Transfer of knowledge and implementation of routine analysis of environmental samples for the presence of naturalised *E. coli.*

E/S/R THE SCIENCE BEHIND THE TRUTH

June 2019

Prepared with support from Envirolink, Ministry of Business Innovation and Employment (MBIE)

PREPARED FOR: **Northland Regional Council**

CLIENT REPORT No: CSC 19009

PREPARED BY: **Megan Devane**

REVIEWED BY: **Brent Gilpin**

ACKNOWLEDGEMENTS

ESR would like to thank the staff at Northland Regional Council for their time in coordinating and carrying out the sampling programme for the Northland Rivers. In particular, ESR appreciate the contributions of Pania Te Whaiti and Jean-Charles Perquin in asking the insightful research questions that led to the formation of this Envirolink Report.

Manager	Peer reviewer	Author
Brent J Gilpin	**Brent J Gilpin**	**Megan Devane**
Manager of Water and Biowaste Group	Manager of Water and Biowaste Group	Senior Scientist

DISCLAIMER

The Institute of Environmental Science and Research Limited (ESR) has used all reasonable endeavours to ensure that the information contained in this client report is accurate. However, ESR does not give any express or implied warranty as to the completeness of the information contained in this client report or that it will be suitable for any purposes other than those specifically contemplated during the Project or agreed by ESR and the Client.

CONTENTS

EXECUTIVE SUMMARY 1

1. INTRODUCTION 6

1.1 WORK PLAN: 7

2. METHODOLOGY FOR CHARACTERISATION OF ESCHERICHIA FROM RIVER WATER AND SEDIMENT 9

2.1 CHARACTERISTICS OF SAMPLING SITES 9

2.1.1 Land cover categories for each site 9

2.2 RAINFALL AND WATER QUALITY PARAMETERS 9

2.3 SAMPLE COLLECTION 11

2.4 ENUMERATION OF *ESCHERICHIA* FROM WATER SAMPLES 12

2.5 ISOLATION OF *ESCHERICHIA* FROM WATER SAMPLES 13

2.5.1 Notes on streak isolation procedure to obtain purified colonies of *E. coli* derived from a single *E. coli* strain 15

2.6 ENUMERATION AND ISOLATION OF *ESCHERICHIA* FROM SEDIMENT SAMPLES 16

2.6.1 Notes on *E. coli* enumeration in sediment 17

2.7 DRY WEIGHT DETERMINATION FOR SEDIMENT 18

2.8 CHARACTERISATION OF PUTATIVE *ESCHERICHIA* ISOLATED COLONIES 18

2.9 STATISTICAL ANALYSIS 19

3. RESULTS 20

3.1 RAINFALL AND WATER QUALITY PARAMETERS 20

3.2 *E. COLI* ENUMERATION IN WATER AND SEDIMENT 21

3.2.1 *E. coli* in water 21

3.2.2 *E. coli* in sediment 22

3.2.3 Faecal source tracking analysis of selected samples 22

3.3 CHARACTERISATION OF *ESCHERICHIA* ISOLATES INTO *E. COLI* PHYLOGROUPS OR CRYPTIC CLADES OF *ESCHERICHIA* 23

3.3.1 Characterisation of selected isolates of *E. coli* phylogroups to determine genetic relatedness 26
3.3.2 Waitangi 26
3.3.3 Victoria 26
3.3.4 Hatea 31
3.3.5 Ahuroa 32
3.3.6 Raumanga 32
3.3.7 Identification of cryptic clades in water and sediment 34

4. DISCUSSION 36

5. CONCLUSIONS 42
5.1 RECOMMENDATIONS 43

APPENDIX A: SAMPLING SITES 45
A.1 GPS COORDINATES OF SAMPLING SITES 45
A.2 PHOTOS OF NORTHLAND RIVER SITES SELECTED FOR SEDIMENT AND WATER SAMPLING TO EVALUATE NATURLAISED *ESCHERICHIA* AND *E. COLI* 46

APPENDIX B: METHOD PREPARATION 49
B.1 CHELEX PREPARATION SOLUTION FOR CELL/COLONY PREPARATION FOR PCR ANALYSIS 49

APPENDIX C: RAINFALL DATA AND WATER QUALITY PARAMETERS 50
C.1 RAINFALL DATA FOR 2017 AND 2018 SAMPLING EVENTS 50
C.2 WATER QUALITY PARAMETERS 54

APPENDIX D: *E. COLI* CONCENTRATIONS 55
D.1 *E. COLI* CONCENTRATIONS IN SEDIMENT AND WATER 55
D.2 CHARACTERISATION OF *ESCHERICHIA* ISOLATES INTO *E. COLI* PHYLOGROUPS AND *ESCHERICHIA* CRYPTIC CLADES 57

6. REFERENCES 60

LIST OF TABLES

TABLE 1: NORTHLAND RIVERS SITES SELECTED FOR SAMPLING EVENTS FOR NATURALISED *ESCHERICHIA* AND *E. COLI* 9

TABLE 2: *E. COLI* CONCENTRATIONS IN SEDIMENT AND WATER 2017-2018 AS DETERMINED BY MEMBRANE FILTRATION 22

TABLE 3: 20TH FEBRUARY 2017 SAMPLING; ATTRIBUTION OF *ESCHERICHIA* ISOLATES TO GROUPS AND RECOMMENDATIONS FOR FURTHER SUBTYPING. 24

TABLE 4: 2018 SAMPLING; ATTRIBUTION OF *ESCHERICHIA* ISOLATES TO GROUPS AND RECOMMENDATIONS FOR FURTHER SUBTYPING 25

TABLE 5: REP-PCR RESULTS FOR DETERMINING GENETIC RELATIONSHIPS AND POTENTIAL FOR NATURALISATION WITHIN DOMINANT *E. COLI* PHYLOGROUPS 28

TABLE 6: OVERALL PREVALENCE (%) OF NATURALISED *ESCHERICHIA* CLADE V ISOLATES IN WATER AND SEDIMENT 35

TABLE 7: EXAMPLES WHERE ACTION OR ALERT CONCENTRATIONS OF *E. COLI* WERE EXCEEDED AND THE EFFECT OF THE PREVALENCE OF CRYPTIC CLADES 35

TABLE 8: CUMULATIVE RAINFALL DATA FOR SAMPLING EVENTS IN 2017 AND 2018. 50

TABLE 9: *E. COLI* ENUMERATION IN SEDIMENT AND WATER DURING 2017 55

TABLE 10: *E. COLI* ENUMERATION IN SEDIMENT AND WATER DURING APRIL AND AUGUST 2018. 56

TABLE 11: CHARACTERISATION OF *E. COLI* ISOLATES AND ESCHERICHIA CRYPTIC CLADES FROM SELECTED WATER AND SEDIMENT SAMPLES COLLECTED 19 FEBRUARY 2017. 57

TABLE 12: CHARACTERISATION OF *E. COLI* ISOLATES AND *ESCHERICHIA* CRYPTIC CLADES FROM SELECTED WATER AND SEDIMENT SAMPLES COLLECTED DURING APRIL 2018 58

TABLE 13: CHARACTERISATION OF *E. COLI* ISOLATES AND *ESCHERICHIA* CRYPTIC CLADES FROM SELECTED WATER AND SEDIMENT SAMPLES COLLECTED DURING AUGUST 2018 59

LIST OF FIGURES

FIGURE 1: MAP OF THE NORTHLAND RIVER SITES SELECTED FOR SAMPLING FOR NATURALISED *ESCHERICHIA* AND *E. COLI* 10

FIGURE 2: COLONY ISOLATION OF *E. COLI*. 16

FIGURE 3: SELECTED *E. COLI* ISOLATES FROM PHYLOGROUPS B1 AND B2 IDENTIFIED IN SEDIMENTS AND WATER AT HATEA AND AHUROA 33

FIGURE 4: FEBRUARY 2017 RAINFALL 51

FIGURE 5: RAINFALL DURING MARCH TO APRIL 2018 52

FIGURE 6: RAINFALL DURING JULY TO AUGUST 2018 53

EXECUTIVE SUMMARY

The faecal indicator bacterium, *Escherichia coli*, is used as an indicator of faecal contamination for water quality monitoring purposes in the freshwaters of New Zealand (Aotearoa) (Ministry for the Environment and Ministry of Health 2003). When the *E. coli* are from fresh faecal sources they have strong associations with pathogens and are, therefore, useful indicators of water quality. However, the potential role of naturalised *Escherichia* as confounders to routine water quality testing has been identified in recent years.

There are two different types of naturalised bacteria belonging to the *Escherichia* genus.

- One type of naturalised *Escherichia* belong to the species of *E. coli* and are derived from a mammalian/avian faecal origin. These naturalised *E. coli* are able to survive post-defecation in the environment and reside in reservoirs such as soil and river sediment where they may replicate.
- The other type of naturalised *Escherichia* are identified as *E. coli* using conventional *E. coli* water quality testing methods but in fact belong to separate species within the *Escherichia* genus. They are termed Cryptic Clades of *Escherichia* and are rarely identified in human clinical or mammalian faecal specimens. Typical biochemical tests used to differentiate *E. coli* from other *Escherichia* genera do not separate these naturalised *Escherichia* Cryptic Clades from *E. coli*. This lack of differentiation means these *Escherichia* have the potential to confound water quality testing methods. *Escherichia* cryptic clade isolates have low levels of antibiotic resistance and virulence factors, which together with the low levels identified in clinical samples suggests that they are unlikely to be pathogens, and less likely to be good indicators for pathogens from faecal sources.

Preliminary studies in Northland Rivers have indicated that both types of naturalised *Escherichia* (persistent *E. coli* and cryptic clades) were likely to be contributing to water

quality monitoring results at some river sites. This means standard methods of *E. coli* enumeration may not accurately reflect recent faecal contamination at certain sites. Northland Regional Council (NRC) sought assistance and guidance from ESR to design and implement a project that would help with the identification of naturalised *Escherichia* in Northland Rivers reporting chronically elevated concentrations of *E. coli*. This project involved ESR undertaking a review of NRC data collected during 2017-2018, where river water and river sediments were evaluated for the prevalence of naturalised *Escherichia*. Once reviewed, ESR assisted NRC in designing a sampling schedule and developing a comprehensive methodology for characterising those *Escherichia* isolates that may or may not belong to the species of *E. coli*.

This report provides the first methodology in Aotearoa applicable to characterising *E. coli* subtypes, and identifying naturalised *E. coli* and other *Escherichia* species that are potentially confounding water quality monitoring. The methodology provides guidance with the design of a testing programme, analytical approach and interpretation of the significance of the results.

The Northland Rivers sampled during 2017 and 2018 represented a range of anthropogenic influences including native bush (Victoria at DOC Reserve Crossing, Waipoua at Swimming Hole), pastoral (Ahuroa at Piroa Falls, Tirohanga at Tirohanga Road, Waitangi at Wakelins), urban (Raumanga at Raumanga Valley Park), and a mixture of all of these including exotic forest (Hatea at Whangarei Falls).

Escherichia were enumerated in water and sediment samples. Purified *Escherichia* isolates (n = ~20 isolates per sample) were characterised into either the *E. coli* phylogroups (A, B1, B2, C, D, E and F) or into the *Escherichia* cryptic clades by the Clermont et al. (2013) quadruplex Polymerase Chain Reaction (PCR) scheme. Where one or more *E. coli* phylogroups dominated at a particular location, these *E. coli* strains were further differentiated by Repetitive Extragenic Palindromic (Rep)-PCR methodology to classify

isolates into groups of genetically related strains. If multiple isolates of a genetically related strain of *E. coli* were identified in a sediment sample, this provided evidence that this particular strain was persisting and potentially replicating in the sediment, supporting its identification as a "naturalised" strain of (faecally-derived) *E. coli*.

The main findings from the survey of sediment and water from Northland Rivers and characterisation of *Escherichia* isolates are presented below:

- *E. coli* concentrations in water (as measured by membrane filtration and growth on an *E. coli* selective medium) were low for the 2017-2018 sampling events, with a maximum of 770 colony forming units (CFU)/100 mL, and 60% of samples contained less than 260 *E. coli* /100 mL.
- *E. coli* concentrations in sediment (as measured by membrane filtration and growth on an *E. coli* selective medium) for the 2017-2018 sampling events were much higher than water with an average of 8,100 CFU/100 gram dry weight of sediment.
- Individual putative *E. coli* isolates from water and sediment were characterised into either *E. coli* or cryptic clades of *Escherichia* with the majority of samples being numerically dominated by the *E. coli* isolates (92% of water samples, n = 12, and 82% of sediment samples, n = 11). Prevalence of *E. coli* in the 20 isolates characterised from individual samples ranged from 52 to 100% in water and 33 to 100% in sediment.
    - Numerically dominant phylogroups of *E. coli* in water and sediment were B1 and B2, followed by the group "D or E". *E. coli* phylogroups B1, B2, D and E include pathogenic forms of *E. coli*. *E. coli* belonging to the commensal phylogroup A were infrequently isolated from sediment.
    - Rep-PCR identified genetically related groupings of *E. coli* subtypes belonging to B1 or B2 or "D and E" phylogroups in sediments. These

results provided evidence for the naturalisation of certain strains of (faecal) *E. coli* in sediments where these particular strains were able to persist and grow. Furthermore, there was evidence of the potential transfer of genetically related *E. coli* strains between sediment and water, and the identification of spatial and temporal persistence of genetically related *E. coli* in sediment between sampling events.

- All of the naturalised *Escherichia* Cryptic Clade isolates recovered from sediment and water in this sampling programme belonged to Clade V. Clade V was identified in 75% of water samples (n = 12) and 91% of sediment samples (n = 11). During the February 2017 sampling event, 0-20% of the isolates in each water sample and 8-15% of the putative *E. coli* isolates in each sediment sample belonged to the naturalised Clade V *Escherichia*. A higher overall frequency of Clade V was identified in the 2018 sampling events with Clade V isolates ranging from 0–48% in water and 0–67% in sediment, but this was in association with, in general, lower *E. coli* concentrations in the water (as measured by routine water quality methods).
    - Highest Clade V prevalence occurred at Victoria which was one of the sites classified as having low anthropogenic influences. Hatea (classified as having mixed pollution sources) was the only site that reported non-detection of Clade V in either sediment or water on two sampling occasions.

Recommendations

Where chronic elevated *E. coli* concentrations are identified during routine water quality monitoring then these sites should be first investigated by sanitary surveys and Faecal Source Tracking (FST) analyses.

When this second tier of investigation does not identify a human / livestock or avian faecal source(s) or the faecal source is not recent, an investigation of the contribution of naturalised

sources of *Escherichia* to *E. coli* concentrations in water may be worthwhile. A high proportion of *Escherichia* cryptic clades in water/sediment, and/or genetically related strains of (faecal) *E. coli* in sediment would be indicative of naturalised *Escherichia*. This report outlines the methodology for investigation of potentially naturalised *Escherichia* from water and sediment.

The report has also highlighted that naturalised sources of *E. coli* in sediments are likely to belong to the *E. coli* phylogroups of B1, B2, D and E. Some strains in these *E. coli* phylogroups have been implicated in clinical cases of infection. The identification of naturalised sources of *E. coli*, may therefore, warrant a health risk investigation into whether other pathogen groups associated with faecal contamination such as *Campylobacter*, *Cryptosporidium*, *Giardia*, and viruses are also residing in the sediments. Some of these pathogens may persist in sediments even after the implementation of mitigation actions to remove a faecal contamination pathway.

