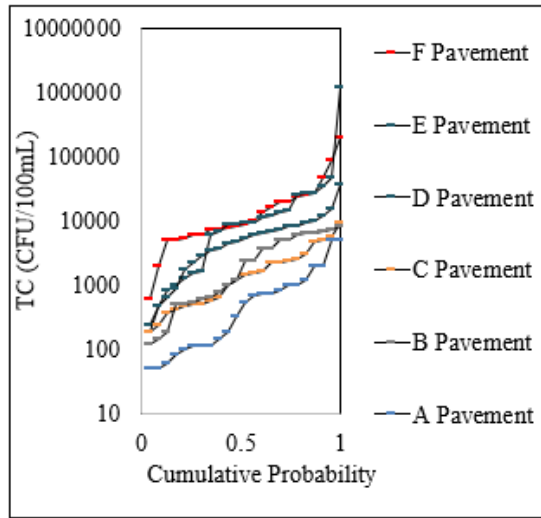


B.

P-Value	A	C	B	D	E	F
A	1	** 8.22×10^{-3}	*** 2.42×10^{-5}	*** 2.43×10^{-10}	*** 3.57×10^{-11}	*** 4.14×10^{-15}
C		1	* 2.01×10^{-2}	*** 4.89×10^{-5}	*** 8.79×10^{-6}	*** 8.12×10^{-11}
B			1	* 2.27×10^{-2}	** 1.98×10^{-3}	*** 4.62×10^{-6}
D				1	0.068	* 1.01×10^{-2}
E					1	0.811
F						1

Figure 3.8. The relationship between *E. coli* concentrations in culvert runoff and road features; (A) Cumulative probability plots for each site; (B) P values for two-group comparisons done by T-Test. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

A.

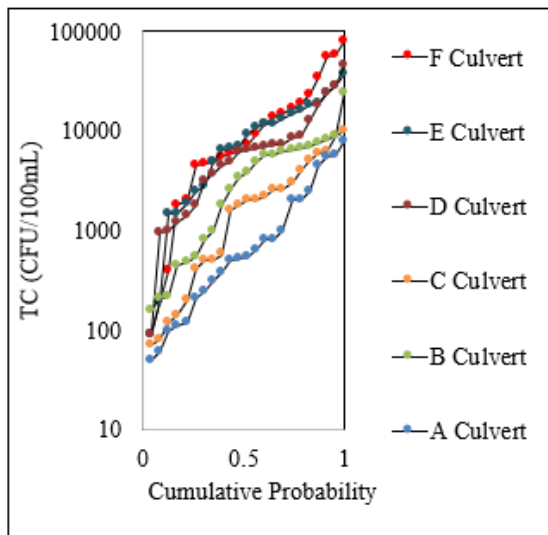


B.

P-Value	A	C	B	D	E	F
A	1	** 1.67×10^{-3}	*** 2.53×10^{-6}	*** 3.14×10^{-10}	*** 1.24×10^{-13}	*** 2.17×10^{-14}
C		1	0.073	*** 6.42×10^{-5}	*** 4.11×10^{-8}	*** 7.22×10^{-10}
B			1	** 1.76×10^{-3}	*** 8.24×10^{-5}	*** 4.76×10^{-9}
D				1	0.124	1.09×10^{-2}
E					1	0.527
F						1

Figure 3.9. The relationship between total coliform concentrations in pavement runoff and road features. (A) Cumulative probability plots for each site; (B) P values for two-group comparisons done by T-Test. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

A.



B.

P-Value	A	C	B	D	E	F
A	1	** 2.53×10^{-3}	*** 1.80×10^{-5}	*** 2.84×10^{-11}	*** 7.10×10^{-12}	*** 1.01×10^{-14}
C		1	0.113	*** 7.56×10^{-6}	*** 6.17×10^{-7}	*** 1.98×10^{-9}
B			1	** 3.68×10^{-3}	*** 2.26×10^{-4}	*** 2.70×10^{-6}
D				1	0.201	1.12×10^{-2}
E					1	0.392
F						1

Figure 3.10. The relationship between total coliform concentrations in culvert runoff and road features. (A) Cumulative probability plots for each site; (B) P values for two-group comparisons done by T-Test. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

were still some factors that appeared to affect the parameters, including PI, ADP, seasonality, surrounding land use patterns, and AADT.

The Pearson correlation and multiple linear regression both showed turbidity had significant relationships with the rainfall features (PI and ADP), though not all sites shared the same significant levels. It is possible that the dust may accumulate on the surface so that the longer ADP causes higher turbidity in runoff; the higher PI has the stronger power to wash off dust and soil from the surface, and then the turbidity in runoff was getting greater. However, FIB, both *E. coli* and TC, were found no relationship with PI and ADP. It is possible that FIB may not be persistent and accumulate on surfaces with dust and soil.

Seasonal differences for FIB were noted in this study. Although the variations of FIB concentration were great in each season, winter FIB concentrations were lowest compared to other seasons based on the results of the Kruskal-Wallis test. According to other similar studies, these seasonal variations of FIB are significant and should be considered for the stormwater best management in the future (Selvakumar and Borst 2006; Hallberg, Renman, et al., 2007).

Surrounding land use patterns and AADT were also related to FIB concentrations in runoff. The results in this study suggest FIB concentrations may increase in urban and high AADT sites. Previous studies also showed surrounding land use patterns and AADT affected other water quality parameters, like heavy metals, PAHs, and nutrients (Andrews and Sutherland 2004; Gnecco, Berretta, et al., 2005; Klimaszewska, Polkowska, et al., 2007; Maniquiz, Lee, et al., 2010; Kayhanian and Paytan 2011; Kayhanian 2012). Thus, public health risks may increase in urban and high AADT sites. Future studies about pathogen contaminations in runoff should focus more on these sites.

There is no correlation between FIB and turbidity, and all these parameters should be monitored in the storm-water best management programs. This study only monitored 6 sites in one city, and further research is needed to confirm the findings in this study for other places.

References

- Andrews S, Sutherland RA (2004) Cu, Pb and Zn contamination in Nuuanu watershed, Oahu, Hawaii. *Science of the Total Environment* 324:173-182
- Aryal RK, Furumai H, Nakajima F, Boller M (2005) Dynamic behavior of fractional suspended solids and particle-bound polycyclic aromatic hydrocarbons in highway runoff. *Water Research* 39:5126-5134
- Athayde DN, Shelley PE, Driscoll ED, Gaboury D, Boyd G (1983) "Results of the nationwide urban runoff program". *Results of the Nationwide Urban Runoff Program*
- Bertrand-Krajewski JL, Chebbo G, Saget A (1998) Distribution of pollutant mass vs volume in stormwater discharges and the first flush phenomenon. *Water Research* 32:2341-2356
- Breslow N (1970) A generalized Kruskal-Wallis test for comparing K samples subject to unequal patterns of censorship. *Biometrika* 57:579-594
- Coulliette AD, Noble RT (2008) Impacts of rainfall on the water quality of the Newport River Estuary (Eastern North Carolina, USA). *Journal of Water and Health* 6:473-482
- Deletic A (1998) The first flush load of urban surface runoff. *Water Research* 32:2462-2470
- Deletic AB, Maksimovic ČT (1998) Evaluation of water quality factors in storm runoff from paved areas. *Journal of Environmental Engineering* 124:869-879
- Driscoll ED, Shelley PE, Strecker EW (1990) Pollutant Loadings and Impacts From Highway Stormwater Runoff Volume III: Analytical Investigation and Research Report. Pollutant Loadings and Impacts from Highway Stormwater Runoff Vol III: Analytical Investigation and Research Report
- Gnecco I, Berretta C, Lanza LG, La Barbera P (2005) Storm water pollution in the urban environment of Genoa, Italy. *Atmospheric Research* 77:60-73
- Grant M (1997) A new membrane filtration medium for simultaneous detection and enumeration of *Escherichia coli* and total coliforms. *Appl Environ Microbiol* 63:3526-3530
- Hallberg M, Renman G, Lundbom T (2007) Seasonal variations of ten metals in highway runoff and their partition between dissolved and particulate matter. *Water, Air, and Soil Pollution* 181:183-191
- Han Y, Lau SL, Kayhanian M, Stenstrom MK (2006a) Characteristics of highway stormwater runoff. *Water Environment Research* 78:2377-2388
- Han YH, Lau SL, Kayhanian M, Stenstrom MK (2006b) Correlation analysis among highway stormwater pollutants and characteristics. *Water Science and Technology* 53:235-243
- Kayhanian M (2012) Trend and concentrations of legacy lead (Pb) in highway runoff. *Environmental Pollution* 160:169-177
- Kayhanian M, Fruchtmann BD, Gulliver JS, Montanaro C, Ranieri E, Wuertz S (2012) Review of highway runoff characteristics: Comparative analysis and universal implications. *Water research* 46:6609-6624

- Kayhanian M, Paytan A (2011) Isotopic fingerprinting approach for nutrient source identification. Proceedings of 10th Annual StormCon Conference, August 21-25, Anaheim, CA
- Kayhanian M, Singh A, Suverkropp C, Borroum S (2003) Impact of annual average daily traffic on highway runoff pollutant concentrations. *Journal of Environmental Engineering* 129:975-990
- Kayhanian M, Stenstrom MK (2005) Mass loading of first flush pollutants with treatment strategy simulations *Transportation Research Record*. p 133-143
- Kayhanian M, Stenstrom MK (2008) First flush highway runoff characterization: for stormwater runoff treatment. *Stormwater* 9:32-45
- Kayhanian M, Suverkropp C, Ruby A, Tsay K (2007) Characterization and prediction of highway runoff constituent event mean concentration. *Journal of Environmental Management* 85:279-295
- Klimaszewska K, Polkowska Z, Namieśnik J (2007) Influence of mobile sources on pollution of runoff waters from roads with high traffic intensity. *Polish Journal of Environmental Studies* 16:889-897
- Krometis LAH, Characklis GW, Drummey PN, Sobsey MD (2010) Comparison of the presence and partitioning behavior of indicator organisms and *Salmonella* spp. in an urban watershed. *Journal of Water and Health* 8:44-59
- Li MH, Barrett ME (2008) Relationship between antecedent dry period and highway pollutant: Conceptual models of buildup and removal processes. *Water Environment Research* 80:740-747
- Maniquiz MC, Lee S, Kim LH (2010) Multiple linear regression models of urban runoff pollutant load and event mean concentration considering rainfall variables. *Journal of Environmental Sciences* 22:946-952
- McCarthy D, Mitchell VG, Deletic A (2006) *Escherichia coli* levels in urban stormwater. Paper Presented to the Urban Drainage Modelling and International Water Sensitive Urban Design Conference, Melbourne, 4–6 April 2006
- Paule-Mercado M, Ventura J, Memon S, Jahng D, Kang J-H, Lee C-H (2016) Monitoring and predicting the fecal indicator bacteria concentrations from agricultural, mixed land use and urban stormwater runoff. *Science of the total environment* 550:1171-1181
- Sansalone JJ, Buchberger SG (1997a) Characterization of solid and metal element distributions in urban highway stormwater. 36:155-160
- Sansalone JJ, Buchberger SG (1997b) Partitioning and first flush of metals in urban roadway storm water. *Journal of Environmental Engineering* 123:134-143
- Selvakumar A, Borst M (2006) Variation of microorganism concentration in urban stormwater runoff with land use and seasons. *Journal of Water and Health* 4:109-124
- Shinya M, Tsuchinaga T, Kitano M, Yamada Y, Ishikawa M (2000) Characterization of heavy metals and polycyclic aromatic hydrocarbons in urban highway runoff. vol 42. *Int Water Assoc, United Kingdom*, p 201-208
- Stotz G (1987) Investigations of the properties of the surface water run-off from federal highways in the FRG. *Science of the Total Environment, The* 59:329-337

Thomson NR, McBean EA, Snodgrass W, Monstrenko IB (1997) Highway stormwater runoff quality: Development of surrogate parameter relationships. *Water, Air, and Soil Pollution* 94:307-347

Wemple BC, Jones JA (2003) Runoff production on forest roads in a steep, mountain catchment. *Water Resources Research* 39

Westerlund C, Viklander M, Bäckström M (2003) Seasonal variations in road runoff quality in Luleå, Sweden *Water Science and Technology*. vol 48, p 93-101

Wissler C (1905) The Spearman correlation formula. *Science* 22:309-311

Chapter 4

Persistence of Environmental *E. coli* Strains from Stormwater Runoff

A version of this chapter was originally submitted for publication by Songyi Liu and Qiang He:

Citation

S. Liu, Q. He (2019)." Persistence of Environmental *E. coli* Strains from Stormwater Runoff ". In: the 28th Tennessee Water Resources Symposium, Nashville, TN.

Abstract

Escherichia coli has been used as an indicator for fecal contamination, but high concentrations of *E. coli* were commonly found in stormwater runoff collected from highway viaducts where fecal materials were rarely reported. It has recently been recorded that the capability of *E. coli* to survive for a long time in some environments, which suggests the detection of *E. coli* with no recent exposure to fecal contaminants. The objective of this study was to evaluate the survival of *E. coli* strains isolated from stormwater runoff under environmentally relevant conditions. One laboratory *E. coli* strain (MG1655) was tested as a control. The *E. coli* strains were exposed to four environmental factors: (i) presence or absence of sunlight; (ii) solid surface versus liquid media for growth; (iii) different temperatures, including 4 °C, 20 °C, and 35 °C; (iv) different levels of relative humidity, i.e., 13%, 50%, and 100%. The environmental strains isolated from stormwater runoff persisted for a longer period of time than the lab strain under all environmental conditions tested. On the solid surface, the lab strain completely lost viability in less than 1 hour at the relative humidity of 13% and 50%. In comparison, the environmental strains were able to remain viable for a prolonged period with the maximum $T_{99.9}$ of nearly 15 days, which is congruent with the persistence of these strains in natural environments; however, the viable counts were limited with the maximum T_{90} of 0.25 days, which illustrates high *E. coli* concentrations in highway runoff have a low probability from the persistence on the pavement. As expected, the control and wild-type *E. coli* strains survived longer at lower temperatures with sufficient water. Additionally, it was found that sunlight exposure resulted in the rapid inactivation of all strains tested in this study.

Findings from this study could provide insight into the interpretation of environmental monitoring data using *E. coli* as the fecal indicator.

1. Introduction

Stormwater runoff is a concern as a non-point source of fecal contamination in receiving water bodies (Cabelli, Dufour, et al., 1982; Jiang 2004; Marsalek and Rochfort 2004; McCarthy, Mitchell, et al., 2006; Surbeck, Jiang, et al., 2006). As rainfall washes over land, it picks up and transports a variety of pollutants, including human and animal fecal waste (Athayde, Shelley, et al., 1983). In order to measure potential fecal contamination and pathogens of enteric origins, *Escherichia coli* (*E. coli*) as an indicator organism for fecal contamination has been used to indicate the potential for the presence of pathogens of fecal origin (Grant 1997; Haile, Witte, et al., 1999).

Numerous studies reported consistently high concentrations of *E. coli* in stormwater runoff (Driscoll, Shelley, et al., 1990; Berhanu Desta, Bruen, et al., 2007; Yufen, Xlaoke, et al., 2008). Chapter 2 in this study also showed high concentrations of *E. coli*, especially in highway runoff, 71% of samples revealed the concentrations were higher than 1,000 CFU/100 mL. However, the runoff samples were collected from highway viaducts where fecal contaminants were rarely reported. Moreover, researchers have found that the capability of *E. coli* to survive for a long time in the environment, which suggests the detection of *E. coli* with no recent exposure to fecal materials (Hoff and Akin 1986; England, Lee, et al., 1993; Sjogren 1994). Therefore, we hypothesized that *E. coli* strains isolated from the highway runoff were robust in relevant environmental conditions.

The environmental factors that affect the persistence of *E. coli* include temperature, sunlight exposure, moisture content, and nutrients (Buchanan and Edelson 1999; Himathongkham, Bahari, et al., 1999; Tao, Bausch, et al., 1999; Rozen and Belkin 2001; Ihssen and Egli 2004; Franz, Van Diepeningen, et al., 2005; Large, Walk, et al., 2005; Bergholz and Whittam 2007; Ihssen, Grasselli, et al., 2007; Semenov, Van Bruggen, et al., 2007; Tabe, Oloya, et al., 2008; Vital, Hammes, et al., 2008; Sidhu, Toze, et al., 2015). Furthermore, different types of *E. coli* strains have been shown to have different survival capabilities in the same environmental conditions (Perkins, Perrow, et al., 2016).

E. coli strains that were isolated in pavement runoff as wild-type *E. coli* may become inactivate rapidly due to harsh environmental conditions such as high temperatures, sunlight exposure, and the loss of moisture on the surface (Rozen and Belkin 2001; Sinton, Hall, et al., 2002; Oliver, Clegg, et al., 2007; Gilpin, Devane, et al., 2013). However, certain conditions like low temperature, the shade portion avoiding sunlight, and receiving water body providing water content, may prolong the survival of *E. coli*. Perhaps, the wild-type *E. coli* could evolve to stand harsh environmental conditions and then survive in a prolonged period (Kiefer, Shelton, et al., 2012). However, very little has been documented about the persistence of wild-type *E. coli* strains from stormwater runoff on environmental surfaces, in the runoff water, and the receiving water body.

The objective of this study was to evaluate the survival of wild-type *E. coli* strains isolated from stormwater runoff under environmentally relevant conditions. One laboratory *E. coli* strain (MG1655) was used as a control. The *E. coli* strains were exposed to four environmental factors: (i) different temperatures, including 4 °C, 20 °C, and 35 °C; (ii) presence or absence of sunlight; (iii) solid surface versus liquid matrices for growth; (iv) different levels of relative humidity (RH), i.e., 13%, 50%, and 100%. Our results show that (i) wild-type strains isolated from stormwater runoff are better adapted for survival than the lab strain, (ii) *E. coli* strains survive longer at lower temperature, (iii) strong sunlight exposure results in the rapid inactivation of all strains, and (iv) *E. coli* strains survive longer under higher water content than more arid conditions.

2. Materials and Methods

2.1 Study site

Stormwater runoff samples were collected from a highway viaduct site without reports of fecal materials, and at least three days without precipitation were required prior to sampling. Sampling was conducted using Nalgene™ Storm Water Samplers, which were installed immediately prior to the anticipated storms. Each sampler collected a full one-liter grab sample representing first-flush stormwater runoff. The runoff sample was the source of wild-type strains and the stormwater runoff matrix for growth.

The stream water matrix for growth was from the runoff receiving stream, which was collected by submerging a sterile 1 L bottle below the stream surface according to standard methods (Lurry and

Kolbe 2000) after at least three antecedent dry days for represented typical base flow. The water samples were stored in the dark at 4 °C and used 4 months after collection.

2.2 Preparation of bacterial strains and inoculum

The persistence of *E. coli* was tested with wild-type and control (lab strain, MG1655) strains of *E. coli*. Strain MG1655 appeared to be one of the most popular *E. coli* in metabolic engineering due to identified genome sequence and unessential mutations (Biriukova, Krylov, et al., 2010). Because of the background levels of *E. coli* in raw water being too low to quantify during inactivation experiments, additional *E. coli* populations were seeded in these matrices.

Wild-type strains of *E. coli* were isolated from roadway runoff using m-Colibblue 24 ® media following the protocol provided by the manufacturer (Hach, Loveland, CO). Five *E. coli* colonies were then aseptically picked. In order to get pure strains, the colonies were diluted with phosphate buffer saline (PBS) and cultivated on m-Colibblue 24 ® media, with the resulting pure colonies being further cultivated in tryptic soy broth (TSB) for 12 hours at 37°C. At the time point, OD₆₀₀ was around 1.2 AU, and bacterial culture enters a stationary phase. The media were autoclaved prior to use.

Before seeding the strains, 1 mL of the bacterial cultures were centrifuged at 12000g for 5 min, and then the supernatant was discarded, followed by the washing of the pellets in PBS twice to remove culture media. All these colonies were confirmed as *E. coli* by the high-throughput 16S rRNA amplicon sequencing method as described previously (Wyckoff, Chen, et al., 2017). Due to the limitation of resolution of the 16S rRNA amplicon sequencing method, only the species level was able to be determined, and the specific strains for wild-type strains of *E. coli* were unknown.

2.3 Bacterial Inoculation

In order to test the effect of different environmental matrices for the *E. coli* persistence, asphalt pavement, polystyrene plastic, stormwater runoff, and stream water were used as environmental matrices. The asphalt pavement was compacted to approximately an air void content 4% using a Superpave gyratory compactor (SGC). The asphalt binder was performance grade (PG) 64-22, and the asphalt content is 6.2%. Limestone aggregates were used in the asphalt concrete with a

nominal maximum aggregate size of 9.5 mm (Brown, Kandhal, et al., 2009). The asphalt pavement test surfaces had a 15cm diameter, and the polystyrene plastic test surfaces had a 6cm diameter. The polystyrene plastic surface served as a neutral material, while the asphalt contained metals and hydrocarbons that may be harmful to *E. coli* survival (McConkey, Duxbury, et al., 1997). Before testing bacterial persistence on these surfaces, the asphalt pavement and polystyrene plastic were sterilized with 70% ethanol for 10 min and then rinsed with distilled water. After evaporation, the surface was marked into two semicircles using a sterile knife. Each sector was inoculated with 0.1 ml of runoff bacteria suspension and evenly spread to nearly 1 cm².

Stormwater runoff and stream water were used as raw water matrices. The *E. coli* pellets that were collected from the previous steps were seeded into 10mL raw field water samples placed in a 20 mL sterile colorless and transparent glass tube, resulting in a concentration of 10⁶ ~ 10⁷ CFU/mL. The stormwater runoff with the seeded strains acts as runoff bacteria suspension for the test for bacteria survival on the previously described surfaces.

2.4 Controlled environmental chamber

After bacterial inoculation, the control treatments were placed in an environmental growth chamber (EGC C3) to determine the effect of temperature, humidity, and light intensity. The different temperatures, included 4 °C, 20 °C, and 35 °C; the different levels of RH, including 13%, 50%, and 100%; the light control includes 0 and 675 µmoles/m²/sec. Before experiments, the chamber was operated for at least 1 hour to reach desired condition levels.

2.5 Microbial recovery and enumerating

Bacterial samples from asphalt pavement surfaces and polystyrene plastic surfaces were collected by a swabbing method that has previously been described (Bokulich, Mills, et al., 2013). Briefly, each sector was swabbed horizontally and vertically three times using sterile polyester-tipped applicators. Prior to swabbing the surface, the tip was soaked in a sterile PBS solution to facilitate the removal of organic matter present on the inoculated surface. After swabbing, swab tips were cut off using sterilized scissors and placed in tubes containing 10 ml of PBS. Tubes were vortexed twice for 1 min each, with a waiting time in between of 1 min. By doing this, cells trapped in the swab tip are released into the PBS, making them available to be enumerated by m-Colibblue24 via membrane filtration (Grant 1997).

The water samples also used m-Colibblue24 media to enumerate *E. coli*, using the appropriate water volumes for filtration. The first sampling time (T_0) immediately followed inoculation, and that count is the initial *E. coli* concentration (N_0). The following sampling time was the 1st hour if the number of bacteria was below the detection limit, the sampling was stopped if the count was lower than a half N_0 , which means bacteria inactivated rapidly, and the sampling continued at the 2nd hour, 3rd hour, or stopped if the number of bacteria fell below the detection limit. If the count at the 1st hour was more than half N_0 , the sampling continued at the 1st day, 3rd day, 5th day, 7th day, or stopped if the number of bacteria fell below the detection limit. If the bacteria still survived after 7 days, the final sampling time would be on the 10th day.

2.6 Statistical analysis and data modeling

Strain MG1655 with three replicates and the five wild-type strains with two replicates under each condition were tested in this study, and the final results showed average values with error bars for lab strain and wild-type strain, respectively. If the time when the number of bacteria was below the detection limit was more than 3 hours, a decay model would be chosen to fit the experimental data. Models were fitted to the data by using the GInaFiT toolbox (Geeraerd, Valdramidis, et al., 2006), including

the first order decay model, which is a commonly used primary classic linear model (Bigelow and Esty 1920):

$$C(t) = C_0 e^{-kt} \quad (1)$$

where t is time (day), C (CFU/mL or CFU/cm²) is the measured concentration at time t , C_0 is the measured concentration at time 0, and k [hour⁻¹ or day⁻¹] is the rate constant.

The Weibull survival model when the curves did not follow log-linear kinetics and showed asymptotic tails (Mafart, Couvert, et al., 2002):

$$C(t) = C_0 \log_{10}^{-\left(\frac{t}{\alpha}\right)^\beta} \quad (2)$$

α (scale parameter) represents the first decimal reduction time (days), and β is the shape parameter. For $\beta < 1$, a concave curve is obtained, while for $\beta > 1$, a convex curve is obtained.

Model performances were estimated by the regression coefficient (R^2) For data analysis. Based on the fitted model, T_{90} , T_{99} and $T_{99.9}$ were calculated, which were the time to reduce the initial inoculum (C_0) 10-, 100-, and 1,000-fold, respectively.

3. Results and Discussion

In order to identify the persistence of *E. coli* strains, 5 potential ecological selective forces were tested: strain types, temperatures, sunlight intensities, habitat matrices, and relative humidities. Treatments where bacterial counts dropped below the detection limit within the first day, are considered the low persistence conditions (Table 4.1), while if the bacterial count of the treatment was detectable after more than 1 day, the treatment was considered as the higher persistence conditions (Table 4.2 and 4.3).

3.1 The low persistence conditions

As expected, all *E. coli* strains survived for a shorter time at higher temperatures, especially at 35°C, where the counts of bacteria were below the detection limit within 1 hour (Table 4.1). While *E. coli* strains survived longer at lower temperatures, RH and sunlight exposure were also factors that restricted the persistence of *E. coli* strains (Sjogren 1994; Hacker and Carniel 2001; Sinton, Hall, et al., 2002). At both 20°C and 4°C, with sunlight exposure of 675 $\mu\text{moles}/\text{m}^2/\text{sec}$, $T_{99.9\text{s}}$ were less than 3 hours for both sets of *E. coli* strains, and less than 1 hour for the *E. coli* strains that were tested on the solid surfaces with 13% and 50% RH.

At lower temperatures and without sunlight exposure, the factor that influenced the persistence of *E. coli* strains was the RH. The lab strain of *E. coli* could not survive with 13% and 50% RH for both asphalt and polystyrene surfaces. This is consistent with previous studies, where it was shown that *E. coli* strain MG1655 had high decay rates on soil, glass, peanut butter, and wood under arid conditions (Byappanahalli, Whitman, et al., 2006; Ihssen, Grasselli, et al., 2007; Zhang and Yan 2012; He, Li, et al., 2013).

