Please note that this is an author-produced PDF of an article accepted for publication following peer review. The definitive publisher-authenticated version is available on the publisher Web site

**Water Research**
August 2007, Volume 41, Issue 16, Pages 3780-3784
http://dx.doi.org/10.1016/j.watres.2007.01.028
© 2007 Elsevier Ltd All rights reserved.

**Archimer**
Archive Institutionnelle de l'Ifremer
http://www.ifremer.fr/docelec/

# 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

Andrew H. Gawler[a], Jean E. Beecher[e], João Brandão[c], Nora M. Carroll[f], Leonor Falcão[c], Michèle Gourmelon[b], Bartholomew Masterson[f], Baltazar Nunes[c], Jonathan Porter[a, e], Alain Rincé[d], Raquel Rodrigues[c], Martin Thorp[f], J. Martin Walters[a] and Wim G. Meijer[f, *]

[a] Environment Agency, National Laboratory Service, Starcross Laboratory, Starcross Exeter, Devon EX6 8PE, UK
[b] Ifremer, EMP, Laboratoire de Microbiologie, Plouzané, France
[c] Instituto Nacional de Saúde Dr. Ricardo Jorge, Avenida Padre Cruz, 1649-016 Lisboa, Portugal
[d] Laboratoire de Microbiologie de l'Environnement, USC INRA 2017, EA 956 Université de Caen Basse Normandie, France
[e] School of Biological and Chemical Sciences, University of Exeter, Exeter, EX4 4PS, UK
[f] UCD School of Biomolecular and Biomedical Science and School of Geography, Planning and Environmental Policy, University College Dublin, Dublin 4, Ireland

*: Corresponding author : Meijer G., Tel.: +353 1 7161364; fax: +353 1 7161183, email address : Wim.Meijer@ucd.ie

**Abstract:**

The recent implementation of the Revised Bathing Water Directive in the European Union has highlighted the need for development of effective methods to differentiate between sources of faecal contamination. It had previously been shown that amplification of 16S rRNA genes of host-specific *Bacteriodales* species using the HF183F and CF128F primers could be used as markers for human and bovine faecal contamination in the United States. This paper determined the sensitivity and specificity of these markers in four Atlantic Rim countries (France, Ireland, Portugal and the United Kingdom) to evaluate their usefulness in determining the origin of faecal contamination. It was shown that the HF183F marker displayed high sensitivity (80–100%) and specificity (91–100%), and is reliable as an indication of human faecal contamination. The CF128F marker displayed 100% sensitivity in all four countries. However, strong regional variations in specificity (41–96%) were observed, highlighting the need for local validation before this marker is employed in source tracking of faecal contamination.

**Keywords:** Bacteriodales; Microbial source tracking; PCR; Specificity; Sensitivity

## 1. Introduction

In March of 2006 the Revised Bathing Water Directive (2006/7/EC) governing the quality of bathing and recreational waters within the European Union came into effect. From a microbiological perspective, the Directive focuses on fewer indicators (intestinal Enterococci and *Escherichia coli*), and sets tighter standards than the previous bathing water directive (76/160/EEC), which have to be met by 2015. The effect of the new legislation on water management is evident when realising that implementation of the new standards will result in a significant increase of non-compliant beaches within the European Union if no action is taken.

A major obstacle in achieving compliance with the standards called for in the new directive is diffuse or non-point source pollution. In contrast to pollution arising from sewage or industrial effluent, non-point source pollution emanates not from a single location, and is therefore hard to manage. For example, pollution arising from agricultural land run off, sewer misconnections or wild animal excrements is hard to rectify as neither location nor nature of the contaminating factor can be readily identified. Under the new directive, non-compliant samples may be discounted when pollution is episodic and predicted, and can be identified and the public warned. Furthermore, the Directive calls for the implementation of a 'bathing water profile', which requires - amongst other matters - identification of pollution. It is clear therefore that tools for identification of the source of pollution are required as it is a necessity to: a) adequately manage bathing water quality to meet the new stringent standards; b) allow for discounting of non-compliant samples; c) establish a 'bathing water profile' as required by the Directive.