We used membrane filtration in conjunction with growth on Brilliance *E. coli*/Coliform agar in this experiment to obtain *E. coli* counts. However, so far in this prevalence pilot study we have not identified naturalised *Escherichia* Cryptic Clades belonging to Clades II, III and IV. Therefore, future studies could incorporate a media trial study to determine if other *E. coli*-specific isolation media are able to isolate all putative *Escherichia* Cryptic Clades from the NZ environment. This could establish whether Brilliance *E. coli*/Coliform agar has a bias towards isolation of only Clade V or whether there is a truly low prevalence of Clades II-IV in the NZ (Aotearoa) environment.

1. INTRODUCTION

Escherichia coli occurs in high numbers in mammalian faeces and is used as a faecal indicator to assess freshwater for faecal contamination for the purpose of advising on public safety (Ministry for the Environment and Ministry of Health 2003).

The potential role of naturalised *Escherichia* species as confounders to routine water quality testing has been discussed by scientists in recent years (Jang et al. 2017, Walk 2015). These bacteria, while present in water samples, may not be a result of recent faecal input. Environmental Science and Research Limited (ESR) carried out a literature review on the sources of naturalised microorganism in streams (Devane 2015). The review highlighted that there are two different types of naturalised microbes belonging to the *Escherichia* genus. One group of the naturalised *Escherichia* belong to the species of *E. coli* and are derived from a mammalian faecal origin. These naturalised *E. coli* have been able to survive post-defecation into the environment and reside in reservoirs such as soil and river sediment. The other group are identified as *E. coli* using conventional *E. coli* water quality testing methods but in fact belong to separate species within the *Escherichia* genus. They are termed cryptic clades of *Escherichia* and are rarely identified in human clinical or mammalian faecal specimens, and therefore, have previously escaped detection and characterisation. Furthermore, biochemcial tests used to differentiate *E. coli* from other microbes do not separate these naturalised *Escherichia* cryptic clades from *E. coli*.

It has only been since the advent of whole genome sequencing that the genetic differences have been revealed between these species of *Escherichia* and the *E. coli* derived from mammalian faeces (Walk et al. 2009). Their low occurrence in clinical settings, low level of antibiotic resistance and virulence factors also suggests that these cryptic clades of *Escherichia* are not highly pathogenic strains but rather may be opportunistic pathogens (Walk 2015).

The literature review of these two groups of naturalised *Escherichia* (the environmentally persistent *E. coli* and the cryptic clades of *Escherichia*) led to regional councils (RC) seeking further information and guidance on addressing the potential for them to act as confounders of microbiological water quality monitoring.

It was considered that the next step in the knowledge transfer from ESR to regional councils would be the demonstration of how the skills and techniques highlighted in the literature review could be applied to distinguish *Escherichia* that originate from a recent faecal input from those *Escherichia* (both faecal and cryptic clades) that may be naturalised, and therefore, not correlate with faecal pathogens and a concomitant health risk.

1.1 WORK PLAN:

Northland Regional Council (NRC) sought assistance and guidance from ESR to design and implement a project that would help with the identification of naturalised *Escherichia* in Northland rivers. This project involved ESR undertaking a review of NRC data. Once reviewed, ESR assisted NRC in designing a sampling schedule and developing a comprehensive methodology for characterising those *Escherichia* isolates that may or may not belong to the *E. coli* species.

The work plan included some initial laboratory analysis on “real samples” collected from Northland rivers during 2017 (prior to the Envirolink) and 2018 by the Northland Regional Council and whose *Escherichia* characterisation was partially funded by NRC. During the summer months, multiple rivers are monitored weekly as part of the council’s swimming water quality monitoring programme. The seven rivers chosen for this project have been selected from the council’s swimming programme as they are popular swimming sites in Northland.

The laboratory analyses demonstrated the steps in the process of characterisation of putative *E. coli* isolates, and assisted with development of a methodology for isolation and characterisation of *Escherichia* from river water and river sediment. The methodology

outlined in this report will act as a guiding protocol for other regional councils willing to undertake similar projects. It will provide assistance with the design of the programme and the interpretation of the significance of the results.

Characterisation of *E. coli* / *Escherichia* strains will enable NRC to determine whether *E. coli* above recreational water quality guidelines are indicative of recent faecal inputs. This information will empower good decision making about whether further investigation is required (such as pathogen identification) to better assess health risk to swimmers.

2. METHODOLOGY FOR CHARACTERISATION OF ESCHERICHIA FROM RIVER WATER AND SEDIMENT

2.1 CHARACTERISTICS OF SAMPLING SITES

Global positioning satellite (GPS) coordinates for each site are provided in Appendix A.1 and a map of the Northland river sites selected for sampling is presented in Figure 1.

2.1.1 Land cover categories for each site

Table 1: Northland rivers sites selected for sampling events for naturalised *Escherichia* and *E. coli*

Site location	Site number	Land Cover Categories	Sampling Events		
			February 2017	**April 2018**	**August 2018**
Ahuroa at Piroa Falls	31759	Pastoral (strong dairy influence)	No	Yes	No
Hatea at Whangarei Falls	105972	Native/exotic/ pastoral/urban	Water only	Yes	Yes
Raumanga at Raumanga Valley Park	103246	Urban	No	Yes	Yes
Tirohanga at Tirohanga Road	102252	Pastoral	No	Yes	Yes
Victoria at DOC Reserve Crossing	104908	Native	Yes	Yes	Yes
Waipoua at Swimming Hole	108613	Native	Yes	Yes	Yes
Waitangi at Wakelins	101752	Pastoral	Yes	Yes	Yes

2.2 RAINFALL AND WATER QUALITY PARAMETERS

Sampling was delayed during late 2017 and early 2018 due to heavy rainfall in the Northland catchments. Rainfall data (approximately two weeks prior to sampling) and water quality parameters of dissolved oxygen, water temperature, pH and specific conductivity (SPC) were collected by Northland Regional Council.

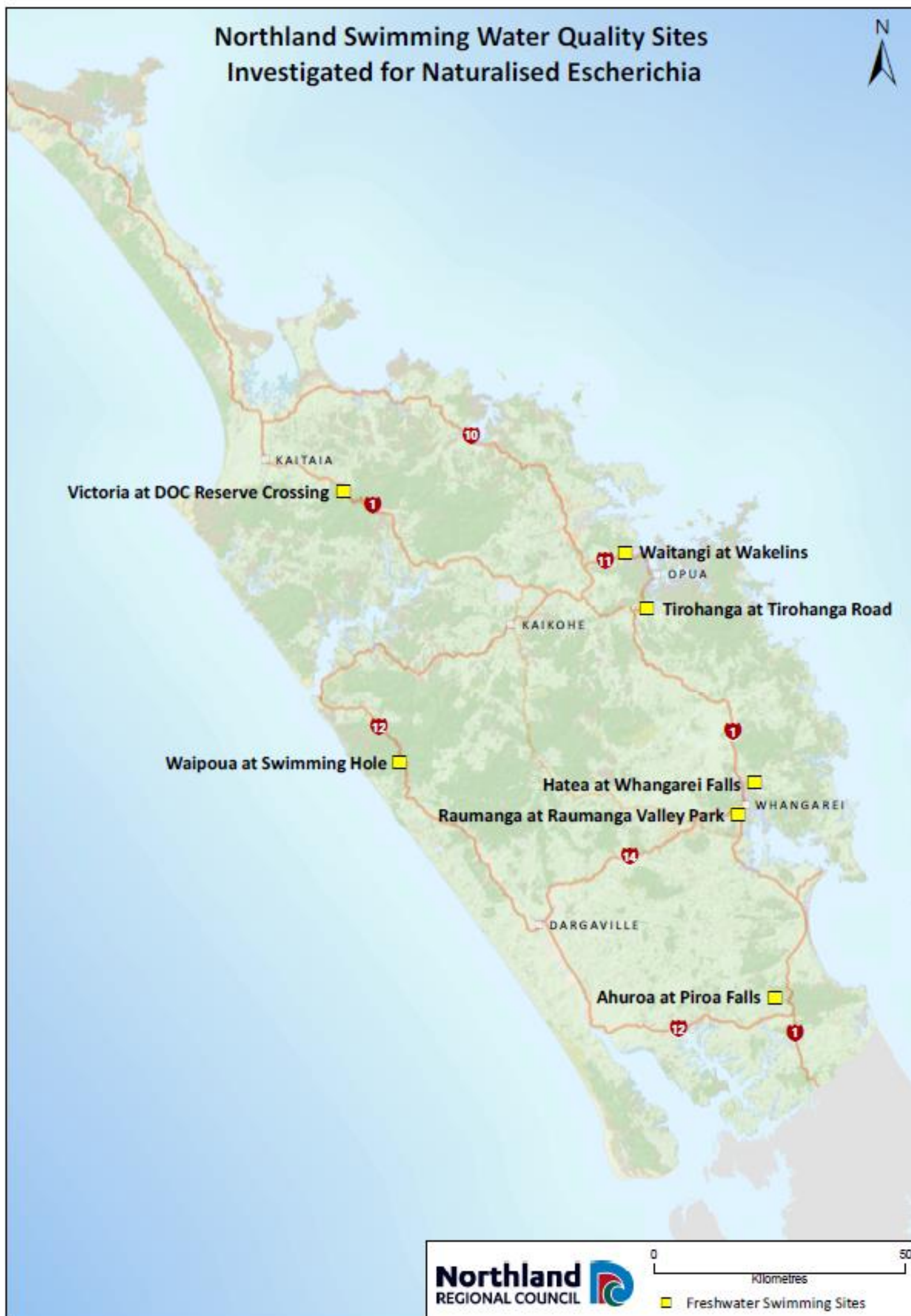


Figure 1: Map of the Northland river sites selected for sampling for naturalised *Escherichia* and *E. coli*

2.3 SAMPLE COLLECTION

Day 1, eg. Monday:

Equipment list:

- Mighty Gripper (The Mighty Gripper Company, Whangarei, NZ)
- 500 mL sterile plastic bottle for water samples, which will fit the Mighty Gripper
- 250 ml sterile pot or container to fit the Mighty Gripper. Container does not need to be deep, but needs to be wide for sample collection.
- Chilly bin with ice packs.
- Water quality instrumentation for measuring dissolved oxygen, water temperature, pH and specific conductivity.

Method for water sampling

1. Collect water samples before sediment so as not to disturb sediment.
2. Water samples are taken standing on the bank using the Mighty Gripper using the built-in extension capability if required. Ensure that during sampling you avoid disturbance of the underlying surface sediment.
3. Use the Mighty Gripper with attached pottle and sample water as per the Ministry for the Environment and Ministry of Health (2003)guidelines : In general, the sample will be taken at approximately 30 cm below the surface at a point where the depth of water is 1 metre for freshwater. However, for these investigations, streams may be shallower so a depth of approximately 10 cm is valid to avoid disturbance of the underlying sediment.
4. Seal the container and label with the sample ID, site name and date & time of collection.
5. Place sample in the chilly bin with ice packs during transport.
6. Ensure sample is placed in the fridge/cool room on arrival back at the office. Samples can be couriered overnight (with ice packs) to laboratory.

Method for sediment sampling

7. Sample the sediment approximately 1 – 2 m upstream of the water samples at each site. Stand on the shoreline or downstream of the sampling area, making sure not to disturb the surface sediment.
8. The majority of the *E. coli* reside within the top 2 cm layer of the sediment. Use the extended pole of the Mighty Gripper with attached 200 mL sterile pottle to sample the upper 1-2 cm of the sediment by dragging the pottle along a 1 metre transect of the surface of the sediment.
9. Seal the container and label with the sample ID, site name and date and time of collection.
10. Place sample in the chilly bin with ice packs during transport. Do not decant water off the surface of the sediment. Ensure the sediment is covered by a layer of the collected water so it does not dry out.
11. Ensure sample is placed in the fridge/cool room on arrival back at the office. Samples can be couriered overnight (with ice packs) to laboratory.

2.4 ENUMERATION OF *ESCHERICHIA* FROM WATER SAMPLES

Day 2 eg Tuesday: Membrane filtration

Analyse 100, 10, 1 ml of water sample in duplicate using the standard membrane filtration method with a 0.45 μm filter (Millipore), APHA 9222 Membrane filter technique for members of the coliform group (APHA 9222- 2017). Filters are placed on Brilliance agar: *E. coli*/Coliform agar (CM0956Thermo Scientific, formerly Oxoid) and incubated at 30°C for 4 hrs and then 37°C for 20 hrs.

Control agar plates

Include positive and negative Brilliance *E. coli* / Coliform media controls of known *E. coli* (positive = purple) and *Klebsiella pneumoniae* (negative = pink colony) and *Enterococcus faecalis* (no growth) plus a sterility media control plate containing no inoculum of bacteria.

Day 3 eg. Wednesday: *E. coli* enumeration

E. coli produce both ß -galactosidase and ß-glucuronidase activity and are, therefore, able to cleave both chromogens present in the Brilliance *E. coli*/Coliform agar to form purple colonies, whereas other coliforms only produce ß -galactosidase and form pink colonies (eg. *Klebsiella pneumoniae*), as they cleave only the galactoside chromogen. The media also inhibits Gram-positive organisms and contains tryptophan, allowing the rapid confirmation of *E. coli* by the formation of indole using a spot indole test if required.

Count and record purple *E. coli* colonies on dilution plates containing between 20 and 200 colonies per plate. Check control agar plates for appropriate growth/non-growth.

2.5 ISOLATION OF *ESCHERICHIA* FROM WATER SAMPLES

The streak-plating method was applied to ensure that the single putative *E. coli* colony selected from the agar plates represented a pure isolate (Liang et al. 2016):

Day 3 eg Wednesday:

Aim: to obtain isolated subtypes of putative *E. coli* for PCR subtyping characterisation by purifying each colony to check we have a single subtype from our original isolated colony. If the colony is not from a single subtype this will affect the characterisation as you will get mixed subtypes contributing to the subtype band pattern when we do the PCR analysis (Clermont et al. 2013). At the end of this procedure we should have approximately 20 isolated colonies from each selected water/sediment sample.

After counting of *E. coli* colonies (normal membrane filtration method), take a single, well isolated purple colony and streak isolate onto Brilliance *E. coli* / Coliform agar. To save on

labour divide agar plate into 4 quadrants and isolate one *E. coli* colony into each quadrant as per the diagram (Figure 2) so that you have four colonies per agar plate. Incubate at 37°C overnight. Can reduce temperature to 30°C to reduce growth and prevent individual colonies from merging on the plate. Repeat this procedure to obtain 20 isolated colonies from each water / sediment sample.

Choose water samples with elevated counts of *E. coli* >200/100 ml for water and any dilution plates from the underlying sediment sample that have greater than 20 colonies of *E. coli*. NB can obtain colonies from multiple plates belonging to the same sample.

Labelling of isolates.

Label each isolate from a single sample with the appropriate identifier for each site and number colonies with the suffix 1 up to 20.

Each time please perform all colony isolation within a Biohazard cabinet both to protect the sterility of the isolation procedure and to protect the worker as some *E. coli* on the plates may be pathogenic and note that *E. coli* O157 is one of those *E. coli* subtypes that is β-glucuronidase negative and therefore will not be purple on the plate, however, other pathogenic *E. coli* subtypes will be purple on Brilliance *E. coli* / Coliform agar.

