Investigation of Faecal Pollution Sources and Bacterial Transfer Hydrodynamics in Rural Catchments

A Thesis Submitted to the National University of Ireland for the Degree of Doctor of Philosophy
By

Sinéad Murphy

Microbial Ecology Laboratory, Discipline of Microbiology, School of Natural Sciences, College of Science, National University of Ireland, Galway.

Research Supervisor: Professor Vincent O’Flaherty
# Table of Contents

## Chapter I

### Introduction

1.1 General Introduction ............................................................... 2

1.2 Faecal Contamination of Water ............................................... 2
   1.2.1 Point and Diffuse Sources ............................................. 3
   1.2.2 Faecal Indicator Organisms .......................................... 6
   1.2.3 *E. coli* as a Faecal Indicator Organism ............................ 7
      1.2.3.1 Methods Used to Detect *E. coli* ................................. 7
      1.2.3.2 Problems with *E. coli* as a FIO ............................... 9

1.3 Alternative Methods Used to Detect Faecal Contamination in Water 10
   1.3.1 Microbial Methods used to Detect Faecal Contamination in Water 10
      1.3.1.1 *Bifidobacterium* spp. .......................................... 10
      1.3.1.2 *Clostridium perfringens* ...................................... 10
      1.3.1.3 *Enterococcus* spp. ............................................ 11
      1.3.1.4 The Ratio of *E. coli* to *Streptococcus* ..................... 11
      1.3.1.5 *Bacteroides fragilis* bacteriophage ................................ 12
   1.3.2 Viral Methods to Detect Faecal Contamination in Water ............. 12
      1.3.2.1 F+ RNA coliphage ................................................ 12
      1.3.2.2 Phage as Non-Host Specific Viral Indicators ................ 13
      1.3.2.3 The Use of Host-Specific Viruses to Detect Faecal Pollution in Water ................................................................. 13
   1.3.3 Chemical Methods to Detect Faecal Contamination in Water .......... 14
      1.3.3.1 Faecal Sterols and Stanols .................................... 14
      1.3.3.2 Optical (laundry) Brighteners .................................. 14
      1.3.3.3 Caffeine ............................................................ 15
1.4 Bacteroidales – An Alternative Faecal Indicator Organism for Use in Microbial Source Tracking

1.4.1 General Bacteroidales Assays

1.4.2 Host-Specific Bacteroidales Assays

1.4.2.1 Human-Specific Bacteroidales Assays

1.4.2.2 Bovine-Specific Bacteroidales Assays

1.4.3 Measuring Survival Rates of Bacteroidales

1.4.4 Survival Abilities of Bacteroidales

1.4.5 Successful Application of Bacteroidales Molecular Assays

1.5 A Tiered Approach to Improving Water Quality Monitoring and Protection

1.6 Nutrients and Their Affect on Water

1.6.1 Monitoring of Nutrients in Water by Grab Sampling

1.6.2 Monitoring Nutrients in Water on a Continuous Basis

1.7 The European Water Framework Directive

1.8 Catchment Science

1.8.1 The Integrated Catchment Science Approach in the EU

1.8.2 The Integrated Catchment Science Approach in Ireland

1.9 The Future of Catchment Science

1.9.1 Researched Needed for Protection of Water Quality

1.9.2 Statement of Project Aims

1.10 References
# Table of Contents

## Chapter II

**Quantifying Faecal Indicator Organism Transfer Pathways in Agricultural Catchments using Loadograph Recession Analysis**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author Contribution</td>
<td>67</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>68</td>
</tr>
<tr>
<td>2.2 Methods</td>
<td>72</td>
</tr>
<tr>
<td>2.2.1 Catchment Sites</td>
<td>72</td>
</tr>
<tr>
<td>2.2.2 Field Monitoring and Water Analysis</td>
<td>74</td>
</tr>
<tr>
<td>2.2.3 Statistical Analysis</td>
<td>76</td>
</tr>
<tr>
<td>2.2.4 Calculation for <em>E. coli</em> load</td>
<td>76</td>
</tr>
<tr>
<td>2.2.5 Loadograph Recession Analysis</td>
<td>76</td>
</tr>
<tr>
<td>2.3 Results</td>
<td>78</td>
</tr>
<tr>
<td>2.4 Discussion</td>
<td>97</td>
</tr>
<tr>
<td>2.5 Conclusions</td>
<td>100</td>
</tr>
<tr>
<td>2.6 References</td>
<td>102</td>
</tr>
</tbody>
</table>

## Chapter III

**Tracking the Sources and Transport Pathways of Faecal Contamination in Catchments during Dynamic Hydrological Phases**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author Contribution</td>
<td>111</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>112</td>
</tr>
<tr>
<td>3.2 Materials and Methods</td>
<td>116</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>3.2.1 Catchment Sites</td>
<td>116</td>
</tr>
<tr>
<td>3.2.2 Water Sample Collection, and Total Coliform and <em>E. coli</em> Analysis</td>
<td>116</td>
</tr>
<tr>
<td>3.2.3 Water Sample Processing and DNA Extraction</td>
<td>116</td>
</tr>
<tr>
<td>3.2.4 Faecal Sample Collection and DNA Extraction</td>
<td>117</td>
</tr>
<tr>
<td>3.2.5 TaqMan® <em>Bacteroidales</em> qPCR Assays</td>
<td>117</td>
</tr>
<tr>
<td>3.2.6 Statistics</td>
<td>118</td>
</tr>
<tr>
<td>3.2.7 Loadograph Recession Analysis</td>
<td>118</td>
</tr>
<tr>
<td>3.2.8 Microbial Source Tracking Loads</td>
<td>118</td>
</tr>
<tr>
<td>3.3 Results</td>
<td>120</td>
</tr>
<tr>
<td>3.3.1 qPCR Assay Specificity</td>
<td>120</td>
</tr>
<tr>
<td>3.3.2 Loadograph Recession Analysis</td>
<td>123</td>
</tr>
<tr>
<td>3.3.3 Transfer Pathways</td>
<td>125</td>
</tr>
<tr>
<td>3.4 Discussion</td>
<td>133</td>
</tr>
<tr>
<td>3.5 Conclusions</td>
<td>137</td>
</tr>
<tr>
<td>3.6 References</td>
<td>139</td>
</tr>
</tbody>
</table>

**Chapter IV**

**Using Subtractive Hybridisation to Design an Ovine - Specific *Bacteroidales* PCR and qPCR Assay**

Author Contribution 147

4.1 Introduction 148

4.2 Materials and Methods 151

4.2.1 Target and Subtracter DNA 151

4.2.2 Sample Collection and DNA Extraction 151

4.2.3 PCR Amplification 153

4.2.4 Restriction Digestion 153

4.2.5 Preparation of Linkers and Ligation 153
4.2.6 Amplification of Target and Subtractor DNA 154
4.2.7 Immobilization of Subtractor in Microplate Well 154
4.2.8 Solution Hybridization 154
4.2.9 Amplification of Subtraction Products, Cloning, Sequencing, and Primer Design 155
4.2.10 PCR Amplification of Plasmids 155
4.2.11 Application of Sheep-Specific PCR Assays to Sheep and Cow Faecal Extracts 155
4.2.12 Amplification of Standard Curves and Determining Assay Specificity 155

4.3 Results 157
4.3.1 Restriction Digestion and Subtractive Hybridisation 157
4.3.2 Sequence Analysis and Alignment 159
4.3.3 Application of BacOvi PCR Assays to Plasmids and Faecal Extracts 161
4.3.4 Establishment of Standard Curves for the BacOvi qPCR Assays 163
4.3.5 Application of qPCR Assay to Various Faecal Extracts 168

4.4 Discussion 172

4.5 Conclusions 175

4.6 References 176

Chapter V

Concluding Remarks

5.1 Overview 181
5.1.1 Targeted High- and Low- Flow Events in Spring/Summer and Autumn/Winter 181
5.1.2 Flow Pathway Separation 182
5.1.3 Cross-amplification of Host-Specific Bacteroidales Assays 182
5.1.4 Determination of Faecal Source Using Multi-tiered Approach 182
5.1.5 Establishment of Pathways Which Transfer Faecal Contamination 183  
5.1.6 Design and Application of Sheep-Specific *Bacteroidales* Assays 184  

5.1.7 Conclusions 184  

5.2 Future Perspectives and Recommendations 186  
5.2.1 Risk Assessment Models 186  
5.2.2 Broadening Numbers of Indicators Analysed 186  
5.2.3 Survival Rates of *Bacteroidales* in Irish Watersheds 186  
5.2.4 Application of Sheep-Specific Assays to Water Samples 187  
5.2.5 Applying Targeted Approach to Catchment on Trial Basis 187  
5.2.6 Standardisation of these Methods in Catchments 187  

5.3 References 188
List of Figures

Chapter I

Figure 1.1 Diagram showing point and diffuse (non-point) sources of faecal pollution (Author’s own).

Figure 1.2 The six catchments that are part of the Teagasc ACP.

Figure 1.3 Karst catchment in west of Ireland (a & b) (Grassland D) and river catchment in the east of Ireland (c & d) (Arable B). a: Turlough quite full after prolonged rain during autumn/winter. b: Autosampler (1) taking and auto bank-analyser evaluating (2) water at spring emergence during autumn/winter. c: Water level high in outlet after heavy rainfall event in autumn/winter. d: Field adjacent to outlet showing proximity of cattle to water during spring/summer.

Chapter II

Figure 2.1 Map of Arable B catchment (a) showing outlet (blue inverted triangle), Met station (red box), rain gauge (red triangle); map of Grassland D catchment (b) showing emergent spring (blue inverted triangle), Met station (red box), rain gauge (red triangle). Map of Ireland showing where each of the two catchments are located (c). (Printed under license number 6155 from the Ordnance Survey Ireland)

Figure 2.2 Time series charts of the parameters measured at high resolution (collated to hourly data) in (a) Grassland D and (b) Arable B. TP – black line; TRP – grey line and dates of sample events (hollow black circles)

Figure 2.3 E. coli catchment load and discharge per event. Circles are Arable B and squares are Grassland D. Closed symbols are events dominated by rising phases, open symbols are events dominated by falling phases. The power trend-line describes the overall E. coli load-discharge relationship for Arable B

Figure 2.4 Grassland D loadographs showing pathway separations for events I. A: Time series of hourly discharge (blue line) and turbidity (purple line) showing pathway separation; B: Time series of hourly TP loads (black line), and E. coli loads (green line with closed symbols).
Figure 2.5 Grassland D loadographs showing pathway separations for events II. A: Time series of hourly discharge (blue line) and turbidity (purple line) showing pathway separation; B: Time series of hourly TP loads (black line), and E. coli loads (green line with closed symbols).

Figure 2.6 Grassland D loadographs showing pathway separations for events III. A: Time series of hourly discharge (blue line) and turbidity (purple line) showing pathway separation; B: Time series of hourly TP loads (black line), and E. coli loads (green line with closed symbols).

Figure 2.7 Arable B loadographs showing pathway separations for events I. A: Time series of hourly discharge (blue line) and turbidity (purple line) showing pathway separation; B: Time series of hourly TP loads (black line), and E. coli loads (green line with closed symbols).

Figure 2.8 Arable B loadographs showing pathway separations for events II. A: Time series of hourly discharge (blue line) and turbidity (purple line) showing pathway separation; B: Time series of hourly TP loads (black line), and E. coli loads (green line with closed symbols).

Figure 2.9 Arable B loadographs showing pathway separations for events III. A: Time series of hourly discharge (blue line) and turbidity (purple line) showing pathway separation; B: Time series of hourly TP loads (black line), and E. coli loads (green line with closed symbols).

Figure 2.10 Arable B loadographs showing pathway separations for events IV. A: Time series of hourly discharge (blue line) and turbidity (purple line) showing pathway separation; B: Time series of hourly TP loads (black line), and E. coli loads (green line with closed symbols).

Figure 2.11 Arable B loadographs showing pathway separations for events V. A: Time series of hourly discharge (blue line) and turbidity (purple line) showing pathway separation; B: Time series of hourly TP loads (black line), and E. coli loads (green line with closed symbols).

Figure 2.12 Arable B loadographs showing pathway separations for events VI. A: Time series of hourly discharge (blue line) and turbidity (purple line) showing pathway separation; B: Time series of hourly TP loads (black line), and E. coli loads (green line with closed symbols).
Chapter III

Figure 3.1 Amplification curves for BacBov assay. Sheep (S), cow (C), human (Hu), horse (H), goat (G) and pig (P) were assayed. The Ct cutoff (CC) was 30 Ct (red line). Sheep and cow both amplified at a Ct of 14. Goat (which was a 1 in 10 dilution) amplified at a Ct of 20. Human and pig came after the Ct cutoff at Ct of 32 and 34, respectively. Horse was close to the Ct cutoff at a Ct of 29.

Figure 3.2 Arable B qPCR loadograph for event I. Time series of hourly TP loads (black line), \textit{E. coli} loads (green line with dots), BacUni loads (orange line with hollow black circles), BacHum loads (black line with blue triangle), and BacBov loads (black line with red squares). Data have been separated into transfer pathways using Loadograph Recession Analysis.

Figure 3.3 Arable B qPCR loadograph for event II. Time series of hourly TP loads (black line), \textit{E. coli} loads (green line with dots), BacUni loads (orange line with hollow black circles), BacHum loads (black line with blue triangle), and BacBov loads (black line with red squares). Data have been separated into transfer pathways using Loadograph Recession Analysis.

Chapter IV

Figure 4.1 Flow chart of process involved in Subtractive Hybridisation (Adapted from Dick \textit{et al.}, 2005).

Figure 4.2 Gel electrophoresis showing target (sheep) and subtracter (cow) products after amplification by \textit{Bacteroidales} primers 32F and 1492R and enzyme restriction by \textit{AcII}. The product sizes were 300 and 500 bp.

Figure 4.3 BacOvi S1-S5 amplicons of plasmids containing target sequence. HLV: Hyperladder™ V 500bp marker, B: Blank; PN: BacOvi Plasmid Neat, PD: BacOvi Plasmid 1:100; NC: Negative Control

Figure 4.4 BacOvi S1-S5 amplicons of plasmids and faecal extracts. Lane 1 HLV: Hyperladder™ V 500bp marker, Lane 2: Blank, Lane 3: BacOvi S1 Plasmid (P) Neat, Lane 4: BacOvi S1 P 1:100, Lane 5: Sheep, Lane 6: Cow, Lane 7: Negative control (-C), Lane 8: BacOvi S2 P Neat, Lane 9: BacOvi S2 P 1:100, Lane 10: Sheep, Lane 11: Cow, Lane 12: -C Lane 13: BacOvi S3 P Neat, Lane 14: BacOvi S3 P 1:100, Lane 15: Sheep, Lane 16: Cow, Lane 17: -C, Lane 18: BacOvi S4 P Neat, Lane 19: BacOvi S4 P 1:100, Lane 20: Sheep, Lane 21: Cow, Lane 22: -C, Lane 23: BacOvi S5 P Neat, Lane 24: BacOvi S5 P 1:100, Lane 25: Sheep, Lane 26: Cow, Lane 27: -C, Lanes 28, 29: Blank, Lane 30: HLV.
Figure 4.5  Standard curves of measured Ct values versus log plasmid marker concentration for the five newly developed BacOvi Assays S1-S5. a: BacOvi S1, b: BacOvi S2, c: BacOvi S3, d: BacOvi S4, e: BacOvi S5.

Figure 4.6  Amplification curves for BacOvi S3 assay tested against various faecal samples. Sheep (S), cow (C), human (Hu), horse (H), goat (G) and pig (P) were assayed in the following dilutions; Neat (N), 1:10 ($10^{-1}$), and 1:100 ($10^{-2}$). The Ct cutoff was 31 (red line) and the threshold was 0.024 (blue line).

Figure 4.7  Standard curve of measured Ct values versus log plasmid marker concentration for BacOvi S6. The annealing temperature was increased from 60°C to 65°C.

Figure 4.8  Amplification curves for BacOvi S6 assay tested against various faecal samples. Sheep (S), cow (C), human (Hu), horse (H), goat (G) and pig (P) were assayed in the following dilutions; Neat (N), 1:10 ($10^{-1}$), and 1:100 ($10^{-2}$). The Ct cutoff was 29 (red line) and the threshold was 0.024 (blue line).
List of Tables

Chapter I

Table 1.1 List of established molecular assays, PCR or qPCR, designed to target various host-specific Bacteroides/Bacteroidales.

Table 1.2 Dates for cattle manure storage periods, and opened and closed periods for spreading (chemical and organic fertilisers, and farmyard manure) for the all the counties (divided into zones) in Ireland.

Chapter II

Table 2.1 Data summaries of sampled event discharge and bacteria data from the karst grassland (Grassland D) and mixed arable (Arable B) catchments. Dates in bold represent those events used for further LRA analysis.

Table 2.2 Pearson’s Correlations of all measured parameters for a. Grassland D and b. Arable B. P-value of P<0.05 = 1 symbol, P<0.01 = 2 symbols; N.C. = no correlation.

Chapter III

Table 3.1 Primers and probes used in this study and the expected size of their amplified product.

Table 3.2 Assay specificity of BacUni-UCD, BacHum-UCD, and BacBov-UCD tested against various pooled faecal samples.

Table 3.3 Transfer pathways for Arable B and Grassland D. Event name, length of sample event, season, and event type.

Table 3.4 Pearson’s correlations matrix, for loads in Arable B during entirety of event I, of E. coli, total phosphorus (TP), BacUni, BacHum and BacBov.

Table 3.5 Pearson’s correlations matrix for loads in Arable B during quick flow of the event I, of E. coli, total phosphorus (TP), BacUni, BacHum and BacBov.

Table 3.6 Pearson’s correlations matrix for loads in Arable B during entirety of event II, of E. coli, total phosphorus (TP), BacUni, BacHum and BacBov.
Table 3.7  Pearson’s correlations matrix, for loads in Arable B during quick flow of event II, of *E. coli*, total phosphorus (TP), BacUni, BacHum and BacBov.

Chapter IV

Table 4.1  Primers, linkers and probes used in this study. Restriction site extensions and overhangs are underlined.

Table 4.2  Expected product size (bp) of each BacOvi PCR assay.

Table 4.3  Novel sheep–specific *Bacteroidales* assay standard curves, amplification efficiency and Ct cutoff.
I, the Candidate, certify that the Thesis is all my own work and that I have not obtained a degree in this University or elsewhere on the basis of any of this work.

Signed: ____________________________ Date: ____________________________
Acknowledgements

There are so many people I need to thank for helping me on my journey through PhD life.

Firstly I would like to thank my supervisor, Prof. Vincent O’Flaherty for giving me this great opportunity and his guidance and support throughout the years of the PhD. There were many occasions to present my research both at home and abroad and he always encouraged me to embrace them.

I would like to thank my co-supervisor Prof. Phil Jordan for all his help, especially with the hydrological aspect of the project, and his guidance through stormy waters.

A thank you to all the people in the Microbial Ecology Lab, especially Niamh Bhreathnach, Dermot Hughes, Eoin Gunnigle and Paul Mc Cay for answering all my questions and helping out with the sampling trips especially the one where the car battery died.... A very special thank you goes to my MEL buddy Ciara Keating! All the tea and all the chats and all the support, we did it!

Thanks to everyone in the Microbiology Dept., for all your support and kind words especially Aileen Gill not just for her baking but for being a good friend too, Dr. Cindy Smith for all her molecular advice and to Maurice Martyn for his efficient ordering.

To everyone I worked with in Teagasc ACP thank you. I would especially like to thank Lucy Crockford, Per-Erik Mellander, Alice Melland, John Kennedy and Maria Merriman for all their guidance and advice over the years. Thank you to all the farmers involved in ACP, without their contribution this project would not have been possible.

To all the people I worked with in UC Davis, Prof. Stefan Wuertz, Dr. Minji Kim, Dr. Alexander Schriewer, and Kaitlyn Hanley, thank you for welcoming me into your lab and giving me all the help I needed.

To my fantastic friends who’ve supported me along the way with text messages, food (very important!), mini-pep talks, coffee, etc., Rebecca (Row the boat!), Lou (supplier of cheese and hugs), Lorraine (chin up, be grand), Catherine (chats, the salt house, napoli), Sarah (my California buddy), Sharon (Queen of pilates) Brian (Don’t Stop Believin’), Max (my office buddy) thank you so very much! A huge thanks to my bestest bud Jess, you really kept me going with Chinese food and awful films!

A massive thank you to my family, my wonderful crazy family who have always supported me. Especially my Nana and Granny who kept me well fed, my Grandad and his infrequent but wise words, my Godmother/second Mam Mary who always made sure I was doing ok, my Aunty Anne who would do whatever she could to help
me with anything, and my Uncle Martin for his motivational speeches. To my brothers, Seán and Ciarán, and my little sister Caoimhe, thank you for all the outdoor adventures, the baking, and the constant reminders that “pressure is for tyres”!

To my wonderful boyfriend Fred, thank you so much for all the support over the last while. It’s been a tough road but you’ve kept me going, kept me sane, kept me smiling and keep surprising me. You always have my back, even when you have to throw cereal bars at me and run away. Thank you.

Last but not least, to the two people who gave me their unyielding motivation and determination, who were always there no matter what whether at the end of a phone or in person, who always believed in me, who answered questions on the nitrates directive or sent me out the door with frozen salmon. There are not enough words to express my sincerest gratitude for everything they’ve done for me. To the two people who brought me into this world, my parents, John and Caroline Murphy, I dedicate this PhD thesis to both of you.
Abstract

Faecal bacteria from point and diffuse source pollution can impact water which poses a serious public health and environmental threat. Faecal pollution contains nutrients such as phosphorus and nitrogen and when in excess these nutrients can cause eutrophication. Observing and establishing the source of faecal pollution is imperative for the protection of water quality and human health. Conventional culture methods to detect such pollution via faecal indicator bacteria have been extensively utilised but do not determine the source of pollution. To combat this, microbial source tracking (MST), an important emerging molecular tool, can be applied to detect host-specific markers in faecally contaminated waters. The main aim of this project was to establish the source of faecal contamination and the pathways by which it was transferred in two rural catchments, Arable B (river catchment) and Grassland D (karst catchment containing turloughs), in Ireland using a multi-tiered approach including elements such as MST. The MST approach was to target ruminant and human-specific faecal Bacteroidales 16S rRNA genes within the two catchments. From a risk assessment perspective, the two catchments differ greatly due to the varying hydrology, agricultural land use, topography, geology, and human density located in each. These two catchments have combinations of potentially high source and transport pressures.

Hydrological connectivity drives the transfer of faecal bacteria and nutrients, of human or agricultural origin, from land to water. In Chapter II, a novel pathway separation technique, Loadograph Recession Analysis (LRA), was applied to identify transfer pathways of faecal bacteria (E. coli). LRA separates the pathways into quick flow - which represents surface overland flow, preferential flow, and tile and ditch drainage; interflow; and delayed flow - which represents baseflow. The study illustrated that high loads of phosphorus and E. coli were transferred during the faster flow pathways in both catchments. Grassland D was found to be a transport limited system and Arable B a less transport limited system. Transfers of E. coli were shown to be dependent on flow and independent of season. Arable B had higher E. coli loads, even in low flow, despite Grassland D having a potentially higher faecal indicator organism FIO source. In Grassland D, higher loads of E. coli were moved during rising phases but this was more subtle in Arable B as the loads
Abstract

mostly remained unchanged during rising and falling phases. Potential chronic point sources in Arable B, were indicated by the asymptotic decline of sustained \textit{E. coli} loads towards low flows.

MST has been used to differentiate between anthropogenic and agricultural faecal sources in faecally contaminated water but has not been used to show how the dominance of faecal sources changes between different phases of hydrological runoff. In Chapter III, it was hypothesised that agricultural ruminant faecal waste would be elevated increased and human faecal waste, where this was present, would be diluted by the quicker flow paths. The universal \textit{Bacteroidales} qPCR assay, BacUni-UCD, and the host- specific qPCR assays BacBov-UCD and BacHum-UCD were applied to water samples taken during different flow phases. The water samples were also assessed for \textit{E. coli} occurrence. Statistical analysis illustrated correlations between \textit{E. coli}, total phosphorus (TP), and BacUni-UCD loads indicating the presence of faecal contamination. BacBov-UCD and BacHum-UCD were detected in Arable B with the agricultural ruminant contamination mostly dominating the quick flow pathway, and human faecal contamination mostly moving through delayed flow. Grassland D also showed low levels of host-specific faecal contamination with ruminant and human sources being highest in quick flows.

There are \textit{Bacteroidales} bovine-specific assays currently available to detect cow faecal pollution but some of these assays have shown cross-amplification with sheep faecal pollution. As of yet, there has been no molecular \textit{Bacteroidales} assay developed to identify and differentiate sheep-specific microbial pollution. The aim of Chapter IV was to use subtractive hybridisation to identify specific DNA sequences to develop such an assay, thus allowing further differentiation of sources of pollution, aiding water quality. Target sheep-specific faecal \textit{Bacteroidales} rRNA gene fragments were differentiated from diverging, though closely related subtracter cow and human faecal \textit{Bacteroidales} sources. The sequences specific to sheep were used to design five PCR assays which worked successfully as demonstrated by their ability to amplify the positive control. Probes were designed to develop qPCR assays to discriminate between sheep and cow faeces. The resultant five qPCR assays were tested against various faecal samples. One qPCR assay was able to successfully differentiate between sheep, cow, human, horse, goat and pig faecal samples.
Abstract

Using combining DNA based analysis of faecal bacteria to discriminate faecal matter sources in conjunction with high-resolution P analysis for hydrological pathway discrimination and *E. coli* analysis, during different events, will add to the understanding and mitigation of FIO transfers from land to water. This will result in a more targeted approach to best management practices which could limit the deterioration of water quality in the most cost effective way. This information can be used by agricultural policy makers or local county councils to help protect water quality.
Chapter I

Introduction
1.1 General Introduction

Water is essential for all life (Gleick, 1998; Vitousek et al., 1997). It covers approximately 71% of the earth’s surface (Sonune and Ghate, 2004); of this, 96.5% is salt water and 2.5% is freshwater, and the majority of this (up to 77%) is locked up in polar ice-caps and glaciers (Greenlee et al., 2009; Postel et al., 1996). This means that less than 1% of freshwater is available to sustain life and for human use (Vitousek et al., 1997; Vorosmarty et al., 1997). Water is vital for all living organisms, and with an increasing global human population, and changes in weather patterns as a result of climate change, there are widening concerns and issues with water quality and quantity for human use. Among these concerns are increasing incidences of microbial and chemical pollution (Schwarzenbach, et al., 2010; Vorosmarty et al., 2000).

1.2 Faecal Contamination of Water

One particular global problem is faecal contamination of water (Soller et al., 2010). All faeces contain varying levels of pathogenic bacteria, such as *Escherichia coli* (*E. coli*), *Enterococcus*, *Lactobacillus*, *Bifidobacterium*, *Clostridium*, *Eubacterium*, *Streptococcus*, and sometimes *Salmonella* or *Campylobacter* (Dowd et al., 2008; Edberg et al., 2000; Franks et al., 1998; Harmsen et al., 1999). A major concern of water contaminated by faecal matter is the adverse effect it can have on human health via drinking and bathing water, and the impact this faecally contaminated water has on the quality of local environments (Allevi et al., 2013; Frey et al., 2013; Marti et al., 2013; Walters et al., 2013). These impacts can result in economic consequences (Walters and Field, 2009).

Approximately 780 million people in the world do not have access to safe drinking water (WHO-UNICEF, 2012). Unclean water storage facilities, poor public health, and inadequate sanitation contribute to unsafe water (Mattioli et al., 2012; Qadri et al., 2005; WHO-UNICEF, 2012). Water containing pathogenic bacteria, such as *E. coli*, viruses, zoonotic waterborne pathogens or parasites can cause diarrheal disease, which is responsible for 1.2 million deaths annually (Austin et al., 2012; Baldursson
and Karanis, 2011; Mattioli et al., 2012). Gastroenteritis is another prevalent illness associated with drinking water that is from faecally contaminated water sources (Soller et al., 2010). For example, in 2007 there was an outbreak of gastroenteritis in Nokia, Finland after the drinking water supply became contaminated with sewage effluent. Nearly 6,500 people became ill after coming into contact with the infected water (Laine et al., 2011). People who bathe or partake in water related recreational activities, such as surfing or swimming, in water that is contaminated with faeces are at a high risk of contracting illnesses such as gastroenteritis; skin irritations; respiratory disease; eye infections; and ear, nose, and throat infections (Prüss, 1998; Soller et al., 2010; Tseng and Jiang, 2012).