3.2 Desiccation resistance and persistence of *E. coli* on surfaces

Although lab strain *E. coli* could not persist at lower humidity, the wild-type strains could be detected for more than 7 days under these conditions. The decay revealed a concave curve that did not follow log-linear kinetics and showed asymptotic tails, and so the Weibull survival model was

Table 4.1. The low persistence conditions

Strain Type	Temperature (°C)	Media Matrix	Relative Humidity (%)	Light (μmoles/m ² /sec)	T _{99.9}
Lab strain	35	Runoff	--	0	Less than 1 hour
Lab strain	35	Streamwater	--	0	Less than 1 hour
Lab strain	35	Polystyrene	50%	0	Less than 1 hour
Lab strain	35	Asphalt	50%	0	Less than 1 hour
Lab strain	20	Polystyrene	50%	0	Less than 1 hour
Lab strain	20	Asphalt	50%	0	Less than 1 hour
Lab strain	4	Polystyrene	50%	0	Less than 1 hour
Lab strain	4	Asphalt	50%	0	Less than 1 hour
Lab strain	35	Polystyrene	13%	0	Less than 1 hour
Lab strain	35	Asphalt	13%	0	Less than 1 hour
Lab strain	20	Polystyrene	13%	0	Less than 1 hour
Lab strain	20	Asphalt	13%	0	Less than 1 hour
Lab strain	4	Polystyrene	13%	0	Less than 1 hour
Lab strain	4	Asphalt	13%	0	Less than 1 hour
Lab strain	20	Streamwater	--	675	Less than 1 hour
Lab strain	20	Runoff	--	675	Less than 2 hours
Lab strain	4	Streamwater	--	675	Less than 1 hour
Lab strain	4	Runoff	--	675	Less than 2 hours
Wild-type strain	35	Runoff	--	0	Less than 1 hour
Wild-type strain	35	Streamwater	--	0	Less than 1 hour
Wild-type strain	35	Polystyrene	50%	0	Less than 1 hour
Wild-type strain	35	Asphalt	50%	0	Less than 1 hour
Wild-type strain	20	Polystyrene	50%	675	Less than 1 hour
Wild-type strain	20	Asphalt	50%	675	Less than 1 hour
Wild-type strain	4	Polystyrene	50%	675	Less than 1 hour
Wild-type strain	4	Asphalt	50%	675	Less than 1 hour
Wild-type strain	35	Polystyrene	13%	0	Less than 1 hour
Wild-type strain	35	Asphalt	13%	0	Less than 1 hour
Wild-type strain	20	Polystyrene	13%	675	Less than 1 hour
Wild-type strain	20	Asphalt	13%	675	Less than 1 hour
Wild-type strain	4	Polystyrene	13%	675	Less than 1 hour
Wild-type strain	4	Asphalt	13%	675	Less than 1 hour
Wild-type strain	20	Streamwater	--	675	Less than 3 hours
Wild-type strain	20	Runoff	--	675	Less than 3 hours
Wild-type strain	4	Streamwater	--	675	Less than 3 hours
Wild-type strain	4	Runoff	--	675	Less than 3 hours

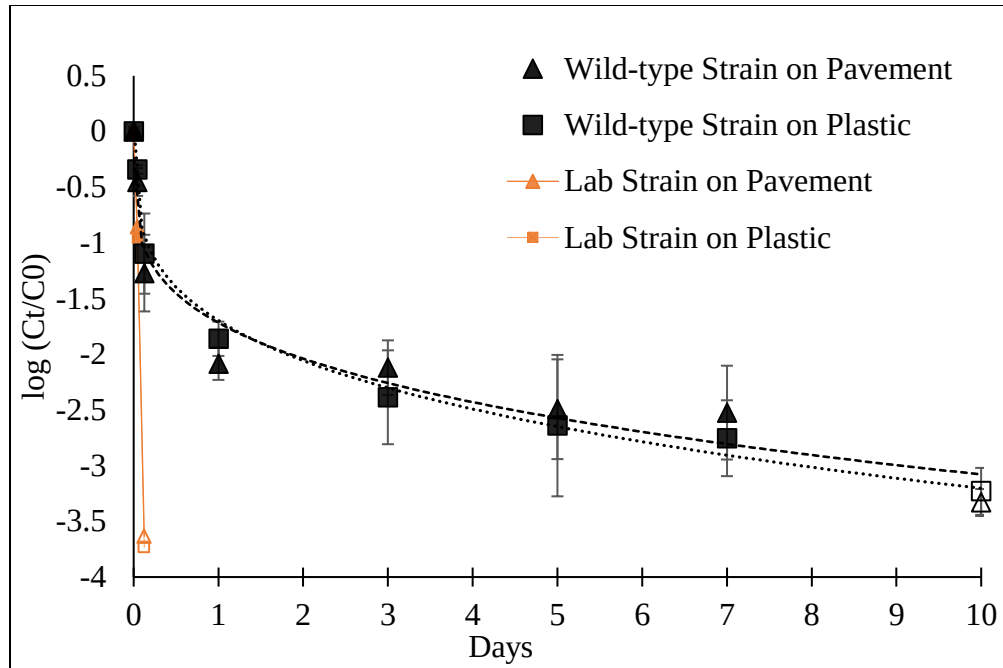


Figure 4.1. *E. coli* strains on solid surfaces at 20°C and 50% RH without sunlight exposure.

Table 4.2. The persistence conditions with the Weibull survival model

Strain Type	Temperature	Media	Relative Humidity	α	β	R^2	T99.9
	(°C)	Matrix	(%)	(days)			days
Wild-type strain	20	Polystyrene	50	0.11	0.25	0.98	8.91
Wild-type strain	20	Asphalt	50	0.17	0.26	0.94	11.63
Wild-type strain	20	Polystyrene	13	0.02	0.17	0.91	12.81
Wild-type strain	20	Asphalt	13	0.01	0.16	0.97	9.59
Wild-type strain	4	Polystyrene	50	0.25	0.27	0.95	14.62
Wild-type strain	4	Asphalt	50	0.21	0.26	0.94	14.36
Wild-type strain	4	Polystyrene	13	0.04	0.19	0.91	12.98
Wild-type strain	4	Asphalt	13	0.04	0.19	0.89	12.98

used to fit the persistent data (Figure 4.1). No surprise, at the lower temperature (4 °C), the wild-type strains survived longer under the same humidity conditions than at 20 °C. On asphalt pavement with 50% RH, T_{90} (α) and $T_{99.9}$ was 0.17 and 11.6 days at 20°C, respectively, while T_{90} and $T_{99.9}$ was 0.21 and 14.4 days, respectively, at 4°C. On polystyrene plastic, with 50% RH, T_{90} and $T_{99.9}$ was 0.11 and 8.9 days at 20°C, respectively, while T_{90} and $T_{99.9}$ was 0.25 and 14.6 days at 4°C, respectively. With 13% RH, T_{90} and $T_{99.9}$ was 0.02 and 12.8 days at 20°C, respectively, while T_{90} and $T_{99.9}$ was 0.04 and 13.0 days at 4°C respectively (Table 4.2).

Because polystyrene plastic is a neutral material and the metals and hydrocarbons in asphalt may be harmful to *E. coli* survival, we hypothesized that *E. coli* strains survived longer on polystyrene plastic than asphalt pavement. However, the results showed the surface materials did not affect the persistence of *E. coli* strains. Other environmental conditions, such as temperature, sunlight exposure, and RH are the essential factors that restrict *E. coli* survival (Figure 4.2).

Compared with 13% and 50% RH, T_{90} s are all smaller in 13% RH than 50% when other controlled conditions were the same, such as 20°C-without sunlight exposure- polystyrene plastic-13% RH VS 20°C -without sunlight exposure- polystyrene plastic-50% RH, 20°C -without sunlight exposure- asphalt pavement -13% RH VS 20°C-without sunlight exposure- asphalt pavement - 50% RH, 4°C-without sunlight exposure- polystyrene plastic-13% RH VS 4°C-without sunlight exposure- polystyrene plastic-50% RH, and 4°C-without sunlight exposure- asphalt pavement -13% RH VS 4°C-without sunlight exposure- asphalt pavement -50% RH. However, the different RHs did not affect $T_{99.9}$. This pattern illustrates that RH only impacts the wild-type strains soon after inoculation, with the lower RH inactivating intolerant bacteria, but the lower RH does not affect the rest of the populations showing the ability of low humidity resistance.

Although the environmental strains were able to remain viable for a prolonged period with the maximum $T_{99.9}$ of nearly 15 days, the counts were limited with the maximum T_{90} of 0.25 days, which illustrates high *E. coli* concentrations in highway runoff have a low probability from the persistence on the pavement. Fecal materials from wild animals or other sources contained high *E. coli* levels may be transferred by vehicles and washed by rainwater. The real reasons that cause high *E. coli* concentrations in highway runoff should be determined in the future.

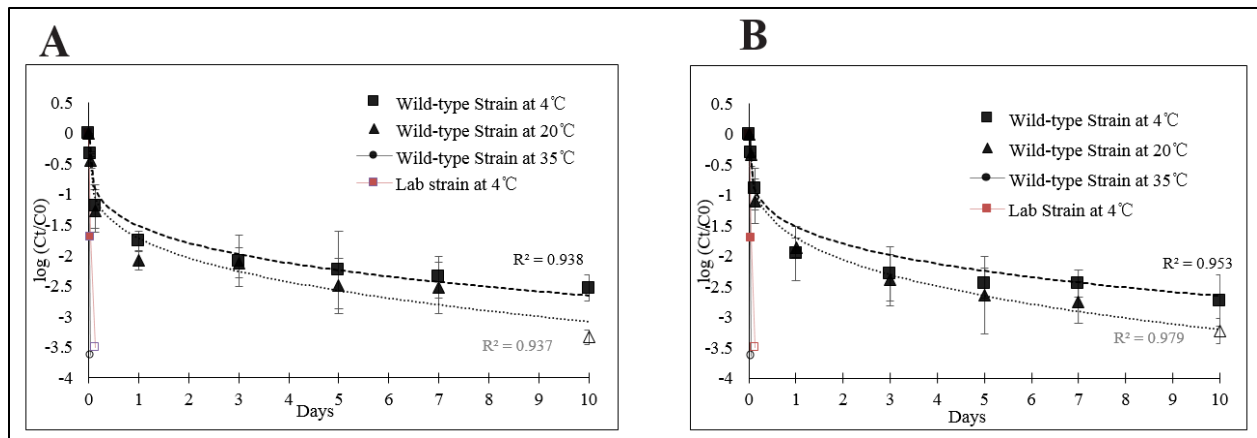


Figure 4.2. *E. coli* strains at 50% RH without sunlight exposure and at different temperatures on surfaces. (A) asphalt surfaces; (B) plastic surfaces.

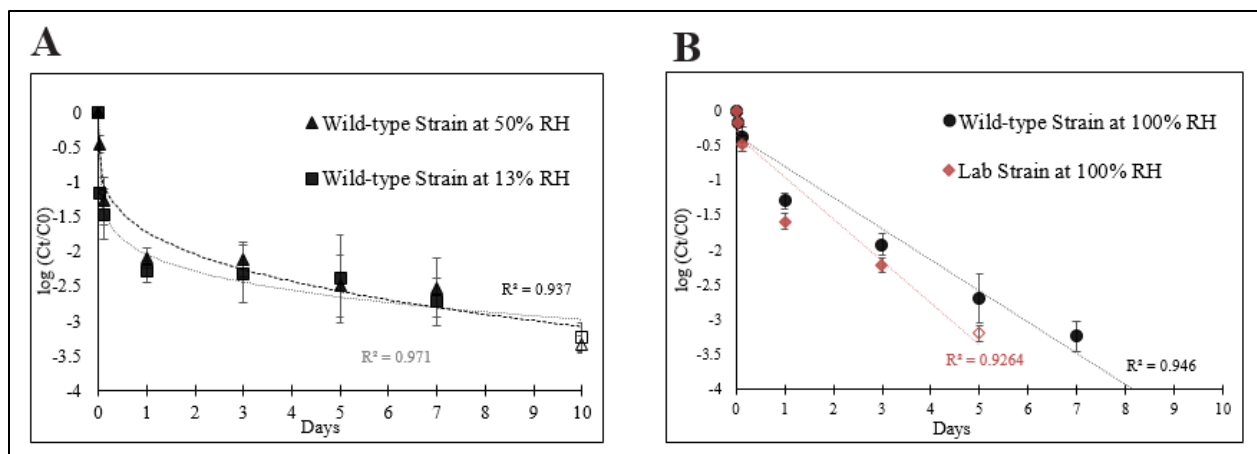


Figure 4.3. *E. coli* strains on asphalt at 20°C without sunlight exposure and under different relative humidities. (A) 13% and 50% relative humidities; (B) 100% relative humidity.

Table 4.3. The persistence conditions with the first-order decay model

Strain Type	Temperature	Media	Relative Humidity	K	R ²	T90	T99.9
	(°C)	Matrix	(%)	day ⁻¹		days	days
Lab strain	20	Runoff	--	0.95	0.95	2.43	7.28
Lab strain	20	Streamwater	--	1.30	0.94	1.76	5.29
Lab strain	20	Polystyrene	100	1.58	0.96	1.45	4.36
Lab strain	20	Asphalt	100	1.60	0.93	1.44	4.33
Lab strain	4	Runoff	--	0.17	0.95	13.46	40.38
Lab strain	4	Streamwater	--	0.31	0.99	7.34	22.01
Lab strain	4	Polystyrene	100	0.68	0.92	3.37	10.12
Lab strain	4	Asphalt	100	0.67	0.93	3.45	10.35
Wild-type strain	20	Runoff	--	0.79	0.93	2.91	8.72
Wild-type strain	20	Streamwater	--	1.15	0.99	2.00	6.01
Wild-type strain	20	Polystyrene	100	1.21	0.91	1.90	5.70
Wild-type strain	20	Asphalt	100	1.18	0.95	1.95	5.84
Wild-type strain	4	Runoff	--	0.18	0.92	13.04	39.11
Wild-type strain	4	Streamwater	--	0.26	0.97	9.02	27.05
Wild-type strain	4	Polystyrene	100	0.46	0.94	5.04	15.12
Wild-type strain	4	Asphalt	100	0.46	0.94	4.98	14.95

3.3 Persistence of *E. coli* in sufficient water

When the RH was increased to 100%, the survival time was increased from less than 1 hour to more than 5 days (Table 4.3). This indicates that water availability is an essential factor in prolonging the survival of the lab strain. Although wild-type strains showed resistance to low humidity, the persistence patterns in higher water content were different from the persistence patterns of low humidity resistance. Instead of the Weibull survival model, the first order decay model was more appropriate to fit the persistence data, which followed log-linear kinetics (Figure 4.4).

Besides the survival of the lab strain on solid surfaces with 100% RH, the stormwater runoff and stream water also showed prolonged persistence of the lab strain. Moreover, the lab strain showed the lowest decay rate in the runoff habitat under the same temperatures. Without a doubt, at the lower temperature (4 °C), the lab strain survived longer than at 20 °C under the same other conditions. On asphalt pavement with 100% RH, T_{90} and $T_{99.9}$ was 1.45 and 4.36 days at 20°C, respectively, while T_{90} and $T_{99.9}$ was 3.45 and 10.3 days at 4°C, respectively. On polystyrene plastic with 100% RH, T_{90} and $T_{99.9}$ was 1.45 and 4.36 days at 20°C, respectively, while T_{90} and $T_{99.9}$ was 3.37 and 10.1 days at 4°C, respectively. In the stormwater runoff, T_{90} and $T_{99.9}$ was 2.43 and 7.28 days at 20 °C, respectively, while T_{90} and $T_{99.9}$ was 13.5 and 40.4 days at 4 °C, respectively. In stream water, T_{90} and $T_{99.9}$ was 1.76 and 5.29 days at 20°C, respectively, while T_{90} and $T_{99.9}$ was 7.34 and 22 days at 4°C, respectively.

Although the surface materials did not affect the persistence of *E. coli* strains, the lab strain showed the lowest decay rate in the runoff habitat and the highest decay rate on the solid surfaces with 100% RH. The reason may be that the nutrients in runoff support *E. coli* survival, and the stream water and solid surfaces are more oligotrophic (Unc, Gardner, et al., 2006; Vital, Hammes, et al., 2008; Solecki, Jeanneau, et al., 2011).

Although the higher water content (i.e., high RH and water matrices) prolonged the survival of lab strain, wild-type strains persisted for an even longer period of time than the lab strain under all the same conditions. On asphalt pavement with 100% RH, the mean T_{90} and $T_{99.9}$ of wild-type strains was 1.95 and 5.84 days at 20°C, respectively, while T_{90} and $T_{99.9}$ was 4.98 and 14.95 days at 4°C respectively. On polystyrene plastic with 100% relative humidity, T_{90} and $T_{99.9}$ was 1.45 and 4.36

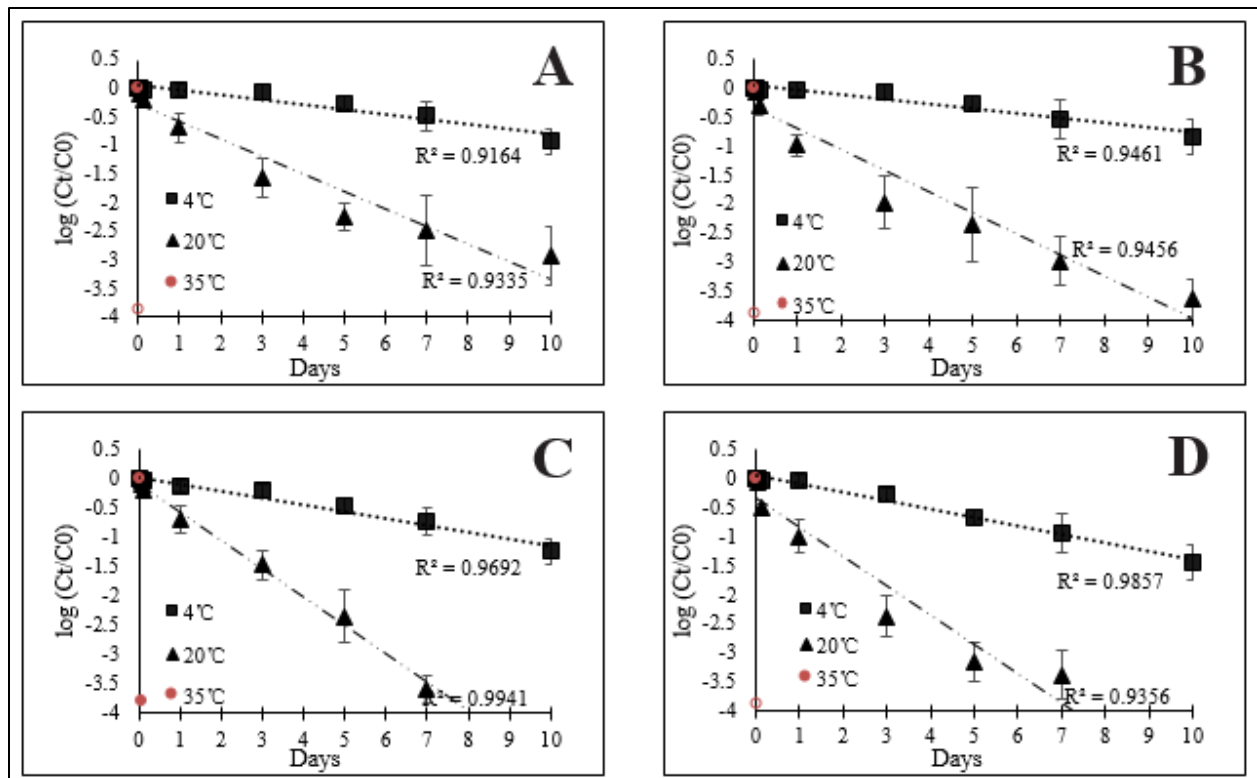


Figure 4.4. *E. coli* strains in liquid matrices. (A) wild-type strain in stormwater; (B) wild-type strain in river water; (C) lab strain in stormwater; (D) lab strain in river water.

days at 20°C, respectively, while T_{90} and $T_{99.9}$ was 3.37 and 10.1 days at 4°C, respectively. In the stormwater runoff, T_{90} and $T_{99.9}$ was 2.43 and 7.28 days at 20°C, respectively, while T_{90} and $T_{99.9}$ was 13.5 and 40.4 days at 4°C, respectively. In stream water, T_{90} and $T_{99.9}$ was 1.76 and 5.29 days at 20°C, respectively, while T_{90} and $T_{99.9}$ was 7.34 and 22 days at 4°C, respectively. In these conditions, lower temperatures also showed the prolonged survival of wild-type strains. Moreover, the same with lab strain, the wild-type strains showed the lowest decay rate in the runoff habitat.

Compared with the low RH conditions, the T_{90} s of wild-type strains in sufficient water content conditions were longer, which means most population was intolerant to desiccation. However, $T_{99.9}$ s in low RH conditions at 20°C were longer than the sufficient water content conditions. The reason may be that the dehydration-tolerant species are in a dormant state for survival but no metabolic activity (Schill, Mali, et al., 2009). The lower temperature also could decrease the metabolic activity of bacteria or even make it into a dormant state, so that prolong the persistence of the bacteria (Semenov, Van Bruggen, et al., 2007).

4. Conclusions

In this study, the results provide clues for the persistence of *E. coli* originating from stormwater runoff in relevant environmental conditions. We examined the decay patterns of lab strain and wild-type strains *E. coli* in four environmental factors: temperature, sunlight exposure, habitats, and RH. Findings indicate that the wild-type strains isolated from runoff persisted for a longer period of time than the lab strain under all environmental conditions tested, expect that the counts of all the strains were below the detection limit within a few hours with high temperature or sunlight exposure.

Although a few populations of the wild-type strains showed the ability of desiccation resistance and might go into a dormant state for prolonged survival, the counts were limited. Therefore, if high concentrations of *E. coli* in pavement runoff were found, these bacteria would have a low probability of persistence on the pavement. The control and wild-type *E. coli* strains could survive a prolonged period in liquid habitats, which means the survival of *E. coli* in shaded water pit may cause unnecessary burdens of high *E. coli* concentrations detected. Moreover, the persistence of *E. coli* in receiving streams may mislead the recent fecal contamination, especially in winter due to the prolonged persistence at a lower temperature. Findings from this study could provide insight

into the interpretation of environmental monitoring data using *E. coli* as the fecal indicator. However, the specific strains and whole genome sequences for wild-type strains of *E. coli* were unknown, and the genome sequences that cause the functional differences between lab strain and wild-type strains should be determined based on metagenome methods in the future.

References

- Athayde, D. N., P. E. Shelley, E. D. Driscoll, D. Gaboury and G. Boyd (1983). ""Results of the nationwide urban runoff program"." Results of the Nationwide Urban Runoff Program.
- Bergholz, T. M. and T. S. Whittam (2007). "Variation in acid resistance among enterohaemorrhagic *Escherichia coli* in a simulated gastric environment." Journal of Applied Microbiology **102**(2): 352-362.
- Berhanu Desta, M., M. Bruen, N. Higgins and P. Johnston (2007). "Highway runoff quality in Ireland." Journal of Environmental Monitoring **9**(4): 366-371.
- Bigelow, W. and J. Esty (1920). "The thermal death point in relation to time of typical thermophilic organisms." The Journal of Infectious Diseases: 602-617.
- Biriukova, I. V., A. A. Krylov, E. M. Kiseleva, N. I. Minaeva and S. V. Mashko (2010). "[Construction of the new *Escherichia coli* K-12 wild-type strain with improved growth characteristics for application in metabolic engineering]." Genetika **46**(3): 349-355.
- Bokulich, N. A., D. A. Mills and M. A. Underwood (2013). "Surface Microbes in the Neonatal Intensive Care Unit: Changes with Routine Cleaning and over Time." Journal of Clinical Microbiology **51**(8): 2617.
- Brown, E. R., P. S. Kandhal, F. L. Roberts, Y. R. Kim, D.-Y. Lee and T. W. Kennedy (2009). Hot mix asphalt materials, mixture design, and construction, NAPA Research and Education Foundation.
- Buchanan, R. L. and S. G. Edelson (1999). "Effect of pH-dependent, stationary phase acid resistance on the thermal tolerance of *Escherichia coli* O157:H7." Food Microbiology **16**(5): 447-458.
- Byappanahalli, M. N., R. L. Whitman, D. A. Shively, M. J. Sadowsky and S. Ishii (2006). "Population structure, persistence, and seasonality of autochthonous *Escherichia coli* in temperate, coastal forest soil from a Great Lakes watershed." Environmental microbiology **8**(3): 504-513.
- Cabelli, V. J., A. P. Dufour, L. J. McCabe and M. A. Levin (1982). "Swimming-associated gastroenteritis and water quality." American Journal of Epidemiology **115**(4): 606-616.
- Driscoll, E. D., P. E. Shelley and E. W. Strecker (1990). "Pollutant Loadings and Impacts From Highway Stormwater Runoff Volume III: Analytical Investigation and Research Report." Pollutant Loadings and Impacts from Highway Stormwater Runoff. Vol. III: Analytical Investigation and Research Report.
- England, L. S., H. Lee and J. T. Trevors (1993). "Bacterial survival in soil: Effect of clays and protozoa." Soil Biology and Biochemistry **25**(5): 525-531.
- Franz, E., A. D. Van Diepeningen, O. J. De Vos and A. H. C. Van Bruggen (2005). "Effects of cattle feeding regimen and soil management type on the fate of *Escherichia coli* O157:H7 and *Salmonella enterica* serovar typhimurium in manure, manure-amended soil, and lettuce." Applied and Environmental Microbiology **71**(10): 6165-6174.

- Geeraerd, A., V. Valdramidis and J. Van Impe (2006). "Erratum to "GInaFiT, a freeware tool to assess non-log-linear microbial survivor curves"[Int. J. Food Microbiol. 102 (2005) 95–105]." International Journal of Food Microbiology **3**(110): 297.
- Gilpin, B. J., M. Devane, B. Robson, F. Nourozi, P. Scholes, S. Lin, D. R. Wood and L. W. Sinton (2013). "Sunlight inactivation of human polymerase chain reaction markers and cultured fecal indicators in river and saline waters." Water Environment Research **85**(8): 743-750.
- Grant, M. (1997). "A new membrane filtration medium for simultaneous detection and enumeration of *Escherichia coli* and total coliforms." Appl. Environ. Microbiol. **63**(9): 3526-3530.
- Hacker, J. and E. Carniel (2001). "Ecological fitness, genomic islands and bacterial pathogenicity." EMBO Reports **2**(5): 376-381.
- Haile, R. W., J. S. Witte, M. Gold, R. Cressey, C. McGee, R. C. Millikan, A. Glasser, N. Harawa, C. Ervin, P. Harmon, J. Harper, J. Dermand, J. Alamillo, K. Barrett, M. Nides and G. Y. Wang (1999). "The health effects of swimming in ocean water contaminated by storm drain runoff." Epidemiology **10**(4): 355-363.
- He, Y., Y. Li, J. K. Salazar, J. Yang, M. L. Tortorello and W. Zhang (2013). "Increased water activity reduces the thermal resistance of *Salmonella enterica* in peanut butter." Appl. Environ. Microbiol. **79**(15): 4763-4767.
- Himathongkham, S., S. Bahari, H. Riemann and D. Cliver (1999). "Survival of *Escherichia coli* O157:H7 and *Salmonella typhimurium* in cow manure and cow manure slurry." FEMS Microbiology Letters **178**(2): 251-257.
- Hoff, J. C. and E. W. Akin (1986). "Microbial resistance to disinfectants: Mechanisms and significance." Environmental Health Perspectives **Vol. 69**: 7-13.
- Ihssen, J. and T. Egli (2004). "Specific growth rate and not cell density controls the general stress response in *Escherichia coli*." Microbiology **150**(6): 1637-1648.
- Ihssen, J., E. Grasselli, C. Bassin, P. François, J. C. Piffaretti, W. Köster, J. Schrenzel and T. Egli (2007). "Comparative genomic hybridization and physiological characterization of environment isolates indicate that significant (eco-)physiological properties are highly conserved in the species *Escherichia coli*." Microbiology **153**(7): 2052-2066.
- Ihssen, J., E. Grasselli, C. Bassin, P. François, J.-C. Piffaretti, W. Köster, J. Schrenzel and T. Egli (2007). "Comparative genomic hybridization and physiological characterization of environmental isolates indicate that significant (eco-) physiological properties are highly conserved in the species *Escherichia coli*." Microbiology **153**(7): 2052-2066.
- Jiang, S. (2004). "Is Urban Runoff a Source of Human Pathogenic Viruses to Recreational Beach Waters? Technical Completion Report, Project No. W-943."
- Kiefer, L. A., D. R. Shelton, Y. Pachepsky, R. Blaustein and M. Santin-Duran (2012). "Persistence of *Escherichia coli* introduced into streambed sediments with goose, deer and bovine animal waste." Letters in Applied Microbiology **55**(5): 345-353.
- Large, T. M., S. T. Walk and T. S. Whittam (2005). "Variation in acid resistance among shiga toxin-producing clones of pathogenic *Escherichia coli*." Applied and Environmental Microbiology **71**(5): 2493-2500.

