Research Paper

Identification of host-specific *Bacteroidales* 16S rDNA sequences from human sewage and ruminant feces

Siobhán Dorai-Raj[1], Justin O'Grady[2], Martin Cormican[3] and Emer Colleran[1]

[1] Environmental Microbiology Research Laboratory, Microbiology, School of Natural Sciences, National University of Ireland, Galway, Ireland
[2] Department of Infection, Windeyer Institute, University College London, England
[3] Centre for Health from Environment, Ryan Institute, National University of Ireland, Galway, Ireland

The need to identify the source of fecal contamination of water has led to the development of various fecal source identification methods, a field known as microbial source tracking (MST). One promising method of MST focuses on fecal members of the order *Bacteroidales,* some of which exhibit a high degree of host-specificity. In order to identify host-specific *Bacteroidales* genetic markers, a ~1060 bp section of *Bacteroidales* 16S rDNA was amplified from human sewage (n = 6), and bovine (n = 6) and ovine fecal (n = 5) samples and used for the generation of three clone libraries. Phylogenetic analysis of sequences from the three clone libraries revealed that the *Bacteroidales* species found in both human sewage and bovine and ovine feces were a highly diverse group of organisms, many of which were not represented by previously characterised 16S rDNA. Ovine and bovine feces appear to host similar populations of *Bacteroidales* species and these species were more diverse and less closely related to cultivated species than the *Bacteroidales* population found in human sewage. Species of *Bacteroidales* from the ruminant and human feces formed isolated clusters containing putatively host-specific sequences. These sequences were subsequently exploited for the design of host-specific primers which were used in MST studies.

Keywords: 16S rRNAm / Phylogenetic tree / Microbial source tracking / Feces / *Bacteroidales*

Received: April 13, 2011; accepted: May 15, 2011

DOI 10.1002/jobm.201100184

Introduction

Water polluted by feces poses a significant risk to human health and can lead to economic losses due to expenditure associated with treatment of patients with waterborne diseases and economic losses due to closure of beaches, lakes and shellfish harvesting areas. Monitoring methods used to detect fecal contamination of water have traditionally been based on the cultivation and enumeration of fecal indicator bacteria (i.e. fecal coliforms, *E. coli* and fecal enterococci). However, since these indicator bacteria are found in the feces of both animals and humans, these methods provide us with no information on the source of fecal contamination. The need to determine the source of contamination has led to the development of various fecal source identification methods, a field which is often called MST.

Recently, a number of studies have been performed on a method of MST based on the detection of host-specific strains of bacteria from the order *Bacteroidales* [1–4]. Members of *Bacteroidales* are among the most numerous bacterial populations in feces, greatly exceeding the numbers of *E. coli* [2]. In addition, members of the *Bacteroidales* order are genetically diverse, with a distribution that is limited to body cavities and, due to their anaerobic nature, are unlikely to survive for long after being released into water [2]. Most importantly, for the purpose of MST, several researchers have concluded that some strains of the microorganisms are of human fecal origin, whereas others are exclusively found in animals feces [5–7].

Fecal contamination of rural water supplies in Ireland is common [8] and apart from human feces, the

Correspondence: Siobhán Dorai-Raj, Microbiology, School of Natural Sciences, National University of Ireland, Galway, Ireland.
E-mail: Siobhandorairaj@hotmail.com
Phone: +353 91 493864
Fax: +353 91 494598

main source of contamination is ruminant in nature (bovine and ovine feces). The pathogens which cause most cases of gastroenteritis in the British isles (*Campylobacter* spp., *Salmonella* spp., verotoxigenic *E. coli* O157, *Listeria monocytogenes* and *Cryptosporidium parvum*) have all been found in livestock manures [9]. Manure production in Ireland is estimated at 100 million tons per annum, with cattle and sheep accounting for about 96% of the total [10]. Given the scale of manure production by cattle and sheep in Ireland, a reliable method to detect ruminant contamination is essential for remediation of the problem of fecal contamination of water.

The aim of this study was to analyse *Bacteroidales* 16S rDNA sequences from human sewage and bovine and ovine feces, with a view to identifying host-specific sequences which could be exploited for the design of host-specific assays for MST. To our knowledge, this is the first phylogenetic analysis of *Bacteroidales* ovine feces and one of very few studies that analyzed human sewage samples for the purposes of MST [11]. Since, in Ireland, there are an estimated 5.1 million sheep [12], it was important that both ovine and bovine feces were analyzed in this study, such that a ruminant-specific assay could be designed that would detect both bovine and ovine feces.