1.2.1 Point and Diffuse Sources

Faecal contamination can arise from both point and diffuse sources (Fig. 1.1). A point source is a single, easily identifiable point of faecal pollution such as a wastewater treatment plant or industrial effluent discharge (Arnscheidt et al., 2007; Converse et al., 2009; Parveen et al., 1997). Diffuse or non-point sources are more complicated as they are dispersed in an irregular and heterogeneous manner with multiple potential inputs of faecal contamination (Converse et al., 2009; Parveen et al., 1997). Faecal pollution from diffuse sources can originate from agriculture where there is runoff from farmyard areas; seepage from manure or slurry incorrectly stored manure; or washings from the milking parlours of dairy farms (Kay et al., 2010; Vinten et al., 2008). Human settlement can contribute to diffuse pollution, through for example, defective septic tank wastewater treatments (Arnscheidt et al., 2007). Domestic and wild animals can also add to the problem of diffuse faecal pollution (Ahmed et al., 2008a; Converse et al., 2009). All of this activity results in a highly varied mixture of animal faecal material and some incidences of human contamination that makes it difficult to determine the dominating source at any one time (Converse et al., 2009; Gawler et al., 2007). This mixture of faecal material can pose a human health risk due to pathogenic bacteria, zoonotic waterborne pathogens, viruses or protozoan parasites (Austin et al., 2012; Baldursson and Karanis, 2011; Mattioli et al., 2012). It is therefore imperative that methods to track faecal pollution can differentiate between human and agricultural sources of faecal contamination.
By determining the origin of the pollution, prevention measures can be used to stop faecal transfers to water (D'Arcy and Frost, 2001; Field and Samadpour, 2007; Scott et al., 2002).
Figure 1.1 Diagram showing point and diffuse (non-point) sources of faecal pollution (Author’s own).
1.2.2 Faecal Indicator Organisms

Faecal indicator organisms (FIO) can be used to identify faecal contamination in water and may suggest the presence of pathogens (Paruch and Mæhlum, 2012; Traister and Anisfeld, 2006). It is easier and more effective to monitor FIO than to monitor pathogens (Edberg et al., 2000; Harwood et al., 2013a). This is because there are a large variety of pathogens, such as bacteria, viruses, zoonotic waterborne pathogens, and protozoa, that can affect human health by gaining access to water via faecal contamination (Austin et al., 2012; Baldursson and Karanis, 2011; Mattioli et al., 2012). They can also be present in low concentrations making their identification a challenge (Harwood et al., 2013a). Furthermore, the methods used to detect pathogens can be difficult to utilise, expensive, labour-intensive, and may not always give accurate results (Edberg et al., 2000). In the late 1890s research began to find a suitable FIO to assess possible faecal contamination in water (Edberg et al., 2000). In order to achieve this, several factors were taken into consideration: the type of FIO that is suitable for tracking faecal pollution; how specific this FIO is; the relationship between water pollution and hydrology; and the methods used to detect this bacterium and associated problems and solutions. Therefore a candidate FIO needed to meet the following criteria:

- Must be universally inhabitant in the intestine of warm-blooded animals
- Should not be pathogenic
- Should be easily detected by simple and cost effective methods
- The concentration of FIO must be of greater value than pathogens
- Should not be able to multiply outside of the host
- Should be resistant to various environmental stresses
- Should have prominent association with pathogens

(Ahmed et al., 2008a; Paruch and Mæhlum, 2012)
Coliforms are gram-negative, rod-shaped, non-spore forming bacteria and have long been used as FIO (Borman et al., 1944; Paruch and Mæhlum, 2012). They are part of the taxonomic family Enterobacteriaceae, which includes the coliform subgroups; Klebsiella, Enterobacter, Serratia, Citrobacter, and thermotolerant coliforms (also known as faecal coliforms). Thermotolerant coliforms are resistant to high temperature regimes, which are supposed to suppress the growth of bacteria that are of non-faecal origin but referring to them as faecal coliforms is technically incorrect as research has shown that they can also come from non-faecal sources (Borman et al., 1944; Foppen and Schijven, 2006; Paruch and Mæhlum, 2012).

1.2.3 Escherichia coli as a Faecal Indicator Organism

*Escherichia coli* is a facultative anaerobe and is the dominant bacterium in the thermotolerant coliform group (Foppen and Schijven, 2006; Paruch and Mæhlum, 2012). It is found in the intestines of humans and the guts of warm blooded animals (Edberg et al., 2000; Foppen and Schijven, 2006). In the 1890s, Theobald Smith studied *E. coli* as a component of the natural flora of mammals and found it to exist in high levels within faecal material (> 10⁹ g⁻¹); he proceeded to suggest it as a suitable FIO for detecting faecal contamination in water (Edberg et al., 2000).

*Escherichia coli* has been the FIO of choice compared to the use of coliforms and thermotolerant coliforms (Foppen and Schijven, 2006; Paruch and Mæhlum, 2012). Apart from *Escherichia*, the other coliform subgroups can be found in the environment in soil or vegetation (Edberg et al., 2000; Paruch and Mæhlum, 2012) and therefore their use can result in false positives for faecal contamination. In the 1970s, research by Dufour confirmed that *E. coli* was the only true “faecal” coliform vindicating its use as a FIO (Dufour, 1977). However, because testing for *E. coli* was expensive and difficult, the faecal coliform test, established by Eijkman in 1904, continued to be used as a surrogate for *E. coli* resulting in many false positives (Edberg et al., 2000).

1.2.3.1 Methods Used to Detect *E. coli*

“*Standard Methods for the Examination of Water and Wastewater*” (1985) states three ways of testing for *E. coli* including: the most-probable-number (MPN)
multiple-tube fermentation; the membrane filtration method (MF); and Defined Substrate Technology (chromogenic substrate method). The first step for both MPN and MF (lactose) methods is to analyse for total coliforms. However, if the water sample analysed yields presumptively positive results for total coliforms, then further experimental investigations need to be performed. For the MPN method, an aliquot of positive sample is added to liquid lactose broth medium containing 4-methyl-umbelliferyl-β-D-glucuronide (MUG). The enzyme β-glucuronidase, which is present in 95% of E. coli, cleaves MUG to produce the fluorescent product 4-methylumbelliferone (Eckner, 1998). After 24 h of incubation, a positive fluorescence determines the presence of E. coli (Eckner, 1998; Edberg et al., 2000).

The MF method allows for enumeration of positive E. coli colonies on the surface of membrane-thermotolerant Escherichia coli (mTEC) agar (Dufour et al., 1981; Edberg et al., 2000). These cheap and simple culture methods allow E. coli to be used as a FIO for regulatory standards (Layton et al., 2006).

Research in the 1980s discovered a new method of measuring water samples for E. coli based on the finding that a positive hydrolysis of MUG resulted in fluorescence; thus Defined Substrate Technology (DST) was developed (Eckner, 1998; Edberg et al., 2000; Traister and Anisfeld, 2006). Powder medium added to the water sample contains a substrate that is used as a food source. The metabolism of the substrate leads to the growth of E. coli that releases and emits a fluorescent signal due to the enzymatic activity of β-glucuronidase (Edberg et al., 2000; Noble and Weisberg, 2005). Multiple instruments can then be used to detect this emitted fluorescence (Noble and Weisberg, 2005). DST was used in the development of Colilert® and Colisure® (IDEXX Laboratories, Inc., Westbrook, ME, USA) where, after an incubation of 18 h for Colilert® or 24 h for Colisure®, total coliforms can be detected by direct one colour visualisation and E. coli can be detected by fluorescence. A positive result for total coliforms is confirmed by the presence of a yellow colour for Colilert® or a magenta colour for Colisure® (Brennan et al., 2010; Edberg et al., 2000). Both experimental tests have a detection limit as low as one colony forming unit of E. coli from a 100 ml sample of water (Edberg et al., 2000). The Irish Environmental Protection Agency (EPA) and the United States (US) EPA both
approve of the Colilert® method for the detection of total coliforms and *E. coli* in water samples, as a result of its cost efficiency, ease of use, and sensitivity for the positive detection of *E. coli* (Edberg *et al.*, 2000; www.epa.gov; www.epa.ie).

### 1.2.3.2 Problems with *E. coli* as a FIO

There have been multiple disadvantages associated with *E. coli* as a FIO in recent years calling into question its suitability as a FIO (Ahmed *et al.*, 2008a; Field and Samadpour, 2007; Layton *et al.*, 2006). There is a need to define and distinguish between origins of faecal contamination for the assessment of water quality (Ahmed *et al.*, 2008a). *Escherichia coli* cannot be used to do this as humans and animals contain diverse numbers of *E. coli* in their faeces, and strains are not genetically unique to one host species or the other (Ahmed *et al.*, 2008a; Field and Samadpour, 2007). As discussed above, two of the criteria a FIO must meet are that it should not have the ability to multiply outside of the host or come from non-faecal sources (Ahmed *et al.*, 2008a). Several studies have shown that *E. coli* can survive and replicate outside of the host in the natural environment and in soil (Ahmed *et al.*, 2008a; Field and Samadpour, 2007). A study of a Lake Michigan stream found that *E. coli* had grown and established a population in the water (Byappanahalli *et al.*, 2003). Research in Ireland showed that *E. coli* can survive for long periods of time in soil, up to 9 years, possibly resulting in false positives for faecal contamination in tested water samples (Brennan *et al.*, 2010). Another disadvantage of this indicator is that it appears to show a weak correlation with pathogens in the water (Ahmed *et al.*, 2008a). In 2004, a study in Finland demonstrated no significant correlation between the detection of *E. coli* and the occurrence of pathogenic *Salmonella* spp., *Giardia* spp., *Campylobacter* spp., *Cryptosporidium* spp., and norovirus (Horman *et al.*, 2004).
1.3 Alternative Methods Used to Detect Faecal Contamination

As detecting *E. coli* might not be a definitive measure of faecal contamination, other methods have been established over the years to detect faecal pollution. There are several other microbial methods that utilise *Bifidobacterium* spp., *Clostridium perfringens*, *Enterococcus*, the ratio of *E. coli* to *Streptococcus*, *Bacteroides fragilis* phage, and F+ RNA coliphage (Ahmed *et al.*, 2008a; Scott *et al.*, 2002). Host-specific viruses, such as norovirus or adenovirus, can be used to establish the source of faecal contamination (Roslev and Bukh, 2011). Chemical methods such as faecal sterols and stanols, optical (laundry brighteners), and caffeine can also be used (Field and Samadpour, 2007; Scott *et al.*, 2002). These methods will be briefly explained in the sections that follow.

1.3.1 Microbial Methods used to Detect Faecal Contamination in Water

1.3.1.1 *Bifidobacterium* spp.

*Bifidobacterium* spp. are obligate anaerobes that have been suggested as suitable candidates for detecting human faecal pollution as they are found in abundance in the human intestine but seldom found in animals (Bonjoch *et al.*, 2004; Mushi *et al.*, 2010). Culturing of these bacteria can be difficult but PCR (polymerase chain reaction) and qPCR (quantitative polymerase chain reaction) methods are now available (Bonjoch *et al.*, 2004). They probably do not replicate outside of the host and thus can be used to provide an evaluation of recent human faecal contamination. However, their die-off rate can be faster than that of many waterborne pathogens, can increase during summer periods when the water temperature is higher, and may also be affected by high levels of predation (Mushi *et al.*, 2010; Rhodes and Kator, 1999). There is also a low sensitivity for detection methods making it difficult to establish definitively the presence of faecal contamination (Ahmed *et al.*, 2008a).

1.3.1.2 *Clostridium perfringens*

*Clostridium perfringens* (*C. perfringens*) are anaerobic pathogenic bacteria which are frequently found in the faeces of humans and animals (Davies *et al.*, 1995; Rood and Cole, 1991). *Clostridium perfringens* were thought to be a good alternative to *E. coli* as they are found in high concentrations in wastewater (10$^3$ to 10$^4$/100 ml) and can
be used to determine the presence of enterovirus (Fujioka and Shizumura, 1985; Tallon et al., 2005). Significant correlations were seen between C. perfringens and Giardia in studies of water in Canada and Australia (Ferguson et al., 1996; Payment and Franco, 1993). Although they do not replicate in the environment they do persist for long periods of time so are not effective as indicators of recent faecal contamination (Desmarais et al., 2002). Another disadvantage is that large volumes of water need to be sampled in order to detect C. perfringens (Farnleitner et al., 2010).

1.3.1.3 Enterococcus spp.
Faecal streptococci include the genus Enterococcus, which consists of at least five species, with E. faecalis and E. faecium being the species most frequently found in humans (Harwood et al., 2004; Scott et al., 2002). Their use as a FIO is limited as they have been found to multiply in the environment (Byappanahalli et al., 2003). A study by Schriewer et al. (2010) found no correlation between Enterococci and waterborne pathogens.

1.3.1.4 The Ratio of E. coli to Streptococcus
Faecal streptococci are found in the faeces of humans, farm animals, dogs, cats and various wild animals (Geldreich and Kenner, 1969; Winslow and Palmer, 1910). Winslow and Palmer (1910) suggested using faecal streptococci as a means of differentiating between human and animal faecal waste. Faecal streptococci are found in lower concentrations in human faecal waste and higher concentrations in animal waste than faecal coliforms (Geldreich and Kenner, 1969; Sinton et al., 1993). Because of this information, some studies suggested using a faecal coliform and faecal streptococci ratio to define the source of pollution as human or animal. A ratio of 4:1 was suggested for human faecal pollution as the concentration of faecal coliforms are approximately four times that of faecal streptococci. For animal faecal pollution, the ratio of 0.7:1 was suggested (Geldreich and Kenner, 1969; Sinton et al., 1993). However, the ratio method must only be used during the initial 24 hour release into water due to a severe decrease in populations of the bacteria, and should be approached with caution for three main reasons (Geldreich and Kenner, 1969; Parveen et al., 1999). First, faecal streptococci shows variable survival rates due to
numerous factors that affect their survival in water such as pH, water temperature, organic nutrients, toxic metal such as copper or zinc, and other ecological factors (Geldreich and Kenner 1969; Parveen et al., 1999; Sinton et al., 1993). Secondly, the die-off rate of faecal coliforms can be faster than faecal streptococci. Lastly, sample storage times can affect the ratio of faecal coliforms to faecal streptococci leading to inaccurate results (Sinton et al., 1993).

1.3.1.5 *Bacteroides fragilis* bacteriophage

*Bacteroides fragilis* (*B. fragilis*) is an obligate anaerobe that is found in great amounts in the intestinal microflora of both humans and animals. The phage that infects *B. fragilis*, specifically HSP40, has been proposed as an indicator for determining human faecal pollution as it is deemed to be specific to human faeces and therefore can distinguish between *B. fragilis* of animals (Harwood et al., 2013b; Tartera and Jofre, 1987). However, their usefulness as an indicator is compromised by the fact that they do not present in all geographic areas, and that it can be challenging to detect the bacteriophage in waters with low levels of faecal contamination (Ahmed et al., 2008a; Scott et al., 2002). GB-124 phage shows potential as an indicator of human faecal pollution but more research is needed on the relationship of the phage with human viruses in numerous water sources (Ebdon et al., 2007; 2011).

1.3.2 Viral Methods to Detect Faecal Contamination in Water

1.3.2.1 F+ RNA coliphage

*Escherichia coli* can be infected by viruses known as coliphages (Ahmed et al., 2008a; Scott et al., 2002). Distinguishing serotypes or genotypes of RNA coliphages are found in animal and human faeces, and in theory, RNA coliphages can be used to differentiate between human and animal faecal pollution (Cole et al., 2003; Jiang et al., 2001). Four sub-groups of F+ RNA coliphage exist and are classified as: groups I, II, III, and IV. Group I is found in humans and animals, II and III are found in human sewage, and group IV are specific to animals (Harwood et al., 2013b; Scott et al., 2002). They have a genetic and physical resemblance to human enteric viruses and therefore can be used as an indicator of viral contamination (Ahmed et al.,
Cole et al. (2003) showed a correlation between F+ RNA coliphage and human adenovirus in water samples taken in Southern California. They do not replicate outside of the host but can occur in low concentrations outside the host and necessitate a need for highly sensitive detection methods (Ahmed et al., 2008a; Scott et al., 2002). Further research is required to elucidate the genetic diversity between the four groups of coliphages, and their individual survival rates in order to warrant them as successful and suitable indicators of human or animal faecal pollution (Ahmed et al., 2008a; Scott et al., 2002).

### 1.3.2.2 Phage as Non-Host Specific Viral Indicators

The PRD1 bacteriophage is an important model virus that aids understanding of the molecular structure and phage host interactions (Daugelavicius et al., 1997; Rydman and Bamford, 2002). It can be used as a tracer and viral surrogate for studies on viral transport in the groundwater environment (Harvey and Ryan, 2004). Advantages of this phage are the functional and structural similarities to the mammalian adenovirus including those that are human pathogens; significant stability in aqueous and geologic media; and it is also easy to cultivate and manage in a laboratory (Harvey and Ryan, 2004; Yahya et al., 1993). PRD1 hosts *E. coli* and *Salmonella* are often found in sewage making this phage a useful indicator of faecal pollution in groundwater (Mazari-Kriart, et al., 1999; Rydman and Bamford, 2002). Nonetheless, more research is needed on surface inactivation and how well PRD1’s transport and survival features correlate with those of disease-causing water borne pathogens before it can be established as a means of testing drinking-water aquifers for vulnerability to microbial pollution (Harvey and Ryan 2004).

### 1.3.2.3 The Use of Host-Specific Viruses to Detect Faecal Pollution in Water

Due to their high host specificity and abundance in the human intestine, there are multiple enteric viruses used for detecting faecal contamination in water (Ahmed et al., 2008a; Scott et al., 2002). Not all of the viruses are detectable by culturing techniques and of those that are, the necessary culturing methods for detection can be difficult to use, time-consuming, and lack sensitivity (Arraj et al., 2005; Baggi et al., 2001). Methods such as PCR and Reverse Transcriptase-qPCR (RT-qPCR) can be used to detect viruses such as human-specific adenovirus, norovirus, and enteric
virus (Griffith et al., 2009; Jiang et al., 2001; Wolf et al., 2010). There are bovine-specific adenovirus, norovirus, enteric virus, and polyomavirus detection assays available (de Motes et al., 2004; Fong et al., 2005; Hundesa et al., 2006). There are also assays to identify porcine- and ovine-specific adenovirus and norovirus (Wolf et al., 2010). A disadvantage to using viruses to detect faecal contamination is that large volumes of water need to be processed in order to identify the low concentrations of virus present (Ahmed et al., 2008a).

1.3.3 Chemical Methods to Detect Faecal Contamination in Water

1.3.3.1 Faecal Sterols and Stanols

Faecal sterols and stanols are popular chemical compounds used for faecal contamination detection (Gregor et al., 2002; Leeming et al., 1996). Animal diet, gut microbes, and metabolism influence the faecal stanols formed from sterols in the intestines of animals (Gregor et al., 2002). Coprostanol is found in the intestines of humans and animals and is often used as a chemical indicator of human faecal contamination (Gregor et al., 2002; Leeming et al., 1996). Problems have arisen with this indicator when it is the only method used to detect contamination as it can produce false results (Shah et al., 2007). Sterol ratios such as cholesterol/coprostanol or coprostanol/epicoprostanol have been used to detect faecal contamination but Shah et al. (2007) showed that faecal sterols do not always successfully determine the percentage that each source contributes in mixed faecal samples. Another drawback to faecal sterols is that no correlation has been seen between them and waterborne pathogens (Ahmed et al., 2008a).

1.3.3.2 Optical (laundry) Brighteners

Optical brighteners or fluorescent whitening agents such as sodium tripolyphosphate, and linear alkyl benzenes are white dyes. They are found in laundry detergents for the purpose of making light colours appear brighter (Ahmed et al., 2008a; Scott et al., 2002; Shu and Ding, 2005). They can be used as indicators of human faecal contamination in water because human sewage contains a large amount of laundry water as plumbing systems generally collect wastewater from both toilets and washing machines (Allevi et al., 2013; Hagedorn and Weisberg, 2009). Optical
brighteners are also found in bleached toilet paper and other products that make their way into the septic tank (Allevi et al., 2013). They can be detected by high-performance liquid chromatography (HPLC); fluorometry detection (sometimes in combination with bacterial counts); and UV exposure of cotton pads that have been soaked in the water to be sampled (Hagedorn and Weisberg 2009). Results can be collected rapidly and easily, but background fluorescence from other compounds can interfere and also concentrations can be low and difficult to detect in large volumes of water (Gregor et al., 2002; Hagedorn and Weisberg 2009). Furthermore optical brighteners can only be used to detect human waste and not animal waste and no correlation has been seen with waterborne pathogens (Scott et al., 2002).

1.3.3.3 Caffeine

Caffeine is one of the most consumed drugs in the world (Buerge et al., 2003). It is found in beverages such as coffee, tea, soft drinks; several food products such as pastries and chocolate; and in many pharmaceutical products such as cough, cold and headache medication (Allevi et al., 2013; Buerge et al., 2003; Scott et al., 2002). Humans who ingest caffeine do not metabolise it but excrete it in urine; therefore the presence of caffeine in water can be used as a chemical marker for human contamination (Buerge et al., 2003; Hagedorn and Weisberg 2009; Scott et al., 2002). It can be identified by high performance liquid chromatography–mass spectrometry (HPLC–MS) or by gas chromatography–mass spectrometry (GC–MS) (Buerge et al., 2003; Hagedorn and Weisberg 2009). However, dilution of caffeine can result in very low levels in water making it difficult to detect, no correlation has been seen with waterborne pathogens, and caffeine can only detect human contamination (Field and Samadpour 2007; Scott et al., 2002).
1.4 *Bacteroidales* – An Alternative Faecal Indicator Organism for Use in Microbial Source Tracking

As there are questions remaining over the suitability of the aforementioned microbial, viral, and chemical methods used to detect faecal contamination, there is a growing need to look for alternative, or additional, faecal indicators (Ahmed *et al*., 2008a; Field and Samadpour, 2007). Research for a new tool used to detect bacteria in water began at the end of the 20th century (Bernhard and Field, 2000b; Harwood *et al*., 2000). This tool was Microbial source tracking (MST) which encompasses several techniques to detect bacteria in water (Field and Samadpour, 2007; Scott *et al*., 2002; Stapleton *et al*., 2007). The most common method used for MST is the detection of host specific bacteria using molecular markers to identify and confirm the source of faecal contamination in water (Field and Samadpour, 2007; Harwood *et al*., 2013a; Scott *et al*., 2002; Stapleton *et al*., 2007). The order *Bacteroidales*, including the genus *Bacteroides*, are one of multiple non-coliform groups that have been recommended for use as an alternative indicator to *E. coli* and other faecal coliforms (Bernhard and Field, 2000b; Fremaux *et al*., 2009; Layton *et al*., 2006). They are strict anaerobes that are, apparently, restricted to the gut of humans and the rumens of warm-blooded animals (Bernhard and Field, 2000a; Fogarty and Voytek, 2005; Kildare *et al*., 2007). Molecular markers that target the 16s rRNA gene sequences have been designed to detect these bacteria in water samples (Bernhard and Field, 2000b; Kildare *et al*., 2007; Lamendella *et al*., 2013; Layton *et al*., 2006).

The potential advantages of these organisms include:

- They compose a considerable portion of the faecal bacteria population, up to 1000-fold higher than coliforms (Bernhard and Field 2000a; Kildare *et al*., 2007; Layton *et al*., 2006)
- There is a decreased prospective for growth after being released into the environment as they are anaerobic and therefore have a low tolerance for oxygen (Bernhard and Field 2000a; Fremaux *et al*., 2009; Layton *et al*., 2006).
They show a high degree of host specificity as a result of molecular diversity between species which gives them the ability to distinguish between animal or human source of faecal pollution (Lamendella et al., 2013; Saunders et al., 2009; Kildare et al., 2007; Layton et al., 2006).

Culturing of *Bacteroidales* or its isolates is very difficult but not necessary as molecular markers are a rapid and effective way of detecting these bacteria (Fremaux et al., 2009; Layton et al., 2006).

Correlations have been observed between the presence of *Bacteroidales* and *Campylobacter, E. coli O157:H7, and Salmonella* (Walters et al., 2007).

The overall advantage of microbial source tracking is that by using culture-independent, host-specific, molecular-based methods, it could be possible to quickly and accurately detect and establish quickly and accurately the source of faecal contamination in water (Field and Samadpour, 2007; Fremaux et al., 2009; Scott et al., 2002).