- Lurry, D. L. and C. M. Kolbe (2000). Interagency field manual for the collection of water-quality data, US Department of the Interior, US Geological Survey.
- Mafart, P., O. Couvert, S. Gaillard and I. Leguérinel (2002). "On calculating sterility in thermal preservation methods: application of the Weibull frequency distribution model." International journal of food microbiology **72**(1-2): 107-113.
- Marsalek, J. and Q. Rochfort (2004). "Urban wet-weather flows: Sources of fecal contamination impacting on recreational waters and threatening drinking-water sources." Journal of Toxicology and Environmental Health - Part A **67**(20-22): 1765-1777.
- McCarthy, D., V. G. Mitchell and A. Deletic (2006). "*Escherichia coli* levels in urban stormwater." Paper Presented to the Urban Drainage Modelling and International Water Sensitive Urban Design Conference, Melbourne, 4–6 April 2006.
- McConkey, B. J., C. L. Duxbury, D. G. Dixon and B. M. Greenberg (1997). "Toxicity of a PAH photooxidation product to the bacteria *Photobacterium phosphoreum* and the duckweed *Lemna gibba*: Effects of phenanthrene and its primary photoproduct, phenanthrenequinone." Environmental Toxicology and Chemistry: An International Journal **16**(5): 892-899.
- Oliver, D. M., C. D. Clegg, A. L. Heathwaite and P. M. Haygarth (2007). "Preferential attachment of *Escherichia coli* to different particle size fractions of an agricultural grassland soil." Water, Air, and Soil Pollution **185**(1-4): 369-375.
- Perkins, T. L., K. Perrow, P. Rajko-Nenow, C. F. Jago, D. L. Jones, S. K. Malham and J. E. McDonald (2016). "Decay rates of faecal indicator bacteria from sewage and ovine faeces in brackish and freshwater microcosms with contrasting suspended particulate matter concentrations." Science of the Total Environment **572**: 1645-1652.
- Rozen, Y. and S. Belkin (2001). "Survival of enteric bacteria in seawater." FEMS Microbiology Reviews **25**(5): 513-529.
- Schill, R. O., B. Mali, T. Dandekar, M. Schnölzer, D. Reuter and M. Frohme (2009). "Molecular mechanisms of tolerance in tardigrades: New perspectives for preservation and stabilization of biological material." Biotechnology Advances **27**(4): 348-352.
- Semenov, A. V., A. H. C. Van Bruggen, L. Van Overbeek, A. J. Termorshuizen and A. M. Semenov (2007). "Influence of temperature fluctuations on *Escherichia coli* O157:H7 and *Salmonella enterica* serovar Typhimurium in cow manure." FEMS Microbiology Ecology **60**(3): 419-428.
- Sidhu, J. P. S., S. Toze, L. Hodgers, K. Barry, D. Page, Y. Li and P. Dillon (2015). "Pathogen decay during managed aquifer recharge at four sites with different geochemical characteristics and recharge water sources." Journal of Environmental Quality **44**(5): 1402-1412.
- Sinton, L. W., C. H. Hall, P. A. Lynch and R. J. Davies-Colley (2002). "Sunlight inactivation of fecal indicator bacteria and bacteriophages from waste stabilization pond effluent in fresh and saline waters." Applied and Environmental Microbiology **68**(3): 1122-1131.
- Sjogren, R. E. (1994). "Prolonged survival of an environmental *Escherichia coli* in laboratory soil microcosms." Water, Air, & Soil Pollution **75**(3-4): 389-403.

- Solecki, O., L. Jeanneau, E. Jardé, M. Gourmelon, C. Marin and A. M. Pourcher (2011). "Persistence of microbial and chemical pig manure markers as compared to faecal indicator bacteria survival in freshwater and seawater microcosms." Water Research **45**(15): 4623-4633.
- Surbeck, C. Q., S. C. Jiang, J. H. Ahn and S. B. Grant (2006). "Flow fingerprinting fecal pollution and suspended solids in stormwater runoff from an urban coastal watershed." Environmental Science and Technology **40**(14): 4435-4441.
- Tabe, E. S., J. Oloya, D. K. Doetkott, M. L. Bauer, P. S. Gibbs and M. L. Khaita (2008). "Comparative effect of direct-fed microbials on fecal shedding of *Escherichia coli* O157:H7 and Salmonella in naturally infected feedlot cattle." Journal of Food Protection **71**(3): 539-544.
- Tao, H., C. Bausch, C. Richmond, F. R. Blattner and T. Conway (1999). "Functional genomics: Expression analysis of *Escherichia coli* growing on minimal and rich media." Journal of Bacteriology **181**(20): 6425-6440.
- Unc, A., J. Gardner and S. Springthorpe (2006). "Recovery of *Escherichia coli* from soil after addition of sterile organic wastes." Applied and Environmental Microbiology **72**(3): 2287-2289.
- Vital, M., F. Hammes and T. Egli (2008). "*Escherichia coli* O157 can grow in natural freshwater at low carbon concentrations." Environmental Microbiology **10**(9): 2387-2396.
- Wyckoff, K. N., S. Chen, A. J. Steinman and Q. He (2017). "Impact of Roadway Stormwater Runoff on Microbial Contamination in the Receiving Stream." Journal of environmental quality **46**(5): 1065-1071.
- Yufen, R., W. Xlaoke, O. Zhiyun, Z. Hua, D. Xiaonan and M. Hong (2008). "Stormwater runoff quality from different surfaces in an urban catchment in Beijing, China." Water Environment Research **80**(8): 719-724.
- Zhang, Q. and T. Yan (2012). "Desiccation Resistance of Soil *Escherichia coli* Populations Correlates to Intracellular Trehalose Concentration." Applied and Environmental Microbiology: AEM. 01904-01912.

Chapter 5

Natural Stabilities and Dynamics in Bacterioplankton Communities in the Unimpaired Streams within an Ecoregion

Abstract

Bacterioplankton in freshwater streams plays a critical role in the assessment of aquatic ecological health. Recent studies have addressed bacterioplankton community dynamics in streams with major anthropogenic impacts. However, the natural dynamics of low impact streams are understudied. The understanding of this natural variation of the bacterioplankton communities in low impact streams is needed before assessing the stream ecological conditions utilizing microbial community indicators. In this study, three unimpaired streams spanning first, second, and fourth order within an ecoregion were utilized to characterize the temporal dynamics through bacterioplankton community analysis. The results showed that the bacterioplankton community compositions were stable within a short sampling period length of one month at each site. Although in the three sites the species abundance distributions were followed by Poisson lognormal model and the dominated taxa were typical freshwater taxa that were well documented, different streams harbored distinct communities, especially for the fourth order stream compared with the first and second order streams. In the fourth order stream, the microbial community was enriched with known freshwater associated bacteria, while the lower order streams hosted more sediment and soil associated taxa. Moreover, these streams were monitored over two seasons, where pronounced seasonal shifts in community compositions for all the three streams were observed. The streams closed to headwater revealed elevated alpha diversities in spring, but the highest order stream showed a reversed seasonal pattern. In the Spring, more diverse sources of bacteria enter the lower order streams, but enrichment processes for select taxa were enhanced in the higher-order stream. Overall, this study shows that the delimitation of stream types and sampling periods is an important step for defining the reference conditions of bacterioplankton communities within an ecoregion.

1. Introduction

Within ecological regions, or ecoregions, the contiguous landscape units have regional similarities in soil, land surface form, land use, and dominant vegetation. The use of the unimpaired stream sites as references has long been a key component to assess the aquatic ecological conditions at

other sites within ecoregions for the last 3 decades (Hughes, Larsen, et al., 1986; Attrill and Depledge 1997; Dale and Beyeler 2001; Martins, Ligeiro, et al., 2018). Metrics and indices were developed from the physical, chemical, and biological characteristics and were used as a basis against other sites (Aazami, Esmaili Sari, et al., 2015). bacterioplankton-based index of biotic integrity provided a useful measurement since high-throughput sequencing techniques were rapidly developed recently (Simonin, Voss, et al., 2019; Laperriere, Hilderbrand, et al., 2020). One of the techniques is 16S rRNA amplicon sequencing which enables researchers to identify the entire microbial community within a sample (Sinclair, Osman, et al., 2015). However, the natural variations of these bacterioplankton communities in the unimpaired streams must be understood before we can fully assess the potential applicability of microbial community index as an assessment of stream ecological conditions (Dale and Beyeler 2001). A challenge for stream ecologists is to determine the extent to which Pelagic stream communities are structured by the natural variability.

Recent studies have addressed bacterioplankton community diversity in streams within a watershed (Staley, Gould, et al., 2014; Hassell, Tinker, et al., 2018; Teachey, McDonald, et al., 2019). While these studies indicate that taxonomic shifts are closely related to seasonality and hydrological conditions, less is understood about the natural range and variation of microbial groups. Moreover, few investigators have examined how stable the bacterioplankton communities are over time within an ecological region.

In this study, temporal dynamics of bacterioplankton communities were quantified at 3 unimpaired streams within an ecoregion. This work spanned independent streams of 1st, 2nd, 4th order for covering different order types. Measurements of bacterioplankton communities were made in spring and fall to evaluate seasonal shifts in community composition at each stream site. We hypothesized that bacterioplankton communities within each stream would exhibit temporal community composition patterns. Community structure fluctuated around a relatively stable state with communities sampled closely and pronounced seasonal shifts in composition with those sampled many months apart. Furthermore, in the three streams, the species abundance distributions (SADs) followed the same model and the dominated taxa were typical freshwater bacteria that were well documented, while the microbial communities in the two streams closed to headwater showed high diversity levels and hosted elevated sediment and soil-associated bacteria.

2. Materials and Methods

2.1 Sites selection

The study was conducted in East Tennessee, USA. Three stream sites were selected within an ecoregion – 67f The Southern Limestone/Dolomite Valleys and Low Rolling Hills (Griffith, Omernik, et al., 1997), including Clear Creek, White Creek, and Powell Creek, which were named as stream I, stream II, and stream III respectively (Figure 5.1). Water samples were collected by dipping a sterile 1 L bottle below the stream surface following Standard Method (Lurry and Kolbe 2000). Five samples at each stream site were collected in September 2018, and three samples at each site were collected in May 2019. After collection, all samples were transported using a cooler with ice and delivered to the laboratory within 6 hours. For high-throughput sequencing, the water samples were filtered through 0.45 µm filters (MilliporeSigma™ HAWG047S6), and the filters were stored at -80°C until DNA extraction.

2.2 Fecal indicator bacteria test

Fecal indicator bacteria (*E. coli* and total coliforms) tests were performed within 8 hours of sample collection. *E. coli* and total coliforms were determined by m-Colibblue 24 ® media following the protocol provided by the manufacturer (Hach Co., Loveland, CO). Briefly, an appropriate volume of each stream water sample were filtered, and the then filter was attached on a pad which was soaked with 2 ml of the broth. The filters were incubated at 37 °C for 24 ± 2 hours.

2.3 DNA extraction and high-throughput 16S rRNA amplicon sequencing

DNA from the colony isolates was extracted by the FastDNA Spin Kit for Soil (MP Biomedicals) following the manufacturer's protocol. High-throughput 16S rRNA Amplicon Sequencing was performed as previously described (Wyckoff, Chen, et al., 2017). The PCR amplification used primers 515-F and 806-R to target the V4 region of the 16S rRNA gene (Caporaso, Lauber, et al., 2012). PCR products were purified using the Wizard SV Gel and PCR Clean-Up System (Promega), and paired-end sequencing was completed using the Illumina MiSeq platform.

2.4 Bioinformatic and statistical analyses.

Analysis of sequence reads was conducted using QIIME 2 software (Bolyen, Rideout, et al., 2019).

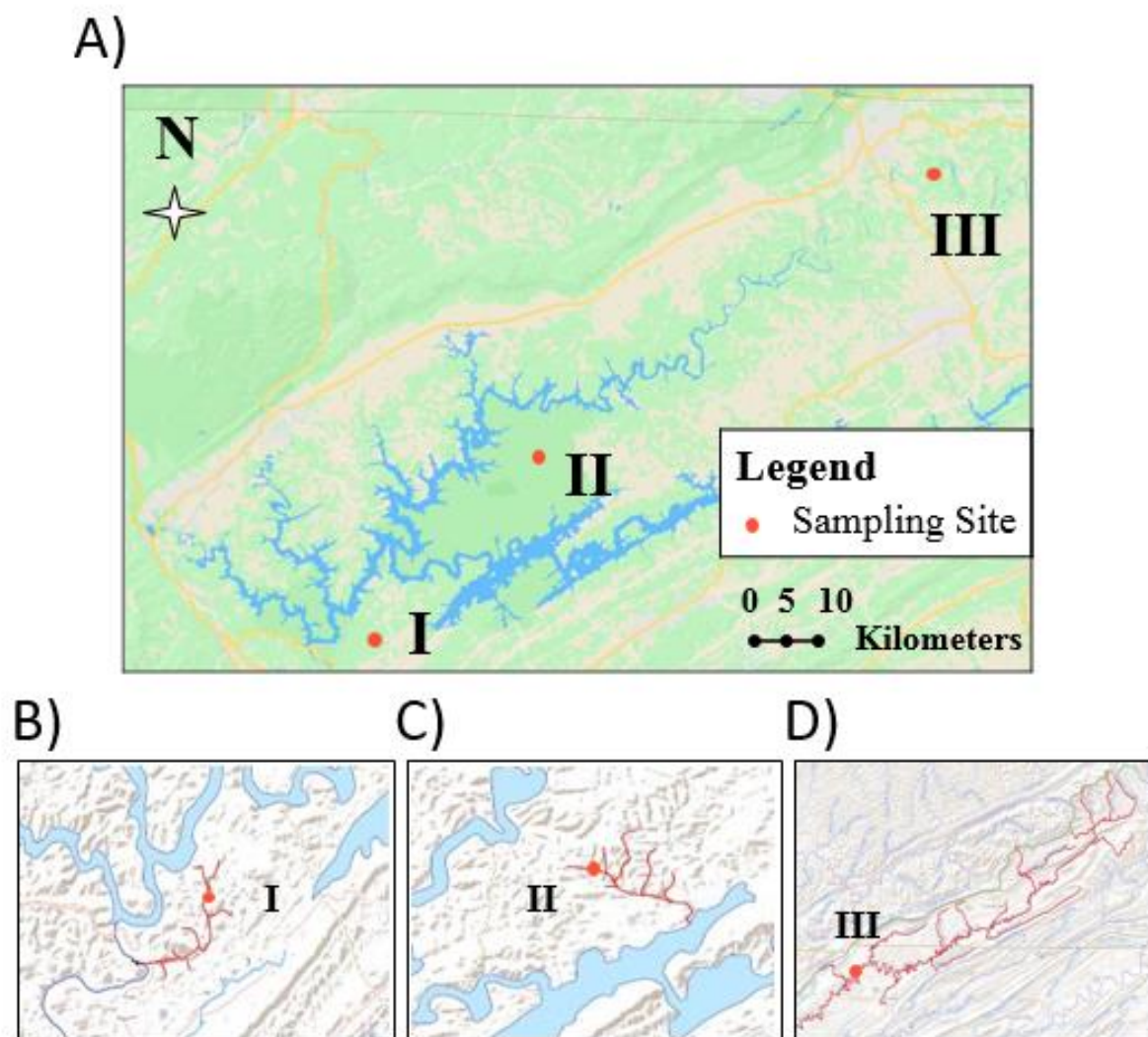


Figure 5.1. The maps of the three streams from 67f eco-region East TN. (A) sampling sites of three streams; (B) The sampling site of stream I; (C) The sampling site of stream II; (D) The sampling site of stream III. A total of 24 water samples were collected in September 2018 and May 2019.

Amplicon sequence variants (ASVs) enumeration was performed using the DADA2 pipeline (Callahan, McMurdie, et al., 2016). The obtained ASVs were then matched to the Greengenes 13.8 database.

Statistical analysis and data visualization were carried out using R version 3.50 (Team 2017). PERMANOVA was applied to assess the difference of community composition, α -diversity index (the Shannon index), and Bray–Curtis (abundance-based) dissimilarity was calculated to represent the taxonomic β -diversity using the vegan R package (Oksanen, Blanchet, et al., 2013).

The “sads” package in R was used to fit the species abundance distribution (SAD) models of samples based on the ASV level composition. Three SAD models, including log-series, Poisson lognormal, and Zipf, were used to fit the raw data in this study. The R-squared was applied to determine the goodness of model fit based on the previous study (Shoemaker, Locey, et al., 2017).

The sequence count table of 16S rRNA gene amplicon studies from the Earth Microbiome Project (EMP) was downloaded (Thompson, Sanders, et al., 2017). EMP samples from freshwater, freshwater sediment and soil were selected and merged with our stream water samples. The proportions of each stream water microbiota attributable to freshwater, sediment, and soil at the genus level were estimated using Sourcetracker (Knights, Kuczynski, et al., 2011), a computational tool that performs Bayesian inference which was using stream water microbiota as the sink and those in the soil as sources.

3. Results and Discussions

3.1 Fecal indicator bacteria in the unimpaired streams

Based on State of Tennessee water quality standards, a geometric mean concentration of *E. coli* shall not exceed 126 CFU/100mL with at least 5 sampling time points (TDEC 2007). The three streams showed low levels of *E. coli*, and the average concentrations of *E. coli* across the three sampling sites were from 52 to 80 CFU/100mL. Meanwhile, the average concentrations of total coliforms were from 728 CFU/100mL to 1428 CFU/100mL (Figure 5.2).

The levels of FIB, including *E. coli* and total coliforms, in the three unimpaired streams, were similar across sampling locations, and the concentrations were not significantly different between seasons at each site ($P > 0.05$). The natural variations of FIB in unimpaired streams were limited

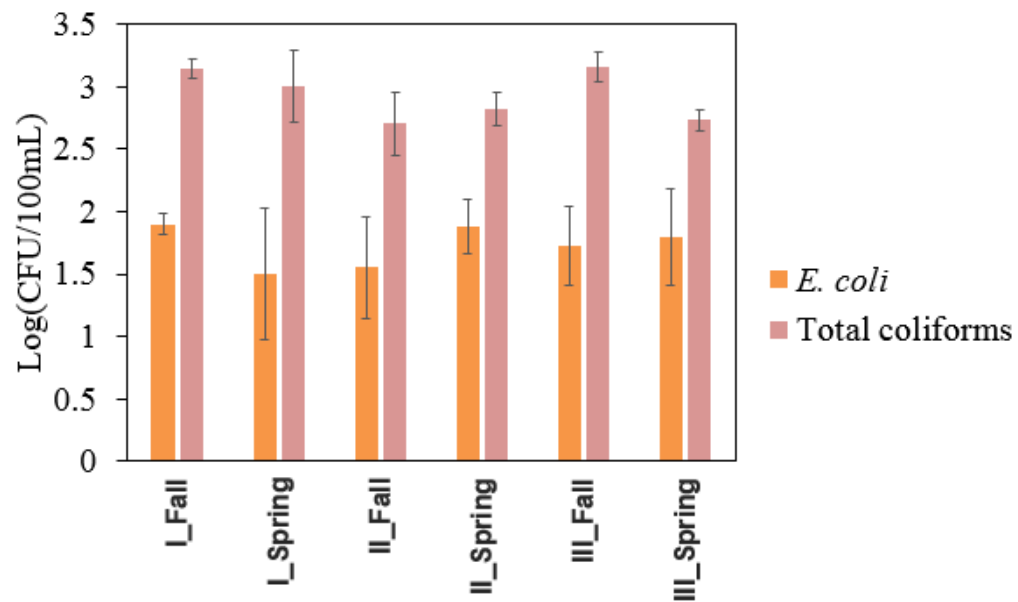


Figure 5.2. *E. coli* and total coliforms in ecoregion streams. FIB showed relatively stable concentrations across sampling locations and times (there is no significant difference across sampling locations and times).

in this study.

3.2 Species abundance distributions for the bacterioplankton communities

The species abundance distribution (SAD) is one of the most important patterns in ecology and biodiversity and widely used in macroecological studies, but it has been neglected in microbial ecology until recently (Shoemaker, Locey, et al., 2017). SAD curves for each stream in two different seasons were plotted (Figure 5.3). We tested common SAD models, including log-series, Poisson lognormal, and Zipf. In all three streams and two seasons, Poisson lognormal clearly provides the best fit. It explained more than 90% of the variation of the bacterioplankton SADs (Table 5.S1). Previous studies also reported Poisson lognormal model gave the best fit for other freshwater ecosystems (Tang, Xie, et al., 2020). Although numerous models were used to predict SADs, lognormal has been more commonly observed with microorganisms in natural habitats (Shoemaker, Locey, et al., 2017). The log-series distribution was well for fitting macroecological SADs; the Zipf model was successfully used to explain extremely uneven abundances (Shoemaker, Locey, et al., 2017).

Although Poisson lognormal model was the best model for all samples in this study, the data set exhibited different patterns among streams within a season, especially for stream III, which showed lower ASVs numbers and higher individuals of common ASVs. In spring, stream III only contained 258 ASVs. Meanwhile, the highest individual number of common ASVs was 777; however, stream II contained 1341 ASVs, and the highest individual number of common ASVs was 123. Moreover, at each site, the two different seasons exhibited different ASVs numbers and different numbers of individuals of common ASVs. Although the Zipf model did not fit the raw data well, ASVs in Stream III fitted the Zipf model (30.1% in Fall and 37.9% in Spring) better than stream I (23.8% in Fall and 20.5% in Spring) and II (15.5% in Fall and 13.1% in Spring), which may be explained by enhanced enrichment processes for freshwater-associated taxa in stream III as high order downstream (Teachey, McDonald, et al., 2019). The patterns revealed the different bacterioplankton diversities in different streams and seasons.

3.3 Short period fluctuations in bacterioplankton community structure within a season

In September 2018, five time points at each stream site were examined for bacterioplankton community diversity and composition. The results showed the relatively stable alpha and beta

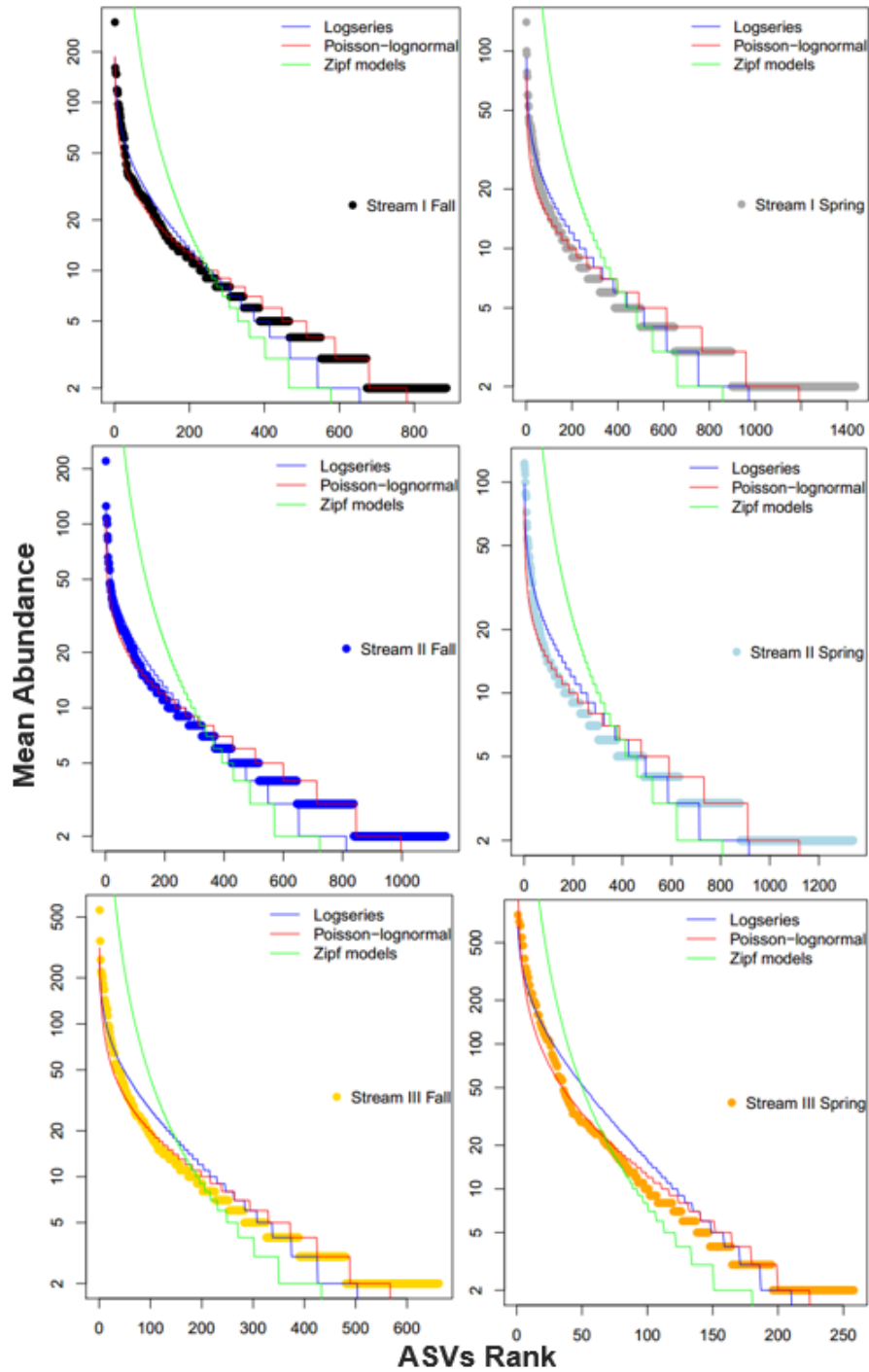


Figure 5.3. Predicting SADs of Bacterioplankton communities at each site. Each model was fit to the observed SAD. Table 5.S1 shows the variations of the SADs explained, which indicated the Poisson lognormal model was the best performance for all the three streams in both Spring and Fall.

bacterioplankton diversities within a one-month sampling period at each stream site (Figure 5.4). Principal-coordinate-analysis (PCoA) showed nearly overlapped sample dots for the same stream (Figure 5.4A), and the mean bray-Curtis dissimilarities were less than 0.5 within the same stream (Figure 5.4B). The Shannon indices were also showed stabilities within the same stream. The Shannon indices were from 5.23 to 6.25, from 5.97 to 6.61, from 4.37 to 5.30 for stream I, II, and III, respectively (Figure 5.4C).

Although the levels of FIB did not show significance across sampling locations, distinguished alpha and beta bacterioplankton diversities were observed (Figure 5.4). PCoA showed a clear separation of stream III from stream I and II, while samples from the two streams clustered together (Figure 5.4A). Bray-Curtis dissimilarity analyses revealed the dissimilarities among streams were higher than the dissimilarities within each stream (Figure 5.4B). Moreover, the dissimilarities between stream III and the other two streams are considerably higher than the dissimilarity between stream I and II. Alpha diversity also showed significant differences from stream III to the other two streams. Stream III was the least diverse versus stream I and II (Shannon index, $P < 0.05$). The differences in alpha diversities of microbial communities between stream I and II were smaller and not statistically significant.

Actually, several studies found headwater hosted higher microbial diversity than higher-order downstream in a stream network (Teachey, McDonald, et al., 2019). The three independent streams also showed similar microbial diverse patterns, stream I and II as small streams are closed to the headwater, and stream III is the highest-order stream (Figure 5.1B-D).

3.4 Short period fluctuations in the abundances of individual ASVs within a season

The top 10 ASVs in Fall at each stream site were showed in figure 5.5. These taxonomy names of dominated ASVs in the three sites were typical freshwater bacteria that were well documented (Newton, Jones, et al., 2011), especially for the taxa named as *Flavobacterium* and *Comamonadaceae*, which were found across all the three streams. Moreover, most top 10 ASVs at each site were at similar levels of relative abundance across 5 sampling time points within the month. These analyses show temporal stabilities in the bacterioplankton community over short time periods.

