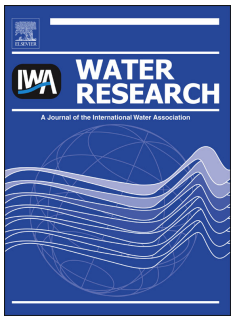
Accepted Manuscript



Title: Specificity and sensitivity evaluation of novel and existing Bacteroidales and Bifidobacteria specific PCR assays on feces and sewage samples and their application for fecal source tracking in Ireland

Authors: Siobhán Dorai-Raj, Justin O' Grady, Emer Colleran

PII: S0043-1354(09)00563-6

DOI: 10.1016/j.watres.2009.08.050

Reference: WR 7609

To appear in: *Water Research*

Received Date: 27 February 2009

Revised Date: 24 August 2009

Accepted Date: 30 August 2009

Please cite this article as: Dorai-Raj, S., Grady, J.O., Colleran, E. Specificity and sensitivity evaluation of novel and existing Bacteroidales and Bifidobacteria specific PCR assays on feces and sewage samples and their application for fecal source tracking in Ireland, Water Research (2009), doi: 10.1016/j.watres.2009.08.050

Specificity and sensitivity evaluation of novel and existing *Bacteroidales* and *Bifidobacteria* specific PCR assays on feces and sewage samples and their application for fecal source tracking in Ireland.

Siobhán Dorai-Raj[a, *], Justin O' Grady[b], Emer Colleran[a].

[a]Environmental Microbiology Research Laboratory, Microbiology, School of Natural Sciences, [b]Molecular Diagnostics Research Group, National Centre for Biomedical Sciences, National University of Ireland, Galway, University Road, Galway, Ireland.

*Corresponding author. Tel.; +35391493864; fax: +35391494598. E-mail address: Siobhan.dorairaj@nuigalway.ie

Abstract

Three novel ruminant-specific PCR assays, an existing ruminant-specific PCR assay and five existing human-specific PCR assays, which target 16S rDNA from *Bacteroidales* or *Bifidobacteria*, were evaluated. The assays were tested on DNA extracted from ruminant (n = 74), human (n = 59) and non-ruminant animal (n = 44) sewage/fecal samples collected in Ireland. The three novel PCR assays compared favourably to the existing ruminant-specific assay, exhibiting sensitivities of 91 - 100% and specificities of 95 - 100% as compared to a sensitivity of 95% and specificity of 94%, for the existing ruminant-specific assay. Of the five human-specific PCR assays, the assay targeting the *Bifidobacterium catenulatum* group was the most promising, exhibiting a sensitivity of 100% (with human sewage samples) and a specificity of 87%. When tested on rural water samples that were naturally contaminated by ruminant feces, the three novel PCR assays tested positive with a much greater percentage (52 - 87%) of samples than the existing ruminant-specific assay (17%). These novel ruminant-specific assays show promise for microbial source tracking and merit further field testing and specificity evaluation.

1. Introduction

The need to determine the source of fecal contamination of water has led to the development of various fecal source identification methods, a relatively new field commonly known as microbial source tracking (MST). The methods of MST used in this study are based on molecular detection of host-specific strains of bacteria from the order *Bacteroidales* and the genus *Bifidobacterium*. Both these groups of bacteria are abundantly present in feces and several researchers have concluded that some strains of the microorganisms are confined to specific hosts (Fiksdal et al., 1985; Gavini et al., 1991; Resnick and Levin, 1981). Host-specific *Bacteroidales* molecular assays have been successfully used as a method of MST by a number of research groups (Gawler et al., 2007; Kildare et al., 2007; Seurinck et al., 2005). Katherine Field and colleagues, in particular, have performed extensive research into the use of *Bacteroidales* 16S rDNA-based PCR assays for MST (Bernhard and Field, 2000a; b; Bernhard et al., 2003; Field et al., 2003; Shanks et al., 2006). The use of molecular methods to detect host-specific species of *Bifidobacteria* is not as well studied as detection methods for *Bacteroidales* but a number of studies have been performed (Bernhard and Field, 2000a; Blanch et al., 2006; Bonjoch et al., 2004; Lynch et al., 2002).

Fecal contamination of rural water supplies in Ireland is common (EPA, 2008) and apart from human feces, the main sources of contamination are cow and sheep feces (both ruminants). The principal aim of this study was the development and evaluation of novel ruminant-specific PCR assays and the use of these assays for MST on contaminated water samples collected from rural water supplies in Ireland. To develop the assays, novel

ruminant-associated *Bacteroidales* 16S rDNA sequences were identified by terminal restriction fragment length polymorphism (TRFLP) analysis of human and ruminant fecal/sewage samples. These sequences were exploited for the design of ruminant-specific PCR primers. A secondary aim of the study was the evaluation of a number of existing putatively host-specific PCR assays on Irish fecal/sewage reference samples and the application of the assays for MST on Irish naturally contaminated water samples.

The human-specific (HF183F & Bac708R) and ruminant-specific (CF128F & Bac708R) PCR assays, designed by Bernhard and Field in 2000 for MST (2000a; b), were the first set of *Bacteroidales* PCR assays chosen for evaluation in this study. The second set were designed to detect *Bacteroides thetaiotaomicron* (BT 1 & 2) and *Bacteroides vulgatus* (BV 1 & 2) and were developed by Wang et al. (1994; 1996) for the detection of anaerobic bacteria in human and animal fecal samples. The authors found that both *Bacteroides thetaiotaomicron* and *Bacteroides vulgatus* were present in large numbers in adult human samples and at lower numbers or absent in animal samples (Wang et al., 1996).