### 1.4.1 General *Bacteroidales* Assays

Bernhard and Field (2000b) designed GenBac, a general *Bacteroidales* PCR assay that detects all faecal samples containing *Bacteroidales*. BacUni-UCD is a universal Taqman-based qPCR that can do the same. When tested, it detected *Bacteroidales* in faecal samples taken from human, influent from a wastewater treatment plant, cow, dog, cat, horse, and seagull samples (Kildare et al., 2007). Total *Bacteroidales* is a Taqman based qPCR assay that, as the name suggests, also amplifies the total *Bacteroidales* present in faecal samples. It amplified human, cow, dog, cat, pig, elk, and gull faeces (Dick and Field, 2004). Layton et al. (2006) designed AllBac, a general *Bacteroidales* assay that could detect human, bovine and equine faecal sources of water contamination.
1.4.2 Host-Specific Bacteroidales Assays

Multiple molecular assays have been designed to detect human and animal faecal contamination via PCR and/or qPCR. Table 1.1 summarises the most widely employed Bacteroidales assays relevant to the pollution of freshwater systems in regions with a temperate climate, which is the focus of this review. Detailed reviews of the development and testing of other assays have recently been provided elsewhere, as have reports of their applications in diverse environments (Ahmed et al., 2008a; Bernhard and Field 2000a,b; Dick et al., 2005; Field and Samadpour, 2007; Green et al., 2012; Harwood et al. 2013a; Jenkins, et al., 2009; Kildare et al., 2007; Kobayashi et al., 2013; Okabe et al., 2007; Reischer et al., 2007; Shanks et al., 2009).
Table 1.1 List of established molecular assays, PCR or qPCR, designed to target various host-specific *Bacteroides/Bacteroidales*

<table>
<thead>
<tr>
<th>Source Association</th>
<th>Assay</th>
<th>Assay Type</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human</td>
<td>HF134, HF183</td>
<td>PCR</td>
<td>(Bernhard and Field 2000b)</td>
</tr>
<tr>
<td></td>
<td>BacHum-UCD***</td>
<td>PCR, qPCR</td>
<td>(Kildare <em>et al.</em>, 2007)</td>
</tr>
<tr>
<td></td>
<td>HuBac</td>
<td>PCR, qPCR</td>
<td>(Layton <em>et al.</em>, 2006)</td>
</tr>
<tr>
<td></td>
<td>BacH</td>
<td>PCR, qPCR</td>
<td>(Reischer <em>et al.</em>, 2007)</td>
</tr>
<tr>
<td></td>
<td>Human-Bac1</td>
<td>PCR, qPCR</td>
<td>(Okabe <em>et al.</em>, 2007)</td>
</tr>
<tr>
<td></td>
<td>HumM2, HumM3</td>
<td>PCR, qPCR</td>
<td>(Shanks <em>et al.</em>, 2009)</td>
</tr>
<tr>
<td>Cows/Cattle (Bovine)</td>
<td>CF128, CF193</td>
<td>PCR</td>
<td>(Bernhard and Field 2000b)</td>
</tr>
<tr>
<td></td>
<td>BacCow-UCD (BacBov)<em>,</em>**</td>
<td>PCR, qPCR</td>
<td>(Kildare <em>et al.</em>, 2007)</td>
</tr>
<tr>
<td></td>
<td>Cow-Bac1, Cow-Bac1, Cow-Bac3</td>
<td>PCR, qPCR</td>
<td>(Okabe <em>et al.</em>, 2007)</td>
</tr>
<tr>
<td></td>
<td>BacR</td>
<td>PCR, qPCR</td>
<td>(Reischer, <em>et al.</em>, 2006)</td>
</tr>
<tr>
<td></td>
<td>CowM2, CowM3</td>
<td>PCR, qPCR</td>
<td>(Shanks <em>et al.</em>, 2008)</td>
</tr>
<tr>
<td></td>
<td>RumB1F and (BacPreR)<strong>, (Bac32F)</strong> RumD1R, (Bac32F)** and RumD2R</td>
<td>PCR</td>
<td>(Dorai-Raj <em>et al.</em>, 2009)</td>
</tr>
<tr>
<td></td>
<td>Rum2Bac</td>
<td>PCR, qPCR</td>
<td>(Mieszkin <em>et al.</em>, 2010)</td>
</tr>
<tr>
<td>Chicken</td>
<td>Chicken-Bac</td>
<td>PCR, qPCR</td>
<td>(Kobayashi <em>et al.</em>, 2013)</td>
</tr>
<tr>
<td>Dog</td>
<td>BacCan-UCD</td>
<td>PCR, qPCR</td>
<td>(Kildare <em>et al.</em>, 2007)</td>
</tr>
<tr>
<td></td>
<td>DF475</td>
<td>PCR</td>
<td>(Dick <em>et al.</em>, 2005b)</td>
</tr>
<tr>
<td>Duck</td>
<td>Duck-Bac</td>
<td>PCR, qPCR</td>
<td>(Kobayashi <em>et al.</em>, 2013)</td>
</tr>
</tbody>
</table>
**Chapter I**

<table>
<thead>
<tr>
<th>Animal</th>
<th>Primer(s)</th>
<th>Method(s)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elk</td>
<td>EF990</td>
<td>PCR</td>
<td>(Dick <em>et al</em>., 2005b)</td>
</tr>
<tr>
<td>Goose</td>
<td>CGPrevF5, CGOF1-Bac, CGOF2-Bac</td>
<td>PCR, qPCR</td>
<td>(Fremaux <em>et al</em>., 2010)</td>
</tr>
<tr>
<td>Horse</td>
<td>HoF597</td>
<td>PCR</td>
<td>(Dick <em>et al</em>., 2005a)</td>
</tr>
<tr>
<td>Pig</td>
<td>Pig Bac-1, Pig Bac-2</td>
<td>PCR, qPCR</td>
<td>(Okabe <em>et al</em>., 2007)</td>
</tr>
<tr>
<td></td>
<td>Pig-1-Bac, Pig-2-Bac</td>
<td>PCR, qPCR</td>
<td>(Mieszkin <em>et al</em>., 2009)</td>
</tr>
<tr>
<td>Poultry</td>
<td>CP-1-10, CP1-24</td>
<td>PCR</td>
<td>(Lu <em>et al</em>., 2007)</td>
</tr>
</tbody>
</table>

* BacCow-UCD is now known as BacBov-UCD (personal communication)

** Primers in brackets were designed by (Bernhard and Field 2000a)

*** Used in this research
Chapter I

1.4.2.1 Human-Specific Bacteroidales Assays

Bernhard and Field (2000b) designed HF183F, which was the first human-associated Bacteroidales forward primer to be used to identify sources of human faecal material in water. They designed the primer using length heterogeneity – PCR (LH-PCR) and terminal restriction fragment length polymorphism (T-RFLP) to screen faecal bacterial DNA extracts from environmental water samples. HF183F was paired with Bac708R, a general Bacteroidales-Prevotella reverse primer, and the result of this analysis, together with further cloning and sequencing, was a cluster-specific primer set (HF183) for PCR that is successful in detecting human faecal contamination in water samples (Bernhard and Field 2000a; Bernhard and Field 2000b; Kildare et al., 2007). The HF183 assay displayed 100% specificity for human faecal samples and sewage. This primer set has been tested successfully in Australia (Ahmed et al., 2008), Canada (Fremaux et al., 2009), France (Gawler et al., 2007; Gourmelon et al., 2007), Ireland (Dorai-Raj et al., 2009; Gawler et al., 2007), the UK (Gawler et al., 2007), Belgium (Seurnick et al., 2005), Portugal (Gawler et al., 2007), Kenya, (Jenkins et al., 2009) and the USA (Bernhard and Field 2000b).

In Australia, HF183 was assayed against 52 sewage and septic tank samples and showed 100% specificity by detecting the presence of human faeces in all samples. HF183 showed good specificity in Saskatchewan, Canada as it was 100% specific for raw sewage and 94% specific for human faecal samples (Fremaux et al., 2009). In 2005, Seurnick et al. used SYBR Green \textsuperscript{®} real-time PCR (also known as qPCR) to evaluate the specificity of the marker in Belgium. This adaptation to the original method was shown to be specific for 5 out of 7 human individual faecal samples tested, and with a 99% specificity in sewage, it demonstrated that it could be successfully used as a tracer of sewage faecal pollution (Ahmed et al., 2008b; Seurnick et al., 2005). In France, one study showed 91% specificity for human faecal samples (Gawler et al., 2007), the other study showed 94% specificity for human faecal samples and 100% specificity for human waste effluents (Gourmelon et al., 2007). In Ireland and the UK, the molecular marker, HF183, displayed 100% specificity for human faecal samples, 96% specificity for human faecal identities in a study in Portugal, 100% specificity by a research group in Kenya, and 100% specificity for human faecal samples analysed in the USA (Dorai-Raj et al., 2009;
Fogarty and Voytek 2005; Gawler et al., 2007; Jenkins et al., 2009). All of the studies also tested against faecal samples from various animals including cow, pig, sheep, duck, deer, goat and horse with no cross-amplification being observed. The high levels of specificity and reproducibility associated with the HF183 human-specific assay, across several countries and continents, demonstrate the great potential its use has for identifying the presence of human faecal contamination in water (Ahmed et al., 2008b; Bernhard and Field 2000b; Dorai-Raj et al., 2009; Fogarty and Voytek 2005; Gawler et al., 2007; Gourmelon et al., 2007; Jenkins et al., 2009; Seurnick et al., 2005).

Kildare et al. (2007) designed BacHum-UCD in silico, the human–specific PCR and Taqman based qPCR assay. Host-specific Bacteroidales DNA sequences for human, cow, and dog were obtained using the Genbank Database and aligned against each other to determine a sequence specific to each species. These sequences were used as templates to design primers and a probe specific for the assay. The specificity has previously been assessed in the USA, and of the 18 human faecal samples, and the 14 raw sewage samples analysed, human-host Bacteroidales amplified 67% and 100%, respectively, and demonstrated a 98% specificity (Kildare et al., 2007). Some cross-amplification was observed with dog, with the authors suggesting horizontal gene transfer as a possible cause (Kildare et al., 2007). In 2009, 96% specificity was obtained when the molecular marker was examined in Australia (Ahmed et al., 2009). A minimal amount of cross-amplification was seen with pig, sheep, horse and dog (Ahmed et al., 2009). The human-specific primer set was further investigated in Kenya, demonstrating 100% specificity for human faeces isolated from collected water samples. The high level of specificity determined by these studies suggests that BacHum-UCD can be used as an evaluation method of sewage pollution in several countries, ranging from the industrialised world to the low-income countries, to aid the assessment of water quality, especially as mixed-human faecal input (sewage) is generally the primary cause of human faecal contamination rather than individual human input (Ahmed et al., 2009; Kildare et al., 2007).

Another human–specific PCR and Taqman based qPCR assay, HuBac, was designed by Layton et al. (2006) through the alignment and amplification of DNA sequences
from faecal samples isolated from human, cow, and dog. The specificity of this molecular marker was first evaluated in the USA where the assay was designed. A number of animal faecal samples obtained from dog and pig tested positive, and the cross-amplification of these lead to a reduction of the specificity for the identification of human faeces. In Australia, the HuBac primer set showed cross-amplification with cattle, pig, sheep, horse, dog, duck, and kangaroo (Ahmed et al., 2009). The cross-amplification with pig, which is non-ruminant, suggests that their intestines might be more similar to those of humans, also non-ruminant, than to ruminant mammals such as cow or sheep (Layton et al., 2006). The HuBac assay can be described as selective for human-associated Bacteroides and not specific, and therefore in cases where multiple sources of faecal pollution were found, the lack of specificity of this molecular marker means that its uses are limited, and may not be a suitable indicator to determine the faecal source in contaminated water (Ahmed et al., 2009; Layton et al., 2006).

1.4.2.2 Bovine-Specific Bacteroidales Assays

The methods used to design each bovine-specific Bacteroidales assay were the same methods as the human-specific assay it was designed alongside. For example; the bovine-specific Bacteroidales primer set, CF128 was designed using the same method described to establish HF183, as both were designed by Bernhard and Field (2000b).

CF128 is a ruminant-specific marker designed by Bernhard and Field (2000b) that successfully distinguished between human and cow faecal samples via PCR. CF128 was able to differentiate between cow and human faecal samples tested by Fogarty and Voytek (2005) in the USA. It showed 100% specificity in four Atlantic rim countries: France, Portugal, UK, and Ireland (Gawler et al., 2007). It was also found to be 98% specific in Saskatchewan, Canada (Fremaux et al., 2009). However, the bovine-specific primer set was unable to differentiate between faecal extracts derived from ruminant; deer, moose, caribou, bison, and goat, and pig, a non-ruminant species (Fremaux et al., 2009). The inability of the CF128 assay to differentiate between bovine and swine faecal samples has also been previously reported in France, Ireland, and the USA (Dorai-Raj et al., 2009; Gourmelon et al., 2007;
Chapter I

Shanks et al., 2010). Cross-amplification was also reported from these studies with sheep, deer and goat. A study conducted by Shanks et al. (2010) demonstrated that CF128 had 76% specificity for bovine but cross-amplified with chicken, dog and duck.

Another bovine-specific Bacteroidales primer, BacCow-UCD (now known as BacBov), was designed by Kildare et al. (2007) in the USA. It was designed with the aforementioned CF128F as the forward primer, and a newly designed reverse primer, and probe (for Taqman based qPCR) and when tested showed 100% specificity for bovine faecal samples but also amplified 38% of horse faecal samples examined and is referred to as a “cow/horse” molecular marker (Kildare et al., 2007). In Kenya cross-reactivity with human or donkey faecal samples was not detected (Kildare et al., 2007). Layton et al. (2006) designed the BoBac PCR and Taqman based qPCR assay that was highly specific for Bacteroides species found in bovine faeces but showed cross-amplification with one of the four canine faecal samples. As some of the bovine-specific assays demonstrate cross-amplification with other ruminants this is an area that needs to be improved and a more bovine-specific molecular marker needs to be established.

1.4.3 Measuring Survival Rates of Bacteroidales

DNA can survive in the water for several days or even weeks (Bae and Wuertz, 2009a; Walters and Field, 2009). As a result, there is a need to distinguish between live and dead cells. Bae and Wuertz (2009a) designed a qPCR method to successfully discriminate between viable and non-viable Bacteroidales DNA using propidium monoazide (PMA), a DNA intercalating dye that only enters dead cells containing a cell envelope structure that is no longer intact. The dye binds to extracellular (free) DNA, and to the dead cells’ DNA, inhibiting it from amplifying via qPCR (Bae and Wuertz, 2009a; Nocker et al., 2006). Bae and Wuertz (2009a) used the PMA qPCR to demonstrate that host-specific Bacteroidales DNA, BacHum-UCD and BacBov-UCD, can survive longer in seawater when compared to DNA viability within host-specific Bacteroidales cells. Bae and Wuertz (2012), performed a continuous flow microcosm experiment to compare the decay rates of Bacteroidales, and the relationship of the bacterium to pathogens such as
Enterococcus spp., Campylobacter jejuni, Salmonella enteric, Serovar Typhimurium, and adenovirus. Similar results recorded from the earlier study showed again that Bacteroidales DNA survived longer than DNA within Bacteroidales cells. They also showed that sunlight had little effect on Bacteroidales DNA but caused Bacteroidales cells to decay faster demonstrating that the amount of available sunlight is an important variable to consider when sampling from the environment to measure DNA content of Bacteroidales for assessing survival rates. It was concluded that the PMA qPCR assay method can be used to determine recent faecal contamination sources and to quantify the presence of waterborne pathogens, that is a direct benefit to human health by identifying potential risks that can be found in water (Bae and Wuertz, 2012).

Measuring RNA content is another approach that can be used to establish the difference between viable and non-viable cells as RNA degrades rapidly allowing for a more accurate evaluation of bacterial survival and growth in varying environmental conditions (Kerkhof and Ward, 1993; Weller and Ward, 1989). Walters and Field (2009) performed a microcosm experiment, with both light and dark incubations, to compare the survival rates of host-specific Bacteroidales DNA and RNA, and the effect light has on both. They studied the human-specific molecular markers, HF134 and HF183, and the ruminant-specific molecular markers, CF128 and CF193. The extracted RNA was first reversed transcribed to cDNA, and both it and DNA were measured by SYBR Green qPCR, demonstrating that the human-specific Bacteroidales DNA survived longer compared to the human-specific Bacteroidales cells, and the ruminant-specific DNA compared to Bacteroidales cells were found to be more variable but the human-specific cells did not survive as long as the ruminant-specific cells (Walters and Field, 2009).

Liang et al. (2012) also set up a microcosm to compare the survival rates of host-specific Bacteroidales measuring DNA and RNA contents. It was found that the decay rates were very similar for the human-specific molecular marker, HF183, and for the ruminant-specific molecular marker, CF193, for both DNA and RNA, and concluded that the decay rates from their study were different to other studies because of different environmental parameters, and stressed the need for measurements of Bacteroidales in a more natural environment.
1.4.4 Survival Abilities of Bacteroidales

There is limited information currently available describing the survival capabilities of Bacteroidales in freshwater and sediment environments (Field and Samadpour, 2007; Scott et al., 2002; Sokolova et al., 2011). Survival of Bacteroides cells is assumed to be temperature, sunlight, and predation dependent (Bernhard and Field, 2000a; Sokolova et al., 2011). It is perceived that they can survive for up to six days under oxygen stress conditions (Bernhard and Field, 2000a). Several studies have been conducted to establish the decay rates of Bacteroidales in the environment or varying environmental settings to understand further this microorganism, as previous research has been restricted to laboratory-based experiments (Bae and Wuertz, 2009b; Sokolova et al., 2011; Tambalo et al., 2012; Walters and Field, 2009).

Sokolova et al. (2011) measured the human-specific molecular marker BacH, and the ruminant-specific molecular marker, BacR in water during the course of spring, summer, and winter (only BacR) conditions in Sweden to determine the effect of temperature on the survival rates of Bacteroidales. They found that both BacH and BacR only survived for a maximum of two days in water during summer as a result of the high temperature of approximately ~20°C. In spring and winter, where the temperature was much lower at 5-6°C, survival increased, with BacH surviving 6-8 days in spring, and BacR surviving for up to 6 days in spring and 9-10 days in winter (Sokolova et al., 2011).

1.4.5 Successful Application of Bacteroidales Molecular Assays

All of the aforementioned research demonstrates that host-specific Bacteroidales molecular markers can be used to determine successfully the source faecal contamination in water. Another advantage for the use of Bacteroidales as a faecal indicator is that Bacteroidales can accurately distinguish between human and ruminant sources of faecal pollution in water due to the high level of specificity associated with the host organism. The use of molecular techniques such as PCR and qPCR, allows for a rapid and sensitive detection of contamination. The issues related to the survival rates of the Bacteroidales have been noted, and successfully overcome using the previously discussed methods, PMA-qPCR or measuring RNA content to assess cell viability.
1.5 A Tiered Approach to Improving Water Quality Monitoring and Protection

No single microbial, chemical or viral method has the characteristics needed to be the ultimate microbial source tracking tool (Boehm et al., 2003; Field and Samadpour, 2007; Gourmelon et al., 2010; Noble et al., 2006). Complex interacting systems and usually competing objectives impact on the successful management of land and water, which in Europe, is vital for the implementation of the EU Water Framework Directive (WFD) regulations (Allan et al., 2006; Letcher et al., 2006). Therefore, using a range of established and novel techniques in unison can help characterise the point and diffuse faecal pollution sources that impact water quality (Allan et al., 2006; Noble et al., 2006). This “toolbox” or tiered approach allows for a more effective and targeted approach to resource management and remediation (Gourmelon et al., 2010; Noble et al., 2006). The “toolbox” approach can be used to effectively establish the source of faecal pollution and determine the level of impact that a particular source has on water which assists an informed decision on which source to prioritise (Plummer and Long, 2007; Boehm and Fuhrman, 2003; Noble et al., 2006). Many of the aforementioned techniques have complemented each other when used in the “toolbox” approach and contributed important information on spatiotemporal variability of pollutants, as well as determining the source (Ahmed et al., 2007; Allan et al., 2006; Boehm and Fuhrman, 2003).

Boehm and Fuhrman (2003) used a three tier approach to determine the source of pollution in Avalon Bay, California. The first step was to review the temporal and spatial irregularity of the pollution signal and try to predict the impact on FIO that sunlight and coastal tides would have. The second step was to use standard FIO tests, such as \textit{E. coli}, to try and establish “hotspots” and pollution sources. The third step involved using human-specific \textit{Bacteroides} and enterovirus PCR assays to determine if the pollution source was human (Boehm and Fuhrman, 2003). Their results illustrated that there was faecal contamination evident in the bay as they detected relatively high levels of \textit{E. coli}. Using the human-specific assays, they were able to conclude that the predominant source of human contamination was close to shore.
and not from inside or outside of the bay which was significant information for management decisions (Boehm and Fuhrman, 2003).

Plummer and Long (2007) also used a multi-tiered approach to analyse water quality in Massachusetts. The study took a “weight of evidence” approach and analysed land use via Geographical Information Systems (GIS) and sanitary surveys; monitored water quality using traditional methods such as FIO, chemical and physical parameters; and determined the source using MST targets (\textit{R. coprophilus}, sorbitol-fermenting \textit{Bifidobacteria}, and coliphages). They found that the overall quality of the water could be effectively assessed by traditional methods and that one of each parameter was adequate. Analysing multiple parameters did not provide any extra information (Plummer and Long, 2007). The MST targets successfully differentiated the source between human and cattle. The authors discussed how the negative results that they found could not be deemed representative until those areas were tested over different hydrological and seasonal conditions. If the results were still negative after this analysis then it could be concluded that the source was being effectively managed (Plummer and Long, 2007).

A multi-disciplinary study across six research institutes in France used microbial, chemical and viral methods in a multi-tiered approach to differentiate between human, bovine and porcine faecal pollution in surface waters (Gourmelon et al., 2010). For the microbiology, FIOs such as \textit{E. coli} plus various \textit{Bacteroidales} PCR assays: a human-specific marker, HF183; two pig-specific markers, Pig-1-Bac and Pig-2-Bac; and a ruminant-specific marker Rum-2-Bac were used. There was also a third pig-specific marker, \textit{Lactobacillus amylovorus} (\textit{L. amylovorus}) and a second human-specific marker, which was a phylotype related to \textit{Bifidobacterium adolescentis} (\textit{B. adolescentis}). The chemical markers used were caffeine; seven synthetic compounds (such as TCEP and benzophene) that have been used to determine human faecal pollution in other studies; and faecal sterol and stanol fingerprints. Viruses were analysed using F-specific RNA Bacteriophage (FRNAPH) with genogroup II and III being human-specific and genogroup I and IV being animal-specific (Gourmelon et al., 2010). Their results illustrated that using a ratio
of coprostanol to 24-ethylcoprostanol was more effective in determining the source of faecal pollution than using just one of the stanols. They established that a ratio of $<60\%$ coprostanol/24-ethylcoprostanol (R1 ratio) could detect animal faecal pollution and that a ratio of $>60\%$ coprostanol/24-ethylcoprostanol was able to determine human faecal pollution. They were then able to further differentiate between bovine and porcine faecal pollution by establishing a ratio of sitostanol to coprostanol (R2 ratio). A ratio of $>1\%$ sitostanol/coprostanol suggested that the source was bovine and a $<1\%$ suggested the source was porcine (Gourmelon et al., 2010). When the $E. \ coli$ counts were over 500 CFU/100 ml, the microbial, chemical and viral markers successfully established the source of the faecal pollution. Human faecal pollution was determined using caffeine, TCEP, benzophene, the R1 and R2 ratios, the qPCR bacterial markers (HF183 and $B. \ adolescents$) and genotype II of FRNAPH. A source of faecal pollution was effectively identified as pig by the use of the steroid ratios R1 and R2, and the qPCR markers, Pig-2-Bac and $L. \ amylovorus$. Ruminant faecal pollution was established by using the steroid ratios and the qPCR marker, Rum-2-Bac. The study found that these methods were the most effective at differentiating between human, porcine and bovine faecal pollution (Gourmelon et al., 2010).

These studies demonstrated that the “toolbox” approach can be used successfully especially when FIO, microbial and chemical MST tools, water parameters and land use are utilised simultaneously to achieve an accurate representation of water quality in a catchment (Allan et al., 2006; Plummer and Long, 2007). In Europe, applying these tools at a catchment level would give a better understanding of faecal pollution sources and give significant data that could be used by managers and researchers trying to implement the regulations of the EU WFD in a successful and cost effective manner (Ahmed et al., 2007; Allan et al., 2006; Gourmelon et al., 2010; Noble et al., 2006).
1.6 Nutrients and their Affect on Water

Anthropogenic and agricultural faecal contamination, from point and diffuse sources, contains harmful bacteria, as discussed earlier, but it also contains nutrients such as phosphorus (P) and nitrogen (N) (Jordan et al., 2012; McDowell, 2006). The EU WFD has set limits for these nutrients in freshwater; in Ireland, for example, this limit for molybdate reactive P (MRP) is 0.035 mg L$^{-1}$ in rivers and groundwater, and a recommendation for N (as Nitrate-N) of 11.3 mg L$^{-1}$ for drinking water (S.I. No. 272, 2009).

Excessive accumulations of these nutrients can cause eutrophication in lakes, rivers, streams and the oceans (Arnscheidt et al., 2007; Preedy et al., 2001). This often results in the proliferation of phytoplankton cells leading to algal blooms or aquatic vegetation, which produces large bacterial growth leading to decreased oxygen content (hypoxia) or a total depletion of oxygen (anoxia) in water (Anderson et al., 2002; Bennett et al., 2001; Conley et al., 2009; Smith et al., 1999). This can cause mass fish kills in water bodies with low flushing rates or of a stagnant nature (Withers and Jarvie, 2008). One of the environmental impacts of toxic algal blooms is the retention of potent toxins by filter-feeding bivalves such as shellfish (Ibelings and Chorus, 2007). Eutrophication in freshwater lakes, rivers, and streams is most commonly caused by disproportionate amounts of P (Smith et al., 1999), and excessive concentrations of N are the main cause of eutrophication in the oceans (Howarth and Marino, 2006).

Septic tanks systems (STS) are the principal method used to treat human faecal waste in many rural areas of the world, including Australia, Europe, Canada, Japan, USA and New Zealand, but can be a potential source of human faecal contamination if they are not properly managed and the tank becomes defective (Ahmed et al., 2005; Allevi et al., 2013; Gilpin et al., 2003; Harwood et al., 2013a; Macintosh et al., 2011; Marti et al., 2013; Nnane et al., 2011; Pronk et al., 2009; Savichtcheva and Okabe, 2006; Shah et al., 2007; Withers et al., 2014). Until recently, septic tanks were not seen as significant contributors to faecal pollution or eutrophication but research has now shown that discharges from faulty STS can significantly impact
water quality as they are a diffuse source of P and N (Macintosh et al., 2011; Withers et al., 2011). Ireland’s rural topography is particularly susceptible to the possible negative impacts of STS as rural localities have large and scattered populations with land areas of low soil permeability (Macintosh et al., 2011). These negative impacts are of great concern as approximately 70% of STS are not functioning correctly across the UK and Ireland (Arnscheidt et al., 2007; Withers et al., 2012). It has been shown that STS not performing correctly can contribute to low-level but continuous sources of P to rural water bodies (Arnscheidt et al., 2007; Macintosh et al., 2011) and that during the ecologically active summer period defective STS can be a considerable source of P. As it is now known that poorly functioning STS can contribute continuous low levels of nutrients, such as P and N, to water it is possible that some of the nutrients originally presumed to be from agriculture could actually be of anthropogenic origin (Withers et al., 2014).

Agriculture, especially if it is intensive, can contribute high levels of P and N to water bodies (Edwards et al., 2008; Jordan et al., 2012; Monaghan et al., 2008). Agricultural faecal runoff that contains high levels of P, will also contain high levels of enteric bacteria such as E. coli (Bowes et al., 2003; Carrillo et al., 1985). Studies in New Zealand found that the highest levels of N loss derives from animal urine patches on grazed land (Monaghan et al., 2008).

Research has shown that farmyards are a strong potential source for runoff that is highly polluted and contains nutrients such as P and N (Edwards et al., 2008; Withers and Jarvie, 2008). Edwards et al. (2008) studied runoff and seepage from farm buildings, the roofs of farm buildings, hard standing areas, cattle walkways, and defective storage units of livestock waste (slurry/manure). They found that runoff from hard standing areas and roofs contained high levels of P and N. This made the areas a high risk source of contamination especially during hydrologically active periods, when the source is limited, where high flows can move the contaminants from the farm through transport routes into receiving areas such as open water bodies. During dry periods, seepage from defective storage units of slurry, or farm buildings has the potential to move pollutants to the transport route where they can
aggregate until moved at a later time (Edwards et al., 2008). Edwards et al. (2008) also discussed how periods of light precipitation can play an important role in the movement of highly contaminated runoff from areas of low permeability into water.

As a source of contamination agriculture is of particular concern for surface water that is fed by groundwater, as P and N can be transferred to water via the soil-aquifer continuum (Mellander et al., 2013). Haygarth et al. (2005) described the nutrient transfer continuum, which is based on the concept that nutrients from agriculture can be moved via hydrological pathways to water bodies such as streams and rivers where an impact on water quality may occur. Phosphorus is most often transferred via overland flow to water either in soluble or particulate form (McDowell, 2006; Preedy et al., 2001) but can also leach to groundwater pathways and use these to travel to streams (Haygarth et al., 1998; Holman et al., 2010; Holman et al., 2008). Usually N is leached using groundwater pathways from surface sources but can also travel via surface runoff (Mellander et al., 2012a).

1.6.1 Monitoring of Nutrients in Water by Grab Sampling
Routine monitoring of nutrients by regulatory bodies is achieved by taking grab samples at known times on a weekly or monthly basis. This can be a problem though as grab sampling makes the assumption that regular sampling allows for an accurate representation of nutrient concentration, which is not always correct (Jordan et al., 2007; Kot et al., 2000; Martin et al., 1992). If a catchment has a “flashy” hydrology, i.e., is characterised by impermeable soil and/or steep slopes, then a high flow event of great magnitude but brief time period could input substantial concentrations of nutrients into a water body, which subsequently might be missed by grab sampling because the event times do not correspond to the sampling times (Cassidy and Jordan, 2011). This results in an underestimation of the nutrient concentrations; an underrepresentation of annual fluxes; and hides the dynamics of sub-annual transfers (Jordan et al., 2007; Kot et al., 2000; Martin et al., 1992; Rozemeijer et al., 2010). A way of combating these issues would be to increase the number of samples and the frequency of the sampling regime but that can be expensive and laborious (Kot et al., 2000; Rozemeijer et al., 2010).
1.6.2 Monitoring Nutrients in Water on a Continuous Basis

Because of the limitations of grab sampling regimes, there is a need for high frequency sampling of nutrients, which can be achieved by continuous monitoring instrumentation such as bank-side auto-analysers (Cassidy and Jordan, 2011; Jordan et al., 2007; Jordan et al., 2005). Jordan et al. (2005) set up a bank-side auto-analyser in the Lough Neagh basin, Northern Ireland, and monitored total P (TP) over a prolonged period of time. The equipment used was Dr. Lange Sigmathax and Phosphax instrumentation which is proficient in taking water samples every 10 min and measuring for TP photometrically using the molybdate-antimony-ascorbic acid method (Jordan et al., 2005). Discharge was measured at the same site and using these data it was possible for TP loads to be calculated. The instruments are also capable of sufficient sampling during low flow. Jordan et al. (2005) found that by using this high temporal resolution sampling they were able to show that storm events were responsible for the transfer of high TP loads into Lough Neagh causing a lake eutrophic impact but that the non-storm TP loads kept the stream water in an eutrophic state between high flows. Had grab samples been used to analyse this water it is possible that these significant results would have been missed.