In this study, in order to increase the likelihood of identifying novel genetic markers, a novel combination of *Bacteroidales*-specific primers (Bac32F & Bacto1080R) was used for clone library generation [2, 13]. The primers were used to amplify a 1060 bp section of 16S rDNA from human sewage, and bovine and ovine fecal samples. These PCR products were used to generate three clone libraries which were analyzed by comparison to sequences in the EMBL nucleotide prokaryote database and by construction of a phylogenetic tree. Novel ruminant-specific markers were identified and these markers were subsequently used to design ruminant-specific PCR primers which have been successfully used, in a preliminary study, to track ruminant fecal contamination in rural drinking water supplies in Ireland [14].

Material and methods

Sample collection

Human sewage samples (untreated primary effluent, $n = 6$) were collected from two wastewater treatment works in County Galway, Ireland. Bovine ($n = 6$) and ovine ($n = 5$) fecal samples were collected from two farms in County Galway. All samples were collected with sterile utensils, placed in sterile containers, transported on ice and stored at −80 °C.

DNA extraction and PCR amplification

DNA was extracted from human sewage and bovine and ovine feces using the Powersoil™ DNA Isolation Kit (MoBio, Carlsbad, CA) as described in Dorai-Raj et al., 2009 [14]. The Bac32F (5′-AACGCTAGCTACAGGCTT-3′) [2] & Bacto1080R (5′-GCACTTAAGCCGACACCT-3′) [13] primer pair was used to amplify an approximately 1060 bp fragment of 16S rDNA from DNA (10 ng) extracted from six bovine fecal samples, five ovine fecal samples and five human sewage samples. The Probe Match program of the Ribosomal Database Project [15] was used to determine the specificity of the primer pair. Each 50 μl PCR mixture contained: 1 X *Taq* polymerase buffer, 200 μM dNTPs (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 °C for 3 min, 35 cycles consisting of 94 °C for 30 s, 58 °C for 1 min, and 72 °C for 1 min, followed by a final 5 min extension at 72 °C. A positive control (fecal DNA from the target source previously found positive) and a negative no-template control were included in every experiment and all PCR assays were performed in triplicate.

Clone library construction and sequencing

PCR products were purified using the 'High Pure PCR Product Purification Kit' (Roche Diagnostics, Mannheim, Germany) and quantified using the PicoGreen dsDNA Quantitation Kit (Molecular Probes, Eugene, OR). After quantification of the PCR products, the concentration of each was adjusted, before pooling of the PCR products from the individuals belonging to each host-species, to ensure an equal ratio of PCR product from each fecal/sewage sample used for cloning. Three pools of PCR products, each representing one of the host species, were cloned using the 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 [16]. A representative clone from each OTU group identified was selected and sequenced by MWG BIOTECH AG (Ebensburg, Germany). Five OTU groups, which each contained at least two clones, were selected at random and two clones from each of these five groups were sequenced to ensure the ARDRA analysis grouped similar sequences together accurately.

Sequence analysis

The sequence data were checked for chimeric properties using Chimera Check on the RDP II website and using the Mallard [17] and Pintail [18] programs of the Bioinformatic Toolkit website (http://www.bioinformatics-toolkit.org/index.html). Non-chimeric sequences were then compared to the EMBL nucleotide prokaryote database using the FASTA3 program on the European Bioinformatic Institute (EBI) website to determine the closest phylogenetic neighbours of cultivated species of bacteria (http://www.ebi.ac.uk/Tools/fasta33/index.html). Sequences were also aligned with each other 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. Library coverage (C), a measure of captured diversity, was calculated as $C = [1 - (n/N)] \times 100$, where n is the number of different OTU types from a clone library that are encountered only once and N is the total number of clones analyzed [19, 20].