Two *Bifidobacteria*-specific PCR assays designed by Matsuki et al. (1999; 1998) for microbial ecology studies of the human gastrointestinal tract were also chosen for evaluation in this study. The first assay was designed to detect *B. adolescentis* (BiADO 1 & 2) and the second was designed to detect *B. catenulatum* and *B. pseudocatenulatum* (BiCAT 1 & 2). These species were found in a high percentage of the human fecal samples tested by Matsuki et al (1999).

All the assays developed and/or evaluated in this study are conventional PCR assays as opposed to real-time PCR assays. A number of *Bacteroidales*-specific real-time

PCR assays have been developed which have the potential to detect and quantify host-specific targets in water (Kildare et al., 2007; Layton et al., 2006; Okabe et al., 2007; Reischer et al., 2007; Reischer et al., 2006; Stricker et al., 2006). However, none of these assays were available for evaluation during the timeframe of this study. The advantages of real-time PCR include increased specificity, sensitivity and the ability to accurately quantify the target. However, since real-time PCR technology is expensive, requires additional expertise and is not generally available in Irish local authority environmental monitoring laboratories, this development and evaluation of conventional PCR assays is still worthwhile.

2. Materials and methods

2.1 Sample collection and determination of *E. coli* densities

Human sewage samples (untreated primary effluent, n = 33) were collected from two different wastewater treatment works in Co. Galway, Ireland. Ruminant (n = 74) and non-ruminant (n = 44) animal fecal and slurry samples were collected from various farms in Co. Galway. The ruminant samples consisted of cow (n = 25), sheep (n = 39), deer (n = 1) and goat (n= 4) fecal samples and five cow slurry samples. The non-ruminant samples consisted of horse (n = 12), donkey (n = 2), dog (n = 2), goose (n = 1), chicken (n = 2), pet pig (n= 2) and farmed pig (n = 8) fecal samples and 15 pig slurry samples. Raw and piped water samples were collected every two weeks from three frequently contaminated rural drinking water supplies in Co. Galway over a six month period. *E. coli* concentrations (most probable number [MPN] per 100-ml of water sample) were

116 measured using Colilert-18 and Quanti-tray/2000 (Idexx, Westbrook, ME). All samples
117 were collected with sterile utensils, placed in sterile containers and transported on ice.
118 Water samples (1-L) were filtered within six hours of collection using 0.2-µm-pore-size
119 cellulose nitrate filter membranes (Sartorius AG, Goettingen, Germany). Turbid water
120 samples were pre-filtered with a 2.7-µm-pore-size glass fibre filter membrane to remove
121 debris before filtration with aforementioned 0.2-µm-pore-size filter membranes.
122 Fecal/sewage samples and filter membranes were stored at -80ºC. DNA from 26
123 individual human stool samples, which were donated by healthy human adults, was
124 kindly provided by the Microbiology Department, University College Cork, Ireland.
125

126 **2.2 DNA extraction**

127 DNA was extracted from 20 mg of sewage sediment, from fecal samples and from filter
128 membranes using the Powersoil™ DNA Isolation Kit (MoBio, Carlsbad, CA) following
129 the manufacturer's protocol, modified to include the use of four washes with solution C5
130 and a 10 min incubation at 70°C after the addition of solution C1. A mock DNA
131 extraction was included each time DNA extractions were performed to test for
132 contamination of kit components. To confirm DNA samples were free of PCR inhibitors,
133 10 ng of DNA was tested using the *Bacteroidales*-specific primer pair Bac32F & Bac
134 708R (Table 1). Amplification indicated the absence of inhibitors.
135

136 **2.3 Clone library construction**

137 The Bac32F & Bacto1080R primer pair (Table 1) was used to amplify a ~ 1060 bp
138 fragment of 16S rDNA from DNA (10 ng) extracted from six cow fecal samples, five

sheep fecal samples and five human sewage samples. Each 50-µl PCR mixture contained: 1 × *Taq* polymerase buffer, 200µM dNTP (dATP, dCTP, dGTP, dTTP), 12.5 pmol of each primer, and 1.25 U *Taq* DNA polymerase (Sigma-Aldrich, St. Louis, MO). Thermal cycling was performed in a Mastercycler personal PCR machine (Eppendorf, Hamburg, Germany) as follows: an initial denaturation step at 94^{o}C for 3 minutes, 30 cycles consisting of 94^{o}C for 30 s, 58^{o}C for 1 min, and 72^{o}C for 1 min, followed by a final 5-min extension at 72^{o}C. A positive control (fecal DNA from the target source previously found positive or plasmid DNA containing the target insert) and a negative no-template control were included in every experiment and all PCR assays were performed in triplicate. After purification ('High Pure PCR Product Purification Kit', Roche Diagnostics, Mannheim, Germany) and quantification of the PCR products (PicoGreen dsDNA Quantitation Kit, Molecular Probes, Eugene, OR) three pools of PCR products, each representing one of the host species, were cloned, (TOPO TA Cloning® kit, Invitrogen, De Schelp, Netherlands). Plasmid DNA was extracted from 100 clones per library (QIAprep Spin Miniprep Kit, Qiagen, Hilden, Germany) and clones were resolved into operational taxonomic units (OTUs) using amplified rDNA restriction analysis (ARDRA) with both *Hae*III and *Alu*I (Roling and Head, 2005). A representative clone from each OTU group identified was sequenced by MWG BIOTECH AG (Ebensburg, Germany). The sequence data were checked for chimeric properties using Chimera Check on the RDP II website and using the Mallard (Ashelford et al., 2006) and Pintail (Ashelford et al., 2005) programs of the Bioinformatic Toolkit website (http://www.bioinformatics-toolkit.org/index.html). Non-chimeric sequences were aligned using the multiple sequence alignment program ClustalW

(http://www.ebi.ac.uk/Tools/clustalw/index.html). Default parameters were used for ClustalW and all other computer programs utilized. Aligned sequences showing > 97% similarity were treated as a single OTU (Okabe et al., 2007) and only one of each OTU was included in further analysis.