Wall et al. (2011) also monitored TP and discharge on a continuous basis, across six catchments in Ireland, but in addition they measured total reactive phosphorus (TRP), total oxidised nitrogen (TON), turbidity, and electrical conductivity. Nitrogen as TON was measured by a Hach-Lange Nitratax SC plus using UV absorbance technology. Turbidity was measured using Hach-Lange Solitax 0-1000 NTU probes (Melland et al., 2012; Wall et al., 2011). Sub-hourly data were collated to hourly average synchronous water quality and discharge parameters, and underwent quality control using the WISKI-7 (Kisters, 2011) database management system. Wall et al. (2011) were able to distinctly measure non-point, surface flow dependent P transfers during specific high-flow events. They also demonstrated that diffuse, subsurface N transfers from permeable sub-surface layers of rock are displayed as signals in baseflow. Similar studies across catchments in Ireland and Northern Ireland demonstrate the benefits of high temporal resolution sampling over grab or “snapshot” sampling (Cassidy and Jordan, 2011; Macintosh et al., 2011; Melland et al., 2012; Mellander et al., 2012a).
1.7 The European Water Framework Directive

In 2000, the EU established the WFD (2000/60/EC) which is an integrated management plan that aims to improve water quality throughout the member states (CEC, 2000; Kay *et al*., 2010; Stapleton *et al*., 2008). It prescribes the examination of the water quality of rivers, lakes, groundwater, and coastal waters with the objective of establishing at least “good ecological status and quality” of European watercourses by 2015 (CEC, 2000; Stapleton *et al*., 2008) or defined targets thereafter. To fulfil this aim, the WFD takes an integrated catchment-wide approach by determining and controlling point and non-point sources of pollution in water that is utilised for recreation, water supply, fisheries, or eco-system maintenance (Kay *et al*., 2010; Stapleton *et al*., 2008). For faecal contamination, this establishes agreement with the microbiological standards set by the Bathing Water Directive (BWD 76/60/EEC) (CEC, 1976). The BWD was revised in 2006 to set stricter microbiological standards, and interpret water quality variation, and faecal indicator fluxes (Kay *et al*., 2010; Wyer *et al*., 2010).

The “health-evidence based” microbiological guidelines for recreational waters instituted by the WFD were a result of the guidelines inaugurated by the World Health Organisation (WHO, 2003) and (Kay *et al*., 1994, 2004, 2006; Stapleton *et al*., 2008). For *E. coli* and enterococci, the EU drinking water standard in Ireland is 0 CFU/100 ml (S.I. 122, 2014). The Federal Water Pollution Control Act in the USA established the “Total Maximum Daily Load” (TMDL) method which gave rise to complementary standards in the WFD catchment-wide management of water quality (Kay *et al*., 2008a; Stapleton *et al*., 2008; Wang *et al*., 2010). If an identified water course does not abide by the specific water quality principles laid out then the state must indicate the total load of a pollutant a water course can receive while still adhering to water quality standards (Kay *et al*., 2008a; Stapleton *et al*., 2008; Wang *et al*., 2010). By defining the point and diffuse sources, which are the origin of pollution loads, procedures can be established to reduce this load such as upgraded treatment of human waste and/or better farming practices (Kay *et al*., 2008a; Stapleton *et al*., 2008).
As part of the EU WFD, an agricultural programme of measures known as the Nitrates Directive (ND) was established (Jordan et al., 2012; Kay et al., 2008a; Wall et al., 2011). It is a National Action Programme (NAP) aimed at protecting water from transfers of N and P during hydrologically active periods by restricting the timing, amount, and where the spreading of inorganic fertilisers and organic manure can take place (Fealy et al., 2010; Jordan et al., 2012; Macintosh et al., 2011). The NAP intends to manage nutrients at farm and field scale as the majority of annual loads in rural river catchments are non-point source nutrient transfers (Fealy et al., 2010; Jordan et al., 2012; Smith et al., 2005). The ND establishes measures on both N and P when N is defined as a water quality risk (Jordan et al., 2012). For intensive grassland farming, the stocking density is limited to 170 kg ha\textsuperscript{-1} yr\textsuperscript{-1} of organic nitrogen (ON). A derogation of up to 250 kg ha\textsuperscript{-1} yr\textsuperscript{-1} of ON can be applied for if a higher stocking density is in place but this is on the condition of stricter nutrient management measures (Fealy et al., 2010; Jordan et al., 2012; Mellander et al., 2012a). Ireland is divided into three zones, based on the intrinsic vulnerability of the topography, which are roughly coordinated with predicted annual rainfall meaning the “fertiliser closed periods” are longer in the north of the country (Jordan et al., 2012). Each zone has open and closed periods for spreading chemical fertilisers, organic fertilisers (slurry), and farmyard manure (e.g., the details of periods applied in Ireland are provided in Table 1.2). Jordan et al. (2012) found that incidental P losses were limited as residual losses from the soil P store were still high, principally in flashy catchments during the winter closed period for slurry spreading.
**Table 1.2** Dates for cattle manure storage periods, and opened and closed periods for spreading (chemical and organic fertilisers, and farmyard manure) for the all the counties (divided into zones) in Ireland.

<table>
<thead>
<tr>
<th>Zone</th>
<th>Counties</th>
<th>Storage period for cattle manure</th>
<th>Prohibited application periods</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Chemical fertilisers</td>
</tr>
<tr>
<td>A</td>
<td>Carlow, Cork, Dublin, Kildare, Kilkenny, Laois, Offaly, Tipperary, Waterford, Wexford, Wicklow</td>
<td>16 weeks</td>
<td>15 September–12 January</td>
</tr>
<tr>
<td>C</td>
<td>Donegal, Leitrim, Cavan*, Monaghan*</td>
<td>20 (*22) weeks</td>
<td>15 September–31 January</td>
</tr>
</tbody>
</table>
1.8 Catchment Science

1.8.1 The Integrated Catchment Science Approach in the EU

As part of the EU WFD (2000/60/EC), there was a move away from point source contamination analysis towards a catchment-wide approach (Haygarth et al., 2005; Kay et al., 2005; Stapleton et al., 2008). This integrated catchment management (ICM) is in agreement with the Driving Forces-Pressure-State-Impact Response (DPSIR) nexus established by Schmid and Sinabell (2007) as the need to establish “good ecological status” by 2015 is driving the policy makers towards an integrated approach to water quality assessment (Collins and Ison, 2010). The catchment, with its environmental heterogeneity, is where this nexus is the unit of education for local, provincial, national and international water policies as it is “naturally established and managed” at this scale (Schmid and Sinabell, 2007). ICM studies the relationship between land and water, and takes into consideration all circumstances and elements such as land-use, rainfall, soil, vegetation, seasons, run-off and groundwater, that could impact on water (Collins and Ison, 2010; Pollard, 2002; Uhlenbrook, 2006; Wheater and Peach, 2004). This approach promotes the monitoring of point and diffuse sources of nutrients and faecal pollution to be undertaken in unison so that good management practices can be put in place to protect the quality of water where it is used for water supply, eco-system maintenance, fisheries, or recreation (Stapleton et al., 2008). Catchment management is an inter-disciplinary science that requires comprehensive insight, meticulous data analysis, effective scientific apparatus, and interactions with closely-related disciplines such as hydrology, ecology, geography, topography, agriculture, microbiology, engineering and soil science (Kay et al., 2008b; MacKinnon and Tetzlaff, 2009; Uhlenbrook, 2006).

Hydrology is a very important part of catchment science as it is imperative to know and understand hydrological mechanisms and how they can vary on a spatial and a temporal scale (Uhlenbrook, 2006; Wheater and Peach, 2004). Catchment hydrology does not focus on specific units but on the catchment as a whole and looks at the complex non-linear interactions between the water and the land (Uhlenbrook, 2006). Watersheds and drainage are an important factor in establishing catchment boundaries and making the catchment a hydrologically, ecologically and
topographically suitable management entity in order to establish a better understanding of these diverse environments and aid in the development of sustainable water resources (MacKinnon and Tetzlaff, 2009; Uhlenbrook, 2006). Haygarth et al. (2005) built on the source-pathway-receptor model and set up a source-mobilisation-delivery-impact model to study the transfer of phosphorus at the catchment scale. They concluded that the source and mobilisation aspects of their model were well defined but that there was a need for more research to be conducted into the delivery and impact of nutrient transfers at a catchment scale (Haygarth et al., 2005).

1.8.2 The Integrated Catchment Science Approach in Ireland
Catchments that were selected for ND research as part of the Agricultural Catchment Programme (ACP) (www.teagasc.ie/agcatchments/), in Ireland, underwent a rigorous selection process and met all of the following criteria (Fealy et al., 2010):

- Must contain a watercourse: lake, river, groundwater pathways, aquifer
- Be representative of various agricultural practices at high intensity
- Have different geology and topography
- Pose a risk of nutrient transfer
- Must not contain significant non-agricultural nutrient sources

Six catchments were selected across the republic of Ireland (Fig. 1.2) and there is ongoing research into the implications of the nitrates Directive, the effect on the environment from nutrient transfer and from land-use, and establishing the sources of agricultural ruminant and residual human pollution: (Jordan et al., 2012; Melland et al., 2012; Mellander et al., 2012a; Mellander et al., 2012b; Mellander et al., 2013; Wall et al., 2011; Wall et al., 2012).
Figure 1.2 The six catchments that are part of the Teagasc ACP (From Mellander et al., 2012a).
Chapter I

1.9 The Future of Catchment Science

1.9.1 Research Needed for Protection of Water Quality

Accurately and successfully assessing water quality is of paramount importance if human health and the environment are to be protected. While there is a large selection of tools already in existence more research is required to establish how practical and effective the tools are in achieving this aim. Many studies have highlighted the requirement for empirical data on \( E. coli \) fluxes in high and low flow in various seasons as there has been limited research on this to date (Kay et al., 2008b; Mellander et al., 2012a; Tetzlaff et al., 2012). Some research has been able to separate the hydrological pathways into their different nutrient flow pathways using discharge and nutrient data (Mellander et al., 2012a; Mellander et al., 2013). Monitoring the \( E. coli \) fluxes over various pathways would be extremely beneficial in providing information on the movement of bacteria in water and the possible source of faecal contamination. It might also be possible to add MST methods, such as host-specific \textit{Bacteroidales} assays to this analysis, which would not only determine the presence of faecal contamination but would also differentiate between human and agricultural ruminant faecal pollution and suggest the pathways that the sources are dominant in. As some studies have shown that the bovine-specific \textit{Bacteroidales} markers demonstrate cross-amplification with sheep and goat, it would be of great benefit to have an assay that is specific to sheep (Bernhard and Field, 2000b; Dorai-Raj et al., 2009; Reischer et al., 2013).

MST has the potential to be a very useful tool for achieving the “good” quality water status established by the European Union Water Framework Directive. Using FIO and MST in conjunction with nutrient monitoring such as N or P in a catchment would also help to obtain a clearer picture of water quality. If these research needs were to be met it would be easier to establish the source of faecal pollution and take a more targeted approach to best management practises. If, for example, the source was established as human contamination, which possibly leaked out from a defective septic tank, the necessary maintenance work could be performed to upgrade the septic tank. If the source was found to be agricultural pollution then a sheep-specific assay could be used to establish further differentiation in determining whether the
source is sheep or cow. This would make it easier to propose suitable management practises, such as mitigation measures to be targeted at farms with cows or sheep. The use of this multi-tiered approach together with an integrated catchment science approach would allow for successful implementation of the regulations set out by the EU Water Framework Directive, Bathing Water directive and Water Safety Plans, and effective protection of water quality.

1.9.2 Statement of Project Aims
The main aim of this project was to establish the source of faecal contamination and the pathways by which it was transferred in two rural catchments (Fig 1.3a-1.3d), one karst (containing turloughs; Fig. 1.3a) and one river, in Ireland using a multi-tiered approach including elements such as MST. The aim was undertaken using the following objectives:

1. Water sampling was conducted using an autosampler targeting high flow events and low flow periods during spring/summer and autumn/winter (Fig. 1.3b1)
2. Water quality was assessed using FIO such as total coliforms and *E. coli*
3. Continuous monitoring of chemical parameters and hydrology was achieved using bank-side auto-analysers (Fig. 1.3b2)
4. The source of faecal contamination was determined using host-specific *Bacteroidales* assays
5. A new technique based on hydrograph separation and “end-member mixing analysis” was used to separate nutrient flow pathways into quick flow, interflow and baseflow
6. This analysis was then used to define *E. coli* and faecal pollution fluxes across the various pathways
7. A novel sheep-specific *Bacteroidales* molecular assay was designed using subtractive hybridisation

Chapter II discusses the use of Loadograph Recession Analysis (LRA) in separating hydrological pathways using discharge and nutrient data such as TP. The analysis was conducted on the spring emergence in the karst catchment (Fig 1.3b) and on the outlet in the river catchment (Fig 1.3c & d). *E. coli* data were converted into loads
and plotted onto the LRA in order to determine the flux and magnitude of *E. coli* across high flow events and low flow periods in various seasons.

Chapter III builds on the information reviewed in chapter II. MST data were added to the LRA in order to differentiate the source of faecal contamination between human and agricultural pollution and during different flow phases. The LRA was also used to determine which source was dominant during which pathway, be it quick flow, interflow or baseflow. This significant information could be used for establishing best management practices in catchments.

Chapter IV addresses the issue of bovine-specific *Bacteroidales* demonstrating cross-amplification between sheep and cow. Subtractive hybridisation is an important technique for establishing specific DNA sequences. Target sheep-specific faecal *Bacteroidales* rRNA gene fragments were differentiated from diverging, though closely related subtracter cow and human faecal *Bacteroidales* sources. The resultant sequences were used to design sheep-specific PCR and qPCR assays.

Chapter V discusses the concluding remarks and makes suggestions for future work.
Figure 1.3 Karst catchment in the west of Ireland (a & b) (Grassland D) and river catchment in the east of Ireland (c & d) (Arable B). a: Turlough quite full after prolonged rain during autumn/winter. b: Autosampler (1) taking and auto bank-analysers evaluating (2) water at spring emergence during autumn/winter. c: Water level high in outlet after heavy rainfall event in autumn/winter. d: Field adjacent to outlet showing proximity of cattle to water during spring/summer.
1.10 References


Chapter I


fluorescent in situ hybridization with group-specific 16S rRNA-targeted oligonucleotide probes. *Applied and Environmental Microbiology, 64*(9), 3336-3345.


50


Chapter I


Paruch, A.M., and Mæhlum, T. (2012). Specific features of *Escherichia coli* that distinguish it from coliform and thermotolerant coliform bacteria and define it as the most accurate indicator of faecal contamination in the environment. *Ecological Indicators, 23*, 140-142.


S.I. No. 272. (2009). European Communities Environmental Objectives (surface waters) regulations

S. I. No. 122. (2014). European Union (Drinking Water) Regulations


Chapter I


Chapter I


Chapter I


www.teagasc.ie/agcatchments/ Accessed on 25.08.2014


Chapter II

Quantifying Faecal Indicator Organism Transfer Pathways in Agricultural Catchments using Loadograph Recession Analysis

Accepted for publication in Science of the Total Environment
Chapter II

Author Contribution

Sínead Murphy
(Microbial Ecology Lab, Microbiology Dept., NUI Galway, Galway)
Sample collection and processing, results analysis, script writing

Per-Erik Mellander
(Agricultural Catchments Programme, Teagasc, Johnstown Castle, Wexford)
Loadograph recession analysis

Phil Jordan
(Co-supervisor; Agricultural Catchments Programme, Teagasc, Johnstown Castle, Wexford)
PhD supervision and guidance

Vincent O’Flaherty
(Supervisor; P.I. of Microbial Ecology Lab, Microbiology Dept., NUI Galway, Galway)
PhD supervision and guidance
2.1 Introduction

Faecal pollution of freshwaters and drinking water can arise from various waste water and agricultural sources, negatively affecting the quality of the water and potentially impacting on human health (Allevi et al., 2013; Frey et al., 2013; Marti et al., 2013; Walters et al., 2013). The World Health Organisation (WHO) states that approximately 780 million people in the world drink unsafe water (WHO and UNICEF, 2012), which in the context of inadequate sanitation and poor hygiene, causes diarrhoeal disease cases leading to 1.2 million deaths annually (Mattioli et al., 2012). Escherichia coli (E. coli), which is found in high numbers \((10^9 \text{ g}^{-1})\) in the faeces of warm blooded animals, is a highly prevalent etiological agent of diarrhoea in the developing world (Edberg et al., 2000; Levine et al., 2012). Faecal indicator organisms (FIOs), such as E. coli, are used to confirm and monitor this contamination from exogenic sources (i.e. originating from catchments). While incidences of serious disease outbreaks due to faecal contamination of water are low in developed countries, there is enough concern for strict management protocols for protection of drinking water and bathing water resources to be implemented (Kallis and Butler, 2001; Bartram and Cairncross, 2010).

Agricultural sources of FIOs include incidental point discharges from, for example, farmyards, milking parlours and slurry/manure storage facilities, where faecal matter is not retained within the managed facility (Collins and Rutherford, 2004; Edwards et al., 2008). Faecal runoff from fields, where manure has been applied, is considered a diffuse incidental source, where the loss to water is via convergent, hydrological flowpaths following heavy rainfall (Preedy et al., 2001). Furthermore, if farm animals graze close to water courses, water quality can be impacted due to direct defecation (Fisher et al., 2000; Eyles et al., 2003). As sheep and cattle produce more FIOs per head and can be present in larger numbers than humans in rural areas, the potential for point and diffuse incidental losses from land to water and direct defecation of faecal matter is higher than losses from the human population (Kay et al., 2007). Unsewered rural houses with rudimentary, or improperly sited, septic tank systems can also,
however, make a contribution to FIO loading of streams and groundwater (Arnscheidt et al., 2007; Crowther et al., 2002; Rodgers et al., 2003).

Point and diffuse sources of faecal pollution can also transfer nutrients, such as phosphorus (P) and nitrogen (N), a surplus of which can cause eutrophication when transferred from land to water in surface and sub-surface runoff (Bowes, et al., 2003; Gburek et al., 2005; Mellander et al., 2012a). Septic tank systems can cause low level, but continuous, P as well as FIO contributions in agricultural catchments (Brownlie et al., 2014; Withers et al., 2014). A high proportion of such P inputs can be in a dissolved and highly bioavailable form (Edwards et al., 2000; Withers and Jarvie, 2008; Withers, et al., 2011).

Hydrological connectivity plays an important role in the movement of incidental FIOs from the land into water. Studies in the UK and New Zealand on high river flows, for example, showed an order of magnitude increase in both volume of flow and FIO concentration, which caused a two order of magnitude increase in FIO delivery to waterbodies, when compared to low flow periods (Muirhead et al., 2004; Kay et al., 2010). Other studies have, however, shown that in-channel sources can occasionally produce FIO concentration peaks that resemble FIO peaks seen during rainfall-runoff events, due to entrainment of organisms from the streambed sediments (McDonald and Kay, 1981; Jenkins et al., 1984; Wilkinson et al., 1995). Catchment FIO fluxes commonly peak in summer when livestock are in the fields and not being housed, and when slurry amendments coincide with summer rainfall, although the rate of bacterial die-off is generally higher in summer than in winter because of increased UV light and higher temperatures (Rodgers et al., 2003). There are fewer studies of winter FIO flux dynamics and how they impact contamination rates (Kay et al., 2008; Tetzlaff et al., 2012). Moreover, the specific flowpath dynamics of FIOs during different runoff phases and flow regimes in rivers has received little attention. This omission is important as flowpath vectors of catchment sourced pollution are relevant to FIO monitoring in general and to the specific type of mitigation measures that might be effective (e.g. Mellander et al., 2012a).
Chapter II

Hydrograph separation (Dingman, 2002) has been used to determine hydrological runoff pathways. These pathway sources have been measured further by End Member Mixing Analysis (EMMA), which links observed hillslope water chemistry to observed stream water chemistry (Hooper et al., 1990; Soulsby et al., 2003). Using EMMA, research conducted in a catchment in the UK, for example, showed that the hydrograph was usually dominated volumetrically by subsurface flow from agricultural drains; field surfaces were contributing to the overland flow that generally dominated the hydrograph peak and baseflows were dominated by groundwater which contributed more to the hydrograph during most storm peaks (Soulsby et al., 2003). This principle has also been used to establish and quantify nutrient loads of P and N in hydrological pathways, using high resolution river discharge and nutrient concentration data (Mellander et al., 2012a; Mellander et al., 2013).

In the European Union (EU), the Water Framework Directive (WFD) calls for a move from traditional point-source effluent quality management to a catchment wide approach, involving integrated river basin districts and, where water is used for ecosystem maintenance, water supply, recreation and fisheries (OJEU, 2006). This requires point and diffuse sources of pollution to be studied in unison in order to establish good water quality status by 2015, or by specified target dates thereafter. Chemical and/or microbial tracer studies, and chemical, biochemical and molecular faecal pollution source tracking techniques are key elements of a toolbox to address such catchment scale challenges (Eyles et al., 2003). The need for targeted studies of high-flow events that give more precise information on FIO concentration and fluxes has been recognised in previous work (McCarthy et al., 2008; Kay et al., 2010). It has also been recognised that more empirical data are needed on transfer pathways and total losses (Edwards, et al., 2012) to support the effective implementation of best management practices.

Given this background, and with an emphasis on aiming to add to the understanding of FIO flowpath dynamics in river systems, the objectives of this research were to: (i) define the flowpath magnitude of FIO transfer using an associated flowpath separation technique in diverse catchments; and (ii) study FIO concentrations and
fluxes during high- and low-flow events, in order to better understand diffuse and point source influences. The study was based in two contrasting catchments in Ireland that differed in hydrology, topography, soils, settlement and agricultural land-use and intensity. Both, however, were deemed at risk, \textit{a priori}, from FIO transfers based on combinations of source and pathway factors.
2.2. Methods

This study was carried out in two experimental catchments, as previously described in detail elsewhere (Fealy et al., 2010; Melland et al., 2012; Mellander et al., 2012; Fig. 2.1). A brief summary is provided here.

2.2.1 Catchment Sites

Grassland D is a karst grassland catchment (zone of contribution) in west-central Ireland, covering approximately 30 km$^2$. The sub-surface is a karstic aquifer of pure Carboniferous limestone and is covered by relatively thin glacial till deposits (0-5 m) with thinner soils toward the west of the catchment, often exposing areas of pavement and epi-karst. The watershed has over 1,800 mapped karstic features, mostly dolines (solution depressions) but also springs, and turlough areas (seasonal ephemeral lakes) (Mellander et al., 2012b). The prevalent soils are shallow Brown Earths (Cambisols) and Rendzinas (Leptosols) with Gleys (Gleysols) and Peats (Histosols) in areas with till or areas which are affected by seasonal standing water. Conduit flow defines the flow regime, with travel times varying from 10 to 123 m h$^{-1}$. These flows emerge as springs into the nearby River Robe and on into Lough Mask, a large lake (83km$^2$ surface area) in Co. Mayo (Fig. 1).

The main spring emergence was used as a monitoring point in this study. There is an annual average rainfall of 1203 mm (the thirty year 1981-2010 mean; Met Éireann, Irish Meteorological Service). The catchment has an approximate housing density of 7.8 houses per km$^2$, which is low-moderate for Irish rural settings (Arnscheidt et al., 2007). Land use is dominated by sheep and beef farming on permanent pasture (92%; grazed for 7-8 months). The average stocking rate is equivalent to an annual input of ca. 120 kg Organic N ha$^{-1}$, which is also equivalent to 17 kg Organic P ha$^{-1}$ (Mellander et al., 2013) and there are 90 farms. Animals are wintered indoors and manure storage follows the guidelines of the Nitrates Directive National Action Programme (NAP; S.I. No. 31, 2014). The risk of FIO transfer in this catchment (zone of contribution) has an emphasis on the agricultural source and the links between the surface and the emergent spring via the conduit and fissure network, as similarly determined in other work (Johnson et al., 2011; Sinreich et al., 2013). This is
especially important in this catchment as the emergence is used as a community water supply following treatment.

Figure 2.1 Map of Arable B catchment (a) showing outlet (blue inverted triangle), Met station (red box), rain gauge (red triangle); map of Grassland D catchment (b) showing emergent spring (blue inverted triangle), Met station (red box), rain gauge (red triangle). Map of Ireland showing where each of the two catchments are located (c). (Printed under license number 6155 from the Ordnance Survey Ireland).
Chapter II

Arable B is a 9.5 km² mixed arable catchment and is located in east-central Ireland. The catchment area consists of poorly to moderately drained, grey brown podzolic brown earths and gley soils with clay loams found more so in the subsoil. Ordovician-Silurian calcareous greywacke and banded mudstone of the Salterstown formation makes up the bedrock, which is a poorly productive aquifer. Arable B is deemed to be dominated by surface flow pathways but, in areas with permeable soil, there is the possibility of groundwater flow pathways (Mellander et al., 2012a). These flow pathways combine to form first and second order streams, and the 9.5 km² outlet point in the east of the catchment was used as a monitoring site in this study (Fig. 1). This river is a tributary of the White River, which eventually joins the Irish sea off Co. Louth. The mean annual rainfall for the area is 758 mm (1981-2010 mean, Met Éireann).

There are 35 farms in this catchment with arable crop production of winter wheat and barley being a notable land-use (24%). There is also beef and sheep grazing (29%) and grassland based dairy farming (19%). The average stocking rate is equivalent to an annual input of ca. 96 kg Organic N ha⁻¹ (Wall et al., 2012) equivalent to 14 kg Organic P ha⁻¹ (Jordan et al., 2012). Cattle are wintered indoors over a 3-5 month period but a small sheep flock grazes outdoors over the year. Overall nutrient inputs are likely to be higher than in Grassland D due to inputs of artificial fertilisers onto cropping systems; however, reporting the organic fraction of nutrients as a metric of intensity pressure is of importance here, as these are likely to be more related to FIO. Housing density is approximately 14 per km² as determined from national statistics (Jordan et al., 2012) and this is moderately high for Irish catchments (Macintosh et al., 2011). The risk of FIO transfer in this catchment was considered from combinations of multiple sources and poor-moderately drained soils (Melland et al., 2012).

2.2.2 Field Monitoring and Water Analysis

Hydrometric and water quality stations, which measured sub-hourly parameters on a near-continuous basis, were installed at the spring in Grassland D and the outlet in Arable B (Melland et al., 2012; Mellander et al., 2013; Fig. 1). Total P (TP) and total
reactive P concentrations (TRP; 0.010 to 5.000 mg l\(^{-1}\)) were measured by Hach-Lange Phosphax Sigma systems. Hach-Lange Nitratax SC-Plus probes were also used to measure nitrate-N as total oxidised nitrogen (TON; 0.1 to 50.0 mg l\(^{-1}\)) and Hach-Lange Solitax probes measured turbidity (0 to 1000 NTU).

Sub-hourly rainfall and standard meteorological parameters were measured by Campbell Scientific (BWS-200) weather stations in each catchment. In Arable B, a Corbett flat-v non-standard weir was used to control flow at the stream outlet and measure discharge by applying a rating curve (in WISKI-SKED; Kisters, 2011) to continuously measured water levels (OTT Orpheus-mini). Spot discharge measurements were made for the rating curve using the area-velocity method with an OTT Acoustic Doppler Flowmeter and an OTT C31 impellor flowmeter. In Grassland D, a Thermo-Fisher time-of-flight area-velocity meter measured discharge at the spring emergence in a specially engineered cross-section operated by the Irish Environmental Protection Agency. The WISKI-7 (Kisters, 2011) database management system was used for final quality control of all hydrochemistry and hydrometeorological data. As part of this analysis, sub-hourly data were collated to hourly average synchronous water quality and discharge parameters and total rainfall measurements.