Phylogenetic tree analysis

A single phylogenetic tree was inferred from the combined 16S rDNA sequence data obtained from the three host-species analyzed (human, bovine and ovine). Prior to phylogenetic tree construction, the 16S rDNA sequences representing each OTU from each clone library were aligned with each other and with sequence data from the closest phylogenetic neighbours (as determined by FASTA3 analysis). The alignment was performed using ClustalX. Regions where the alignment were ambiguous were removed (~70 bp in total) and not used in the construction of the phylogenetic tree. The phylogenetic inference package Paup* 4.0 [21] was used to construct all phylogenetic trees. Trees were inferred using a neighbour joining method with a Kimura-2 parameter correction. Bootstrap values were obtained from a consensus of 1000 neighbour joining trees. The 16S rDNA sequence of *Cytophaga fermentans* was used as an outgroup to root the tree. The alignment and phylogenetic tree construction was performed according to the instructions for neighbour joining trees in the manual "Phylogenetic Trees Made Easy" [22].

Results

Clone library construction and sequencing

Sequence data obtained for the two representatives of the same OTU were aligned for all five control sets of OTUs. The sequences were all ≥98% homologous which confirmed the accuracy of the ARDRA analysis (data not shown). Analysis of sequences from the bovine feces, ovine feces and human sewage clone libraries with the programs Chimera check, Mallard and Pintail revealed that 26% of the sewage derived sequences, 15% of the bovine feces derived sequences and 6% of the ovine feces derived sequences were chimeric. Chimeric sequences were not included in any further analysis. Alignment of the non-chimeric sequences with each other using ClustalW revealed varying degrees of similarity. Sequences showing >97% similarity were judged to be the same OTU and only one of each OTU was included in further analysis [20]. After removal of chimeric sequences and regrouping of OTUs the human library contained 33 different OTUs, the bovine library contained 55 different OTUs and the ovine library contained 58 different OTUs. Eleven of the OTUs contained cloned 16S rRNA genes from both the ovine and bovine clone libraries. None of the human sewage derived clones were grouped into OTUs that contained clone sequences derived from either ovine or bovine feces. The percentage coverage for the human, ovine and bovine libraries was 71%, 48% and 49%, respectively. All of the sequences have been submitted to Genbank under the following accession numbers; Bovine clone library: EU573790 – EU573833; Human clone library: EU573834 – EU573866; Ovine clone library: EU573867 – EU573924.

16S rDNA sequence analysis

16S rDNA sequence analysis using FASTA3 revealed that the percentage similarity between cloned bovine- and ovine-derived sequences and sequences from cultivated species was between 83% and 94%, with over 73% of the cloned ruminant-derived 16S rDNA fragments having a percentage similarity of less than 90%. The percentage similarity between cloned human-derived sequences and sequences from cultivated species was between 85% and 100%, with nearly 72% cloned human-derived 16S rDNA fragments having a percentage similarity >95%. Thirty two percent of the cloned 16S rDNA fragments isolated from human sewage shared at least 98% sequence similarity to *Prevotella copri*. None of the ovine or bovine-derived clones had sequence similarity of 98% or more with cultivated species from the database. Overall, 48% of the cloned 16S rDNA fragments isolated from human sewage shared at least 98% sequence similarity with cultivated species from the database.

There was a notable difference in the make-up of the human-derived and ruminant-derived clone libraries with 59% of the cloned human-derived sequences most closely related to *Prevotella* species and 11% most closely

related to *Bacteroides* species. This was in contrast to the results for the bovine and ovine-derived libraries, which gave similar results to each other, with 71% of the cloned ovine-derived sequences and 73% of the cloned bovine-derived sequences most closely related to *Bacteroides* species.

Phylogenetic tree analysis

The phylogenetic tree of the human sewage- and bovine and ovine feces-derived 16S rDNA sequences is shown in Fig. 1. In general, the bovine and ovine feces-derived sequences did not cluster with either cultivated sequences or with human sewage-derived sequences, forming several exclusively ruminant clusters and subclusters, but there were some exceptions. Of the ruminant feces-derived sequences, one ovine OTU, which represented seven (~8%) ovine feces-derived and one bovine (~1.2%) feces-derived 16S rDNA cloned sequences, clustered with *Bacteroides* species (*Bacteroides pyogenes*). Three ovine OTUs, representing three (~3.4%) 16S rDNA cloned sequences, and one bovine feces-derived OTU (~1.2%) clustered within a clade containing *Prevotella* species. Three bovine feces-derived sequences and one ovine feces-derived sequence were associated with a cluster that contained a sequence from *Barnesiella viscericola.* One ovine-derived sequence was associated with a cluster that contained *Parabacteroides merdae* and *Paludibacter propionicigenes*. A bovine feces-derived sequence, CD3, formed a deeply branched node with a human sewage-derived clone, HJ2, which was very distantly related to all other sequences. There were three other ruminant sequences (all ovine feces-derived sequences) that clustered with, or near, to human sewage-derived sequences, and all clustered with deep branching, indicating a distant relationship. There were no clusters containing bovine sequences exclusively and only one small cluster containing ovine sequences exclusively.