2.4 TRFLP analysis

DNA from the six cow and five sheep feces samples and six human sewage samples used in clone library construction was amplified as previously described (Section 2.3) this time using fluorescently labelled primers Bac32F (5'- hexachlorofluorescein [HEX] labelled) and Bacto1080R (5'- [6]-carboxyfluorescein [FAM] labelled). The PCR reaction was carried out in triplicate for each sample and the products were pooled and purified as described in Section 2.3. PCR products were digested with *Hae*III and fragment sizes were measured by polyacrylamide gel electrophoresis in an automated ABI Prism 310 Genetic Analyzer using the GS2500 TAMRA size marker (performed by Gene Analysis Service GmbH, Berlin, Germany). Eight of the samples were analyzed in duplicate to assess the reproducibility of the TRFLP profiles.

2.5 Ruminant-specific PCR primer design

Putative ruminant-specific clone sequences were aligned with all sequences from the human fecal DNA clone library and putative ruminant-specific PCR primers were designed adhering to general primer design guidelines (Dieffenbach et al., 1995). Primers

were checked with Premier Biosoft International's free online primer analysis program Netprimer (http://www.premierbiosoft.com/netprimer/index.html).

2.6 Evaluation of host-specific PCR assays

The host-specific PCR assays were tested on all the samples listed in Table 2, using the PCR reaction components and conditions detailed in Section 2.3. The PCR cycling conditions described by Field et al. (2003) for the HF183F & Bac708R primer pair and the CF128F & Bac708R primer pair were used without modification. The optimum annealing temperatures for the novel ruminant-specific PCR assays were determined empirically (data not shown) as were the annealing temperatures for the other existing PCR assays since these PCR assays were not originally developed for use in MST. The final annealing temperatures used are listed in Table 1. The sensitivity and the specificity of all the PCR assays was calculated using standard definitions (Gawler et al., 2007).

2.7 Evaluation of the sample limit of detection (SLOD) and method detection limit (MDL) of host-specific PCR assays

The SLOD of the assays was defined as the minimum dry weight of feces per filter membrane that could be detected using the assays. This was assessed by extracting DNA in triplicate from decimal dilutions of 1-L water samples to which 100 mg of cow feces or human sewage had been added. The MDL was defined as the minimum number of copies of the 16S rRNA gene template that could be detected using the ruminant-specific assays. For this method, plasmid DNA containing the target *Bacteroidales* 16S rDNA

fragment was decimally diluted to give a range of DNA from 10^9 to 1 copy of plasmid DNA per µl and tested with the PCR assays.

3. Results and Discussion

3.1 Host-specific sequence identification

3.1.1 Clone library analysis

The first step in host-specific sequence identification was the construction of clone libraries using DNA extracted from cow, sheep and human feces/sewage samples. One hundred clones from each library were resolved into OTUs and representative clones from each OTU were sequenced. All sequences were submitted to Genbank under the following accession numbers; cow clone library: EU573790 – EU573833; human clone library: EU573834 – EU573866; sheep clone library: EU573867 – EU573924.

3.1.2 TRFLP analysis

The next step in host-specific sequence identification was TRFLP analysis of the same DNA samples used to generate the clone libraries. Analysis of the TRFLP profiles performed in duplicate confirmed the reproducibility of the method. As illustrated by the representative TRFLP profiles presented in Figure 1, there were HEX-labelled ruminant-associated TRFs at 190 - 191 bp and 222 - 224 bp, and sheep-associated peaks at 105 - 106 bp, 110 bp and 146 bp. Figure 2 illustrates the two FAM-labelled ruminant-specific TRFs which were identified in the profiles, one at 69 - 70 bp and one at 81 bp.

Several cloned *Bacteroidales* sequences corresponding to the ruminant-specific TRFs were identified from the cow and sheep fecal DNA clone libraries. The lengths of experimentally determined TRFs, as compared to sequence-determined TRFs, were inaccurate by up to 4 bp (data not shown) which concurs with other studies (Bernhard and Field, 2000a; Clement et al., 1998; Pandey et al., 2007). The putative ruminant-specific sequences were used to design ruminant-specific primers as described in Section 2.5.

3.2 Host-specific PCR assay evaluation

3.2.1 Specificity and sensitivity evaluation

Following an initial evaluation of novel ruminant-specific PCR assays on a small number of fecal/sewage samples, three assays emerged as potentially useful; Bac32F & RumD1R (product 979 bp), Bac32F & RumD2R (product 997 bp) and RumB1F & BacPreR (product 714 bp) (Table 1). These three PCR assays were then evaluated using a full range of target and non-target fecal samples (Table 2).

In general, the sensitivity and specificity of the novel ruminant-specific PCR assays was high, ranging from 91 – 100% sensitivity and 95 – 100% specificity (Table 2). The novel ruminant-specific PCR assays compared well to the CF128F & Bac708R assay (which had a sensitivity of 95% and a specificity of 94%). Although none of the novel ruminant-specific PCR assays developed exhibited 100% sensitivity and 100% specificity, results were consistent with other studies where ruminant-specific PCR assays were developed or tested (Gawler et al., 2007; Gourmelon et al., 2007; Kildare et al., 2007).

All three novel ruminant-specific PCR assays amplified DNA from all of the 25 individual cow fecal DNA samples and from all or most of the other ruminant fecal samples (Table 2). Most importantly, none of the novel ruminant-specific PCR assays amplified DNA from any of human sewage or fecal samples and so could be used as a tool to differentiate between human and animal contamination.

The human-specific HF183F & Bac708R assay was 100% specific. However, the assay was positive for only three of the 26 individual human fecal samples (r = 12%) and only 23 of the 33 human sewage samples (r = 70%). This assay was tested on human sewage samples from four European countries, including Ireland, by Gawler et al. (2007). The sensitivity results obtained by these authors varied between 76% and 100%, with a sensitivity of 88% for Irish samples. There appears to be a degree of variation in the quantity of this target in Irish sewage samples when the results obtained in this study are compared to the results obtained by Gawler et al. This would suggest a need to validate this assay in the location it is to be used before its application as a MST tool.