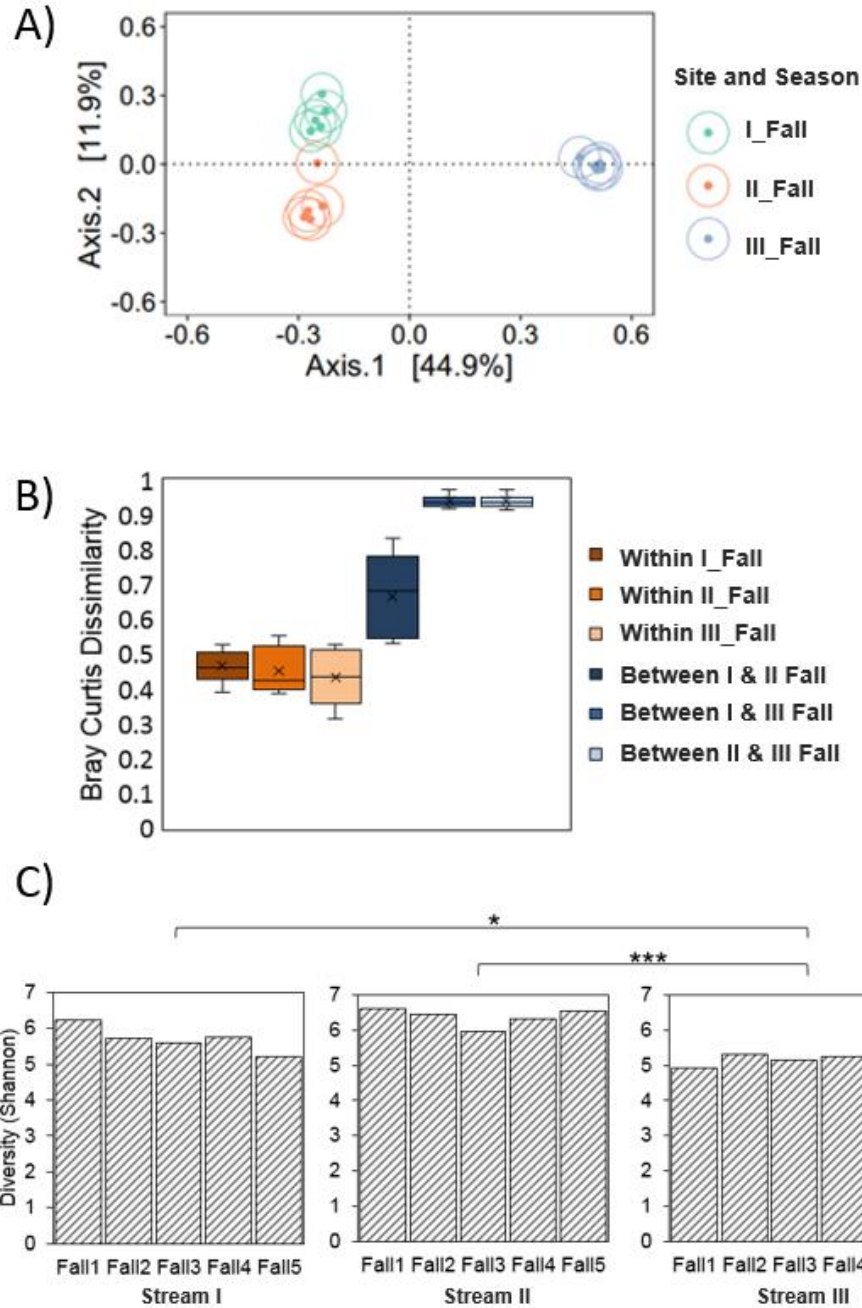


Figure 5.4. Microbial diversity and community structure within a month. (A) PCoA plot; (B) Bray–Curtis dissimilarity within each stream and between streams in Fall; (C) Alpha diversity (Shannon index) for each of the three selected sites in Fall. Fall1, Fall2, Fall3, Fall4, and Fall5 presented the sampling time on September 6th, 11th, 19th, 21st, and 30th, respectively. *, P < 0.05; **, P < 0.01; ***, P < 0.001.

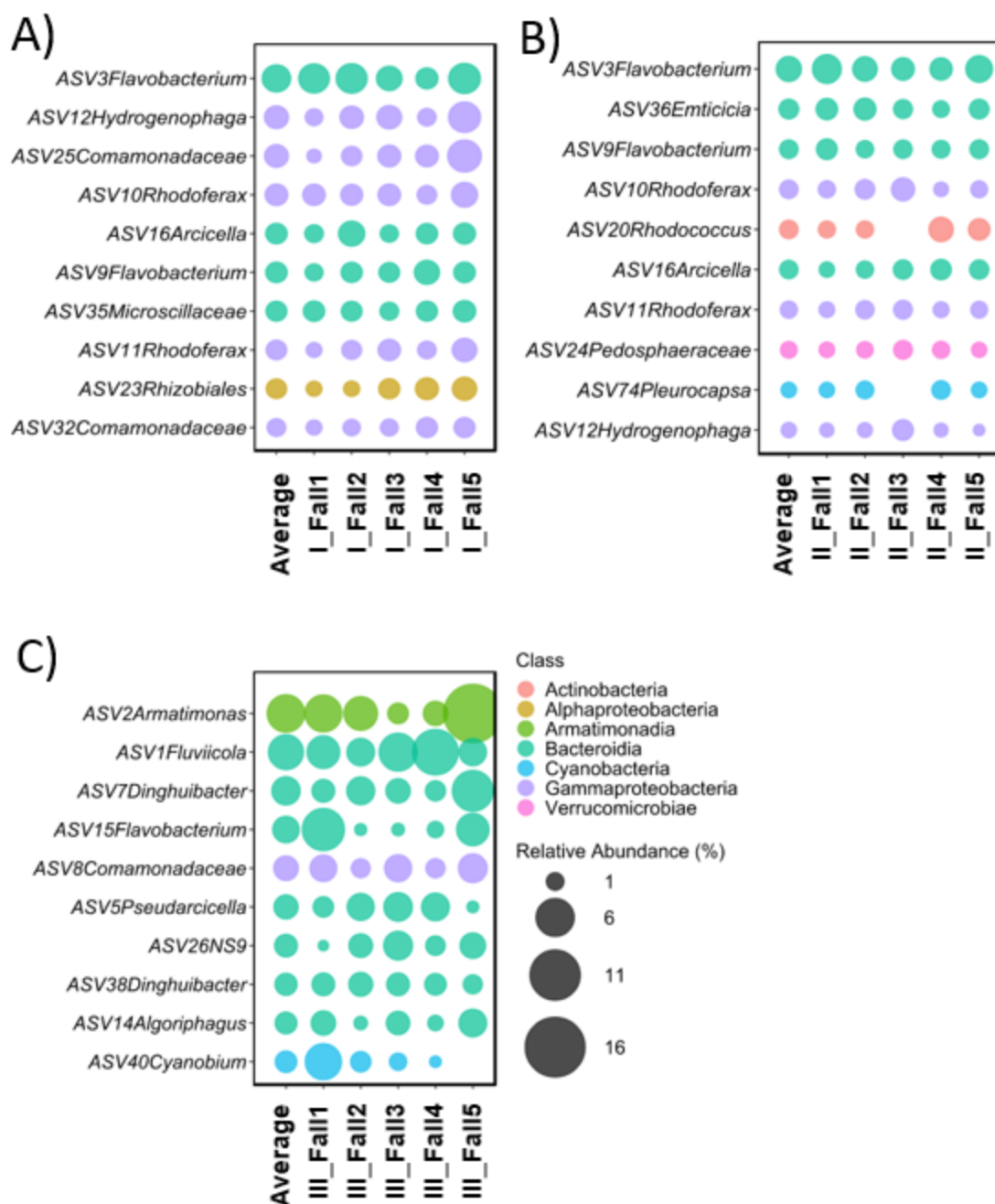


Figure 5.5. Top 10 ASVs in Fall at each sampling site; (A) The sampling site of stream I; (B) The sampling site of stream II; (C) The sampling site of stream III.

Decreases in alpha diversities in stream III were accompanied by increases in the relative abundances of top 10 ASVs. All 5 samples from stream III hosted the ASVs that were more than 5% relative abundance, but none of the ASVs was more than 5% in samples from the stream I and II.

Some top 10 ASVs, such as ASV3 *Flavobacterium*, ASV9 *Flavobacterium*, ASV10 *Rhodoferrax*, ASV11 *Rhodoferrax*, ASV12 *Hydrogenophaga*, and ASV16 *Arcicella*, were found both in stream I and stream II, but none of the top 10 ASVs in stream I and II was found in stream III. This suggests that the distinct bacterioplankton community in stream III compared to stream I and II, although the taxonomy names of the ASVs were commonly found in freshwater ecosystems.

3.5 Seasonal shifts in bacterioplankton community compositions at each site.

Although all streams in this study showed stable community compositions within a month, Bray-Curtis dissimilarities between seasons were higher than the dissimilarities within seasons for all the three streams (Figure 5.6A). Further, PCoA revealed a clear separation of spring from fall at each site (Figure 5.6B-D). Meanwhile, though not all top 10 ASVs in

spring showed seasonal dynamics, and all the three streams contained some of the top 10 ASVs that showed pronounced seasonal different abundance patterns (Figure 5.7A-C).

There are limited researches about specific taxa of seasonal preferences, but in stream III the first and second dominant ASVs are *Sporichthyaceae*, which increased in spring and Summer in other freshwater ecosystems (Fisher, Newton, et al., 2015; Li, Zhang, et al., 2019; Ozbayram, Koker, et al., 2020).

There were just 3 ASVs and 5 ASVs that were higher than 1% relative abundances in spring at stream I and II, respectively, but all top 10 ASVs were higher than 1% in fall at both the two stream sites. Meanwhile, elevated alpha diversities (Shannon indices) in spring were accompanied by the decreased relative abundances of top 10 ASVs at the stream I and II sites (Figure 5.8A). However, stream III showed a reversed seasonal pattern. No doubt, Bray-Curtis dissimilarities between stream I/II and III were sharper in spring than in fall (Figure 5.8B).

3.6 Habitat associations for bacterioplankton

The different bacterioplankton structures among sites and between seasons may be explained by

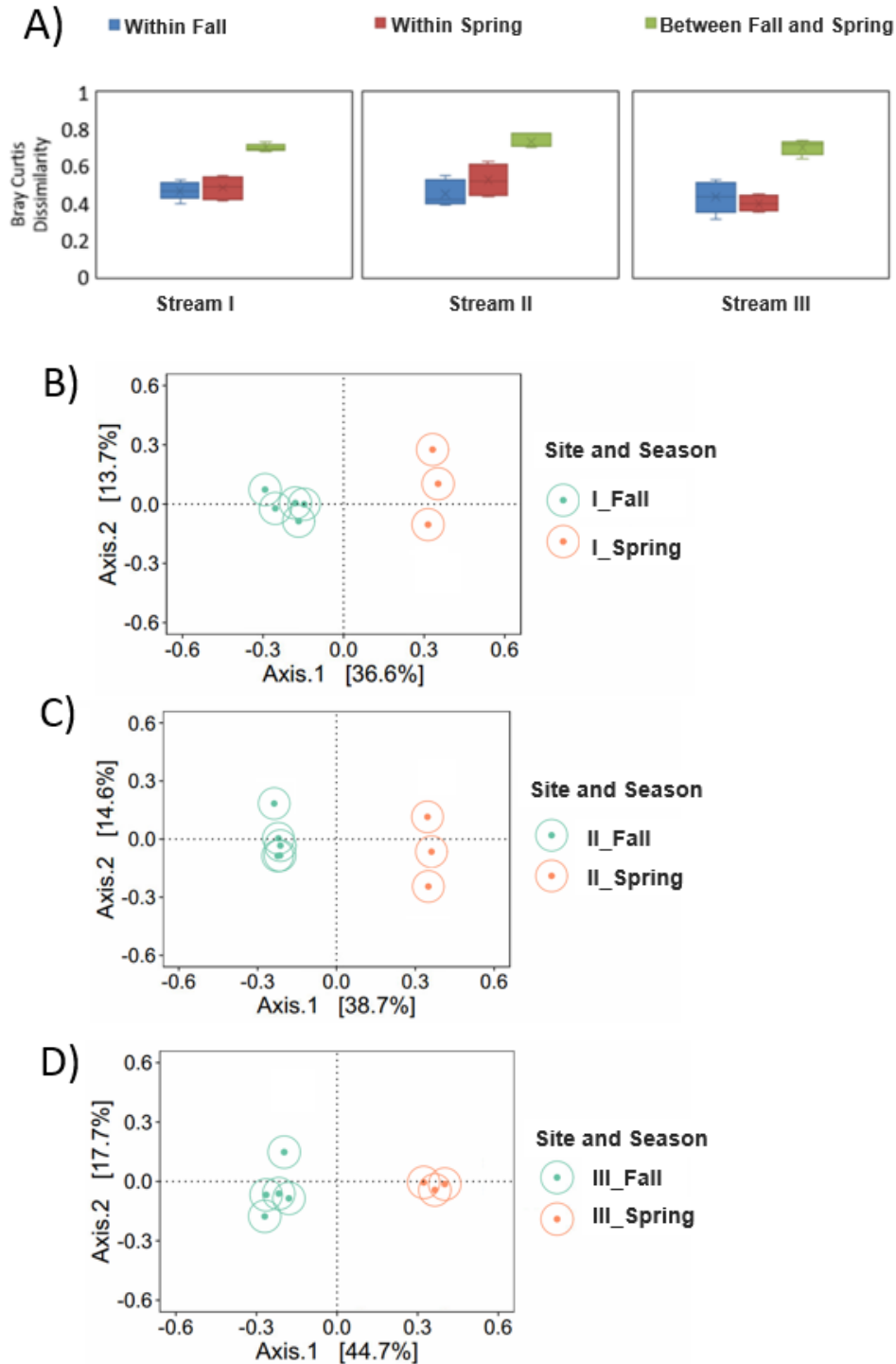


Figure 5.6. Beta-diversity showed samples from Fall and Spring at each site. (A) Bray–Curtis dissimilarity between seasons and within seasons at each site. (B) PCoA plot for samples at the sampling site of stream I; (C) PCoA plot for samples the sampling site of stream II; (D) PCoA plot for samples the sampling site of stream III.

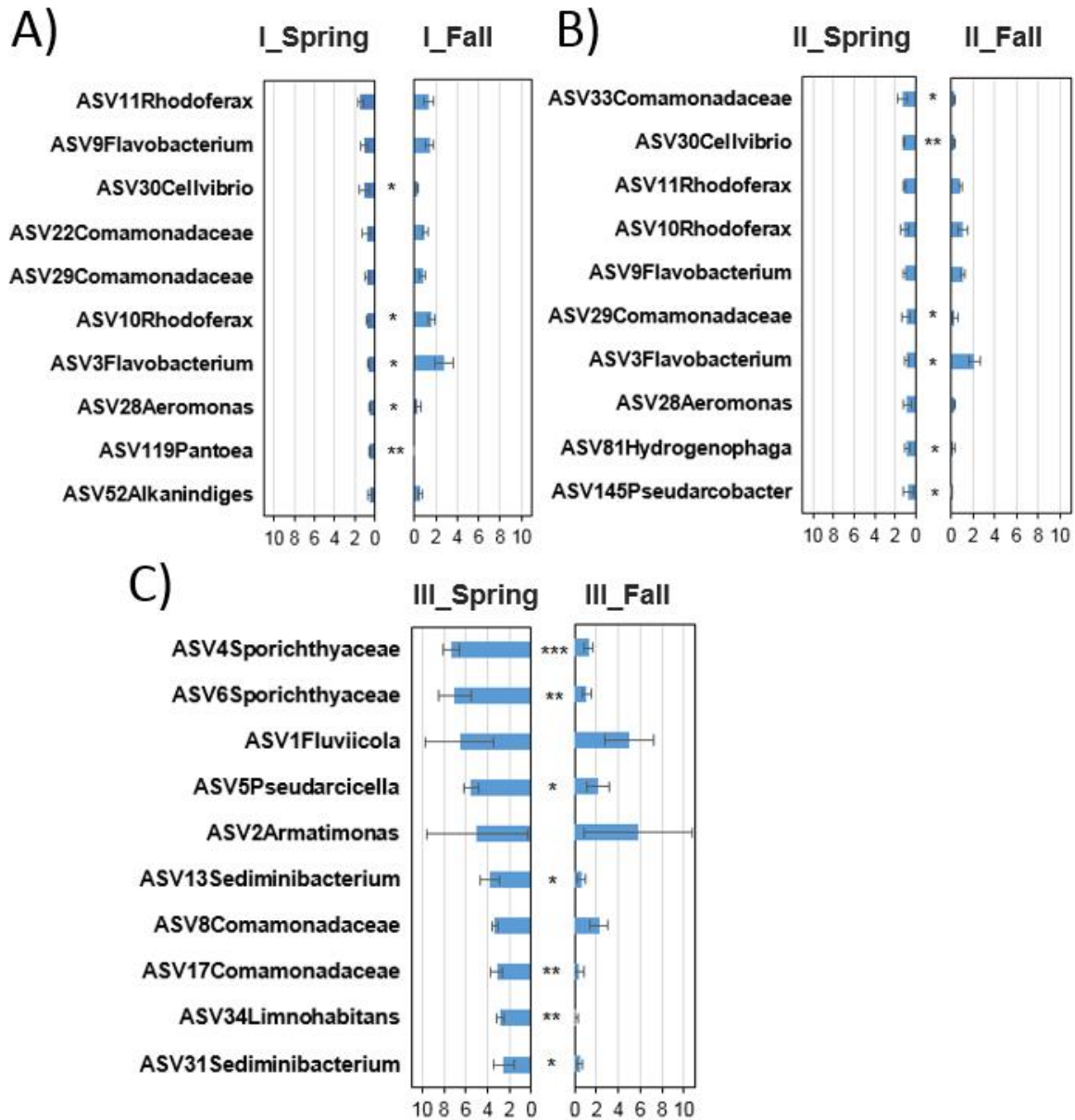


Figure 5.7. The relative abundance (%) of Top 10 ASVs in Spring at each site compared to Fall season; (A) The sampling site of stream I; (B) The sampling site of stream II; (C) The sampling site of stream III. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

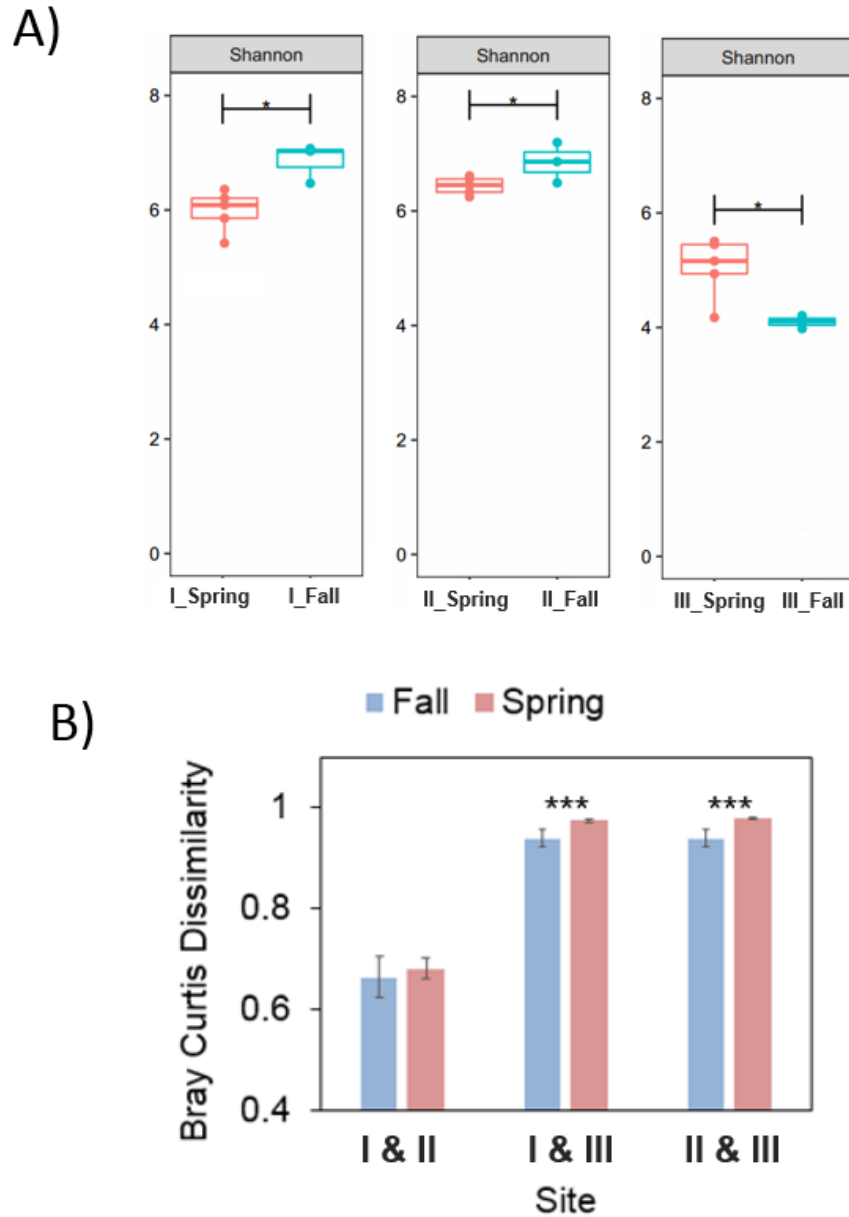


Figure 5.8. Microbial diversity and community structure between fall and spring at each site. (A) Alpha diversity (Shannon index) at each site; (B) Bray-Curtis dissimilarities among sites comparisons for Fall and Spring. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

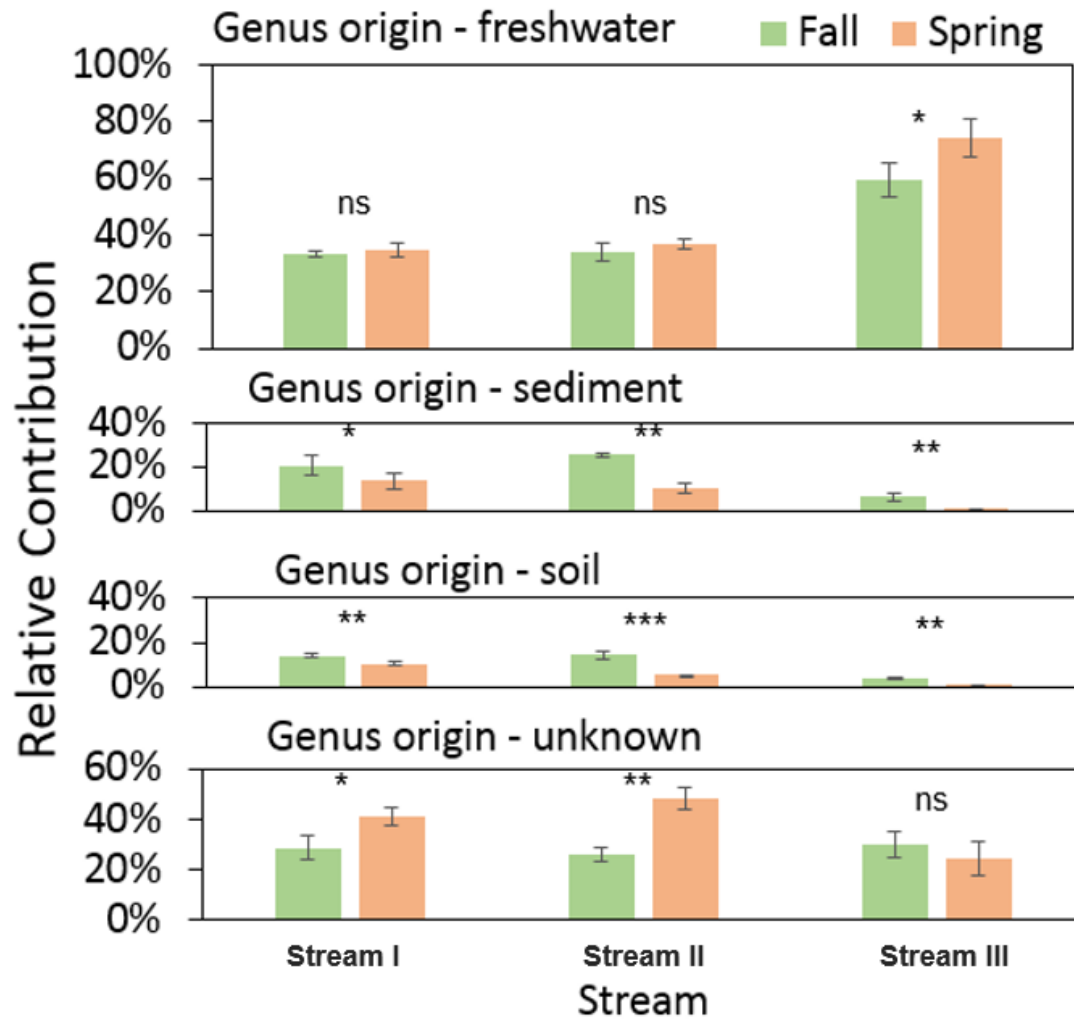


Figure 5.9. The contributions of potential sources of genera (%) in spring and fall at each stream.
 *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

the interactions of the potential microbial sources. In order to explore habitat associations for bacterioplankton in the three streams, a Bayesian method was used to determine putative source environments of bacterioplankton communities at the genus level (Figure 5.9). No doubt, the most dominant known source was freshwater, and stream III showed the highest representation of freshwater microbial communities, especially in spring -- attributing on average 74% of genera. However, microbial communities of stream I and II were high representations of soil and sediment-associated genera, which may explain the compositionally diverse structure. Soil and sediment-associated genera may enter the freshwater ecosystems from headwater, therefore stream I and II as low order streams and were closed to headwater contained higher soil and sediment-associated microbial populations. Meanwhile, enhanced enrichment processes for freshwater-associated taxa in stream III as high order downstream. Although few studies reported putative source environments of bacterioplankton communities, macroecological researches have long been assessed for the last 4 decades (Statzner and Higler 1985). The previous study showed headwater streams contained large amounts of allochthonous community, as stream size increase, the significance of autochthonous primary production was enhanced as well as organic transport from upstream (Statzner and Higler 1985).

Sediment and soil as the potential sources were significantly more dominant in fall than spring across all streams in this study ($P < 0.05$). In stream III, freshwater-associated taxa were significantly higher in Spring than Fall, and the enrichment processes may be amplified in Spring compared to Fall. In both stream I and II sites, there is no significant difference of freshwater-associated genera between spring and fall, but the attributions of unknown sources were significantly increased in Spring, which means taxa from more potential sources into the freshwater ecosystems besides sediment and soil.

It is challenging to infer the reasons for the seasonal shifts of bacterioplankton communities because of too many impact factors, such as precipitation, temperature, growth of vegetation, and sunlight intensity. The influence of each factor should be determined in the future. For macroinvertebrate communities in freshwater ecosystems, energy inputs and organic substrates are the key factors for seasonal shift (Statzner and Higler 1985). Next step, the functions of bacterioplankton communities based on metagenomic methods will be used to figure out the response for energy and substrates.

4. Conclusions

Bacterioplankton communities in three unimpaired streams within an ecoregion were studied. The SADs were all followed by the Poisson lognormal model, which was consistent with other freshwater ecosystems in previous studies. Meanwhile, the dominated taxa in the three sites were typical freshwater taxa that were well documented (Newton, Jones, et al., 2011).

In this work, although concentrations of FIB and relative abundances of some dominant taxa showed relatively stable levels across sampling times at each site, the two streams closed to headwater were distinct from the highest-order stream. The streams closed to headwater host elevated sediment and soil-associated bacteria and diversity levels compared to the highest-order stream in which the microbial community repeatedly enriches for select taxa.

Although relatively stable bacterioplankton community compositions within a one-month sampling period at each unimpaired stream, there were pronounced seasonal shifts for all three sites. In spring, more potential bacterial sources other than sediment and soil enter into the stream headwaters, and the streams closed to headwater showed higher diversity levels than Fall. Meanwhile, the highest-order stream showed enhanced enrichment processes for select taxa and decreased diversity levels in Spring compared to Fall.