Day 4 eg Thursday: Isolation and purification of 20 *E. coli* colonies per water sample

Pick another single, well isolated, purple colony and streak onto a non-selective agar such as TSA (Tryptic soy agar) or CMB Columbia blood agar to check that the colonies truly are pure isolates. Do all the colonies from a single colony look the same or are there different morphologies? If different morphologies are present or there are no well-isolated colonies to choose then re-streak overnight on a half agar plate to provide more room to obtain well-isolated colonies that are likely to be derived from a single colony. Incubate at 30°C or 37°C overnight.

Day 4 When purified *E. coli* colonies are obtained (so that one isolate represents one *E. coli* strain) - then proceed with putting the individual isolates into storage for Step a, keeping *E. coli* viable in broth +glycerol and Step b, preparing isolate for DNA extraction for PCR identification.

a) Storage of viable putative *E. coli* isolates

   Take an individual colony from the plate and put into sterile Brain Heart Infusion (BHI) broth containing 20% glycerol (1 ml in Eppendorf) and freeze at -80°C in freezer.

b) Preparation of *E. coli* isolate for PCR identification

   Take another individual colony and put into chelex solution (Appendix B.1), heat at 96°C for 10 minutes and store at -20°C. This chelex preparation will be used as the DNA template for the Subtyping PCR of Clermont et al. (2013) and Rep-PCR.

2.5.1 Notes on streak isolation procedure to obtain purified colonies of *E. coli* derived from a single *E. coli* strain

Below is the picture of a Brilliance *E. coli* / Coliform agar plate subdivided into four sections for individual colonies from each water/sediment sample (Figure 2). The final aim is to be able to isolate 4 colonies per plate but it may be necessary to use half a plate per isolate in order to obtain discrete single colonies at the end of the streak – this will be up to the individual preference of the worker. Take a single colony and spread colony over the top of the section so as to uniformly spread out the bacteria in that colony. Then get <u>a new loop</u> out and <u>draw it once through the heavy streak</u> and sweep side to side as per the drawing to draw out the individual bacteria so they grow into isolated colonies.

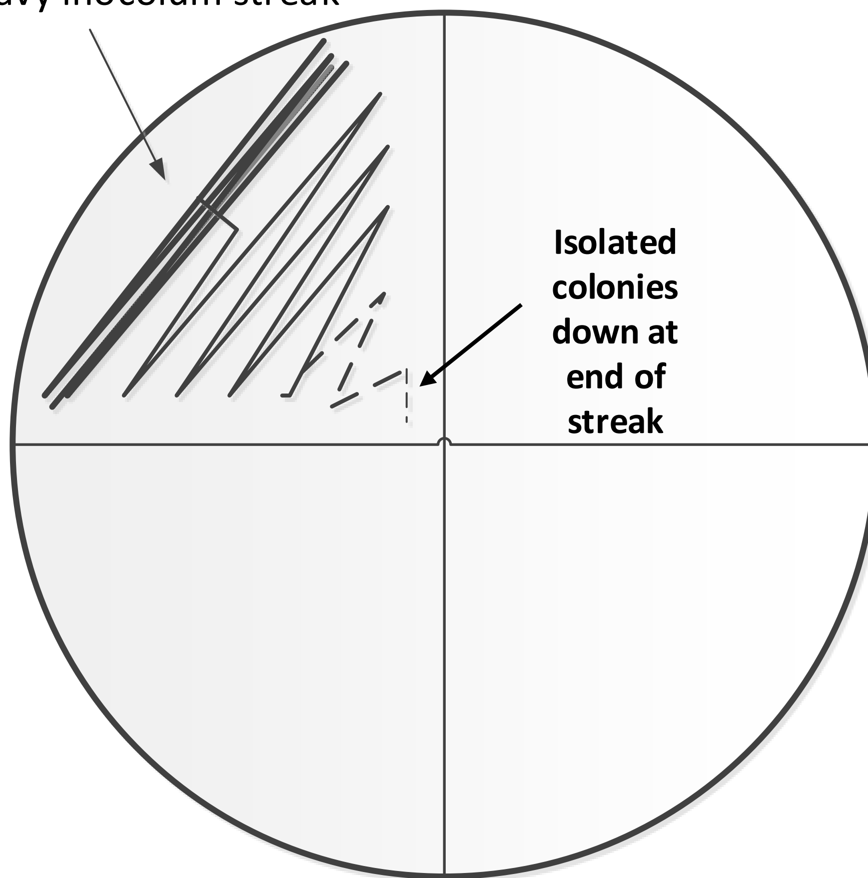


Figure 2: Colony isolation of *E. coli*. Plate is divided into four and one colony is streak isolated onto each quarter of the plate

2.6 ENUMERATION AND ISOLATION OF *ESCHERICHIA* FROM SEDIMENT SAMPLES

Dilutions of sediment depend on the likely levels of *E. coli* in the sediment based on previous samplings and/or prior weather conditions. Therefore, the dilutions recommended below are a guide only.

1. The sediments were allowed to stand for 30 min to allow the bulk sediment to settle, then overlying water was decanted and discarded.
2. Sediment samples were mixed to ensure a homogenous suspension and a subsample (20 g) placed into a sterile 120 mL yellow topped pottle (from WM3 lab).
3. Create a 1:1 (0.5) dilution of the sediment:
    a. Add Ringers Solution (¼ strength) to sediment to create a 1:1 (0.5) dilution (20 g sediment and Ringers Solution added to final weight of 40 g).

4. The 1:1 resuspension is mixed by hand for 2 min, allowed to settle (< 5 min) and a volume of the supernatant eluted and tenfold dilution made by taking 5 mL of the 1:1 resuspension and diluting to 10 mL with 5 mL of Peptone water = 10^{-1} dilution. Then a 1:10 dilution series was made from 10^{-2} to 10^{-4} using Peptone water as follows:

10^{-2} dilution: Pipette 1 ml of **10^{-1} dilution** into 9 mL of peptone water. Gently shake to disperse bacteria.

10^{-3} dilution: pipette 1 ml of **10^{-2} dilution** into 9 mL of peptone water. Gently shake.

10^{-4} dilution: pipette 1 ml of **10^{-3} dilution** into 9 mL of peptone water. Gently shake.

Filter a 1 mL volume of the initial sediment supernatant (10^{-1} dilution) into 99 mL of sterile water in the filter funnel and repeat for other dilutions using the membrane filtration method (APHA 9222- 2017). All dilutions were done in duplicate. Filters were placed onto Brilliance *E. coli* / Coliform agar and incubated as per *E. coli* water samples. After enumeration of putative *E. coli*, then 20 colonies per sediment were isolated for purification as for the water samples (Sections 2.4 - 2.5).

All sediment concentrations were reported as counts per 100 gram dry weight of sediment (Section 2.7).

2.6.1 Notes on *E. coli* enumeration in sediment

If comparing enumeration of *E. coli* between sediment and the overlying water column, it may be important in future studies to report the *E. coli* in sediment as colony forming units (CFU)/100 g of wet weight (ww) of sediment. This is recognised as not being a perfect comparison but is based on water weighing 100 g/100 mL and allows a direct comparison between water and underlying sediment (Nguyen et al. 2018). However, dry weight of sediment should also be calculated as this enables comparison of *E. coli* concentrations between sediments at different sites which will have different characteristics and water holding capacity and thus results can be reported as CFU/ 100 gram of sediment dry weight (dw).

2.7 DRY WEIGHT DETERMINATION FOR SEDIMENT

Dry weight was determined by drying a subsample of sediment (10 – 20 g) in a 105°C oven overnight and repeating until there was no significant weight change over several days (American Public Health Association (APHA 2005)).

2.8 CHARACTERISATION OF PUTATIVE *ESCHERICHIA* ISOLATED COLONIES

Isolated colonies of putative *E. coli* were analysed by ESR with the conventional Polymerase Chain Reaction (PCR) of the revised Clermont quadruplex PCR (Clermont et al. 2013), (which, herein, will be referred to as the quadruplex Clermont PCR). This quadruplex Clermont PCR detects all phylogenetic groups of *E. coli* and also the members of the environmental *Escherichia* cryptic clades. The *Escherichia* cryptic clades are initially clustered into non-specific Clade groupings (for example one amplicon in the quadruplex Clermont PCR detects members of Clades III, IV and V) and can be further differentiated into individual clade designations by the specific PCR markers developed by Clermont et al. (2011). No further analysis of these cryptic clade groupings by Rep-PCR is required as their identification is sufficient for confirming the presence of naturalised cryptic clades of *Escherichia*.

E. coli isolates designated as one of the seven phylogroups A-F by the quadruplex Clermont PCR were collated and where the same dominant phylogroup occurred at the same location in the water and the underlying sediment (defined as approximately ≥40% of the ~20 isolates in one matrix at a particular location) then isolates were further subtyped using the Repetitive Extragenic Palindromic (Rep)-PCR methods (Rademaker et al. 2004). Only the *Escherichia* colonies from the second sampling event in February 2017 (20th February) were isolated and characterised. The selected isolates of *E. coli* phylogroups collected during 20th February 2017 from NRC sites were subtyped using the Rep-PCR GTG primer system (Mohapatra and Mazumder 2008). However, selected *E. coli* phylogroups collected during 2018 were subtyped by the Box A1R primer Rep-PCR method (Rademaker and de Bruijn 1997, Versalovic et al. 1994) because this second method improved the strain discrimination

between isolates by producing a greater number of band patterns, which showed improved clarity compared with the GTG primer system. The use of two different Rep-PCR methods for the 2017 (February) and 2018 (April and August) *E. coli* phylogroup isolates meant that isolates subtyped by these two methods could not be compared.

It should be noted that the B2 phylogroup can have several PCR patterns as was noted for Victoria in April 2018, where three of the B2 in sediment had one PCR pattern and eight had a second pattern. Therefore, although in this sample there were 11 isolates in this B2 phylogroup, they were not all genetically related and only the eight isolates with the same PCR pattern in sediment and five in water were further subtyped by Rep-PCR for comparison. In addition, the PCR pattern designated as “D or E” was not subjected to additional PCR to differentiate whether isolates belonged to either phylogroup D or E. Instead the Rep-PCR analysis was performed to differentiate the isolates into genetically related clusters.

Computer-assisted Rep-PCR DNA fingerprint analysis (Bionumerics version 7.6, Applied Maths, Belgium) was performed on a selection of *E. coli* isolates from a single sample (water or sediment) that belonged to the same *E. coli* phylogroup as determined by the PCR of Clermont et al. (2013). Cluster analysis was performed using curve based Pearson’s product moment correlation coefficient with 1.5% optimisation and 0.5% curve smoothing and hierarchal clustering by unweighted pair group method with arithmetic mean (UPMGA). Genetically related subtypes capable of persistence and/or replication in the environment were suspected if the Rep-PCR DNA fingerprint profiles of *E. coli* isolates shared a genetic similarity of ≥ 90% (Liang et al. 2016).

2.9 STATISTICAL ANALYSIS

Statistical analysis was performed using RStudio (R Core Team 2018) including Student t-test for comparison of seasonal data using the Welch Test for unequal variance.

3. RESULTS

Northland Regional Council collected water and/or underlying sediment samples from four river locations on two occasions in February, 2017. In 2018, the same four rivers and locations were targeted for the collection of water and sediment on 9th April and 6th August. An additional three rivers were also targeted for collection of water and underlying sediment in April but only two of those rivers were sampled in August 2018. View Table 1 for details on which rivers were sampled.

3.1 RAINFALL AND WATER QUALITY PARAMETERS

Heavy rainfall was identified in the two weeks prior to the February sampling events in 2017. Therefore, during the summer of 2017-2018, sampling was delayed on multiple occasions due to continuous rainfall. Sampling was delayed during times of substantial rainfall to minimise disturbance of the sediments in order to gain a picture of what *Escherichia* types were residing in the sediments.

Rainfall data for February 2017 and for 2018 can be viewed in Appendix APPENDIX B: Figures 4-6. Cumulative rainfall data during February 2017 and 2018 sampling periods is presented in Table 8 and shows the higher rainfall during the 2017 February sampling events compared with 2018. During the period from 1st to 21st February, 2017, Ahuroa received the lowest rainfall (67 mm) and Tirohanga the highest rainfall (149 mm). In the two weeks prior to the 20th February sampling event there were two major rain events that would have impacted resuspension of sediments. This included heavy rainfalls of 20-40 mm at all sites 3-5 days before sampling on the 20th February.

Cumulative data for rainfall 17 days prior and during the first sampling in April 2018 recorded levels of rainfall ranging from 29 mm (at Tirohanga) to 79 mm (Ahuroa and at Hatea) (Table 8). In April 2018, sampling was initiated on April 9th after a period of negligible rainfall for 11 days. Rainfall had occurred on 21st -24th March 2018 (approximately 5 to 41 mm on 24th

March) and 29th March when approximately 5-13 mm of rain fell at most of the sites except Waitangi (Figure 5). Waipoua was the site least affected by rainfall during this April period.

Cumulative data for rainfall 14 days prior and during the second sampling event in August 2018 recorded lower levels of rainfall (range 2 to 38 mm) compared with the April 2018 event (Table 8). For the second sample event on 6th August 2018, rain fell three days prior to sampling at most sites excluding Waitangi with up to 10.5 mm at Victoria. Ahuroa was not sampled during this second event.

Water quality parameters were only measured during the 2018 sampling events and are presented in Appendix C.2. Dissolved oxygen (DO) ranged from 86 to 107 % at all sites and 9 to 11 mg/L . These levels for DO are well above the recommended levels for the health of freshwater fish (www.water-research.net/Watershed/dissolvedoxygen). The range of pH at all sites was 6.1 to 8.1. and water temperatures during April 2018 were 15 to 18 °C , and from 9.3 to 12.1° C in the winter sampling of August 2018.

3.2 *E. COLI* ENUMERATION IN WATER AND SEDIMENT

E. coli in water and sediment were enumerated by membrane filtration in each of the matrices. The overall mean concentration of *E. coli* in water for 2017-2018 samplings was just below the Alert level (Table 2) and concentrations were higher in summer compared with winter time. A breakdown of all *E. coli* concentrations in water samples can be found in Appendix APPENDIX D:.

3.2.1 *E. coli* in water

In general, *E. coli* concentrations in river water were low with the Action level of 550 colony forming units (CFU)/100 mL being exceeded on four occasions (four different locations) out of 20, and all water samples were below the national bottomline of 1000 CFU/100 mL.

In 2017, *E. coli* concentrations in water (E. *coli* concentrations in sediment and water

Table 9) were below the Alert guidelines of 260 *E. coli* CFU/100 ml at Waitangi and Hatea, and at Waipoua on the first sampling occasion, but exceeded the Alert level on one occasion at Victoria and the Action level of 550 CFU/100 mL at Victoria and Waipoua on the second sampling occasion.

In 2018, *E. coli* concentrations in water were below the Alert guidelines of 260 *E. coli* CFU/100 ml except at Ahuroa (365 *E. coli*) in April and on both sampling occasions at Hatea when concentrations exceeded the Action level. However, the maximum concentration identified on any occasion and location was 770 *E. coli* at Hatea in April 2018.