The BV 1 & 2 PCR assay (Wang et al., 1994) and the BiADO 1 & 2 PCR assay (Matsuki et al., 1998) had similar sensitivities for individual human fecal samples (88% and 85% respectively) and both were 100% sensitive for human sewage samples (Table 2). These assays also had comparable specificities, amplifying DNA from many of the pig samples and at least one ruminant fecal sample.

The BiCAT 1 & 2 PCR assay (Matsuki et al., 1998) had a sensitivity of 46% for individual human fecal samples and a sensitivity of 100% for human sewage samples (Table 2). The low sensitivity for individual human fecal samples is not a significant drawback since from an environmental monitoring context, the ability to detect mixed

sewage samples is more important than the ability to detect individual human fecal samples. The assay did not test positive with any ruminant fecal samples and the only non-human fecal samples which tested positive were pig slurry samples, making it potentially the most useful human-specific assay tested for MST, since in many cases, pig fecal contamination can be ruled out as a potential source of contamination.

3.2.2 SLOD and MDL of host-specific PCR assays

The SLOD of all PCR assays was evaluated and the MDL was evaluated for the ruminant-specific assays only (because the development of ruminant-specific assays was the focus of this study).

The ruminant-specific PCR assay with the lowest SLOD was the Bac32F & RumD1R assay with a SLOD of 7.3×10^{-6} g (dry weight) of feces per filter (Table 3). The assays with the next lowest SLOD were Bac32F & RumD2R, followed by RumB1F & BacPreR and CF128F & Bac708R (Table 3).

The ruminant-specific PCR assays with the lowest MDL were RumB1F & BacPreR and Bac32F & RumD2R. Both of these PCR assays tested positive with 10^2 - 10^3 copies of target plasmid per 50-µl PCR reaction (Table 3). The MDL for the two other ruminant-specific PCR assays (Bac32F & RumD1R and CF128F & Bac708R) was tenfold higher; at 10^3 - 10^4 copies of the plasmid (Table 3). The MDL of the CF128F & Bac708R PCR assay was also performed in a study by Shanks et al. (2006). The authors reported that the assay routinely detected 10^2 copies of target plasmid DNA, which is a 10 to 10^2 fold lower MDL than determined in this study. This MDL variation is possibly due to the different PCR reagents, *Taq* polymerase and PCR thermal cyclers used. In a study

by Gawler et al. (2007) on the CF128F & Bac708R PCR assay, a detection limit of ~ 2.5 $\times 10^4$ copies of target plasmid DNA was determined, similar to the results obtained in this study. Inter-laboratory variation has been observed by other groups (Griffith et al., 2003) which emphasizes the need to establish standard operating procedures for MST methods and suggests that each new MST method should be validated in the laboratory where it is to be used.

The human-specific PCR assays with the lowest SLODs were the BiADO 1 & 2 and BV 1 & 2 assays (Table 3). These assays both had SLODs of 6.6×10^{-6} g (dry weight) of human sewage per filter. The BiCAT 1 & 2 assay had a SLOD of 6.6×10^{-4} g and the HF183F & Bac708R assay had the highest SLOD of 6.6×10^{-2} g (dry weight) of human sewage per filter. This is a relatively high SLOD which could mean the assay would only be useful in cases of heavily contaminated water. However, a number of other researchers have successfully used this assay in the field (Gourmelon et al., 2007; Lamendella et al., 2007; Shanks et al., 2006) so this high SLOD may be due to the particular sample set tested in this study.

3.3 Application of host-specific PCR assays on naturally contaminated rural water samples

Both the novel and existing host-specific PCR assays were applied as a method of MST to samples taken from contaminated rural drinking water supplies. All the rural water sources were located in pasture lands for cows and sheep and so were expected to be contaminated by ruminant feces. *E. coli* contamination in the rural water samples varied from 0 to 2203 *E. coli* per 100 ml but only water samples with greater than 50 *E. coli* per

318 100 ml (n = 23) were tested with the host-specific assays. The ruminant-specific PCR
319 assays with the highest detection rates were the Bac32F & RumD2R assay and the
320 RumB1F & BacPreR assay which were positive for ruminant DNA in 87% (20/23) of the
321 rural water samples tested. The Bac32F & RumD1R assay was positive for approximately
322 52% (12/23) of the rural water samples. The ruminant-specific PCR assay CF128F &
323 Bac708R assay was positive for ~ 17% (4/23) of the rural water samples. Possible
324 reasons for the superior performance of the novel ruminant-specific PCR assays include
325 the possibility that the novel assays detect microorganisms which persist for longer in the
326 environment or that the novel assays target an indigenous microorganism that is abundant
327 in local fecal pollution. The latter may indicate that methods developed in the region
328 where they are to be used may perform better than methods developed in other regions or
329 countries.

330 Two samples with relatively high levels of *E. coli* contamination (77 and 178 *E.*
331 *coli* per 100 ml) which were pre-filtered with 2.7 µm-pore-size glass fibre filter
332 membranes (Section 2.1) tested negative with all the ruminant-specific PCR assays. It is
333 possible that pre-filtration removed target species of *Bacteroidales* that may have been
334 attached to particles of manure or debris. However, these negative results could also be
335 explained by the inconsistent correlation observed in this study between the level of *E.*
336 *coli* contamination and PCR detection of *Bacteroidales* in the water (data not shown). A
337 study by Shanks et al. (2006) also found that there was poor correlation between *E. coli*
338 counts and presence of ruminant-specific *Bacteroidales.*

339 Two of the putatively human-specific assays (HF183F & Bac708R and BiCAT 1
340 & 2) did not amplify DNA from any of the rural water samples, while the other two

assays (BV 1 & 2 and BiADO 1 & 2) tested positive with five and three of the rural water samples, respectively. The HF183F & Bac708R and BiCAT 1 & 2 PCR assays are also the only two putatively human-specific assays that did not amplify DNA from any ruminant DNA samples (Table 3). It is possible that the BV 1 & 2 and BiADO 1 & 2 assays were detecting ruminant fecal contamination in the rural water samples rather than a human fecal source of contamination. Nonetheless, the possibility of contamination of the water by human waste cannot be ruled out since there may have been leaking septic tanks in the area.