A series of high and low flow events were sampled over time to investigate the relationship between bacterial transfer hydro-dynamics and continuously measured water quality parameters. Autosamplers (Hach Sigma 900 Max Portable Sampler and ISCO 6712 Portable Sampler) were deployed at both hydrometric sites to take a 1 L sample every 2 hours over a 36-96 hour time period when deployed manually. Samples were stored in the autosampler with ice packs for a maximum time period of 52 hours, from first sample, before analysis. This follows a standard protocol that shows no significant difference in *E. coli* density between samples analysed at 0 h and 48 h when samples are stored at less than 10°C (Pope *et al*., 2003). Samples were analysed for coliforms and *E. coli* occurrence by the IDEXX Colisure® Quanti Tray®/2000 system. Briefly, Colisure reagent (IDEXX laboratories, ME, USA) was aseptically added to 100 ml of water sample and mixed to dissolve. The solution was
poured into a 97-well quanti-tray, sealed and incubated at 37°C for 24 hours. Red/magenta wells indicated detection of total coliforms and red/magenta wells that fluoresce under UV light indicated detection of *E. coli*. Ambiguously coloured wells (pink/orange) were further incubated for up to 48 h and re-tested. Faecal Indicator Organism most probable number (MPN) was calculated using the MPN table provided by IDEXX.

2.2.3 Statistical Analysis
GraphPad Prism (Version 5.00 for Windows, GraphPad Software, San Diego California, USA, www.graphpad.com) was used to apply Pearson correlations between microbial and chemical parameters from both catchments in order to determine if correlations were statistically significant. P-values of P<0.05 were all considered statistically significant.

2.2.4 Calculation for *E. coli* load
To compare nutrient and FIO as the same hourly loads, the *E. coli* concentration data were corrected to MPN/m³, multiplied by mean hourly discharge (m³/hr) and reported as *E. coli* hourly load (MPN/hr), or normalised as hourly catchment loads (MPN/km²). Event median loads were compared during sample periods in both catchments to highlight differences in magnitude and load trajectories.

2.2.5 Loadograph Recession Analysis
Loadograph Recession Analysis, a method developed by Mellander *et al.* (2012a) working in Arable B and then applied to storm events in Grassland D (Mellander *et al.*, 2013), was used to separate the major flow pathways using nutrient loads. This method is similar to hydrograph separation (Dingman, 2002), which uses water discharge to graphically separate an event. The release of TP loads from various sources and their transfer via the different dominant flow pathways was estimated to be represented by the separated segments in the Loadograph Recession Analyses; made possible by the high resolution and synchronous TP and flow data. The separation of transfer pathways was validated by increased turbidity in the spring (Grassland D) or outlet (Arable B; data not shown). In this study, conduit and fissure
nutrient loads (ranging from storm to baseflow conditions) were separated in Grassland D and the proportions of quickflow, interflow and delayed flow (also ranging from storm to baseflow conditions) were separated in Arable B.
2.3 Results

Water discharge and quality data were collated over the study period and are shown as time series in Figure 2.2a and 2.2b, together with the time points at which samples for FIO analysis were taken. Table 2.1 provides a summary of the FIO and discharge data.

The FIO dataset indicated generally larger magnitudes of FIO concentrations in samples from Arable B than from Grassland D throughout the study period. For all events, the overall median total coliforms and *E. coli* concentrations were 295.9 MPN/100 ml and 93.5 MPN/100 ml in Grassland D (the highest values of 2419.6 MPN/100 ml for both total coliforms and *E. coli* were noted during the sampling period of 01-03/11/2010) and 2139.6 MPN/100 ml and 1344.6 MPN/100 ml in Arable B, respectively (the highest values of 120980.0 and 19863 MPN/100 ml for total coliforms and *E. coli*, respectively were noted during the sampling period of 23-25/10/2011).

Pearson’s correlations (Table 2.2a and b) between synchronous water quality data indicated that there was a significant positive correlation between *E. coli* and TP concentration in Grassland D (r = 0.32; P<0.01), and significant but lower positive correlations between *E. coli* and TRP and between *E. coli* and turbidity (r = 0.27 and 0.19, respectively; P<0.01). No correlation was found between total coliforms and TON in Grassland D. For Arable B, there was also a significant positive correlation between *E. coli* and TP (r = 0.65; P<0.01), between *E. coli* and TRP (r = 0.33; P<0.01), and between *E. coli* and turbidity (r = 0.71; P<0.01). There were significantly negative correlations between the P fractions and TON.
Figure 2.2a Time series charts of the parameters measured at high resolution (collated to hourly data) in Grassland D. TP – black line; TRP – grey line and dates of sample events (hollow black circles).
Figure 2.2b Time series charts of the parameters measured at high resolution (collated to hourly data) in Arable B. TP – black line; TRP – grey line and dates of sample events (hollow black circles).
Table 2.1 Data summaries of sampled event discharge and bacteria data from the karst grassland (Grassland D) and mixed arable (Arable B) catchments. Dates in bold represent those events used for further LRA analysis.

<table>
<thead>
<tr>
<th>Catchment and Event Date</th>
<th>Discharge, m³/s</th>
<th>Total Coliforms, MPN/100 ml</th>
<th>E. coli, MPN/100 ml</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Min</td>
<td>Max</td>
<td>Median</td>
</tr>
<tr>
<td><strong>Grassland D</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14-18.07.2010</td>
<td>0.005</td>
<td>0.045</td>
<td>0.037</td>
</tr>
<tr>
<td>28-30.09.2010</td>
<td>0.084</td>
<td>0.100</td>
<td>0.093</td>
</tr>
<tr>
<td>26-28.10.2010</td>
<td>0.102</td>
<td>0.177</td>
<td>0.162</td>
</tr>
<tr>
<td>01-03.11.2010</td>
<td>0.212</td>
<td>0.355</td>
<td>0.307</td>
</tr>
<tr>
<td><strong>28-30.04.2011</strong></td>
<td>0.113</td>
<td>0.142</td>
<td>0.127</td>
</tr>
<tr>
<td>15-17.07.2011</td>
<td>0.001</td>
<td>0.018</td>
<td>0.008</td>
</tr>
<tr>
<td><strong>26-28.10.2011</strong></td>
<td>0.279</td>
<td>0.319</td>
<td>0.300</td>
</tr>
<tr>
<td>23-25.01.2012</td>
<td>0.190</td>
<td>0.196</td>
<td>0.192</td>
</tr>
<tr>
<td><strong>16-18.04.2012</strong></td>
<td>0.104</td>
<td>0.144</td>
<td>0.130</td>
</tr>
<tr>
<td>27-29.07.2012</td>
<td>0.084</td>
<td>0.102</td>
<td>0.093</td>
</tr>
<tr>
<td><strong>Arable B</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>04-06.11.2010</td>
<td>0.326</td>
<td>1.373</td>
<td>0.482</td>
</tr>
<tr>
<td>23-25.04.2011</td>
<td>0.011</td>
<td>0.013</td>
<td>0.012</td>
</tr>
<tr>
<td><strong>07-09.05.2011</strong></td>
<td>0.009</td>
<td>0.034</td>
<td>0.011</td>
</tr>
<tr>
<td>15-17.07.2011</td>
<td>0.004</td>
<td>0.007</td>
<td>0.004</td>
</tr>
<tr>
<td><strong>23-25.10.2011</strong></td>
<td>0.052</td>
<td>11.122</td>
<td>0.636</td>
</tr>
<tr>
<td>26-28.01.2012</td>
<td>0.134</td>
<td>0.201</td>
<td>0.154</td>
</tr>
<tr>
<td>21-23.05.2012</td>
<td>0.048</td>
<td>0.058</td>
<td>0.052</td>
</tr>
<tr>
<td><strong>26-28.07.2012</strong></td>
<td>0.114</td>
<td>0.153</td>
<td>0.129</td>
</tr>
</tbody>
</table>
Table 2.2 Pearson’s Correlations of all measured parameters for a. Grassland D and b. Arable B. P-value of P<0.05 = 1 symbol, P<0.01 = 2 symbols; N.C. = no correlation.

Table 2.2a

<table>
<thead>
<tr>
<th></th>
<th>Total coliforms</th>
<th>E. coli</th>
<th>TP</th>
<th>TRP</th>
<th>TON</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total coliforms</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E. coli</td>
<td>0.54</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>●●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP</td>
<td>0.32</td>
<td>0.32</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>●●</td>
<td>●●</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRP</td>
<td>0.38</td>
<td>0.27</td>
<td>0.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>●●</td>
<td>●●</td>
<td>●●</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TON</td>
<td>0.13</td>
<td>0.05</td>
<td>0.54</td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td></td>
<td>●</td>
<td>N.C.</td>
<td>●●</td>
<td>●●</td>
<td></td>
</tr>
<tr>
<td>Turbidity</td>
<td>0.09</td>
<td>0.19</td>
<td>0.67</td>
<td>0.35</td>
<td>0.46</td>
</tr>
<tr>
<td></td>
<td>N.C.</td>
<td>●●</td>
<td>●●</td>
<td>●●</td>
<td>●●</td>
</tr>
</tbody>
</table>

Table 2.2b

<table>
<thead>
<tr>
<th></th>
<th>Total coliforms</th>
<th>E. coli</th>
<th>TP</th>
<th>TRP</th>
<th>TON</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total coliforms</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E. coli</td>
<td>0.73</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>●●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP</td>
<td>0.87</td>
<td>0.66</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>●●</td>
<td>●●</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRP</td>
<td>0.42</td>
<td>0.33</td>
<td>-0.66</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>●●</td>
<td>●●</td>
<td>●●</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TON</td>
<td>-0.06</td>
<td>-0.07</td>
<td>-0.34</td>
<td>-0.76</td>
<td></td>
</tr>
<tr>
<td></td>
<td>N.C.</td>
<td>N.C.</td>
<td>●●</td>
<td>●●</td>
<td></td>
</tr>
<tr>
<td>Turbidity</td>
<td>0.89</td>
<td>0.71</td>
<td>0.93</td>
<td>0.41</td>
<td>-0.10</td>
</tr>
<tr>
<td></td>
<td>●●</td>
<td>●●</td>
<td>●●</td>
<td>N.C.</td>
<td></td>
</tr>
</tbody>
</table>

A significant positive correlation between E. coli and TP would be indicative of faecal pollution from agricultural or anthropogenic sources (Arnscheidt et al., 2007; Carrillo et al., 1985; Vinten et al., 2008), due to pollution associated with particulate matter. The significant TP-FIO correlation was noted for both Grassland D and Arable B
(P<0.01). E. coli data were used for further interpretation in the Loadograph Recession Analyses to assess which nutrient pathways were associated with E. coli transfer during each event sampled (total coliform data were not used in further interpretation due to E. coli being more aligned with faecal matter transfer; Edberg et al., 2000).

Comparison of E. coli loads and discharge showed distinctive differences between the two catchments (Fig 2.3). The median values of event loads (product of E. coli concentrations and discharge in Table 2.1 normalised to catchment area) in Grassland D indicated higher loads per unit discharge, indicative of higher E. coli concentrations, on the events dominated by rising periods. Grassland D, falling limb event E. coli loads appeared to be more variable owing to rapid drop off in concentration as discharges decreased. This was more subtle on the Arable B comparisons owing to the flashier nature of runoff from this catchment and the shorter rising hydrograph times (Melland et al., 2012), leading also to less event data gathered. Nevertheless, E. coli loads were higher in Arable B per unit discharge and the events in both catchments (i.e. each data-point on Figure 2.3 is an individual event) appeared to be related to discharge rate as well as phase.
Figure 2.3 *E. coli* catchment load and discharge per event. Circles are Arable B and squares are Grassland D. Closed symbols are events dominated by rising phases, open symbols are events dominated by falling phases. The power trend-line describes the overall *E. coli* load-discharge relationship for Arable B.
Three of the events were possible to analyse in this manner in Grassland D and six events were analysed in Arable B (dates shown in bold in Table 2.1). Events that were omitted from the analysis were either too small and/or with unclear separations. In Grassland D (where events sampled were separated into conduit, large fissure, medium and/or small fissure flow) conduit flow was determined to be represented by the first quick nutrient flow pathway and large fissure flow was determined to be represented by the second quick nutrient flow pathway. The delayed flow pathway represented medium sized fissures and (occasionally) small sized fissures depending on the event. In Arable B, the TP transfer pathways were separated into quick flow, making up the peak of the hydrograph, followed by interflow and delayed flow. Quick flow was assumed to be represented by overland flow, preferential flow, and tile and ditch drainage. The delayed flow pathway was assumed to include baseflow.

The *E. coli* loads were transposed onto the TP time series for comparison (Figs 2.4 to 2.6) using time series adjoining the particular sampling period. In Grassland D, Figure 2.4 shows that the FIO sampling in the first separated event during April and May 2011 took place during small fissure flow following a period of conduit and large to medium fissure flow. The peak *E. coli* loading during the small fissure flow was $2.21 \times 10^7$ MPN/hr. Figure 2.5 shows that the FIO sampling period in October 2011 took place during medium to small fissure flows, again following a period of conduit and large fissure flows. The peak *E. coli* loading was $3.29 \times 10^9$ MPN/hr, over two orders of magnitude higher than the peaks in the first sample in discharges that were approximately two to three times higher than the first sampled event. Similar peak *E. coli* loading was recorded in the conduit flow of the third sampled event in April 2012, shown in Figure 2.6, with a peak of $2.96 \times 10^9$ MPN/hr, but with discharges similar in magnitude (approximately 0.1 m$^3$/s) to the first sampled event (Fig 2.4).
Figure 2.4 Grassland D loadographs showing pathway separations for events I. a: Time series of hourly discharge (blue line) and turbidity (purple line) showing pathway separation; b: Time series of hourly TP loads (black line), and E. coli loads (green line with closed symbols).
Figure 2.5 Grassland D loadographs showing pathway separations for events II. a: Time series of hourly discharge (blue line) and turbidity (purple line) showing pathway separation; b: Time series of hourly TP loads (black line), and *E. coli* loads (green line with closed symbols).
Figure 2.6 Grassland D loadographs showing pathway separations for events III. a: Time series of hourly discharge (blue line) and turbidity (purple line) showing pathway separation; b: Time series of hourly TP loads (black line), and *E. coli* loads (green line with closed symbols).
In Arable B, the first event studied took place in November 2010 during quick flow, interflow and delayed flow phases (Fig 2.7) and the FIO sampling took place during all three flow phases. *E. coli* loading peaked during the quick flow phase at $1.17 \times 10^{11}$ MPN/hr, subsequent to a discharge peak of $1.373 \text{ m}^3/\text{s}$, and decreased steadily to a median of $5.31 \times 10^9$ MPN/hr by the end of the sampling period, during delayed flow. A smaller storm event period was sampled during May 2011 (Fig 2.8) with samples collected during all three flow phases. The smaller discharges measured during this event (peak 0.034 m$^3$/s) during the quick flow phase were associated with a smaller *E. coli* load, which peaked, again subsequent to the discharge peak, at $1.67 \times 10^{10}$ MPN/hr. The third sampled event in Arable B took place during October 2011 during an intense and prolonged storm (Fig 2.9). All *E. coli* samples were collected during the quick flow part of the event, which gave the highest discharges of the time-series at 11.12 m$^3$/s. In this third event, the *E. coli* loading peaked before the discharge peak, although this discharge peak had two notable intense rainfall periods with an earlier shoulder in the rising limb of the hydrograph and two distinct TP load peaks. The *E. coli* load peak, the highest recorded at $7.31 \times 10^{12}$ MPN/hr, may, therefore, have been related more to the earlier rainfall rather than the subsequent rainfall.

The fourth, fifth and sixth events studied in Arable B were associated with delayed, low flow periods, and following small storm periods in January, May and July 2012, respectively (Figs 2.10, 2.11 and 2.12). Delayed, low *E. coli* loadings (the lowest, on average, of all the Arable B samples measured during the entire study) had peaks of $1.26 \times 10^{10}$ MPN/hr and $3.17 \times 10^{10}$ MPN/hr and medians $2.47 \times 10^9$ MPN/hr and $4.33 \times 10^9$ MPN/hr for the fourth (Fig 2.10) and sixth (Fig 2.12) events, respectively.

Sampling during the fifth Arable B event (Fig 2.11), however, took place during a notable period of turbidity peaks that were independent of rainfall or storm discharges. Peak *E. coli* loading was $6.64 \times 10^9$ MPN/hr and the median over the sample period was $3.70 \times 10^9$ MPN/hr. The outlet monitoring site of Arable B is situated downstream from a periodically used animal ford, which allows cattle access from one field to another and access to drinking water. Anecdotal evidence at this time showed that the ford was in operation and the diurnal signal of turbidity time-series would support this with habitual animal trafficking possibly causing bed
disturbance (although temporary increases in other point source pressures are equally possible).
Figure 2.7 Arable B loadographs showing pathway separations for events I. a: Time series of hourly discharge (blue line) and turbidity (purple line) showing pathway separation; b: Time series of hourly TP loads (black line), and E. coli loads (green line with closed symbols).
Figure 2.8 Arable B loadographs showing pathway separations for events II. A: Time series of hourly discharge (blue line) and turbidity (purple line) showing pathway separation; B: Time series of hourly TP loads (black line), *E. coli* loads (green line with closed symbols), TP loads during sample event (red line).
Figure 2.9 Arable B loadographs showing pathway separations for events III. A: Time series of hourly discharge (blue line) and turbidity (purple line) showing pathway separation; B: Time series of hourly TP loads (black line), and E. coli loads (green line with closed symbols).
Figure 2.10 Arable B loadographs showing pathway separations for events IV. A: Time series of hourly discharge (blue line) and turbidity (purple line) showing pathway separation; B: Time series of hourly TP loads (black line), and *E. coli* loads (green line with closed symbols).
Figure 2.11 Arable B loadographs showing pathway separations for events V. A: Time series of hourly discharge (blue line) and turbidity (purple line) showing pathway separation; B: Time series of hourly TP loads (black line), and *E. coli* loads (green line with closed symbols).
Figure 2.12 Arable B loadographs showing pathway separations for events VI. A: Time series of hourly discharge (blue line) and turbidity (purple line) showing pathway separation; B: Time series of hourly TP loads (black line), and *E. coli* loads (green line with closed symbols).
2.4 Discussion

In the context of *E. coli* source pressure, Grassland D had the greater potential diffuse faecal matter sources with a higher organic nutrient loading; although the point source storm independent pressure (i.e. at low flows) is likely to be higher in Arable B (Melland *et al.*, 2012). However, the magnitude and trajectories of event loads indicated in Fig 2.3 show Arable B having much higher loads per unit discharge, indicative of higher FIO concentrations, and generally much higher magnitudes, indicative of higher discharges – overall suggesting a source limited system. The strong relationship with discharge between events was particularly marked in Arable B with very similar loads per unit discharge in both rising and falling phases of sampled events. The sudden drop in *E. coli* load on the falling phases of discharge in Grassland D, however, was very apparent and the variability is likely as a result of these phases occurring following events of dissimilar peak magnitude. As karst catchments, or zones of contribution, are known to have both short and long response times in terms of rainfall-runoff processes (Labat *et al.*, 2000), this may have implications for routine monitoring, as longer period recessions are likely to have much lower loads and therefore concentrations per unit receding discharge. Nevertheless, the data from Grassland D indicate a more transport limited system for FIO transfers despite greater potential FIO pressures. While this is not always the case for karst systems where direct connections to source pressures can elevate the timing and magnitude of FIO transfers (Butscher *et al.*, 2011), the data from this study firmly points to transport limitation as being of high importance in both catchments (i.e. less transport limitation in Arable B).

The LRA analyses indicated that the majority of high loads were associated with the faster nutrient flow paths in each catchment – conduit and larger fissure flow in Grassland D and the surface flows associated with quick flow in Arable B. Negative correlations with TON in earlier analyses, supports the notion of deeper groundwaters in each catchment having less influence on FIO transfers. As highlighted, other studies have shown karst landscapes as being particularly vulnerable to FIO transfers (Dura *et al.*, 2010; Pronk *et al.*, 2007) and particularly when in close vicinity to specific source pressures (Heinz *et al.*, 2009). While there were significant (in terms
of potential impacts on human health) FIO peaks and loads in the Grassland D catchment, the magnitudes were much smaller than those in the Arable B catchment. Previous work in the Grassland D catchment has noted that, despite high P source pressures from livestock and soil P status, P transfers in the spring emergence were very low (Mellander et al., 2012b). This was explained by a low specific P vulnerability risk in highly buffered clay soils over limestone fissures, despite comparatively (with other the study catchment) thin soils and several hundred karst surface features (Mellander et al., 2013). This subsurface tortuosity over 30km$^2$ is likely to also account for a buffering of FIO transfers between soil surfaces and the spring emergence and which is largely flushed over the several days of discharge in rising limb hydrographs. However, although comparatively lower, the short periods of higher $E.\ coli$ concentrations and loads in the karst catchment were above safe limits of 0 CFU/100 ml (OJEU, 1998), and thus presented a risk to human health if drinking water was untreated or undertreated; vigilance here is, therefore, especially recommended on rising phases of discharge.

In Arable B, the phases of sample events followed a power-relationship (trendline shown in Fig 2.3), mirrored to some extent by the rising phases of the Grassland D sampled events (trendline not shown). However, the close association between rising and falling phases of $E.\ coli$ loads in Arable B and the asymptotic decline to low flows and loads indicates sustained high hourly low flow loads in the order 1 x 10$^6$ MPN/km$^2$ which equate to an approximate average 800 MPN/100ml concentration at discharges less than 4m$^3$/hr. While recorded low discharges were only rarely observed at this rate, $E.\ coli$ concentrations during delayed flow events were typically high and at least in the 1 x 10$^2$ or 1 x 10$^3$ MPN/100ml range (Table 2.1). Mellander et al. (2012a) indicated that P in delayed flow in Arable B was less related to actual concentrations in groundwater as recorded in a series of monitoring wells, and more related to surface pollution of delayed (groundwater) conditions in the main stream. Jordan et al. (2012) and Mellander et al. (2012a) speculated that (at least some) of this ‘chronic’ P pollution during low flows was due to rural point sources and possibly from non-agricultural sources, as found elsewhere (Withers et al. 2014). The $E.\ coli$ loads, although much smaller than quick flow events, appear to follow a similar pattern in Arable B, with sustained concentrations high enough to pose a human
health issue if used for drinking water and, more likely, an animal health issue if used as a supply for drinking troughs or via direct stream access (Beede, 2006; Tetzlaff et al., 2012). The increase in *E. coli* loads during what may have been an animal trafficking period (Davies-Colley et al., 2004) were, overall, small when compared with those related to high flow events but, for a short period, increased the concentration above what was already a high concentration. The high *E. coli* loads, especially during high, quick flow events, could also impact on the adjacent coastal area if occurring during the June to September bathing season (OJEU, 2006).

It is widely accepted that in many rural catchments, the source of high event flow *E. coli* transfers is of animal origin as a direct transfer from recently applied manures or from residual applications where the *E. coli* population has remained (Oliver et al., 2012). From this study, in the series of events studied, it is clear that these are dependent on the quicker flow events, in either catchment. For forward assessment and mitigation, therefore, a next useful stage will be to know how the source predominance of FIO transfers change during different parts of hydrological flow events. For example, if the correlation with P in the hydrological data is also correlated with the different P sources, then the high *E. coli* concentrations observed at lower delayed flows (and likely to be longer duration) in Arable B may well be from different sources than the (higher) *E. coli* concentrations and loads observed during quick flows (of shorter duration). Other studies have used source apportionment of FIOs to define their sources either by proxy markers to indicate presence or absence of human sewage such as chemical markers (e.g. caffeine - Buergeet al., 2003) or more directly using genetic fingerprinting (e.g. host-specific Bacteroidales molecular assays - Bernhard and Field, 2000; Kildare et al., 2007). Using these techniques during different events, especially between quick flow and delayed flow in catchments similar to Arable B and more specifically during rising and falling phases in catchments such as Grassland D, will add to the understanding and mitigation of FIO transfers from land to water. A more targeted approach, such as this, to best management practices could limit the deterioration of water quality in the most cost effective way (D'Arcy and Frost, 2001; O'Brien et al., 2013).
2.5 Conclusions

This study targeted high- and low-flow events in order to provide understanding on FIO concentration and flowpath dynamics across various transfer pathways using Loadograph Recession Analysis and whole event comparisons. Important findings in the two contrasting karst grassland and mixed arable catchments were that:

- The karst grassland catchment, although having a comparatively higher FIO source pressure, in terms of animal loading, indicated buffering of *E. coli* between the soil surface and the emergent catchment spring and so was transport limited. Hence, like P, the specific vulnerability of this catchment for FIO transfers was low and due to hydrological response and not source pressure. The mixed arable catchment was source limited and a less transport limited system for *E. coli* transfer with higher *E. coli* loads across all flow ranges.

- The karst grassland *E. coli* loads were higher during the periods when hydrological events were rising and, despite being comparatively low in terms of load, would pose a human health risk if this drinking water emergence was not treated, or was under treated.

- The Loadograph Recession Analysis illustrated that the high loads were transferred during the faster flow paths associated with nutrient (TP) transfers i.e. surface runoff and preferential flow in the mixed arable, and conduit and large fissure flow in karst grassland catchment, and were dependent on flow rates.

- An asymptotic *E. coli* load decrease to low flows was evident in the mixed arable catchment indicating sustained FIO pollution at low flows – possibly due to point sources which could be of non-agricultural origin.

The methods and results of this study indicate that, in these catchment systems at least, FIO risks associated with transport limitation are more potent pressures compared with source limitation (i.e. if the transport limitation is low, the risk of FIO transfer is higher); both within hydrological events and between catchments. However, further work on discriminating *E. coli* sources in both rising and falling
phases, within specific flowpaths and during asymptotic low FIO load periods will help to refine risk assessments and mitigation efforts and monitoring.
2.6 References


Chapter II


Chapter II


Water Framework Directive 2006/7/EC.

Chapter II


Chapter III

Tracking the Sources and Transport Pathways of Faecal Contamination in Catchments during Dynamic Hydrological Phases

Submitted to Water Research
Chapter III

Author Contribution

Sinéad Murphy
(Microbial Ecology Lab, Microbiology Dept., NUI Galway, Galway
Department of Civil and Environmental Engineering, UC Davis, California)
Sample collection and processing, results analysis, script writing

Minji Kim
(Department of Civil and Environmental Engineering, UC Davis, California)
Laboratory supervision

Per-Erik Mellander
(Agricultural Catchments Programme, Teagasc, Johnstown Castle, Wexford)
Loadograph recession analysis

Phil Jordan
(Co-supervisor; Agricultural Catchments Programme, Teagasc, Johnstown Castle, Wexford)
PhD supervision and guidance

Stefan Wuertz
(Department of Civil and Environmental Engineering, UC Davis, California)
Project supervision and guidance

Vincent O’Flaherty
(Supervisor; P.I. of Microbial Ecology Lab, Microbiology Dept., NUI Galway, Galway)
PhD supervision and guidance
3.1 Introduction

There is a widespread problem with contamination of water by faecal matter from point sources, such as wastewater treatment plants; and from diffuse sources such as agriculture, residential areas, and industry (Lee et al., 2014; Walters et al., 2011; Wilcock et al., 2011). These sources also contain nutrients, such as phosphorus (P) and nitrogen (N) (Mellander et al., 2012; Withers et al., 2011). This is a concern as contaminated drinking water can cause major human health issues (Nnane et al., 2011) and can also impact on the local environment by causing eutrophication (Newton et al., 2013; Withers et al., 2011). It is, therefore, imperative that the source of faecal pollution is identified. Faecal indicator organisms (FIO) can be used for this purpose but must meet some of the following criteria: they should be universally inhabitant in the intestine of warm-blooded animals; they should not be able to multiply outside of the host; and they should have prominent association with pathogens (Ahmed et al., 2008; Prüss, 1998).