3.2.2 *E. coli* in sediment

Sediment concentrations of *E. coli* ranged between 300 to 60,700 *E. coli* / 100 g dry weight (dw) of sediment with a mean of 8,100 CFU/100 g dw for all sediment samples collected in 2017 and 2018 and higher average of *E. coli* in sediments during summer/autumn compared with winter (Table 2). A breakdown of all *E. coli* concentrations in sediment samples can be found in Appendix APPENDIX D:.

Table 2: *E. coli* concentrations in sediment and water 2017-2018 as determined by membrane filtration

	Mean *E. coli* concentrations (number of samples)			
	Water CFU/100 mL	SD	Sediment CFU/100 g dry wtg.	SD
All samples 2017-18	**248** (20)	*232*	**8,100** (20)	*14,000*
Feb-17	**325** (6)	*254*	**7,000** (7)	*8,300*
Apr-18	**228** (7)	*261*	**14,200** (7)	*21,700*
Aug-18	**181** (6)	*178*	**2,300** (6)	*2,200*

3.2.3 Faecal source tracking analysis of selected samples

Faecal source tracking analysis was carried out on selected water samples from Victoria and Waipoua in February 2017 where *E. coli* concentrations were ≥445 CFU /100 mL. The specific faecal sources targeted were human, ruminant and avian PCR markers. Levels of the general non-specific faecal polymerase chain reaction (PCR) marker was identified at levels ranging from 8,000 to 81, 000 gene copies /100 mL but no specific faecal sources

(ruminant/human or avian) were identified in any samples (ESR report to NRC, 26 July 2017). These findings were consistent with observations that these sites had low anthropogenic influences, and were classified as native bush sites.

3.3 CHARACTERISATION OF *ESCHERICHIA* ISOLATES INTO *E. COLI* PHYLOGROUPS OR CRYPTIC CLADES OF *ESCHERICHIA*

Isolates purified from selected water and sediment samples were further characterised into either the *E. coli* phylogroups (A-F) or into the *Escherichia* cryptic clades by the Clermont et al. (2013) quadruplex PCR scheme and then into the individual cryptic clade groups by the Clermont et al. (2011) PCR, where appropriate. The prevalence of isolates of *E. coli* and *Escherichia* cryptic clades in selected samples are shown in Table 3 for the 2017 sampling events and in Table 4 for the 2018 events. These tables also contain the recommendations for which *E. coli* isolates were chosen for further differentiation into clonal groups by the genetic subtyping technique of Rep-PCR.

Cryptic clade *Escherichia*, which are not thought to naturally reside in animal guts, were not characterised beyond their clade group as by definition they are "naturalised *Escherichia*" whose main habitat is that of non-animal associated environments. Results for the determination of cryptic clade groups is presented in Section 3.3.7.

During 2017, the dominant *E. coli* phylogroup identified in 3 of the 4 locations selected for Rep-PCR belonged to the B1 phylogroup (Appendix D.2 Table 11). In 2018, *E. coli* phylogroups B1 abd B2 were frequently isolated from water and sediment followed by the phylogrouping of "D and E" (Appendix D.2, Table 12 and Table 13). The other phylogroups were infrequently isolated and the commensal Group A (Clermont et al. 2000) was isolated very infrequently from sediment in this study. In comparison to the prevalence of the cryptic clades of *Escherichia*, the majority of water samples were numerically dominated by the *E. coli* phylogroups (92% in water samples, n = 12), and 82% *E. coli* in sediment samples (n = 11). Prevalence of these *E. coli* in sediment, ranged from 75 to 100% in water and 33 to 100%.

Table 3: 20th February 2017 sampling; Attribution of *Escherichia* isolates to groups and recommendations for further subtyping analysis of *E. coli* by Rep-PCR to determine if these *E. coli* isolates were genetically similar and hence potentially persisting in the environment. NB two sampling events with enumeration of *E. coli* occurred during February 2017, but additional *Escherichia* characterisation was performed on only one of those events (20th February).

Site ID	ESR Sample Number	Location	Sample Type	*E. coli* /100 mL/ 100 g dry wtg.	*E. coli* (%)	*Escherichia* Cryptic Clades (%)	Recommendations for additional analyses by Rep-PCR of *E. coli*
101752	CMB170232	Waitangi	Water	47	90	10	**B1** was the dominant *E. coli* phylogroup for water (50%) and sediment (30%) and selected for rep-PCR to see if these *E. coli* were genetically similar and hence persisting in the environment
	CMB170236		Sediment	24,200	85*	15	
104908	CMB170233	Victoria	Water	585	80	15	**D/E** was the dominant *E. coli* phylogroup for water (35%) and sediment (30%) and selected for rep-PCR to see if these *E. coli* belonged to D or E and are genetically similar and hence persisting in the environment
	CMB170238		Sediment	9,800	80	15	
105972	CMB170234	Hatea	Water	255	85	0	**B1** was the dominant *E. coli* phylogroup for water (45%) and selected for rep-PCR to see if these *E. coli* were genetically similar and hence persisting in the environment
108613	CMB170235	Waipoua	Water	685	75	20	No further analyses
	CMB170239		Sediment	1,500	75	8	

*Where percentages of *E. coli* and *Escherichia* Cryptic Clades do not add up to 100% this is due to some isolated colonies not being identified as belonging to either of these groups (refer to Appendix D.2; Table 11) .

Table 4: 2018 sampling; Attribution of *Escherichia* isolates to groups and recommendations for further subtyping analysis of *E. coli* by Rep-PCR to determine if these *E. coli* isolates were genetically similar and hence potentially persisting in the environment

Site ID	Location	Month 2018	Sample Type	*E. coli* /100 mL/ 100 g dry wtg.	*E. coli* (%)	*Escherichia* Cryptic Clades (%)	Recommendations: additional analyses by Rep-PCR of *E. coli*
103246	Raumanga	April	Water	142	100	0	**B2** was the dominant *E. coli* phylogroup for water (40%) and sediment (35%) and the dominant B2 PCR pattern (n = 8 water and 5 in sediment) was selected for rep-PCR
			Sediment	5,200	90	10	
		August	Water	127	79	21	In August, **B1** was the dominant phylogroup in water (42%) and **Clade V** in the sediment (60%); however, because **B2** was dominant in April 2018, it was decided to select the **B2** group from the August matrices and type these *E. coli* to compare with the **B2** group in the April 2018
			Sediment	2,500	35*	60	
104908	Victoria	April	Water	62	100	0	**B2** was the dominant *E. coli* phylogroup for water (50%) and sediment (55%) and the dominant B2 PCR pattern (n = 5 water and 8 in sediment) was selected for rep-PCR
			Sediment	21,600	90	10	
		August	Water	63	52	48	High abundance of **Clade V** isolates were noted at Victoria in August samples particularly in the sediments, therefore no further analyses of *E. coli* isolates was performed on August samples**
			Sediment	1,300	33	67	
105972	Hatea	April	Water	770	100	0	**B1** was the dominant *E. coli* phylogroup for water (85%) **B2** was the dominant *E. coli* phylogroup for sediment (90%) and these two phylogroups were selected for rep-PCR
			Sediment	60,700	100	0	
		August	Water	530	90	10	**B1** was the dominant *E. coli* phylogroup for water (40%) and sediment (40%) and this phylogroup was selected for rep-PCR from both matrices
			Sediment	6,500	72	28	
317597	Ahuroa	April	Water	365	95	5	**B1** was the dominant *E. coli* phylogroup for sediment (55%) and there was no single dominant PCR pattern in the water, although the total phylogroup B2 made up 65% of the isolates in the water. The sediment **B1** isolates were selected for Rep-PCR
			Sediment	7,300	90	10	
101752	Waitangi	August	Water	169	80	20	There were no dominant phylogroups in either the water or sediment, and therefore, no further analyses of *E. coli* isolates was performed on August samples
			Sediment	1,000	86	14	

*1 isolate in this sample was not characterized as *E. coli* or an *Escherichia* cryptic clade; **A maximum of 100 isolates including controls were costed for Rep-PCR subtyping, and therefore selection of isolates was rationalised, based on those isolates that were most likely to provide useful information on potential naturalisation of *E. coli* in sediments.

3.3.1 Characterisation of selected isolates of *E. coli* phylogroups to determine genetic relatedness

The aim was to identify clusters of isolates belonging to a single *E. coli* phylogroup in the sediment that were potentially naturalised *E. coli*. Multiple isolates of the same subtype would suggest that this numerically dominant subtype(s) was able to persist and potentially grow in the sediment.

Multiple isolates of the same clonal group within a sediment sample provided evidence for a potential source of *E. coli* for resuspension into the water column. In general, the dominant *E. coli* phylogroups within the sediment and water, at a particular location and time, were selected for further characterisation by Rep-PCR (Table 3 and Table 4). Results for the various locations are presented below based on the order of results as presented in Table 5.

3.3.2 Waitangi

Waitangi isolates in sediment in 2017 were mostly represented by phylogroups B1, B2 and the grouping “D or E”, whereas in the water B1 was the dominant phylogroup (50%) (Table 11). Waitangi samples in April 2018 were not further characterised but August 2018 isolates were represented again by a cross section of B1, B2 and the grouping “D or E”, with members of the A and “A or C” also prevalent in water at 30%, collectively. Additional characterisation by Rep-PCR was not performed on these Waitangi August samples as there was no numerically dominant *E. coli* phylogroup detected.

3.3.3 Victoria

At Victoria in 2017, the combined phylogrouping “D or E” was the dominant group observed in both sediment and water (Appendix D.2, Table 11). This combined group “D or E” was not further refined by individual PCR characterisation, because it was assumed they would be differentiated by the Rep-PCR profiling. One clonal group of four of these “D or E” isolates (57%) were identified in the water and three in the sediment (50%), with these three sediment isolates also being related at a similarity of 90-94% to one of the water isolates.

In 2018 at Victoria, the numerically dominant isolate groupings were phylogroup B2 in water (50%) and sediment (55%) during the April sampling; and Clade V in water (48%) and sediment (67%) during August (Appendix D.2, Table 12 and Table 13). No further analysis was undertaken in the August sampling at Victoria as the naturalised Clade V was the dominant group with highest abundance for any location and sampling time. Rep-PCR was performed on those April 2018 B2 isolates that had the same quadruplex Clermont PCR pattern in water and sediment. Three groups of the B2 isolates in water were identified as being clonally related within the grouping, with two of these groups showing 84-94% similarity to sediment isolates at Victoria.

Overall, the phylogroup B2 (2018) and the grouping of “D or E” (2017) at Victoria showed the potential for the transfer of related *E. coli* strains between water and sediment and the identification of clonally related strains within the sediments suggested the sediment was acting as a habitat for these particular *E. coli* strains.

Table 5: Rep-PCR results for determining genetic relationships and potential for naturalisation within dominant *E. coli* Phylogroups collected in sediment and water at the same location during 2018 sampling events. Cells shaded in the same colour represent related subtypes between different samples/locations/times. Note well subtypes in 2017 cannot be compared with 2018 subtypes due to improved methodology used in 2018 for Rep-PCR. For simplicity, those isolates that were not related to any other isolates are not included in this table summary.

Date of sampling	Location	Sample Type	*E.coli* Phylogroups (Rep-PCR isolate numbers)	Number isolates related (% of the phylogroup)	Genetic similarity within same sample	No. of isolates related between other sample types (% similarity)*	Conclusions for *E. coli*
Feb 2017	Waitangi	Water	B1 = 10	2 (20%)	90%	unrelated	Some evidence for genetically related *E. coli* in sediment supporting presence of naturalised *E. coli*
				3 (30%)	90 – 91%	3 isolates similar (88%) to 3 water isolates at Hatea Feb 2017	
		Sediment	B1 = 6	2 (33%)	87%	unrelated	
Feb 2017	Victoria	Water	D&E = 7	4 (57%)	90 – 100%	3 sediment related to 1 water (90 – 94%) at Victoria Feb 2017	Potential *E. coli* transfers between water and sediment. Genetically related *E. coli* in sediment supports presence of naturalised *E. coli*.
		Sediment	D&E = 6	3 (50%)	90 – 96%		
April 2018	Victoria	Water	B2 = 5	2 (40%)	88%	unrelated	Potential *E. coli* transfers between water and sediment. Genetically related *E. coli* in sediment supports presence of naturalised *E. coli*.
				2 (40%)	89%	related to 2 Victoria sediment isolates in April 2018 (89-94%)	
				1 (20%)	-	Related to 3 Victoria sediment isolates in April 2018 (90%)	
April 2018	Victoria	Sediment	B2 = 8	2 (25%)	94%	refer above cell	
				3 (38%)	90%	refer above cell	
Feb 2017	Hatea	Water	B1 = 9	6 (67%)	84-91%	3 isolates similar (88%) to 3 water isolates at Waitangi Feb 2017	Potentially similar faecal source as 6 of 9 isolates in phylogroup B1 were genetically similar. Note that

Date of sampling	Location	Sample Type	*E.coli* Phylogroups (Rep-PCR isolate numbers)	Number isolates related (% of the phylogroup)	Genetic similarity within same sample	No. of isolates related between other sample types (% similarity)*	Conclusions for *E. coli*
							there was no sediment collected from Hatea in Feb 2017
Apr 2018	Hatea	Water	**B1** = 17	2 (12%)	94%	unrelated	Identified *E. coli* subtype groups with multiple members in water. Potential *E. coli* transfers between water and sediment identified within same sampling period and also between April and August events.
				5 (29%)	91%	89% related to 1 isolate in sediment, August 2018 at Hatea	
				9 (53%)	88-96% (2 groups closely related to each other)	unrelated	
Apr 2018	Hatea	Sediment	**B1** = 2	2 (100%)	Unrelated to each other	unrelated	100% of 18 isolates of B2 phylogroup were genetically related in sediment supporting presence of naturalised *E. coli* in April sampling but there was limited evidence for replication within sediment during August sampling with multiple subtypes detected with 2 or less isolates per group. Refer to Ahuroa samples below for persistence of genetically related naturalised *E. coli* subtypes identified in sediment at two locations.
			B2 = 18	18 (100%)	>90%	unrelated	
Aug 2018	Hatea	Water	**B1** = 8	7 (88%)	90% (1 isolate 88% to these others)	88-90% related to 2 sediment isolates at Hatea in August	
Aug 2018	Hatea	Sediment	**B1 = 10**	1 (10%)	unrelated	89% related to 5 water isolates at Hatea in April	
				2 (20%)	93%	unrelated	
				2 (20%)	98%	>90% related to 4 sediment isolates from Ahuroa Apr 2018	
				2 (20%)	90%	88-90% related to 7 water isolates at Hatea in August	

Date of sampling	Location	Sample Type	*E.coli* Phylogroups (Rep-PCR isolate numbers)	Number isolates related (% of the phylogroup)	Genetic similarity within same sample	No. of isolates related between other sample types (% similarity)*	Conclusions for *E. coli*
Apr 2018	Ahuroa	Sediment	**B1 = 11**	4 (36%)	90%	>90% related to 2 sediment isolates from Hatea in August*	Evidence of spatial and temporal persistence of naturalised *E. coli* subtypes in sediment at different locations and times
				4 (36%)	90%	unrelated	
Apr 2018	Raumanga	Water	**B2 = 8**	5 (63%)	92%	89% related to one isolate from Aug 2018 water sample at Raumanga	Several occasions where genetically related *E. coli* subtypes were identified between April and August sampling events at Raumanga. Potential *E. coli* transfers between water and sediment and identification of the potential temporal persistence of *E. coli* in sediment between sampling events.
				2 (25%)	99%	unrelated	
Apr 2018	Raumanga	Sediment	**B2 = 5**	2 (40%)	99%	91% related to one sediment isolate at Raumanga in Aug2018	
				1 (20%)	unrelated	92% related to one water isolate at Raumanga in Aug2018	
				2 (40%)	98%	unrelated	
Aug 2018	Raumanga	Water	**B2 = 3**	1 (33%)	unrelated	92% related to one Raumanga sediment isolate from Apr 2018	
				1 (33%)	unrelated	89% related to five Raumanga water isolates from Apr2018	
Aug 2018	Raumanga	Sediment	**B2 = 5**	2 (40%)	99%	unrelated	
				1 (20%)	unrelated	91% related to one sediment isolate at Raumanga in Apr2018	

*This is the only comparison identifying genetically related subtypes in sediments between sites (Ahuroa and Hatea).