Overall, the results of the testing of the contaminated rural water samples tentatively reveal, that as anticipated by land-use patterns, the main source of contamination of the raw water of these three rural drinking water supplies is ruminant in origin.

4. Conclusions

Of the five putatively human-specific published PCR assays evaluated in this study, the BiCAT 1 & 2 PCR assay, which targets *Bifidobacterium catenulatum* and *Bifidobacterium pseudocatenulatum*, shows most promise for use as a method of detecting human fecal contamination.

The ruminant-specific PCR assays developed in this study show good specificity, sensitivity, have low SLODs and MDLs and have been used to amplify putatively ruminant-specific *Bacteroidales* strains from naturally contaminated water samples. All of the assays developed in this study compared favourably to the CF128F & Bac708R

PCR assay. The novel ruminant-specific PCR assays show promise for use in MST studies but require more extensive evaluation both *in vitro* and in field studies before they could be employed as an unambiguous method of identifying ruminant fecal pollution. As mentioned in the introduction, all the PCR assays developed and/or tested in this study are conventional PCR assays as opposed to real-time or quantitative PCR (qPCR) assays. While qPCR offers the possibility of elucidating the quantities of different fecal inputs in a contaminated water source, for true quantification of fecal sources with qPCR, not only will the distribution of *Bacteroidales* in different types of feces need to be established, but also the persistence of host-specific *Bacteroidales*, the stability of their relative ratios in the environment and their resistance to waste treatment (Santo Domingo *et al.*, 2007). Nonetheless, it would be advantageous, in the future, to convert the ruminant-specific conventional PCR assays developed in this study into real-time PCR assays.

Acknowledgements

This research was funded by the Environmental Research Technological Development and Innovation Programme (ERTDI) 2000–2006 administered by the Irish Environmental Protection Agency (Project Ref: 2005/EH/CD/E). The programme is financed by the Irish Government under the National Development Plan 2000–2006. The authors wish to thank Owen Doherty for help with fieldwork and the Environmental Change Institute, National University of Ireland, Galway for the use of its facilities.

References

Ashelford, K.E., Chuzhanova, N.A., Fry, J.C., Jones, A.J. and Weightman, A.J. (2005) At Least 1 in 20 16S rRNA Sequence Records Currently Held in Public Repositories Is Estimated To Contain Substantial Anomalies. Appl. Environ. Microbiol. 71(12), 7724-7736.

Ashelford, K.E., Chuzhanova, N.A., Fry, J.C., Jones, A.J. and Weightman, A.J. (2006) New Screening Software Shows that Most Recent Large 16S rRNA Gene Clone Libraries Contain Chimeras. Appl. Environ. Microbiol. 72(9), 5734-5741.

Avgustin, G., Wright, F. and Flint, H. (1994) Genetic diversity and phylogenetic relationships among strains of *Prevotella* (*Bacteroides*) *ruminicola* from the rumen. Int J Syst Bacteriol 44(2), 246-255.

Bernhard, A.E. and Field, K.G. (2000a) Identification of Nonpoint Sources of Fecal Pollution in Coastal Waters by Using Host-Specific 16S Ribosomal DNA Genetic Markers from Fecal Anaerobes. Appl. Environ. Microbiol. 66(4), 1587-1594.

Bernhard, A.E. and Field, K.G. (2000b) A PCR Assay To Discriminate Human and Ruminant Feces on the Basis of Host Differences in *Bacteroides-Prevotella* Genes Encoding 16S rRNA. Appl. Environ. Microbiol. 66(10), 4571-4574.

Bernhard, A.E., Goyard, T., Simonich, M.T. and Field, K.G. (2003) Application of a rapid method for identifying fecal pollution sources in a multi-use estuary. Water Research 37(4), 909-913.

Blanch, A.R., Belanche-Munoz, L., Bonjoch, X., Ebdon, J., Gantzer, C., Lucena, F., Ottoson, J., Kourtis, C., Iversen, A., Kuhn, I., Moce, L., Muniesa, M., Schwartzbrod, J., Skraber, S., Papageorgiou, G.T., Taylor, H., Wallis, J. and Jofre, J. (2006) Integrated Analysis of Established and Novel Microbial and Chemical Methods for Microbial Source Tracking. Appl. Environ. Microbiol. 72(9), 5915-5926.

Bonjoch, X., Balleste, E. and Blanch, A.R. (2004) Multiplex PCR with 16S rRNA Gene-Targeted Primers of *Bifidobacterium* spp. To Identify Sources of Fecal Pollution. Appl. Environ. Microbiol. 70(5), 3171-3175.

Clement, B.G., Kehl, L.E., DeBord, K.L. and Kitts, C.L. (1998) Terminal restriction fragment patterns (TRFPs), a rapid, PCR-based method for the comparison of complex bacterial communities. Journal of Microbiological Methods 31(3), 135.

Dieffenbach, C.W., Lowe, T.M.J. and Dveksler, G.S. (1995) PCR Primer - A Laboratory Manual. Dieffenbach, C.W. and Dveksler, G.S. (eds), pp. p. 133 - 142, Cold Spring Harbor Laboratory Press.

