Science for Communities

Quantitative Microbial Risk Assessment

Phase 2.1 – Initial Data Collection and Recommendations

Authors:

Margaret Leonard (ESR)

Beverley Horn (ESR)

Bridget Armstrong (ESR)

Iain Haysom (ESR)

Carla Eaton (ESR)

Kirstin Thom (ESR)

Beth Robson (ESR)

Alice Chechi (ESR)

Anthony Pita (Massey University)

Brent Gilpin (ESR)

Sarah Coxon (ESR)

Paula Scholes (ESR)

Bindu Priya (ESR)

Angela Cornelius (ESR)

Ashley Orton (ESR)

Susan Lin (ESR)

David Hayman (Massey University)

PREPARED FOR: **Ministry for the Environment**

CLIENT REPORT No: **FW21019**

REVIEWED BY: **Megan Devane**

ACKNOWLEDGEMENTS

We would like to acknowledge and thank the council staff who undertook the sampling, braving some unpleasant weather and the council managers who have supported this project and made the resources available.

We would like to thank staff from the following regional, district and city councils, for their time and efforts in undertaking sample collection for this study.

Auckland Council
Bay of Plenty Regional Council
Environment Canterbury
Environment Southland
Gisborne District Council
Greater Wellington Regional Council
Hawkes Bay Regional Council
Horizons Regional Council
Marlborough District Council
Nelson City Council
Northland Regional Council
Tasman District Council
Taranaki Regional Council
Waikato Regional Council
West Coast Regional Council

We would also like to thank iwi and hapū who generously gave their time and knowledge of their awa. We acknowledge the generosity of the hapū who have been working on different aspects of cultural assessment of their awa. We particularly thank the iwi and hapū who helped with sampling and participated in the cultural assessments and kaupapa Māori monitoring from Ngāti Tama ki Te Waipounamu Trust, Manawhenua ki Mōhua, Ngāti Kahungunu ki Wairarapa, Pukerangiora Hapū, Otaraua hapū and Te Kotahitanga o Te Atiawa and Mel McColgan.

Manager	Peer reviewer	Author
		
Wim Nijhof	Megan Devane	Margaret Leonard
Group Leader, Human and Ecological Health	Senior Scientist, Health & Environment	Senior Scientist, Health & Environment

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

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

The presence of microbial pathogens (bacteria, viruses, and protozoa) in freshwaters poses a health risk to those using the water for drinking, food gathering, swimming and other activities which involve contact such as kayaking or waka ama. The majority of waterborne pathogens that cause human illness are associated with human and/or animal faeces (Gerba 2009). Contamination of waterways with faecal pollution can cause gastrointestinal illness, respiratory illnesses, and skin infections.

The 2003 Ministry for the Environment (MfE) and Ministry of Health (MoH) Microbiological Water Quality Guidelines for Marine and Freshwater Recreational Areas (the Guidelines) (MfE 2003) were developed to help water managers monitor, report on, and control the public health risk posed by microbiological contamination of recreational waters. The data is based on a survey undertaken over 20 years ago (McBride et al 2002), during which time there have been significant changes to land use, waste management practices and analytical techniques.

This project is one of a series of projects that have been undertaken since 2018 to refresh the science on which the recreational water quality guidelines are based, using current data (Milne et al 2018, Gilpin et al 2018, Lake et al 2018, Horn et al 2018, Moriarty et al 2018). A pilot study was undertaken in 2020 which refined the logistics, methods, and costs for the full study.

The objective of this Phase 2.1 project is to:

- confirm and extend the number of sites that are monitored to ensure there is good geographical representation, covering different land uses.
- engage with iwi to select culturally significant sites and to incorporate cultural health assessment (cultural health index (CHI) or similar).
- commence sample collection and analysis as part of the full study to build the database for the revised Quantitative Microbial Risk Assessment.

Councils were highly supportive of the sampling programme which enabled 256 samples to be collected and analysed over a 16-week period. Due to the short timeframe for planning, these samples were drawn mostly from routine weekly/fortnightly recreational water quality monitoring samples taken over the summer and monthly State of the Environment samples with some additional samples being collected where resources permitted. Thirty rivers and one lake were sampled and analysed for the pathogens *Campylobacter*, *Salmonella*, STEC (shiga-toxin producing *E. coli*), *Giardia* and *Cryptosporidium*. Engagement with iwi supported cultural health assessment at four awa and/or tributaries.

Faecal source tracking (FST) was used to assess the prevalence of ruminant, human and avian faecal contamination in each sample. The results showed there can be a mixture of faecal sources between sampling events within sites and at six sites the observed land use never matched the dominant faecal source identified by FST. Overall avian faecal sources were identified as a dominant faecal source. Faecal contamination, therefore, needs to be assessed using FST data, not observed land use.

Most sites were selected based on historical elevated concentrations of *E. coli* in order to target the pathogens. The results are therefore not reflective of water quality in general across New Zealand. The prevalence of the bacterial pathogens was similar to the pilot study, but there were more samples with elevated *Campylobacter* concentrations. In the pilot study (Leonard et al 2020) it was determined that the pathogen which was likely to cause infection was *Campylobacter*, and at concentrations above 10 MPN/100 mL there is the

greater potential for infection if ingested. Based on the prevalence of *Campylobacter* above 10 MPN/100 mL in the pilot study, it was estimated that 1040 samples would need to be taken to obtain 30 such samples from each of the four land uses. Collection of 256 samples during this Phase 2.1 project indicates good progress has been made in building the database for the QMRA. However, only two samples associated with the human faecal source grouping had *Campylobacter* concentrations above 10 MN/100 mL. Therefore, additional Urban sites have been added to the site list to increase the likelihood of meeting the criteria and this will be assessed as the results become known. Low prevalence of *Salmonella* and STEC confirmed the modified methodology of presence/absence and serotyping was an effective use of resources. However, as concentrations were low there was not good agreement between the quantitative PCR and presence/absence tests.

Engagement with iwi on four cultural health assessments highlighted the need to allow significant time to start communications and project design at the pre-planning stage. It was also evident that supporting a range of kaupapa Māori monitoring tools will better meet the needs and priorities of iwi and hapū. The more formalised Cultural Health Index (CHI) may not be required nor appropriate. The process needs to align with the priorities of iwi, hapū and whanau. Sites were assessed to determine if they were culturally significant and engagement with iwi increased the number of such sites to 33% to the total.

The list of 32 sites is presented in Table 5 and includes 11 sites of cultural significance. It is proposed that samples be taken at each site on a three weekly rotation to be confirmed with councils. At critical points data will be assessed to confirm sites and analyses.

The following recommendations are given below:

- Continue to plan for collection of 1040 samples across all faecal source groups with additional sites likely to have human FST
- A routine sampling pattern based on a three-week rotation using the listed sites.
- Continue to analyse *Salmonella* and STEC by presence/absence supported by qPCR.
- Ensure there is a period of sampling in the spring to confirm if the other target pathogens (*Salmonella*, STEC, *Cryptosporidium*, *Giardia*) are likely to be significant.
- Extensive time is needed for iwi engagement to understand their priorities and co-design an approach. Cultural assessment needs to align with iwi and hapū priorities and their tikanga, whether that is kaupapa Māori monitoring or CHI. This will support iwi and hapū to apply and build capacity in mātauranga Māori in the most appropriate way.
- At critical time points assess data to ensure sites fit with the aims of the project and to assess the need for ongoing monitoring of *Salmonella*, STEC, *Cryptosporidium* and *Giardia*. It is possible that a smaller dataset may provide sufficient statistical robustness.

1. INTRODUCTION

1.1 BACKGROUND

The presence of microbial pathogens (bacteria, viruses, and protozoa) in freshwaters poses a health risk to people using the water for drinking, food gathering, swimming and other activities which involve contact such as kayaking or waka ama. The majority of waterborne pathogens that cause human illness, including *Campylobacter*, *Salmonella*, enteric viruses, *Giardia* and *Cryptosporidium*, are associated with human and/or animal faeces (Gerba 2009). Contamination of waterways with faecal material may result from the discharge of inadequately treated sewage, leaking sewage pipes, combined sewage-stormwater discharges, septic tank discharges or leaks, run-off from urban and/or agricultural land, and direct deposition from farm or wild animals. Health effects resulting from contact with contaminated water include gastrointestinal illness, respiratory illnesses, and skin infections.

The 2003 Ministry for the Environment (MfE) and Ministry of Health (MoH) Microbiological Water Quality Guidelines for Marine and Freshwater Recreational Areas (the Guidelines) were developed to help water managers monitor, report on and control the public health risk posed by microbiological contamination of recreational waters (MfE 2003). The freshwater component of the Guidelines uses the indicator organism *E. coli* to assess water quality, with numeric guideline values developed from the findings of the 1998-2000 Freshwater Microbiology Research Programme (FMRP) (McBride et al 2002). The FMRP included a nationwide survey of microbial water quality, monitoring 25 sites representing different land uses and associated faecal impacts (dairy farming, beef and sheep farming, municipal, wildfowl and forested/undeveloped), with water samples collected fortnightly for 15 months, and analysed for 10 pathogens and indicators.

The Guidelines were developed by matching (i) the percentiles of the risk of campylobacteriosis illness derived from a Quantitative Microbial Risk Assessment[1] (QMRA) using the *Campylobacter* data recorded by the study, with (ii) the percentiles of *E. coli* concentrations recorded by the study. *Campylobacter* was chosen because this pathogen was most frequently detected in the FMRP study and a medium correlation was observed with the indicator *E. coli* using Spearman's Rank Correlation Coefficient.

A previous project (Moriarty et al 2018) reviewed the science necessary to design and inform a QMRA Study and designed a QMRA study suitable for updating the FMRP. The proposed Freshwater QMRA study has three objectives:

- To complete a nationally representative survey of freshwater in New Zealand that provides quantitative data on the concentration of a range of pathogenic microorganisms as well as indicator microorganisms.
- To supplement these data with information on potential contamination sources (faecal source tracking) and ancillary environmental data from sampling sites.
- To incorporate the survey and supplemental data into a QMRA providing human health risk estimates from recreational activity exposures to pathogens, and to determine the relationships between pathogen and indicator presence and concentration.

[1] A QMRA is a framework to combine information on the particular pathogen(s) and the potential dose or exposure (a function of the concentration of pathogens in the water and the volume of water that might be ingested during recreation), to estimate the risk of infection and illness.

1.2 PILOT STUDY

A pilot study was conducted during February-March 2020 at 16 sites (Leonard et al 2020). The Executive Summary of this study is provided in Appendix A. Briefly, the seven-week pilot study identified that *Campylobacter* was present in 10% of samples at concentrations greater than 10 MPN/100 mL, which was estimated to potentially be hazardous to public health if ingested. *Campylobacter* was also the target pathogen identified in the FMRP (McBride et al 2002). The pathogens *Giardia*, *Cryptosporidium*, *Salmonella*, and STEC were also present but in low concentrations. While *Giardia* was the most prevalent pathogen in the pilot study, the very low concentrations measured may not pose a significant health risk. Based on the findings of the pilot study, changes were made to the determine the scope of the project, modify methodologies and refine sampling logistics.

The original FMRP study by McBride et al (2002) highlighted that land use was indicative of the concentrations and types of pathogens. Four land uses categories were applied in the pilot study: urban, dairy, sheep & beef and wildfowl/natural. However, faecal source tracking (FST), which wasn't available during the FMRP, showed that sites contained faecal contamination from a mixture of sources. A large database would allow statistical analysis to determine if there was an association between land use and pathogen presence and concentration. Rainfall was also identified as a potential indicator of contamination, as runoff is a transport pathway for faecal contamination into water bodies.

The steps recommended for phase 2 are given below.

- Collect 1040 samples from 40 rivers monthly over two years.
- Use four categories of land use - sheep & beef, dairy, urban and a combined natural/wildfowl category.
- Collect 260 samples per land use category.
- Broaden the criteria for selection so that iwi can select sites of cultural significance which cover these land uses.
- Engage with iwi and councils to increase the number of sites from the 16 analysed in the pilot study to 40, using the 30 sites selected in 2018 (Milne et al 2018) as the basis, while recognising that sites selected may differ to match sampling resources and iwi may choose different sites.
- Engage iwi in site selection to ensure that sites of cultural significance are included.
- Conduct a cultural health assessment such as Cultural Health Index (CHI), or similar assessment, in partnership with iwi.
- Enumerate by culture *E. coli*, enterococci, and *Campylobacter* with non-enumerative isolation of *Salmonella* and STEC, and whole genome sequencing (WGS) on a selection of isolates.
- Enumeration of *Giardia* and *Cryptosporidium* by microscopy.
- Enumeration of indicator organisms, human, ruminant and wildfowl FST markers and pathogens by quantitative PCR (qPCR).
- Collect field data for temperature (air and water), pH, dissolved oxygen and turbidity, and collate flow and rainfall data to assess as explicatory variables.
- Data analysis for association(s) between faecal indicator bacteria (FIB) and pathogens, and faecal source as explanatory variables.
- QMRA populated by data from rivers to provide human health risk estimates from recreational activity exposures to pathogens.

2. 2021 SAMPLING PROGRAMME

Funds became available in late 2020 to commence sampling to build the database of pathogen and indicator data from January 2021. Councils contributed through their routine sampling programme, with a few councils able to undertake additional sampling due to the proximity of sites. Additional sites were sought by liaising with iwi to provide more culturally sensitive sites and develop a Cultural Health Index, or similar.

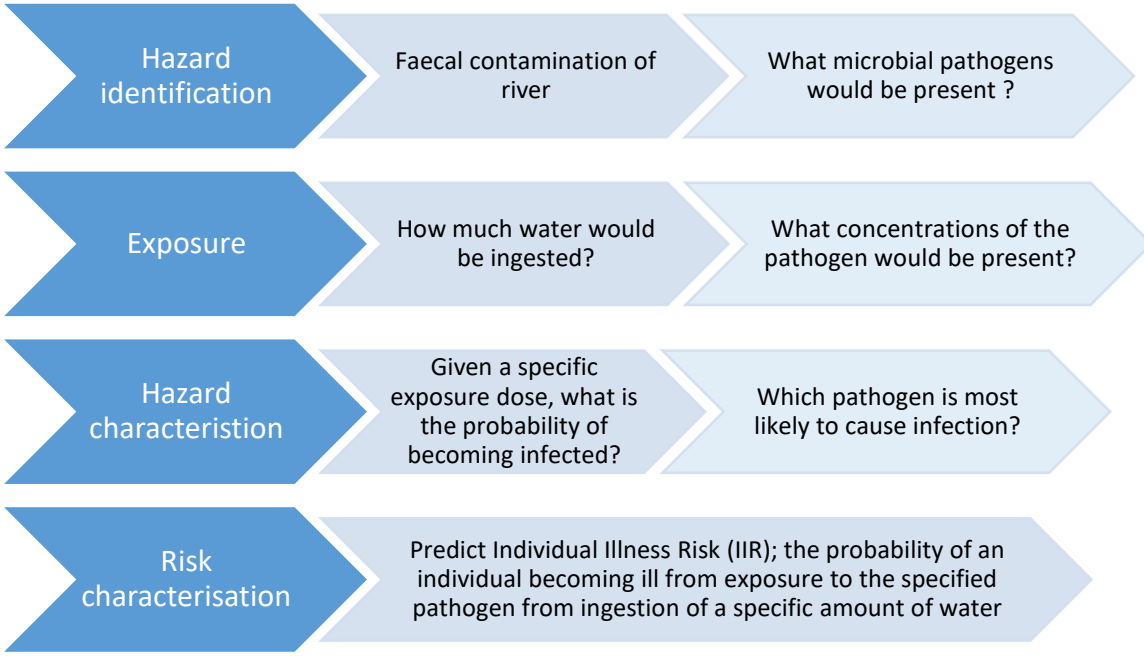

Figure 1: Key steps of a QMRA

Hazard identification
Faecal contamination of river
What microbial pathogens would be present ?
Exposure
How much water would be ingested?
What concentrations of the pathogen would be present?
Hazard characteristion
Given a specific exposure dose, what is the probability of becoming infected?
Which pathogen is most likely to cause infection?
Risk characterisation
Predict Individual Illness Risk (IIR); the probability of an individual becoming ill from exposure to the specified pathogen from ingestion of a specific amount of water

Figure 1 summarises the key steps in a QMRA. This project is building a database of pathogen and FIB concentrations from field data (Steps 1-3). This data will be used as an input to the model (Step 4).

The purpose of Phase 2.1 is to:

- confirm and extend the number of sites which are monitored to ensure there is good geographical representation, covering different land uses
- engage with iwi, to select culturally significant sites and to incorporate cultural health assessment (CHI or similar)
- complete project planning and commence sample collection and analysis.

2.1 SAMPLING

Intensive sampling was planned over the bathing season, with some sampling extending to late autumn. This extension captured some of the seasonal variation in the presence of pathogens in the environment, including their attenuation on land and transport to waterways under different rainfall patterns.

Due to the short lead-in period before sampling commenced, it was not possible to visit each council, so instead a short video was made and distributed to each council sampling team. Sampling staff and managers were invited to a Zoom meeting to discuss the sampling in more detail and respond to any questions. Site visits were also made to conduct sanitary inspections. Delays with courier delivery were mostly remedied by changing courier and by spreading out sample arrival times. However, some samples were still significantly delayed. Microbial culture analysis is time sensitive, therefore, any delivery delays were noted in the results. No microbial culture analysis was undertaken if there had been more than 48 hours delay in delivery.

The sampling period was over 16 weeks from 25 January-31 May 2021 and timed to fit with State of the Environment monitoring (sampled monthly), or the bathing season recreational water quality sampling programme (weekly or fortnightly sampling January-March/April depending on the local “bathing season”). Additional sampling occurred after that the end of the recreational water quality sampling programme depending on the resources available with the aim of collecting a minimum of 200 samples for analysis and to inform cultural health assessment.

A sampling application which had been developed to enable data to be collected in the field and loaded into the database was trialled in the field.

2.2 SITE SELECTION

In addition to the 16 sites monitored in 2020, another eight sites were selected from the initial site selection report (Milne et al 2018). Low impact sites were also identified by councils, and iwi nominated sites of cultural significance. Other sites had been identified based on elevated *E. coli* concentrations.

The dominant land uses were targeted with the following number of sites:

- Nine urban sites, two of which are likely to be influenced by wildfowl.
- Twelve sites where the major influence is likely to be sheep & beef farming.
- Thirteen sites where the major influence is likely to be dairy farming.
- Three sites which did not have intensive land use or development and not expected to have significant faecal sources.

The sampling sites were geographically distributed across New Zealand and are listed with their observed land use in Appendix B. Eight of the sites listed above were selected after consultation with iwi on cultural engagement.

2.3 INCORPORATION OF CULTURAL VALUES

2.3.1 Culturally significant sites

The cultural significance of sites is a commonly used indicator for kaupapa Māori monitoring and cultural health assessment. The concept of cultural landscape has been described by McGregor and Begley (2014) as landscape "characterised by its natural and physical aspects but also its sites, whakapapa, stories, mahinga kai, rock art and wāhi tapu". Feedback from earlier consultation with iwi indicated that sites of cultural significance had not been included in the selection of sites. Further engagement has been undertaken with iwi regarding culturally significant sites. Iwi management plans were also reviewed to identify sites of cultural significance in the selection of the original list of council nominated sites (Milne et al 2018).

Engagement with iwi on cultural health assessment of selected sites was commenced for four awa and/or their tributaries. Sites for these awa were also included in the sampling programme. They were not selected on the basis of elevated *E. coli* concentrations and sampling of the awa was not confined to one location. Our initial approach was to use these as case studies, leveraging relevant work being undertaken in the rohe. The awa and iwi hapū were:

- Wakapuaka - Ngāti Tama ki Te Waipounamu Trust.
- Motupipi - Manawhenua ki Mōhua representing Te Ātiawa, Ngāti Tama and Ngāti Rārua in Mōhua.
- Ruamahanga tributaries: Kopuaranga, Waipoua, Whangaehu - Ngāti Kahungunu ki Wairarapa.
- Waitara and tributary Manganui - Pukerangiora Hapū, Otaraua hapū and Te Kotahitanga o Te Atiawa, the iwi post-settlement governance entity.

2.3.2 Cultural health assessment

While western science provides a perspective on the suitability of a site for swimming, Te Mana o Te Wai changes the focus to the health of the wai as paramount and takes a more holistic view of the wider environment. A formalised approach to including mātauranga Māori in assessment of the health of an awa has been the use of the Cultural Health Index (CHI) tool which quantitatively and qualitatively measures indicators of relevance to local iwi and hapū in the rohe and derives scores. The safety of mahinga kai, measurement of the quality of water and traditional practices (such as swimming) are indicators frequently used in kaupapa Māori monitoring. *E. coli* is indicative of faecal contamination and consequently the risk to health from mahinga kai and contact activities such as swimming. While laboratory analysis is available for *E. coli,* it may not be practical as there are requirements around the sterility of containers, limited time between collection and delivery to the laboratory and a time lapse between a sample being taken and the results received. A compartment bag test (CBT) for *E. coli* is a test kit which can be used outside the laboratory. These were provided for iwi to use in conjunction with the kaupapa Māori monitoring for cultural health assessment. It provided more flexibility as they can be used at times convenient to iwi and hapū and results are more rapid.

2.4 ANALYSIS AND ENVIRONMENTAL DATA

2.4.1 Analytical methods

Each sample collection event involved the collection of 6 L and 10 L water samples. Figure 2 provides an overview of the analytical approach, with details of the methods in Appendix C.

For each sampling event, the following analytical tests were undertaken on the water samples:

- *E. coli* and enterococci enumerated by Colilert and Enterolert.
- *Campylobacter* enumerated by MPN using a 10-tube MPN, including a 1 L sample to compare prevalence with *Salmonella* and STEC
- *Salmonella* and STEC from 1 L samples enriched for these pathogens.
- *Cryptosporidium* and *Giardia* enumerated by microscopy from a 10 L sample.
- Pathogens and indicators including human, ruminant, and wildfowl FST markers, enumerated by qPCR by filtering up to 2 L of water and DNA extraction from the filtered water sample.
- WGS on selected isolates of *Campylobacter*, *Salmonella* and STEC to confirm if the species are pathogenic to humans.