*Escherichia coli* (*E. coli*) has been well established as a FIO as a result of its association with faeces (Layton et al., 2006). It has been used by regulatory bodies for many years to identify faecal contamination in water, as the methods used to culture it are relatively cheap and simple, especially with the advent of defined substrate technology in the 1980s (Edberg et al., 2000; Layton et al., 2006; Noble and Weisberg, 2005). It has since been demonstrated, however, that some conventional FIO, such as *E. coli*, can replicate outside of the host, and come from non-faecal sources, such as soil, resulting in false-positives for faecal contamination (Ahmed et al., 2008; Brennan et al., 2010). Additional methods, which aim to discriminate faecal from non-faecal sources of contamination, such as microbial source tracking (MST), are therefore potentially important tools for water quality monitoring and management (Fogarty and Voytek, 2005).

MST can be used in conjunction with *E. coli* analysis; if high levels of *E. coli* are found in water, MST can provide additional information by confirming whether faecal pollution is present and defining the source of this contamination (Boehm et al., 2010; Gourmelon et al., 2007; Walters et al., 2007). MST includes molecular
techniques to detect host-specific bacteria such as the faecal anaerobe group, *Bacteroidales* (Ahmed *et al*., 2008; Bernhard and Field, 2000; Kildare *et al*., 2007). As these strict anaerobes display a high level of host specificity, they can be used to distinguish between anthropogenic and agricultural animal faecal pollution, such as that arising from ruminants (Bernhard and Field, 2000; Kildare *et al*., 2007; Schriewer *et al*., 2010). Polymerase chain reaction (PCR) and quantitative PCR (qPCR) allow for rapid and highly specific detection of the *Bacteroidales* group from water samples (Bernhard and Field, 2000; Kildare *et al*., 2007).

A number of studies have demonstrated the successful application of *Bacteroidales* as a means of detecting and defining faecal source contamination in several countries, including Ireland (Dorai-Raj *et al*., 2009); the UK (Stapleton *et al*., 2009); several other European countries (Gawler *et al*., 2007); Kenya (Jenkins *et al*., 2009); and the USA (Dick and Field, 2004; Kildare *et al*., 2007; Layton *et al*., 2006). Large recent studies that took place across several continents, also demonstrated that *Bacteroidales* are a good indicator of faecal source pollution (Layton *et al*., 2013; Reischer *et al*., 2013).

To help establish further sources of point and diffuse faecal pollution, it is important to develop a better understanding of nutrient transfer pathways (Mellander *et al*., 2013). Loadograph recession analysis (LRA), a technique established by Mellander *et al*. (2012), is based on hydrograph separation (Dingman, 2002) and End-Member Mixing Analysis (EMMA) (Hooper *et al*., 1990; Soulsby *et al*., 2003) but uses additional high-resolution nutrient data. It separates the pathways based on nutrient load, with each segment representing the release of the nutrient load into the flow pathway by which it is transferred (Mellander *et al*., 2012). The pathways are broken down into quick flow, interflow and delayed flow.

Dry periods, or rainfall events of varying intensity, from drizzle to heavy rainfall, can impact on the transfer of nutrients across the various pathways. For example, a storm in a flashy catchment (one that responds quickly to rainfall) could result in the majority of P being transferred via quick flow (Jordan *et al*., 2007). Chronic P contamination of rivers and streams during baseflow in rural catchments could be
caused either by ill-managed septic tanks systems (STS) or be a result of leakages from a poorly managed farmyard (Arnscheidt et al., 2007; Edwards et al., 2008).

Different regions and nations have specified management protocols to preserve water quality. For example, as part of the European Union’s (EU) Water Framework Directive (WFD), a “good” water quality status must be achieved by 2015 (WFD 2006/7/EC) or by defined six year cycles afterwards. In order to help achieve this, the EU has specified the need to take a multi-disciplinary approach to better understand catchment science, in order to address the issues associated with water quality (Jordan et al., 2012). The EU have also brought about legislation as part of the Nitrates Directive National Action Programme (NAP) which, among other regulations, includes closed periods for land spreading of manures or slurries (between mid October and mid-late January in Ireland, for example; (S.I. No. 31, 2014)). The time-frame of the open periods is based on the intrinsic vulnerability of the landscape with regard to rainfall, soil moisture and growth patterns. This allows farmers to incorporate good farming practices, such as effective mitigation techniques, and spread slurry at times when it is deemed to have less impact on water (Burt et al., 2013; Wilcock et al., 2009).

However, there is a specific need to distinguish between human and agricultural faecal pollution in complex catchments. An increased knowledge regarding the links between hydrological connectivity and pollution sources would allow for more effective pollution prevention measures (Mellander et al., 2012). Incorporating MST with other water quality parameters, such as P, could facilitate a more accurate definition of water quality than using only stand alone methods (Burt et al., 2013). This link is important as P transfer in catchments, for example, with both point and diffuse nutrient patterns, can be dependent on different river discharge and runoff pathway phases, as previously highlighted (Jordan et al., 2007).

Based on the agricultural and human settlement profiles of each catchment in this study, and building on the nutrient pathway analysis information from previous research (Mellander et al., 2012; 2013) the aim of this study was to define the sources of faecal contamination in changing hydrological phases and flow pathways
using MST. The objectives were to determine the dominance of the faecal sources in the pathway or phase, as defined by partitioning of nutrient pathways in high flows and in low flows.
3.2 Materials and Methods

3.2.1 Catchment Sites

The two catchment sites evaluated in this study were a mixed arable area, Arable B (9.5 km$^2$) in east-central Ireland, and a karst grassland, Grassland D (30 km$^2$) in west-central Ireland. Descriptions of the catchments and field monitoring and water analysis can be found in section 2.2.1 and 2.2.2 in chapter II.

3.2.2 Water Sample Collection, and Total Coliform and E. coli analysis

An autosampler (Hach Sigma 900 Max Portable Sampler or ISCO 6712 Portable Sampler) was deployed at the spring (Grassland D) and outlet (Arable B) to take a 1 L sample every 2 h over a 36-96 h time period. The water was analysed for total coliforms and E. coli occurrence by the IDEXX Colisure® Quanti Tray®/2000 system (described in more detail in section 2.2.2 in chapter II).

3.2.3 Water Sample Processing and DNA Extraction

Samples were filtered through 0.2 µm cellulose nitrate membrane filters to concentrate bacterial cells, which then underwent chemical extraction of total nucleic acids. The following nucleic acid extraction method was adapted from (Carrigg et al., 2007): the filter was cut into six segments; each segment was added to a 1.5 ml micro-centrifuge tube. To each tube, 500 µl of CTAB buffer, 500 µl of lysis buffer (50 µM Tris–HCl [pH 8]; 40 µM ethylene diamine tetraacetic acid [EDTA; pH 8]; 750 µM filter sterilised sucrose) and 20 µl of lysozyme (10 mg ml$^{-1}$) were added; mixtures were briefly vortexed (30 seconds) and incubated at 37°C for 15 min and then vortexed again before a further incubation at 37°C for 15 min. Sodium dodecyl sulphate was added to a final concentration of 2% (w/v); the samples were again vortexed and then incubated at 70°C for 1 h. After this, 6 µl of proteinase K were added. Samples were then vortexed and incubated at 50°C for a further 30 min, followed by centrifugation for 15 min (10,000×g). The supernatants were transferred to fresh 2 ml micro-centrifuge tubes and the aqueous phase was extracted by mixing with an equal volume of chloroform–isoamyl alcohol (24:1), followed by centrifugation (10,000×g) for 15 min. Nucleic acids were then precipitated from the extracted aqueous layer with addition of 0.6 vol of isopropanol in fresh 1.5 ml micro-
centrifuge tubes, overnight incubation at room temperature, followed by centrifugation (10,000×g) for 60 min and then a further centrifugation step (12,000×g) for 15 min. The pelleted nucleic acids were washed with 70% (v/v) ice-cold ethanol by centrifugation (12,000×g) for 15 min and air dried before re-suspension in 30 μl DEPC-treated water. Extracts from the six filter segments for each sample were recombined after re-suspension.

3.2.4 Faecal Sample Collection and DNA Extraction
DNA was extracted from various animal faecal sources from within each catchment to build a DNA archive of potential faecal pollution sources and also to evaluate qPCR assay specificity. Animal faecal samples from both catchments included cow (n = 7), sheep (n = 10), horse (n = 4), pig (n = 6) and goat (n = 4). Extracted DNA from human faecal samples (n = 10) was kindly donated by Galway University Hospital. The animal faecal samples (25 mg) underwent nucleic acid extraction using the PowerSoil DNA™ Isolation Kit (MoBio, Carlsbad, CA, USA) with the following modifications to the method: after solution C1 was added, a 10 min incubation at 70°C was introduced, there were two washes with solution C5. Each faecal extract was mixed into species-specific genomic pools.

3.2.5 TaqMan® Bacteroidales qPCR Assays
For all TaqMan assays, 10 μl of nucleic acid extract were assayed in a final reaction volume of 25 μl. Each 25 μl reaction contained a final concentration of 1x TaqMan® Environmental Master Mix 2.0 (Applied Biosystems®, Carlsbad, CA, USA) with 400 nM each of forward and reverse primer, and 80 nM probe. The samples were placed in 96-well plates and amplified in an ABI StepOnePlus™ Real-Time PCR System (Applied Biosystems®). Standard amplification conditions were used: 2 min at 50°C and 10 min at 95°C, followed by 40 cycles of 15 s at 95°C and 1 min at 60°C. Cycle threshold (Ct) values were evaluated with a threshold that was specific to each assay; 0.018 for universal Bacteroidales marker (BacUni-UCD) detection, 0.026 for human-specific Bacteroidales marker (BacHum-UCD) detection, and 0.022 for ruminant-specific Bacteroidales marker (BacBov-UCD) detection. The assays were
originally developed by Kildare et al. (2007). Table 3.1 summarises the assays used in this study.

3.2.6 Statistics
Statistical analysis was performed using GraphPad Prism (Version 5.00 for Windows, GraphPad Software, San Diego California, USA, www.graphpad.com). Pearson correlations were applied to *E. coli* counts, hydrochemical parameters and qPCR results from both catchments in order to determine if correlations were statistically significant. *P*-values of *P*<0.05 (*), *P*<0.01 (**) or *P*<0.001 (***) were all considered statistically significant.

3.2.7 Loadograph Recession Analysis
Loadograph Recession Analysis (LRA), developed by Mellander et al. (2012) is similar to hydrograph separation (Dingman, 2002) and End Member Mixing Analysis (EMMA), where water discharge is used to graphically separate an event. LRA can separate the major flow pathways using nutrient loads. Previous work on these two catchments used LRA to separate conduit and fissure nutrient loads in Grassland D; and the proportions of quick flow, interflow and delayed flow in Arable B (Mellander et al., 2012; 2013).

3.2.8 Microbial Source Tracking Loads
The qPCR data for the MST *Bacteroidales* assays were converted into loads by multiplying the data, which is given as gene copy (gc)/ml, by mean hourly discharge (m³/h) and correcting to gc/m³ and then gc/h. This study incorporated qPCR data to the LRA in order to distinguish conclusively whether human or animal faecal pollution was the dominant source during particular transfer pathways and phases.
Table 3.1. Primers and probes used in this study and the expected size of their amplified product.

<table>
<thead>
<tr>
<th>Primer/Probe</th>
<th>Sequence (5’-3’)</th>
<th>Product size (bp)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bacteriodales</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Universal</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BacUni-520f</td>
<td>CGTTATCCGGATTTATTGTTTA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BacUni-620r</td>
<td>AATCGGAGTTCCTCGTGATATCTA</td>
<td>170</td>
<td>Kildare et al., 2007</td>
</tr>
<tr>
<td>BacUni-656p</td>
<td>6-FAM-TGGTGTAGCGGTGAAA-MGB</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Human-specific genes</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BacHum-160f</td>
<td>TGAGTTTCACATGTCCGCATGA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BacHum-241r</td>
<td>CGTTACCCCGCTACTATCTAATG</td>
<td>81</td>
<td>Kildare et al., 2007</td>
</tr>
<tr>
<td>BacHum-193p</td>
<td>TCCGGTAGACGATGGGATCGTT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*BacCow-CF128f</td>
<td>CCAACYTCCCCGTACTCT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*BacCow-305r</td>
<td>GGACCGTGTCTCAGTCCAGTG</td>
<td>177</td>
<td>Kildare et al., 2007</td>
</tr>
<tr>
<td>*BacCow-257p</td>
<td>6-FAM-TAGGGGTTCTGAGAGGAAGGTC</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Referred to as BacBov in the text as the assay is now known as BacBov (Personal Communication).
3.3 Results

3.3.1 qPCR Assay Specificity

BacUni-UCD was assayed against cow, sheep, goat, human, horse, and pig faecal samples. It detected *Bacteroidales* in all faecal samples (Table 3.2). Both of the host-specific *Bacteroidales* qPCR assays were tested against various faecal samples in order to establish assay specificity (Table 3.2). The BacHum-UCD detected all human faecal samples and did not detect sheep, cow, pig, goat or horse faecal samples. The BacBov-UCD assay detected cow but it also detected sheep samples at similar Ct values of 14 (Fig. 3.1), and goat (1 in 10 dilution) was detected shortly after at a Ct value of 20. As a consequence of this high level of cross-reactivity between cow, sheep, and goat, the BacBov-UCD assay was determined to be more suitable for detecting agricultural ruminant faecal pollution rather than just bovine faecal pollution. BacBov-UCD did not detect human, horse or pig.
Table 3.2 Assay specificity of BacUni-UCD, BacHum-UCD, and BacBov-UCD tested against various pooled faecal samples. + = detection, - = no detection

<table>
<thead>
<tr>
<th>qPCR assays</th>
<th>Cow (n = 7)</th>
<th>Sheep (n = 10)</th>
<th>Human (n = 10)</th>
<th>Horse (n = 4)</th>
<th>Goat (n = 4)</th>
<th>Pig (n=6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BacUni-UCD</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>BacHum-UCD</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>BacBov-UCD</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>
Figure 3.1 Amplification curves for BacBov assay. Sheep (S), cow (C), human (Hu),
horse (H), goat (G) and pig (P) were assayed. The Ct cutoff (CC) was 30 Ct (red
line). Sheep and cow both amplified at a Ct of 14. Goat (which was a 1 in 10
dilution) amplified at a Ct of 20. Human and pig came after the Ct cutoff at Ct of 32
and 34, respectively. Horse was close to the Ct cutoff at a Ct of 29.
3.3.2 Loadograph Recession Analysis

Loadograph Recession Analysis was used to separate the Total Phosphorus (TP) transfer pathways for all catchment events (Table 3.3). It was deduced that the individual segments represented the release of TP loads from numerous inputs and their movement via the various major flow pathways.

For Arable B, six events were sampled. The TP transfer pathways were separated into quick flow, which made up the peak of the hydrograph, which was deemed to be represented by overland flow, preferential flow, and tile and ditch drainage. This was followed by interflow and delayed flow, which was assumed to include baseflow.

For Grassland D, two events were sampled. These events were separated into conduit flow, which was assumed to be represented by the first quick flow pathway. This was followed by large fissure flow, which was assumed to be represented by the second quick flow pathway. Medium sized fissures and sometimes small sized fissures, depending on the event, were represented by the delayed flow pathway.
### Table 3.3 Transfer pathways for Arable B and Grassland D. Event name, length of sample event, season, and event type. ER & F: Event Rising and Falling, EF: Event falling, BF: Baseflow, ER: Event Rising.

**Catchment: Arable B**

<table>
<thead>
<tr>
<th>Event</th>
<th>Season</th>
<th>Length of sample event (hours)</th>
<th>Event Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Autumn/Winter</td>
<td>34</td>
<td>ER &amp; F</td>
</tr>
<tr>
<td>II</td>
<td>Spring/Summer</td>
<td>37</td>
<td>ER &amp; F</td>
</tr>
<tr>
<td>III</td>
<td>Autumn/Winter</td>
<td>16</td>
<td>ER &amp; F</td>
</tr>
<tr>
<td>IV</td>
<td>Autumn/Winter</td>
<td>48</td>
<td>EF</td>
</tr>
<tr>
<td>V</td>
<td>Spring/Summer</td>
<td>48</td>
<td>BF</td>
</tr>
<tr>
<td>VI</td>
<td>Spring/Summer</td>
<td>37</td>
<td>EF</td>
</tr>
</tbody>
</table>

**Catchment: Grassland D**

<table>
<thead>
<tr>
<th>Events</th>
<th>Season</th>
<th>Length of sample event (hours)</th>
<th>Event Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>VII</td>
<td>Autumn/Winter</td>
<td>48</td>
<td>ER</td>
</tr>
<tr>
<td>VIII</td>
<td>Spring/Summer</td>
<td>48</td>
<td>ER</td>
</tr>
</tbody>
</table>
3.3.3 Transfer Pathways

During event I in Arable B, the loads were transferred via quick flow, interflow and delayed flow (Fig. 3.2). Significant positive correlations between *E. coli* loads and TP loads (*P*<0.001), and between BacUni loads and TP loads (*P*<0.001) were noted for the duration of the event (Table 3.4). This would suggest the possibility of agricultural runoff or human contamination, such as sewage seepage, which would contain high levels of *E. coli*, TP and faecal pollution. Pearson statistical analysis indicated that there were significant positive correlations between *E. coli* loads and BacUni loads (*P*<0.01); between *E. coli* loads and BacBov loads (*P*<0.01); and between BacUni loads and BacBov loads (*P*<0.001). Animal contamination was shown as one source of faecal pollution for this particular event caused by the presence of the BacBov marker. A significant positive correlation between BacUni loads and BacHum loads (*P*<0.05) demonstrates that human faecal contamination was another pollution source for the entirety of Arable B event I, but the concentration was much lower than that of the ruminant contamination (Fig. 3.2).

When the data representing the quick flow period were analysed as a discrete set, significant positive correlations were noted between TP loads and BacUni loads (*P*<0.01); and between TP loads and *E. coli* (*P*<0.01; Table 3.5). A significant correlation was also found between BacUni loads and BacBov loads (*P*<0.05; Table 3.5). This is illustrated in Figure 3.2, where the peak in BacUni loads of $1.24 \times 10^{15}$ gc/h at 7 h corresponded to the peak in BacBov loads, of $3.01 \times 10^{11}$ gc/h. The peak in TP of 4982.76 g/h at 6 h just preceded the peak in BacUni at 7 h (Fig. 3.2). As shown in Figure 3.2, high levels of TP, *E. coli*, BacUni and BacBov strongly suggest animal faecal contamination was present. The data strongly suggest that this agricultural pollution was the dominant source of faecal contamination during quick flow for Arable B, event I. There were not enough data to statistically analyse the interflow period. Analysis of the delayed flow period revealed no correlations between any of the MST loads, but a significant positive correlation was seen between *E. coli* loads and TP loads (*P*<0.01). Figure 3.2 illustrates that there was still ruminant contamination present, albeit at lower levels than during the quick flow period, and that a small amount of human contamination was also evident.
**Table 3.4** Pearson’s correlations matrix for loads in Arable B during entirety of event I, of *E. coli*, total phosphorus (TP), BacUni, BacHum and BacBov. *P*-value of $P<0.05 = *, P<0.01 = **, P<0.001 = ***$, N.C. = No correlation

<table>
<thead>
<tr>
<th></th>
<th><em>E. coli</em></th>
<th>TP</th>
<th>BacUni</th>
<th>BacHum</th>
<th>BacBov</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. coli</em></td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP</td>
<td>***</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BacUni</td>
<td>**</td>
<td>***</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BacHum</td>
<td>N.C.</td>
<td>N.C.</td>
<td>*</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>BacBov</td>
<td>**</td>
<td>***</td>
<td>***</td>
<td>N.C.</td>
<td>-</td>
</tr>
</tbody>
</table>

**Table 3.5** Pearson’s correlations matrix for loads in Arable B during quick flow of the event I, of *E. coli*, total phosphorus (TP), BacUni, BacHum and BacBov. *P*-value of $P<0.05 = *, P<0.01 = **$, N.C. = No correlation

<table>
<thead>
<tr>
<th></th>
<th><em>E. coli</em></th>
<th>TP</th>
<th>BacUni</th>
<th>BacHum</th>
<th>BacBov</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. coli</em></td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP</td>
<td>**</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BacUni</td>
<td>N.C.</td>
<td>**</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BacHum</td>
<td>N.C.</td>
<td>N.C.</td>
<td>N.C.</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>BacBov</td>
<td>N.C.</td>
<td>*</td>
<td>*</td>
<td>N.C.</td>
<td>-</td>
</tr>
</tbody>
</table>
Figure 3.2 Arable B qPCR loadograph for event I. Time series of hourly TP loads (black line), *E. coli* loads (green line with dots), BacUni loads (orange line with hollow black circles), BacHum loads (black line with blue triangle), and BacBov loads (black line with red squares). Data have been separated into transfer pathways using Loadograph Recession Analysis.
For event II in Arable B, Pearson’s correlations showed significant positive correlations between *E. coli* loads and BacUni loads ($P<0.001$); between *E. coli* loads and TP loads ($P<0.05$) and between TP loads and BacUni loads ($P<0.05$; Table 3.6), which would all indicate faecal contamination from animal or human sources. Figure 3.6 illustrates that the concentration of human-derived faecal contamination was higher than the ruminant-derived contamination for the entirety of this event. The rate of discharge was quite low (data not shown) across this event, resulting in almost baseflow conditions. It is interesting to note that the peaks in BacHum with similar concentrations showed diurnal patterning (Fig. 3.6). BacHum at 2 h and 26 h had a concentration of $1 \times 10^9$ gc/h and $9.52 \times 10^8$ gc/h respectively, which corresponded to 22:00 on both days. The peaks in BacHum of $2.85 \times 10^9$ gc/h and $2.58 \times 10^9$ gc/h occurred at 6 h and 30 h respectively and corresponded to 02:00 on both days. The largest BacHum peak at 14 h was $3.37 \times 10^9$ gc/h which corresponded to 10:00. The peak observed 24 h later, of $3.65 \times 10^8$ gc/h, was one magnitude lower than the 14 h peak, but still quite high.

For the quick flow in Arable B, event II, there was a positive significant correlation between *E. coli* loads and BacUni loads ($P<0.05$; Table 3.7). A positive significant correlation was also noted between BacUni loads and BacBov loads ($P<0.05$) for the quick flow (Table 3.7). At 10 h (Fig. 3.3), the peak in BacUni loads, of $3.30 \times 10^{10}$ gc/h, coincided with the peak in BacBov loads ($1.60 \times 10^9$ gc/h) and with the peak in *E. coli* loads ($1.67 \times 10^{10}$ gc/h). These results suggest that agricultural animal contamination was again the dominant cause of faecal pollution during the quick flow period of event II. The amount of data for interflow was insufficient for statistical analysis. For the delayed flow, there was just one significant positive correlation noted, between *E. coli* loads and TP loads ($P<0.05$).
Table 3.6  Pearson’s correlations matrix for loads in Arable B during entirety of event II, of *E. coli*, total phosphorus (TP), BacUni, BacHum and BacBov. *P*-value of $P<0.05 = \ast$, $P<0.001 = \ast\ast\ast$, N.C. = No correlation

<table>
<thead>
<tr>
<th></th>
<th>E. coli</th>
<th>TP</th>
<th>BacUni</th>
<th>BacHum</th>
<th>BacBov</th>
</tr>
</thead>
<tbody>
<tr>
<td>E. coli</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BacUni</td>
<td>***</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BacHum</td>
<td>N.C.</td>
<td>N.C.</td>
<td>N.C.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BacBov</td>
<td>N.C.</td>
<td>N.C.</td>
<td>N.C.</td>
<td>N.C.</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.7  Pearson’s correlations matrix for loads in Arable B during quick flow of event II, of *E. coli*, total phosphorus (TP), BacUni, BacHum and BacBov. *P*-value of $P<0.05 = \ast$, N.C. = No correlation

<table>
<thead>
<tr>
<th></th>
<th>E. coli</th>
<th>TP</th>
<th>BacUni</th>
<th>BacHum</th>
<th>BacBov</th>
</tr>
</thead>
<tbody>
<tr>
<td>E. coli</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP</td>
<td>N.C.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BacUni</td>
<td>*</td>
<td>N.C.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BacHum</td>
<td>N.C.</td>
<td>N.C.</td>
<td>N.C.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BacBov</td>
<td>N.C.</td>
<td>N.C.</td>
<td>*</td>
<td>N.C.</td>
<td></td>
</tr>
</tbody>
</table>
Figure 3.3 Arable B qPCR loadograph for event II. Time series of hourly TP loads (black line), *E. coli* loads (green line with dots), BacUni loads (orange line with hollow black circles), BacHum loads (black line with blue triangle), and BacBov loads (black line with red squares). Data have been separated into transfer pathways using Loadograph Recession Analysis.
For the remainder of the events sampled in Arable B, the loads were only transferred via one pathway. During event III, the loads were moved during quick flow, where a significant positive correlation was seen between TP loads and *E. coli* loads ($P<0.001$), between BacUni loads and BacHum loads ($P<0.01$) and between BacBov loads and BacHum loads ($P<0.05$). Event IV was mostly characterised by delayed flow and only two samples were transferred via interflow. Pearson’s correlations indicated significant positive correlations between TP loads and *E. coli* loads ($P<0.001$), and between BacUni loads and BacBov loads ($P<0.001$), showing that agricultural faecal contamination was the dominant source during this autumn/winter event.

Event V was an unusual event as cows had access to the river during the delayed flow period, when all of the microbial loads were transferred. Significant correlations were noted between BacUni loads and TP loads ($P<0.001$), and BacUni loads and BacBov loads ($P<0.05$). The significant positive correlation between BacHum loads and BacBov loads ($P<0.001$) shows that agricultural and human contamination was present.

All loads were transferred during delayed flow for event VI. Statistical analysis illustrates significant positive correlations between TP loads and *E. coli* loads ($P<0.001$), and between BacHum loads and TP loads ($P<0.05$), showing that human contamination was present. The significant positive correlations between *E. coli* loads and BacBov loads ($P<0.001$), and BacUni loads and BacBov loads ($P<0.001$) demonstrate that agricultural faecal contamination was also impacting this event.

There were only two events in Grassland D (Table 3.3). Event VII was an autumn/winter event rising, with all loads being transferred via the medium fissure and small fissure flow. Only BacUni-UCD was assayed against samples for this event. For the overall event, Pearson’s correlations showed significant positive correlations between TP loads and *E. coli* loads ($P<0.01$), and TP loads and BacUni loads ($P<0.05$) proposing that faecal contamination was present but as the host-specific markers were not applied to these water samples, it is not possible to differentiate between anthropogenic and agricultural sources. For the individual
medium fissure and small fissure flows there were no significant correlations observed.

For event VIII, all loads were moved during the conduit flow. Significant positive correlations were seen between BacUni loads and BacBov loads ($P<0.05$), BacUni loads and BacHum loads ($P<0.05$), and TP loads and *E. coli* loads ($P<0.05$). These results show that human and agricultural faecal contamination was present.
3.4 Discussion

Defining sources of faecal contamination in large catchments can be challenging (Crowther, et al., 2002). Microbial source tracking, in conjunction with E. coli loads and loadograph recession analysis (LRA) can be used to determine what the source of faecal contamination is and whether it is transferred by a quick flow, interflow or baseflow pathway or phase. As of yet, no study has utilised LRA and MST for this purpose. Using LRA, and differentiating between anthropogenic and agricultural faecal pollution and determining the route by which it travelled, will be important for limiting the trophic impact of farms and sewage systems on water (Kay et al., 2010; Kildare et al., 2007; Mellander et al., 2012). This practical and efficient toolkit could be used to prioritise best management practices and mitigation techniques, in order to protecting human health and the environment (Arnscheidt et al., 2007; Edwards et al., 2008; Wilcock et al., 2009).