Overall, the human sewage-derived sequences formed clusters associated with cultivated sequences, which

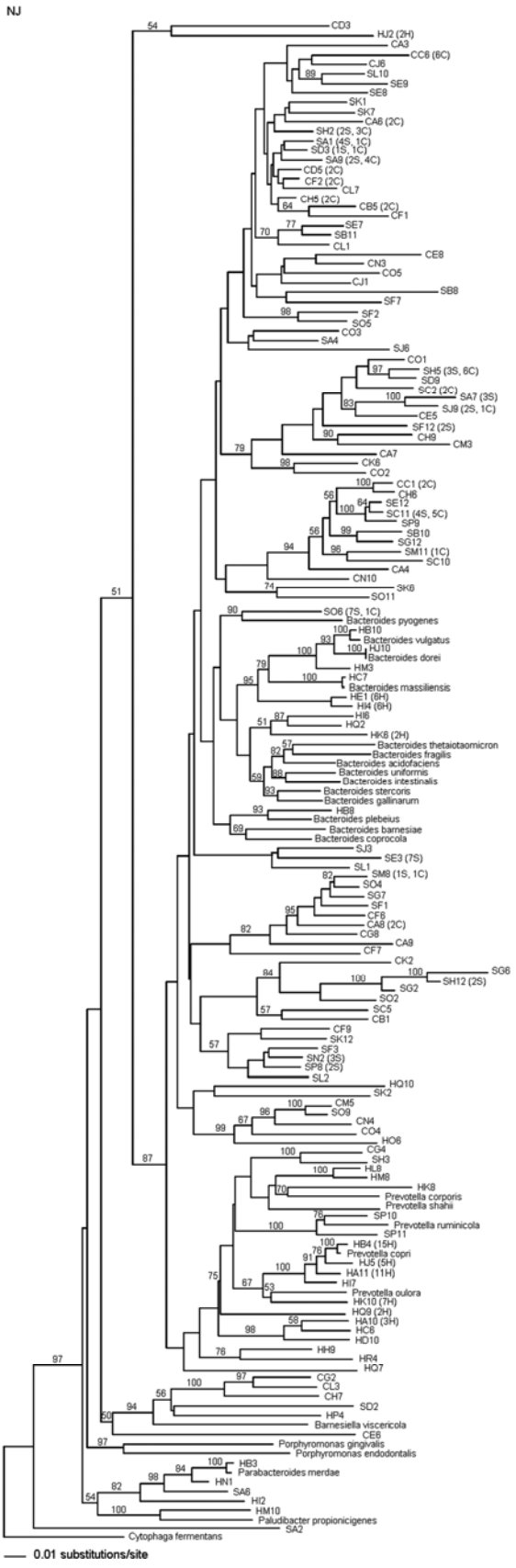

Figure 1. The number in brackets indicates the number of OTUs represented by the sequence in question. Clone sequences which represent a single OTU are not followed by any numbers. The letter "S" denotes sequences derived from sheep (ovine) feces, "C" denotes sequences derived from cow (bovine) feces and "H" denotes sequences derived from human sewage. Numbers above the internal branches are percentages of bootstrap replicates that support the branching order. Bootstrap values below 50% are not shown. The scale bar represents 1% estimated sequence divergence.

did not contain any sequences derived from ruminant feces. Ten human sewage-derived OTUs, representing 21 (~26%) 16S rDNA cloned sequences, clustered with *Bacteroides* species, and eight human sewage-derived OTUs, representing 38 (~47%) 16S rDNA cloned sequences clustered closely with *Prevotella* species. Three cloned human sewage-derived sequences were associated with a cluster that contained a sequence from *Parabacteroides merdae* and one cloned human sewage-derived sequence clustered with *Paludibacter propionicigenes*, albeit with deep branching. Seven human sewage-derived sequences formed small sub-clusters which did not contain sequences from any cultivated species but which were most closely related to *Prevotella* species. In general, the bootstrap values for the tree were low but it is worth noting that most of the branches which contained human sewage-derived sequences had a much higher bootstrap value than most of the branches containing ruminant feces-derived sequences.