Overall, this work suggests that delimitation of stream types and sampling periods is an important step for defining reference conditions of bacterioplankton communities within an ecoregion. It should also be mentioned that, due to the limited sampling stream sites in this study, observations need to be validated in different streams and ecoregions.

References

- Aazami, J., A. Esmaili Sari, A. Abdoli, H. Sohrabi and P. J. Van den Brink (2015). "Assessment of Ecological Quality of the Tajan River in Iran Using a Multimetric Macroinvertebrate Index and Species Traits." Environmental Management **56**(1): 260-269.
- Attrill, M. J. and M. H. Depledge (1997). "Community and population indicators of ecosystem health: Targeting links between levels of biological organisation." Aquatic Toxicology **38**(1-3): 183-197.
- Bolyen, E., J. R. Rideout, M. R. Dillon, N. A. Bokulich, C. C. Abnet, G. A. Al-Ghalith, H. Alexander, E. J. Alm, M. Arumugam and F. Asnicar (2019). "Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2." Nature biotechnology **37**(8): 852-857.
- Callahan, B. J., P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson and S. P. Holmes (2016). "DADA2: high-resolution sample inference from Illumina amplicon data." Nature methods **13**(7): 581.
- Caporaso, J. G., C. L. Lauber, W. A. Walters, D. Berg-Lyons, J. Huntley, N. Fierer, S. M. Owens, J. Betley, L. Fraser and M. Bauer (2012). "Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms." The ISME journal **6**(8): 1621.
- Dale, V. H. and S. C. Beyeler (2001). "Challenges in the development and use of ecological indicators." Ecological Indicators **1**(1): 3-10.
- Fisher, J. C., R. J. Newton, D. K. Dila and S. L. McLellan (2015). "Urban microbial ecology of a freshwater estuary of Lake Michigan." Elementa (Washington, DC) **3**.
- Griffith, G., J. Omernik and S. Azevedo (1997). "Ecoregions of Tennessee: US Environmental Protection Agency National Health and Environmental Effects Research Laboratory EPA."
- Hassell, N., K. A. Tinker, T. Moore and E. A. Ottesen (2018). "Temporal and spatial dynamics in microbial community composition within a temperate stream network." Environmental Microbiology **20**(10): 3560-3572.
- Hawkins, C. P., J. R. Olson and R. A. Hill (2010). "The reference condition: predicting benchmarks for ecological and water-quality assessments." Journal of the North American Benthological Society **29**(1): 312-343.
- Hughes, R. M., D. P. Larsen and J. M. Omernik (1986). "Regional reference sites: a method for assessing stream potentials." Environmental Management **10**(5): 629-635.
- Knights, D., J. Kuczynski, E. S. Charlson, J. Zaneveld, M. C. Mozer, R. G. Collman, F. D. Bushman, R. Knight and S. T. Kelley (2011). "Bayesian community-wide culture-independent microbial source tracking." Nature Methods **8**(9): 761-763.
- Laperriere, S. M., R. H. Hilderbrand, S. R. Keller, R. Trott and A. E. Santoro (2020). "Headwater stream microbial diversity and function across agricultural and urban land use gradients." Applied and Environmental Microbiology **86**(11).
- Li, Y., J. Zhang, J. Zhang, W. Xu and Z. Mou (2019). "Microbial Community Structure in the Sediments and Its Relation to Environmental Factors in Eutrophicated Sancha Lake." International journal of environmental research and public health **16**(11): 1931.

- Lurry, D. L. and C. M. Kolbe (2000). Interagency field manual for the collection of water-quality data, US Department of the Interior, US Geological Survey.
- Martins, I., R. Ligeiro, R. M. Hughes, D. R. Macedo and M. Callisto (2018). "Regionalisation is key to establishing reference conditions for neotropical savanna streams." Marine and Freshwater Research **69**(1): 82-94.
- Newton, R. J., S. E. Jones, A. Eiler, K. D. McMahon and S. Bertilsson (2011). "A guide to the natural history of freshwater lake bacteria." Microbiology and molecular biology reviews **75**(1): 14-49.
- Oksanen, J., F. G. Blanchet, R. Kindt, P. Legendre, P. R. Minchin, R. B. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens and H. Wagner (2013). "Vegan: Community Ecology Package. R package version 2.0-10." Vegan: Community Ecology Package.
- Ozbayram, E. G., L. Koker, R. Akçaalan, F. Aydın, A. Ertürk, O. Ince and M. Albay (2020). "Contrasting the Water Quality and Bacterial Community Patterns in Shallow and Deep Lakes: Manyas vs. Iznik." Environmental Management: 1-7.
- Shoemaker, W. R., K. J. Locey and J. T. Lennon (2017). "A macroecological theory of microbial biodiversity." Nature ecology & evolution **1**(5): 1-6.
- Simonin, M., K. A. Voss, B. A. Hassett, J. D. Rocca, S. Y. Wang, R. L. Bier, C. R. Violin, J. P. Wright and E. S. Bernhardt (2019). "In search of microbial indicator taxa: shifts in stream bacterial communities along an urbanization gradient." Environmental Microbiology **21**(10): 3653-3668.
- Sinclair, L., O. A. Osman, S. Bertilsson and A. Eiler (2015). "Microbial community composition and diversity via 16S rRNA gene amplicons: Evaluating the illumina platform." PLoS ONE **10**(2).
- Staley, C., T. J. Gould, P. Wang, J. Phillips, J. B. Cotner and M. J. Sadowsky (2014). "Bacterial community structure is indicative of chemical inputs in the Upper Mississippi River." Frontiers in Microbiology **5**(OCT).
- Statzner, B. and B. Higler (1985). "Questions and comments on the river continuum concept." Canadian Journal of Fisheries and Aquatic Sciences **42**(5): 1038-1044.
- Tang, X., G. Xie, K. Shao, Y. Hu, J. Cai, C. Bai, Y. Gong and G. Gao (2020). "Contrast diversity patterns and processes of microbial community assembly in a river-lake continuum across a catchment scale in northwestern China." Environmental Microbiome **15**: 1-17.
- Teachey, M. E., J. M. McDonald and E. A. Ottesen (2019). "Rapid and stable microbial community assembly in the headwaters of a third-order stream." Applied and Environmental Microbiology **85**(11).
- Team, R. C. (2017). "R Foundation for Statistical Computing; Vienna, Austria: 2016. R: A language and environment for statistical computing." URL <http://www.R-project.org/>[Google Scholar].
- Thompson, L. R., J. G. Sanders, D. McDonald, A. Amir, J. Ladau, K. J. Locey, R. J. Prill, A. Tripathi, S. M. Gibbons and G. Ackermann (2017). "A communal catalogue reveals Earth's multiscale microbial diversity." Nature **551**(7681): 457-463.

Wyckoff, K. N., S. Chen, A. J. Steinman and Q. He (2017). "Impact of Roadway Stormwater Runoff on Microbial Contamination in the Receiving Stream." Journal of environmental quality **46**(5): 1065-1071.

Chapter 6

Microbial Diversity in Unimpaired Streams across Ecoregions

Abstract

With the development of the high-throughput 16S rRNA Amplicon Sequencing techniques, the use of microorganisms for the ecological assessment of natural water bodies has greatly increased. Typically, this sort of assessment is done with streams that have major anthropomorphic impacts without a significant understanding of the unimpaired stream dynamics. The understanding of the natural spatial variation of the microbial diversities in these unimpaired streams is important to identify reference bacterial composition prior to the assessment of stream ecological conditions. In this study, microbial community composition was utilized to characterize the spatial variations among 15 streams across three different ecoregions, where these streams corresponded to two stream types: low-order streams (i.e., closest to the headwater) and high-order streams. The results indicated that culture-based fecal indicator bacteria (FIB) showed low concentrations and relatively stable across these unimpaired streams, but the microbial communities revealed spatial variations. The microbial communities were distinct between low-order streams and high-order streams, which corresponded to low-order streams having elevated sediment and soil-associated bacteria, as well as higher community diversity levels. In contrast, freshwater-associated bacteria were more prevalent in high-order streams. Moreover, microbial communities in both low-order streams and high-order streams exhibited significant distance-decay patterns, where this trend was stronger in low-order streams. These results show impacts of stream size and order have on the microbial community structure, as well as the geographical variability in streams with virtually no anthropomorphic impact.

1. Introduction

More recent studies have used microorganisms to assess the ecological conditions of freshwater ecosystems due to the rapid development of the high-throughput 16S rRNA Amplicon Sequencing techniques in microbial community analyses (Lau, Washington, et al., 2015; Li, Li, et al., 2017; Niu, Li, et al., 2018; Simonin, Voss, et al., 2019; Hilderbrand, Keller, et al., 2020). The use of natural water microbial community as background bacterial composition has been a critical component to identify deviations from the high ecological quality in the absence of anthropogenic

stresses at other stream sites (Hägglund, Bäckman, et al., 2018; Breton-Deval, Sanchez-Flores, et al., 2019). However, the microbial characteristics of streams vary according to their biogeographic features (Tolkkinen, Mykrä, et al., 2016; Vilmi, Karjalainen, et al., 2016; Hosen, Febria, et al., 2017; Teachey, McDonald, et al., 2019). Moreover, even within a stream network, bacterial compositions were different between headwater and higher-order downstream (Hassell, Tinker, et al., 2018; Teachey, McDonald, et al., 2019). Thus, spatial and order-type-specific variations of reference conditions of microorganisms in the streams with high ecological status must be understood.

Some recent studies have used local stream water samples with low fecal indicator bacteria (FIB) counts from the same watershed as background bacterial composition (Hägglund, Bäckman, et al., 2018; Breton-Deval, Sanchez-Flores, et al., 2019). However, the low and stable levels of FIB were not able to present the stability of the microbial community. Meanwhile, few studies described the natural variation of background bacterial composition. In this way, strong variations may occur for characterizing the ecological condition of freshwater streams based on sequence-based microorganism approaches.

This study aimed to evaluate the influences of spatial and stream order variations on the microbial community compositions in reference streams. This work spanned independent streams of different order types across three ecoregions. We hypothesized that FIB showed low concentrations and relatively stable across sampling locations while the microbial communities revealed spatial variation. Furthermore, Habitat filtering by stream order level, and not ecoregion, was the major driver shaping bacterioplankton communities.

2. Materials and Methods

2.1 Sites selection

The study was conducted in East Tennessee, USA. 15 reference streams across three ecoregions were selected (Figure 6.1) – 67f The Southern Limestone/Dolomite Valleys and Low Rolling Hills, 68a Cumberland Plateau, and 71g Eastern Highland Rim ((Griffith, Omernik, et al., 1997). The stream names, order levels, and Geographical distance from headwater were shown in table 6.S1. The sampling streams with 1 or 2 orders were defined as low order streams closed to headwater, and 3 or 4 order streams were defined as high order streams. Water samples were collected by

dipping a sterile 1 L bottle below the stream surface following Standard Method (Lurry and Kolbe 2000) from May to June 2019 within 30 days in order to exclude temporal variations. After collection, all samples were transported using a cooler with ice and delivered to the laboratory within 6 hours. For high-throughput sequencing, the water samples were filtered through 0.45 μ m filters (MilliporeSigma™ HAWG047S6), and the filters were stored at -80°C until DNA extraction.

2.2 Fecal indicator bacteria test

Fecal indicator bacteria (*E. coli* and total coliforms) tests were performed within 8 hours of sample collection. *E. coli* and total coliforms were determined by m-Colibblue 24 ® media following the protocol provided by the manufacturer (Hach Co., Loveland, CO) as described in previous chapters.

2.3 DNA extraction and high-throughput 16S rRNA amplicon sequencing

DNA from the colony isolates was extracted by the FastDNA Spin Kit for Soil (MP Biomedicals) following the manufacturer's protocol.

High-Throughput 16S rRNA Amplicon Sequencing was performed as previously described (Wyckoff, Chen, et al., 2017). The PCR amplification used primers 515-F and 806-R to target the V4 region of the 16S rRNA gene (Caporaso, Lauber, et al., 2012). PCR products were purified using the Wizard SV Gel and PCR Clean-Up System (Promega), and paired-end sequencing was completed using the Illumina MiSeq platform.

2.4 Bioinformatic and statistical analyses.

Analysis of sequence reads was conducted using QIIME 2 (Bolyen, Rideout, et al., 2019). Amplicon sequence variants (ASVs) enumeration was performed using the DADA2 pipeline (Callahan, McMurdie, et al., 2016). The obtained ASVs were then matched to the Greengenes 13.8 database.

Statistical analysis and data visualization were carried out using R version 3.50 (Team 2017). PERMANOVA was applied to assess the difference of community composition, α -diversity indices (the Shannon index, the Simpson index, and Pielou's evenness), and Bray–Curtis (abundance-based) dissimilarity was calculated to represent the taxonomic β -diversity using the vegan R package (Oksanen, Blanchet, et al., 2013).

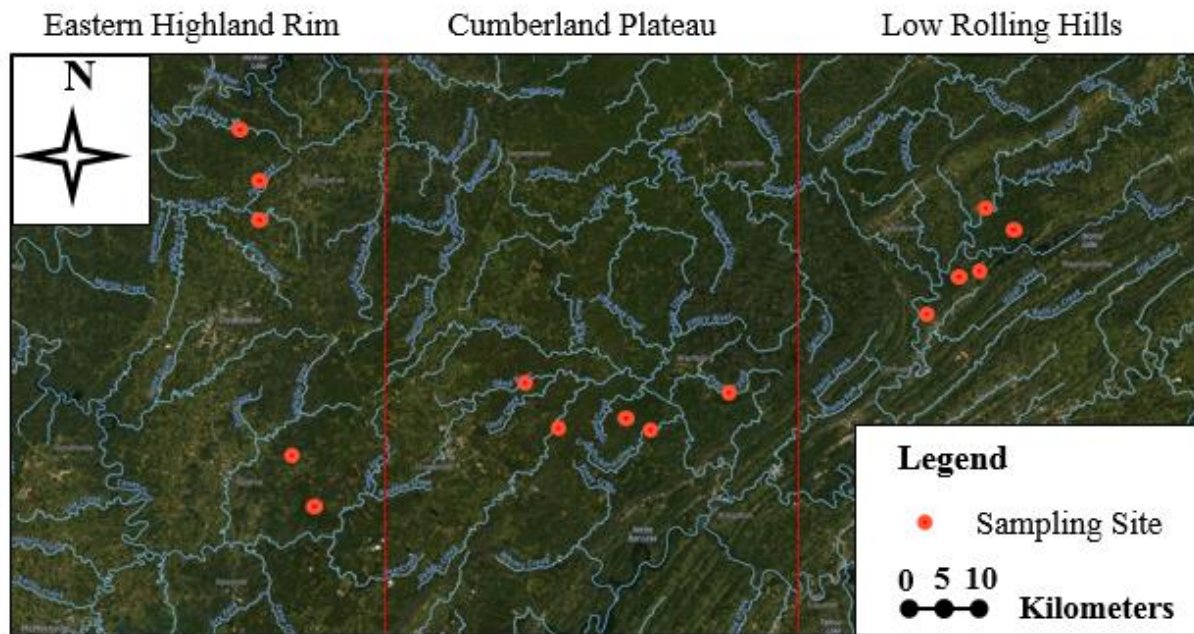


Figure 6.1. The maps of the 15 streams across three ecoregions in East TN -- 67f The Southern Limestone/Dolomite Valleys and Low Rolling Hills, 68a Cumberland Plateau, and 71g Eastern Highland Rim 67f eco-region.

The sequence count table of 16S rRNA gene amplicon studies from the EMP was downloaded (Thompson, Sanders, et al., 2017). EMP samples from freshwater, freshwater sediment, and soil, were selected and merged with our stream water samples. The proportion of each stream water microbiota attributable to freshwater, sediment, and soil at the genus level were estimated using Sourcetracker (Knights, Kuczynski, et al., 2011), which was using stream water microbiota as the sink and those in the soil as sources.

3. Results and Discussions

3.1 Fecal indicator bacteria in the ecoregion streams

All 15 streams in this study showed the average concentrations of *E. coli* were from 20 to 102 CFU/100mL. Meanwhile, the average concentrations of total coliforms were from 377 CFU/100mL to 1452 CFU/100mL (Figure 6.2). The levels of *E. coli* in the 15 streams were all less than 126 CFU/100mL, which is the threshold of State of Tennessee water quality standards (TDEC 2007).

The levels of FIB, including *E. coli* and total coliforms, in the reference streams were not significantly different across sampling locations ($P > 0.05$). The natural spatial variations of FIB in streams were limited in this study.

3.2 Habitat filtering by stream order level, and not ecoregion, is the major driver shaping microbial structures.

Although the levels of FIB were not significantly different across these unimpaired streams, distinguished alpha and beta microbial diversities were observed between different stream order levels (Figure 6.3 and Figure 6.4). Principal-coordinate-analysis (PCoA) showed a clear separation of low-order streams and high-order streams, while the same order streams clustered together even though they were from different ecoregions (Figure 6.3A). Cluster Dendrogram (Figure 6.3B) also showed similar patterns. Low-order streams and high-order streams were clustered as two groups. Moreover, Bray-Curtis similarity analyses revealed the similarities between the two type streams are considerably lower than the similarities within the same order level streams (Figure 6.3C). It has long been assessed for the last 4 decades that stream community structure and function distributed along river gradients (Statzner and Higler 1985). Although most researches studied macroecology, the river continuum concept may also be applied to the

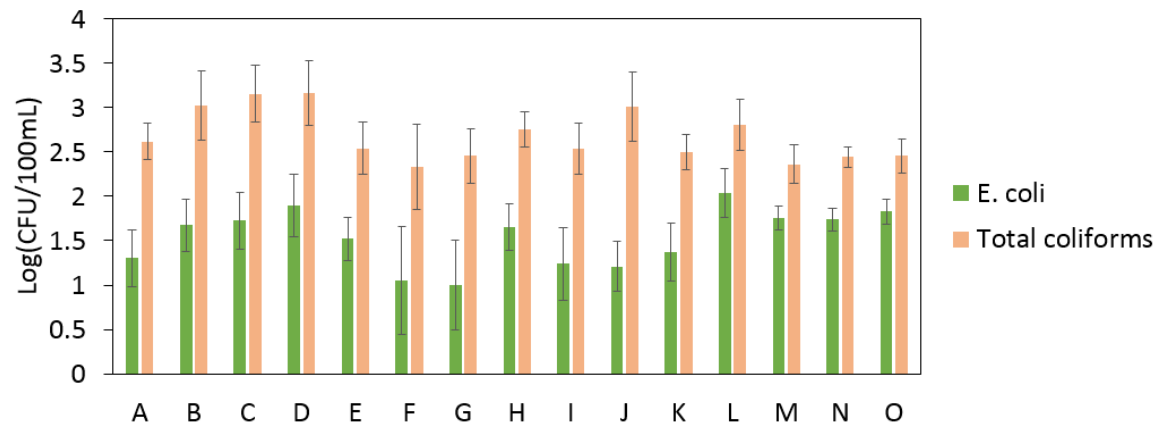


Figure 6.2. *E. coli* and total coliforms in ecoregion streams. FIB levels were relatively low and stable across sampling locations.

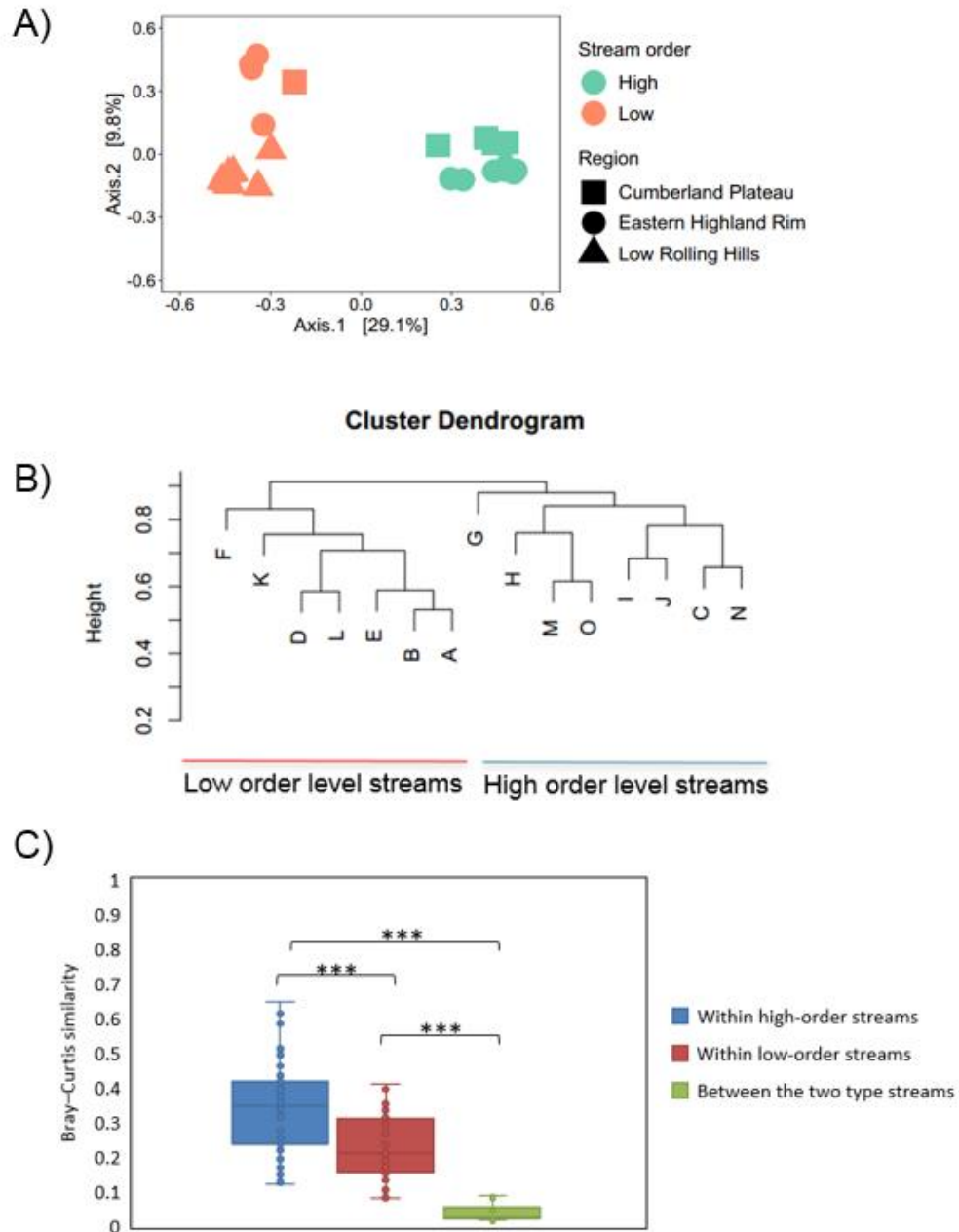


Figure 6.3. Beta-diversity showed Habitat filtering by stream order level, and not ecoregion, is the major driver shaping bacterioplankton communities. (A) PCoA; (B) Cluster Dendrogram; (C) Bray–Curtis similarity. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

microbial community: headwater represent the maximum interface with the landscape, and the inputs are comprised of short- and long-chain organics, but the more refractory and relatively high molecular weight compounds caused the incorporation of new functional species into the community for an efficiency adjustment towards the utilization of energy in downstream (Statzner and Higler 1985). In the future, the functions of bacterioplankton communities based on metagenomic methods will be used to figure out the response for energy and substrates.

Alpha diversity (the Shannon index, the Simpson index, and Pielou's evenness) also showed a significant difference between low-order and high-order streams. The group of high-order streams was significantly less richness, evenness, and Shannon diverse than the group of low-order streams ($P < 0.001$). Several studies found headwater hosted higher microbial diversity than the high-order downstream in a stream network (Hassell, Tinker, et al., 2018; Teachey, McDonald, et al., 2019), the independent streams in this study also showed similar microbial diverse patterns.

3.3 The abundances of individual ASVs across stream sites

Decreases in alpha diversities in the high-order streams were accompanied by increases in the relative abundance of top ASVs (Figure 6.5). All samples from high-order streams hosted the ASVs that were more than 7% relative abundances, and even some ASVs were higher than 20%. However, none of the ASVs was more than 5% in samples from low-order streams.

All abundant ASVs in the high-order streams were also found in low-order streams, though the relative abundances might be less in the low-order streams. This phenomenon may result by enriching the microbial community repeatedly for select taxa across high-order freshwater ecosystems. Meanwhile, the high abundant taxonomies in the high-order streams, like *Flavobacterium*, *Comamonadaceae*, *Planktophila*, *Rhodoluna*, *Pseudarcicella*, *Sporichthyaceae*, *Polynucleobacter*, and *Fluviicola*, were also generally found in other freshwater ecosystems (Newton, Jones, et al., 2011).

3.4 Putative source environments for the microbial community in the reference stream water

In order to explore habitat associations for the microbial community in the reference streams, a Bayesian method was used to determine putative source environments of bacterioplankton communities at the genus level (Figure 6.6), including soil, sediment, and freshwater, based on the Earth Microbiome Project (EMP) datasets (Thompson, Sanders, et al., 2017). No doubt, the most

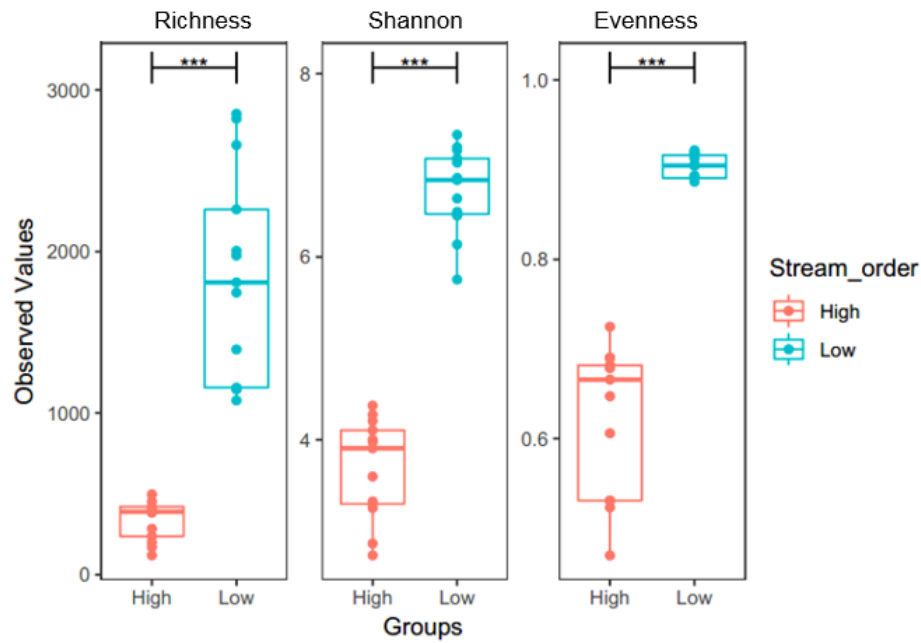


Figure 6.4. Alpha diversities (the Shannon index, the Simpson index, and Pielou's evenness) for low-order streams and high-order streams. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

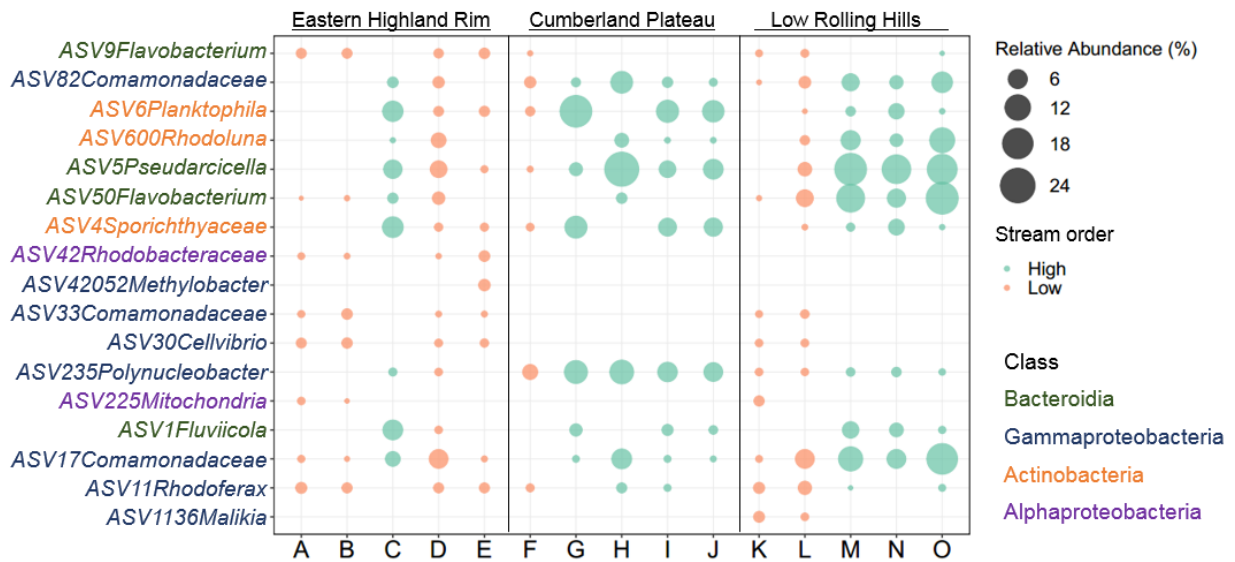


Figure 6.5. Balloon plot showing the relative abundance of the ASVs found in different streams. The top 3 most abundant ASVs in each stream are displayed.