3.3.4 Hatea

Only water was collected from Hatea during 2017 and the numerically dominant *E. coli* phylogroup was B1 at 45% with no B2 identified. Six of the nine B1 isolates were related at levels of 84-91% in this water sample (Table 5). In addition, three of these isolates were related by 88% to water samples at Waitangi collected during the same event in February 2017. These similarities are not all ≥90%, and therefore, some fall below the strict criterion of clonally related. Furthermore, the six B1 isolates identified in the Waitangi sediment (2017) were not related to any in the Waitangi water (or Hatea water), perhaps suggesting some isolates may be capable of persistence and transmission via water but are less likely to reside in the sediments.

In April 2018, the water isolates at Hatea were numerically dominated by B1 (85%), whereas the sediment was dominated by B2 (90%) of isolates (Appendix D.2, Table 12). This result showed a large disconnect between the two matrices. Within the B1 grouping of water isolates there were three clonal groups and one of those groups was temporally related (89%) to a sediment isolate also collected from Hatea but later in August 2018 (Table 5). Strong evidence of naturalisation was evident for the 18 B2 isolates in the Hatea April sediments, which all had >90% similarity to each other (Figure 3).

At Hatea in August, the numerically dominant phylogroup in water and sediment was B1 (40% in each matrix), which was similar to the 2017 and April 2018 events. In the water, 88% of these B1 isolates were clonally related, whereas in the sediment, the B1 phylogroup was differentiated into four small unrelated groups of isolates. Two of the sediment isolates and seven of the water isolates were related to each other for this sampling event, and another two sediment isolates were related (>90% similarity) to 4 sediment isolates from Ahuroa in April 2018. However, 25% of isolates in the August water sample belonged to phylogroup A but this phylogroup was not identified in the sediments during 2018 at Hatea (Appendix D.2, Table 13) supporting the observation that commensal phylogroup A was detected at low frequency in sediments.

3.3.5 Ahuroa

Due to logistical difficulties during the August 2018 sampling event, samples of sediment and water were collected from Ahuroa only during the April 2018 event, (note that this location was not sampled in 2017). Phylogroup B2 was numerically dominant (65%) in the Ahuroa water isolates and belonged to two different PCR patterns as designated by the quadruplex Clermont PCR (Appendix D.2, Table 12). In the sediment sample, however, the B1 phylogroup was dominant (55%), and therefore, these were the isolates characterised by Rep-PCR. Eleven B1 isolates were subtyped and two groups of four were each related at a level of >90% similarity. In addition, one of these groups was related to two sediment isolates from Hatea collected in August 2018 (Figure 3 and Table 5) supporting spatial stability of certain B1 *E. coli* isolates.

3.3.6 Raumanga

The *E. coli* phylogroup B2 was numerically dominant in water (40%) and sediment (35%) during the April 2018 sampling at Raumanga (note that this location was not sampled in 2017). In contrast, for the August event, B1 was the dominant phylogroup in the water (42%) and the *Escherichia* Cryptic Clade V in the sediment (60%). The quadruplex Clermont PCR recognised two PCR patterns for the B2 but only the numerically dominant PCR pattern was subjected to Rep-PCR for comparison between both sampling events and matrices. In general, the B2 subtypes from April and August samples were clustered into small groups of genetically related isolates with only one group of more than 2 isolates (n = 5, 92% similarity), and these isolates were identified in water.

However, there were two occurrences of genetically related isolates between sediments (91% similarity) in April and August, and between sediment and water (92% similarity) at Raumanga at different times. These occurrences add to the body of evidence for potential *E. coli* transfers between water and sediment and the identification of temporal persistence of *E. coli* in sediment between sampling events.

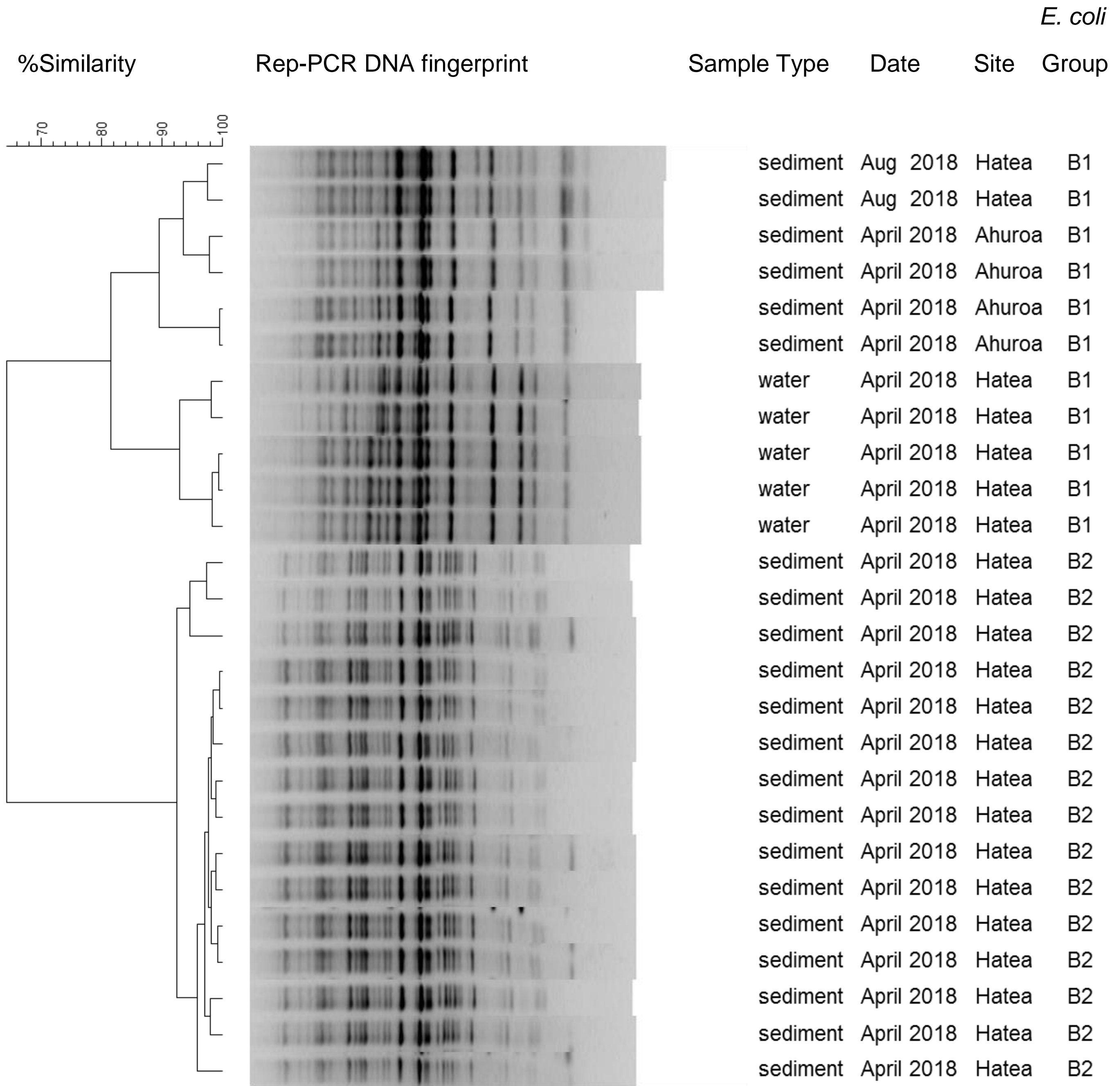


Figure 3: Selected *E. coli* isolates from phylogroups B1 and B2 identified in sediments and water at Hatea and Ahuroa during April and August 2018

3.3.7 Identification of cryptic clades in water and sediment

All *Escherichia* cryptic clade isolates from 2017 and 2018 that were identified by the quadruplex Clermont PCR were then typed into individual cryptic clade groups using the Clermont et al. (2011) PCR scheme. All cryptic clade isolates during this study for 2017 and 2018 events were classified as Clade V.

The only sample that did not contain any *Escherichia* Cryptic Clade V in the 20th February 2017 was the water sample from Hatea, otherwise 10-20% of the isolates in water and 8-15% of the isolates in the sediments belonged to Clade V (Table 3). A higher overall frequency of Clade V was identified in the 2018 sampling events with Clade V isolates ranging from 0–48% in water and 0–67% in sediment (Table 4). Similar to 2017, no Clade V was isolated at Hatea in April from either water or sediment, however, at Hatea in August, percentages were 10 and 28%, respectively.

Table 6 provides an overview of the average prevalence of Clade V *Escherichia* isolated from the water and sediments at all of the sites. Overall averages note a contribution of 12% of Clade V *Escherichia* to the concentration of *E. coli* in water and 22% in sediment. There is also a potential seasonal contribution to the concentrations of naturalised *Escherichia* where Clade V concentrations were higher in both sediment and water during the winter sampling event than either February 2017 or April 2018 events. There were too few samples collected in winter from either sediment or water to make statistically relevant statements about seasonal differences in the individual matrices. However, when the data for sediment and water were combined, using a Student t-test there was a statistically significant difference ($p = 0.012$) between the prevalence of the Clade V *Escherichia* observed in winter to prevalence in non-winter (summer and autumn samples). As noted above however, the rainfall events in February 2017 will also have contributed to the observed lower Clade V frequency. Additional sampling events would be required to test if there is a difference between seasons in Northland, and also in other more temperate NZ environments.

Table 6: Overall prevalence (%) of naturalised *Escherichia* Clade V isolates in water and sediment compared with all *Escherichia* detected at the sampling sites

Date	Average prevalence (%)	
	Water	Sediment
All samples 2017-18	11.5	21.5
February 2017	11.3	12.7
April 2018	1.3	7.5
August 2018	24.8	42.3

The highest contributions of the naturalised Cryptic Clade V to the *E. coli* concentrations in water, were at times when the Alert levels (260 CFU/100 mL) in water were not exceeded. Therefore, the occasions where Action levels (550 CFU/100 mL) were exceeded and Cryptic Clade V detected in the water are presented in Table 7. The calculation of the contribution of the non-faecal cryptic clades to that *E. coli* concentration is also estimated in the table.

Table 7: Examples where Action or Alert concentrations of *E. coli* were exceeded and the effect of the prevalence of cryptic clades determined on *E. coli* concentration in those water samples

Site	Date	*E. coli* (CFU/100 mL)	Prevalence of Clade V in water	*E. coli* (CFU/100 mL) minus Clade V contribution
Victoria	February 2017	585*	15%	497
Waipoua	February 2017	685	20%	548
Ahuroa	April 2018	365**	5%	347
Hatea	April 2018	770	0	770
Hatea	August 2018	530	10%	477

*Red highlight = Action level of 550 CFU/100 mL
**Orange highlight = Action level of 260 CFU/100 mL

4. DISCUSSION

Characterisation of *Escherichia* in water and sediments from Northland rivers revealed that during 2017, the numerically dominant *E. coli* phylogroup identified in 3 of the 4 locations selected for Rep-PCR belonged to the B1 phylogroup (Table 11). In 2018, *E. coli* phylogroups B1 and B2 were frequently isolated from water and sediment followed by the phylogrouping of “D and E” (Table 12 and Table 13). The other phylogroups were infrequently isolated and the commensal Group A (Clermont et al. 2000) was isolated in very low frequency (maximum 1 isolate/location). Group A, however, was identified in up to 25% of water isolates at any one location suggesting it may not be well suited to a non-host environment such as sediment.

The review of Tenaillon et al. (2010) outlines that phylogroup B1 is generally associated with animals, where this strain dominates the *E. coli* flora. In contrast, B2 is considered to be the dominant *E. coli* phylogroup in the human intestinal microflora of those living in developed countries (Ambrosi et al. 2019, Tenaillon et al. 2010). Phylogroups D and E, include pathogenic *E. coli* such as *E. coli* O157:H7 (Gilbert et al. 2008, Samadpour et al. 2002). Isolates belonging to B1 have been isolated from Australian waterbodies and associated with blooms of *E. coli* in eutrophic Australian waters (Power et al. 2005) as reviewed by Gordon (2013). Previous studies have also noted that the survivability of the B1 phylogroup in aquatic environments is better when compared with other phylogroups (Ratajczak et al. 2010, Walk et al. 2007). Berthe et al. (2013) studied the persistence of the various phylogroups of *E. coli* and noted that those strains that survived 4-14 days in water with a low level of faecal contamination mainly belonged to the B1 phylogroup, with reduced persistence attributed to B2 strains. However, these experiments of Berthe et al. (2013) in water microcosm studies did not include survivability in sediment.

E. coli concentrations in sediment in this study, are provided as *E. coli*/100 g dry weight for comparison between sites and studies. Future research could allow for *E. coli*/100 g wet

weight in sediments as used by other researchers (Haller et al. 2009, Lee et al. 2006, Nguyen et al. 2018). Nguyen believes wet weight of 100 g sediment is a better comparator between sediment and water. Lee et al. (2006) observed that a resuspension event of 100 mg/L suggests that a sediment concentration of 4 x 10^6 *E. coli* /100 g wet weight of sediment would be required to approximate 400 *E. coli* /100 mL in the water column above the sediment. These concentrations of *E. coli* in sediment were not exceeded during this pilot study (Table 9 and Table 10). The approximations of *E. coli* resuspension from sediments would also need to factor in the sediment type and surface properties of the bacteria as both can impact on resuspension rates (Bai et al. 2016, Liang et al. 2016).

The strongest evidence for naturalisation of *E. coli* was seen at Hatea (April 2018) for the 18 B2 clonally related isolates (>90% subtype similarity) recovered from the sediment. This high level of identity suggests that this B2 strain is not only persisting in the sediment but also replicating within the sediment. Site survey information suggests that the waters at Hatea are subjected to a mix of pollutant sources (native/exotic/pastoral/urban). The pastoral and urban elements may be the sources of the B1 and B2 isolates, which were frequently isolated from this location.

From this current study, it appears that the phylogroups of B1 and B2 followed by the grouping of “D and E” were frequently isolated from New Zealand sediments as well as water (Table 3 and Table 4). As these *E. coli* phylogroups are implicated in intestinal and extra-intestinal infections, this may raise concerns about the virulence of “naturalised” *E. coli* in sediment reservoirs, and further research may focus on whether they maintain their virulence potential for causing disease after persisting in sediments. In comparison the commensal phylogroup of A, which is not strongly associated with pathogenic potential, was identified in low frequency in sediments.