From a risk assessment perspective, Arable B and Grassland D are two very different catchments. As Grassland D has more nutrient and faecal inputs from more intensive agriculture, it was assumed to be more subject to agricultural contamination, however transport limitation, as determined for specific P vulnerability (Mellander et al., 2013), appears to also reduce FIO transfer. Arable B, on the other hand, is source limited in terms of animal agriculture, yet large amounts of ruminant agricultural pollution were detected in the outlet. This is a result of the faster transport pathways that are available in Arable B, such as surface and near-surface flow, which connect the sources.

The human population was presumed to have a higher impact in Arable B (although less than the agricultural animal pressure) as the housing density is greater and it is likely, as a result of less permeable soil, to have STS effluents that are not retained as efficiently as in areas with more permeable soils. It was hypothesised and demonstrated here that signals would be particularly strong during periods of delayed flow, when diffuse sources would not be connected. The diurnal cycling noted in the BacHum results for one Arable B delayed flow event is synonymous with other studies indicating nutrient diurnal cycling at small (Jordan et al., 2007)
Chapter III

and large (Wade et al., 2012) scales and is indicative of wastewater discharges. While this kind of chronic wastewater pollution is likely to be continuous and evident in times of water stress in Arable B, the results also show that human sources of *E. coli* were present in quick flow events. This is likely to be due to misconnections of storm waters through STS or via deposition and reconnection processes in the more ephemeral or low gradient streams receiving STS effluents, which are flushed during rain events (Arnscheidt et al., 2007; Withers et al., 2014).

Several studies have shown that the presence of high *E. coli* counts or elevated levels of P indicates faecal pollution, the source of which can then be determined by host-specific assays. Boehm et al. (2010) demonstrated that human faecal contamination was found in water with high concentrations of soluble reactive P and *E. coli*. Jenkins et al. (2009) showed that when there were high levels of *E. coli* in the water (up to 20,000 CFU/100 ml), the BacUni and BacBov markers detected agricultural ruminant faecal contamination. Kinzelman and McLellan, (2009) found that the human-specific end-point PCR assay, HF183, detected human pollution in water that was associated with counts of *E. coli* over $10^4$ CFU/100 ml. The disadvantage of using end-point PCR is that as PCR is qualitative, it only shows absence or presence and therefore cannot be correlated to quantitative data, such as *E. coli* counts. Hence, qPCR measurement of *Bacteroidales* RNA can be used to establish viable counts of faecal contamination (Walters et al., 2009). When Lee et al. (2014) measured faecal contamination using host-specific *Bacteroidales* qPCR assays, they found that *E. coli* numbers correlated well with the assays. Their research also took into consideration the various land use parameters of their study sites, allowing them to define accurately the source of contamination. In the findings presented here, for most of the sample events the host-specific qPCR assays correlated well with *E. coli* and/or TP. As the assays were used in alliance with loadograph recession analysis, it was possible to determine, for the first time, during which pathway the source was dominant.

For Arable B, it was observed that agricultural ruminant pollution is more likely to dominate the quick flow. As quick flow is deemed to be derived from surface and
Chapter III

near-surface flow and, as some of the events (I and III) coincided with the closing of the slurry spreading period, it is likely that at least some of this faecal transfer was from slurry (Mellander et al., 2012; Monaghan et al., 2007; S.I. No. 31, 2014). Defining agriculture as the source of faecal pollution during quick flow allows for specific mitigation techniques to be put in place to limit mobilisation potential. Landscape features that can attenuate pollutants in quick flow pathways, such as ditch networks (Shore et al., 2014), wetlands, attenuation features (Ockenden et al., 2012) and buffer strips (Anbumozhi et al., 2005; Dawson and Smith, 2010), may be similarly applicable to faecal matter transfers and retention. Such features could be used as part of multi-objective schemes and as part of paid eco-system service provision (e.g. Buckley et al., 2012).

Event V noted that a short period of cattle trafficking in the Arable B river was clearly defined by the BacBov results during a delayed, low flow period. This, however, was likely due to direct defecation by cattle, bed disturbance and/or wash-off of any faecal matter (Davies-Colley et al., 2004), which occurred during a period where there was already a high background level of faecal matter present.

In Grassland D, for event VII, the statistically significant positive correlations between TP, E. coli and BacUni would strongly suggest that faecal contamination was present during the medium and small fissure flows, which are both delayed flow. As the host-specific (BacBov and BacHum) assays were not applied to the water samples in this event, it was not possible to determine the specific source of faecal contamination. Human and agricultural faecal contamination travelled through the conduit flow in high numbers for event VIII (which was during a deluge of rain), showing the rapid response of the karstic aquifer, directly connected to the surface, to a heavy rainfall event (Pronk et al., 2009).

From a regulatory perspective water that is found to have high levels of faecal contamination does not comply with the drinking or bathing water standards established by the EU (BWD 76/60/EEC; 2000/60/EC) (OJEU, 1998; 2006). Therefore, determining what source is dominant during which pathway or flow phase
will be important for the protection of water through specific measures and via targeted policing. Faecal contamination of water from agricultural or anthropogenic activities can impact greatly on human health and the environment (Nnane et al., 2011; Withers et al., 2011). From a health perspective zoonotic waterborne pathogens such as *Cryptosporidium*, which have been a problem in Irish catchments in the past, can be found in faecally contaminated water and drinking this water can cause diarrhoea (Cummins et al., 2010). Gastroenteritis can be caused by drinking polluted water and polluted bathing waters can cause respiratory and skin infections (Prüss, 1998; Soller et al., 2010). Applying Quantitative Microbial Risk Assessment (QMRA) to both of these catchments would be a potential way of establishing the level of risk to human health that zoonotic waterborne pathogens or pathogenic bacteria, associated with faecally contaminated water, can cause (Cummins et al., 2010).
3.5 Conclusions

This research took a catchment-wide approach to analysing water for the discrimination of faecal contamination from different human and agricultural animal sources using an MST analysis of water samples and from an analysis of correlated nutrient pathways. The methods and results established that:

- Faecal contamination from agricultural animals dominated the transfers from two contrasting catchments; one (karst grassland) transport limited with a comparatively high animal stocking density and one (mixed arable) with a less transport limited system and a lower animal stocking density.

- The less transport limited system indicated more faecal matter transfer overall and was also related to all sources.

- In both catchments, quick flows, separated by a P load recession analysis (where P was highly correlated with FIO parameters), indicated that this pathway transferred faecal loads from both agricultural and human sources more so than other pathways.

- In the mixed arable catchment, MST showed a strong signal of faecal contamination from human origins (combined with agricultural animal *E.coli*), which followed a diurnal cycle. This is indicative of wastewater inputs over 24 hour cycles and is similar to nutrient cycling, as found in other studies.

- Increases of human derived faecal contamination in quick flows, rather than dilution, indicated that sewage sources are either mis-connected to storm drains or are accumulated in ephemeral ditches and then flushed during rain events.

- The presence of ruminant faecal matter at low flows also indicated that this source was not disconnected during these flow phases and was not only dependent on storm quick flows for transfer from land to water.
The techniques used in this study, by combining DNA based analysis of faecal bacteria to discriminate faecal matter sources as well as high-resolution P analysis for hydrological pathway discrimination, show the potential for catchment risk assessment of faecal matter pollution, mitigation appraisal and policy monitoring. The use of qPCR to determine the dynamically changing nature of source influence at different hydrological phases is an important development for water resources management. Further analysis would benefit from addressing the issue of the specificity of the BacBov assay. This assay was used to discriminate agricultural animal Bacteroidales in this study, but it was not specific to cow faecal pollution, as it also detected sheep.
3.6 References


Chapter III


Chapter III


in an intensively farmed catchment in southern New Zealand. *Agriculture, Ecosystems and Environment, 118*(1-4), 211-222.


Water Framework Directive 2006/7/EC.


Chapter IV

An Ovine-Specific *Bacteroidales* PCR and qPCR Assay for Microbial Source Tracking

Manuscript in preparation for submission to Water Research
Author Contribution

Sinéad Murphy
(Microbial Ecology Lab, Microbiology Dept., NUI Galway, Galway
Department of Civil and Environmental Engineering, UC Davis, California)
Sample collection and processing, results analysis, script writing

Minji Kim
(Department of Civil and Environmental Engineering, UC Davis, California)
Sample processing and results analysis (up to sequence analysis phase)

Stefan Wuertz
(Department of Civil and Environmental Engineering, UC Davis, California)
Project supervision and guidance

Vincent O’Flaherty
(Supervisor; P.I. of Microbial Ecology Lab, Microbiology Dept., NUI Galway, Galway)
PhD supervision and guidance
4.1 Introduction

The current global population of sheep stands at over 1 billion (www.thesheepsite.com). According to the Ireland National Sheep and Goat Census, Ireland has a large population of approximately 3.58 million sheep (Department of Agriculture, 2012). Sheep farming is very important in Ireland, coming in at fourth for most valuable financial enterprise (www.teagasc.ie/sheep/). High levels of bacteria, such as *Campylobacter coli* (*C. coli*), *Escherichia coli* (*E. coli*), and *Salmonella* spp., and protozoans such as *Cryptosporidium parvum* (*C. parvum*), *Cryptosporidium hominis* (*C. hominis*), and *Giardia* spp. can be found in sheep faeces (Devane et al., 2005; Ministry for the Environment, 2003; Xiao et al., 2004). Sheep faeces also contain nutrients, such as nitrate and phosphorus (McGoll and Gibson, 1979). Therefore, if sheep faecal pollution transfers into waterways, it can have a negative impact on human health causing diarrheal disease; respiratory illnesses; eye and skin infections; and impinge the local environment by causing eutrophication (Arnscheidt et al., 2007; Nnane et al., 2011; Prüss, 1998).

Grazing animals that defecate directly onto the land; and slurry or manure from housed livestock that is spread on the land are two major contributors of non-human faecal indicator organisms (Kay et al., 2008; Tian et al., 2002). Faecal pollution from sheep or cattle can cause a pathogenic toxigenic *E. coli* outbreak in private water supplies (Licence et al., 2001). It is therefore very important to know whether sheep or cattle contribute more to faecal pollution in order to establish which has the higher risk for water contamination. Research in New Zealand showed that, during events when stocking densities are similar, sheep pasture produces greater *E. coli* loading rates than cattle pastures (Wilcock, 2006). This is in agreement with studies in Scotland, which indicated that sheep grazing resulted in a higher risk of *E. coli* leaching to field drains than from slurry spreading in either wet summer conditions or autumn/spring (Vinten et al., 2004). Moriarty et al. (2011) demonstrated that sheep faeces contain higher levels of *E. coli* and *Cryptosporidium* spp. than cow faeces (Moriarty et al., 2011). As sheep can be grazed on hills and will defecate directly onto the land, there is an increased possibility of faecal bacteria transfer into
waterways as a result of the steep topography which can cause increased overland flow when there is a heavy or prolonged rainfall event (Collins et al., 2005).

Detecting faecal pollution and distinguishing its source is imperative for the protection of water quality and ergo human health and the environment (Allevi et al., 2013; Walters, et al., 2011). Microbial source tracking can be used to achieve this as it uses bacterial organisms that are found in human or animal faeces to identify faecal pollution in water (Kay et al., 2008). Bacteroidales are enteric bacteria that are strict anaerobes and are found in the human intestine and the rumen of warm-blooded animals (Bernhard and Field, 2000a; Kildare et al., 2007; Layton et al., 2006). They comprise a considerable portion of the faecal bacteria population (Ahmed et al., 2008; Bernhard and Field, 2000a; Kildare et al., 2007; Layton et al., 2006). As a result of a high degree of host specificity, they can be used to distinguish between human and agricultural faecal contamination (Kildare et al., 2007; Layton et al., 2006).

Multiple host-specific Bacteroidales 16S molecular markers for humans and cows have been designed using various techniques and approaches (Bernhard and Field, 2000b; Dorai-Raj et al., 2009; Kildare et al., 2007). Terminal restriction fragment length polymorphism (TRFLP) analysis has been used to define human- and ruminant-specific 16S rRNA sequences from which host-specific PCR primers were subsequently designed (Bernhard and Field, 2000b; Dorai-Raj et al., 2009). Kildare et al. (2007) produced universal, human, and bovine qPCR primers and probes by taking Bacteroidales sequences from GenBank and aligning them in order to establish a host-specific sequence. While each of these assays display high specificity for distinguishing between human and ruminant, cross-amplification between cow and sheep has been shown in the studies that have tested these markers (Bernhard and Field, 2000b; Dorai-Raj et al., 2009; Murphy et al., 2015 - in preparation; Reischer et al., 2013). At the time of writing, no sheep-specific Bacteroidales PCR or qPCR assay has been designed.
Subtractive hybridisation is an important technique for establishing specific DNA sequences by differentiating between target gene fragments from diverging, though closely related subtracter gene fragments (Dick et al., 2005; Zwirglmaier et al., 2001). Dick et al. (2005) used subtractive hybridisation to design Bacteroidales host-specific PCR primers that successfully distinguished between cow and elk which would be highly comparable faecal sources. Green et al. (Green et al., 2012) were able to use the technique to establish a very specific gull qPCR assay from a small number of sequences. Based on the fact that subtractive hybridisation can produce unique DNA sequences and that there is a demand for a sheep-specific assay that can differentiate between cow and sheep faecal pollution, the aim of this study was to use subtractive hybridisation to design and establish a Bacteroidales sheep-specific qPCR assay.
4.2 Materials and Methods

4.2.1 Target and Subtracter DNA
This protocol used sheep faeces as the target source and cow and human faeces as the subtracter sources. The protocol is summarised in the flow chart (Fig. 4.1). Sheep and cow faecal samples were collected from fields that contained only sheep or cow so as to avoid cross-contamination. Types of sheep from which samples were collected were Cheviot, Blackface Mountain and Suffolk. For the sheep composites, 10 samples were collected, and for the cow composite 7 samples were collected. Human faecal DNA extracts (10) were kindly donated by Galway University Hospital.

4.2.2 Sample Collection and DNA Extraction
Sheep and cow faecal samples were collected from multiple farms and stored at -20°C until extraction. DNA was extracted from 25 mg of faecal sample using the Powersoil™ DNA Isolation Kit (Cambio, Cambridgeshire, UK) following the manufacturer’s protocol, which was modified to include a 10 min incubation at 70°C after the addition of solution C1, and two washes with solution C5. gDNA, in each individual DNA extract, was measured using a NanoDrop 2000 (Thermo Fisher Scientific Inc, Wilmington, DE, USA). The extracts were quantified using the universal Bacteroidales assay, BacUni (Kildare et al., 2007) to confirm that the faecal extracts contained the universal Bacteroidales sequence. Only extracts that contained the sequence were used for the process of subtractive hybridisation. Each DNA extract (which were verified with the BacUni assay) was diluted to an appropriate concentration (20 ng/µl for sheep and cow, and 3 ng/µl for human). Individual sheep faecal extracts were mixed together to make a pool of genomic DNA (Figure 4.1). This was repeated for cow and human faecal extracts.
Figure 4.1 Flow chart of process involved in Subtractive Hybridisation (Adapted from Dick et al., 2005).
4.2.3 PCR Amplification

PCR primers extended with AcI restriction sites were used to amplify target (sheep) and subtracter (cow and human) genomic pools (Fig. 4.1). The Bacteroidales-specific 16S rRNA primer AcI Bac32F and the bacterial 16S rRNA primer 1492R amplifies approximately 1460 bp. Each 50 µl reaction contained 1X Taq polymerase buffer, each primer at a concentration of 1 µM, dNTP at a concentration of 200 µM, 2 mM MgCl₂, 1.25 U Taq polymerase and 5 µl template. Cycling parameters were as follows: initial denaturation of 94°C for 10 min followed by 30 cycles of 94°C for 1 min, 60°C for 1 min, and 72°C for 2 min, and final extension at 72°C for 15 min on an Eppendorf Mastercycler Gradient (Eppendorf, Hamburg, Germany).

4.2.4 Restriction Digestion

The products of PCR underwent AcI restriction enzyme digestion (Fig. 4.1). Each 30 µl reaction contained 6 µl nuclease-free water, 3 µl 10X AcI buffer, 20 µl PCR product and 1 µl AcI enzyme. The reaction was incubated at 37°C for 60 min. The products underwent gel electrophoresis using 15-20 µl of each sample.

4.2.5 Preparation of Linkers and Ligation

Separate linkers for subtracter and target DNA (S1/S2 and T1/T2) were obtained from Invitrogen Corp. (Carlsbad, CA, USA). Each linker pair was diluted, 8 µg in 40 µl 10 mM Tris-HCl, pH 8.5. The linkers were heated to 65°C for 10 min, 37°C for 10 min, and 20°C for 20 min using a water bath. This allowed gradual hybridization of double-stranded linkers. The hybridized linkers contain AcI-compatible 5’ overhangs. Target and subtracter restriction fragments were ligated to their respective linkers using T4 DNA ligase (M202S; New England Biolabs, Ipswich, MA, USA). Each 20 µl reaction contained 2 µl 10X T4 DNA ligase buffer, 3 µl target or subtracter DNA (200 ng/µl), 1 µl linker T/S, 13 µl nuclease-free water and 1 µl T4 DNA Ligase. Ligation took 16 h at 16°C. After incubation there was a heat inactivation step at 65°C for 10 min. Target and subtracter restriction fragments were ligated to their respective linkers using T4 DNA ligase (M202S; New England Biolabs). The ligation products then underwent PCR purification to remove excess linker (QIAquick® PCR purification kit).
4.2.6 Amplification of Target and Subtractor DNA

T1 and S1 oligonucleotides were used as PCR primers to amplify target and subtractor ligation products, respectively (Fig. 4.1). Each 50 µl reaction contained 1X *Taq* polymerase buffer, T1/S1 primer (1 µM), each dNTP (200 µM), 2 mM MgCl₂, and 1.25 U of *Taq* polymerase. Cycling parameters were the same as previously but with a final extension step at 72°C for 10 min. The products underwent gel electrophoresis using 6 µl of each sample. The PCR products (from cow and human samples) were re-amplified using the same PCR mixture and cycling parameters to increase the DNA concentration after purification.

4.2.7 Immobilization of Subtractor in Microplate Well

Equal amounts of subtracter DNA from the two host sources were mixed and diluted in 50 µl of PBS buffer (8.0 mM Na₂HPO₄, 1.5 mM KH₂PO₄, 137 mM NaCl, 27 mM KCl, pH 7.2; and MgCl₂ added to a final concentration of 100 mM)(Fig. 4.1). The mixture was heated to 99°C for 10 min to denature the DNA. The mixture, 50 µl, was immediately placed into a MaxiSorp microplate well (Nalge Nunc, Naperville, IL, USA). The plate was incubated at 37°C for 1 h, the buffer was removed and the plate was dried for 2 h at 70°C.

4.2.8 Solution Hybridization

The target DNA was diluted 1:10 in hybridization buffer to a final volume of 40 µl (2.5X SSC [1X SSC is 0.15 M NaCl plus 0.015 M sodium citrate], 1% blocking reagent, 0.01% sodium dodecyl sulphate (w/v%), 0.05% *N*-lauroylsarcosine (w/v%) and formamide (29%, 35%, 41%)). The different concentrations of formamide were used for stringency optimization. The target DNA mixture was heated to 99°C for 10 min to denature the DNA. The DNA was then immediately placed on ice for 2 min. The mixture, 40 µl, was added directly to the microplate wells (Fig. 4.1). Hybridisation took 2 h at 70°C.
4.2.9 Amplification of Subtraction Products, Cloning, Sequencing, and Primer Design

Subtracted, unhybridized target DNA (2 µl) was removed from the supernatant and diluted (1:10, 1:100, 1:1000) and amplified with primer T1 (Fig. 4.1). The mixture and cycling parameters were the same as used previously for primer T1. PCR fragments were extracted using QiaQuick® gel purification kit (QIAGEN). The purified products were diluted, reamplified (additional 10 min extension step at 72°C) and cloned into TOPO TA vectors (Invitrogen™ Corp.). Ten individual clones were randomly selected for bidirectional sequencing. The unique sequences were used to design sheep-specific primers and probes for PCR and qPCR.

4.2.10 PCR Amplification of Plasmids

Sheep-specific PCR assays (BacOvi S1-S5), using the designed primers, (Table 4.1) were used to amplify plasmids containing the target sequences. Each 50 µl reaction contained 1X Taq polymerase buffer, each primer (1 µM), dNTP (200 µM), 2 mM MgCl₂, 1.25 U Taq polymerase, and 5 µl template. Cycling parameters were as follows: initial denaturation of 95°C for 5 min followed by 30 cycles of 95°C for 30 sec, 55°C for 30 sec, and 72°C for 30 sec, and final extension at 72°C for 5 min.

4.2.11 Application of Sheep-Specific PCR Assays to Sheep and Cow Faecal Extracts

The BacOvi S1-S5 PCR assays were applied to the sheep and cow pooled genomic samples to determine assay specificity. The same reaction components and concentration, and the same cycling parameters were used as in section 4.2.10.

4.2.12 Amplification of Standard Curves and Determining Assay Specificity

For the standard curve, the concentration of plasmid containing the target sequence was measured and the plasmid was diluted to give a starting concentration of 9.55 x \(10^9\) gene copy (gc)/µl. A serial dilution was performed where the concentration decreased 10-fold each time to give a final concentration of 9.55 x \(10^{-2}\) gc/µl. For each of the standard curves, 10 µl of each concentration of plasmid \((10^9-10^{-2}\) gc/µl) was assayed in a final reaction volume of 25 µl. This resulted in a 12 point standard curve which was replicated 6 times on the 96-well plates. Each 25 µl reaction
Chapter IV

contained a final concentration of 1X TaqMan® Environmental Master Mix 2.0 (Applied Biosystems®, Carlsbad, CA, USA) with 400 nM each of the forward and reverse primer, and 80 nM of probe. The standard curves were amplified in an ABI StepOnePlus™ Real-Time PCR System (Applied Biosystems®). Standard amplification conditions were used: 2 min at 50°C and 10 min at 95°C, followed by 40 cycles of 15 s at 95°C and 1 min at 60°C. Each of the BacOvi qPCR assays was tested against sheep, cow, human, horse, pig and goat faecal extracts (faecal samples were collected from various farms and DNA was extracted using the aforementioned DNA extraction method).
4.3 Results

4.3.1 Restriction Digestion and Subtractive Hybridisation
Following PCR amplification of the target and subtracter DNA using *Bacteroidales*-specific 16S rRNA primer *AciBac32F* and the bacterial 16S rRNA primer 1492R, restriction digestion was performed using *AciI*. The products of restriction digestion were 300 bp and 500 bp for sheep, cow and human (human not shown; Fig. 4.2). These products then underwent ligation followed by subtractive hybridisation. The resultant subtracted, unhybridised DNA was removed and amplified via PCR using primer T1 as forward and reverse primer – the product of which was 500 bp. The target fragments with 35% formamide stringency were cloned and sequenced as this was the optimum stringency.
Figure 4.2 Gel electrophoresis showing target (sheep) and subtracter (cow) products after amplification by *Bacteroidales* primers 32F and 1492R and enzyme restriction by *AciI*. The product sizes were 300 and 500 bp. The second subtracter (human) underwent the same process but results are not shown here.
4.3.2 Sequence Analysis and Alignment

The returned forward and reverse target sequences were assembled into contigs using the CAP contig assembly program in the BioEdit sequence alignment editor 7.2.3 (www.mbio.ncsu.edu/bioedit/bioedit.html). The contigs then underwent a NCBI BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) and none of the top hits gave matches for cow or human. More sequences were downloaded for Bacteroidales in sheep from NCBI BLAST. Separate Bacteroidales cow and human sequences were also downloaded from NCBI BLAST. All of the sheep, cow and human sequences were aligned using the BioEdit sequence alignment editor 7.2.3 (http://www.mbio.ncsu.edu/bioedit/bioedit.html). Sequences that were specific to sheep were selected to design primers and probes. Table 4.1 shows the sequences for the primers and probes designed in this study. These sequences showed no self-complementarity as confirmed by the Oligocalc website (www.basic.northwestern.edu/biotools/oligocalc.html). The melting temperature ($T_m$) of each primer and probe was calculated using the Applied Biosystems $T_m$ calculator (www6.appliedbiosystems.com/support/techtools/calc/). Following the instructions in Molecular Microbial Ecology (Osborn and Smith, 2005): the probe had a $T_m$ that was 8-10°C higher than the $T_m$ of the primers; the sequence picked was positioned as close as possible to the forward primer but did not overlap it; guanine residues were avoided at the 5’ end as they are natural quenchers; and there were more guanine residues than cytosine residues.
Table 4.1 Primers, linkers and probes used in this study. Restriction site extensions and overhangs are underlined.

<table>
<thead>
<tr>
<th>Primer/linker/probe</th>
<th>Oligonucleotide sequence (5′- 3′)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BacUni Assay</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BacUni-520f</td>
<td>CGTTATCCGGATTTATTGTTTA</td>
<td>(Kildare et al., 2007)</td>
</tr>
<tr>
<td>BacUni-620r</td>
<td>AATCGGAGTTCCTCGTGATATCTA</td>
<td>(Kildare et al., 2007)</td>
</tr>
<tr>
<td>BacUni-656p</td>
<td>6-FAM-TGGTGTAGCGGTTAAA-MGB</td>
<td>(Kildare et al., 2007)</td>
</tr>
<tr>
<td><strong>AciBac32F</strong></td>
<td>AATATAAACCAGCAACGCTAGCTACAGGCTT</td>
<td>(Bernhard and Field, 2000a)</td>
</tr>
<tr>
<td><strong>Aci1492R</strong></td>
<td>AATATAAACCAGCTACCTTTGTTACGACTT</td>
<td>(Wilson et al., 1990)</td>
</tr>
<tr>
<td><strong>Linker S1</strong></td>
<td>CGCCAGGGAACACCCAGTCAGCAC</td>
<td>(Zwirglmaier et al., 2001)</td>
</tr>
<tr>
<td><strong>Linker S2</strong></td>
<td>CGGTCGGTGACTGGGTGTCCCTGGCG</td>
<td>(Zwirglmaier et al., 2001)</td>
</tr>
<tr>
<td><strong>Linker T1</strong></td>
<td>AGGGGATAAAACCAATCACACACCA</td>
<td>(Zwirglmaier et al., 2001)</td>
</tr>
<tr>
<td><strong>Linker T2</strong></td>
<td>CGTGGTGTGTAATGGTTATCCCCCT</td>
<td>(Zwirglmaier et al., 2001)</td>
</tr>
<tr>
<td><strong>BacOvi S1 assay</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S2-4F</td>
<td>AACGATGAATAACTCGCTGTTG</td>
<td>This study</td>
</tr>
<tr>
<td>S2-4R1</td>
<td>CTTAATGTTTTCCCTACGTAC</td>
<td>This study</td>
</tr>
<tr>
<td>S2-4P</td>
<td>6-FAM-ATACAATGTCAAGGCTATGTC-TAMRA</td>
<td>This study</td>
</tr>
<tr>
<td><strong>BacOvi S2 assay</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S2-4F</td>
<td>AACGATGAATAACTCGCTGTTG</td>
<td>This study</td>
</tr>
<tr>
<td>S2-4R2</td>
<td>CGTGGCAGGCCGTACT</td>
<td>This study</td>
</tr>
<tr>
<td>S2-4P</td>
<td>6-FAM-ATACAATGTCAAGGCTATGTC-TAMRA</td>
<td>This study</td>
</tr>
<tr>
<td><strong>BacOvi S3 assay</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sp2F</td>
<td>AACGAGCGCAACCCTT</td>
<td>This study</td>
</tr>
<tr>
<td>Sp3R</td>
<td>GCAGTCCCGGTGAGT</td>
<td>This study</td>
</tr>
<tr>
<td>Sp3P</td>
<td>6-FAM-TAGTTGCCATCAGGTTGATGCT-TAMRA</td>
<td>This study</td>
</tr>
<tr>
<td><strong>BacOvi S4 assay</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sp1BF</td>
<td>ATGGAGGGTGACAGAAAGCC</td>
<td>This study</td>
</tr>
<tr>
<td>Sp1BR</td>
<td>CCGAATCGAGAAGGGTTTT</td>
<td>This study</td>
</tr>
<tr>
<td>Sp3P</td>
<td>6-FAM-TCAGTGCGACACTGACTCCA-TAMRA</td>
<td>This study</td>
</tr>
<tr>
<td><strong>BacOvi S5 assay</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sp1BF</td>
<td>ATGGAGGGTGACAGAAAGCC</td>
<td>This study</td>
</tr>
<tr>
<td>Sp2BR</td>
<td>CGGATCGAGACTCCA</td>
<td>This study</td>
</tr>
<tr>
<td>Sp3P</td>
<td>6-FAM-TCAGTGCGACACTGACTCCA-TAMRA</td>
<td>This study</td>
</tr>
</tbody>
</table>
4.3.3 Application of BacOvi PCR Assays to Plasmids and Faecal Extracts

The plasmids that contained the target sequence from which the BacOvi S1-S5 primers and probes were designed were used as positive controls in order to establish whether the assays would work successfully. The BacOvi PCR assays S1-S5 were applied to an aliquot of neat plasmid, and a dilution of 1:100. All five assays successfully detected both the neat plasmid and the 1:100 dilution, as illustrated in Figure 4.3. The expected bp size of each product is shown in Table 4.2. As all five assays were effective, the assays were then applied to sheep and cow faecal extracts in order to ascertain assay specificity (Fig. 4.4). BacOvi S1 and S4 did not detect either sheep or cow. BacOvi S2 amplified products but they could not be determined as sheep or cow as they were not the expected size of 101 bp. For BacOvi S3 a faint band was seen for sheep and a fainter band was visible for cow. BacOvi S5 amplified both sheep and cow.