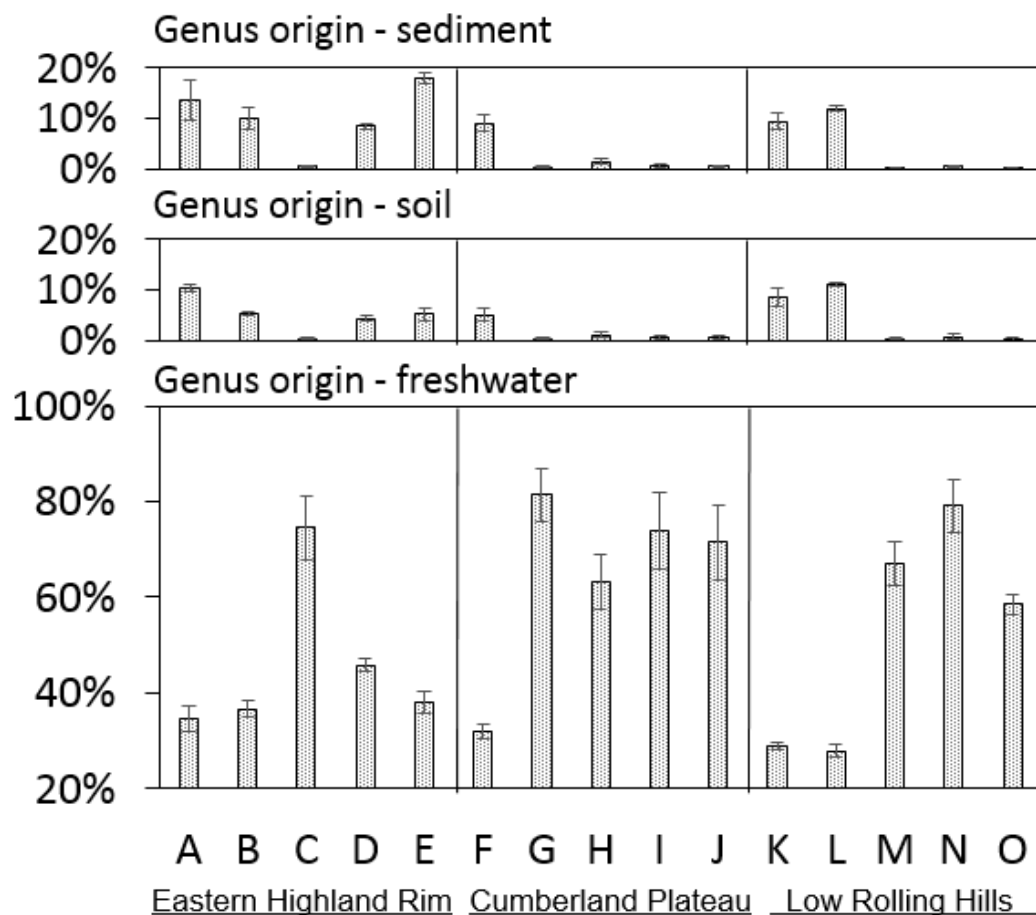


Figure 6.6. Percentage of stream water bacterial genera attributable to sediment, soil, and freshwater ecosystems as determined by Sourcetracker.

dominant known source was freshwater, attributing from 27% to 81%, especially for high-order streams where the representations of freshwater microbial communities were higher than 63%. However, the representations of freshwater microbial communities in low-order streams were less than 45%. Further, higher representations of sediment-associated genera (8% -- 18%) in low-order streams than high-order streams (less than 1%) were found in this study, as the similar as sediment-associated genera patterns, higher representations of soil-associated genera (4% -- 11%) in low-order streams than high-order streams (less than 0.7%).

Low-order streams are closed to headwater that acts as a primary entry point for soil and sediment associated microorganisms into freshwater ecosystems, which may explain the compositionally diverse structure. Meanwhile, high-order streams, as downstream in a river network, enriched the selected freshwater-associated taxa and showed a reduction in microbial diversity.

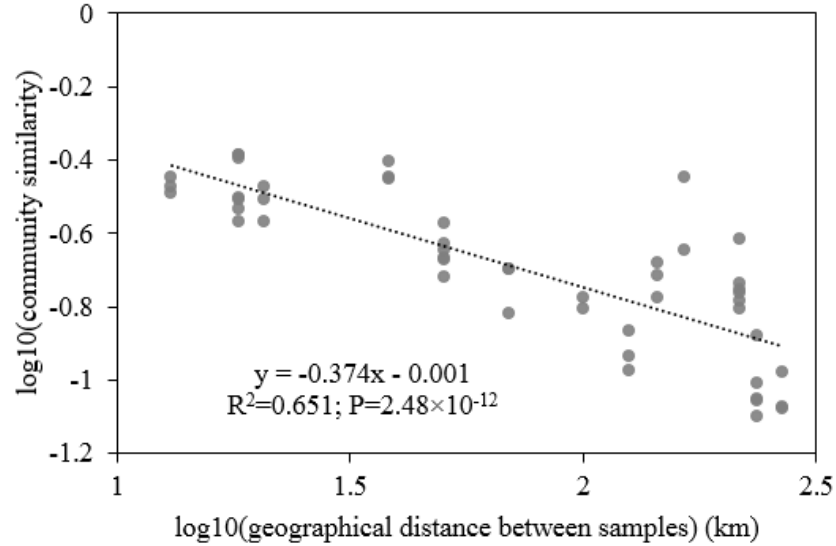
3.5 spatial factors related to microbial community composition.

Actually, Bray-Curtis similarity analyses (Figure 6.3C) not only showed the similarities between the two order-type streams are considerably lower than the similarities within the same order level streams, but also showed the similarities within the low-order streams were significantly lower than the high-order streams. Meanwhile, Figure 6.5 showed ASV42 *Rhodobacteraceae* was exclusively positive in the low-order streams from the Eastern Highland Rim ecoregion, and ASV1136 *Malikia* was exclusively positive in the low-order streams from the Low Rolling Hills ecoregion. These results indicated the spatial variations within the low-order streams across the three ecoregions.

Although all abundant ASVs in the high-order streams can be found across different ecoregions, some of them showed different relative abundances between different regions. For example, the relative abundances of ASV235 *Polynucleobacter* were significantly higher in the streams from the Cumberland Plateau ecoregion than others, and the relative abundances of ASV50 *Flavobacterium* were significantly higher in the streams from the Low Rolling Hills ecoregion than others. These patterns illustrated the local specific contribution may be a driven factor to shape the distribution of microorganisms in high-order freshwater ecosystems.

The spatial predictors of microbial community compositions showed similar results with the distribution of abundant ASVs (Figure 6.7). Significantly negative relationships between the Bray–

A) Low-order level streams



B) High-order level streams

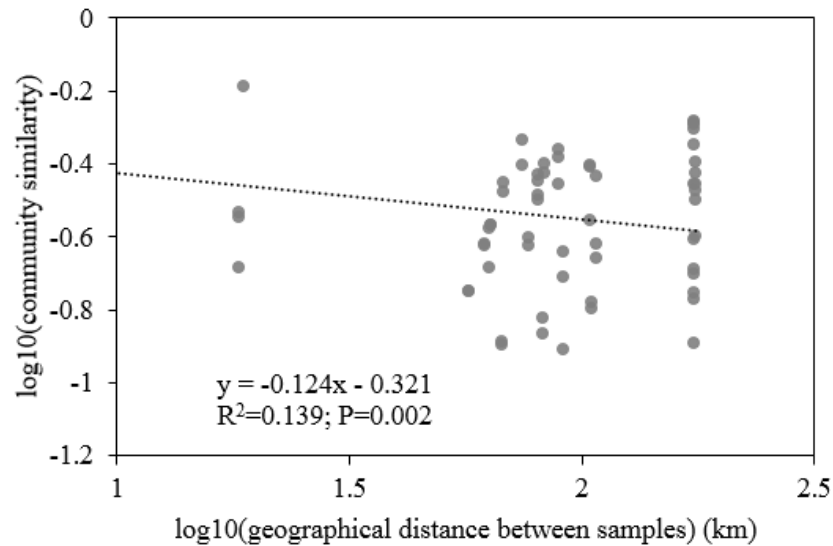


Figure 6.7. Distance-decay relationships between the bacterial community similarity; (A) samples from low-order level streams; (B) samples from high-order level streams. The bacterial communities in two types of streams showed significant distance decay across ecoregions (the slope of the distance-decay relationship estimated by least squares regression models were significantly less than zero), the distance-decay slopes of bacterial communities in low-order level streams (slope = -0.374, $P < 0.0001$) were significantly steeper than high-order level streams (slope = -0.124, $P < 0.002$).

Curtis similarity of microbial communities and the geographical distance were found in both the low-order streams and the high-order streams ($P < 0.01$). This was not surprising because multiple studies found niche partitioning processes to bacterial communities in freshwater ecosystems (Tolkkinen, Mykrä, et al., 2016; Vilmi, Karjalainen, et al., 2016; Hosen, Febria, et al., 2017; Teachey, McDonald, et al., 2019). Moreover, the low-order streams showed stronger correlations between spatial distances and microbial communities than the high-order streams. The reason may be that the enrichment process of the selected freshwater-associated microbes in high-order streams offset spatial heterogeneity at the east Tennessee geographical scale in this study.

4. Conclusions

In recent years, various microbial community-based methods have been developed, tested, and suggested for the assessment of freshwater ecological health. However, challenges were imposed by the spatial variability of the microbial community in these ecosystems. This study investigated microorganisms in 15 reference streams across three ecoregions. The results showed that low and relatively stable concentrations of FIB across sampling locations, but the microbial communities revealed spatial variation patterns. While there was spatial variation observed, the stream order level was the more significant driver for stream microbial community. The low-order streams closed to headwater host elevated sediment and soil-associated bacteria and diversity levels compared to the high-order streams in which the microbial community repeatedly enriches freshwater-associated taxa in this study. Moreover, microbial communities in both low-order and high-order streams exhibited significant distance-decay patterns, but the communities had a stronger distance-decay relationship in low-order streams. Overall, this work suggests a selection of nearby and similar order type reference streams is essential for the establishment of reference microbial conditions.

References

- Bolyen, E., J. R. Rideout, M. R. Dillon, N. A. Bokulich, C. C. Abnet, G. A. Al-Ghalith, H. Alexander, E. J. Alm, M. Arumugam and F. Asnicar (2019). "Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2." Nature biotechnology **37**(8): 852-857.
- Breton-Deval, L., A. Sanchez-Flores, K. Juárez and R. Vera-Estrella (2019). "Integrative study of microbial community dynamics and water quality along The Apatlaco River." Environmental Pollution **255**: 113158.
- Callahan, B. J., P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson and S. P. Holmes (2016). "DADA2: high-resolution sample inference from Illumina amplicon data." Nature methods **13**(7): 581.
- Caporaso, J. G., C. L. Lauber, W. A. Walters, D. Berg-Lyons, J. Huntley, N. Fierer, S. M. Owens, J. Betley, L. Fraser and M. Bauer (2012). "Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms." The ISME journal **6**(8): 1621.
- Griffith, G., J. Omernik and S. Azevedo (1997). "Ecoregions of Tennessee: US Environmental Protection Agency National Health and Environmental Effects Research Laboratory EPA."
- Häggglund, M., S. Bäckman, A. Macellaro, P. Lindgren, E. Borgmästars, K. Jacobsson, R. Dryselius, P. Stenberg, A. Sjödin and M. Forsman (2018). "Accounting for bacterial overlap between raw water communities and contaminating sources improves the accuracy of signature-based microbial source tracking." Frontiers in microbiology **9**: 2364.
- Hassell, N., K. A. Tinker, T. Moore and E. A. Ottesen (2018). "Temporal and spatial dynamics in microbial community composition within a temperate stream network." Environmental Microbiology **20**(10): 3560-3572.
- Hilderbrand, R. H., S. R. Keller, S. M. Laperriere, A. E. Santoro, J. Cessna and R. Trott (2020). "Microbial communities can predict the ecological condition of headwater streams." PloS one **15**(8): e0236932.
- Hosen, J. D., C. M. Febria, B. C. Crump and M. A. Palmer (2017). "Watershed urbanization linked to differences in stream bacterial community composition." Frontiers in Microbiology **8**(AUG).
- Knights, D., J. Kuczynski, E. S. Charlson, J. Zaneveld, M. C. Mozer, R. G. Collman, F. D. Bushman, R. Knight and S. T. Kelley (2011). "Bayesian community-wide culture-independent microbial source tracking." Nature Methods **8**(9): 761-763.
- Lau, K. E. M., V. J. Washington, V. Fan, M. W. Neale, G. Lear, J. Curran and G. D. Lewis (2015). "A novel bacterial community index to assess stream ecological health." Freshwater Biology **60**(10): 1988-2002.
- Li, J., Y. Li, B. Qian, L. Niu, W. Zhang, W. Cai, H. Wu, P. Wang and C. Wang (2017). "Development and validation of a bacteria-based index of biotic integrity for assessing the ecological status of urban rivers: A case study of Qinhuai River basin in Nanjing, China." Journal of Environmental Management **196**: 161-167.
- Lurry, D. L. and C. M. Kolbe (2000). Interagency field manual for the collection of water-quality data, US Department of the Interior, US Geological Survey.

- Newton, R. J., S. E. Jones, A. Eiler, K. D. McMahon and S. Bertilsson (2011). "A guide to the natural history of freshwater lake bacteria." Microbiology and molecular biology reviews **75**(1): 14-49.
- Niu, L., Y. Li, P. Wang, W. Zhang, C. Wang, J. Li and H. Wu (2018). "Development of a microbial community-based index of biotic integrity (MC-IBI) for the assessment of ecological status of rivers in the Taihu Basin, China." Ecological Indicators **85**: 204-213.
- Oksanen, J., F. G. Blanchet, R. Kindt, P. Legendre, P. R. Minchin, R. B. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens and H. Wagner (2013). "Vegan: Community Ecology Package. R package version 2.0-10." Vegan: Community Ecology Package.
- Simonin, M., K. A. Voss, B. A. Hassett, J. D. Rocca, S. Y. Wang, R. L. Bier, C. R. Violin, J. P. Wright and E. S. Bernhardt (2019). "In search of microbial indicator taxa: shifts in stream bacterial communities along an urbanization gradient." Environmental Microbiology **21**(10): 3653-3668.
- Statzner, B. and B. Higler (1985). "Questions and comments on the river continuum concept." Canadian Journal of Fisheries and Aquatic Sciences **42**(5): 1038-1044.
- Teachey, M. E., J. M. McDonald and E. A. Ottesen (2019). "Rapid and stable microbial community assembly in the headwaters of a third-order stream." Applied and Environmental Microbiology **85**(11).
- Team, R. C. (2017). "R Foundation for Statistical Computing; Vienna, Austria: 2016. R: A language and environment for statistical computing." URL <http://www.R-project.org/>[Google Scholar].
- Thompson, L. R., J. G. Sanders, D. McDonald, A. Amir, J. Ladau, K. J. Locey, R. J. Prill, A. Tripathi, S. M. Gibbons and G. Ackermann (2017). "A communal catalogue reveals Earth's multiscale microbial diversity." Nature **551**(7681): 457-463.
- Tolkkinen, M. J., H. Mykrä, R. Virtanen, M. Tolkkinen, T. Kauppila, L. Paasivirta and T. Muotka (2016). "Land use impacts on stream community composition and concordance along a natural stress gradient." Ecological Indicators **62**: 14-21.
- Vilmi, A., S. M. Karjalainen, T. Nokela, K. Tolonen and J. Heino (2016). "Unravelling the drivers of aquatic communities using disparate organismal groups and different taxonomic levels." Ecological Indicators **60**: 108-118.
- Wyckoff, K. N., S. Chen, A. J. Steinman and Q. He (2017). "Impact of Roadway Stormwater Runoff on Microbial Contamination in the Receiving Stream." Journal of environmental quality **46**(5): 1065-1071.

Chapter 7

The Impact of Storm Events on Receiving Stream Water Bacterial Community

Abstract

Microbial contaminants in stormwater runoff are subsequently transported to receiving streams and may be responsible for serious ecological impacts. The objective of this study was to investigate the microbial community composition-based method for assessment of the impacts of stormwater runoff on receiving streams at three sites in East Tennessee, USA. No doubt, turbidity and FIB levels were higher in the wetflow than baseflow at each receiving creek in this study, and the 16S rRNA amplicon sequencing method also highlighted the impact that stormwater runoff had on general water quality. The dominated taxa in baseflow water samples were typical freshwater taxa that were also found in nearby unimpaired reference streams, but the runoff-associated microorganisms that were rare in baseflow dramatically increased in wetflow. The Sourcetracker analysis revealed that stormwater runoff and the receiving stream baseflow were the major contributors to the overall microbial community in wetflow across the three study sites. In general, findings in this study could be helpful for developing stormwater runoff control measures.

1. Introduction

Microbial contaminants are major concerns in the pollution of water bodies (Ashbolt and Bruno 2003; Federigi, Verani, et al., 2019). Those contaminations can be found in runoff and subsequently transported to water bodies, and studies have reported a high prevalence of microbial contaminants in sewer overflows, defective septic systems, roof-harvested rainwater, and agricultural runoff. However, the researches of microbial contamination in stormwater runoff from the impervious area were rarely (Saget, Chebbo, et al., 1996; Jiang, Noble, et al., 2001; Thurston, Gerba, et al., 2001; Ahmed, Hamilton, et al., 2018; Boehm, Graham, et al., 2018; Hamilton, Reyneke, et al., 2019). It is impossible to monitor all microorganisms in water bodies for the determination of microbial contamination because of the great diversity of microbes. Therefore, fecal indicator bacteria (FIB), like *Escherichia coli* and total coliforms, were selected to assess the microbial quality due to high levels in feces and the likely presence of potential pathogens (USEPA 1986). However, FIB cannot identify the sources of microbial pollution on the receiving water bodies. Apportioning sources is essential for developing best management practices.

Advances in next-generation sequencing have potentially shown new opportunities for microbial contamination assessment through analysis of microbial communities (Shendure, Porreca, et al., 2005; Wang, Garrity, et al., 2007; Lazarevic, Whiteson, et al., 2009; Claesson, Wang, et al., 2010; McLellan, Huse, et al., 2010; Kim and Bae 2011; Ashbolt 2015; Eren, Morrison, et al., 2015; Madoui, Engelen, et al., 2015; Nshimiyimana, Freedman, et al., 2015; Federigi, Verani, et al., 2019). The high-throughput 16S rRNA amplicon sequencing is one of the culture-independent methods used for microbiome analyses (Liu, Lozupone, et al., 2007; Kennedy, Hall, et al., 2014; Logares, Sunagawa, et al., 2014; Schloss, Westcott, et al., 2015; Sinclair, Osman, et al., 2015). Moreover, Sourcetracker, one of the computational models based on sequencing data, takes contamination sources as entire source communities into the contaminated samples, where the mixing proportions of contributors are trying to be determined (Knights, Kuczynski, et al., 2011).

The objective of this study was to investigate microbial community composition-based methods for the assessment of the impacts of stormwater runoff to receiving stream water. Three sites were selected for sampling in East Tennessee. For each stream, samples were taken under baseflow and wetflow conditions; additionally, samples of stormwater runoff entering each stream were collected. Turbidity, FIB, and microbial communities were analyzed to address the following questions: (1) How do storm events impact turbidity, FIB, and microbial communities in the receiving streams across the study sites? (2) What are the relationships of bacteria among stormwater runoff, receiving stream baseflow, and wetflow at the three sampling sites? (3) What are the relative contributions of microbes from stormwater runoff and receiving stream baseflow to wetflow?

2. Materials and Methods

2.1 Sites selection

The study was conducted in East Tennessee, USA (Figure 7.1). Three types of water samples, including stormwater runoff, the receiving stream baseflow, and wetflow at the immediate downstream of runoff effluent, were collected from October 2018 to February 2020.

Stream water samples were collected by dipping a sterile 1 L bottle below the stream surface following Standard Method (Lurry and Kolbe 2000). The baseflow water samples were collected prior to the anticipated storms.

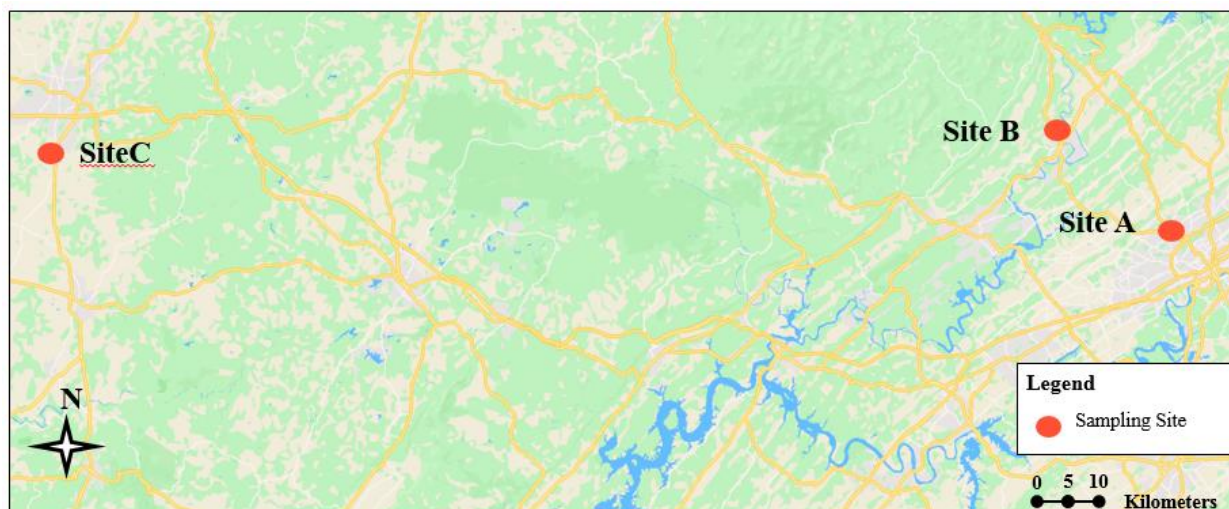


Figure 7.1. The map of sampling sites. Three sites, including Site A, Site B, and Site C, were chosen for this study. At each site, three types of water samples, including stormwater runoff, the receiving stream baseflow, and wetflow at the immediate downstream of runoff effluent, were collected.

Stormwater runoff sampling was conducted using Nalgene™ Storm Water Samplers installed prior to the anticipated storms, and the sampler collected a full one-liter grab sample of the first flush stormwater runoff. Meanwhile, the receiving stream wetflow water samples were collected following the same sampling methods of baseflow.

After collection, all samples were transported using a cooler with ice and delivered to the laboratory within 6 hours.

2.2 FIB and turbidity test

FIB and turbidity analyses were generally performed within 8 hours of sample collection. Turbidity was measured by the turbidity meter following the manufacturer instrument (Thermo Scientific™ AQ3010). *E. coli* and total coliforms were determined by growth on the selective and differential membrane filtration m-Colibblue24® broth (Hach Co., Loveland, CO). The filters were incubated at 37 °C and growth was observed at 24 h.

2.3 DNA extraction and high-throughput 16S rRNA amplicon sequencing

DNA from the colony isolates was extracted by the FastDNA Spin Kit for Soil (MP Biomedicals) following the manufacturer's protocol.

High-Throughput 16S rRNA Amplicon Sequencing was performed as previously described (Wyckoff, Chen, et al., 2017). The PCR amplification used primers 515-F and 806-R to target the V4 region of the 16S rRNA gene (Caporaso, Lauber, et al., 2012). PCR products were purified using the Wizard SV Gel and PCR Clean-Up System (Promega), and paired-end sequencing was completed using the Illumina MiSeq platform.

2.4 Bioinformatic and statistical analyses.

Analysis of sequence reads was conducted using QIIME 2 (Bolyen, Rideout, et al., 2019). Amplicon sequence variants (ASVs) enumeration was performed using the DADA2 pipeline (Callahan, McMurdie, et al., 2016). In contrast to the traditional cluster-based methods based on 97% sequence identity, DADA2 is an error model that can differentiate ASVs based on as little as a single nucleotide. This method is preferred for this study because the 16S V4 region amplicon is short, and some TC species differ at only one position. If traditional 97% identification methods were used, these taxa may be mistakenly grouped together.

Statistical analysis and data visualization were carried out using R version 3.50 (Team 2017). Bray–Curtis (abundance-based) dissimilarity was calculated to represent the taxonomic β -diversity using the vegan R package.

Sourcetracker was used to determine the relative importance of stormwater runoff versus receiving stream baseflow in contributing the microbes in wetflow (Knights, Kuczynski, et al., 2011). Wetflow water samples were considered as sinks, while the receiving stream baseflow and nearby stormwater runoff were considered as sources of microbes.

3. Results and Discussions

3.1 FIB and Turbidity in stormwater runoff and receiving stream water

FIB (*E. coli* and total coliforms) and Turbidity results across the three sites were shown in figure 7.2, and the pairwise t-test results were shown in table 7.S1. FIB and Turbidity in receiving stream baseflow water samples were significantly lower than runoff and wetflow water samples ($P < 0.01$), although they show considerable variations in each type of water samples. Numerous researches also reported FIB and turbidity were significantly higher in impaired streams after storm events (Makepeace, Smith, et al., 1995). There is no doubt that storm events highly impacted the levels of FIB and Turbidity for the receiving streams in this study.

The average of turbidities was 1.4 NTU in baseflow, 9.6 NTU in wetflow, and 28.4 NTU in runoff across the three streams. Not only turbidities in receiving stream baseflow were significantly lower than wetflow, but also turbidities in wetflow were significantly lower than runoff at Site A and Site B.

There is no significant difference of FIB between stormwater runoff and receiving stream wetflow water samples across the three sites. The ranges of *E. coli* concentrations were huge, from 10 to 2400 CFU/100mL in baseflow, from 240 to 120000 CFU/100mL in wetflow, and from 60 to 180000 CFU/100mL in Runoff. Meanwhile, the ranges of total coliforms concentrations were also huge, from 70 to 12000 CFU/100mL in baseflow, from 380 to 210000 CFU/100mL in wetflow, and from 850 to 240000 CFU/100mL in Runoff. Although the variations of FIB concentrations were high, the pairwise t-test showed significantly lower concentrations in baseflow, which means the concentrations of FIB were consistently higher in the wetflow and runoff after the storm events compared to the prior baseflow sampling points.