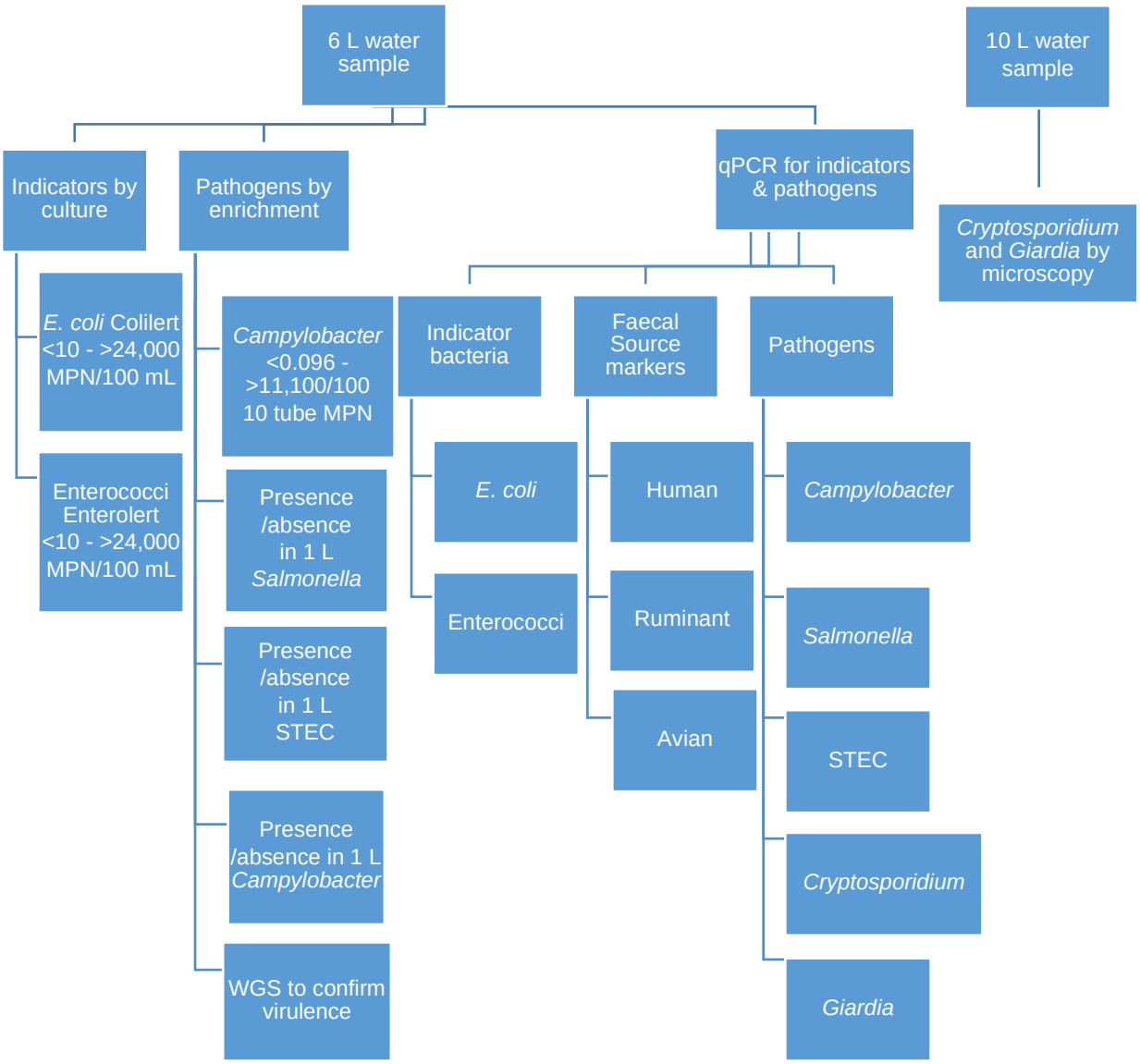

Figure 2: Summary of analyses

2.4.2 Environmental data

Environmental data has been recorded from physio-chemical measurements taken in the field/laboratory or from records and is summarised in Table 1.

Table 1: Environmental data

Parameter	Source
Photo of river at sampling time	Cell phone
Flow	Gauged site (or taken from flow recorded, or modelled)
Turbidity	Council equipment if equipment available
Conductivity	Council meter if equipment available
pH	Council meter if equipment available
Dissolved Oxygen	Council meter if equipment available
Water temperature	Council meter if equipment available
Air temperature	Phone or meter or thermometer
Time sampling started and finished	Watch/cell phone
Presence of animals	Observation
Rainfall at time of sampling	Observation
Wind direction and strength	Observation
Sunlight	Observation
Rainfall previous 24, 48 and 72 hours	Records from nearest weather station

3. RESULTS

3.1 SAMPLING SITES

The sites selected provided good coverage across New Zealand. The location of the 30 rivers and one lake which were sampled are shown in Figure 3 and the 37 sites are listed in Appendix B.

Figure 3: Location of sampling sites

Except for the 10 sites selected by iwi/hapū for cultural significance and three sites with only bush or wildfowl impacts, sites were selected based on elevated concentrations of *E. coli*, observed land use and/or previous FST data. This was done to include freshwater bodies likely to be contaminated by pathogens from either urban, or dairy and beef & sheep farming sources. The results are therefore not reflective of water quality in general across New Zealand.

Table 2 presents the number of samples in the different land use categories for each week of sampling. In total, 256 samples were collected between 25 January and 31 May 2021. To provide anonymity, the samples are grouped by observed land use. There are not equal numbers of samples for the different categories as the sampling plan was based on current sampling programmes. Due to the short timeframe available to commence sampling, there was insufficient time for councils to resource non-routine sampling.

Table 2: Summary of sampling

	Urban	Sheep & Beef	Dairy	Low Impact	Total
Week commencing	Number of samples	Number of samples	Number of samples	Number of samples	Number of samples
25-Jan	4	6	3	2	15
1-Feb	6	3	2	1	12
15-Feb	4	6	4	3	17
22-Feb	6	6	3	2	17
1-Mar	5	8	3	3	19
8-Mar	7	3	5	2	17
15-Mar	7	7	2	3	19
22-Mar	7	3	4	2	16
29-Mar	8	5	3	3	19
12-Apr	4	5	6	1	16
19-Apr	3	2	7	3	15
3-May	4	5	7	0	16
10-May	3	1	6	0	10
17-May	6	3	5	2	16
24-May	4	5	6	1	16
31-May	2	8	6	0	16
Total	80	76	72	28	256
Number of sites	9	12	13	3	37

3.2 FAECAL SOURCE TRACKING (FST) MARKERS

The general faecal source marker GenBac was analysed in 256 samples and results were available for 250 (DNA was degraded by an unknown contaminant in five samples and there was an extraction problem with one sample). This non-specific marker indicates likely faecal contamination by a range of animals including human, cow, sheep, deer, goat, pig, possum, rabbit, cat, dog, horse, and birds. Faecal source identification was undertaken for each sample using FST markers specific for human, ruminant, and avian faecal contamination. FST data is not available for six samples due to extraction and analytical issues. Five of these samples were from one river where the internal control showed that DNA had deteriorated during processing and appears to be related to some contaminant in the river which interfered with the DNA extraction.

3.2.1 Faecal Source Tracking marker prevalence

Three human FST markers were used: HF183, crAssphage and BiADO. Low concentrations of one or more of the human indicative markers in samples may be the result of low level human faecal sources or may the consequence of cross reaction with non-human sources. Since these possibilities cannot be distinguished, in this study, samples with at least two markers and a combined total of 1,000 copies/100 mL or more were characterised as containing a significant human source of faecal pollution. There were 72 samples (29%) which met these criteria.

The BacR marker indicates the presence of ruminant sources of faecal pollution and samples with >1,000 copies/100 mL were considered significant. In addition, samples with >100 copies/100 mL BacR and where the ratio of BacR copies to the total GenBac copies was >0.3% were reported. The BacR marker was detected in 171 samples (67%) of which 107 samples had >1,000 copies/100 mL and 45 had >10,000 copies/100 mL. A maximum of 7,100,000 copies/100 mL was observed.

For avian sources, the GFD marker was used. Avian FST marker concentrations greater than 100 copies/100 mL were reported in 206 samples. The GFD marker was the most frequently detected FST marker, and was detected in 226/250 samples (90%), 102 samples had greater than 1,000 copies/100 mL and 24 samples (10%) were below the detection limit. Therefore, this marker was not as ubiquitous as in the 2020 pilot study (Leonard et al 2020).

3.2.2 Alignment of FST marker prevalence and land use

The raw data for each FST marker in each sample, $\geq$100 copies/100 mL, is presented in Figures 4-7 to illustrate the temporal variation in the types of FST markers and their concentrations within a site. The sites are grouped according to the observed land use. The distribution of the FST markers within and across the sites is discussed below. The dominant FST was calculated for each sample and is presented in Figure 8[2].

[2] Note that there needs to be two types of Human FST markers to confirm a human faecal source

Urban land use

Urban land use had all FST markers present (Figure 4), with the avian marker present in at least one sample from each site at significant concentrations (>1,000 copies/100 mL). All sites, except site 9, had human FST markers. Ruminant FST markers were also detected in samples from Urban sites, with significant concentrations in 15 samples (from sites 1, 2, 4, 6 and 8), implying a likely rural influence in the upper catchment of these Urban sites.

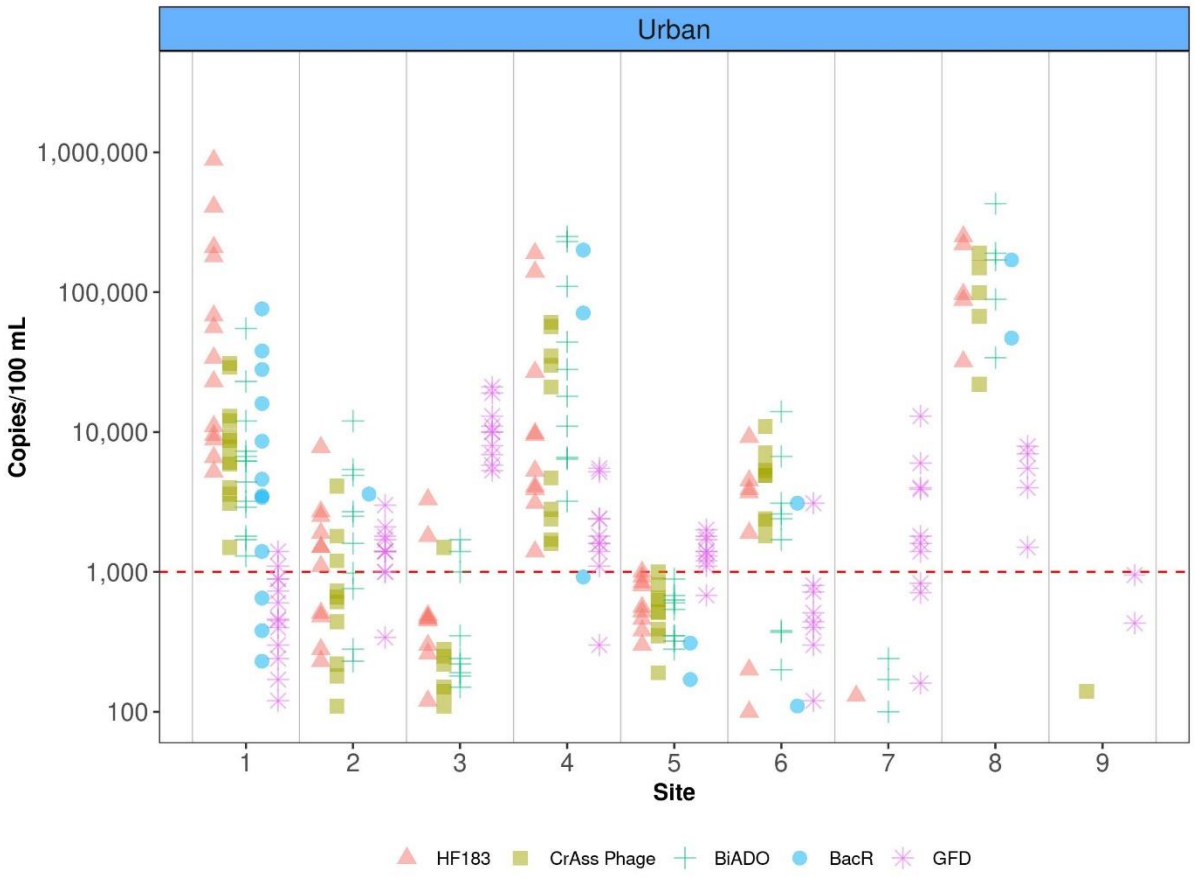

Figure 4: Concentration of FST markers at Urban land use sites.

Dotted line represents the concentration where detection of a FST marker was indicative of a significant faecal source.

Sheep & Beef and Dairy land uses

Ruminant only FST markers were dominant in 88 samples, 33 Dairy and 44 Sheep & Beef. Only one Sheep & Beef site (site 15) had no dominant ruminant FST markers (Figure 5) and in 19 samples the concentrations were below the threshold for a dominant ruminant FST marker. All sites in Dairy land use had at least some samples with ruminant FST markers above the threshold, except for site 33 (Figure 6).

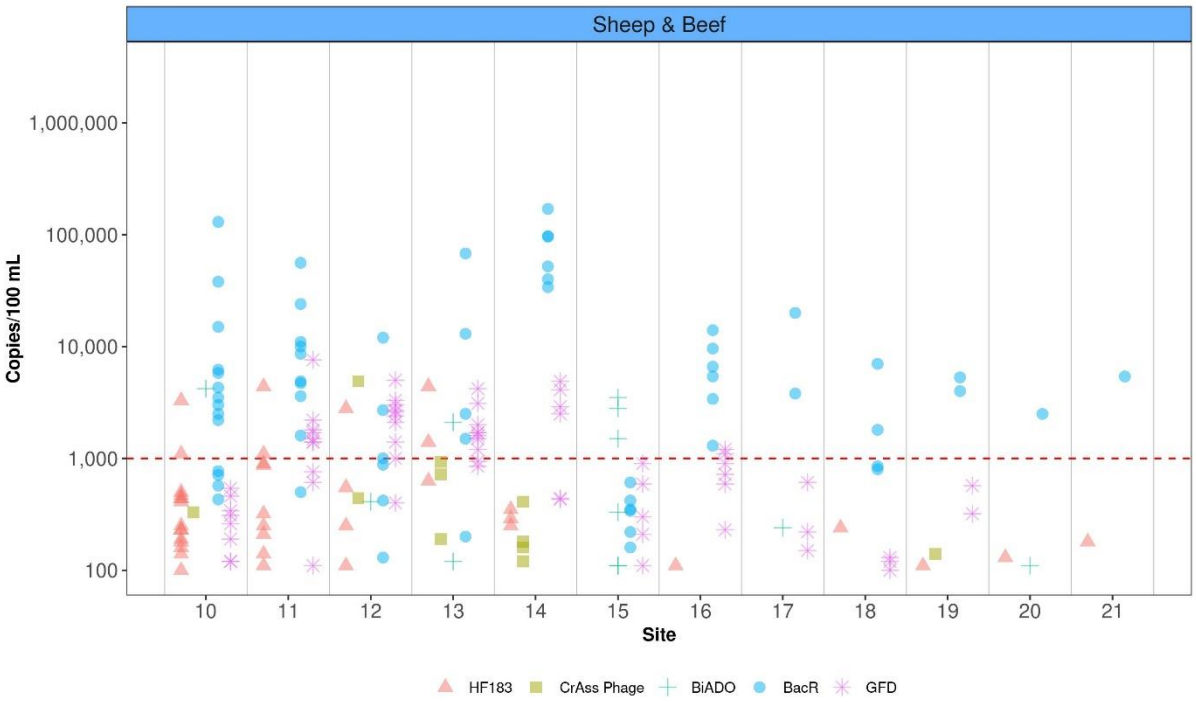

Figure 5: Concentrations of FST markers at Sheep & Beef land use sites

Dotted line represents the concentration where detection of a FST marker was indicative of a significant faecal source.

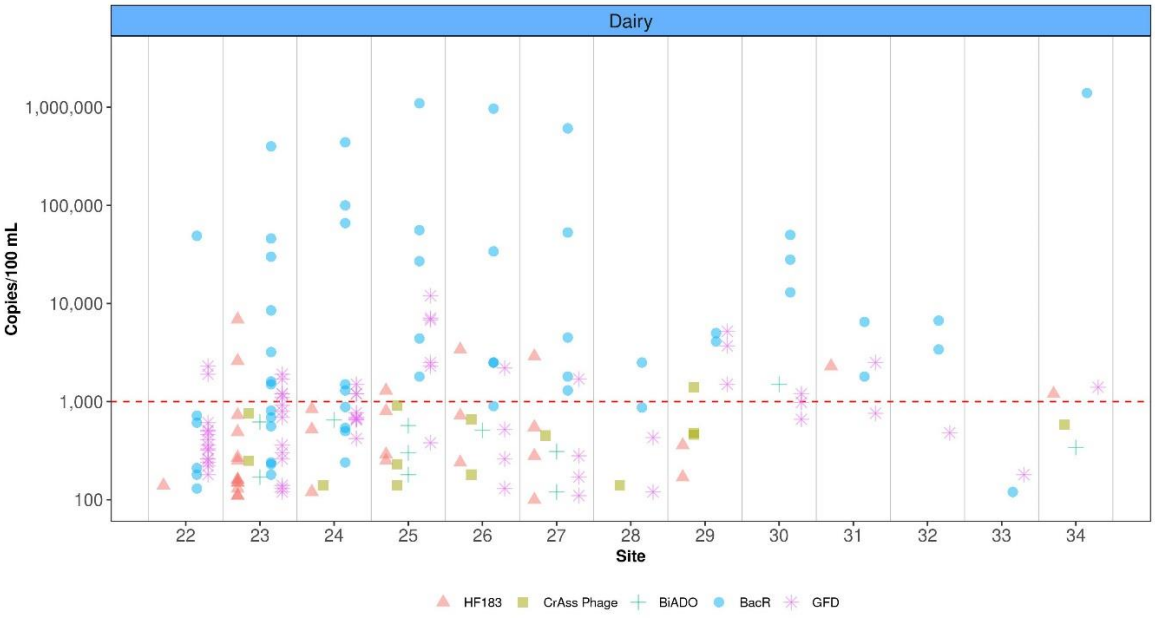

Figure 6: Concentrations of FST markers at Dairy land use sites

Dotted line represents the concentration where detection of a FST marker was indicative of a significant faecal source.

Avian FST markers (>100 copies/100 mL) were present at all Dairy sites in at least some samples and at all but two Sheep & Beef sites. High concentrations of the human FST marker HF183 were present at four Sheep & Beef sites (seven samples) and six Dairy sites (seven samples). There were seven sites (eight samples) where it was supported by the

presence of at least one of the other human FST markers. Two of these sites were in the Sheep & Beef land use grouping and five sites in Dairy.

Low Impact land use

These Low Impact sites were selected with the expectation that the only FST markers present would be avian as there was no intensive farming or settlement. However, only one site had one sample with avian FST markers greater than 100 copies/100 mL. Instead, ruminant FST were present in at least one sample for all Low Impact sites, indicating there could be (feral) deer or goats in the catchment. Site 35 had high concentrations of the human FST marker HF183. However, it was not supported by the presence of the other human FST markers, suggesting that the source of the HF183 in this river is a non-human faecal source. The HF183 FST marker can also be indicative of possums or goats.

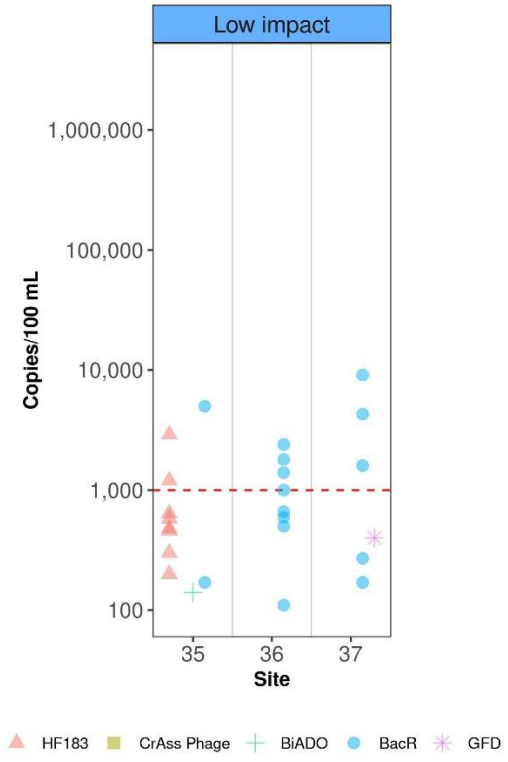

Figure 7: Concentrations of FST markers at Low Impact sites

Dotted line represents the concentration where detection of a FST marker was indicative of a significant faecal source.

3.2.3 Classification of sites

A summary of the dominant FST marker for each sample is shown in Figure 8 with the sites grouped by observed land use. The summary FST data confirms that there is a mix of dominant faecal sources across all observed land uses over time. At six sites, the dominant FST markers from all samples were inconsistent with the observed land use. One site had been selected for its cultural significance rather than high *E. coli* from land use impact. Two Urban sites (7 and 9) and one Dairy site (33) only had avian FST. The three Low Impact sites (35, 36, 37) all had ruminant FST markers but no dominant avian FST markers, which had been expected to be the dominant faecal source. All but 4/37 sites had some samples with avian FST markers $\geq$ 100 copies /100 mL over the sampling period. In most samples avian faecal material is likely to be a contributor to faecal contamination.

Twenty-two samples did not meet any criteria for a dominant faecal source and are grouped as Unidentified. In 18 of these samples, *E. coli* concentrations were below 550 MPN/100 mL and would be suitable for contact recreation[3]. However, four samples from Low Impact site 35, had high concentrations of *E. coli* (>900 MPN/10-0 mL), with concentrations of GenBac ranging from 65,000 to 310,000 copies/100 mL, and one with a high concentration of one of the three human FST markers (Figure 7) indicating another faecal source was present (i.e. not detectable with the FST markers used in this study). As shown in Figure 8, 12/22 samples (50%) which were grouped as Unidentified were from two rivers (sites 35 and 36) that were expected to be Low Impact, six samples were from Sheep & Beef and four were from sites grouped as Dairy land use.