This study also sought to ascertain the contribution of the recently recognised *Escherichia* cryptic clades (Walk 2015, Walk et al. 2009) to water quality monitoring of *E. coli* as an

indicator of faecal contamination. Overall at these locations in Northland, there appears to be a consistent contribution of the naturalised *Escherichia* Cryptic Clade V isolates to the enumeration of *E. coli* in water (mean 12%) and sediment (mean 22%) when using the membrane filtration method with growth on Brilliance *E. coli* / Coliform agar (Oxoid) medium (Table 6). The prevalence observed in the sediments may suggest the sediments are acting as a habitat for the cryptic *Escherichia* species, which could potentially be transferred into the water column during resuspension events. It is not feasible to compare *E. coli* and *Escherichia* cryptic clade concentrations using the Colilert method of *E. coli* enumeration because Colilert employs a most probable number (MPN) enrichment method which may enhance the growth of certain *Escherichia* strains over others in individual Colilert wells. Therefore, in comparison to the direct plating method of membrane filtration, it would require multiple isolations from the many wells of the Colilert test to ensure the detection of all representative individual *Escherichia* subtypes present in the original water sample.

Based on the identification of only Clade V from the potential four environmental Clades (II-V) of *Escherichia* future studies of cryptic clade prevalence may want to trial different agar types for isolation, to investigate whether this increases the detection of the Clades II-IV in the NZ environment. Furthermore, Alm et al. (2011) states that the number of colonies (putative *E. coli*) chosen from a single sample for additional characterisation is an important factor in determining how many different subtypes will be identified. If only 5 colonies per sample are selected, this represents a 90% probability of identifying a subtype that is present in the sample at a frequency of 40%. Whereas, selection of 45 colonies will enable detection of those subtypes present at a frequency of 5% in the sample. Obviously, as isolation and purification of the putative *E. coli* colonies is the time and cost limiting step, then 20 colonies provide a mid-point between cost efficacy and detection of minor subtypes within a single sample. The other consideration is that multiple samplings of the same location/sediment will also increase the likelihood of identifying minor subtypes within a sample and may also provide a temporal aspect to the characterisation.

In this pilot study, there may be a seasonal response to detection of the cryptic clades of *Escherichia* with significantly higher prevalence noted in the August 2018 winter sampling event compared with February 2017 and April 2018 events. However, confounding this observation is the higher than expected cumulative rainfall, which occurred during the summer months of the 2017 – 2018 season. The first sampling events in February 2017 were impacted by preceding high rainfall (Figure 4), and therefore the proposed sampling in the 2017 - 2018 summertime was delayed till April 2018 due to difficulty in finding a 2-week window of negligible rainfall prior to sampling. The aim of the sampling programme was to minimise sediment disturbance and hence disturbance of the microbial communities in the sediment. Prior to April 2018, the many months of persistent rainfall may have contributed to the scouring out of sediments reducing the naturalised *Escherichia* cryptic clade populations in comparison to the *E. coli*. Liang et al. (2016) have noted differences in the surface properties (protein and polysaccharide components) of *E. coli* in sediment compared with those isolates in water, therefore, there may be advantages for different *Escherichia* populations in colonisation of sediments and in resuspension properties. These intrinsic differences in bacterial strains may contribute to the dominant types of phylogroups (eg B1) identified in the sediments in this study and others (Ratajczak et al. 2010, Walk et al. 2007).

During the 2017 sampling, the native bush sites of Victoria and Waipoua reported levels of elevated *E. coli* (>Alert and Action levels) with an 8-20% contribution of Cryptic Clade V in water and sediment. Furthermore, faecal source tracking analysis at these two sites in 2017 did not identify ruminant/avian/human faecal sources. The prevalence of Clade V at these two sites, however, was similar to their prevalence at non-native sites during 2017. During 2018, levels of *E. coli* in water at these two native bush sites were less than 70 CFU/100 ml, but during the winter of 2018, the Victoria site recorded the highest contributions of Clade V in water (48%) and sediment (67%) for the entire study. Recent international evidence is suggesting that Clade V may be associated with some feral animals and with avian sources

(Blyton et al. 2015, Vadnov et al. 2017). Therefore, the prevalence of this Clade V in New Zealand indigenous avian and in feral populations needs further investigation.

Isolates from Hatea and Ahuroa (April 2018) were examples where the dominant *E. coli* strain in the water (B1) was not the dominant strain identified in the sediment (B2). This shows the disconnect between these two matrices, where water contamination is transient compared with sediment populations which are likely to be more stable. These observations are supported by the study of Liang et al. (2016) who noted genetic differences between water and sediment isolates of *E. coli*. The sediment isolates had greater hydrophobicity, and enhanced polysaccharide and protein associated with extracellular polymeric substance, with these factors affecting surface properties of bacteria (Liang et al. 2016). In addition, these two locations provided an example of temporal and spatial stability where a genetically related strain of B1 was identified in the sediment of both locations but at different sampling times (April and August 2018), suggesting that this B1 strain may be better adapted to a non-host environment compared with other *E. coli* strains. At Raumanga, there were also occurrences of genetically related isolates of phylogroup B2 between sediments (91% subtype similarity) in April and August 2018, and between sediment and water (92% subtype similarity) at different times. These occurrences add to the body of evidence for potential *E. coli* transfers between water and sediment and the identification of temporal persistence of *E. coli* in sediment between sampling events. The August sampling at Raumanga also identified a high relative prevalence of the *Escherichia* Clade V (60%) in the sediment, when phylogroup B1 (42%) was dominant in the water with 21% of isolates also attributed to Clade V (Table 4). Notably, Raumaunga was categorised as urban land cover, however, similar to Victoria, the *E. coli* concentration in the water (127 CFU/100 mL) was below the Alert level when Clade V was prevalent.

This pilot study has highlighted that in certain waterways, there can be substantial (20-48% contributions) of non-*E. coli* identified in waters by *E. coli* enumeration methods. In addition, these strains of *Escherichia* reside in sediments, making them potentially available for re-

mobilisation into the water column. The highest contributions of the naturalised Cryptic Clade V to the *E. coli* concentrations in water, were, however, at times when the Alert levels were not exceeded. Therefore, based on the assumption that these cryptic clades are confounding water quality tests, the occasions where Action levels were exceeded and Cryptic Clade V detected in the water are presented in Table 7, along with the potential contribution of the cryptic clade to the concentration of *E. coli* in the water. The impacts on water management decisions from the cryptic clade contribution would be different for two of these five exceedances. It should be noted that this revised concentration (ie minus *Escherichia* cryptic clade contribution) does not include any contribution from the naturalised (faecally-derived) *E. coli* identified in the sediments in this pilot study and potentially re-suspended into the water column. If re-suspended naturalised *E. coli* were also present in this water sample, this would further reduce the *E. coli* concentration contribution from recent faecal inputs. However, caution is required here because naturalised *E. coli* may be an indicator of pathogen presence in sediments from historical/aged faecal events and potentially still represent a health risk (Devane et al. 2014).

Victoria and Waipoua are the two sites classified as land category: native bush, and they provide the two occasions where recognition of the cryptic clades would move the *E. coli* exceedance from Action to Alert level (Table 7). This pilot study is, therefore, providing preliminary evidence that further investigation of naturalised sources of potentially non-faecal *Escherichia* should be explored particularly for those sites where livestock and human impacts are not detected by FST analysis and site surveys, and yet exceedances of *E. coli* still occur.

5. CONCLUSIONS

The Northland Regional Council's (NRC) strategy was to assess faecal contamination in Northland Rivers. This report provides a methodology applicable to characterising *E. coli* subtypes and identifying naturalised *E. coli* and other *Escherichia* species that are potentially confounding water quality monitoring. This methodology will assist NRC with its strategy to effectively mitigate faecal contamination in rivers for public safety.

This Envirolink provides evidence that allows a better understanding of the proportion of naturalised historical *E. coli* compared to the proportion of recent *E. coli* contributing to bacteriological contamination in Northland's rivers. Using this information, the council will be able to establish strategies and mitigation measures to reduce input of fresh *E. coli* in waterways as well as developing an understanding of typical locations of 'historical' *E. coli* concentrations associated with non-recent faecal inputs. Furthermore, the contributions of non-*E. coli* populations of *Escherichia* to standard methods of water quality monitoring was assessed in this report. These are *Escherichia* species (the so-called "Cryptic Clades") that are not normally derived from livestock, human or avian faeces but occur in environmental habitats such as water and river sediment, and are identified as having the same biochemical signature as *E. coli* in the tests used to monitor *E. coli* concentrations in fresh water.

The outcomes of this project will assist councils to target those rivers that have elevated *E. coli* concentrations in water but low numbers of naturalised *E. coli* and/or *Escherichia,* suggesting that faecal contamination is a result of recent faecal inputs. This information will enable NRC to locate the source of faecal contamination and implement appropriate mitigating strategies and management actions. Ongoing monitoring and management will ensure continuous water quality improvements in the future.

If significant numbers of historical naturalised *E. coli* are identified as persisting at some river sites in Northland, then it may be beneficial for NRC to conduct testing to identify pathogenic *E. coli* or other pathogens such as *Cryptosporidium* in the sediments. This will help to inform the council and public of the associated health risks for recreational use of these rivers.

5.1 RECOMMENDATIONS

Where chronic elevated *E. coli* concentrations are identified during routine water quality monitoring then these sites should be first investigated by sanitary surveys and Faecal Source Tracking (FST) analyses.

When this second tier of investigation does not identify a human / livestock or avian faecal source(s) or the faecal source is not recent, an investigation of the contribution of naturalised sources of *Escherichia* to *E. coli* concentrations may be worthwhile. A high proportion of *Escherichia* cryptic clades in water/sediment, and/or genetically related strains of (faecally-derived) *E. coli* in sediment would be indicative of naturalised *Escherichia*. This report outlines the methodology for investigation of potentially naturalised *Escherichia* species from water and sediment.

The report has also highlighted that naturalised sources of *E. coli* in sediments are likely to belong to the *E. coli* phylogroups of B1, B2, D and E. Some strains in these *E. coli* phylogroups have been implicated in clinical cases of infection. The identification of naturalised sources of *E. coli*, may therefore, warrant a health risk investigation into whether other pathogen groups associated with faecal contamination such as *Campylobacter*, *Cryptosporidium*, *Giardia*, and viruses are also residing in the sediments. Some of these pathogens may persist in sediments even after the implementation of mitigation actions to remove a faecal contamination pathway.

We used membrane filtration in conjunction with growth on Brilliance *E. coli*/Coliform agar in this experiment to obtain *E. coli* counts. However, so far in this prevalence pilot study we

have not identified naturalised *Escherichia* Cryptic Clades belonging to Clades II, III and IV. Therefore, future studies could incorporate a media trial study to determine if other *E. coli*-specific isolation media are able to isolate all putative *Escherichia* cryptic clades from the NZ environment. This could establish whether Brilliance *E. coli*/Coliform agar has a bias towards isolation of only Clade V or whether there is a truly low prevalence of Clades II-IV in the NZ (Aotearoa) environment.

APPENDIX A: SAMPLING SITES

A.1 GPS COORDINATES OF SAMPLING SITES

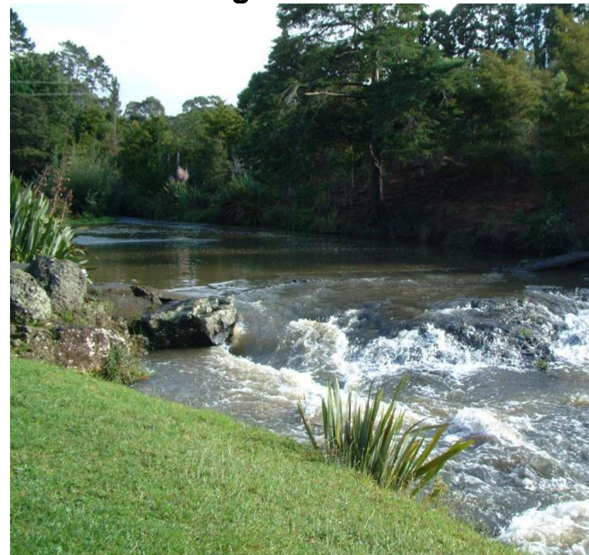
Site Name	Northing	Easting
Ahuroa at Piroa Falls	6007913	1725149
Raumanga at Raumanga Valley Park	6044187	1717608
Tirohanga at Tirohanga Road	6084784	1699502
Victoria at DOC Reserve Crossing	6108122	1639482
Waipoua at Swimming Hole	6054513	1650503
Waitangi at Wakelins	6095708	1695269
Hatea at Whangarei Falls	6050300	1720857

A.2 PHOTOS OF NORTHLAND RIVER SITES SELECTED FOR SEDIMENT AND WATER SAMPLING TO EVALUATE NATURLAISED *ESCHERICHIA* AND *E. COLI*

Ahuroa

Hatea at Whangarei Falls

Hatea at Whangarei Falls



Raumanga

Raumanga

Tirohanga

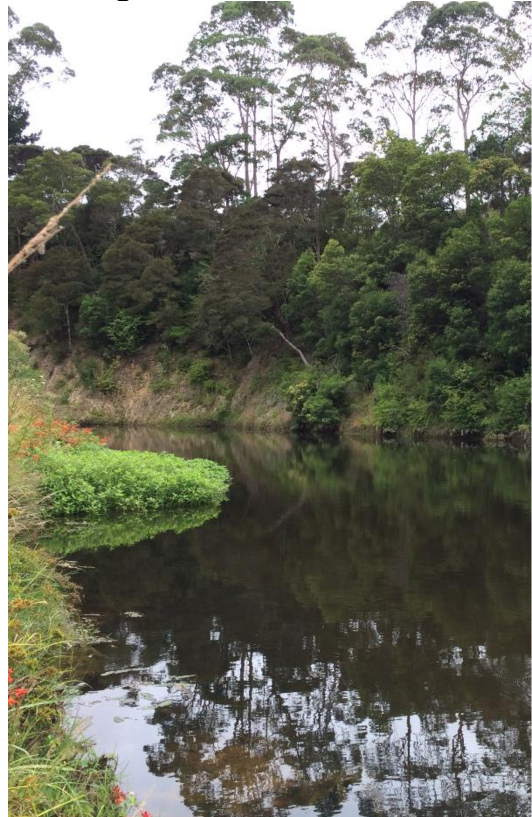
Tirohanga

Victoria

Waipoua at swimming hole

Waitangi at Wakelins

Waitangi at Wakelins



APPENDIX B: METHOD PREPARATION

B.1 CHELEX PREPARATION SOLUTION FOR CELL/COLONY PREPARATION FOR PCR ANALYSIS

- Turn on water bath to 95°C.
- Always wear gloves when handling tubes/cultures for DNA preps
- Label two sterile eppendorf tubes with the culture's laboratory ID.
- To one tube aseptically add 500 µl of 2 % Chelex* in distilled water (NB: Mix Chelex solution well before dispensing).
- To the water/Chelex-containing tube add 1-2 pure colonies of an individual putative *E. coli* isolate.
- Place in the water bath and heat for 10 minutes
- Cool at room temp for @ 1min.
- After cooling, spin in an eppendorf centrifuge maximum speed (approximately 16,000 g) for 5 minutes.
- Remove 250 µl of the supernatant and transfer to a sterile tube.
- Store at 4°C or -20°C depending on how soon the PCR/mBiT analysis will be performed.

*Preparation of Chelex solution

Make up final 2% working solution by adding 2 g Chelex™ 100 Sodium form (Sigma Cat#C7901) into 100 mL sterile distilled water and adjust to pH 7.2 +/- 0.2. Do not autoclave. Add magnetic stir bar to bottle, always have solution stirring when dispensing the Chelex solution.