**Table 4.2** Expected product size (bp) of each BacOvi PCR assay.

<table>
<thead>
<tr>
<th>Assay</th>
<th>Bp</th>
</tr>
</thead>
<tbody>
<tr>
<td>BacOvi S1</td>
<td>71</td>
</tr>
<tr>
<td>BacOvi S2</td>
<td>101</td>
</tr>
<tr>
<td>BacOvi S3</td>
<td>62</td>
</tr>
<tr>
<td>BacOvi S4</td>
<td>66</td>
</tr>
<tr>
<td>BacOvi S5</td>
<td>84</td>
</tr>
</tbody>
</table>
Figure 4.3 BacOvi S1-S5 amplicons of plasmids containing target sequence. HLV: Hyperladder™ V 500bp marker, B: Blank; PN: BacOvi Plasmid Neat, PD: BacOvi Plasmid 1:100.

Figure 4.4 BacOvi S1-S5 amplicons of plasmids and faecal extracts. Lane 1 HLV: Hyperladder™ V 500bp marker, Lane 2: Blank, Lane 3: BacOvi S1 Plasmid (P) Neat, Lane 4: BacOvi S1 P 1:100, Lane 5: Sheep, Lane 6: Cow, Lane 7: Negative control (-C), Lane 8: BacOvi S2 P Neat, Lane 9: BacOvi S2 P 1:100, Lane 10: Sheep, Lane 11: Cow, Lane 12: -C Lane 13: BacOvi S3 P Neat, Lane 14: BacOvi S3 P 1:100, Lane 15: Sheep, Lane 16: Cow, Lane 17: -C, Lane 18: BacOvi S4 P Neat, Lane 19: BacOvi S4 P 1:100, Lane 20: Sheep, Lane 21: Cow, Lane 22: -C, Lane 23: BacOvi S5 P Neat, Lane 24: BacOvi S5 P 1:100, Lane 25: Sheep, Lane 26: Cow, Lane 27: -C, Lanes 28, 29: Blank, Lane 30: HLV.
4.3.4 Establishment of Standard Curves for the BacOvi qPCR Assays

The concentrations for each BacOvi qPCR assay standard were determined by calculating what percentage of the plasmid was specific to each assay. Calculation for BacOvi qPCR assays (explanation shown is specific for BacOvi S1 and all other assays followed the same method):

The proportion (percentage) of the plasmid that the target makes was calculated by dividing the size of the target amplicon by the size of the total plasmid (vector + insert):

BacOvi S1: Total plasmid = 4627 bp
Target = 71 bp

\[
\frac{71}{4627} \times 100 = 1.53\%
\]

Therefore the target was 1.53% of the plasmid. The concentration of the plasmid was 36.5 ng/μl meaning 1.53% of this was 0.56 ng/μl.

The following equation was used to convert the concentration of the target template to copies/μl:

\[
6.023 \times 10^{14} \text{(Da/ng)} \times \text{concentration of plasmid (ng/μl)} \\
\text{Relative molecular mass (g/mol)}
\]

\[
6.023 \times 10^{14} \text{(Da/ng)} \times 0.56 \text{ ng/μl} \\
71 \text{ bp} \times 660 \text{ Da} \\
= 7.20 \times 10^9 \text{ copies/μl}
\]
BacOvi S2: Total plasmid = 4627 bp
Target = 101 bp

\[
\frac{101}{4627} \times 100 = 2.18
\]

2.18% of 36.5 ng/μl = 0.8 ng/μl

\[
6.023 \times 10^{14} (\text{Da/ng}) \times 0.8 \text{ ng/μl} \
101 \text{ bp} \times 660 \text{ Da}
= 7.23 \times 10^9 \text{ copies/μl}
\]

BacOvi S3: Total plasmid = 4622 bp
Target = 62 bp

\[
\frac{62}{4622} \times 100 = 1.34
\]

Concentration of the plasmid for BacOvi S3, S4 and S5 = 48.5 ng/μl

1.34% of 48.5 ng/μl = 0.65 ng/μl

\[
6.023 \times 10^{14} (\text{Da/ng}) \times 0.65 \text{ ng/μl} \
62 \text{ bp} \times 660 \text{ Da}
= 9.55 \times 10^9 \text{ copies/μl}
\]

BacOvi S4: Total plasmid = 4622 bp
Target = 66 bp

\[
\frac{66}{4622} \times 100 = 1.43\%
\]
1.43% of 48.5 ng/μl = 0.69 ng/μl

\[
6.023 \times 10^{14} \text{(Da/ng)} \times 0.69 \text{ ng/μl}
\]
\[
66 \text{ bp} \times 660 \text{ Da}
\]
\[
= 9.55 \times 10^9 \text{ copies/μl}
\]

BacOvi S5: Total plasmid = 4622 bp
Target = 84 bp

\[
\frac{84 \times 100}{4622} = 1.81\%
\]

1.81% of 48.5 ng/μl = 0.88 ng/μl

\[
6.023 \times 10^{14} \text{(Da/ng)} \times 0.88 \text{ ng/μl}
\]
\[
84 \text{ bp} \times 660 \text{ Da}
\]
\[
= 9.56 \times 10^9 \text{ copies/μl}
\]

All five BacOvi qPCR assays amplified their standard curves successfully as shown in Figure 4.5. BacOvi S1, S3, S4 and S5 had good efficiencies of 107%, 103%, 104%, and 104% respectively (Table 4.3). BacOvi S2 had a poor efficiency of 111%. All five assays had a good \(R^2\) value of 0.999 (Fig. 4.5). The Ct cutoffs ranged from 28 to 32 as illustrated in Table 4.3.
Figure 4.5 Standard curves of measured Ct values versus log plasmid marker concentration for the five newly developed BacOvi Assays S1-S5. a: BacOvi S1, b: BacOvi S2, c: BacOvi S3, d: BacOvi S4, e: BacOvi S5.
Table 4.3 Novel sheep–specific *Bacteroidales* assay standard curves, amplification efficiency, and Ct cutoff.

<table>
<thead>
<tr>
<th>Sheep–specific Bacteroidales assay</th>
<th>Standard Curve</th>
<th>Amplification Efficiency (%)</th>
<th>Ct cutoff</th>
</tr>
</thead>
<tbody>
<tr>
<td>BacOvi S1</td>
<td>$y = -3.159 + 40.46$</td>
<td>107</td>
<td>28</td>
</tr>
<tr>
<td>BacOvi S2</td>
<td>$y = -3.088 + 40.27$</td>
<td>111</td>
<td>30</td>
</tr>
<tr>
<td>BacOvi S3</td>
<td>$y = -3.245 + 41.55$</td>
<td>103</td>
<td>31</td>
</tr>
<tr>
<td>BacOvi S4</td>
<td>$y = -3.235 + 41.90$</td>
<td>104</td>
<td>32</td>
</tr>
<tr>
<td>BacOvi S5</td>
<td>$y = -3.234 + 41.82$</td>
<td>104</td>
<td>31</td>
</tr>
<tr>
<td>BacOvi S6</td>
<td>$y = -3.128 + 39.73$</td>
<td>109</td>
<td>29</td>
</tr>
</tbody>
</table>
4.3.5 Application of qPCR Assay to Various Faecal Extracts

The BacOvi qPCR assays (S1-S5) were applied to sheep, cow, human, horse, goat, and pig faecal extracts in three dilutions; neat, 1:10, and 1:100 (Fig. 4.6). BacOvi S1 and S2 did not amplify any of the faecal extracts (results not shown). BacOvi S4 and S5 did not amplify horse, pig or human but did amplify sheep, cow and goat at very similar Ct values (results not shown). The BacOvi S3 assay had a Ct cutoff of 31. It detected all three dilutions of sheep, cow and goat with a difference of 6 Ct between sheep and cow, and sheep and goat (Fig. 4.6). The assay did not detect human, horse or pig at any dilution. In order to try increase the assay specificity, the annealing temperature was increased from 60°C to 65°C and the assay was renamed BacOvi S6. The standard curve was amplified successfully as illustrated in Fig. 4.7. The Ct cutoff went from 31 to 29. The R² value stayed at 0.999 (Fig. 4.7) and the amplification efficiency (Table 4.3) changed from 103% to 109%. The assay subsequently detected sheep faecal extracts at neat and 10⁻¹ dilutions but did not detect sheep 10⁻² or any dilutions for cow, human, horse, goat or pig faecal extracts (Fig. 4.8). The annealing temperature was increased from 65°C to 67.5°C to try increase specificity further but the R² value and the amplification efficiency were both poor (results not shown) demonstrating that 65°C is the optimum annealing temperature for BacOvi S6.
Figure 4.6 Amplification curves for BacOvi S3 assay tested against various faecal samples. Sheep (S), cow (C), human (Hu), horse (H), goat (G) and pig (P) were assayed in the following dilutions; Neat (N), 1:10 (10^{-1}), and 1:100 (10^{-2}). The Ct cutoff was 31 (red line) and the threshold was 0.024 (blue line).
Figure 4.7 Standard curve of measured Ct values versus log plasmid marker concentration for BacOvi S6. The annealing temperature was increased from 60°C to 65°C.
Figure 4.8 Amplification curves for BacOvi S6 assay tested against various faecal samples. Sheep (S), cow (C), human (Hu), horse (H), goat (G) and pig (P) were assayed in the following dilutions; Neat (N), 1:10 ($10^{-1}$), and 1:100 ($10^{-2}$). The Ct cutoff was 29 (red line) and the threshold was 0.024 (blue line).
4.4 Discussion

Sheep faecal pollution can contain a high number of zoonotic bacteria, protozoa, and nutrients that can impair water quality and impact on human health and the environment (Arnscheidt et al., 2007; Devane et al., 2005; McGoll and Gibson, 1979; Prüss, 1998). There are assays available for the detection of cow faecal pollution but studies have shown that many of these assays show cross-amplification with sheep (Bernhard and Field, 2000b; Dorai-Raj et al., 2009). To date, no sheep-specific Bacteroidales PCR or qPCR assay has been designed. The aim of this research was to use subtractive hybridisation to develop such an assay.

Subtractive hybridisation is a complex molecular tool that can differentiate between highly similar gene fragments to find a specific sequence (Dick et al., 2005; Zwiglmaier et al., 2001). Cow faecal extracts were used as one of the subtracters as it was important that the developed assay could establish between two species with similar diets and therefore similar faeces. It was especially important that the faecal samples collected were from fields that only had cows or sheep so as to establish specific sequences and not introduce cross-contamination at the start of the experiment. Faecal samples were taken from various parts of Ireland to ensure diversity between samples. Human DNA was used as the second subtracter in order to ensure there was no cross-amplification between sheep and human sequences. The use of the Bacteroidales-specific 16S rRNA forward primer and the bacterial 16S rRNA reverse primer allowed for a large amount of sequence data to be collected from which the sheep-specific primers and probes were designed. A large amount of sequence data for sheep, cows and humans was downloaded from BLAST and added to the data obtained from this research to build a library to ensure that the sequences specific only to sheep, and that did not show any similarity to cow or human, were selected.

The sheep-specific Bacteroidales PCR assays designed in this study, BacOvi S1-S5, successfully amplified their positive controls. When they were applied to sheep and cow faecal extracts, only two of the assays, BacOvi S3 and S5, amplified the faecal extracts but unfortunately they did not differentiate between sheep and cow. As
qPCR is more sensitive and specific due to the addition of a probe, sequences were selected to design hydrolysis probes (Mackay, 2004). The standard curves amplified substantially for each of the five assays. The efficiencies varied between the five assays with BacOvi S2 having the poorest efficiency. Sheep, cow, human, horse, goat and pig were chosen as the faecal extracts to test the specificity of the qPCR assays as each of these species can impact on water in agricultural areas of Ireland and other countries (Dorai-Raj et al., 2009; Kildare et al., 2007; Mieszkin et al., 2009). The BacOvi S3 qPCR assay showed the most potential. It did not detect human, horse or pig faecal signals. It detected all three dilutions of sheep faeces tested but it also detected the neat and the 1:10 dilution for cow and goat. There were 6 Ct values between sheep neat and cow neat, and sheep neat and goat neat. It was decided to increase the temperature in order to increase assay specificity thus the annealing temperature went from 60°C to 65°C and the assay was renamed BacOvi S6. The efficiency changed slightly but was still suitable. The assay detected sheep neat and 1:10 dilution but not detect sheep 1:100 or any of the three dilutions for cow, human, horse, pig or goat. These results illustrate that the BacOvi S6 qPCR assay is capable of successfully differentiating between two similar species such as sheep and cow, and two even more similar species such as sheep and goat. The assay does not show any cross-amplification with human, horse or pig. This demonstrates that this assay has significant potential for use in an agricultural area to determine if the source of faecal pollution is sheep. In addition, this could be used in conjunction with the cow-specific Bacteroidales assays that show cross-amplification with sheep. If the cow-specific assay detects faecal contamination than BacOvi S6 could be applied to those samples. If BacOvi S6 detects sheep faecal contamination, then the source is sheep. If it does not, the source could be deemed to be cow. If the source can be established to be sheep or cow, best management practices such as mitigation measures could be put in place in order to limit the impact of this faecal pollution on water.

For this study, the BacOvi S6 assay has been only applied to faecal samples but in order to establish fully its use in microbial source tracking, the assay needs to be tested on water samples and spiked water samples. The study used subtractive
hybridisation to develop a molecular assay that can accurately distinguish between sheep and cow. When tested to its full potential the BacOvi S6 assay could be imperative in establishing the source of sheep faecal pollution in water and effectively aiding in the protection of human health and the environment.
4.5 Conclusions

The research conducted in this study showed that:

- Subtractive hybridisation resulted in specific DNA sequences for sheep
- BacOvi S6 assay successfully differentiates between sheep and cow, goat, human, horse and pig faecal samples
- The assay needs to be tested on water samples and spiked water samples

The preliminary results from the BacOvi S6 assay demonstrate that the assay has the potential to be used in MST.
Chapter IV

4.6 References


Chapter IV


www6.appliedbiosystems.com/support/techtools/calc/

www.basic.northwestern.edu/biotools/oligocalc.html

www.mbio.ncsu.edu/bioedit/bioedit.html

www.teagasc.ie/sheep/

www.thesheepsite.com


Chapter V

Concluding Remarks
5.1 Overview

Faecal contamination of water poses a major risk to human health and the environment. Methods used to determine this contamination have evolved throughout the years to provide more rapid and accurate results. As time goes on, the analysis of spot samples of water has expanded to an integrated catchment approach to fully characterise and evaluate all of the factors contributing to water pollution. Applying a “toolbox” approach to this strategy results in a clearer picture of water contamination and allows for informed decisions to be made about mitigation techniques and best management practices in order to protect water quality.

The main objective of this study was to combine a multi-tiered approach of using faecal indicator organisms (FIO), chemical and hydrological parameters, with catchment research and other tools such as microbial source tracking (MST) in order to determine the source of faecal pollution in two contrasting catchments in Ireland. The results presented in this thesis have utilised and built on years of significant information established by research on water quality, with the aim of furthering this knowledge and protecting the precious resource, which is essential for all life.

5.1.1 Targeted High- and Low- Flow Events in Spring/Summer and Autumn/Winter

The results presented in Chapter II illustrated that important information relevant to water quality protection was produced by targeting high- and low-flow events, across all seasons. Despite Grassland D having more potential sources of *E. coli*, Arable B returned higher loads of *E. coli* in all events sampled. Buffering of *E. coli* transfer between the soil and the spring was evident in Grassland D, deeming it to be a “transport limited system”. The specific vulnerability of this catchment for *E. coli* transfers was low as a result of hydrological response and not source pressure. In Grassland D higher loads of *E. coli* were evident when the event was in the rising phase which yields significant information for best management practices. It shows that extra caution must be exercised regarding production of drinking water in this karst catchment during rising phases of discharge. Much higher loads per unit discharge in Arable B, imply increased concentrations of *E. coli* and higher magnitudes overall, and this would illustrate that Arable B is a less transport limited system. The levels of *E. coli* were higher overall in Arable B and the hourly load remained high even at low discharges, possibly as a result of point sources that could be of non-agricultural origin. This
highlights that measures must be put in place to order to combat this faecal contamination issue during high- and low-flows. As this water drains into the nearby coastal area and could impact the bathing water, it is especially important that the problem is rectified. The water in Arable B might not suitable as potable source either due to the high levels of \textit{E. coli} present.

\subsection*{5.1.2 Flow Pathway Separation}
Use of the novel pathway separation technique, Loadograph Recession Analysis (LRA), used in conjunction with \textit{E. coli} loads, in Chapter II resulted in the separation of the various flow pathways and the determination of which pathways were responsible for the transport of \textit{E. coli} loads. It showed that high loads were transferred during the faster nutrient (TP) flow paths, which were quick flow in Arable B, and conduit and large fissure flow in Grassland D. This observation has significant implications for best management practices as it illustrates that the mitigation techniques should target surface flow, which is the quick flow pathway, to reduce the transfer load of \textit{E. coli}.

\subsection*{5.1.3 Cross-amplification of Host-Specific \textit{Bacteroidales} Assays}
Part of Chapter III entailed checking the host-specificity of the \textit{Bacteroidales} assays. The human-specific assay, BacHum, showed no cross-amplification with cow, sheep, goat, pig or horse. However, the bovine-specific assay, BacBov, did not detect human, horse or pig, but it did show cross-amplification with sheep and goat. These results meant that while BacHum could be used to successfully establish the source of human faecal pollution, BacBov could only determine agricultural ruminant faecal pollution. This has ramifications for catchment management as it cannot be clearly defined if the source of faecal contamination, in an agricultural area, is cow or sheep making it difficult to institute suitable mitigation techniques.

\subsection*{5.1.4 Determination of Faecal Source Using Multi-tiered Approach}
In Chapter III, the universal \textit{Bacteroidales} assay, BacUni, and the FIO, \textit{E. coli} were used to determine if faecal contamination was present. Phosphorus (TP) was analysed continuously in the spring, Grassland D, and the outlet, Arable B. Significant correlations were seen between BacUni, \textit{E. coli} and TP indicating that faecal pollution was evident. The source of the faecal pollution was then differentiated into human, using BacHum, and agricultural ruminant, using BacBov. Human and agricultural faecal pollution was found to be present in
both catchments, albeit at lower levels in Grassland D. In Arable B, the levels of human contamination were higher than agricultural contamination during low flow, and also exhibited diurnal patterning possibly related to septic tank usage. For high-flow, higher levels of agricultural pollution were evident. This information will be useful for making decisions about mitigation techniques.

5.1.5 Establishment of Pathways Which Transfer Faecal Contamination

Building on the LRA from Chapter II, the pathways were separated from both catchments into quick flow (conduit and large fissure flow), interflow (medium fissure flow), and delayed flow (small fissure flow) and the MST data were added to define the pathways through which source contamination was being moved in Chapter III. This multi-tiered method garnered some significant information. It illustrated that for Arable B, the less transport limited system transferred more faecal waste overall which originated from both agricultural ruminant and human sources. Agricultural ruminant pollution dominated the quick flow in Arable B. Including important land use knowledge showed that some of the events coincided with the closure of the slurry spreading period so it is probable that this contamination was from slurry. This important information allows for a more targeted approach to be taken for best management practices, such as installing riparian buffer zones. Research across various catchments demonstrated that riparian buffer zones reduced the movement of agricultural pollutants and faecal bacteria into water. Another important mitigation technique for agriculture is not to allow livestock to have access to a river as this increases the risk of faecal contamination of the water. As agricultural ruminant faecal waste was also detected at low flows, this illustrates that this source was not disconnected during these flow phases. It also shows that the source is not just dependent on storm quick flows for transfer from land to water.

In Arable B, delayed flow or events with low rates of discharge were dominated by human faecal pollution, which is associated with defective septic tanks. This is especially true as one of the events showed diurnal patterns of human faecal contamination, which would indicate flows associated with septic tank usage. Therefore management practices must be put in place to audit septic tanks for inefficient wastewater management in order to make the necessary improvements. This would include upgrading the percolation (soak away) areas to be more
efficient by applying vertical peat biofilters or combined willow and reed beds. The increase, rather than dilution, of human faecal contamination during quick flows, indicated the possibility that wastewater sources are either disconnected to storm drains or concentrate in ephemeral ditches and are then flushed during rainfall events. If human faecal pollution is transferred via overland flow, such as in a high flow event, then riparian buffer zones might help to reduce the amount of faecal contamination impacting on the water. Being able to establish the flow pathways by which faecal pollution is transferred results in a much more targeted approach to best management practices. Mitigation techniques must be established to reduce the level of not just agricultural faecal contamination but human faecal contamination as well. This would allow for better compliance with standards set by the European Union Water Framework Directive, the Bathing Directive and water safety plans.

5.1.6 Design and Application of Sheep-Specific Bacteroidales Assays
As demonstrated in Chapter III, there is cross-amplification evident, between cow, sheep and goat, for the BacBov molecular assay. Subtractive hybridisation was used to design and establish a sheep-specific PCR and qPCR assay. The PCR assay was not effective in distinguishing between sheep and cow but the advent of a probe resulted in a more specific qPCR assay. The qPCR assay could successfully differentiate between sheep, and cow, goat, horse, pig and human faecal samples. This assay will be of great use for determining whether the source of faecal pollution is cow or sheep.

5.1.7 Conclusions
The ability to separate flow pathways and assign the pathway that transfers the majority of human or agricultural faecal pollution will allow for informed decisions to be made for the policy making and practices that are imperative for protecting water quality. This study demonstrated that transfer pathways of FIO could be assigned using the LRA approach. Moreover, the source or sources (human or ruminant) of the faecal material could also be indicated through the deployment of MST assays. If agricultural ruminant faecal pollution can be further differentiated into sheep and cow, this will also add significant instruction to the policy makers.

Comprehensive water quality information on the sources of faecal pollution and how faecal material is transferred during different hydrological phases will ensure that more successful
policies aimed at improving water quality can be developed and that scarce economic resources can be targeted effectively. Preventing faecal contamination from gaining access to waterways might be more beneficial than treating water with expensive techniques to remove the contamination and to make it fit for human consumption. This might prove especially true in countries that cannot afford to treat their water with costly, energy-intensive methods. Building on the Millennium Development Goals (MDG), the United Nations (UN) have established a post-2015 Global Goal for Water with the aim of achieving universal access to safe drinking water and basic sanitation and establishing sustainable use of water resources. They plan to achieve this through initiatives such as sustainable sanitation and ending open defecation (www.unwater.org). As weather patterns continue to shift as a result of climate change, and the global human population continues to grows, putting a greater demand on water supply, it will become more imperative to conserve, improve and protect water (Vorosmarty et al., 2000). The highly original contribution to knowledge provided by this PhD thesis can greatly help to achieve this. From all the inter-disciplinary research conducted in this PhD, it can be concluded that methods such as MST added to a multi-tiered approach and applied in a catchment, can efficiently accumulate important information. MST has the potential to be very useful in helping to achieve the water quality standards established by the EU WFD and BD. Before MST can be used in this way, limitations such as survival rates needs to be addressed. The information presented by this PhD thesis could be used for establishing the risk to human health and putting in place the necessary measures to prevent that risk.
Chapter V

5.2 Future Perspectives and Recommendations

The results presented in this thesis have provided important information for the continuing protection of water quality, however more research is still needed.

5.2.1 Risk Assessment Models
Models such as SWAT (Soil and Water Assessment Tool) can simulate a broad array of best management practices (BMP) (Lam et al., 2011), should be used on both catchments to assess the suitability of various BMPs, so as to establish which one would be most effective. Other risk assessment models, such as QMRA (Quantitative Microbial Risk Assessment) would may prove useful if applied to the two catchments (Soller et al., 2010). This model will provide information on the possible risk that microbial pathogens could constitute to human health. If the models are successful in both catchments then this approach could be applied to other catchments. QMRA will be especially useful in Irish catchments that are impacted by the zoonotic waterborne pathogen, Cryptosporidium (Cummins et al., 2010).

5.2.2 Broadening Numbers of Indicators Analysed
Faecal stanols and sterols, notably the coprostanol/cholesterol ratio, have proved successful in differentiating between human, carnivore and herbivore faecal pollution (Shah et al., 2007). It might be of benefit to use it in conjunction with the Bacteroidales molecular assays in a “toolbox” approach in both catchments and other catchments in Ireland.

5.2.3 Survival Rates of Bacteroidales in Irish Catchments
Studies have been conducted on survival rates of Bacteroidales in water and have shown that they can survive for up to 6 days depending on oxygen stress and temperature (Bernhard and Field, 2000; Sokolova et al., 2011). No similar study has been carried out in Ireland. Such a study would be very beneficial for testing the validity of these assays in Irish water as a result of the different environmental conditions. This would be important if Bacteroidales was to be used as a tool for regulatory compliance. Survival rate analysis of BacHum in septic tanks and their leachate would provide practical information for dealing with the issues of defective septic tanks.
5.2.4 Application of Sheep-Specific Assays to Water Samples

The BacOvi assay designed, as part of this PhD, needs to be applied to water samples and evaluated in an environmental context. The first step would be to apply it to water samples spiked separately with sheep, cow and goat faecal samples to test specificity and sensitivity. If this is successful, then the assay should be applied to water samples that are spiked with pooled samples of sheep, cow and goat faecal extracts. If the assay proves effective in only detecting sheep faecal contamination then the assay should be applied to environmental samples from areas with known sheep faecal contamination and areas with no such contamination.

5.2.5 Applying Targeted Approach to Catchment on Trial Basis

The majority of results presented by this PhD thesis strive towards a more targeted approach to best management practices (BMPs). Taking the information provided by this thesis and applying it to a catchment, with the permission of the farmers and land owners, on a trial basis could yield some interesting results about the effectiveness of the chosen BMPs. As the characteristics of the catchment would already have been established, applying BMPs suitable to that catchment should not be very difficult or expensive. If the trial is successful in one catchment, other catchments could then be trialled in the same way with the overall aim being to apply this method to watersheds across Ireland.

5.2.6 Standardisation of these Methods in Catchments

While there have been many successful techniques established for testing water in research laboratories, not many of these techniques will be used by regulatory agencies for water quality analysis. Standardisation of these methods, across various laboratories in Ireland, is a necessity, not only for accurate water quality analysis, but for effective promotion of their use to regulatory agencies. Along with this is the need to know the environment from which the water sample derives. Continuing research of catchments that have been efficiently characterised in terms of their hydrology, microbiology, geology, topography, human settlement and agricultural land-use will be imperative for practical and successful protection of water quality. The laboratory techniques will only be verified as beneficial when they have been thoroughly analysed in “real world” environments.
5.3 References


www.unwater.org