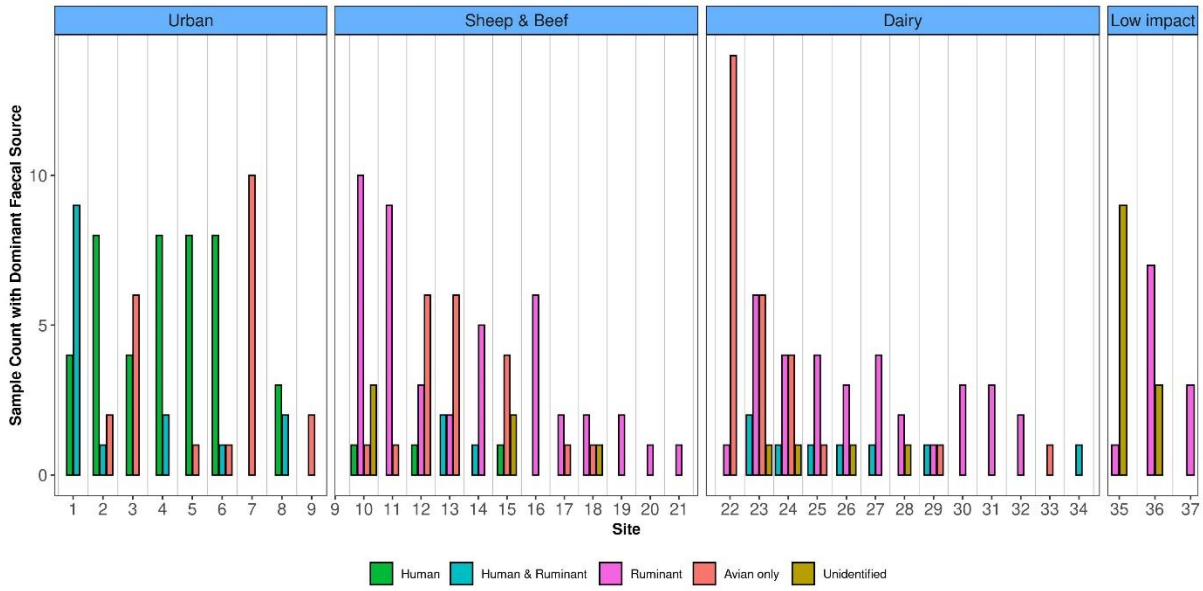

Figure 8: Summary of the dominant FST for each sample by observed land use

Human only FST markers were dominant in 46 samples mostly at six Urban sites, except for one sample each from sites 10, 12 and 15 which were Sheep & Beef sites. Sites expected to have dominant human FST markers also had ruminant FST markers. A mixture of human and ruminant FST markers (26 samples) was found at 13 sites: five Urban sites (15 samples), one Sheep & Beef site (2 samples) and seven Dairy sites (8 samples).

Due to this significant mismatch of observed land use and the dominant faecal source, the rest of the report classifies samples based on dominant faecal source rather than observed land use.

The FST data highlights that observed land use is not always a useful indicator of the source of faecal contamination. As the purpose of land use observation and testing for FST markers is to assist in determining mitigation steps, FST will provide more accurate information. Therefore, FST markers are favoured for reporting on faecal sources. Where both human and ruminant FST markers are dominant they are assigned to a human & ruminant faecal source grouping. The new grouping, Unidentified, is used where avian, human, and ruminant FST markers are present in either low concentrations or not detected.

[3] At 260 MN/100 mL extra sampling and investigation is required by the Guidelines to investigate the source of faecal pollution (MfE 2003).

3.3 FAECAL INDICATOR BACTERIA (FIB)

3.3.1 *E. coli*

E. coli were detected in 252 (98%) of samples using Colilert assays. There was a range in *E. coli* concentrations with two samples above the limit of detection (24,200 MPN/100 mL), four samples below the detection limit (<10 MPN/100 mL), 10 samples at the detection limit and 43 samples (17%) ≥1,000 MPN/100 mL. *E. coli* concentrations are presented for the different dominant faecal sources in Figure 9, with the criterion for recreational water quality marked at 550 MPN/100 mL. There are 75 samples >550 MPN/100 mL and 130 above 260 MPN/100 mL, which is the trigger for daily sampling in the Guidelines (MfE 2003).

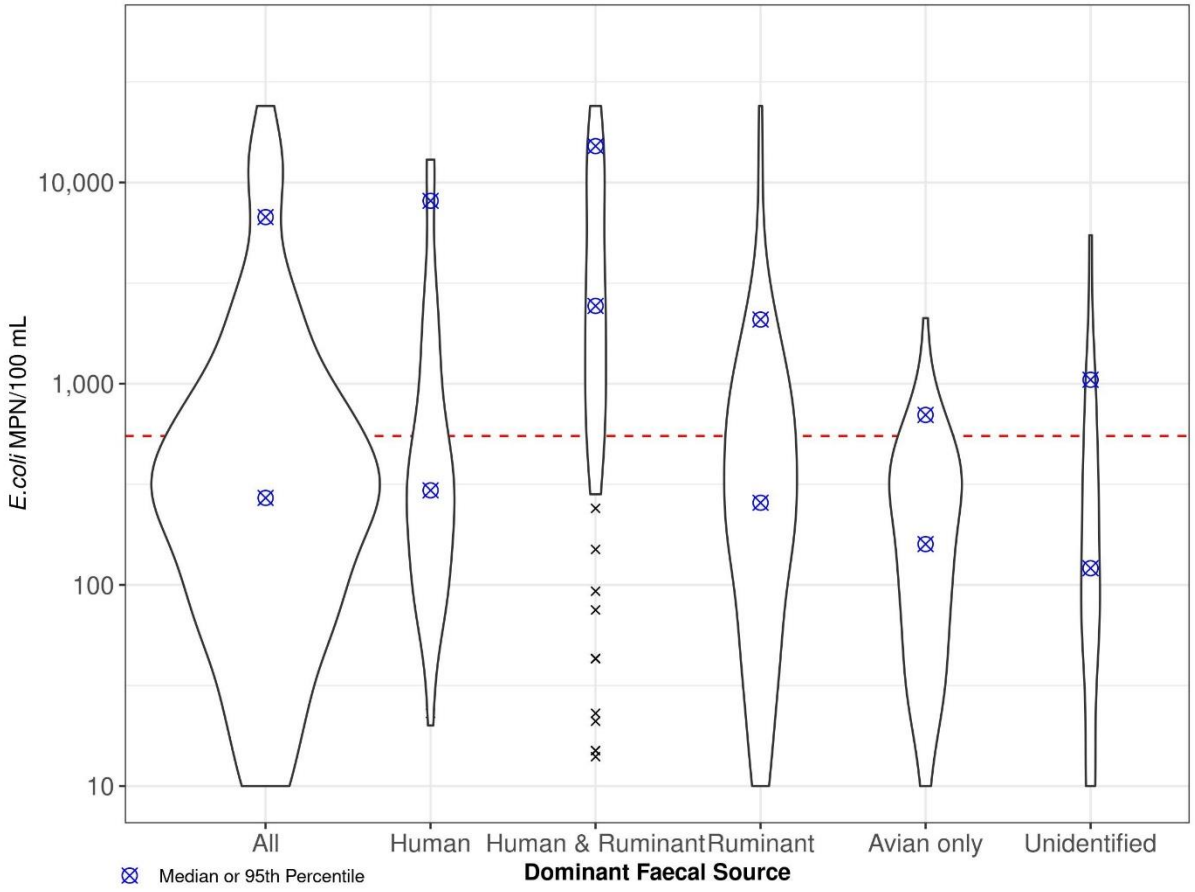

Figure 9: Concentrations of *E. coli* against dominant faecal source

The dotted line represents 550 *E. coli*/100 mL, which is criterion at which the public should be notified of a health risk (MfE, 2003). Not shown on the graph are the two samples above the limit of detection (24,200 MPN/100 mL), one in ruminant and one in human.

While the median *E. coli* concentration for all the samples is 267 MPN/100 mL, when sorted by faecal source the human & ruminant group has the highest median of 2,489 MPN/100 mL and 95th percentile of 15,525 MN/100 mL. The median *E. coli* concentrations with dominant human FST markers and dominant ruminant FST markers are both similar to the overall median at 295 MPN/100 mL and 256 MPN/100 mL, respectively. However, there were more samples at high concentrations in the human group as the 95th percentile is 9,168 compared to 2,073 MPN/100 mL in the ruminant group. Concentrations of *E. coli* in the avian group had the least spread with a median of 158 MPN/100 mL and 95th percentile of 692 MPN/100 mL. The lowest median was the Unidentified group at 81 MPN/100 mL with a 95th percentile of 1,041 MPN/100 mL.

Quantitation of *E. coli* was undertaken by the two methods: culture using the traditional Colilert method and by qPCR. There is a moderately strong correlation between the two methods with an R^2 value of 0.641 (Figure 10) and is highly significant with $p < 0.001$.

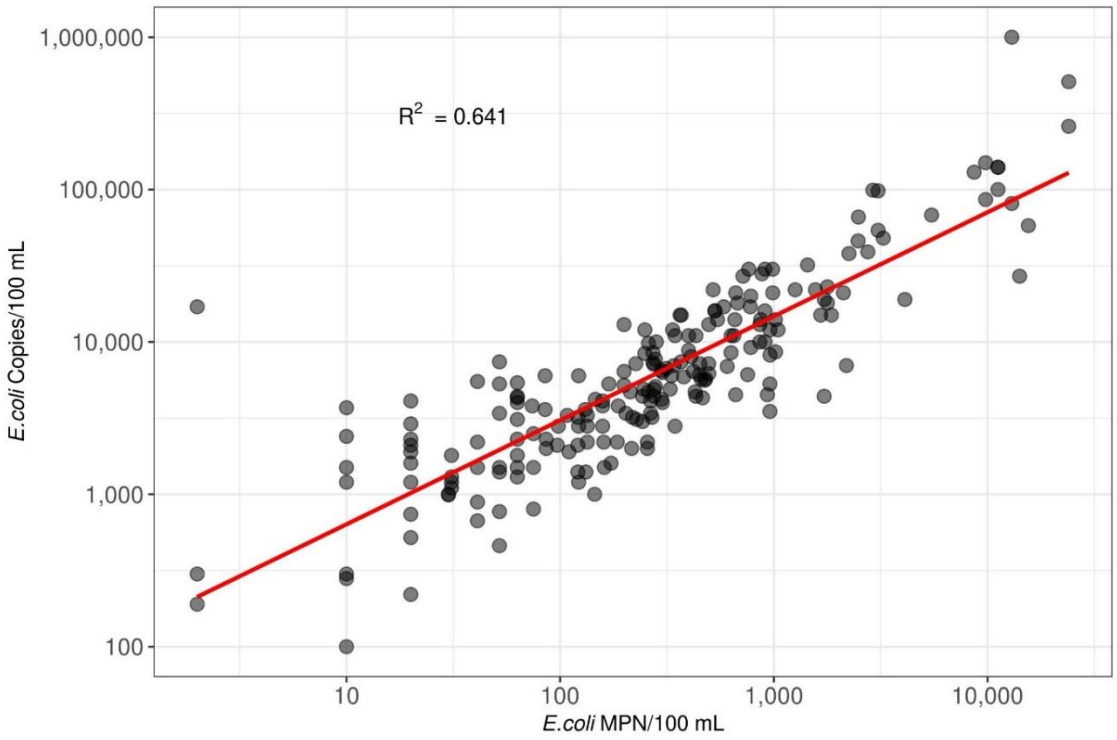

Figure 10: Correlation between culture and qPCR for *E. coli*

3.3.2 Enterococci

Enterococci was analysed in 252 samples and detected in 233 samples (92%), with 19 samples below the limit of detection (10 MPN/100 mL), 22 samples at the level of detection and none above the upper limit of detection of 24,200 MPN/100 mL. Concentrations of enterococci ranged from <10-24,196 MPN/100 mL, with a median of 97 MPN/100 mL and 95th percentile of 2,023 MPN/100 mL (Figure 11).

Again, the human & ruminant group had the highest median concentration of 780 MPN/100 mL (Figure 11) and highest 95th percentile of 23,329 MPN/100 mL. Figure 11 highlights the wide spread of data in this grouping. Although the medians are similar for the human (110 MPN/100 mL), ruminant (73 MPN/100 mL) and avian (86 MPN/100 mL) groups, there is a wider range of concentrations in the ruminant group as seen by the 95th percentile value of 2,076 MPN/100 mL compared with 1,630 MPN/100 mL and 480 MPN/100 mL for human and avian groups, respectively. As with *E. coli,* the median concentration for Unidentified was lowest at 41 MPN/100 mL.

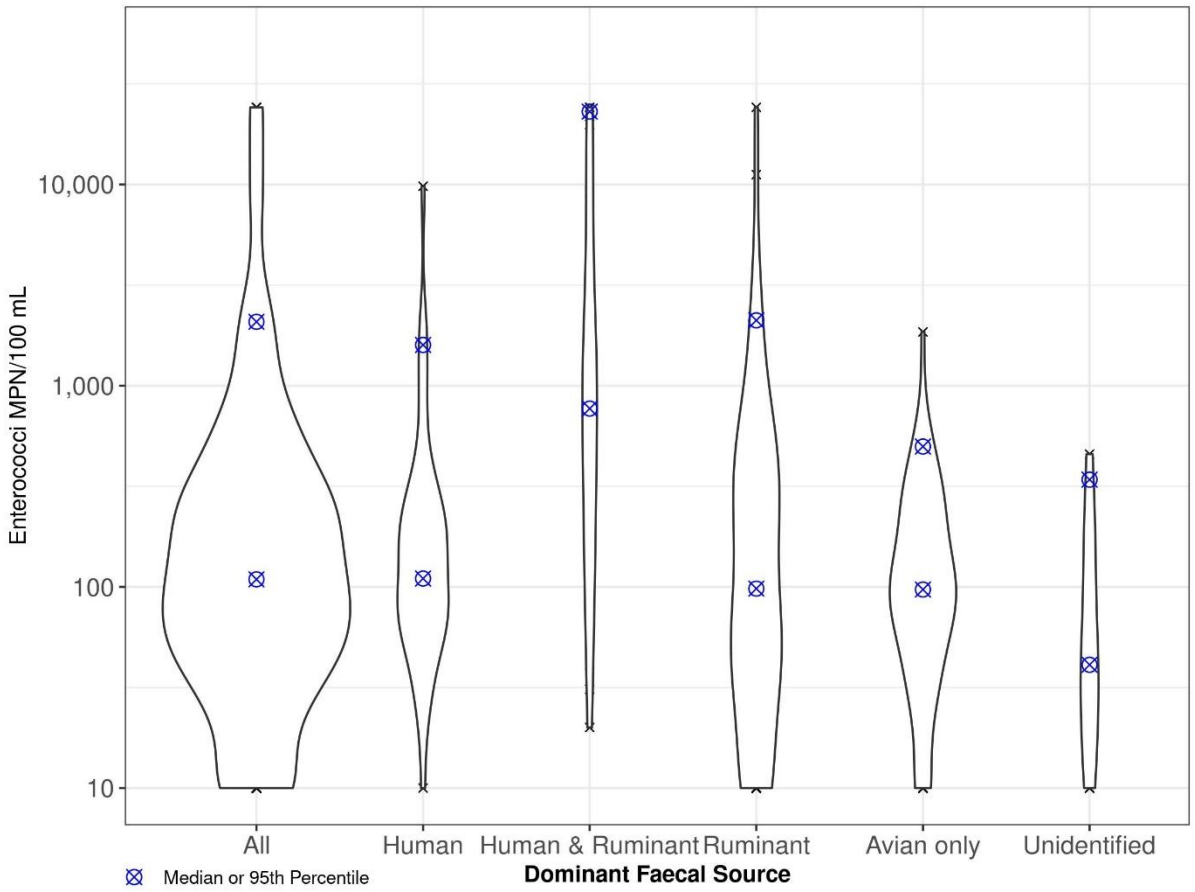

Figure 11: Concentrations of enterococci by dominant faecal source

Quantitation of enterococci was undertaken by two methods for samples with dominant human FST markers. The two methods were culture using the traditional Enterolert method (MPN/100 ML) and by qPCR (copies/100 mL). The Pilot Study 2020 indicated poorer correlation with other faecal sources. The correlation observed between culture and qPCR was moderate R^2 =0.458 (Figure 12) and is highly significant with $p < 0.001$.

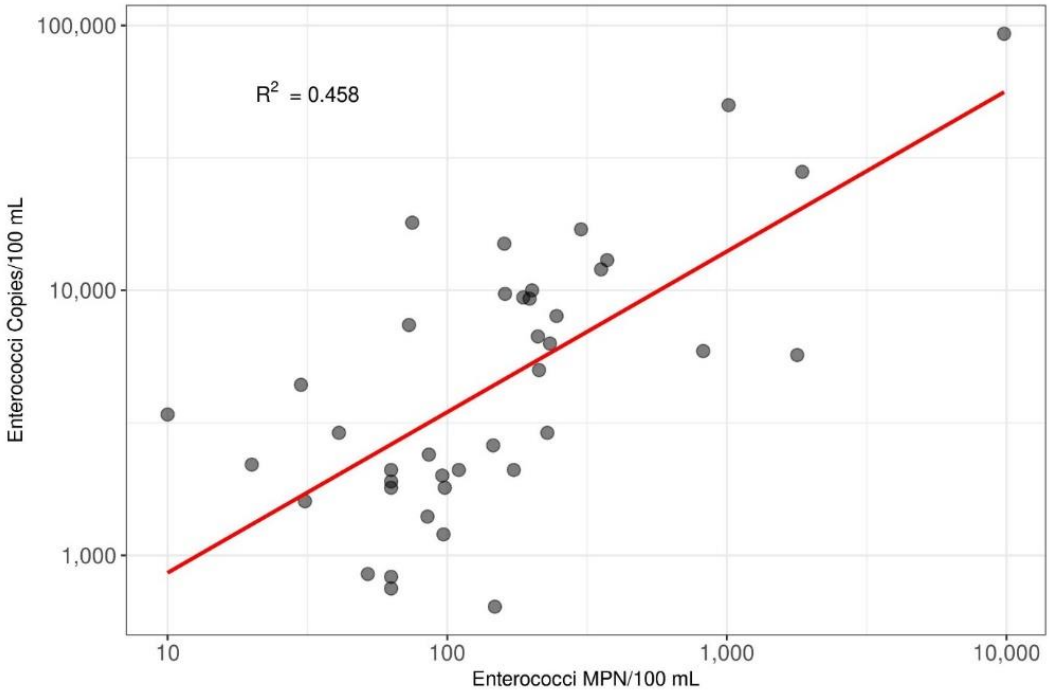

Figure 12: Correlation of enterococci by Enterolert and qPCR for human faecal source dominant samples

3.4 BACTERIAL PATHOGENS

3.4.1 Presence/absence of bacterial pathogens

The presence of pathogenic bacteria was determined by enrichment in selective broths followed by either biochemical confirmation or PCR detection on bacterial colonies. The frequency of detection across all samples is shown in Table 3 for 252 samples (four arrived too late for analysis by culture) and using faecal source category for 250 samples (there were technical problems with the FST marker qPCR analysis for six samples). *Campylobacter* spp. were most frequently detected (65% of samples), followed by *Salmonella* spp. (16%) then STEC (2.8%). In 10 samples *Salmonella* spp. were detected in the absence of *Campylobacter* but only once was STEC detected in the absence of *Campylobacter*. Both *Salmonella* and *Campylobacter* were present in 30 samples. In four samples, all three pathogens were present with three of these samples having human & ruminant FST as the dominant markers, the other had dominant ruminant FST markers.

Table 3: Frequency of detection of culturable pathogenic bacteria by faecal source group

Total and faecal source category	Number of samples	*Campylobacter*		*Salmonella*		*STEC*	
		Number positive	%	Number positive	%	Number positive	%
Total	252	161	64	40	16	7	2.8
Human & Ruminant*	25	24	96	12	46	4	15
Avian*	69	45	65	4	6	0	0
Ruminant*	85	58	68	10	12	3	3
Human*	45	26	58	14	31	0	0
Unidentified*	22	4	18	0	0	0	0

* Excludes six samples positive for *Campylobacter* which had technical problems during FST determination, and therefore, were unable to be assigned to a faecal source group.

Table 3 also presents the frequency of detection of each pathogen for each faecal source group. All three pathogens are most frequently detected in the human & ruminant group. *Campylobacter* is next most frequently detected, at similar rates, in the avian, human and ruminant groups, while *Salmonella* is the next most frequently detected in the human group.

Campylobacter spp. are the most widely distributed pathogens, detected at all but four sites, although two of these sites had only two samples taken. *Salmonella* is less prevalent, at 18 sites (17 freshwater bodies) and STEC is detected in only seven samples, at six sites (five freshwater bodies). *Campylobacter* is detected in four samples in the Unidentified Source group, with 3/4 at low concentrations (<10 MPN/100 mL), but one sample had 250 MPN/100 mL, although the *E. coli* concentrations in this sample was relatively low (75MPN/100 mL).

3.4.2 Quantification o*f Campylobacter*

Quantitation of Campylobacter *by culture and MPN*

As it was the most important pathogen identified in the pilot study, *Campylobacter* spp. were quantified using culture enrichment and MPN on 252 samples[4]. *Campylobacter* spp. were enumerated by culture and MPN in 165 samples (66%), which includes one sample greater than the limit of detection (11,000 MPN/100 mL). There were 54 samples $\geq$10 MPN/100 mL and 66 samples <1 MPN/100 mL (the detection limit was 0.096 MPN/100 mL). The frequency of *Campylobacter* detection by culture was similar to that determined in the absence/presence test in Table 3, except for four samples which were enumerated by culture and MPN (0.096-0.19 MPN/100 mL) but not by presence/absence test.

Figure 13 shows that *Campylobacter* was mostly detected at low concentrations with a median of 0.34 MPN/100 mL and a 95th percentile of 134 MPN/100 mL. The highest median was 9.2 MPN/100 mL for human & ruminant, with a 95th percentile of 89 MPN/100 mL. While data in the avian dominant faecal source had a lower median of 0.34 MPN/100 mL, the spread of data was the greatest of all sources with a 95th percentile of 240 MPN/100 mL. High concentrations of *Campylobacter* were also observed in ruminant faecal source grouping, with a median of 0.34 MPN/100 mL and 95th percentile of 222 MPN/1000 mL.

Figure 13 highlights that the low concentrations measured were associated with human faecal sources with a median of 0.34 MPN/100 mL and 95th percentile of 9.2 MPN/100 mL. The few data points for Unidentified are spread widely with the median being less than the detection limit, 0.096 MPN/100 mL and 95th percentile of 3.6 MPN/100 mL.