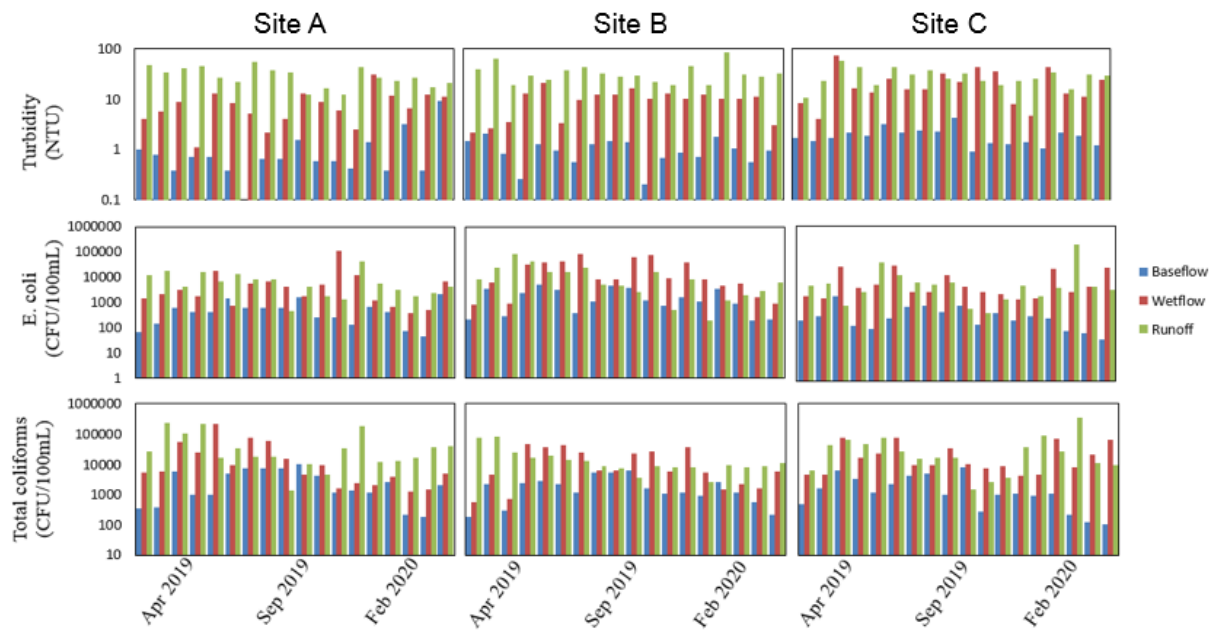


Figure 7.2. FIB and Turbidity in stormwater runoff and receiving stream water across the three study sites.

3.2 Bacterial community structure in the receiving stream baseflow, wetflow, and stormwater runoff

Non-metric Multi-dimensional Scaling analysis (NMDS) showed a clear separation of stormwater runoff and receiving stream baseflow, while samples from receiving stream wetflow placed in the middle of baseflow and runoff (Figure 7.1A). Meanwhile, both stormwater runoff and stream wetflow water from different sampling sites clustered together, but samples of baseflow from Site C separated from Site A and Site B. The dissimilarity among baseflow samples may be resulted in the geographical distances, and Site C is far away from Site A and Site B; The geographical distance between A and B sampling site is 37 km, while the distance between A and C is 238 km, and the distance between B and C is 249 km. Multiple studies reported the significant distance decay relationships of bacterial communities in freshwater ecosystems (Tolkkinen, Mykrä, et al., 2016; Vilmi, Karjalainen, et al., 2016; Hosen, Febria, et al., 2017; Teachey, McDonald, et al., 2019).

Bray-Curtis dissimilarity analyses revealed that the dissimilarities among water types were significantly higher than within each water type ($P < 0.001$). Meanwhile, the dissimilarities between stormwater runoff and receiving stream baseflow were highest at each site. However, the dissimilarities between baseflow and wetflow were not found significant difference compared to the dissimilarities between wetflow and runoff at each location. The results illustrated both stormwater runoff and receiving stream baseflow influenced the wetflow in this study.

The relative abundances of top taxonomies at the genus level among different water types also showed similar patterns with the Beta-diversity of the microbial community (Figure 7.4). The dominant genera like *Massilia*, *Acinetobacter*, and *Pseudomonas* were rarely found in the receiving stream baseflow, but they were increased in the wetflow. Meanwhile, the dominant genera in runoff were abundant across the three sample sites. However, the dominant genera in baseflow showed high variant across the three streams. For example, the mean relative abundance of *Pseudarcicella* was 13.4% in stream C, but 1.2% and 0.1% in stream A and stream B, respectively. Moreover, the dominant genera in both runoff and receiving stream baseflow were all found in wetflow.

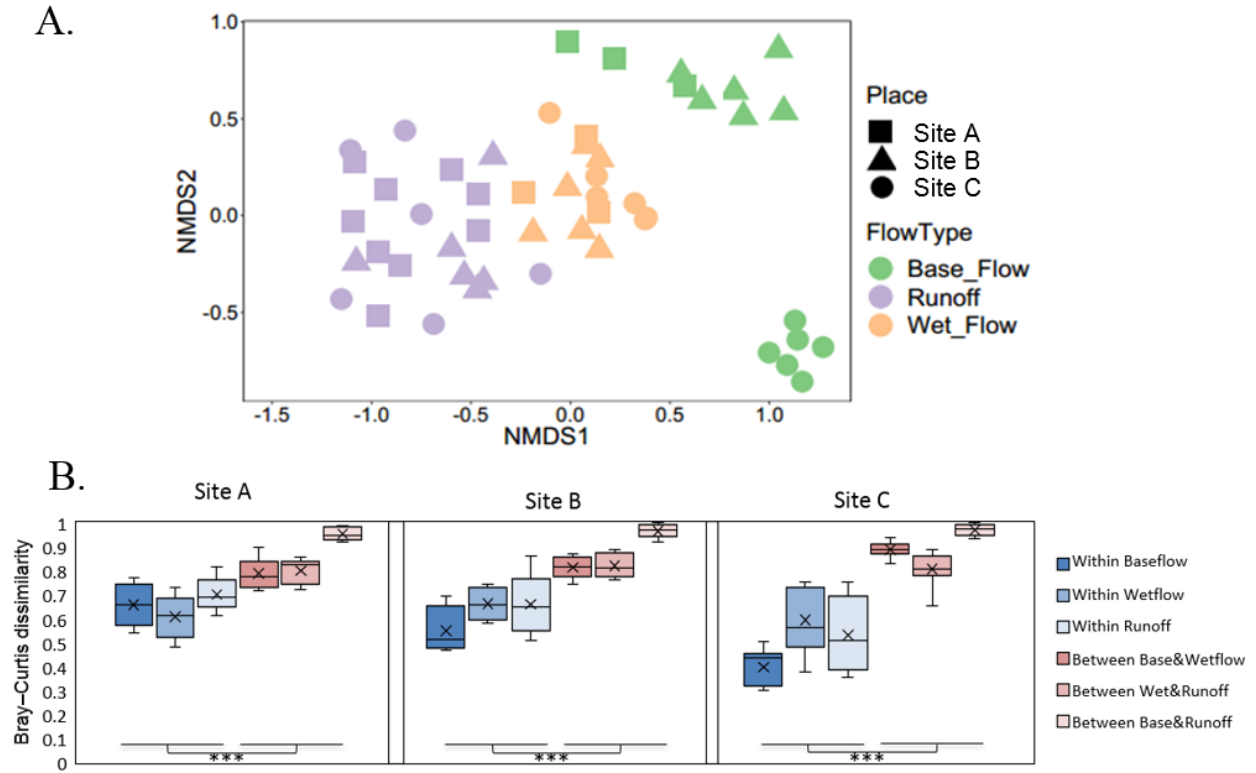


Figure 7.3. Beta-diversity of the microbial community in this study. (A) Non-metric Multi-dimensional Scaling analysis (NMDS) of samples based on Bray–Curtis distance from the microbe cluster by water types(color) and sampling sites; (B) Bray–Curtis dissimilarities within water types and among water types at each site. Two-group comparisons were made by Permutation T-Test. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

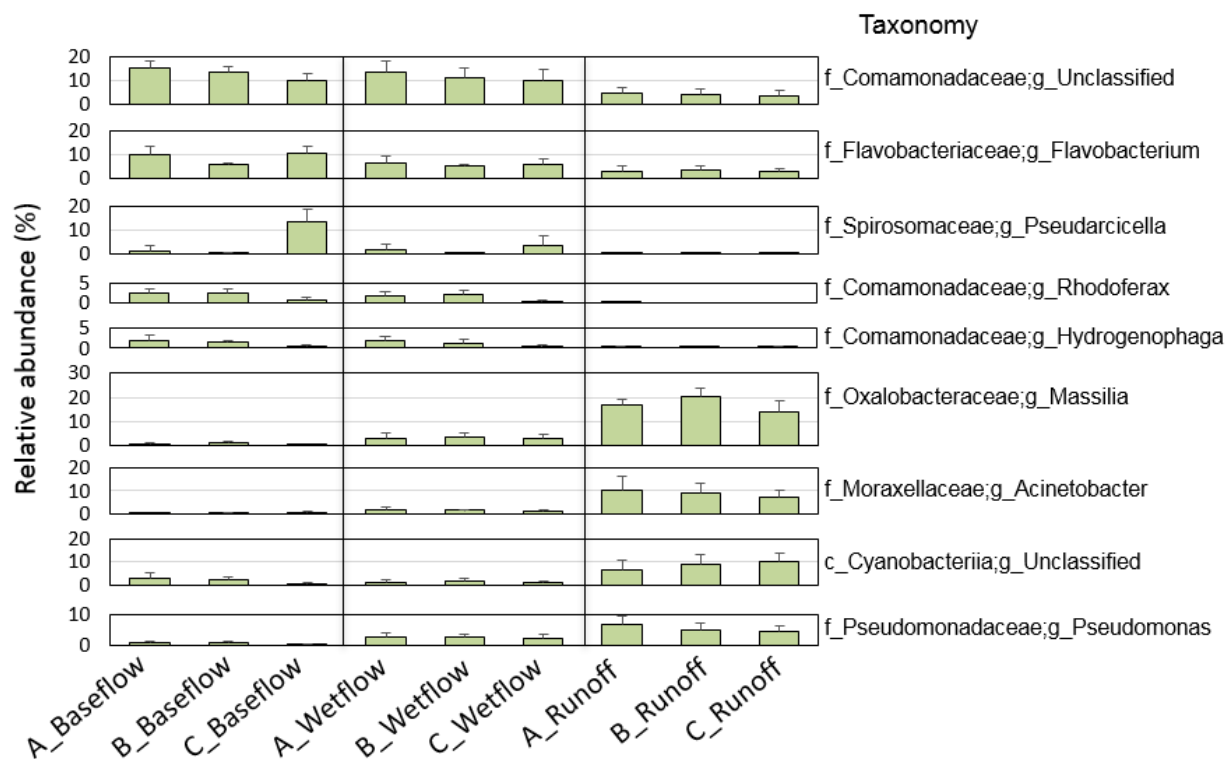


Figure 7.4. Relative abundances of top taxonomies at the genus level were found in different water types at each site.

3.3 The abundances of individual ASVs in the receiving stream baseflow, wetflow, and stormwater runoff

Figure 7.5 showed the top 10 ASVs in the receiving stream baseflow compared to wetflow, runoff, and nearby unimpaired reference streams. As expected, most of the ASVs were not found in stormwater runoff, or the relative abundances of them were much smaller in runoff than baseflow. However, a few ASVs, like ASV17 *Comanomadaceae*, ASV12 *Hydrogenophaga*, ASV22 *Comanomadaceae*, and ASV28 *Aeromonas*, were found similar relative abundances in baseflow and runoff. These ASVs were also presented in nearby unimpaired reference streams, which means they may exist in different habitats in this region. The results denied the hypothesis that these ASVs in baseflow were introduced from runoff and then persisted in the streams.

All top 10 ASVs in baseflow at each site were also found in wetflow, and most of them showed similar relative abundances across dry and wet weather conditions. Meanwhile, most of the top 10 ASVs in the receiving stream wetflow revealed similar abundances in baseflow (Figure 7.6). ASV505 *Massilia* and ASV2441 *Massilia* are the only two ASVs were abundant in Wetflow and runoff, while low relative abundances in baseflow. The results illustrated the impacts of storm events for the top 10 populations in stream water were limited.

However, the dominant populations in stormwater runoff dramatically increased in stream water wetflow, while they were rare in baseflow (Figure 7.7). For example, ASV5174 *Massilia* as one of the top 10 ASVs in runoff at both Clinton and Cook, whose mean relative abundance was 2.2% and 0.75% in Site B runoff and Site C runoff, respectively; meanwhile, the mean relative abundance of ASV5174 *Massilia* was 0.48% and 0.11% in Clinton wetflow and Cook wetflow respectively, but this taxon did not exist in baseflow and reference stream water samples.

All top 10 ASVs in runoff at each site were also found in wetflow, even though some of them did not exist in baseflow. ASV1037 *Rheinheimera* was the only population with similar relative abundances in baseflow and wetflow, but nonexistent in reference streams. The results demonstrated the stormwater runoff heavily influenced wetflow of receiving streams as exogenous microbial communities, while almost all exogenous microorganisms could not persistent in stream water.

Ternary plots showed denser nearby the triangle sides and thin in the middle of the graphics, which

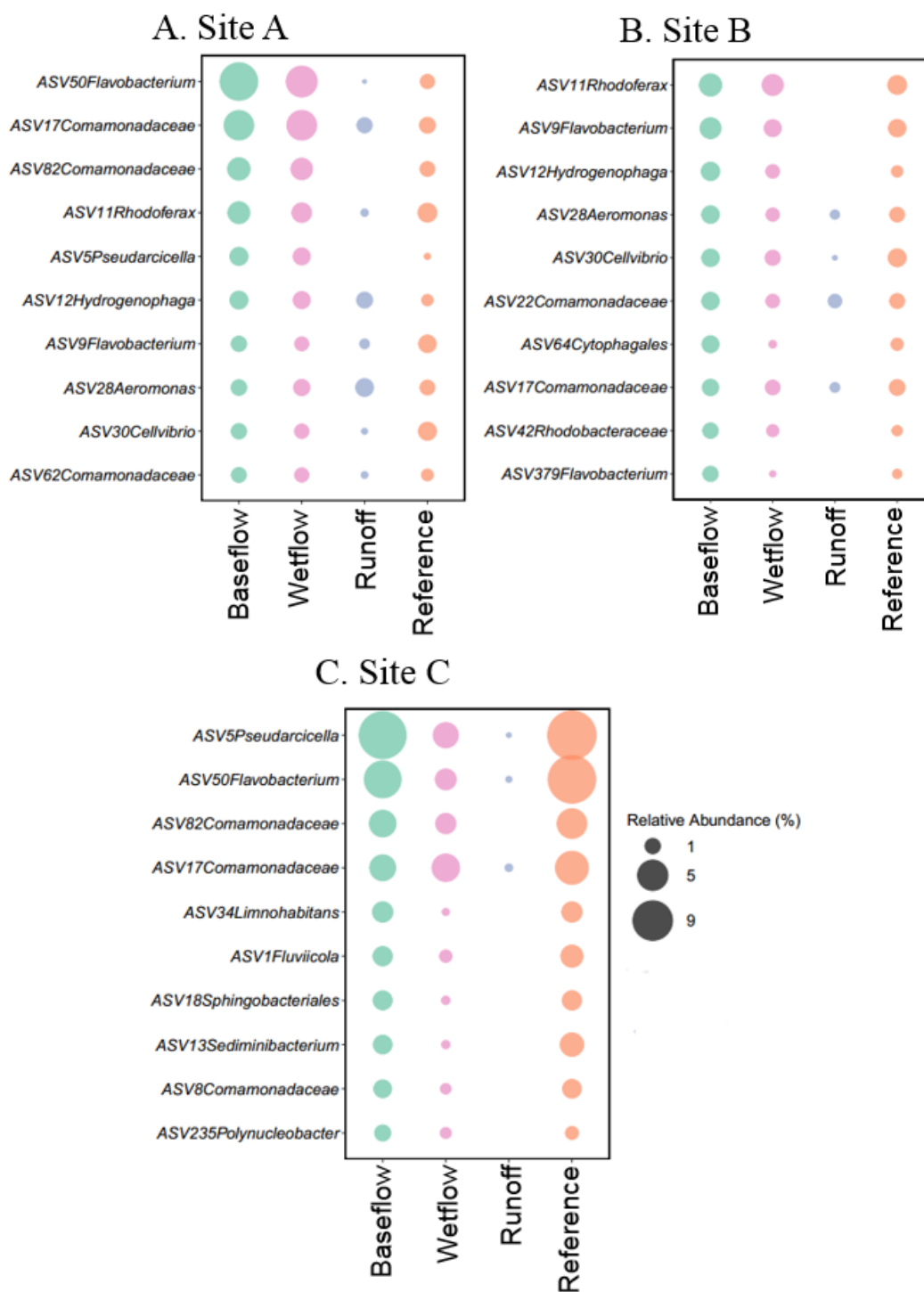


Figure 7.5. Top 10 ASVs in the receiving stream baseflow compared to wetflow, runoff, and nearby unimpaired reference streams at each sampling site. (A) The sampling site A; (B) The sampling site B; (C) The sampling stream C.

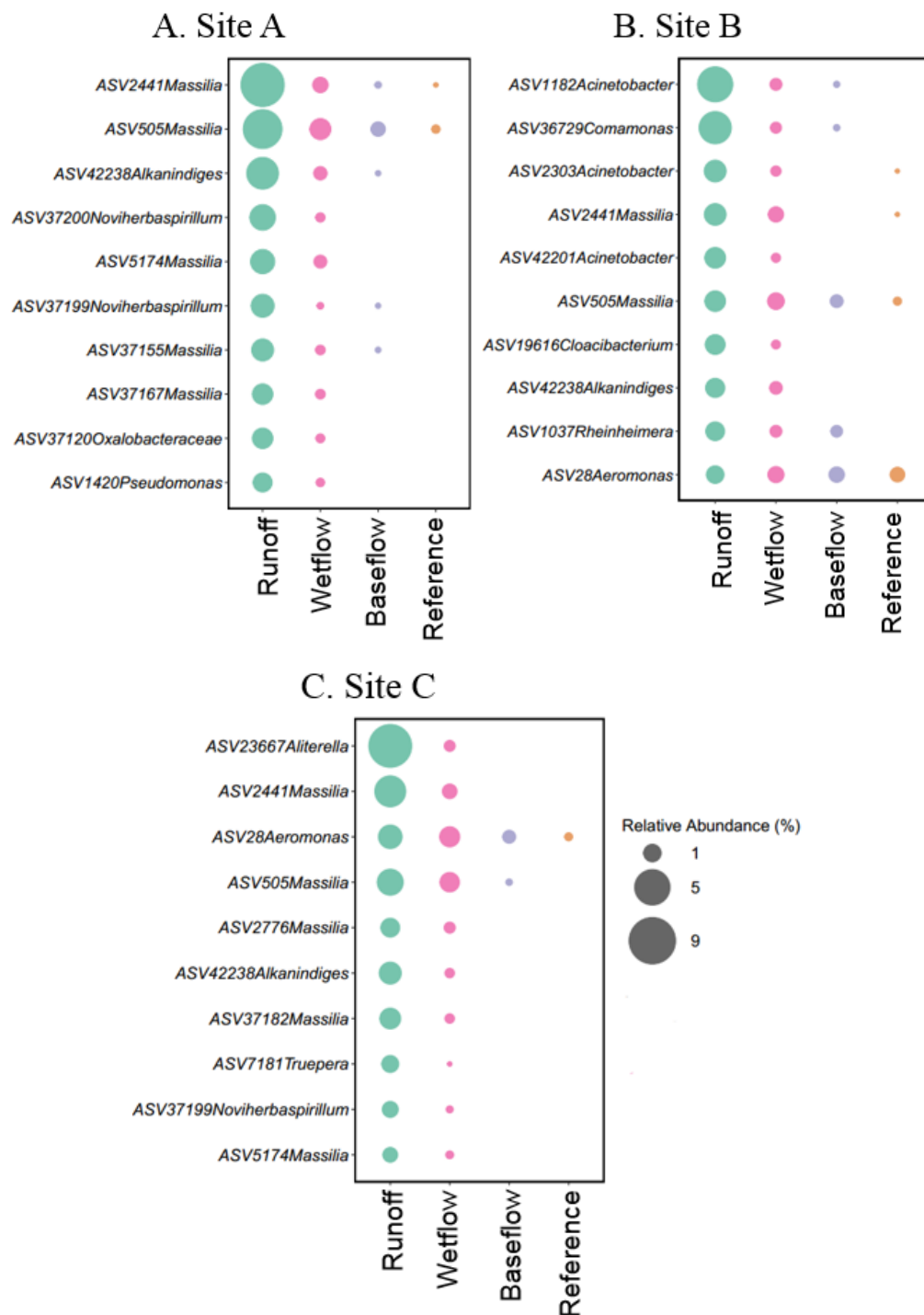


Figure 7.6. Top 10 ASVs in the receiving stream wetflow compared to baseflow, runoff, and nearby unimpaired reference streams at each sampling site. (A) The sampling site A; (B) The sampling site B; (C) The sampling stream C.

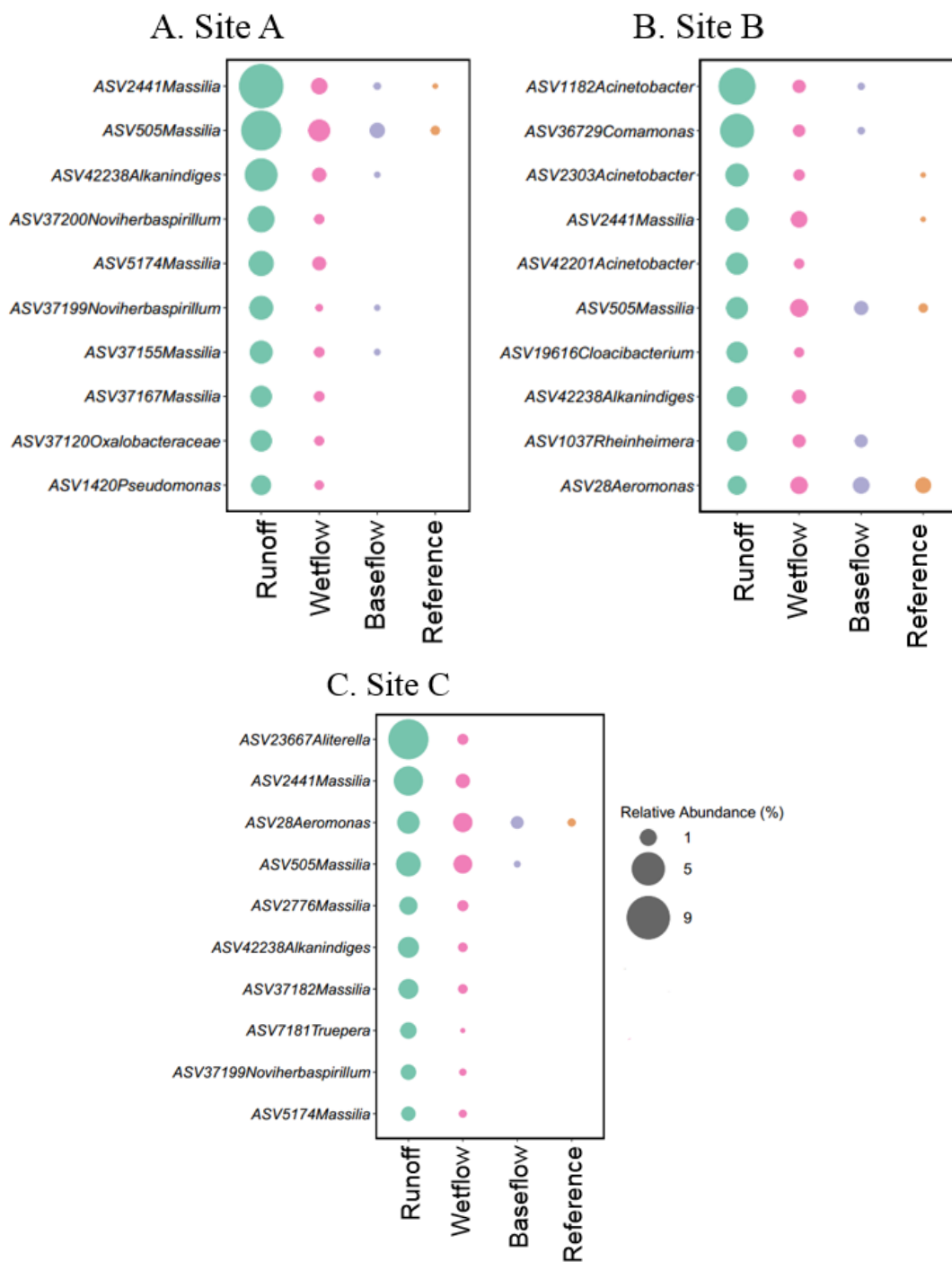


Figure 7.7. Top 10 ASVs in the runoff, compared to the receiving stream baseflow, wetflow, and nearby unimpaired reference streams at each sampling site. (A) The sampling site A; (B) The sampling site B; (C) The sampling stream C.

further illustrated that most of the microbiomes in the receiving stream wetflow samples were either from baseflow or runoff samples (Figure 7.8). Meanwhile, A few dominant ASVs presented across the three types of water, which belonged to *Proteobacteria* mostly.

Interestingly, a few dominant ASVs were found on the triangle points, which means they were uniquely presented in one type of water. Figure 7.9 was further revealed the pattern. ASV440 *Hydrogenophaga* was the only dominant ASV that uniquely present in receiving stream wetflow samples, which illustrated baseflow and runoff are the major contributors for microorganisms in wetflow, and the contributions of other unknown sources were limited.

There are 19 dominant ASVs were unique in stormwater runoff, and 6 dominant ASVs were unique in baseflow; meanwhile, the unique ASVs that were in runoff or baseflow showed high variant across sites and events, and there is no unique ASV that was found across all sites and events in each type of water. These results revealed that both receiving stream baseflow and stormwater runoff existed some of the dominant ASVs that showed susceptibility and could not be detected after receiving runoff.

Moreover, the evenness of microbial structure in receiving stream wetflow was significantly higher than baseflow and runoff ($P < 0.001$) (Figure 7.10), which is reasonable due to the dilution process in both baseflow and runoff. However, the richness in wetflow did not significantly increase compared to baseflow, even though wetflow received exogenous microbes from runoff. This phenomenon further explained that some susceptible ASVs in stream water could not be detected after receiving runoff.

3.4 Sourcetracker analysis

Sourcetracker, a Bayesian method, was used to determine putative source environments of microbial communities in receiving stream wetflow. Wetflow water samples were considered as sinks, while baseflow and stormwater runoff were regarded as sources of microbes (Figure 7.11). Based on the results, baseflow and stormwater runoff were the major sources that contributed to the microbes in wetflow, and they are significantly higher than other unknown sources. The baseflow with the average values of the relative contributions were higher than 49%, and the microbes in runoff accounted for higher than 24% in wetflow. In comparison, the mean contributions of the unknown sources were less than 16% across stream sites.