Discussion

The results from any clone library analysis must be cautiously interpreted since each physical, chemical and biological step involved in the molecular analysis of an environment is a potential source of bias which may skew the results in one way or another [16, 23, 24]. Because of these biases, the actual diversity of a community cannot be reliably estimated from clone distributions. However, since the three clone libraries analyzed in this study where all prepared and analyzed using the same molecular methods, and the same definition of OTU was applied to each, the relative diversity of these clone libraries can be compared.

Most other studies which have analyzed *Bacteroidales* 16S rDNA sequences in feces or sewage, with a view to identifying host-specific sequences, have used the Bac32F and Bac708R primer pair [2] which amplify an approximately 700 bp section of the 16S rRNA gene in *Bacteroidales* [1–3, 11]. The present study used the Bac32F forward primer but paired it with a probe (Bacto1080) which was designed to target the 16S rRNA gene in species of *Bacteroides-Porphyromonas-Prevotella* (the three largest genera in the order *Bacteroidales*) [13]. In this study, the Bacto1080 probe was used as a reverse primer and designated "Bacto1080R". This Bac32F & Bacto1080R primer pair amplified an approximately 1060 bp region of the 16S rRNA gene which gave scope for a more thorough analysis of 16S rDNA sequence diversity and gave potential for the amplification of a larger region of host-specific sequence that could be used for the subsequent design of host-specific PCR primers.

In this study, 11 OTUs defined by ARDRA and sequence alignment contained cloned 16S rRNA genes from both the ovine and bovine clone libraries. None of the human sewage-derived clones were grouped into OTUs that contained clone sequences derived from either ovine or bovine feces. This was as expected since the two ruminants' fecal flora are likely to be similar to each other and less likely to be similar to human fecal flora, as humans have an entirely different diet and digestive system. This close relationship between the two ruminant species fecal flora can also be seen in the phylogenetic tree (Fig. 1).

Clone library coverage, which is a measure of how far the actual species composition of the fecal samples is captured by the clone library, was calculated by a very simple method, since a thorough analysis of species richness was not within the scope of this study. Therefore, the results are not definitive but are nonetheless useful to compare the values of the three libraries. The coverage figures of 71%, 49% and 48% for the human, ovine and bovine clone libraries indicate that the probability of the next cloned sequence falling into a novel OTU was 29%, 51% and 52%, respectively [19]. These figures indicate that the species diversity of the samples has not been thoroughly covered by the clone libraries, but since the purpose of the study was to identify host-specific species of *Bacteroidales*, rather than to provide a complete analysis of species richness, the clone libraries were of an adequate size and are a comparable to those used in similar studies [2, 3, 11]. The library coverage figures are also an indication that ruminant feces has a more diverse *Bacteroidales* population than human sewage. The higher number of OTUs in the bovine and ovine clone libraries (55 and 58, respectively) when compared to the human library (33) also reflects the higher microbial diversity in the ruminant libraries. While the lesser diversity of the human sewage flora could be attributed to the fact that the sewage samples used in the construction of the clone library were not fresh fecal samples, this finding is in agreement with studies reporting on the microbial diversity in the GI systems of ruminants and humans which did use fresh human stool samples [19, 25].

A low percentage similarity between sequences derived from ruminant feces and human sewage and sequences from cultivated species was observed in this study and has also been observed in a number of other studies on clone libraries originating from human feces, rumen fluid or ruminant feces [3, 19, 26–28]. A study by Dick *et al.* [3], using *Bacteroidales* specific PCR

primers to construct a clone library, found that the range of the ruminant (bovine and elk) sequence identity with the closest known species was 87 to 91% which supports the results of this study. Kobayashi *et al*. [28] state that previous studies have shown that among 16S rDNA sequences isolated from the rumen, only 2–31% show a close relationship (97% or more sequence similarity) with previously described species.

Fifty two percent of the human sewage-derived clones in this study had sequence similarity of ≤98% with cultivated species from the database. This is similar to the results of Lamendella *et al*. [1], who analyzed sequences derived from septic tanks, and found that approximately 40% of the sequences had 96% or less sequence similarity to *Bacteroidales*-like sequences from databases. This is also in agreement with the results of Suau *et al*. [19] who found that of the clones in their human feces library which fell within the *Bacteroides* group, 62% had less than 98% sequence similarity with cultivated sequences in public databases and were thus derived from potentially new species.