[4] 252 samples were analysed as four samples arrived outside the critical period for analysis by culture

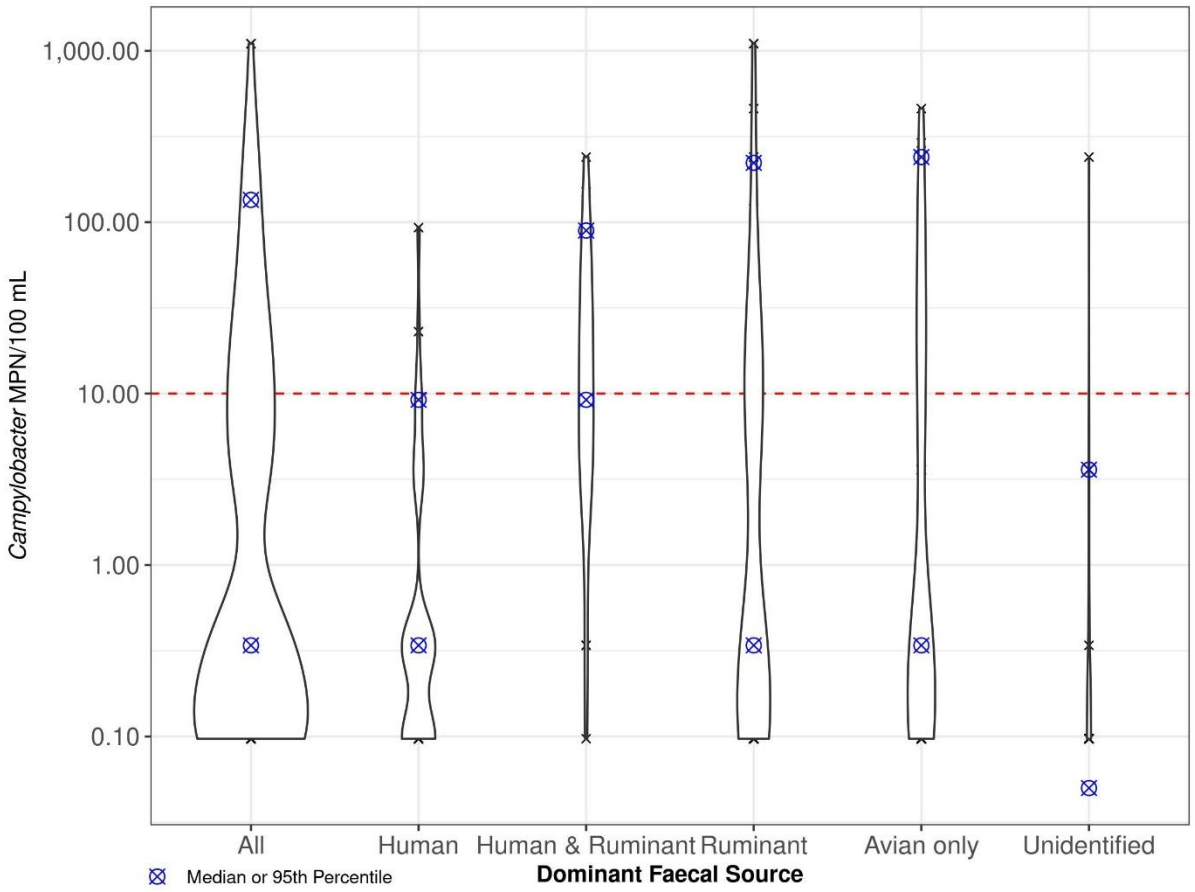

Figure 13: Concentrations of *Campylobacter* by dominant faecal source

Quantitation of Campylobacter *by qPCR*

Direct qPCR analysis of *C. jejuni* and *C. coli* was performed on DNA extracts from water samples. Quantitative PCR results for *Campylobacter* are available for 214 samples[5]. Figure 14 shows that *Campylobacter* spp. were detected at least once at all sites by qPCR, except for sites 32, 33 and 35.

[5] Data are not available for all samples due to machine error.

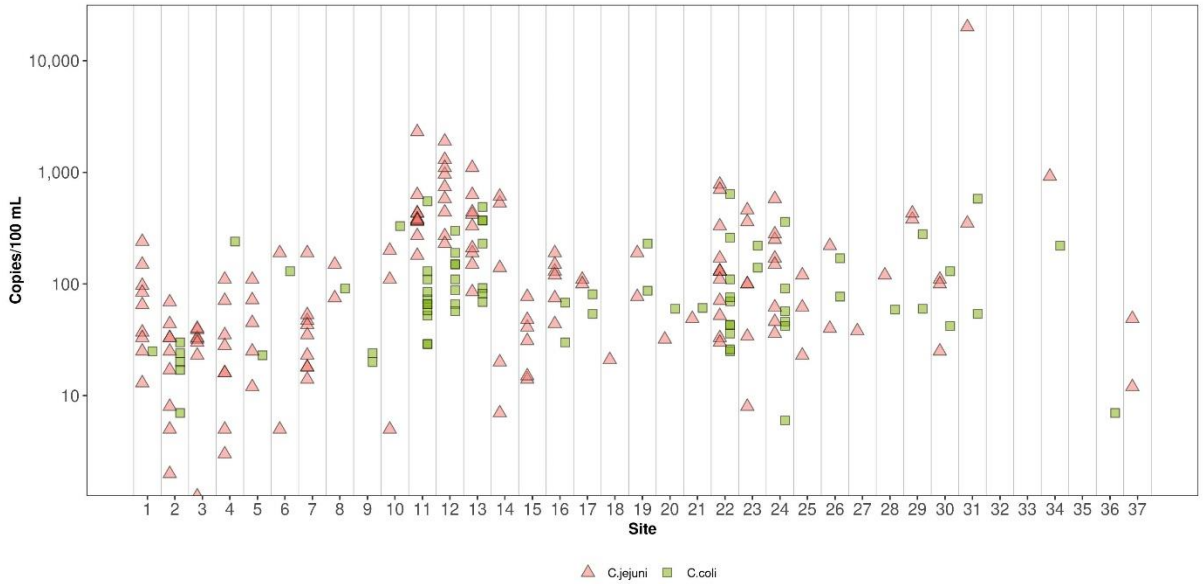

Figure 14: Quantitation of *Campylobacter* species by qPCR at all sites

C. jejuni was detected in 147/214 samples (2-20,000 copies /100 mL) and *C. coli* was detected in 74 samples (6-640 copies/100 mL) with *C. jejuni* and *C. coli* concurrently detected in 68/214 samples.

There were 18 samples where *Campylobacter* was detected by MPN culture (0.34-23 MPN/100 mL), but not by qPCR. In 29 samples, *C. jejuni* and/or *C. coli* were detected by qPCR but not by MPN culture. Three of these samples had *C. coli* only (7-24 copies/100 mL) and 20 had *C. jejuni* only (8-77 copies/100 mL). Where there was comparative data, all but one of the 42 samples which had more than 10 MPN/100 mL were detected by qPCR.

Analysis of *Campylobacter* by WGS (Table 9, Appendix D) confirmed 108/131 *Campylobacter* isolates as *C. jejuni,* three were possible *C. jejuni* and seven were possible *Campylobacter species* and 10 as *C. coli.* Serotyping analysis showed that there were 30 isolates were unable to be serotyped, but were assessed as likely to be *Campylobacter*, and one isolate identified as not *Campylobacter*.

A comparison of the traditional culture MPN method and qPCR for quantitation of *Campylobacter* showed that the linear correlation between qPCR and the traditional method for *Campylobacter* was weak ($R^2 = 0.2664$).

3.4.3 *Salmonella*

Direct qPCR analysis on DNA extracts from water samples detected the *Salmonella* virulence gene *inv*A in 23/214 samples from 14 sites at concentrations ranging from 22-380 copies/100 mL (Figure 15). There were 29 samples which were positive by the presence/absence test but below the detection limit by qPCR. Twenty samples were negative by the presence/absence test but detected by qPCR at concentrations ranging from 22 to 220 copies/100 mL. Only three samples were detected by both methods and the concentrations by qPCR were 33-380 copies/100 mL.

WGS data is given in Appendix E, Table 10 and shows that the isolates were identified as *Salmonella enterica.* The most common serotype was Typhimurium (22 isolates), which is commonly found in human clinical samples and in cattle. The serotype Bovismorbificans (2 isolates) is also found in human clinical samples and in cattle but is less common. The serotype Enteritidis was identified in five isolates and is commonly associated with human clinical samples and poultry. Two Stanley serotypes were identified. This is a common human *Salmonella* serotype normally associated with travel to Asia. One isolate was a Schwarzengrund serotype, which is more common in pigs and poultry.

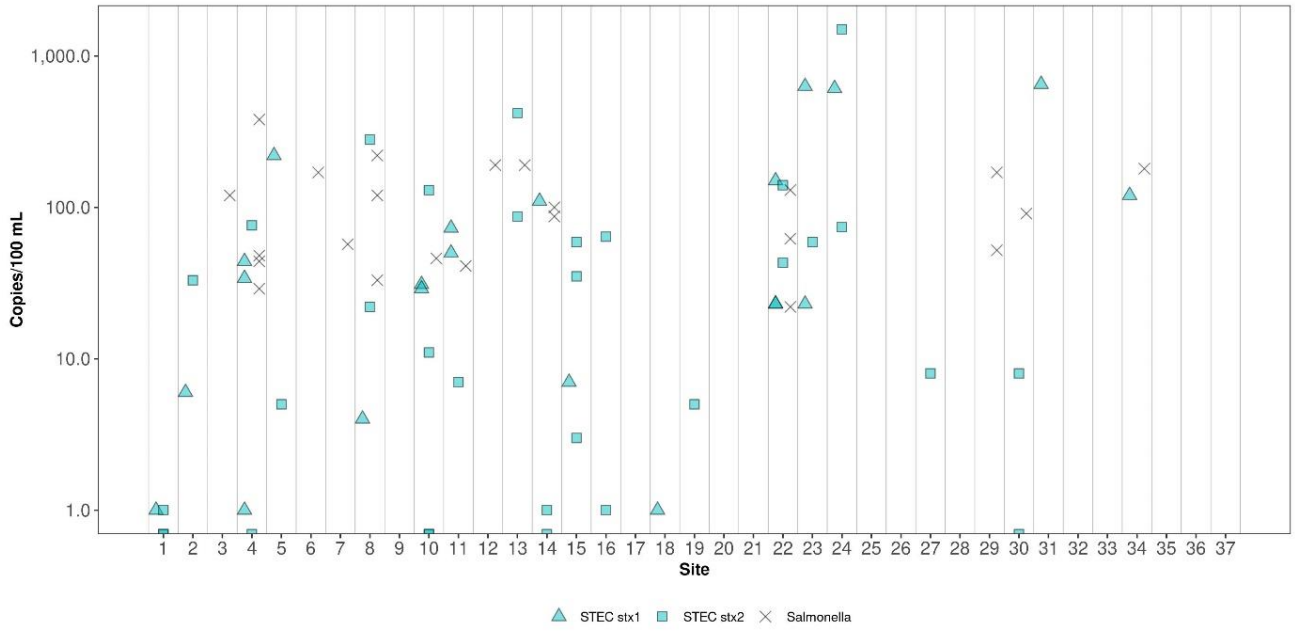

Figure 15: Virulence genes for STEC and *Salmonella* at all sites

3.4.4 STEC

Direct qPCR analysis on DNA extracts from water samples was used to detect the virulence genes for shiga toxin, *stx*1 and *stx*2. Identification of either of these two genes in *E. coli* isolates leads to the nomenclature shiga toxin-producing *E. coli* or STEC. The raw data is presented in Appendix E, Table 11. The *stx*1 gene was detected in 22 samples, ranging in concentration from 1 to 650 copies/100 mL, with six samples having concentrations of less than 10 copies/100 mL. The qPCR showed that six samples also contained the *stx*2 gene at concentrations between 3 to 1,500 copies/100 mL.

The *stx2* gene was detected in 25 samples in total, ranging in concentration from 1 to 1500 copies/100 mL, with nine samples at concentrations of less than 10 copies/100 mL.

The *stx*1 gene was detected at 15 sites and *stx*2 gene at 17 sites. Isolates from the six samples which contained both *stx*1 and *stx*2 were identified as *E. coli* and one is a mixture.

Three *E. coli* serotypes were O84:H2, one was O9:H30, two were ONT:H21 and one was ONT:H7. The O84:H2 serotype is commonly found in human clinical samples and was the 13th most common serotype in human clinical cases in 2020 in New Zealand. No serotypes belonging to O9:H30 have been confirmed in clinical cases in last five years (Wright J, 2021 pers. com email 24 June).

3.5 PROTOZOA

Protozoa were detected by filtering 10 L of water in the laboratory. Protozoa were eluted from the filter and concentrated before a subsample was placed on to a slide for visual detection. There were 253 samples tested for protozoa. The recovery rate through filtration, elution from the filter and microscopy is estimated to be between 15-55%. Results presented here are not adjusted by recovery rates. *Cryptosporidium* and *Giardia* analysis by traditional microscopy was undertaken on 253 samples and by qPCR on 214 samples.

3.5.1 *Cryptosporidium*

Generic *Cryptosporidium* was detected by microscopy in 104 (41%) of samples at low concentrations, ranging from 1 to 14 oocysts/10 L, with a median of below the level of detection (1 oocyst/10 L). As seen in Figure 16, all samples, except the maximum, were less than or equal to six oocysts/10 L. Most samples where *Cryptosporidium* spp. were detected had concentrations of one to two oocysts/10 L (76 samples), 22 samples had three or four oocysts/10 L and five samples had five or six oocysts/10 L. The highest concentration of *Cryptosporidium* spp. was associated with the avian faecal source group at 14 oocysts/10 L.

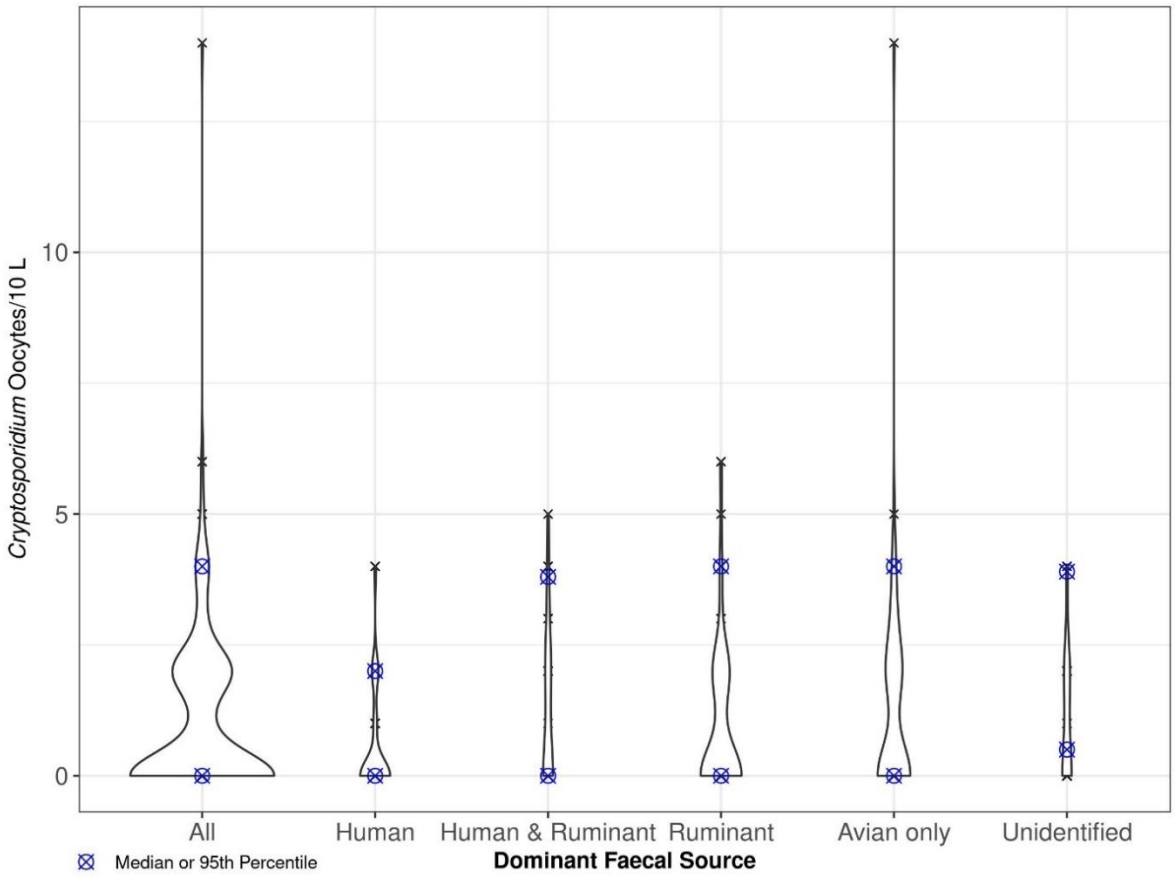

Figure 16: Concentrations of *Cryptosporidium* by dominant faecal source

Quantitative PCR for the *Cryptosporidium* species commonly found associated with illness in humans, *C. parvum* and *C. hominis,* did not detect these species in any sample.

3.5.2 *Giardia*

Giardia spp. were detected by microscopy in 129 samples (51%), ranging in concentration from one to 21 cysts/10 L (Figure 17). The median concentration was one cyst/10 L and the 95th percentile was seven cysts/10 L. There were six samples with concentrations greater than or equal to 10 cysts/10 L and 114 samples between 2 to 9 cysts/10 L. The human faecal source group had the sample with the highest concentration and human and human & ruminant had the highest median of 2 cysts/10 L.

Giardia spp. were detected by qPCR in two samples at 100 and 350 copies/10 L, but no cysts of *Giardia* were detected in either of these samples by microscopy.

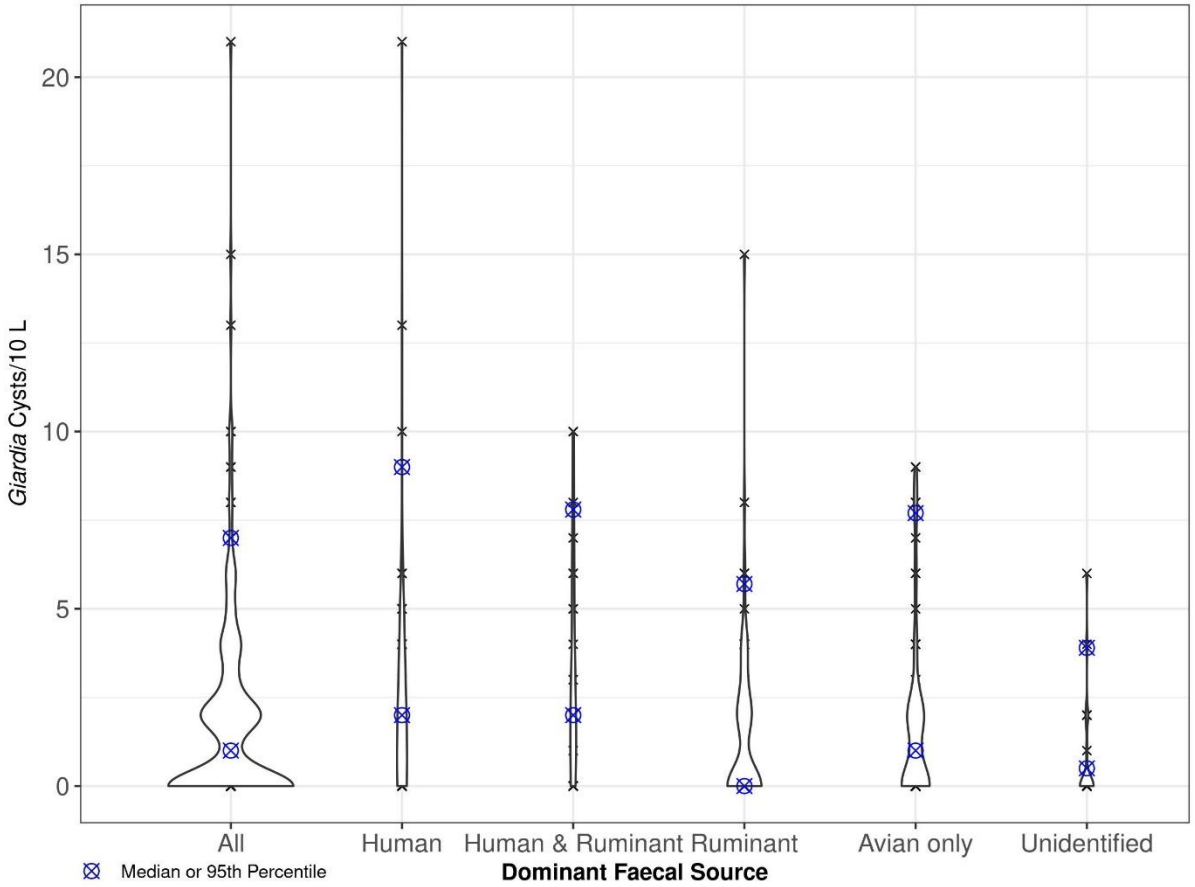

Figure 17: Concentrations of *Giardia* by dominant faecal source

3.6 FIELD AND ENVIRONMENTAL DATA

Not all councils and iwi had meters that could measure the range of parameters assessed in this study. Based on the variation observed between field and laboratory analysed in the pilot study, it was decided to use field data unless there was a laboratory that could analyse samples locally to overcome changes in physiochemical parameters during transport. Rainfall and flows were collated from council or publicly available data. As discussed in the Pilot Study a large dataset is required for robust statistical analysis, so no further analysis is undertaken at this interim stage.

The sampling app was field tested with field data being transferred directly to the database. Analytical results are also available from the database via the app.

3.7 CULTURAL HEALTH

One objective of this project was to engage with iwi to develop CHI using a Case Study approach. However, the short period over which the project was undertaken had significant implications for working holistically with our iwi partners, as we needed to be mindful of their own priorities and timeframes. We engaged with iwi who expressed an interest in cultural health assessment, or similar, and who had existing projects, concerns and/or objectives around four awa where samples for this project were already being collected. This ensured that the work aligned and supported current iwi and hapū priorities.

Iwi nominated the awa and/or tributaries and locations and ESR staff met with iwi before the sites were visited. To understand the health of the awa more than one site was selected. Due to lack of access to some culturally sensitive sites, alternative sites which were more accessible were selected. In addition, the application of CHI varies, and it was noted that a hapū may wish to develop a specific approach which better met their local context and aspirations. Some iwi and hapū participants indicated that there are a range of approaches ranging from kaupapa Māori monitoring to the resource intensive CHI. Resources may be limited; therefore, priority becomes an important factor to consider. This project presented opportunities for kaupapa Māori assessment at some sites and the more formalised CHI at other sites. This was particularly useful where iwi and hapū were developing their own kaupapa and tikanga on cultural health monitoring.

It is important to note that some sites may be wāhi tapu (sacred) and the data collected may be sensitive. For the purposes of this report, and after discussion with the Ministry, an overview is given of the four case studies rather than the scores or details of the findings This protects the Intellectual Property associated with the cultural health assessment or kaupapa Māori monitoring. For this reason, the names of the sites are not provided nor the results of the assessment. The intention was to provide the opportunity to support iwi and hapū. The case studies reflect a range of responses from well-developed CHI to a developmental approach. At two awa, a cultural health index framework had been established or adapted from previous assessments. At the other two awa cultural health assessment was being developed and a kaupapa Māori approach was taken.

3.7.1 Case Study 1

This awa had been assessed twice previously based on a CHI developed by Gail Tipa and council and refined by tangata whenua to reflect local conditions. The cultural history and context had already been documented and scores given to selected sites. After a site visit, iwi chose to re-assess three of the original five sites leveraging the expanded microbiological data available from this project to determine if there had been an improvement. As well as taking samples for microbiological analysis and undertaking physio-chemical analysis to inform elements of the atua Tangaroa, hapū members met at the sites to undertake the other aspects of assessment including other indicators of Tangaroa, Tāne Mahuta, Haumia-tiketike and Rongo-mā-Tāne, Tūmatauenga and Tāwhirimātea to determine the CHI. Compartment Bag Tests (CBT) were used in combination with laboratory data. Additional CBT were provided for more analysis of *E. coli* to support cultural health assessment at other times of the year e.g., seasonal monitoring. The CHI identified improvements across different indicators at the sites along the awa.