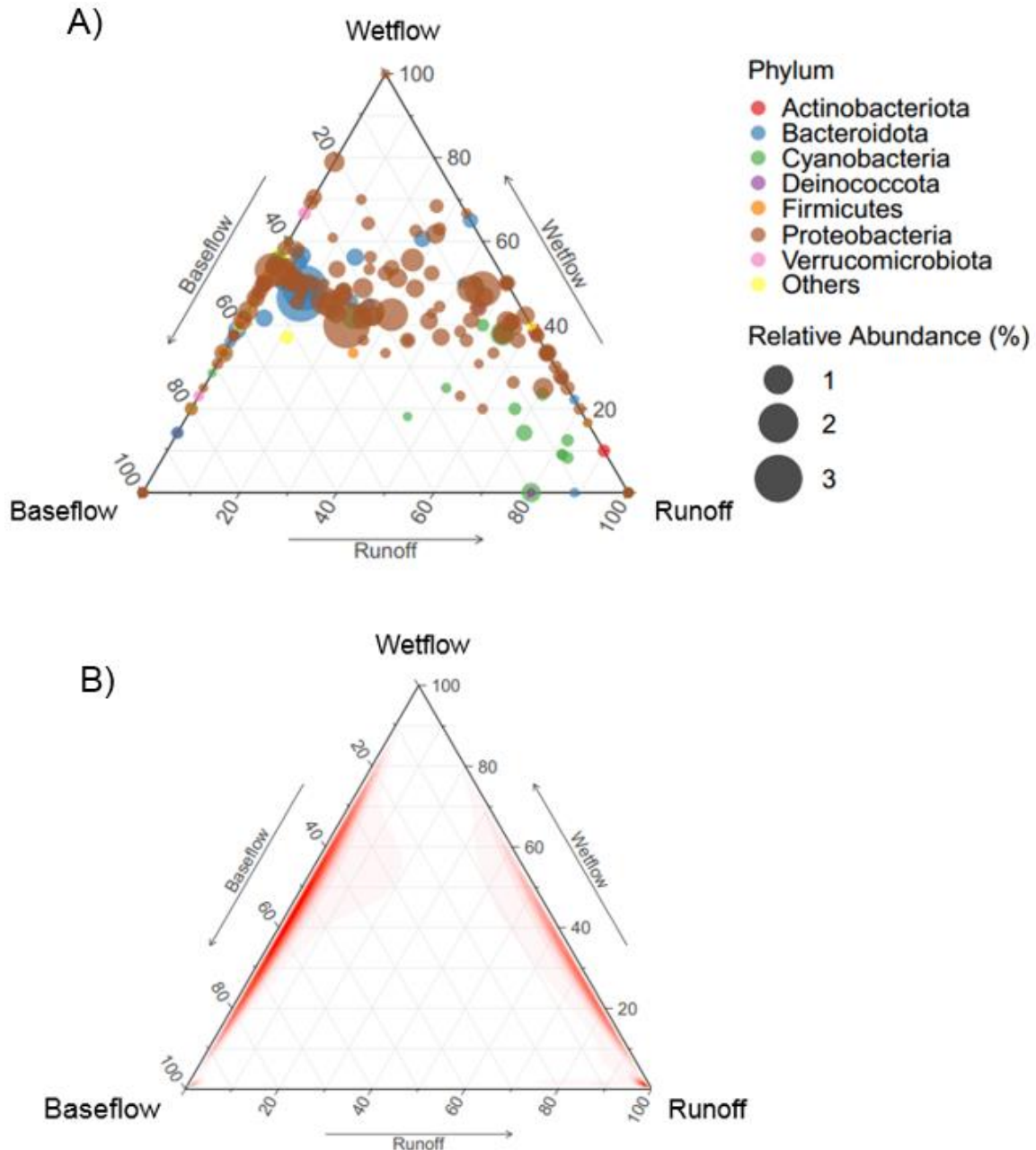


Figure 7.8. Ternary plots of dominant ASVs across the three types of water. (A) Ternary plots showing the distributions of dominant ASVs (> 1% in at least one sample, the total number is 219) among baseflow, wetflow, and runoff. (B) The density pattern of ternary plots.

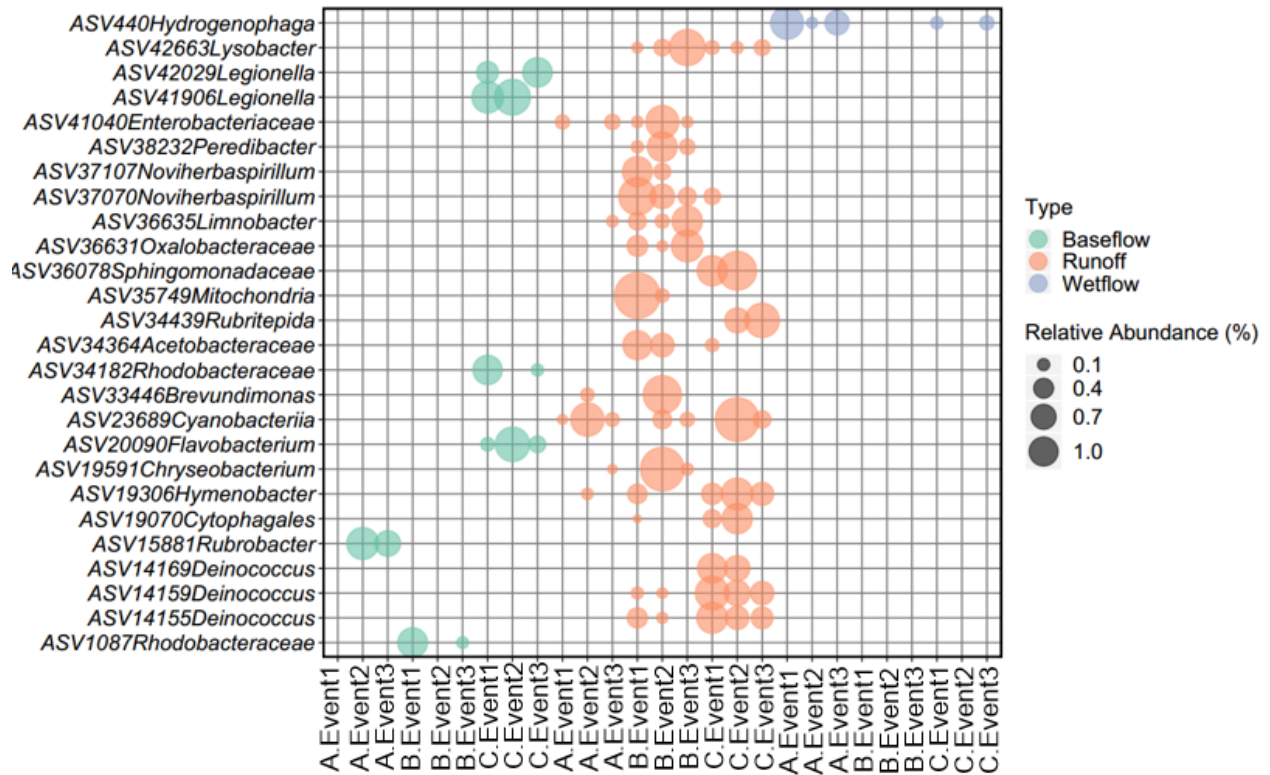


Figure 7.9. The unique dominant ASVs (> 1% in at least one sample) in each type of water.

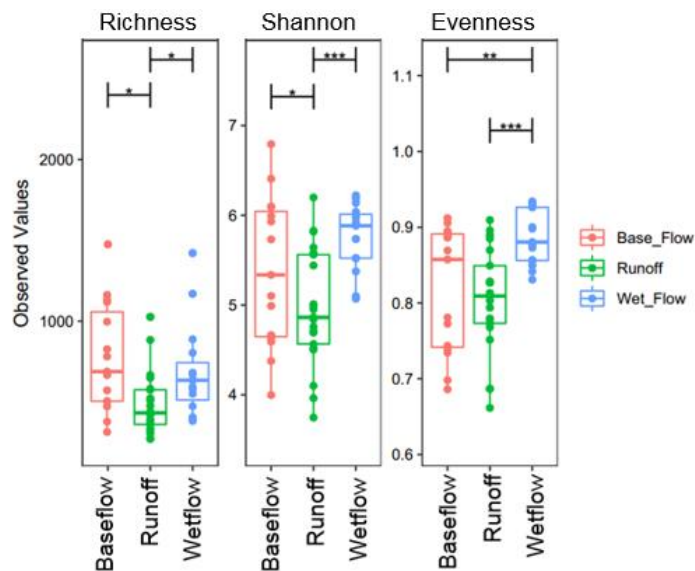


Figure 7.10. Compared alpha diversity (Richness, Shannon index, and Evenness) of microbial structures among three types of water samples.

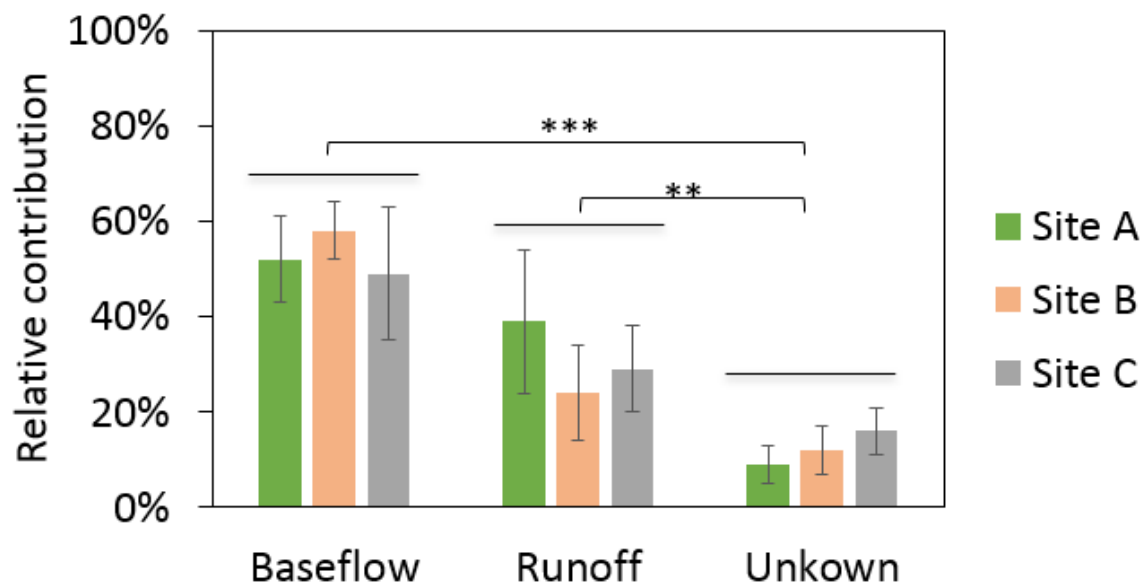


Figure 7.11. Sourcetracker analysis showing the relative contributions of stormwater runoff and receiving stream baseflow to the overall microbial community of wetflow water samples at each site. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

Although the methods identified that the increased microorganisms in wetflow were sourced from the stormwater runoff, the relative contributions showed high variations across stream sites and sampling events. The range of baseflow contributions was from 32% to 66%, and the range of stormwater runoff contributions was from 14% to 59%. Although the means of the relative contributions of stormwater runoff were less than baseflow across the three sites, there is no significant difference due to high variations. The contributions of runoff associated stream features and weather conditions need to be determined based on more sampling efforts in the future.

4. Conclusions

Turbidity, FIB, and Bacterial communities in the receiving stream baseflow, wetflow, and stormwater runoff were studied at three sites. The high Turbidity and FIB levels in the wetflow of receiving creeks likely resulted from nonpoint runoff sources. The dominated taxa in baseflow water samples were typical freshwater taxa that were also found in nearby unimpaired reference streams, but the runoff-associated microorganisms that were rare in baseflow dramatically increased in wetflow. The Sourcetracker analysis also revealed the stormwater runoff and the receiving stream baseflow as the major sources for the bacteria in each receiving stream after storm events. Overall, findings in this study could be helpful for developing stormwater runoff control measures. In the future, more sites and storm events should be studied in order to figure out the contributions of runoff for microbial contamination associated with stream features and weather conditions.

References

- Ahmed, W., K. A. Hamilton, A. Lobos, B. Hughes, C. Staley, M. J. Sadowsky and V. J. Harwood (2018). "Quantitative microbial risk assessment of microbial source tracking markers in recreational water contaminated with fresh untreated and secondary treated sewage." Environment International **117**: 243-249.
- Ashbolt, N. J. (2015). "Microbial Contamination of Drinking Water and Human Health from Community Water Systems." Current environmental health reports **2**(1): 95-106.
- Ashbolt, N. J. and M. Bruno (2003). "Application and refinement of the WHO risk framework for recreational waters in Sydney, Australia." Journal of Water and Health **1**(3): 125-131.
- Boehm, A. B., K. E. Graham and W. C. Jennings (2018). "Can We Swim Yet? Systematic Review, Meta-Analysis, and Risk Assessment of Aging Sewage in Surface Waters." Environmental Science and Technology **52**(17): 9634-9645.
- Bolyen, E., J. R. Rideout, M. R. Dillon, N. A. Bokulich, C. C. Abnet, G. A. Al-Ghalith, H. Alexander, E. J. Alm, M. Arumugam and F. Asnicar (2019). "Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2." Nature biotechnology **37**(8): 852-857.
- Callahan, B. J., P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson and S. P. Holmes (2016). "DADA2: high-resolution sample inference from Illumina amplicon data." Nature methods **13**(7): 581.
- Caporaso, J. G., C. L. Lauber, W. A. Walters, D. Berg-Lyons, J. Huntley, N. Fierer, S. M. Owens, J. Betley, L. Fraser and M. Bauer (2012). "Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms." The ISME journal **6**(8): 1621.
- Claesson, M. J., Q. Wang, O. O'Sullivan, R. Greene-Diniz, J. R. Cole, R. P. Ross and P. W. O'Toole (2010). "Comparison of two next-generation sequencing technologies for resolving highly complex microbiota composition using tandem variable 16S rRNA gene regions." Nucleic Acids Research **38**(22).
- Eren, A. M., H. G. Morrison, P. J. Lescault, J. Reveillaud, J. H. Vineis and M. L. Sogin (2015). "Minimum entropy decomposition: Unsupervised oligotyping for sensitive partitioning of high-throughput marker gene sequences." ISME Journal **9**: 968-979.
- Federigi, I., M. Verani, G. Donzelli, L. Cioni and A. Carducci (2019). "The application of quantitative microbial risk assessment to natural recreational waters: A review." Marine Pollution Bulletin **144**: 334-350.
- Hamilton, K. A., B. Reyneke, M. Waso, T. Clements, T. Ndlovu, W. Khan, K. GiGiovanni, E. Rakestraw, F. Montalto, C. N. Haas and W. Ahmed (2019). "A global review of the microbiological quality and potential health risks associated with roof-harvested rainwater tanks." NPJ Clean Water **2**(7).
- Hosen, J. D., C. M. Febria, B. C. Crump and M. A. Palmer (2017). "Watershed urbanization linked to differences in stream bacterial community composition." Frontiers in Microbiology **8**(AUG).
- Jiang, S., R. Noble and W. Chu (2001). "Human adenoviruses and coliphages in urban runoff-impacted coastal waters of Southern California." Applied and Environmental Microbiology **67**(1): 179-184.

- Kennedy, K., M. W. Hall, M. D. J. Lynch, G. Moreno-Hagelsieb and J. D. Neufeld (2014). "Evaluating bias of Illumina-based bacterial 16S rRNA gene profiles." Applied and Environmental Microbiology **80**(18): 5717-5722.
- Kim, K. H. and J. W. Bae (2011). "Amplification methods bias metagenomic libraries of uncultured single-stranded and double-stranded DNA viruses." Applied and Environmental Microbiology **77**(21): 7663-7668.
- Knights, D., J. Kuczynski, E. S. Charlson, J. Zaneveld, M. C. Mozer, R. G. Collman, F. D. Bushman, R. Knight and S. T. Kelley (2011). "Bayesian community-wide culture-independent microbial source tracking." Nature Methods **8**(9): 761-763.
- Lazarevic, V., K. Whiteson, S. Huse, D. Hernandez, L. Farinelli, M. Østerås, J. Schrenzel and P. François (2009). "Metagenomic study of the oral microbiota by Illumina high-throughput sequencing." Journal of Microbiological Methods **79**(3): 266-271.
- Liu, Z., C. Lozupone, M. Hamady, F. D. Bushman and R. Knight (2007). "Short pyrosequencing reads suffice for accurate microbial community analysis." Nucleic Acids Research **35**(18).
- Logares, R., S. Sunagawa, G. Salazar, F. M. Cornejo-Castillo, I. Ferrera, H. Sarmento, P. Hingamp, H. Ogata, C. de Vargas, G. Lima-Mendez, J. Raes, J. Poulain, O. Jaillon, P. Wincker, S. Kandels-Lewis, E. Karsenti, P. Bork and S. G. Acinas (2014). "Metagenomic 16S rDNA Illumina tags are a powerful alternative to amplicon sequencing to explore diversity and structure of microbial communities." Environmental Microbiology **16**(9): 2659-2671.
- Lurry, D. L. and C. M. Kolbe (2000). Interagency field manual for the collection of water-quality data, US Department of the Interior, US Geological Survey.
- Madoui, M. A., S. Engelen, C. Cruaud, C. Belser, L. Bertrand, A. Alberti, A. Lemainque, P. Wincker and J. M. Aury (2015). "Genome assembly using Nanopore-guided long and error-free DNA reads." BMC Genomics **16**(1).
- Makepeace, D. K., D. W. Smith and S. J. Stanley (1995). "Urban Stormwater Quality: Summary of Contaminant Data." Critical Reviews in Environmental Science and Technology **25**(2): 93-139.
- McLellan, S. L., S. M. Huse, S. R. Mueller-Spitz, E. N. Andreishcheva and M. L. Sogin (2010). "Diversity and population structure of sewage-derived microorganisms in wastewater treatment plant influent." Environmental Microbiology **12**(2): 378-392.
- Nshimiyimana, J. P., A. J. E. Freedman, P. Shanahan, L. C. H. Chua and J. R. Thompson (2015). ""Variation of Bacterial Communities and Pathogen Taxa as a Function of Land Use and Water Quality in an Urban Tropical Catchment of Singapore"." Proceedings of the 115th General Meeting of American Society for Microbiology.
- Saget, A., G. Chebbo and J. L. Bertrand-Krajewski (1996). "The first flush in sewer systems." Water Science and Technology **33**(9): 101-108.
- Schloss, P. D., S. L. Westcott, M. L. Jenior and S. K. Highlander (2015). "Sequencing 16S rRNA gene fragments using the PacBio SMRT DNA sequencing system." PeerJ PrePrints **3**.
- Shendure, J., G. J. Porreca, N. B. Reppas, X. Lin, J. P. McCutcheon, A. M. Rosenbaum, M. D. Wang, K. Zhang, R. D. Mitra and G. M. Church (2005). "Molecular biology: Accurate multiplex polony sequencing of an evolved bacterial genome." Science **309**(5741): 1728-1732.

- Sinclair, L., O. A. Osman, S. Bertilsson and A. Eiler (2015). "Microbial community composition and diversity via 16S rRNA gene amplicons: Evaluating the illumina platform." PLoS ONE **10**(2).
- Teachey, M. E., J. M. McDonald and E. A. Ottesen (2019). "Rapid and stable microbial community assembly in the headwaters of a third-order stream." Applied and Environmental Microbiology **85**(11).
- Team, R. C. (2017). "R Foundation for Statistical Computing; Vienna, Austria: 2016. R: A language and environment for statistical computing." URL <http://www.R-project.org/>[Google Scholar].
- Thurston, J. A., C. P. Gerba, K. E. Foster and M. M. Karpiscak (2001). "Fate of indicator microorganisms, giardia and cryptosporidium in subsurface flow constructed wetlands." Water Research **35**(6): 1547-1551.
- Tolkkinen, M. J., H. Mykrä, R. Virtanen, M. Tolkkinen, T. Kauppila, L. Paasivirta and T. Muotka (2016). "Land use impacts on stream community composition and concordance along a natural stress gradient." Ecological Indicators **62**: 14-21.
- USEPA (1986). Ambient water quality criteria for bacteria-1986.
- Vilmi, A., S. M. Karjalainen, T. Nokela, K. Tolonen and J. Heino (2016). "Unravelling the drivers of aquatic communities using disparate organismal groups and different taxonomic levels." Ecological Indicators **60**: 108-118.
- Wang, Q., G. M. Garrity, J. M. Tiedje and J. R. Cole (2007). "Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy." Applied and Environmental Microbiology **73**(16): 5261-5267.
- Wyckoff, K. N., S. Chen, A. J. Steinman and Q. He (2017). "Impact of Roadway Stormwater Runoff on Microbial Contamination in the Receiving Stream." Journal of environmental quality **46**(5): 1065-1071.

Conclusion

As a significant source of microbial pollution, stormwater runoff is mainly discharged into receiving waters and may be responsible for serious ecological impacts. For this dissertation, Culture-based enumeration of FIB and sequencing-based microbial community served as the microbial analysis for stormwater runoff and streams.

With the widespread use of FIB, the efficacy of culture-based coliform testing was assessed. A variety of environmental matrices, including stream water, stormwater runoff, sewage, and stream sediments, were used to explore the performance of the coliform testing method through a comparison with high-throughput 16S rRNA amplicon sequencing. The results suggested that this culture-based method was generally accurate at correctly identifying coliforms in the environmental samples in this study.

With the efficacy of FIB culture-based detection being validated, this method was used to monitor stormwater runoff, including pavement runoff and culvert runoff, over 23 storm events at 6 sites. It was revealed that while there was a moderate relationship between storm characteristics and FIB levels in the runoff, the concentrations of coliforms were more related to roadway characteristics and seasonal impacts.

This study also evaluated the survival of *E. coli* strains isolated from stormwater runoff under environmentally relevant conditions. Findings indicate that the wild-type strains isolated from runoff persisted for a more extended period than the lab strain (MG1655) under all environmental conditions tested, but even for wild-type strains, the counts were below the detection limit within a few hours under high temperature or sunlight exposure conditions. Moreover, the persistence of *E. coli* in receiving streams may mislead the recent fecal contamination, especially in winter due to the prolonged persistence at a lower temperature.

With the concern that FIB might not actually be an indication of fecal contamination, the microbial water quality was assessed utilizing 16S rRNA amplicon sequencing for a more in-depth microbial community analysis. The understanding of natural variations of the bacterioplankton communities in the streams where the absence of anthropogenic sources is needed before assessing the potential applicability of microbial community as assessments of stream ecological conditions. It was revealed that while there was spatial and temporal variation observed, the stream sizes and characteristics were larger drivers for stream microbial community.

This study also investigated microbial community composition-based methods for assessment of the impacts of stormwater runoff to receiving streams where the impacts of stormwater runoff were highlighted by FIB culture-based approaches. The results showed that the dominated taxa in receiving stream baseflow water samples were typical freshwater taxa that were also found in nearby unimpaired reference streams, and no doubt distinct from the bacterial communities in stormwater runoff. Moreover, the microbes sourced from the stormwater runoff dramatically increased in wetflow. Stormwater runoff and the receiving stream baseflow as the major contributors for the overall microbial community in wetflow were determined across the three study sites. Findings in this study could be useful for developing stormwater runoff control measures.

Appendix

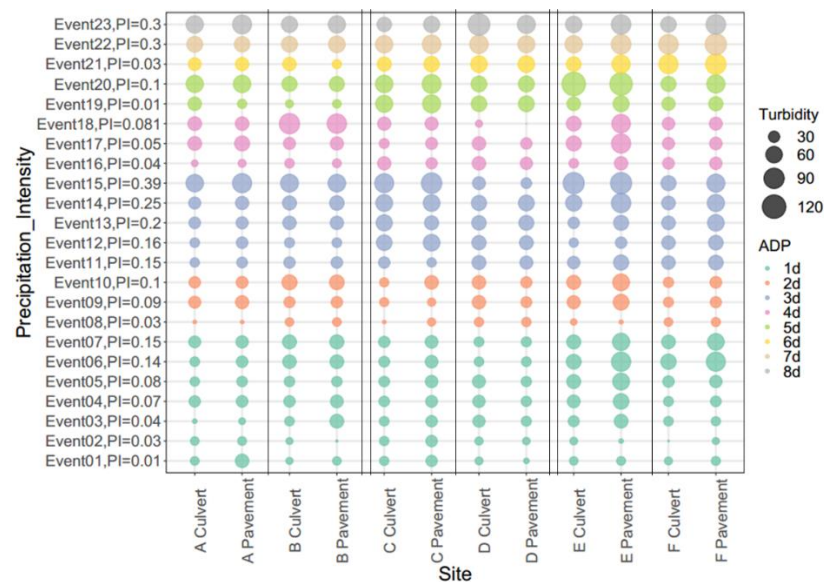


Figure 3.S1. Turbidity in runoff ranked by PI and ADP

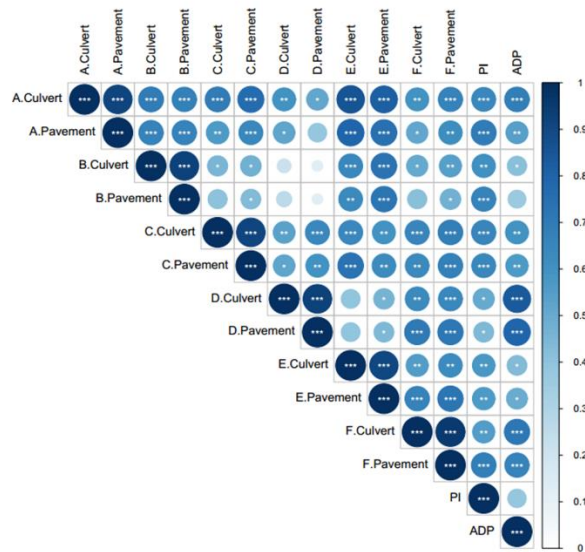


Figure 3.S2. Turbidity distribution and Correlation with rainfall characteristics

Table 5.S1. Comparison of the performance of SAD models for the bacterioplankton communities at each site.

The model types	R-squared					
	Stream I Fall	Stream I Spring	Stream II Fall	Stream II Spring	Stream III Fall	Stream III Spring
Lognormal	0.914	0.934	0.968	0.967	0.946	0.991
Log-series	0.785	0.675	0.522	0.506	0.628	0.786
Zipf	0.238	0.205	0.155	0.131	0.302	0.379

Table 6.S1. Stream Features.

Ecoregion	Stream ID	Stream Name	Stream order	Geographical distance from headwater (km)
67f The Southern Limestone/Dolomite Valleys and Low Rolling Hills	A	White Creek	1	0.5
	B	Clear Creek	2	1.1
	C	Powell River	4	75.8
	D	Big Barren Creek	2	6.2
	E	Blairs Creek	1	1.5
68a Cumberland Plateau	F	Island Creek	2	2.8
	G	Daddy's Creek	3	16.2
	H	Crab Orchard Creek	3	17.6
	I	Clear Creek	3	41.6
	J	Emory River	4	54.4
71g Eastern Highland Rim	K	Hurricane Creek	1	0.5
	L	Post Oak Creek	1	5.8
	M	Spring Creek	3	14.4
	N	Roaring River	3	16.2
	O	Flat Creek	3	11.2

Table 7.S1. Statistical analysis of flow type differences in water quality (P values based on pairwise T-test)

Parameters	Site A			Site B			Site C		
	Base-Wet	Base-Runoff	Wet-Runoff	Base-Wet	Base-Runoff	Wet-Runoff	Base-Wet	Base-Runoff	Wet-Runoff
Turbidity	6.91×10^{-9}	2.57×10^{-16}	7.56×10^{-13}	5.60×10^{-9}	3.83×10^{-15}	1.63×10^{-12}	3.43×10^{-4}	5.09×10^{-9}	7.66×10^{-2}
<i>E. coli</i>	2.61×10^{-3}	6.53×10^{-4}	3.26×10^{-1}	5.44×10^{-5}	1.70×10^{-5}	4.22×10^{-1}	1.21×10^{-3}	7.21×10^{-3}	2.47×10^{-1}
Total coliforms	2.45×10^{-3}	8.99×10^{-4}	1.05×10^{-1}	2.51×10^{-5}	1.14×10^{-5}	1.17×10^{-1}	9.35×10^{-4}	1.46×10^{-3}	1.50×10^{-1}

VITA

Songyi Liu was born in August 1990 in Ganzhou, China. He received his B. S. in Environmental Engineering from Beijing University of Chemical Technology in 2012. After graduation, He worked as a Research Assistant in Beijing DQY Agriculture Technology Co., Ltd. After two years of work experience, he started his master's study in Civil and Environmental Engineering at the Missouri University of Science and Technology. After graduation, he started his Ph.D. study at the University of Tennessee, Knoxville, under the direction of Dr. Qiang He in May 2016. His research is focused on Microbial Source Tracking for Detection of Fecal Contamination in Water Bodies and Differential Decay of Fecal Indicator Bacteria Originating from different Sources.