None of the ruminant clones had ≥98% sequence similarity with cultivated sequences. The higher percentage similarity of human sequences to cultivated sequences in the database when compared to the percentage similarity of ruminant cloned sequences with cultivated sequences was also observed in a number of other studies [1–3, 29]. This is probably because most cultivated species would originally have been isolated from human feces as there have been a greater number of studies performed on human fecal flora than on ruminant fecal flora [2].

There was a notable difference in the make-up of the human and ruminant clone libraries with most of the human clones being most closely related to *Prevotella* species rather than *Bacteroides* species and a majority of ruminant sequences being most closely related to *Bacteroides* species rather than *Prevotella* species. However, since the percentage similarities of the ruminant clone sequences to the cultivated sequences are so low, it can be argued that the similarity of the ruminant clone sequences to either *Bacteroides* or *Prevotella* cultivated sequences is insignificant, given that the two genera are so closely related phylogenetically. According to Tajima *et al*. [27] sequence similarity between clone sequences and cultivated species of less than 90% is too low to allocate sequences to taxa with a reasonable degree of confidence. Paster *et al*. [30] suggested that phylogenetic clustering of bacterial groups in phylogenetic trees, rather than specific similarity values to cultivated sequences, should be used as a guide for defining bacterial taxa.

Since the percentage similarities of the human clone sequences to cultivated sequences is significantly higher, the results of database analysis of these sequences is worthy of discussion. The results of this study are in direct contrast with the results of a recent study by Mieszkin *et al*. [11] who found that 67% of cloned *Bacteroidales* sequences characterised from human sewage samples were *Bacteroides*-like rather than *Prevotella*-like. A number of factors could account for the disparity between the results of the two studies. Firstly, the difference in the nature of the human sewage samples tested may have affected the outcome. The human sewage samples used in the study by Mieszkin *et al.* [11] were water samples collected at the outlets of wastewater treatment plants and the human sewage samples tested in this study were untreated primary effluent, collected from the primary settling tanks of wastewater treatment plants. Secondly, the *Bacteroidales* 16S rDNA fragments cloned in the study by Mieszkin *et al*. [11] were amplified from sewage samples using the same forward primer as used in this study (Bac32F) but a different reverse primer (Bac708R) [2]. These primers amplify a smaller fragment of DNA (approximately 690 bp) than the fragment amplified in this study (approximately 1060 bp). *In silico* sequence analysis using this longer fragment of 16S rDNA may account, in part, for the difference observed between the two studies. More significantly, the fact that a different reverse primer (Bacto1080R) was used in this study, designed to target the *Bacteroides-Porphyromonas-Prevotella* group [13] rather than the *Bacteroides-Prevotella* group (as the Bac708R primer was designed to detect [2]), would mean that a slightly different population of *Bacteroidales* would have been amplified and analyzed.

Thirty two percent of the human clones in this study share at least 98% sequence similarity to *Prevotella copri*. A recent study by Wéry *et al*. [31] which involved analysis of effluent samples from wastewater treatment plants also found 16S rRNA sequences that were closely related to *Prevotella copri*. This species is not mentioned as showing sequence similarity to human feces-derived clones in many other previous studies, as *Prevotella copri* is a novel species, first described by Hayashi *et al.* in 2007 [32]. In previous studies of the human gut, most of which date from the late 1990s, very few of the sequences from 16S rDNA clone libraries show a high similarity to any *Prevotella* sequences but this may be due to the under characterisation of the *Prevotella* genus when these studies were performed. In addition, since most other studies analyzed fresh human stool samples rather than human sewage samples, the prevalence of *P. copri* in the human sewage-derived clone library is

this study could be due to superior persistence of this particular species outside the GI tract. However, as discussed, since this species is very recently characterised, there is no published evidence, thus far, to support this theory.

In the phylogenetic tree (Fig. 1), in accordance with the sequence analysis results, most of the human sewage-derived sequences clustered with cultivated species and very few of the ruminant-derived sequences did. There were a high percentage of human sequences clustering with *Prevotella* species, which has not been observed in many other studies. As discussed above, this could be due to the fact that sewage samples rather than fresh human stools were analyzed in this study and also may be due to the under characterisation of the *Prevotella* genus when these studies were performed.