3.7.2 Case Study 2

The second CHI was undertaken on a river with limited public access. One site in the Phase 2.1 routine sampling programme was a tributary to the awa. As there had been no previous monitoring of the awa, this was noted as being a valuable opportunity to use cultural health monitoring tools to acquire quantitative and qualitative data. The first phase of the assessment was to understand the history of the awa. A format which had been applied previously to other awa was adapted to suit local context using qualitative and quantitative measures. Key components included: what species were living in the catchment in the past, what lives there now, and what species whānau would like to see in future. Two site visits were undertaken, one in conjunction with routine sampling for this project with hapū and the key assessor but was hampered by poor weather. A subsequent visit was made to complete cultural health monitoring. CBT for *E. coli* was used in combination with laboratory analysis. Additional CBT were provided for cultural monitoring at other times of the year e.g., seasonal monitoring or different weather conditions.

3.7.3 Case Study 3

The culturally sensitive sites in this case study were identified as part of an ongoing project between ESR and iwi and hapū. Of 20 potential sites, two were chosen based on information from LAWA (lawa.org.nz) that identified sampling sites which had a history of *E. coli* analysis and previously known to have elevated *E. coli* concentrations. Iwi also selected a further awa which had particular significance. Two sites were visited by hapū and ESR following a hui to discuss the project in the morning. Due to limited time, the third site was visited subsequently. Training was provided by ESR and council to support iwi sampling for the project. The first two sites were changed subsequently as access was problematic. As well as monitoring the water quality, an iwi team member made a video recording of iwi descriptions of cultural stream health. Eleven samples were taken at three awa and the data discussed with iwi.

This project has supported iwi goals of developing kaupapa Māori monitoring and cultural assessment across awa in the region. It is planned that some sites will have a CHI, but it is important to use different levels of monitoring to target different objectives and respond to different pressures on the awa. Some awa will require weekly monitoring, while monitoring at other awa can be less frequent. Very active community groups have been nurtured to build capacity and capability that supports assessment and subsequent improvement in the health of the awa in the region. A key goal is to upskill hapū and to have an intergenerational approach and engage tamariki in cultural monitoring. The CBT for *E. coli* has been very successful in engaging and empowering school aged tamariki and provision of the CBT will support ongoing cultural assessment.

3.7.4 Case Study 4

A hui was held with iwi, hapū, regional council and ESR to discuss the inclusion of culturally significant sites in the sampling programme. Additional sites (i.e., further to the council-nominated site) were selected, and samples collected from each of these additional sites, with hapū members (including an environmental monitor) and council staff visiting the sites together on each occasion. Iwi and hapū are currently undertaking cultural health monitoring at other sites in their rohe in relation to other kaupapa, and were able to follow a similar, intuitive framework in assessing the mauri and condition of the awa and surrounding whenua during these site visits.

In reflecting on this project, hapū members noted that they enjoyed being involved in the mahi and learned a lot from council staff about the technical aspects of data and sample collection. Moreover, they reflected that these site visits provided a valued opportunity for individuals to connect or reconnect with these places, and for cross-hapū and cross-generation sharing of mātauranga, including tikanga, the history of the awa and its resources, and understanding how that has changed. The addition of the microbiological

data collected through this project was considered important in supporting the cultural monitoring of the awa, particularly with regards to mahinga kai and whether it is safe to harvest or consume. There is considerable interest in the ‘next steps’ for this kaupapa and a desire to explore opportunities for further whānau involvement, from tamariki and rangitahi, to kaumātua.

4. DISCUSSION

4.1 FAECAL SOURCE TRACKING AND LAND USE CATEGORIES

The 1998-2000 FMRP study (McBride et al 2002) on which the New Zealand Guidelines (MfE 2003) are based on land usage assessed by observation, as FST was not available. The inconsistency between observed land use and FST markers observed in the current study illustrates the value of using FST to identify sources of faecal contamination more accurately, to ensure mitigation is correctly targeted.

Due to the mix of faecal sources within sites it is not possible to assume that sampling at a specific site will consistently provide samples with the predicted dominant faecal sources. Urban sites for example, are a mix of avian only, human and human & ruminant FST markers, so it is likely that more samples will be required in these areas to obtain 30 samples with high *Campylobacter* concentrations. It is also evident that upper catchment and weather have an impact. For example, in the urban sites there were ruminant contributions identified on six of nine occasions when there was significant rainfall prior to sampling. In contrast, when there was no significant rain prior to sampling events, human FST markers were dominant in these urban sites. Site 1, in particular, is strongly impacted by upper catchment ruminant contamination during rainfall events.

These results highlight the importance of repetitive sampling when elevated *E. coli* levels are detected and at different times, flow rates, and/or weather conditions to confirm the inclusion or exclusion of livestock or human contamination as faecal sources. Meijer (2012) recommended that a minimum of eight samples per location be tested before undertaking remediation. This multiple sampling regime avoids missing critical sources (human and livestock) and maximises the cost-effectiveness of mitigations.

It is evident that land use observations provide a coarse assessment of the potential sources of faecal contamination and FST markers should be used. Selection of urban sites with less farmed animals in the upper catchment could provide more data with dominant human FST markers. Therefore, data should be assessed using results from FST rather than observed land use. A summary of the presence of the pathogens across dominant faecal source groupings and the Unidentified group is:

- *Campylobacter* dominated all groups and was detected in all except one sample with human & ruminant FST markers (25/26 samples)
- STEC by presence/absence was only detected in samples which have strong ruminant FST markers (ruminant or human & ruminant)
- *Salmonella* by presence/absence was present across all faecal source groupings except Unidentified; most frequently present in samples containing human FST markers (human and human & ruminant) (26 samples)
- *Giardia* and *Cryptosporidium* were detected across all faecal source groups and the Unidentified group.

4.2 PATHOGENS

4.2.1 Bacterial pathogens

Across the 16 weeks of sampling, *Campylobacter* was the most prevalent pathogen detected in 66% of samples by culture, and the most dominant pathogen across all faecal source groups. There were 54/252 (21%) samples with *Campylobacter* concentrations of 10 MPN/100 mL or more. This was higher than the 10% found in the 2020 study (Leonard et al 2020).

The increase in concentrations of *Campylobacter* highlights the importance of sampling over different seasons and in different years. From the 1998-2000 study, it would be expected that *Campylobacter* prevalence and concentrations would peak during spring (McBride et al 2002), a period which has not yet been sampled within this programme and is discussed in section 4.5.

The bacterial pathogens *Salmonella* and STEC were detected at similar prevalence levels to the 2020 study. *Salmonella* was detected in 16% of samples compared to 18% in 2020, with STEC at 2.8% compared to 2% in 2020. This continued low prevalence of *Salmonella* and STEC and the differences between the results of the presence/absence tests and qPCR highlights the low concentrations of these pathogens in the samples. If the pathogens are present infrequently and/or in low concentrations, then the probability of detection of the pathogen in any single sample is low which leads to the variability seen in the results. It supports the modification of the analytical method to enrich and test for STEC presence/absence rather than via culture and quantification by MPN. This significantly reduces time and resources.

The Pilot Study (Leonard et al 2020) identified that a statistically robust QMRA analysis would require samples which had concentrations of *Campylobacter* at ≥10 MPN/100 mL for each land use/dominant faecal source. From Table 4 it can be seen that good progress has been made in collecting samples across a variety of sites with different faecal source contamination.

Table 4: Prevalence of samples with *Campylobacter* concentration greater than, or equal to 10 MPN/100 mL

FST	Total number of samples	Number of samples with *Campylobacter* ≥10 MPN/100 mL *	Number of sites where *Campylobacter* is > 10 MPN/100 mL *
Avian	69	18	8
Human	45	2	1
Ruminant	84	20	9
Human & Ruminant	26	12	9
Unidentified Source	22	1	1
Total	246*	53	20

*246 samples - excludes six samples positive for *Campylobacter* which had technical problems during FST determination, and therefore, were unable to be assigned to a faecal source group and four samples which arrived too late for analysis

However, as shown in Figure 8, the dominant source of faecal contamination is highly varied within sites and at some sites the observed land use is always inconsistent with the FST marker results. While sampling has targeted sites likely to have a range of dominant faecal sources, results in Table 4 highlight the difficulty in targeting sites where the faecal source is solely attributed to human faecal sources such as urban sites. Only one site and 2/45 samples with dominant human markers had *Campylobacter* concentrations of 10 MPN/100 mL or more.

It is recommended that the size of the sampling programme remain at 1040 samples and more sites with human dominant faecal sources and elevated *Campylobacter* concentrations need to be identified. However, following the next set of data this should be reviewed, and it is possible that a smaller dataset may provide sufficient statistical robustness.

4.2.2 Protozoan pathogens

Cryptosporidium had a similar prevalence in the 2021 study at 41% compared to 42% in 2020. However, the prevalence of *Giardia* reduced to 51% from 81% prevalence seen in the 2020 study. While the median concentrations by microscopy had increased from 2020 (Leonard et al 2020), they remain low. The maximum concentration of *Giardia* spp. was 21 cysts/10 L and for *Cryptosporidium* spp. was 14 oocysts/10 L, which is only 2.1 and 1.4 cysts per litre respectively.

The qPCR method used directly targets the human infectious *Cryptosporidium* species *C. hominus* and *C. parvum*. So, as in the pilot study (Leonard et al 2020), although *Cryptosporidium* spp. were detected by microscopy, none of these species were measured by qPCR. *Giardia* was detected twice by qPCR, but neither of these two samples had *Giardia* cysts using the microscopy method. This is likely due to the low concentration of pathogens in the water samples, as discussed previously.

4.3 CULTURAL HEALTH

Mātauranga Māori embodies a unique and extensive body of knowledge of ecosystems and there is a continuum of tools, frameworks and methods for Mātauranga Māori monitoring from kaupapa Māori monitoring to the formalised CHI. Tools have been developed by iwi since the early 2000s. Rainforth and Harmsworth (2019) note that CHI has generally been developed with iwi and council, and may not be reflective of kaupapa Māori monitoring undertaken by many iwi and hapū. A formal CHI involves qualitative and quantitative assessment of site status, mahinga kai and cultural stream health. It requires significant investment of time and resources including organising a team with historical knowledge of the rohe, often over the weekend, access to sites of cultural significance, and setting nets to quantify mahinga kai. There are many examples of the application of formalised cultural health assessment to support major iwi objectives. However, Māori need to be able to employ the range of mātauranga Māori tools that are appropriate to their needs in their particular rohe.

During this project it was recognised that many iwi and hapū had a vast array of demands on their time and resources. Prioritisation of resources is important. Local priorities and local tikanga means that one approach is not transferable across Aotearoa or even for all situations within the rohe. This is reflected in the four case studies presented here. The project objective of *seeking culturally significant sites* provided a sound basis for engagement. The project was able to support iwi and hapū in their aspirations to build intergenerational capacity and capability, gain knowledge on awa which had not been assessed before, provide training in western science monitoring and funding for cultural assessment.

Each iwi has a different approach and is at a different phase in terms of development and application of cultural health assessment. CHI was not the level of assessment required at two of the sites. Instead facilitating kaupapa Māori monitoring, empowering hapū through training to take samples, and supporting intergenerational transfer of knowledge were key benefits from the activity. The process of determining *E. coli* using the CBT was reported to

strongly engage and empower tamariki. Working with tamariki is an important aspect of building intergenerational knowledge.

Timing is also important. In the awa where the case studies were situated, iwi had already identified issues and had concerns about the health of the awa and were working towards their goals for improved freshwater outcomes. Another key observation is that engagement is based on building trusted relationships with iwi, supported by a commitment to an ongoing partnership that extends to other issues of awa health such as mahinga kai, which was frequently mentioned as being of the highest priority.

In future, support for cultural assessment needs to retain this flexibility to ensure that iwi and hapū are able to apply and build capacity in mātauranga Māori in a way that suits local requirements and tikanga. It is important to note that undertaking CHI would not be as beneficial as the broader objective to support kaupapa Māori monitoring to achieve hapū goals.

4.4 CULTURALLY SENSITIVE SITES

4.4.1 Sites identified in the current monitoring plan

Some of the freshwater bodies selected by Milne et al (2018) are culturally sensitive at the sampling sites identified. These include:

- Hatea River at Whangarei Falls – The river forms at the confluence of the Mangakino and Waitaua streams at the Whangarei Falls. There is a long history of habitation in the area and gathering of tuna (eels) and koura (crayfish) and the Pehiāweri Marae is nearby. This site has a history of cultural practices associated with healing and preparing the dead for burial. The base of the Falls was once tapu. The falls are the tribal boundary of Ngati Hau – “Otuihau”. Public recreational use of the area is one of the objectives of the Parihaka and Hatea River Reserves Management Plan.
- Arahura River at State Highway 6. This river is famous as a source of greenstone. Te Rūnanga o Ngāti Waewae is based at Arahura and the marae overlooks the river near the State Highway 6 bridge. The Arahura river is of high cultural significance to Ngāi Tahu, and there is a long history of settlement in the area. The title to the bed of the Arahura River was vested by the Crown to the Mawhera Incorporation and became part of the Ngāi Tahu Treaty Settlement in 1996.
- Powell Creek at Golden Bay. This is a tributary to the Motupipi River, which is of high cultural significance to local iwi in Mohua Golden Bay. Its name indicates its significance as a source of food at the mouth (pipi) of the awa.
- Kaiate River at Kaiate Falls Rd. The Kaiate Falls (Te Rerekawau) are identified in the Ngā Pōtiki iwi environmental plan (Conroy and Donald 2019). The issue with poor water quality at this site is noted and there is a policy for inclusion in research and monitoring at this site.
- Henley Lake at Masterton. This lake is artificial, created from wetlands in 1988 on the northern banks of the Ruamahanga River. The Ruamahanga River is of high cultural significance to Te Ngāti Kahungunu Ki Wairarapa supporting habitation, food gathering and cultural practices.
- Wakapuaka River at Paremata Flats Reserve is a pubic area and has been identified as an important traditional site for Ngāti Tama for food gathering and other cultural practices.

- Wairoa River at State Highway 2. The Takitimu marae is located opposite the sampling site at the ramp for the Water Ski Club. There is a history of settlement in the area, with archaeological sites along the bank.

4.4.2 Additional sites

Ngāi Tahu ki Murihiku Freshwater Objectives September 2020 highlight that water quality improvement is an objective for both Moffatt Creek and Otepuni Creek. A different site, Iron Bridge on the Ōreti River, was proposed as suitable for the monitoring programme as it is a site of cultural significance and has recreational value.

Pukerangiora Hapū, Otaraua hapū and Te Kotahitanga o Te Atiawa, the iwi post-settlement governance entity, have identified other sites on the Waitara River which are culturally sensitive. They are included in the cultural health assessment.

Ōwairoa is identified by Ngā Pōtiki as a culturally sensitive site and the Environmental plan policy is for participation in research and monitoring.

4.5 FUTURE SAMPLING

An important feature of this project is that there is coverage of diverse geographic locations and land uses. This ensures that there is good coverage across different environmental factors such as terrain, soils, rainfall, temperature, and land use.

Data from McBride et al (2002) highlighted temporal variation in microorganism prevalence. *Giardia* concentrations peaked over late autumn and again in early spring, while *Cryptosporidium* peaked in late spring and *Salmonella* peaked in spring. *Campylobacter* peaks tend to be in the summer months. Even though protozoa concentrations are low, it is important to ensure that sampling covers the peaks expected in spring in at least one year of sampling. Lambing and calving periods will differ depending on climate, so it is important that there is a routine sampling pattern established to ensure the peaks are captured across the country.

Due to the high variation in faecal sources at most of the different land use sites, it is necessary to continue to have a broad coverage. Therefore, even though Sheep & Beef have a good number of samples with high *Campylobacter* concentrations, continued sampling across all faecal sources is important until the required number of samples is achieved. The number of urban sites has been increased by using the original site data reported by Milne et al (2018). Three sites have been added to increase the chances of having an increased number of samples with concentrations of *Campylobacter* at 10 or more MPN/100 mL.

Site selections also need to include culturally sensitive sites. Eleven sites have been chosen from discussion with iwi and status in iwi management plans. The Arahura River has a high level of cultural significance. However, it has been removed from the list due to unknown interference in PCR analyses.

The following list of 32 sites is recommended as there is good knowledge of their characteristics from the 2020 and 2021 studies (Table 5). Two Low Impact sites are included as controls and 11 culturally significant sites (33%) have been included.

Table 5: Proposed sites for future sampling in Phase 2.2

Site	Region	Land use
Moffat Creek at Moffat Road	Southland	Dairy
Murchison (Neds) Creek at Murchison	Tasman	Dairy
Piako River at Paeroa-Tahuna Rd	Waikato	Dairy
Powell Creek at Golden Bay	Tasman	Dairy
Selwyn River at Coes Ford	Canterbury	Dairy
Waiotu River at State Highway 1	Northland	Dairy
Waitara River at Bertrand Rd	Taranaki	Dairy
Manakau Stream at State Highway 1 Bridge	Manawatu-Whanganui	Sheep & Beef
Hatea River at Whangarei Falls	Northland	Sheep & Beef
Kaiate River at Kaiate Falls Rd	Bay of Plenty	Sheep & Beef
Papakura River at Alfriston-Ardmore Road	Auckland	Sheep & Beef
Wairoa River at State Highway 2	Hawkes Bay	Sheep & Beef
Wakapuaka River at Paremata Flats Reserve	Nelson	Sheep & Beef
Wharekopae River at Rere Rockslide	Gisborne	Sheep & Beef
*Nuhaka River at Opoutama Rd	Hawkes Bay	
Heathcote River at Catherine St	Canterbury	Urban
Otepuni Creek at Nith Street	Southland	Urban
Porirua Stream at Town Centre	Wellington	Urban
Sawyers Creek at Dixon Park	West Coast	Urban
Taylor River at Riverside Park	Marlborough	Urban
Avon River at Antigua Boatsheds	Canterbury	Urban
Oteha Stream at Days Bridge	Auckland	Urban
*Opanuku River at Candia Rd	Auckland	Urban
*Waiwhetu Stream at Whiles Line East	Wellington	Urban
*Maitai River Collingwood St bridge	Nelson	Urban
Henley Lake at Masterton	Wellington	Urban (Birds)
Hutt River at Kaitoke	Wellington	Low Impact
Ōwairoa River, Waitoa Rd	Bay of Plenty	Low Impact (mostly)
Manganui River	Taranaki	Dairy
Motupipi River	Tasman	Dairy
Tributaries to Ruamāhanga River	Wellington	Sheep & Beef /Dairy
*Ōreti River Iron Bridge	Southland	Dairy

*Sites which are additional to the sampling undertaken in Phase 2.1.
Culturally significant sites are indicated by shading.

It is proposed that samples be taken at each site on a routine basis using a three weekly rotation, to be confirmed with councils, to manage the logistics for councils and laboratories. At critical time points, data will be assessed to confirm sites and analyses. Other sites from Milne et al (2018) may be used to augment data for certain FST markers as required. Further engagement with iwi may also identify suitable sites which can also support kaupapa Māori monitoring. Therefore, the sampling sites may change in the future depending on alignment of the results with the requirements of the project.

5. RECOMMENDATIONS

5.1 PHASE 2.1

During this phase of the project, we have established the database and commenced sampling and data analysis. The purpose of Phase 2.1 was to:

- Confirm and extend the number of sites that are monitored to ensure there is good geographical representation, covering different land uses.
- Engage with iwi to select culturally significant sites and to incorporate cultural health assessment (CHI or similar).
- Complete project planning and commence sample collection and data analysis.

In this interim report, an overview of the data has been undertaken to inform the next phases of the project. Robust statistical analysis will occur after sample collection is complete and the QMRA has commenced.

5.2 SITE SELECTION

Land use was identified as an indicator of the concentration and types of pathogens in freshwater in the 1998-2000 FMRP study. The Recreational Water Quality Guidelines (MfE 2003) incorporated land use observations in the sanitary inspection used to categorise recreational sites and to identify the most likely source of faecal contamination in the event of elevated *E. coli* concentrations. Selection of sites across a range of land uses and different geographies was, therefore, a key feature of the original sampling design (Milne et al 2018). However, data from the Pilot Study (Leonard et al 2020) and this Phase 2.1 study highlights that sources of faecal contamination will vary temporally and spatially and observed land use may not be consistent with the FST markers detected. This study has also highlighted that at on some occasions and at some sites (e.g., site 35) the source of faecal contamination is not identifiable, despite significant *E. coli* concentrations and no intense land use impact.

Thus, while land use observations can help identify sites most likely to have specific types of contamination, there needs to be an adaptive approach with ongoing assessment of the data collected to ensure that there is good coverage of significant contamination by different faecal sources. More sites with human faecal contamination as the dominant source and concurrent elevated *Campylobacter* concentrations are required, so three additional Urban sites have been added to the site list (Table 5).

Until the minimum number of samples for each dominant faecal source grouping are reached planning should be for routine sampling and collection of 1040 samples. Critical assessment points should be included in future sampling programmes to determine progress with a recommendation on the need to add or subtract sampling sites taking into account the need to have good geographical and seasonal coverage. Sampling in spring is necessary as increased *Campylobacter* concentrations are likely due to lambing and calving, and generally higher rainfall, which initiates land runoff of contaminants into water bodies.

5.3 ENGAGEMENT WITH IWI FOR CULTURAL HEALTH ASSESSMENT

The deeper engagement with iwi on four cultural health assessments highlighted the need to align proposed cultural health assessment with iwi and hapū priorities. Significant time is needed to start communications and iwi need to be involved in design of kaupapa Māori monitoring at the pre-planning stage.

It was also evident that supporting a range of kaupapa Māori monitoring tools will better meet the needs and priorities of iwi and hapū. The more formalised CHI may not be required, nor align with iwi and hapū priorities and objectives.

The list of 32 sites in Table 5 includes 11 sites of cultural significance. Unfortunately, the Arahura River, which has very high cultural significance, has not been included as there is some characteristic of the river at high turbidity which interferes with the analyses.