Dore, J., Sghir, A., Hannequart-Gramet, G., Corthier, G. and Pochart, P. (1998) Design and evaluation of a 16S rRNA-targeted oligonucleotide probe for specific detection and quantitation of human faecal *Bacteroides* populations. Syst Appl Microbiol 21(1), 65-71.

EPA (2008) The Quality of Drinking Water in Ireland: A Report for the Year 2006, Environmental Protection Agency, Ireland.

Field, K.G., Chern, E.C., Dick, L.K., Fuhrman, J., Griffith, J., Holden, P.A., LaMontagne, M.G., Le, J., Olson, B. and Simonich, M.T. (2003) A comparative study of culture-independent, library independent genotypic methods of fecal source tracking. Journal of Water and Health 1(4), 181-194.

Fiksdal, L., Maki, J.S., LaCroix, S.J. and Staley, J.T. (1985) Survival and detection of *Bacteroides* spp., prospective indicator bacteria. Applied and Environmental Microbiology 49(1), 148-150.

Gavini, F., Pourcher, A., Neut, C., Monget, D., Romond, C., Oger, C. and Izard, D. (1991) Phenotypic differentiation of bifidobacteria of human and animal origins. Int J Syst Bacteriol 41(4), 548-557.

Gawler, A.H., Beecher, J.E., Brandao, J., Carroll, N.M., Falcao, L., Gourmelon, M., Masterson, B., Nunes, B., Porter, J., Rince, A., Rodrigues, R., Thorp, M., Martin Walters, J. and Meijer, W.G. (2007) Validation of host-specific *Bacteriodales* 16S rRNA genes as markers to determine the origin of faecal pollution in Atlantic Rim countries of the European Union. Water Research 41(16), 3780.

Gourmelon, M., Caprais, M.P., Segura, R., Le Mennec, C., Lozach, S., Piriou, J.Y. and Rince, A. (2007) Evaluation of two library-independent microbial source tracking methods to identify sources of fecal contamination in French estuaries. Appl Environ Microbiol 73(15), 4857-4866.

Griffith, J.F., Weisberg, S.B. and McGee, C.D. (2003) Evaluation of microbial source tracking methods using mixed fecal sources in aqueous test samples. Journal of Water and Health 1(4), 141-151.

Kildare, B.J., Leutenegger, C.M., McSwain, B.S., Bambic, D.G., Rajal, V.B. and Wuertz, S. (2007) 16S rRNA-based assays for quantitative detection of universal, human-, cow-, and dog-specific fecal *Bacteroidales*: A Bayesian approach. Water Research 41(16), 3701.

Lamendella, R., Domingo, J.W.S., Oerther, D.B., Vogel, J.R. and Stoeckel, D.M. (2007) Assessment of fecal pollution sources in a small northern-plains watershed using PCR and phylogenetic analyses of *Bacteroidetes* 16S rRNA gene. FEMS Microbiology Ecology 59(3), 651-660.

Layton, A., McKay, L., Williams, D., Garrett, V., Gentry, R. and Sayler, G. (2006) Development of *Bacteroides* 16S rRNA Gene TaqMan-Based Real-Time PCR Assays for Estimation of Total, Human, and Bovine Fecal Pollution in Water. Appl. Environ. Microbiol. 72(6), 4214-4224.

Lynch, P.A., Gilpin, B.J., Sinton, L.W. and Savill, M.G. (2002) The detection of *Bifidobacterium adolescentis* by colony hybridization as an indicator of human faecal pollution. J Appl Microbiol 92(3), 526-533.

Matsuki, T., Watanabe, K., Tanaka, R., Fukuda, M. and Oyaizu, H. (1999) Distribution of Bifidobacterial Species in Human Intestinal Microflora Examined with 16S rRNA-Gene-Targeted Species-Specific Primers. Appl. Environ. Microbiol. 65(10), 4506-4512.

Matsuki, T., Watanabe, K., Tanaka, R. and Oyaizu, H. (1998) Rapid identification of human intestinal bifidobacteria by 16S rRNA-targeted species- and group-specific primers. FEMS Microbiology Letters 167(2), 113 -121.

Okabe, S., Okayama, N., Savichtcheva, O. and Ito, T. (2007) Quantification of host-specific *Bacteroides–Prevotella* 16S rRNA genetic markers for assessment of fecal pollution in freshwater. Applied Microbiology and Biotechnology 74(4), 890.

Pandey, J., Ganesan, K. and Jain, R.K. (2007) Variations in T-RFLP profiles with differing chemistries of fluorescent dyes used for labeling the PCR primers. Journal of Microbiological Methods 68(3), 633.

Reischer, G.H., Kasper, D.C., Steinborn, R., Farnleitner, A.H. and Mach, R.L. (2007) A quantitative real-time PCR assay for the highly sensitive and specific detection of human faecal influence in spring water from a large alpine catchment area. Letters in Applied Microbiology 44, 351.

Reischer, G.H., Kasper, D.C., Steinborn, R., Mach, R.L. and Farnleitner, A.H. (2006) Quantitative PCR Method for Sensitive Detection of Ruminant Fecal Pollution in Freshwater and Evaluation of This Method in Alpine Karstic Regions. Appl. Environ. Microbiol. 72(8), 5610-5614.

Resnick, I. and Levin, M. (1981) Assessment of bifidobacteria as indicators of human fecal pollution. Applied and Environmental Microbiology 42(3), 433-438.

Roling, W.F.M. and Head, I.M. (2005) Molecular Microbial Ecology. Osborn, M.A. and Smith, C.J. (eds), pp. 25-57, Taylor & Francis Group, Abingdon.

Santo Domingo, J.W., Bambic, D.G., Edge, T.A. and Wuertz, S. (2007) Quo vadis source tracking? Towards a strategic framework for environmental monitoring of fecal pollution. Water Research 41(16), 3539.