As the sequence analysis results revealed, in general, the human and ruminant sequences formed host-specific clusters, apart from a few small clusters containing both ruminant and human sequences. Similar host-specific clusters were observed in a number of similar studies [1, 3, 4, 11]. In addition, the finding that most of the ruminant feces-derived sequences did not cluster with any cultivated species of *Bacteroidales* has also been supported by a number of other studies [1, 3, 4, 11]. Dick *et al*. concluded that fecal *Bacteroidales* in ruminants were taxonomically diverse and displayed an endemic distribution of fecal *Bacteroidales* as compared to the cosmopolitan distribution of the human-derived fecal *Bacteroidales* [3]. They postulated that the unique nature of the ruminant digestive system may have led to a evolutionary pathway different to that followed by species inhabiting non-ruminant hosts.

Overall, the analysis of the *Bacteroidales* 16S rDNA cloned sequences from human, bovine and ovine feces revealed a highly diverse group of organisms, many of which were not represented by previously characterised 16S rDNA. The population of *Bacteroidales* in ruminant feces was even more diverse than the *Bacteroidales* population found in human sewage and the bacteria were less closely related to cultivated species. It appears that members of the order *Bacteroidales* are underrepresented among cultured strains, which may be a reflection of the difficulties involved in culturing strictly anaerobic species of bacteria. Ovine and bovine feces appear to host a similar population of *Bacteroidales* species. Species of *Bacteroidales* from the ruminant and human feces formed isolated clusters suggesting a level of host-specificity that could be exploited for use in microbial source tracking studies. Ruminant-feces-derived sequences isolated in this study have been used successfully for the design of ruminant-specific PCR assays as described in Dorai-Raj *et al*. [14]. These assays have been used to identify ruminant fecal contamination in rural drinking water supplies in Ireland [14].

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.

Conflict of Interest Statement

The authors declare no financial or commercial conflicts of interest.

References

[1] Lamendella, R., Domingo, J.W.S., Oerther, D.B., Vogel, J.R. *et al.*, 2007. Assessment of fecal pollution sources in a small northern-plains watershed using PCR and phylogenetic analyses of *Bacteroidetes* 16S rRNA gene. FEMS Microbiol. Ecol., **59**, 651–660.

[2] Bernhard, A.E., Field, K.G., 2000. 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**, 1587–1594.

[3] Dick, L.K., Bernhard, A.E., Brodeur, T.J., Santo Domingo, J.W. *et al.*, 2005. Host distributions of uncultivated fecal *Bacteroidales* bacteria reveal genetic markers for fecal source identification. Appl. Environ. Microbiol., **71**, 3184–3191.

[4] Kildare, B.J., Leutenegger, C.M., McSwain, B.S., Bambic, D.G. *et al.*, 2007. 16S rRNA-based assays for quantitative detection of universal, human-, cow-, and dog-specific fecal *Bacteroidales*: A Bayesian approach. Water Res., **41**, 3701–715.

[5] Fiksdal, L., Maki, J.S., LaCroix, S.J., Staley, J.T., 1985. Survival and detection of *Bacteroides* spp., prospective indicator bacteria. Appl. Environ. Microbiol., **49**, 148–150.

[6] Allsop, K., Stickler, D., 1985. An assessment of *Bacteroides fragilis* group organisms as indicators of human faecal pollution. J. Appl. Bacteriol., **58**, 95–99.

[7] Kreader, C., 1995. Design and evaluation of *Bacteroides* DNA probes for the specific detection of human fecal pollution. Appl. Environ. Microbiol., **61**, 1171–1179.

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

[9] Hutchison, M.L., Walters, L.D., Avery, S.M., Munro, F. *et al.*, 2005. Analyses of livestock production, waste storage, and pathogen levels and prevalences in farm manures. Appl. Environ. Microbiol., **71**, 1231–1236.

[10] Lee, J., 1999. Securing a balance between competitive agriculture and environmental protection. Proceedings of the 1999 Agri-Food Millennium Conference., 116–129.

[11] Mieszkin, S., Yala, J.-F., Joubrel, R., Gourmelon, M., 2010. Phylogenetic analysis of *Bacteroidales* 16S rRNA gene sequences from human and animal effluents and assessment of ruminant faecal pollution by real-time PCR. J. Appl. Microbiol., **108**, 974–984.

[12] CSO, 2009. Statistical Yearbook of Ireland, Stationery Office, Dublin, Ireland, 2009, pp. 171–185.

[13] Dore, J., Sghir, A., Hannequart-Gramet, G., Corthier, G. *et al.*, 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**, 65–71.