5.4 RECOMMENDATIONS

The following recommendations are made:

- Continue to plan for collection of 1040 samples across all faecal source groups with additional sites likely to have human FST.
- A routine sampling pattern based on a three-week rotation using the listed sites.
- Continue to analyse *Salmonella* and STEC by presence/absence supported by qPCR.
- Ensure there is a period of sampling in the spring to confirm if the other target pathogens (*Salmonella*, STEC, *Cryptosporidium*, *Giardia*) are likely to be significant.
- Extensive time is needed for iwi engagement to understand their priorities and co-design an approach. Cultural assessment needs to align with iwi and hapū priorities and their tikanga, whether that is kaupapa Māori monitoring or CHI. This will support iwi and hapū to apply and build capacity in mātauranga Māori in the most appropriate way.
- At critical time points assess data to ensure sites fit with the aims of the project and to assess the need for ongoing monitoring of *Salmonella*, STEC, *Cryptosporidium* and *Giardia*. It is possible that a smaller dataset may provide sufficient statistical robustness.

GLOSSARY

CBT Compartment bag tests. A portable water quality test that can be used to easily and rapidly assess the presence of *E. coli* in a water sample.

CHI Cultural Health Index. A tool for assessing the cultural health that qualitatively and quantitatively measures indicators of importance to iwi and hapū and produces scores.

FIB Faecal indicator bacteria. Collective term used to describe a group of bacteria, including coliforms, *E. coli* and enterococci, that are present in high concentrations in human and animal faeces, and therefore can be used to indicate contamination.

FST Faecal source tracking. Molecular methods used to identify the source of faecal contamination by targeting organisms or molecules specific to certain faecal sources e.g., human, ruminant, avian.

MPN Most probable number. An estimate of the number of viable (live) bacteria in a sample based on sub-dividing a sample into a series of dilutions and assessing the presence/absence of viable bacteria in each subdivision.

QMRA Quantitative microbial risk assessment. A framework to combine information on target pathogen(s) and the potential dose or exposure to estimate the risk of infection and/or illness. Exposure is a function of the concentration of pathogens in the water and the volume of water that might be ingested during recreation.

STEC Shiga-toxic producing *E. coli*.

WGS Whole genome sequencing. The process by which the entire DNA sequence of an organism is sequenced at a single time.

APPENDIX A: QUANTITATIVE MICROBIAL RISK ASSESSMENT PILOT STUDY 2020 - EXECUTIVE SUMMARY

Current Freshwater Recreational Guidelines and the National Policy Freshwater Policy Statement (2020) are based on data collected during the 1998-2000 Freshwater Microbiology Research Programme (FMRP), a Quantitative Microbial Risk Assessment (QMRA) and an association found between the pathogen, *Campylobacter,* and faecal indicator bacteria, *Escherichia coli.* (McBride et al 2002). There have been significant changes in land use and waste management practices in the 20 years since the FMRP raising some uncertainty over the relevance of the FMRP-derived relationships between pathogens and indicators. In view of the fundamental importance of these relationships, and the need for further confidence in the guidelines they underpin, an understanding of the current prevalence of pathogens, and relevance to human health risk is required.

This report describes a pilot study undertaken to inform the options, number of samples, costs, and logistics for a large-scale replacement study for the 1998-2000 FMRP and QMRA. The pilot also enabled a selection of new methodologies to be trialled. Due to the limited number of samples collected, the aims of the pilot did not include establishing the current state of concentrations of pathogens and faecal indicators in the rivers sampled.

There has also been a shift to acknowledge Māori values, their intergeneration perspectives and to incorporate these in science/ policy responses. The new National Policy Statement Freshwater (NPS-FW 2020 (New Zealand Government 2020) is centred on giving effect to Te Mana o Te Wai, which needs to be taken into account in the formulation of the next stage of this project.

Pilot study results

Between February and March 2020, 52 water samples from 16 rivers (initially characterised as six urban, five dairy and five sheep & beef) from around New Zealand were analysed for the levels of *Escherichia coli*, enterococci, *Campylobacter*, *Salmonella*, Shiga toxin-producing *E. coli* (STEC), *Cryptosporidium*, *Giardia*, norovirus, enterovirus, adenovirus and a set of faecal source tracking (FST) markers (human, ruminant and wildfowl). The river sampling sites were selected on the basis the site regularly had elevated *E. coli* concentrations recorded during previous monitoring.

Key results of the pilot study were:

- *E. coli* were detected in all samples with 17 samples from eight different rivers with ≥1,000 *E. coli* MPN/100 mL.
- All the rivers contained wildfowl markers, and were the only source identified in 10 of the samples. Four of these samples were taken from two of the rivers, which based on observed land use would have been characterised as Sheep & Beef or Dairy.
- All six urban rivers contained human FST markers, in all the samples tested. However, five of the samples from two of the urban rivers also contained significant concentrations of ruminant FST markers.
- Conversely while the ruminant FST marker confirmed the observed land use of four dairy, and four sheep & beef rivers, two of these rivers also contained human markers in at least some of the samples tested.

- Using a combination of culture and quantitative PCR (qPCR) based methodologies, *Campylobacter* were detected in 37 samples from 14 rivers, *Salmonella* in 12 samples from 8 rivers, and shiga toxin-producing *E. coli* (STEC) in 13 samples from 11 rivers. Using culture methods, the maximum level of *Campylobacter* was 92 MPN /100 mL, with 85% of detections ≤4.1 MPN/100 mL, while *Salmonella* and STEC detections were all ≤0.25 MPN/100 mL.
- Whole genome sequencing confirmed likely pathogenic strains of *Campylobacter, Salmonella* and STEC.
- Using traditional microscopy techniques *Giardia* was detected in 42 samples from 15 rivers, with 40 samples having ≤ 24 cysts/100 L, and a maximum of 250 cysts/100 L. C*ryptosporidium* was detected in 22 samples from 12 rivers with a maximum of 31 oocysts/100 L. Quantitative PCR only detected *Giardia* in one sample, and *C. parvum* and *C. hominis*, were not detected in any sample by qPCR.
- Human adenovirus (HAdV) was not detected in any sample, while noroviruses were detected in five rivers, and enterovirus in two rivers. However, the concentration of viruses in the samples were too low to quantify and only one sample contained both viruses.
- A good correlation was achieved with qPCR and *E. co*li (R^2= 0.85) and a reasonable correlation with enterococci and qPCR for samples which had high concentrations of human FST.

Direct comparisons with the 1999-2000 Freshwater Microbiological Research programme (FMRP) survey are limited by differences in detection levels, methodology, pilot study sample size and targeting of more contaminated rivers. With those caveats in mind:

- The pilot study samples had higher average levels of *E. coli*, being more like the samples from FMRP with highest levels of *E. coli*.
- The pilot study had a higher prevalence of Giardia and Cryptosporidium, however the concentrations detected were similar to FMRP.
- Adenovirus were detected in a third of FMRP samples but were not detected in the 2020 pilot study.
- In the FMRP 9% of samples had *Campylobacter* >110 MPN/100 mL, while the maximum level in the pilot study was 92 MPN/100 mL. Adjusting for differences in detection limits, the frequency of detection of *C. jejuni* was higher in the pilot study (46%), than the FMRP (30%).

The key conclusions from these observations are:

- Potentially pathogenic micro-organisms were detected in 94% of the samples, but generally at very low levels. Above water quality criterion for *E. coli* of 540 MPN/100 mL, the number of types of pathogens and their prevalence increased compared to below the criterion.
- Quantitative PCR is likely to detect pathogens when present at higher levels and does allow estimation of pathogen presence at lower levels. The lack of an isolate however means whole genome sequencing cannot be used to confirm pathogenic potential. Further calibration work is required to convert qPCR results to a concentration for input into the QMRA.

- Half the samples with viruses detected were associated with human contaminated samples, but the other half were from rivers without obvious human sources of contamination.
- Faecal source tracking markers were readily detected in almost all the samples from all the rivers allowing a source identification.
- Land use has been shown to influence variations in pathogens and provides information on potential sources important for management. Faecal source tracking was a tool unavailable in 1998-2000, and in this pilot study confirmed the frequent occurrence of more than one source of contamination, and also that observed land use does not always match the important sources of contamination. This is important information for effective management of water quality Future studies should include both visual confirmation of land use and FST.

A large-scale replacement for the 1998-2000 FMRP

A large-scale study akin to the 1998-2000 FMRP could be undertaken. We would recommend analysis as illustrated in Figure 1 enumerating by culture *E. coli*, enterococci, and *Campylobacter*, isolating *Salmonella* and STEC, genome sequencing of isolates from each, and quantitative PCR for indicator organisms and pathogens as illustrated below. Protozoa are included as a desirable option but challenging to achieve due to sampling logistics.

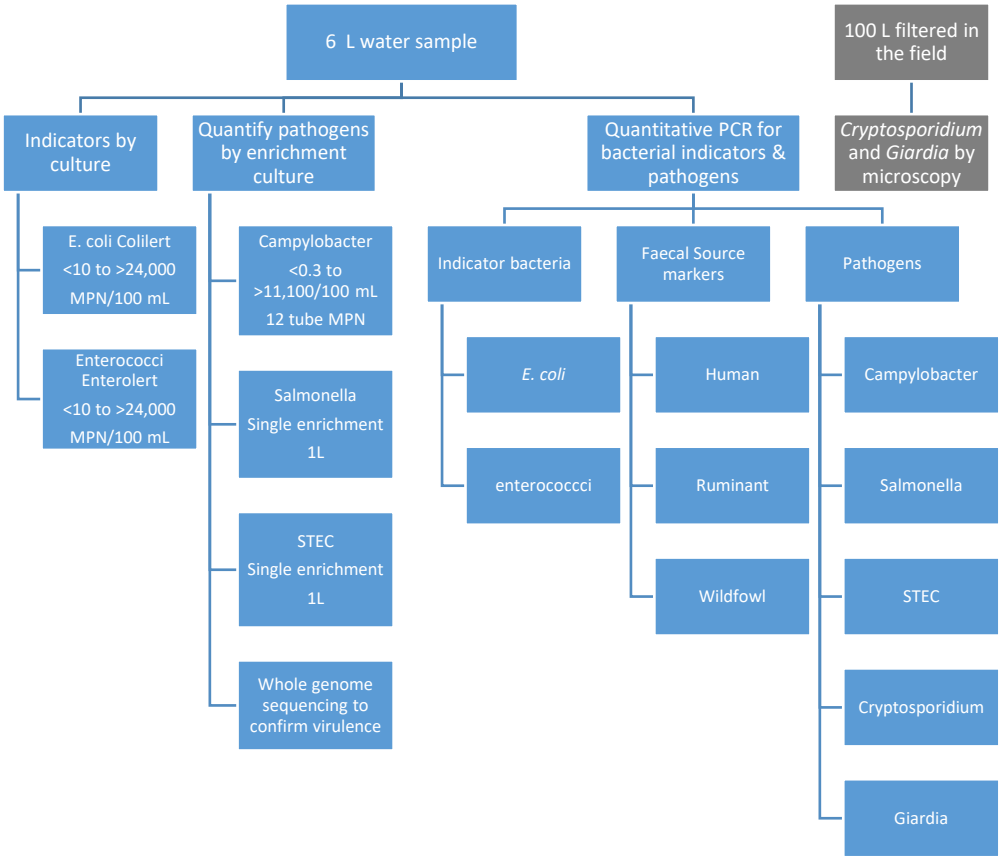

Figure 1: Analysis recommendations for survey

For this study we would recommend sampling 40 rivers monthly over a 2-year period to provide at least 26 samples from each river. We propose four categories of rivers (human, dairy, sheep & beef, wildfowl/natural), which with ten rivers in each category would provide a minimum of 260 river samples per category and 1,040 samples in total. Iwi should be engaged in site selection to ensure that sites of cultural significance are chosen. These sites may not have a history of testing as Council selection criteria for monitoring may differ.

Although the levels were low, the frequent detection of protozoa in this pilot study highlights the importance of this group of pathogens. This may reflect the choice of sites known to have elevated *E. coli* concentrations. Logistically they add significantly to the workload for samplers and timing, necessitating at least an hour and a half at each site. We have included them in this full study proposal, but recognise that many councils would need additional resourcing to undertake this. Separation of the protozoa sampling into a dedicated study may be a more practical option.

The exclusion of the analysis of samples for viruses from this proposed full study reduces analytical costs considerably. There is limited evidence from this study or others that virus levels will correlate with other indicator organisms, except at very high levels (Korajkic et al. 2018). At those levels (where faecal source tracking markers and indicators are high and easily detected) the risks from viral and protozoa can be predicted much more readily based on assessment of faecal indicators and sources using risk assessment approaches.

Previous study costs based on 1200 samples were estimated at $7.9 million, with $5.5 million directly related to laboratory analysis costs. Following this pilot study, we estimate based on 1,040 samples, including the protozoa component, revised laboratory analysis costs of $3.4 million, and a total study cost of $5.25 million. This doesn't allow for sampling costs, which particularly for protozoa may need direct funding.

This is significantly less than estimates prior to the pilot study. However, it is important to consider the impact this work would have.

Alternative considerations

To make a significant impact in the near future on improving recreational water quality management, we propose consideration of complimentary work streams to strengthen existing freshwater recreational water quality guidelines to develop a framework. The three-step framework would support evidence based decisions on how to investigate water quality when guidelines are exceeded, supported by worked examples and could be undertaken in phases.

There are three key components to this.

- Step 1: Detection of faecal contamination using indicators such as *E. coli* (possibly in conjunction with enterococci) and other chemical or environmental measures to compare against guidelines.
- Step 2: Explicit guidance on how to investigate water quality to determine sources of contamination, and therefore appropriate mitigations or interventions.
- Step 3: If the source analysis suggests low risk sources of pollution, or sources which mitigations or interventions are politically, socially or economically challenging, then site specific health risk assessment using tools such as Quantitative Microbial Risk Assessment (QMRA) and scenario modelling may be required to reclassify risk, or to support necessary interventions when the guidelines expectations are not met.

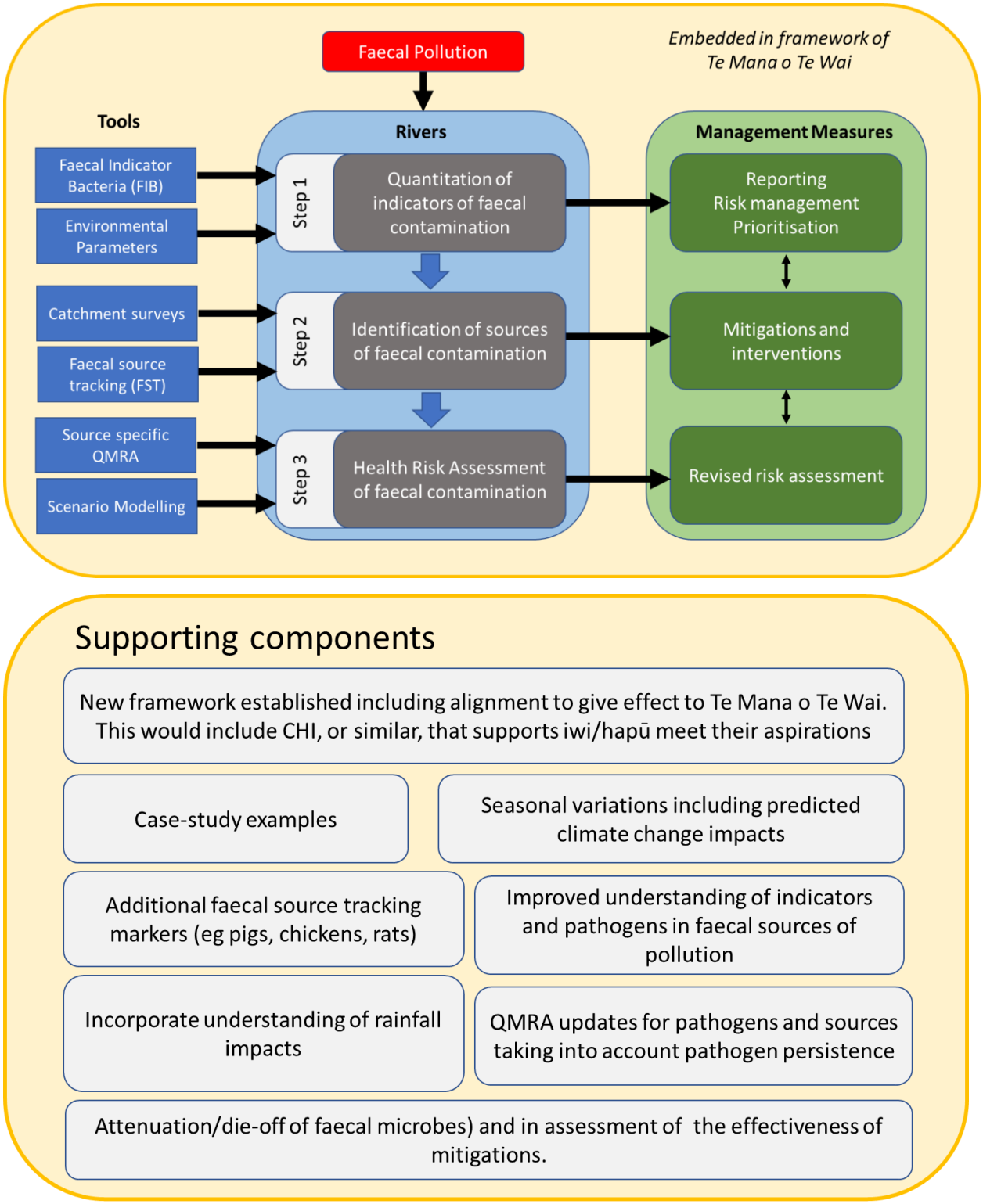

Figure 2: Conceptual framework for water quality assessment. QMRA, quantitative microbial-risk assessment.

This needs direct linkage with industry and research programmes such as Our Land & Water to guide intervention options, and ongoing monitoring programmes for trend analysis of water quality. Most importantly this needs to give effect to Te Mana o Te Wai and incorporate Māori values and approaches. The pilot study-initiated engagement with iwi/hapū at each site, which needs to be developed and extended to co-design a new framework which encompasses Mātauranga Māori.

APPENDIX B: SITES

The sites selected for the Phase 2.1 study are summarised in Table 6 below:

Table 6: Sampling sites and dominant land use

Site	Region	Land use
Moffat Creek at Moffat Road	Southland	Dairy
Murchison (Neds) Creek at Murchison	Tasman	Dairy
Piako River at Paeroa-Tahuna Rd	Waikato	Dairy
Powell Creek at Golden Bay	Tasman	Dairy
Selwyn River at Coes Ford	Canterbury	Dairy
Waiotu River at State Highway 1	Northland	Dairy
Waitara River at Bertrand Rd	Taranaki	Dairy
Manakau Stream at State Highway 1 Bridge	Manawatu-Whanganui	Sheep & Beef
Hatea River at Whangarei Falls	Northland	Sheep & Beef
Kaiate River at Kaiate Falls Rd	Bay of Plenty	Sheep & Beef
Papakura River at Alfriston-Ardmore Road	Auckland	Sheep & Beef
Wairoa River at State Highway 2	Hawkes Bay	Sheep & Beef
Wakapuaka River at Paremata Flats Reserve	Nelson	Sheep & Beef
Wharekopae River at Rere Rockslide	Gisborne	Sheep & Beef
Ashburton River at State Highway 1	Canterbury	Sheep & Beef
Heathcote River at Catherine St	Canterbury	Urban
Otepuni Creek at Nith Street	Southland	Urban
Porirua Stream at Town Centre	Wellington	Urban
Sawyers Creek at Dixon Park	West Coast	Urban
Taylor River at Riverside Park	Marlborough	Urban
Avon River at Kerrs Reach	Canterbury	Urban
Oteha Stream at Days Bridge	Auckland	Urban
Avon River at Antigua Boatsheds	Canterbury	Urban (Birds)
Henley Lake at Masterton	Wairarapa	Urban (Birds)
Arahura River at State Highway 6	West Coast	Low Impact
Hutt River at Kaitoke	Wellington	Low Impact
Kaiate River Control	Bay of Plenty	Low Impact
Manganui River	Taranaki	Dairy
Waitara River	Taranaki	Dairy
Motupipi	Tasman	Dairy
Wakapuaka - 2 locations in addition to Paremata Flats Reserve	Nelson	Sheep & Beef
Waipoua River	Wairarapa	Sheep & Beef
Kopuaranga River - 2 locations	Wairarapa	Sheep & Beef
Whangaehu River - 2 locations	Wairarapa	Dairy

The shaded sites were selected for cultural health assessment.

APPENDIX C: METHODS

C.1 Bacterial Methods

C.1.1 *E. coli* and enterococci

Water (10 mL) was analysed for *E. coli* and enterococci using IDEXX Colilert and Enterolert assays respectively (APHA 2017a, APHA 2017b). This provides a detection range of <10 MPN/100 mL up to >24,000 MPN/100 mL.

C.1.2 *Campylobacter*

For quantitative analysis of *Campylobacter* by MPN, water volumes of 1 x 1000 mL, 3 x 10 mL 3 x 1 mL and 3 x 0.1 mL were analysed (APHA 2017c; Medeiros et al 2002; ISO 2017). The 1000 mL and 100 mL aliquots were filtered through 0.22 µm filters and placed into 100 mL and 25 mL of Preston broth (OPM 322), respectively. The 1 mL and 0.1 mL aliquots were directly added to 25 mL Preston broth. Broths were incubated at 41.5 +/- 1˚C for 24 h, and then a loopful plated onto modified charcoal-cefoperazone-deoxycholate agar (mCCDA) plates. Plates were incubated at 41.5 +/- 1˚C for 44 h, and then 2-4 putative *Campylobacter* colonies were transferred to Columbia Blood Agar (CBA) plates.

Where putative *Campylobacter* colonies were observed, conventional PCR using the *Campylobacter* multiplex assay of Wong et al (2004) was performed on DNA extracted from the original Preston broth culture. Briefly, 1 mL of broth culture was centrifuged at 7,500 g for 10 min. The pellet was then resuspended in 1 mL PBS and centrifuged at 7,500 g for 10 min. The pellet was then resuspended in 0.4 mL PBS and heated at 95°C for 10 min. Samples were then centrifuged at 16,000 g for 5 min and the supernatant used in the PCR. Visualisation of PCR amplicons was carried out on the MultiNA Microchip Electrophoresis System (Shimadzu) using fluorescence detection of the DNA products.

Where the PCR identified a pure *Campylobacter* culture, the result was validated via PCR from one colony on the CBA plate. Briefly, colonies were resuspended in 0.5 mL PBS then centrifuged at 5,000 g for 10 min at 4˚C. The pellet was then resuspended in 0.5 mL of 2% Chelex 100 Sodium form (Sigma Cat#C7901) in sterile Milli-Q water (Millipore, Merck). The Chelex solution was then heated at 95°C for 10 min, centrifuged at 16,000 g for 5 min and the supernatant used in the PCR. Where the PCR identified a mixed culture, colony PCR was performed on multiple colonies from the CBA plate. Figure 18 below shows a typical result for *Campylobacter* PCR detection.