APPENDIX C: RAINFALL DATA AND WATER QUALITY PARAMETERS

C.1 RAINFALL DATA FOR 2017 AND 2018 SAMPLING EVENTS

Table 8: Cumulative rainfall data for sampling events in 2017 and 2018. Collection of water and sediment occurred on 20th February in 2017, and on 9th April and 6th August in 2018.

Site	February 2017 Rainfall (mm) 1st Feb to 21st Feb	April 2018 Rainfall (mm) 19th March to 9th April	August 2018 Rainfall (mm) 23rd July to 6th August
Ahuroa	66.5	79.0	10.5
Hatea	107.0	79.0	9.0
Raumanga	76.6	64.0	5.6
Tirohanga	148.5	29.0	9.5
Victoria	59.5	41.5	38.0
Waipoua	102.5	34.5	27.5
Waitangi	99.0	48.5	1.5

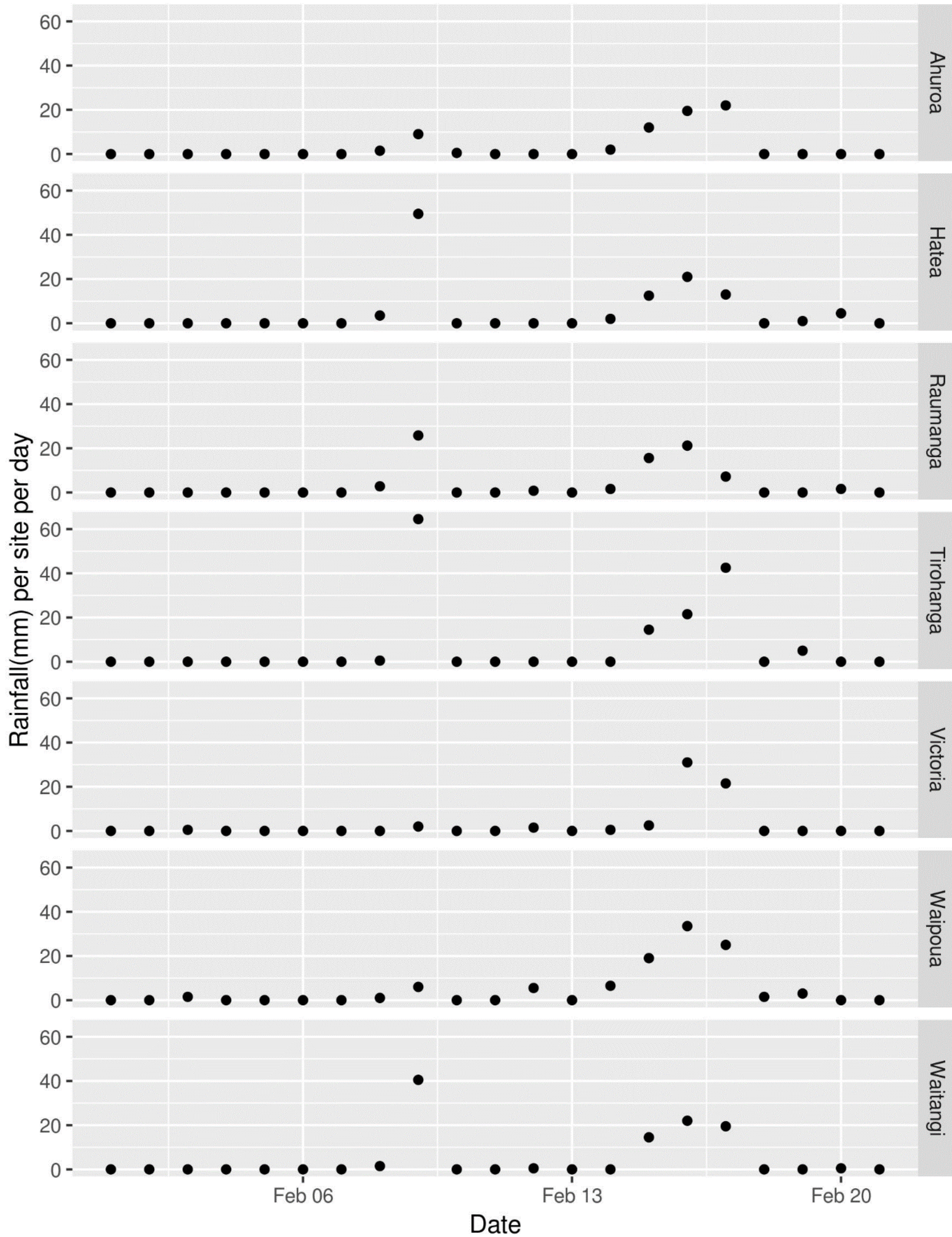


Figure 4: February 2017 rainfall

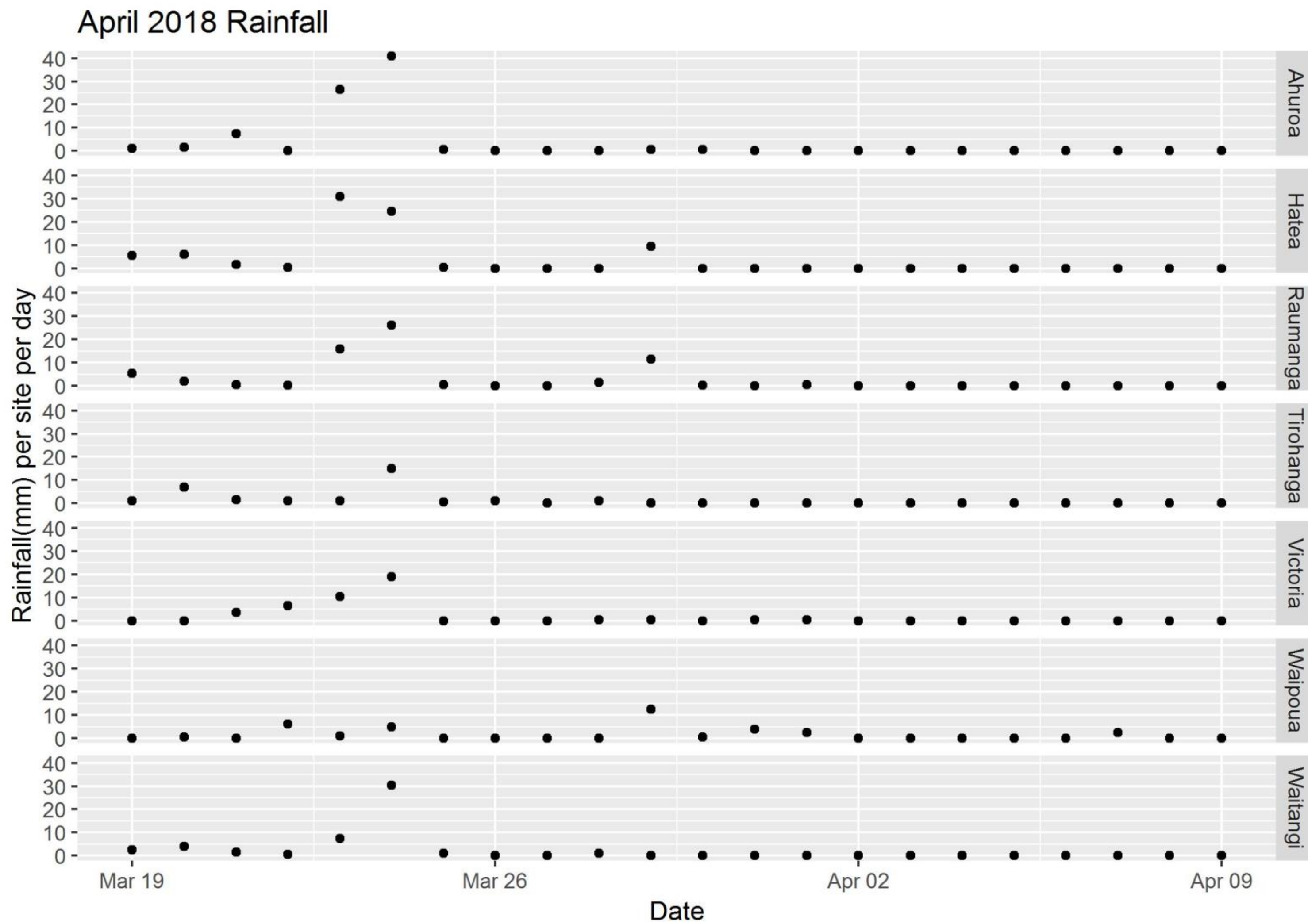


Figure 5: Rainfall during March to April 2018

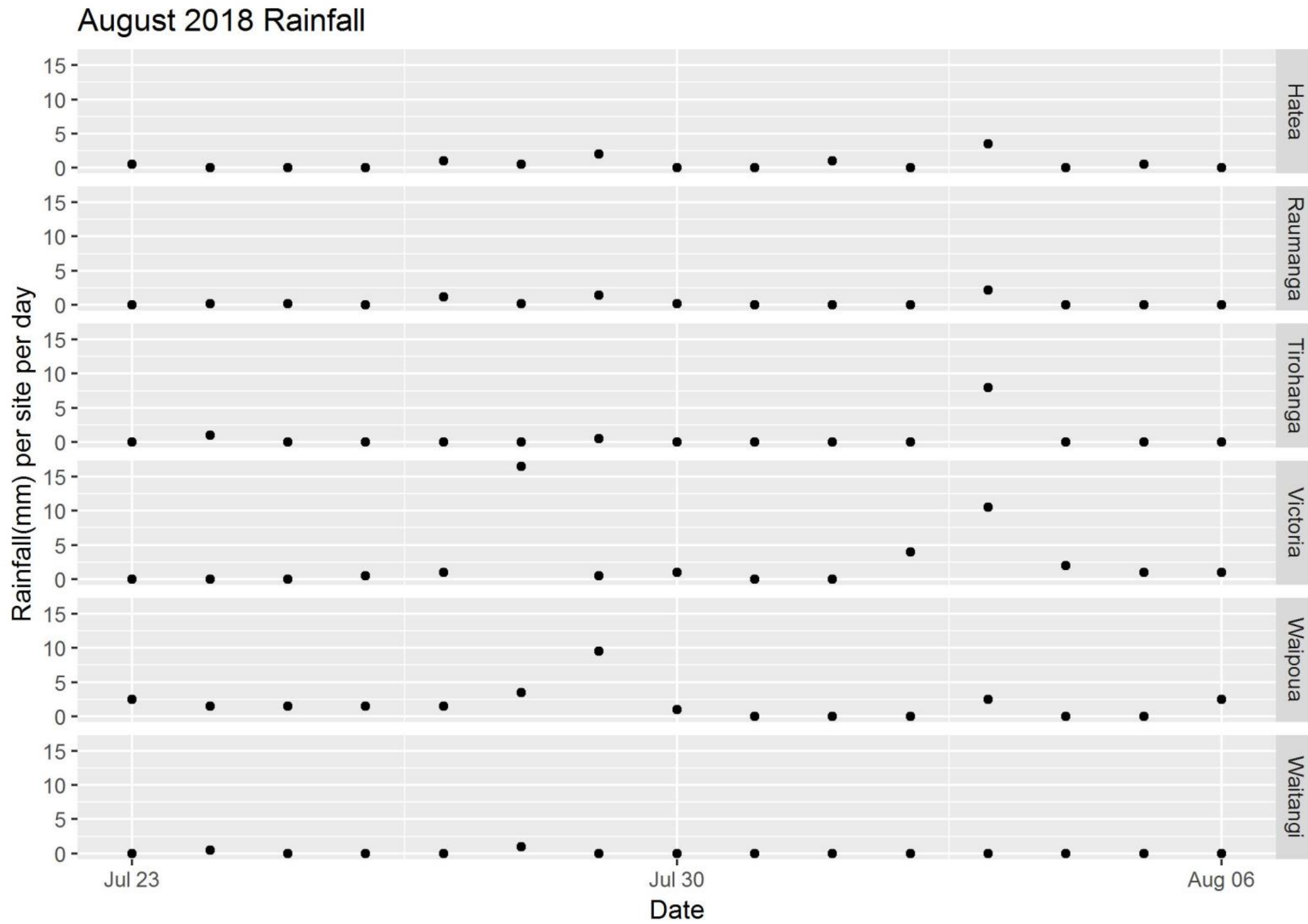


Figure 6: Rainfall during July to August 2018. Ahuroa was not sampled during the August 2018 event.

C.2 WATER QUALITY PARAMETERS

Location		Tirohanga		Victoria		Waitangi		Waipoua		Hatea		Raumanga		Ahuroa at Piroa Falls	
Variable	Units	April	August	April	August	April	August	April	August	April	August	April	August	April	August
DO%	%	107.3	94.6	99.4	99.0	96.6	95.7	90.9	91.2	92.4	85.6	102.9	102.8	103.2	N/A
pH		6.26	6.94	7.77	7.5	6.99	8.07	7.48	6.94	7.42	7.49	8.11	8.07	7.71	N/A
DO	mg/L	10.71	10.73	9.62	10.63	9.2	10.67	9.16	10.51	9.12	9.42	10.23	11.08	10.42	N/A
Temp.	celsius	15.4	9.9	16.8	12.1	18.1	10.6	15.0	9.3	16.0	11.1	15.7	12.0	14.9	N/A
SPC	uS/cm	81.1	78.8	82.7	148.5	153.6	114.6	106.8	99.9	199.9	187.1	203.7	202.2	152.1	N/A

APPENDIX D: *E. COLI* CONCENTRATIONS

D.1 *E. COLI* CONCENTRATIONS IN SEDIMENT AND WATER

Table 9: *E. coli* enumeration in sediment and water during 2017

Site ID	ESR lab Number	Date Received	Sample type	Sampling location	*E. coli* /100 mL / 100 g dry weight of sediment
101752	CMB170206	14/02/2017	water	Waitangi at Wakelins	215
	CMB170209	14/02/2017	Sediment		600
	CMB170232	20/02/2017	water		47
	CMB170236	20/02/2017	Sediment		24,200
	CMB170237	20/02/2017	Sediment		1,500
104908	CMB170207	14/02/2017	water	Victoria at DOC reserve crossing	445
	CMB170210	14/02/2017	Sediment		5,800
	CMB170233	20/02/2017	water		585
	CMB170238	20/02/2017	Sediment		9,800
108613	CMB170208	14/02/2017	water	Waipoua at swimming hole	40
	CMB170211	14/02/2017	Sediment		5,400
	CMB170235	20/02/2017	water		685
	CMB170239	20/02/2017	Sediment		1,500
105972	CMB170234	20/02/2017	water	Hatea at Whangarei Falls	255

Table 10: *E. coli* enumeration in sediment and water during April and August 2018.