[14] Dorai-Raj, S., O'Grady, J., Colleran, E., 2009. Specificity and sensitivity evaluation of novel and existing *Bacteroidales* and *Bifidobacteria*-specific PCR assays on feces and sewage samples and their application for microbial source tracking in Ireland. Water Res., **43**, 4980–4988.

[15] Maidak, B.L., Cole, J.R., Lilburn, T.G., Parker, C.T., Jr. *et al.*, 2001. The RDP-II (Ribosomal Database Project). Nucl. Acids Res., **29**, 173–174.

[16] Roling, W.F.M., Head, I.M., 2005. Prokaryotic systematics: PCR and sequence analysis of amplified 16S rRNA genes. In: M.A. Osborn, C.J. Smith (eds.), Molecular Microbial Ecology. Taylor & Francis Group, Abingdon, pp. 25–57.

[17] Ashelford, K.E., Chuzhanova, N.A., Fry, J.C., Jones, A.J. *et al.*, 2006. New screening software shows that most recent large 16S rRNA gene clone libraries contain chimeras. Appl. Environ. Microbiol., **72**, 5734–5741.

[18] Ashelford, K.E., Chuzhanova, N.A., Fry, J.C., Jones, A.J. *et al.*, 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**, 7724–7736.

[19] Suau, A., Bonnet, R., Sutren, M., Godon, J.-J. *et al.*, 1999. Direct analysis of genes encoding 16S rRNA from complex communities reveals many novel molecular species within the human gut. Appl. Environ. Microbiol., **65**, 4799–4807.

[20] Leven, L., Eriksson, A.R.B., Schnurer, A., 2007. Effect of process temperature on bacterial and archaeal communities in two methanogenic bioreactors treating organic household waste. FEMS Microbiol. Ecol., **59**, 683–693.

[21] Swofford, D.L., 2001. Phylogenetic analysis using parsimony (*and other methods), version 4.0b6. Sinauer Associates, Sunderland, MA.

[22] Hall, B.G., 2004. Phylogenetic trees made easy: a how-to manual. Sinauer Associates, Sunderland, MA.

[23] von Wintzingerode, F., Gobel, U., Stackebrandt, E., 1997. Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. FEMS Microbiol. Rev., **21**, 213–229.

[24] Rajendhran, J., Gunasekaran, P., 2011. Microbial phylogeny and diversity: small subunit ribosomal RNA sequence analysis and beyond. Microbiol. Res., **166**, 99–110.

[25] Ozutsumi, Y., Hayashi, H., Sakamoto, M., Itabashi, H. *et al.*, 2005. Culture-independent analysis of fecal microbiota in cattle. Biosci., Biotech., Biochem., **69**, 1793–1797.

[26] Whitford, M.F., Forster, R.J., Beard, C.E., Gong, J. *et al.*, 1998. Phylogenetic analysis of rumen bacteria by comparative sequence analysis of cloned 16S rRNA genes. Anaerobe, **4**, 153.

[27] Tajima, K., Aminov, R.I., Nagamine, T., Ogata, K. *et al.*, 1999. Rumen bacterial diversity as determined by sequence analysis of 16S rDNA libraries. FEMS Microbiol. Ecol., **29**, 159–169.

[28] Kobayashi, Y., 2006. Inclusion of novel bacteria in rumen microbiology: Need for basic and applied science. Anim. Sci., **77**, 375–385.

[29] Okabe, S., Okayama, N., Savichtcheva, O., Ito, T., 2007. Quantification of host-specific *Bacteroides–Prevotella* 16S rRNA genetic markers for assessment of fecal pollution in freshwater. Appl. Microbiol. Biotech., **74**, 890.

[30] Paster, B.J., Dewhirst, F.E., Weisburg, W.G., Tordoff, L.A. *et al.*, 1991. Phylogenetic analysis of the spirochetes. J. Bacteriol., **173**, 6101.

[31] Wery, N., Monteil, C., Pourcher, A.M., Godon, J.J., 2010. Human-specific fecal bacteria in wastewater treatment plant effluents. Water Res., **44**, 1873–1883.

[32] Hayashi, H., Shibata, K., Sakamoto, M., Tomita, S. *et al.*, 2007. *Prevotella copri* sp. nov. and *Prevotella stercorea* sp. nov., isolated from human faeces. Int. J. Syst. Evol. Microbiol., **57**, 941–946.