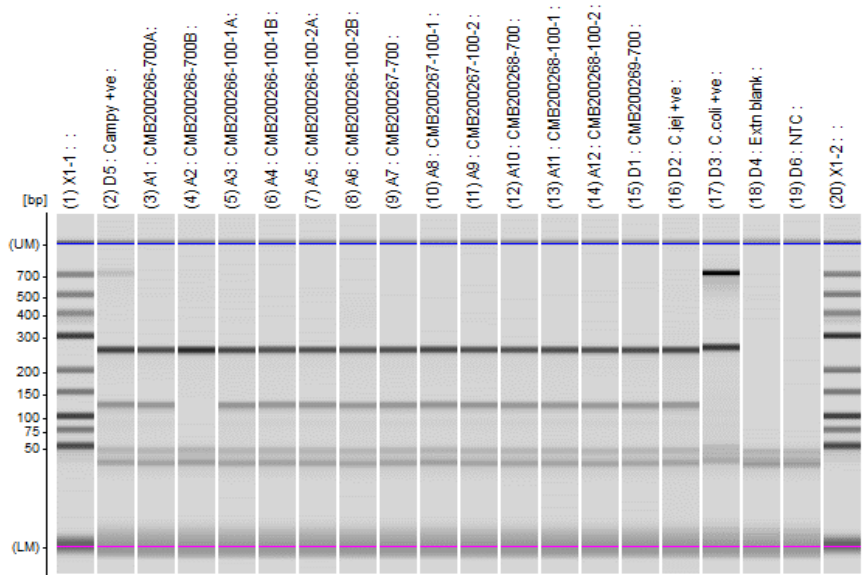

Figure 18: An example of *Campylobacter* PCR detection results

Expected band sizes for the targets are 695 bp for *C. coli* (695 bp), 246 bp for thermotolerant *Campylobacter* and 99 bp for *C. jejuni*. Note there is a slight bp offset for targets relative to the size markers (lanes 1 and 20) as displayed on the image.

C.1.3 *Salmonella*

For presence/absence analysis of *Salmonella*, water was filtered through 0.45 µm filters, and placed into 25 mL of buffered peptone water (BPW) broth. Broths were incubated at 37˚C for 18 h and then 1 mL of BPW was transferred into 10 mL of Muller-Kauffmann Tetrathionate Novobiocin Supplement (MKTTn) broth, and 0.1 mL of BPW was transferred into 10 mL of Rappaport-Vassiliadis Soya (RVS) peptone broth, which were then incubated at 37˚C and 41.5˚C respectively for 24 h. A loopful of each broth was plated onto Xylose Lysine Deoxycholate (XLD) and Hektoen Enteric agar and incubated for 24 h at 37˚C. Putative *Salmonella* colonies were plated on MacConkey agar and tryptic soy agar (TSA), and inoculated into tryptone broth and urease broth, and onto triple sugar iron agar (TSI) or Lysine Iron Agar (LIA) slopes. Isolates that were consistent with *Salmonella* then had polyO and polyH serology undertaken. If still indicative of *Salmonella*, Microgen biochemical testing was undertaken. *Salmonella* isolates were then whole genome sequenced.

C.1.4 STEC

For presence/absence analysis of STEC, water was filtered through 0.45 µm filters, and placed into 25 mL of modified Trypticase Soy Broth (TSB) with novobiocin (ISO 2012). Broths were incubated at 37˚C for 24 h and then tested by STEC multiplex PCR for the *eae*, *hly*A, *stx*1 and *stx*2 genes (Paton and Paton, 1998). Broths that tested positive were plated on MacConkey Agar with Sorbitol, Cefixime, and Tellurite (CT-SMAC) and MacConkey STEC chrome agar and incubated for 24 h at 37˚C. Twenty colonies from each plate were streak isolated and purified on CBA plates, and then multiplex PCR performed on pools of 5 isolates (Figure 19). If a pool was positive, then individual isolates were tested by multiplex PCR. This provided information on presence or absence of STEC strains.

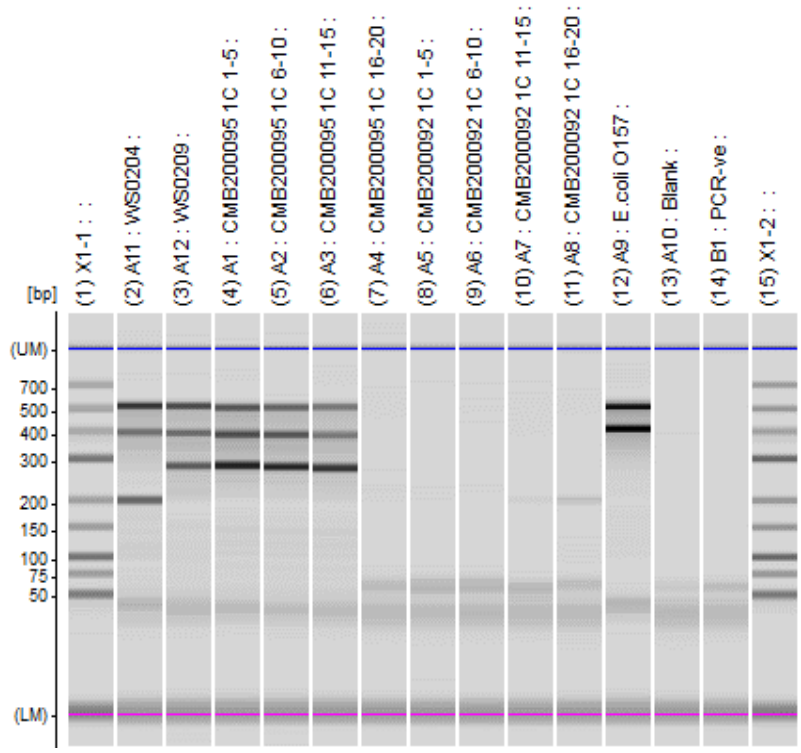

Figure 19: An example of STEC PCR detection results

Lanes 2, 3, 12 are positive controls, lanes 13 and 14 are negative controls. The remaining lanes contain pools of isolates being tested. Expected band sizes for the targets are 534 bp for *hly*A, 384 bp for *eae*, 255 bp for *stx*2, and 180 bp for *stx*1. Note there is a slight bp offset for targets relative to the size markers (lanes 1 and 15) as displayed on the image.

C.2 Protozoa

Protozoa were analysed using USEPA Method 1623: *Cryptosporidium* and *Giardia* in Water by Filtration/IMS/FA (USEPA 2005). Briefly, 10 L of water was filtered through an IDEXX Filta-Max® filter. *Cryptosporidium* oocysts and *Giardia* cysts retained on the filter were eluted and the eluate centrifuged to pellet the (oo)cysts which were isolated using anti-*Cryptosporidium* and anti-*Giardia* immunomagnetic beads and separated from the other material. The (oo)cysts were stained on well slides with a fluorescent label and DAPI[6]. The stained samples were then examined using UV fluorescence and differential interference contrast microscopy. The number of objects on the slide that met the size, shape and fluorescence characteristics of *Cryptosporidium* and *Giardia* (oo)cysts is reported. Massey University have determined that the recovery rate from river water varies from 15-55%. The data is reported as the actual count.

[6]DAPI 4',6-diamidino-2-phenylindole

C.3 Quantitative PCR

Quantitative PCR methodology is described in more detail in Gilpin et al (2018). Two litres of water were collected from each river and up to 1 L was filtered in duplicate through a 0.45 µm Millipore filter (range of 400 mL to 2 L total). CD1 buffer from the PowerSoil Pro kit (Qiagen), spiked with Salmon Sperm DNA (Sigma) was added to the filter and vortexed. Filters were stored at -20˚C. Sterile beads were added to the filter, and tubes placed in a Biospec Minibead beater, where they were beaten for 3 min at 2,500 rpm, centrifuged at 3,500 g for 5 min and supernatant transferred to a new tube. The volume was adjusted to 650 µL using CD1 buffer, if required. Samples were extracted using the PowerSoil Pro protocol on the QiaCube extraction robot. Filters were eluted in 80 µL of elution buffer and then duplicate filters were combined to a final volume of 160 µL for each water sample.

Quantitative PCR analysis was undertaken on a LightCycler 480 (Roche), with each amplification performed in duplicate using 2 µL of DNA extract. The PCR targets are given in Table 7. Each run included negative and positive controls, and gBlock or Ultramer standard curves. All samples were initially tested using primers specific for Sketa salmon sperm as a process and inhibition control. Significant inhibition was only detected in five samples from the Arahura River when turbidity was elevated, and this issue could not be resolved. No further adjustment of samples was undertaken for the other samples. PCR assays were confirmed to have an efficiency of >90%. To enable relative comparisons between assays, a common slope (1.926) and y intercept (Cycle threshold (CT) 40) was used in calculations to convert CT to copy number equivalents/100 mL of filtered water.

Copies/100 mL = (1.9263626137313^(40-CT))*((160/2)*(100/Volume filtered))

Table 7: Target bacterial genes and methods for qPCR

Micro-organism	Target	Reference
qPCR Bacteria		
E. coli	23S rDNA	Chern et al 2011
Enterococci	23S rDNA	USEPA 2013, Haugland et al 2012
C. jejuni	*map*A	Best et al 2003
C. coli	*ceu*E	Best et al 2003
Salmonella spp.	*inv*A and *ttr*	Hoorfar et al 2000, Malorny et al 2004
STEC	*eae*, *stx*1, *stx*2	Derzelle et al 2011, Ibekwe et al 2004
Cryptosporidium parvum, C. *hominis*	18S rDNA	Mary et al 2013
Giardia	beta-giardin gene	Baque et al 2011
Faecal Source Markers		
General	Bacteriodales 16S rRNA	Shanks et al 2009, Shanks et al 2010
Human	Bacteroides HF183 crAssphage CPQ_056 (56rass) *Bifidobacterium adolescentis* (BiADO)	Ahmed et al 2019
Ruminant	Bacteroidales 16S rRNA (BacR)	Reischer et al 2006
Avian	GFD – Unclassified *Helicobacter* spp. 16S rRNA gene	Devane et al 2007, Green et al 2012

C.4 Whole Genome Sequencing

All bacterial pathogen isolates were recovered from -80°C storage by streak plating onto Columbia blood agar (CBA) and incubation at 37°C for 24-48 h. A single colony was inoculated into 10 mL tryptone soya broth and incubated at 37°C for 18 h prior to DNA extraction. One mL of broth culture was used for DNA extraction using the Qiagen Dneasy Blood and Tissue Kit QiaCube (Qiagen, Hilden, Germany). DNA quality and concentration were assessed using Nanodrop (Thermo Fisher Scientific, Waltham, MA. USA), Qubit™ and PicoGreen® (Quant-iT; Thermo Fisher Scientific). Sequencing libraries containing 1 ng of DNA were prepared using Nextera XT chemistry (Illumina, San Diego, CA, USA) for 150 bp pair-end sequencing run on an Illumina NextSeq sequencer according to the manufacturer's recommendations (Illumina).

Sequence quality and species identification was determined using the Nullarbor pipeline (Seemann et al accessed 2020). Sequence quality was evaluated on a per genome basis using BioNumerics version 7.6.3 (Applied Maths, NV, Belgium). All genomes passed the basic quality metrics for raw sequence data from Illumina sequencers of average Q-score >30 in both reads and at least 40 X average coverage with expected genome sizes (Lindsey et al 2016). Genetic variability within each taxa was evaluated using Multi-Locus Sequence Typing (MLST), and whole genome MLST (wgMLST). The wgMLST schema was assessed within BioNumerics. Phylogenetic cluster analysis of the isolates was investigated using wgMLST (categorical data values) and analysed using single-linkage algorithm.

APPENDIX D: CAMPYLOBACTER

D.1 MPN Table and confidence levels

The following data was used to determine the concentrations of *Campylobacter* and the confidence levels.

Table 8: MPN calculations and confidence intervals

Volume (mL)	1000	10	1	0.1		Confidence interval from Asympt. LogNormal	
Tubes	1	3	3	3	MPN /100 mL	CI lower limit	CI upper limit
Possible combinations:	0	0	0	0	<0.096	-	0.29
	0	1	0	0	0.096	0.014	0.18
	1	0	0	0	0.34	0.016	7.4
	1	1	0	0	3.6	0.5	26
	1	2	0	0	9.2	2.2	39
	1	3	0	0	23	5.7	94
	1	1	1	0	7.4	1.8	30
	1	2	1	0	15	4.4	49
	1	3	1	0	43	9.8	190
	1	2	2	0	21	7.3	610
	1	3	2	0	93	22	320
	1	3	3	0	240	58	990
	1	2	1	1	21	7.3	61
	1	3	2	0	93	22	390
	1	3	3	0	240	58	990
	1	2	1	1	20	7.1	59
	1	2	2	1	28	10	73
	1	2	3	1	36	15	89
	1	3	1	1	75	19	290
	1	3	2	1	150	45	500
	1	3	3	1	460	99	2200
	1	3	1	2	120	36	360
	1	3	2	2	210	74	620
	1	3	3	2	1100	260	4700
	1	3	3	3	>1100	4700	-

D.2 Whole genome sequencing data for *Campylobacter*

Isolates from 113 river samples were whole genome sequenced, with two isolates analysed from eight samples, three from another sample, four from another sample and five from another. Table 9 describes the species and sequence types (STs) determined for the samples which tested positive for *Campylobacter.* It was possible to determine STs for 101

isolates. *Campylobacter jejuni* was identified in 108 isolates and 10 were identified as *C. coli:* five were replicates from site 7, two from site 12, two from site 35 and one from site 29. There were 26 different MLST types and 15 new *C. jejuni* types. Three samples were possibly *C. jejuni,* seven were possibly *Campylobacter* species and one was determined not to be *Campylobacter*.

Table 9: Whole genome sequencing data for *Campylobacter*

Asterisks indicate MLST ST that have been inferred with a cut off of < 100 loci differences to an isolate with an assigned ST.

Site	Identification	Species	MLST ST
1	1-01/03/2021	*C. jejuni*	ST991*
1	1-08/03/2021	*C. jejuni*	ST2381
1	1-12/04/2021	*C. jejuni*	ST3655
1	1-15/02/2021	*C. jejuni*	ST1326
1	1-15/03/2021	*C. jejuni*	ST2619
1	1-22/03/2021	*C. jejuni*	ST2381
1	1-29/03/2021	*C. jejuni*	-
2	2-01/03/2021	*C. jejuni*	ST137
2	2-08/03/2021	*C. jejuni*	STnew3
2	2-12/04/2021	*C. jejuni*	ST3655
2	2-15/03/2021	*C. jejuni*	ST3655
2	2-22/03/2021	*C. jejuni*	ST2381
2	2-29/03/2021	*C. jejuni*	ST177
3	3-08/03/2021	*C. jejuni*	STnew2
4	4-01/03/2021	*C. jejuni*	ST45
4	4-08/03/2021	*C. jejuni*	ST991*
4	4-15/02/2021	*C. jejuni*	ST2381
4	4-15/03/2021	*C. jejuni*	ST583
4	4-22/02/2021	*C. jejuni*	ST2381
4	4-22/03/2021	*C. jejuni*	ST2381
4	4-29/03/2021	Possible *Campylobacter* species	-
4	4-29/03/2021	Possible *C. jejuni*	-
4	4-29/03/2021	*C. jejuni*	ST22
5	5-15/03/2021	*C. jejuni*	ST583
5	5-22/03/2021	*C. jejuni*	ST3660
6	6-12/04/2021	Possible *C. jejuni*	-
6	6-12/04/2021	*C. jejuni*	ST696
6	6-29/03/2021	*C. jejuni*	ST45
7	7-01/03/2021	*C. jejuni*	ST190
7	7-08/03/2021	*C. jejuni*	STnew2
7	7-22/03/2021	*C. coli*	-
7	7-22/03/2021	*C. coli*	-
7	7-22/03/2021	*C. coli*	-
7	7-22/03/2021	*C. coli*	-
7	7-22/03/2021	*C. coli*	-

Site	Identification	Species	MLST ST
7	7-29/03/2021	*C. jejuni*	STnew5
8	8-09/03/2021	*C. jejuni*	ST45
8	8-13/04/2021	Not *Campylobacter*	-
8	8-13/04/2021	*C. jejuni*	STnew6
8	8-29/03/2021	*C. jejuni*	ST2539
11	11-01/03/2021	*C. jejuni*	ST3655
11	11-08/03/2021	*C. jejuni*	STnew13
11	11-12/04/2021	Possible *Campylobacter* species	-
11	11-12/04/2021	*C. jejuni*	STnew14
11	11-15/02/2021	*C. jejuni*	ST2381
11	11-15/03/2021	*C. jejuni*	ST2381
11	11-15/03/2021	*C. jejuni*	ST2539*
11	11-22/02/2021	*C. jejuni*	ST2381
11	11-25/01/2021	*C. jejuni*	-
11	11-29/03/2021	*C. jejuni*	ST3655
12	12-01/03/2021	*C. jejuni*	ST3655
12	12-12/04/2021	*C. coli*	-
12	12-12/04/2021	*C. coli*	-
12	12-12/04/2021	*C. jejuni*	-
12	12-12/04/2021	*C. jejuni*	ST8398
12	12-15/03/2021	*C. jejuni*	ST3640
12	12-22/03/2021	*C. jejuni*	ST3655
12	12-25/01/2021	*C. jejuni*	-
12	12-29/03/2021	*C. jejuni*	ST3640
12	12-29/03/2021	*C. jejuni*	ST3845
12	12-29/03/2021	*C. jejuni*	ST3655*
13	13-03/05/2021	*C. jejuni*	ST2381
13	13-12/04/2021	Possible *C. jejuni*	-
13	13-12/04/2021	*C. jejuni*	ST21
13	13-15/02/2021	*C. jejuni*	STnew11
13	13-22/02/2021	*C. jejuni*	ST2381
13	13-22/03/2021	*C. jejuni*	STnew8
13	13-25/01/2021	*C. jejuni*	STnew4
14	14-01/02/2021	*C. jejuni*	STnew9
14	14-01/02/2021	Possible *Campylobacter* species	-
14	14-01/03/2021	*C. jejuni*	STnew12
14	14-03/05/2021	*C. jejuni*	ST3655
14	14-12/04/2021	Possible *Campylobacter* species	-
14	14-12/04/2021	*C. jejuni*	ST1965
14	14-15/02/2021	Possible *Campylobacter* species	-
14	14-15/03/2021	*C. jejuni*	STnew12
14	14-29/03/2021	*C. jejuni*	-

Site	Identification	Species	MLST ST
15	15-01/03/2021	*C. jejuni*	STnew10
15	15-22/02/2021	*C. jejuni*	ST1225
16	16-01/03/2021	*C. jejuni*	ST583
16	16-15/03/2021	*C. jejuni*	ST2381
16	16-22/02/2021	*C. jejuni*	ST2381
22	22-08/03/2021	*C. jejuni*	ST583
22	22-12/04/2021	*C. jejuni*	ST2619
22	22-15/02/2021	*C. jejuni*	ST45
22	22-15/03/2021	*C. jejuni*	ST2381
22	22-22/02/2021	*C. jejuni*	ST45
22	22-22/03/2021	*C. jejuni*	ST2381
23	23-01/03/2021	*C. jejuni*	STnew7
23	23-03/05/2021	Possible *Campylobacter* species	-
23	23-12/04/2021	*C. jejuni*	ST991
23	23-15/02/2021	*C. jejuni*	STnew1
23	23-19/04/2021	*C. jejuni*	ST4502
23	23-22/02/2021	*C. jejuni*	STnew1
23	23-25/01/2021	Possible *Campylobacter* species	-
23	23-29/03/2021	Being repeated	-
24	24-08/03/2021	*C. jejuni*	ST2381
24	24-12/04/2021	*C. jejuni*	ST5128
24	24-22/02/2021	*C. jejuni*	ST2381
24	24-22/03/2021	*C. jejuni*	ST2381
24	24-25/01/2021	*C. jejuni*	ST2539
25	25-08/03/2021	*C. jejuni*	ST45
25	25-12/04/2021	*C. jejuni*	ST991
25	25-15/02/2021	*C. jejuni*	ST177*
25	25-29/03/2021	*C. jejuni*	-
26	26-03/05/2021	*C. jejuni*	ST2381
26	26-03/05/2021	*C. jejuni*	ST2381
26	26-19/04/2021	*C. jejuni*	ST3655
28	28-03/05/2021	*C. jejuni*	ST8715
28	28-03/05/2021	*C. jejuni*	ST3655
29	29-09/03/2021	*C. jejuni*	STnew15
29	29-13/04/2021	*C. jejuni*	ST3655
29	29-13/04/2021	*C. coli*	-
30	30-08/03/2021	*C. jejuni*	ST45
30	30-12/04/2021	*C. jejuni*	ST45
31	31-19/04/2021	*C. jejuni*	ST2381
31	31-22/03/2021	*C. jejuni*	ST2381
33	33-19/04/2021	*C. jejuni*	ST9820
35	35-12/04/2021	*C. jejuni*	ST2539*

Site	Identification	Species	MLST ST
35	35-15/02/2021	*C. coli*	-
37	37-01/03/2021	*C. jejuni*	ST1225
37	37-15/02/2021	*C. jejuni*	ST3674
37	37-15/03/2021	*C. jejuni*	ST2381
37	37-19/04/2021	*C. jejuni*	ST1256
37	37-29/03/2021	*C. jejuni*	-

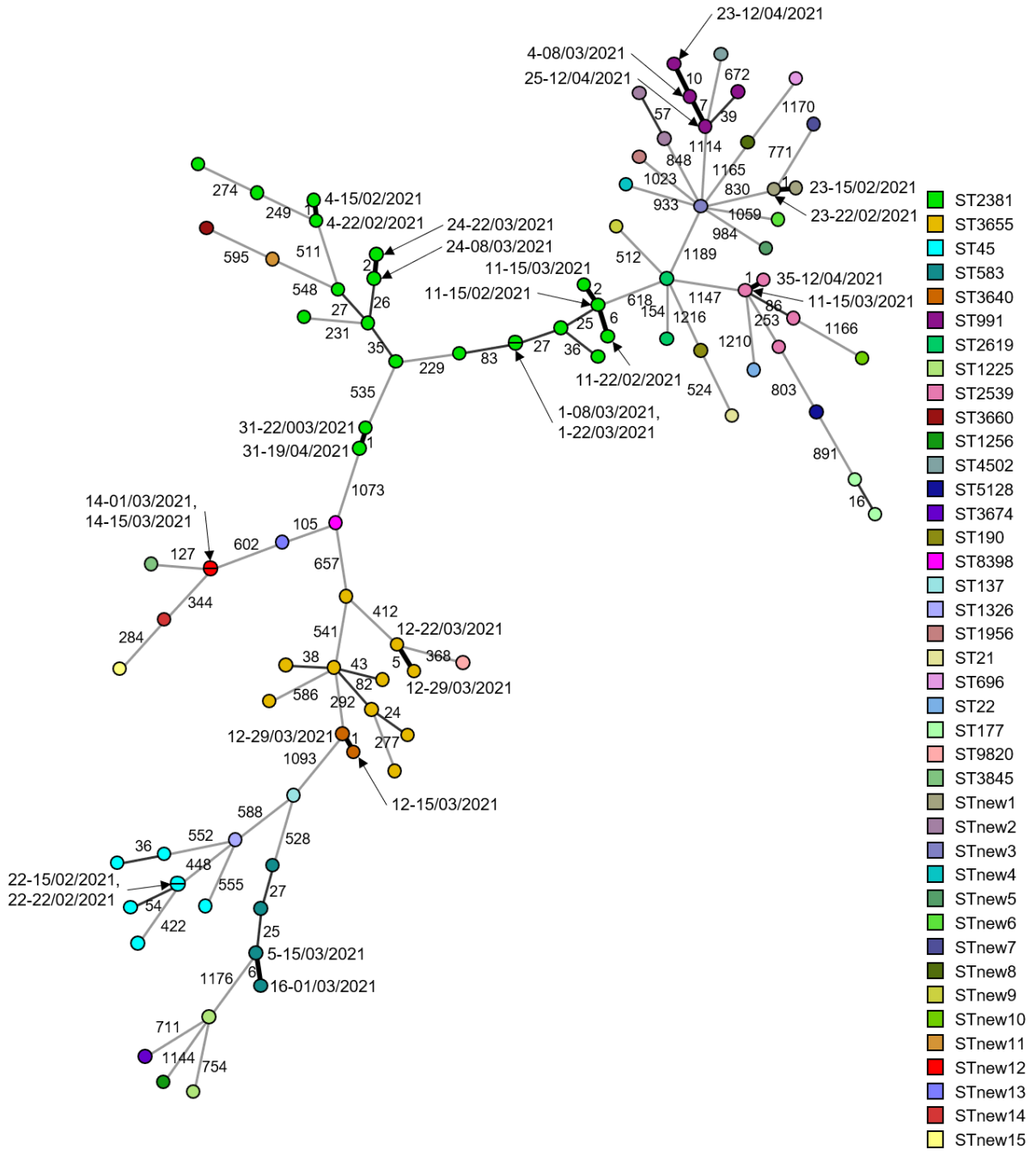

Figure 20: wgMLST comparison of *C. jejuni isolates*

Labelled based on site. Numbers of branches are the number of wgMLST differences between isolates. The colour relates to the MLST sequence.