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

Shanks, O.C., Nietch, C., Simonich, M., Younger, M., Reynolds, D. and Field, K.G. (2006) Basin-Wide Analysis of the Dynamics of Fecal Contamination and Fecal Source Identification in Tillamook Bay, Oregon. Appl. Environ. Microbiol. 72(8), 5537-5546.

Stricker, A.R., Wilhartitz, I., Farnleitner, A.H. and Mach, R.L. (2006) Development of a Scorpion probe-based real-time PCR for the sensitive quantification of *Bacteroides* sp. ribosomal DNA from human and cattle origin and evaluation in spring water matrices. Microbiol Res.

Wang, R.-F., Cao, W.-W., Campbell, W.L., Hairston, L., Franklin, W. and Cerniglia, C.E. (1994) The use of PCR to monitor the population abundance of six human intestinal bacterial species in an in vitro semicontinuous culture system. FEMS Microbiology Letters 124(2), 229.

Wang, R., Cao, W. and Cerniglia, C. (1996) PCR detection and quantitation of predominant anaerobic bacteria in human and animal fecal samples. Appl. Environ. Microbiol. 62(4), 1242-1247.

511 Table 1 - Primers used in this study

512

Primers[a]	Sequence (5'-3')	Target	Annealing temp (ºC)	Reference
CF128F	CCAACYTTCCCGWTACTC	*Bacteroidales*	62ºC	(Bernhard and Field, 2000b)
HF183F	ATCATGAGTTCACATGTCCG	*Bacteroidales*	63ºC	(Bernhard and Field, 2000b)
Bac32F	AACGCTAGCTACAGGCTT	*Bacteroidales*	Variable[b]	(Bernhard and Field, 2000a)
Bac708R	CAATCGGAGTTCTTCGTG	*Bacteroidales*	Variable[b]	(Bernhard and Field, 2000a)
Bacto1080R	GCACTTAAGCCGACACCT	*Bacteroidales*	58ºC	(Dore et al., 1998)
BT 1	GGCAGCATTTCAGTTTGCTTG	*Bacteroides*	50ºC	(Wang et al., 1994)

		thetaiotaomicron		
BT 2	GGTACATACAAAATTCCACACGT	*Bacteroides thetaiotaomicron*	50ºC	(Wang et al., 1994)
BV 1	GCATCATGAGTCCGCATGTTC	*Bacteroides vulgatus*	50ºC	(Wang et al., 1994)
BT 2	TCCATACCCGACTTTATTCCTT	*Bacteroides vulgatus*	50ºC	(Wang et al., 1994)
BiADO 1	CTCCAGTTGGATGCATGTC	*Bifidobacterium adolescentis*	55ºC	(Matsuki et al., 1998)
BiADO 2	CGAAGGCTTGCTCCCAGT	*Bifidobacterium adolescentis*	55ºC	(Matsuki et al., 1998)
BiCAT 1	CGGATGCTCCGACTCCT	*Bifidobacterium catenulatum* and *Bifidobacterium pseudocatenulatum*	61ºC	(Matsuki et al., 1998)
BiCAT 2	CGAAGGCTTGCTCCCGAT	*Bifidobacterium catenulatum* and *Bifidobacterium pseudocatenulatum*	61ºC	(Matsuki et al., 1998)
RumD1R	ATCTCTGAGCCTGTCCAG	*Bacteroidales*	60ºC	This study

RumD2R	TGGTCCGAAGAAGGGCCC	*Bacteroidales*	63ºC	This study
RumB1F	CTCCGCATGGAGTTTCCAC	*Bacteroidales*	62ºC	This study
BacPreR	AGGTGTCGGCTTAAGTGC	*Bacteroidales*	62ºC	(Avgustin et al., 1994)

513

514 [a] The primer pairs CF128F & Bac708R, RumB1F & BacPreR, Bac32F & RumD1R and Bac32F & RumD2R are putatively ruminant-
515 feces-specific. The primer pairs HF183F & Bac708R, BT 1 & BT2, BV 1 & BV2, BiADO 1 & BiADO 2, BiCAT 1 & BiCAT 2 are
516 putatively human-feces-specific. The primer pairs Bac32F & Bacto1080R and Bac32F & Bac708R were non-host-specific.
517 [b]When Bac32F was paired with Bac708R, the annealing temperature of 53ºC was used, when paired with Bacto1080R, RumD1R or
518 RumD2R the annealing temperature listed for these reverse primers was used. When Bac708R was paired with CF128F or HF183F
519 the annealing temperature listed for these forward primers was used.

520

521 Table 2 - Sensitivity and specificity of host-specific PCR assays

Sensitivity and Specificity	RumB1F & BacPreR	Bac32F & RumD1R	Bac32F & RumD2R	CF128F & Bac708R	HF183F & Bac708R	BT 1 & 2[b]	BV 1 & 2	BiADO 1 & 2	BiCAT 1 & 2
% Sensitivity (r)	97	91	100	95	12/70[c]	65/39[c]	88/100[c]	85/100[c]	46/100[c]
% Specificity (s)	97	100	95	94	100	NT	86	84	87
Sample Type	No. of positive PCR results/No. of samples tested								
Human sewage	0/33	0/33	0/33	0/33	23/33	13/33	33/33	33/33	33/33
Human feces	0/26	0/26	0/26	0/26	3/26	17/26	23/26	22/26	12/26
Cow[a]	25/25	25/25	25/25	25/25	0/25	NT	0/25	4[w]/25	0/25
Cow[a] Slurry	5/5	4/5	5/5	1/5	0/5	NT	0/5	0/5	0/5