Site ID	ESR lab Number	Date Received	Sample type	Sampling location	*E. coli* /100 mL / 100 g dry weight of sediment
101752	CMB180677	10/04/2018	water	Waitangi at Wakelins	84
	CMB180678	10/04/2018	sediment		300
	CMB181338	7/08/2018	water		169
	CMB181344	7/08/2018	sediment		1,000
104908	CMB180673	10/04/2018	water	Victoria at DOC reserve crossing	62
	CMB180674	10/04/2018	sediment		21,600
	CMB181341	7/08/2018	water		63
	CMB181347	7/08/2018	sediment		1,300
108613	CMB180671	10/04/2018	water	Waipoua at swimming hole	69
	CMB180672	10/04/2018	sediment		800
	CMB181342	7/08/2018	water		47
	CMB181348	7/08/2018	sediment		2,000
105972	CMB180679	10/04/2018	water	Hatea at Whangarei Falls	770
	CMB180680	10/04/2018	sediment		60,700
	CMB181339	7/08/2018	water		530
	CMB181345	7/08/2018	sediment		6,500
102252	CMB180675	10/04/2018	water	Tirohanga at Tirohanga Road	102
	CMB180676	10/04/2018	sediment		3,600
	CMB181337	7/08/2018	water		148
	CMB181343	7/08/2018	sediment		400
317597	CMB180681	10/04/2018	water	Ahuroa at Piroa Falls	365
	CMB180682	10/04/2018	sediment		7,300
103246	CMB180683	10/04/2018	water	Raumanga at Raumanga Valley Park	142
	CMB180684	10/04/2018	sediment		5,200
	CMB181340	7/08/2018	water		127
	CMB181346	7/08/2018	sediment		2,500

D.2 CHARACTERISATION OF *ESCHERICHIA* ISOLATES INTO *E. COLI* PHYLOGROUPS AND *ESCHERICHIA* CRYPTIC CLADES

The data from February 2017 is taken with permission from a 2017 report to NRC.

Table 11: Characterisation of *E. coli* isolates and Escherichia Cryptic Clades from selected water and sediment samples collected 19 February 2017. Only the samples from 19th February were selected for subtyping based on higher concentrations of *E. coli* in water and/or sediment observed during this second sample collection.

Site ID	ESR Sample Number	Sample Type	Location	No. *Esche-richia* isolates tested	*E. coli* phylogroup							Not *E. coli*	Unknown *E. coli*	¥Cryptic Clades III-V	*Cryptic Clades I or II
					B1	B2	A	A or C	D or E	E or *Clade I	F				
101752	CMB170232	Water	Waitangi	20	10	4	ND	1	1	2	ND	ND	ND	2	ND
	CMB170236	Sediment		20	6	4	ND	1	6	ND	ND	ND	ND	3	ND
104908	CMB170233	Water	Victoria	20	ND	3/6**	ND	ND	7	ND	ND	ND	1	3	ND
	CMB170238	Sediment		20	1	3/4**	ND	1	6	1	ND	ND	1	3	ND
108613	CMB170235	Water	Waipoua	20	2	5/4**	ND	ND	4	ND	ND	1	ND	4	ND
	CMB170239	Sediment		12	4	2	ND	ND	3	ND	ND	2	ND	1	ND
105972	CMB170234	Water	Hatea	20	9	ND	1	2	3	2	ND	2	1	ND	ND

*Cryptic Clades I-V belong to the group of so-called "naturalised" cryptic clades of *Escherichia*, however, Cryptic Clade I is now thought to be a commensal in the intestinal tract of some animals;

** the B2 phylogroup can have several DNA PCR patterns, for example, in sample CMB170235, five of the B2 had one PCR pattern and four had a second PCR pattern, therefore although there are nine isolates in this phylogroup, they will not all be genetically related

¥All Cryptic Clade grouping III-V identified by Clermont et al. (2013) were identified as belonging to Cryptic Clade V by the Clermont et al. (2011) PCR.

B2 phylogroup can have several PCR patterns, for example, in sample CMB170235, five of the B2 had one PCR pattern and four had a second PCR pattern, therefore, although there are nine isolates in this phylogroup, they will not all be genetically related

Table 12: Characterisation of *E. coli* isolates and *Escherichia* Cryptic Clades from selected water and sediment samples collected during April 2018

Site ID	ESR Sample Number	Sample Type	Location	Isolate count	B1	B2	A	A or C	D or E	E or Clade I	F	Unknown *E. coli*	**Cryptic Clades III-V	Cryptic Clades I or II
103246	CMB180683	Water	Raumanga	**20**	4	8	0	2	4	0	0	2	0	0
103246	CMB180684	Sediment	Raumanga	**20**	4	7 (5/2)	0	1	5	0	1	0	2	0
104908	CMB180673	Water	Victoria	**20**	1	10 (5/5)	0	0	5	0	2	2	0	0
104908	CMB180674	Sediment	Victoria	**20**	ND	11 (3/8)	0	0	6	0	1	0	2	0
105972	CMB180679	Water	Hatea	**20**	17	ND	0	0	3	0	0	0	0	0
105972	CMB180680	Sediment	Hatea	**20**	2	18	0	0	0	0	0	0	0	0
317597	CMB180681	Water	Ahuroa	**20**	2	13 (6/7)	1	1	2	0	0	0	1	0
317597	CMB180682	Sediment	Ahuroa	**20**	11	2 (1/1)	1	ND	3	0	1	0	2	0

¥ Two of the *E. coli* isolates were not able to be phylotyped by the Clermont et al. (2013) quadruplex PCR; All isolates in 2018 were identified as either *E. coli* or *Escherichia* Clades (therefore not *E. coli* column was removed)

Table 13: Characterisation of *E. coli* isolates and *Escherichia* Cryptic Clades from selected water and sediment samples collected during August 2018

Site ID	ESR Sample Number	Sample Type	Location	Isolate count	B1	B2	A	A or C	D or E	E or Clade I	F	Unknown *E. coli*	*Cryptic Clades III-V	Cryptic Clades I or II
103246	CMB181340	Water	Raumanga	**19**	8	4 (3/1)	0	0	3	0	0	0	4	0
103246	CMB181346	Sediment	Raumanga	**20***	1	6 (5/1)	0	0	0	0	0	1	12	0
104908	CMB181341	Water	Victoria	**25**	1	10 (9/1)	0	0	2	0	0	0	12	0
104908	CMB181347	Sediment	Victoria	**9**	0	2	0	0	1	0	0	0	6	0
105972	CMB181339	Water	Hatea	**20**	8	2 (1/1)	5	0	1	1	1	0	2	0
105973	CMB181345	Sediment	Hatea	**25**	10	3	0	1	4	0	0		7	0
101752	CMB181338	Water	Waitangi	**20**	5	2	4	2	3	0	0	0	4	0
101752	CMB181344	Sediment	Waitangi	**7**	2	2	0	0	2	0	0	0	1	0

*1 isolate in CMB181346 was not characterized as *E. coli* or an *Escherichia* cryptic clade

**All Cryptic clade isolates typed as Clade III-V by Clermont et al. (2013) were confirmed as Clade V when using singleplex Clade PCRs

6. REFERENCES

Alm, E.W., Walk, S.T. and Gordon, D.M. (2011) Population Genetics of Bacteria. Walk, S.T. and Feng, P.C.H. (eds), ASM Press, ProQuest Ebook Central, .

Ambrosi, C., Sarshar, M., Aprea, M.R., Pompilio, A., Di Bonaventura, G., Strati, F., Pronio, A., Nicoletti, M., Zagaglia, C., Palamara, A.T. and Scribano, D. (2019) Colonic adenoma-associated *Escherichia coli* express specific phenotypes. Microbes Infect.

APHA 9222- (2017) Standard Methods For the Examination of Water and Wastewater, American Public Health Association, Washington D.C., USA.

APHA (2005) Standard methods for the examination of water and wastewater, American Public Health Association, Washington D.C., USA.

Bai, H.J., Cochet, N., Pauss, A. and Lamy, E. (2016) Bacteria cell properties and grain size impact on bacteria transport and deposition in porous media. Colloids and Surfaces B-Biointerfaces 139, 148-155.

Berthe, T., Ratajczak, M., Clermont, O., Denamur, E. and Petit, F. (2013) Evidence for coexistence of distinct *Escherichia coli* populations in various aquatic environments and their survival in estuary water. Applied and Environmental Microbiology 79(15), 4684-4693.

Blyton, M.D., Pi, H., Vangchhia, B., Abraham, S., Trott, D.J., Johnson, J.R. and Gordon, D.M. (2015) Genetic structure and antimicrobial resistance of *Escherichia coli* and cryptic clades in birds with diverse human associations. Applied and Environmental Microbiology 81(15), 5123-5133.

Clermont, O., Bonacorsi, S. and Bingen, E. (2000) Rapid and simple determination of the *Escherichia coli* phylogenetic group. Applied and Environmental Microbiology 66(10), 4555-4558.

Clermont, O., Christenson, J.K., Denamur, E. and Gordon, D.M. (2013) The Clermont *Escherichia coli* phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. Environmental Microbiology Reports 5(1), 58-65.

Clermont, O., Gordon, D.M., Brisse, S., Walk, S.T. and Denamur, E. (2011) Characterization of the cryptic *Escherichia* lineages: rapid identification and prevalence. Environ Microbiol 13(9), 2468-2477.

Devane, M.L. (2015) The sources of “natural” microorganisms in streams, Institute of Environmental Science and Research (ESR) Ltd., Envirolink Report 1533 ESRC269.

Devane, M.L., Moriarty, E.M., Wood, D., Webster-Brown, J. and Gilpin, B.J. (2014) The impact of major earthquakes and subsequent sewage discharges on the microbial quality of water and sediments in an urban river. Science of the Total Environment 485-486C, 666-680.

Gilbert, M., Monk, C., Wang, H.L., Diplock, K. and Landry, L. (2008) Screening policies for daycare attendees: lessons learned from an outbreak of E. coli O157:H7 in a daycare in Waterloo, Ontario. Canadian Journal of Public Health. Revue Canadienne de Sante Publique 99(4), 281-285.

Gordon, D.M. (2013) *Escherichia coli* (Second Edition). Donnenberg, M.S. (ed), pp. 3-20, Academic Press, Boston.

Haller, L., Poté, J., Loizeau, J.-L. and Wildi, W. (2009) Distribution and survival of faecal indicator bacteria in the sediments of the Bay of Vidy, Lake Geneva, Switzerland. Ecological Indicators 9(3), 540-547.

Jang, J., Hur, H.-G., Sadowsky, M.J., Byappanahalli, M.N., Yan, T. and Ishii, S. (2017) Environmental *Escherichia coli*: ecology and public health implications—a review. Journal of Applied Microbiology 123(3), 570-581.

Lee, C.M., Lin, T.Y., Lin, C.C., Kohbodi, G.A., Bhatt, A., Lee, R. and Jay, J.A. (2006) Persistence of fecal indicator bacteria in Santa Monica Bay beach sediments. Water Research 40(14), 2593-2602.
Liang, X., Liao, C.Y., Thompson, M.L., Soupir, M.L., Jarboe, L.R. and Dixon, P.M. (2016) *E. coli* surface properties differ between stream water and sediment environments. Frontiers in Microbiology 7.
Ministry for the Environment and Ministry of Health (2003) Microbiological Water Quality guidelines for Marine and Freshwater Recreational Areas. Health, M.f.t.E.a.M.o. (ed), p. 155, Ministry for the Environment and Ministry of Health, Wellington.
Mohapatra, B.R. and Mazumder, A. (2008) Comparative efficacy of five different rep-PCR methods to discriminate Escherichia coli populations in aquatic environments. Water Science and Technology 58(3), 537-547.
Nguyen, K.H., Senay, C., Young, S., Nayak, B., Lobos, A., Conrad, J. and Harwood, V.J. (2018) Determination of wild animal sources of fecal indicator bacteria by microbial source tracking (MST) influences regulatory decisions. Water Research 144, 424-434.
Power, M.L., Littlefield-Wyer, J., Gordon, D.M., Veal, D.A. and Slade, M.B. (2005) Phenotypic and genotypic characterization of encapsulated *Escherichia coli* isolated from blooms in two Australian lakes. Environ Microbiol 7(5), 631-640.
R Core Team (2018) R: A language and environment for statistical computing, https://www.R-project.org/, Vienna, Austria.
Rademaker, J.L.W. and de Bruijn, F.J. (1997) DNA Markers: Protocols, Applications, and Overviews. Caetano-Anoll´es, G. and Gresshoff, P.M. (eds), pp. 151–171, John Wiley & Sons, New York.
Rademaker, J.L.W., Louws, F.J., Versalovic, J. and de Bruijn, F.J. (2004) Molecular microbial ecology manual. Kowalchuk, G.A., de Bruijn, F.J., Head, I.M.A., A.D. and van Elsas, J.D. (eds), pp. 1-33, Kluwer Academic Publishers, Dordrecht.
Ratajczak, M., Laroche, E., Berthe, T., Clermont, O., Pawlak, B., Denamur, E. and Petit, F. (2010) Influence of hydrological conditions on the *Escherichia coli* population structure in the water of a creek on a rural watershed. BMC Microbiology 10(1), 1-10.
Samadpour, M., Stewart, J., Steingart, K., Addy, C., Louderback, J., McGinn, M., Ellington, J. and Newman, T. (2002) Laboratory investigation of an *E. coli* O157:H7 outbreak associated with swimming in Battle Ground Lake, Vancouver, Washington. Journal of Environmental Health 64(10), 16-20, 26, 25.
Tenaillon, O., Skurnik, D., Picard, B. and Denamur, E. (2010) The population genetics of commensal *Escherichia coli*. Nature Rev Microbiol 8(3), 207-217.
Vadnov, M., Barbič, D., Žgur-Bertok, D. and Erjavec, M.S. (2017) *Escherichia coli* isolated from feces of brown bears (*Ursus arctos*) have a lower prevalence of human extraintestinal pathogenic *E. coli* virulence-associated genes. Canadian Journal of Veterinary Research 81(1), 59-63.
Versalovic, J., Schneider, M., de Bruijn, F.J. and Lupski, J.R. (1994) Genomic fingerprinting of bacteria with repetitive sequence based polymerase chain reaction Methods Mol Cell Biol 5, 25-40.
Walk, S.T. (2015) The "Cryptic" *Escherichia*. EcoSal Plus 6(2).
Walk, S.T., Alm, E.W., Calhoun, L.M., Mladonicky, J.M. and Whittam, T.S. (2007) Genetic diversity and population structure of *Escherichia coli* isolated from freshwater beaches. Environmental Microbiology 9(9), 2274-2288.
Walk, S.T., Alm, E.W., Gordon, D.M., Ram, J.L., Toranzos, G.A., Tiedje, J.M. and Whittam, T.S. (2009) Cryptic lineages of the genus *Escherichia*. Applied and Environmental Microbiology 75(20), 6534-6544.

INSTITUTE OF ENVIRONMENTAL
SCIENCE AND RESEARCH LIMITED

Kenepuru Science Centre
34 Kenepuru Drive, Kenepuru, Porirua 5022
PO Box 50348, Porirua 5240
New Zealand
T: +64 4 914 0700 **F:** +64 4 914 0770

Mt Albert Science Centre
120 Mt Albert Road, Sandringham, Auckland 1025
Private Bag 92021, Auckland 1142
New Zealand
T: +64 9 815 3670 **F:** +64 9 849 6046

NCBID – Wallaceville
66 Ward Street, Wallaceville, Upper Hutt 5018
PO Box 40158, Upper Hutt 5140
New Zealand
T: +64 4 529 0600 **F:** +64 4 529 0601

Christchurch Science Centre
27 Creyke Road, Ilam, Christchurch 8041
PO Box 29181, Christchurch 8540
New Zealand
T: +64 3 351 6019 **F:** +64 3 351 0010

www.esr.cri.nz