WGS of *C. coli* is shown in Figure 21. Three of the five isolates from site 7 were indistinguishable, one was very similar and one is quite diverse (1387 MLST). The difference in wgMLST shows the other isolates are diverse.

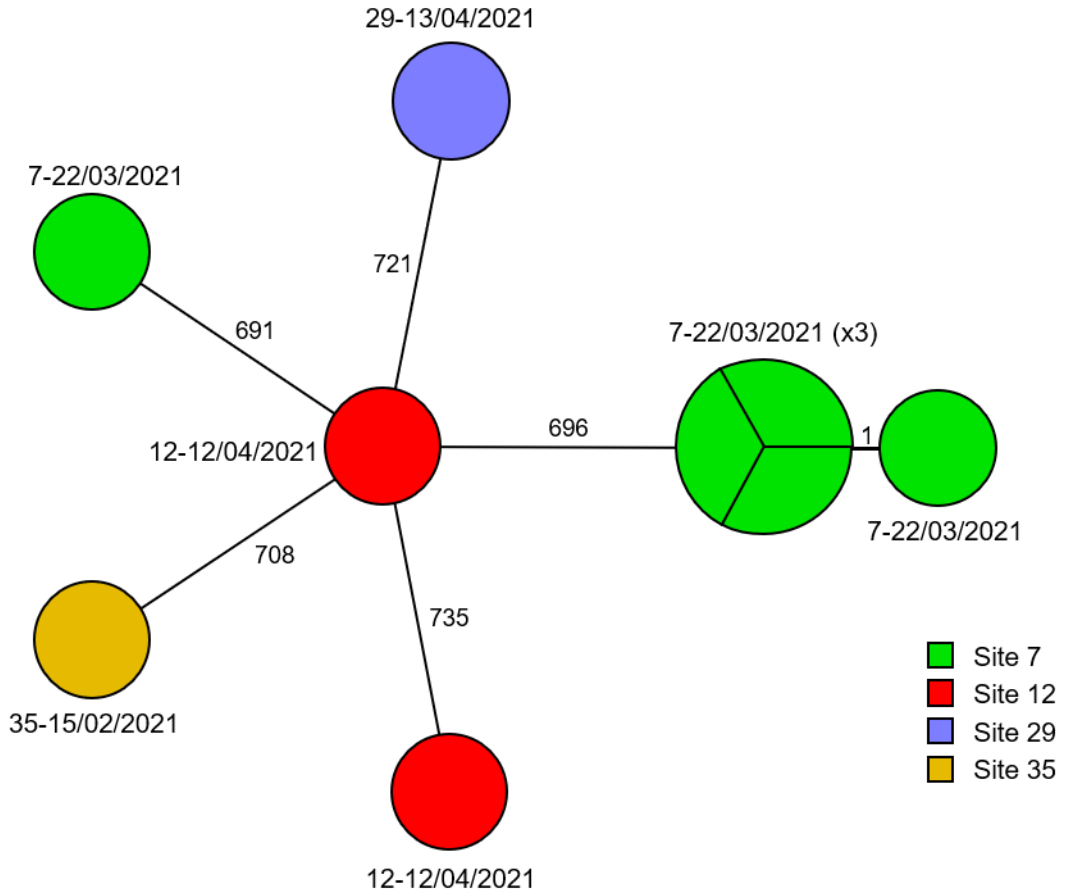

Figure 21: wgMLST comparison of *C. coli* isolates

Labelled based on site. Numbers of branches are the number of wgMLST differences between isolates. The colour relates to the MLST sequence.

APPENDIX E: *SALMONELLA* AND STEC

E.1 Whole genome sequencing data for *Salmonella*

Salmonella serotypes identified from positive enrichments are presented in Table 10.

Table 10: Whole genome sequencing data for *Salmonella*

Site	Serovar	MLST ST
1	Typhimurium	ST568
1	Typhimurium	ST568
2	Typhimurium	ST19
3	Typhimurium	ST19
4	Typhimurium	ST568
4	Typhimurium	ST2297
4	Typhimurium	ST568
4	Typhimurium	ST19
5	Typhimurium	ST19
5	Typhimurium	ST19
5	Typhimurium	ST568
5	Enteritidis	ST183
6	Enteritidis	ST183
6	Enteritidis	ST183
6	Enteritidis	ST183
6	Typhimurium	ST568
6	Typhimurium	ST568
7	Bovismorbificans	ST377
8	Typhimurium	ST568
8	Typhimurium	ST568
11	Typhimurium	ST19
11	Typhimurium	ST568
15	Typhimurium	ST568
23	Stanley	ST29
23	Typhimurium	ST568
23	Enteritidis	ST183
23	Schwarzengrund	ST96
26	Typhimurium	-
27	Stanley	ST29
29	Typhimurium	ST19
29	Typhimurium	ST19
36	Bovismorbificans	ST377

Most of the isolates were Typhimurium ST568 (12 isolates) and ST19 (8 isolates), and wgMLST indicates that there is not much variety amongst these isolates. The two Stanley isolates were only separated by 2 wgMLST and were from different locations on the same river (23 and 27). The Enteritdis isolates were from three different rivers (5, 6 and 23) and are also similar, being separated by 17 wgMLST (Figure 22).

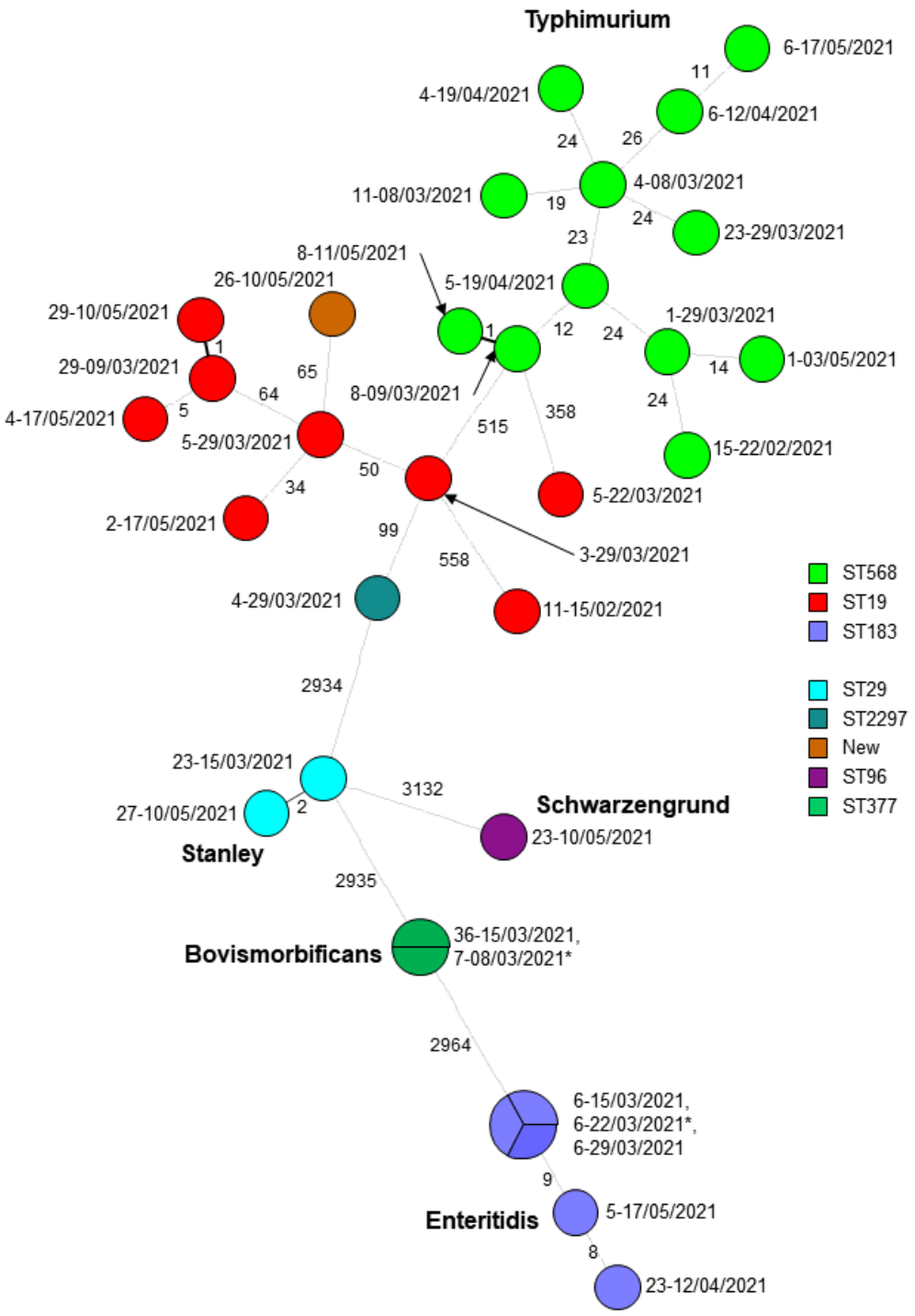

Figure 22: wgMLST comparison of *Salmonella* isolates

Labelled based on site. Numbers of branches are the number of wgMLST differences between isolates. The colour relates to the MLST sequence.

E.2 Whole genome sequencing data for STEC

STEC serotypes and pathogenicity is shown in Table 11. Although STEC identified from sites 18, 23 and 24 had no *eae* or *ehx*A genes they were confirmed as STEC and are pathogenic.

Table 11: Shiga toxin producing *Escherichia* isolated from MPN enrichments and characterised by whole genome sequencing

Site	MPN enrichment	Identification	Serotype	Virulence genes	Pathotype
1	*stx*1, *eae*, *ehx*A	*Escherichia coli*	O84:H2	15	STEC
1	*stx*1, *eae*, *ehx*A	*Escherichia coli*	O84:H2	17	STEC
1	*stx*1, *eae*, *ehx*A	*Escherichia coli*	O84:H2	17	STEC
18	*stx*2	*Escherichia coli*	O9:H30	3	STEC
23	*stx*1	*Escherichia coli*	ONT:H21	4	STEC
23	*stx*2, *eae*, *ehx*A	*Escherichia coli* + *Morganella morganii*	ONT:H7	11	STEC
24	*stx*2	*Escherichia coli*	ONT:H21	4	STEC

The wgMLST analysis indicates that the three isolates from site 1 on the 12/04/21 are the same. Even though two isolates were from site 23 they are diverse (Figure 23).

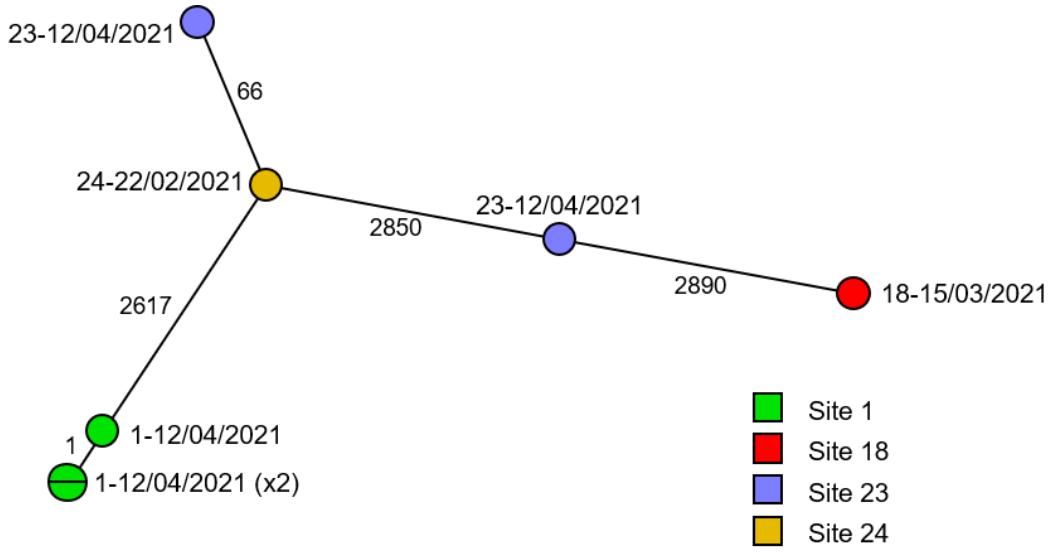

Figure 23: wgMLST comparison of STEC isolates

Labelled based on site. Numbers of branches are the number of wgMLST differences between isolates. The colour relates to the MLST sequence.

REFERENCES

Ahmed W, Payyappat S, et al. 2019. A duplex PCR assay for the simultaneous quantification of Bacteroides HF183 and crAssphage CPQ_056 marker genes in untreated sewage and stormwater. *Environment International* 126, 252-259.

APHA. 2017a. 9223 Enzyme substrate coliformtest. 9223 B Enzyme substrate test. *Standard methods for the examination of water and wastewater.* American Public Health Association, Washington, DC.

APHA. 2017b. 9230 Fecal enterococcus/streptococcus groups. 9230 D Fluorogenic substrate enterococcus test. *Standard methods for the examination of water and wastewater.* American Public Health Association, Washington, DC.

APHA. 2017c. 9260 Detection of pathogenic bacteria 9260 G *Campylobacter*. *Standard methods for the examination of water and wastewater.* American Public Health Association, Washington, DC.

Baque RH, Gilliam AO, et al. 2011. A real-time RT-PCR method to detect viable *Giardia lamblia* cysts in environmental waters. *Water Research* 45, 3175-3184.

Best EL, Powell EJ, et al. 2003. Applicability of a rapid duplex real-time PCR assay for speciation of *Campylobacter jejuni* and *Campylobacter coli* directly from culture plates. *FEMS Microbiology Letters* 229, 237-241.

Chern E, Siefring S, et al. 2011. Comparison of quantitative PCR assays for *Escherichia coli* targeting ribosomal RNA and single copy genes. *Letters in Applied Microbiology* 52**,** 298-306.

Conroy E, Donald M. 2019. *Tūhoromatanui: Ngā Pōtiki Environmental Plan 2019-2029.* Ngā Pōtiki ā Tamapahore Trust.

Derzelle S, Grine A, et al. 2011. A quantitative PCR assay for the detection and quantification of Shiga toxin-producing *Escherichia coli* (STEC) in minced beef and dairy products. *International Journal of Food Microbiology* 151, 44-51.

Devane ML, Robson B, et al. 2007. A PCR marker for detection in surface waters of faecal pollution derived from ducks. *Water Research* 41, 3553-3560.

Gerba CP. 2009. *Environmentally transmitted pathogens*. In Environmental Microbiology ed. IL Pepper, CP Gerba, TJ Gentry, pp. 445-484.

Gilpin B, Hewitt J, et al. 2018. *Freshwater microbiological sciences review (FMSR): Stage one. Work stream B: Microbial methods*. Christchurch: Institute of Environmental Science and Research.

Green HC, Dick LK, et al. 2012. Genetic markers for rapid PCR-based identification of gull, Canada goose, duck, and chicken fecal contamination in water. *Applied and Environmental Microbiology* 78**,** 503-510.

Haugland RA, Siefring S, et al. 2012. Influences of sample interference and interference controls on quantification of enterococci fecal indicator bacteria in surface water samples by the qPCR method. *Water Research* 46, 5989-6001.

Hoorfar J, Ahrens P, et al. 2000. Automated 5′ nuclease PCR assay for identification of *Salmonella enterica*. *Journal of Clinical Microbiology* 38, 3429-3435.

Ibekwe AM, Watt PM, et al. 2004. Fate of *Escherichia coli* O157:H7 in irrigation water on soils and plants as validated by culture method and real-time PCR. *Canadian Journal of Microbiology* 50, 1007-1014.

ISO. 2017. *ISO 10272-1:2017 Microbiology of the food chain — horizontal method for detection and enumeration of* Campylobacter *spp. — part 1: Detection method*. International Organization for Standardization.

ISO. 2012. *ISO/TS 13136:2012, Microbiology of Food and Animal Feed—Real-Time Polymerase Chain Reaction (PCR)-Based Method for the Detection of Food-Borne Pathogens—Horizontal Method for the Detection of Shiga Toxin-Producing* Escherichia coli *(STEC) and the Determination of O157, O111, O26, O103, and O145 Serogroups.* [Online]. Available: *http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=53328*

Korajkic A, McMinn BR, et al. 2018. Relationships between microbial indicators and pathogens in recreational water settings. *International Journal of Environmental Research and Public Health* 15, 2842.

Lake R, Horn B, et al. 2018. *Freshwater microbiological sciences review (FMSR): Stage one work stream C: Dose response recommendation for dose response*. Christchurch: Institute of Environmental Science and Research.

Leonard M, Gilpin B, et al. 2020. *Quantitative microbial risk assessment pilot.* Christchurch: Institute of Environmental Science and Research. A report for the Ministry for the Environment.

Lindsey RL, Pouseele H, et al. 2016. Implementation of whole genome sequencing (WGS) for identification and characterization of shiga toxin-producing *Escherichia coli* (STEC) in the United States. *Frontiers in Microbiology*, 7, 766.

Malorny B, Paccassoni E, et al. 2004. Diagnostic real-time PCR for detection of *Salmonella* in food. *Applied and Environmental Microbiology* 70, 7046-7052.

Mary C, Chapey E, et al. 2013. Multicentric evaluation of a new real-time PCR assay for quantification of *Cryptosporidium* spp. and identification of *Cryptosporidium parvum* and *Cryptosporidium hominis*. *Journal of Clinical Microbiology* 51, 2556-2563.

McBride G, Till D, et al. 2002. *Freshwater microbiology research programme report: pathogen occurrence and human health risk assessment analysis*. Wellington: Ministry for the Environment and Ministry of Health.

McGregor S, Begley C. 2014. *Cultural values report*. Prepared on behalf of Te Rūnanga o Kaikōura.

Medeiros D, Hofmann L. 2002. Isolation of *Campylobacter* from food, MFLP-46. In: *Compendium of analytical methods, vol. 3*. Ontario: Health Canada.

Meijer K. (2012) Otepuni Faecal Source Investigation, Environment Southland. Publication No. 2012-15.

Milne J, Moriarty E, et al. 2018. *Freshwater microbiological sciences review (FMSR): Stage one. Work stream A: site selection. Recommended sampling sites*. Christchurch: Institute of Environmental Science and Research.

MfE. 2003. *Microbiological water quality guidelines for marine and freshwater recreational areas*. Wellington: Ministry for the Environment.

Moriarty E, Milne J, et al. 2018. *Freshwater microbiological sciences review (FMSR): Stage one work stream D: Survey design and QMRA project plan*. Christchurch: Institute of Environmental Science and Research.

New Zealand Government. 2020. National Policy Statement for Freshwater Management 2020. Wellington: Ministry for the Environment.

Paton AW, Paton JC. 1998. Detection and characterization of Shiga toxigenic *Escherichia coli* by using multiplex PCR assays for *stx*1, *stx*2, *eaeA*, enterohemorrhagic *E. coli hly*A, rfb_{O111} and rfb_{O157}. *Journal of Clinical Microbiology* 36, 598-602.

Rainforth H, Harmsworth G. 2019. *Kaupapa Māori freshwater assessments. A summary of iwi and hapū-based tools, frameworks and methods for assessing freshwater environments*, Perception Planning Ltd.

Reischer GH, Kasper DC, et al. 2006. Quantitative PCR method for sensitive detection of ruminant fecal pollution in freshwater and evaluation of this method in alpine karstic regions. *Applied and Environmental Microbiology* 72, 5610-5614.

Seemann T, Goncalves Da Silva A, et al. *Nullarbor github* [Online]. Available: https://github.com/tseemann/nullarbor [Accessed 2020].

Shanks OC, Kelty CA, et al. 2009. Quantitative PCR for genetic markers of human fecal pollution. *Applied and Environmental Microbiology* 75, 5507-5513.

Shanks OC, White K, et al. 2010. Performance assessment PCR-based assays targeting *Bacteroidales* genetic markers of bovine fecal pollution. *Applied and Environmental Microbiology* 76, 1359-1366.

USEPA. 2005. *Method 1623: Cryptosporidium and Giardia in Water by Filtration/IMS/FA*. Washington DC: United States Environmental Protection Agency.

USEPA. 2013. *Method 1609: Enterococci in water by taqman® quantitative polymerase chain reaction (qPCR) with internal amplification control (IAC) assay*. Washington DC: United States Environmental Protection Agency.

Wright, J. Senior Scientist Enteric Reference and Leptospira Reference Laboratories. Pers. Com. Email 24/6/21.

Wong T, Devane ML, et al. 2004. Validation of a PCR method for *Campylobacter* detection on poultry packs. *British Food Journal* 106, 642-650.

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