Sheep[a]	37/39	34/39	39/39	39/39	0/39	NT	1/39	1[w]/39	0/39
Deer[a]	1/1	1/1	1/1	1/1	0/1	NT	0/1	0/1	0/1
Goat[a]	4/4	3/4	4/4	4/4	0/4	NT	0/4	0/4	0/4
Horse	1[w]/12	0/12	2[w]/12	0/13	0/13	NT	0/13	0/13	0/13
Donkey	0/2	0/2	0/2	0/2	0/2	NT	0/2	0/2	0/2
Dog	0/2	0/2	0/2	0/2	0/2	NT	0/2	0/2	0/2
Goose	0/1	0/1	0/1	0/1	0/1	NT	0/1	0/1	0/1
Chicken	0/2	0/2	0/2	0/2	0/2	NT	0/2	0/2	0/2
Pet Pig	2 [w]/2	0/2	0/2	2/2	0/2	NT	2/2	0/2	0/2
Pig	0/8	0/8	0/8	4/8	0/8	NT	2/8	0/8	0/8
Pig Slurry	0/15	0/15	3 [w]/15	0/15	0/15	NT	11/15	14/15	15/15

[a]Ruminant species

[b]These primer pairs were not tested on all the samples as they did not exhibit adequate specificity or sensitivity in initial assays.

524 [c]Sensitivity evaluation results for human-specific PCR assays based on testing the human-specific PCR assays with individual human

525 feces (n=26)/human sewage (n=33) separately

526 [w]Indicates there was a very weak band on the gel, reflecting poor PCR product yield

527 NT – not tested

528 Table 3 - SLOD and MDL of host-specific PCR assays

PCR Assay	SLOD[a] - g of dry feces/sewage [g of wet feces/sewage]	MDL[a] - no. of copies of plasmid per 50-µl PCR reaction
HF183F & Bac708R	6.6×10^{-2} [1×10^{-1}]	NT
BiCAT 1 & 2	6.6×10^{-4} [1×10^{-3}]	NT
BiADO 1 & 2	6.6×10^{-6} [1×10^{-5}]	NT
BV 1 & 2	6.6×10^{-6} [1×10^{-5}]	NT
CF128F & Bac708R	7.3×10^{-3} to 7.3×10^{-4} [1×10^{-2} to 1×10^{-3}]	1×10^{3} to 1×10^{4}
Bac32F & RumD1R	7.3×10^{-6} [1×10^{-5}]	1×10^{3} to 1×10^{4}
Bac32F & RumD2R	7.3×10^{-5} [1×10^{-4}]	1×10^{2} to 1×10^{3}
RumB1F & BacPreR	7.3×10^{-3} to 7.3×10^{-4} [1×10^{-2} to 1×10^{-3}]	1×10^{2} to 1×10^{3}

529 [a]10^{x} to 10^{y} indicates that all of the triplicate samples amplified at 10^{x} and some amplified

530 at 10^{y}

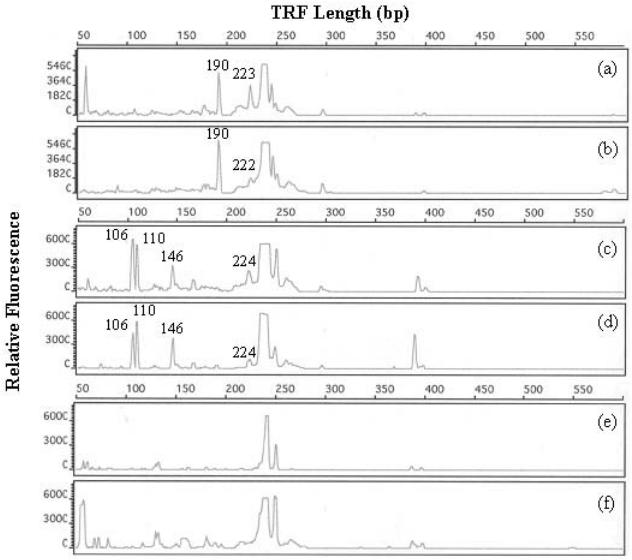
531 NT – not tested

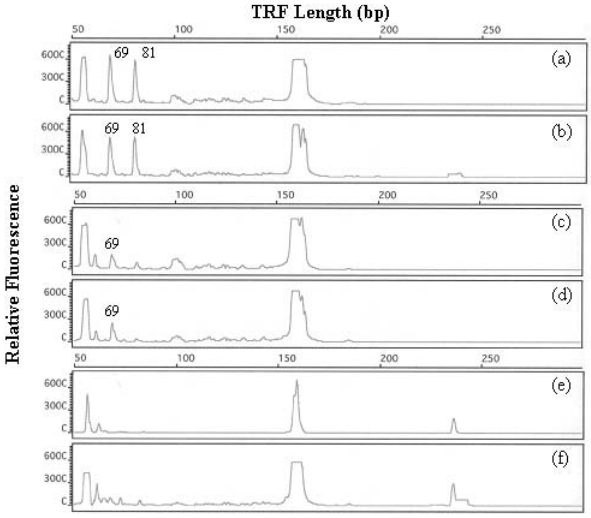
532 Fig. 1 - HEX labelled TRFLP profiles of 16S rDNA from two of each cow [(a) and (b)]
533 and sheep fecal [(c) and (d)]and human sewage DNA samples [(e) and (f)], amplified
534 with Bac32F-HEX and Bacto1080R-FAM and digested with *Hae*III.

535

536 Fig. 2 - FAM labelled TRFLP profiles of 16S rDNA fragments two of each cow [(a) and
537 (b)] and sheep fecal [(c) and (d)] and human sewage DNA samples [(e) and (f)],
538 amplified with Bac32F-HEX and Bacto1080R-FAM and digested with *Hae*III.

539
540


TRF Length (bp)
Relative Fluorescence
50 100 150 200 250 300 350 400 450 500 550
190
223
(a)
190
222
(b)
106
110
146
224
(c)
106
110
146
224
(d)
(e)
(f)


TRF Length (bp)
50
100
150
200
250
Relative Fluorescence
600C
300C
C
69
81
(a)
(b)
(c)
(d)
(e)
(f)