To date considerable efforts have been made to establish methods to identify the source of faecal contamination. These include $F^+$ RNA coliphage genotyping, *E. coli*

48 genotyping techniques, and the use of enterotoxin genes as biomarkers (Scott *et al.*, 2002).
49 A promising emerging methodology for microbial source tracking is based on detection
50 of *Bacteriodales* species associated with humans or animals. It had been noted that
51 although *Bacteriodales* sp. are abundantly present in the intestinal track of mammals,
52 differences in species composition of *Bacteriodales* populations exist, which led to their
53 use as markers for human and animal pollution (Allsop & Stickler, 1985; Kreader, 1995).
54 This method was significantly improved upon following analysis of the *Bacteriodales*
55 populations in human and bovine faeces by terminal restriction fragment length
56 polymorphism and clonal libraries. This gave rise to the development of primer
57 combinations that amplify 16S rRNA genes of *Bacteriodales* species specifically
58 associated with either human or bovine faeces (Bernhard & Field, 2000b; Dick *et al.*,
59 2005). This technology was proven to be robust when tested in the United States
60 (Bernhard & Field, 2000a; Boehm *et al.*, 2003; Bernhard & Field, 2000b; Bernhard *et al.*,
61 2003; Dick *et al.*, 2005; Layton *et al.*, 2006). However, a comprehensive validation for
62 application within the European Union has not been undertaken. The aim of this paper
63 was therefore to determine whether the amplification of 16S rRNA genes from host
64 specific *Bacteriodales* can be used in source tracking of faecal pollution in the European
65 Union.
66
67 **2. Materials and methods**
68 **2.1. DNA extraction**
69 Individual faecal samples (at least 25 samples from respectively bovine and humans and
70 at least 10 samples from other animals; Tables 1 and 2) were collected from human
71 volunteers, cattle, pig and chickens in France (Brittany), Ireland (Co. Wicklow), Portugal

(Alentejo) and the United Kingdom (Devon) in sterile containers and stored at -80$^{o}$C. The samples were obtained from at least two independent farms in each country. DNA was extracted from 250 mg faecal sample with either the FastDNA Spin Kit (Q-BIOgene; France, Ireland, United Kingdom) or the QIAamp mini Kit (Qiagen; Portugal) according to the manufacturers' instructions.

**2.2. PCR amplification**

16S rDNA of *Bacteriodales* species specific for bovine faeces was amplified using the oligonucleotide pair CF128F [5'-CCAACYTTCCCGWTACTC-3'] and Bac708R [5'-CAATCGGAGTTCTTCGTG-3'], whereas the combination of HF183F [5'-ATCATGAGTTCACATGTCCG-3'] and Bac708R was used to amplify sequences for *Bacteriodales* species associated with humans (Bernhard & Field, 2000a; Bernhard & Field, 2000b). The former primer combination is expected to generate an amplicon of 580 bp, whereas the latter yields a 525 bp product. The Bac32F primer [5'-AACGCTAGCTACAGGCTT-3'] in combination with Bac708R was used to amplify all *Bacteriodales* 16S rRNA genes, which served as a control to verify the presence of amplifiable DNA in every individual samples (Bernhard & Field, 2000a; Bernhard & Field, 2000b). DNA (10 ng per reaction) was amplified using *Taq* DNA polymerase in the presence of 200 ng $\mu l^{-1}$ non-acetylated BSA as recommended by the manufacturer (Takara). Non-acetylated BSA was included to overcome potential problems caused by inhibitors of the PCR reaction (Kreader, 1996). The reaction mixture was incubated at 94$^{o}$C for 3 min and subsequently subjected to 35 cycles of 94$^{o}$C for 1 min, 62$^{o}$C (CF128F) or 63$^{o}$C (HF183F) for 1 min, 72$^{o}$C for 1.5 min, followed by an incubation of at 72$^{o}$C for 7 min. Samples were analyzed by agarose gel electrophoresis as described (Sambrook &

Russell, 2001). Plasmids containing the target sequences for the CF128 and HF183 markers were used as positive controls in PCR reactions of every individual sample, whereas reactions without addition of sample were used as negative controls. Gel imaging and documentation systems of similar specifications were used in all laboratories. Participants met on a three monthly basis to compare data and set standards to ensure that results were interpreted in the same manner in all laboratories.

### 2.3. Statistical analysis

Sensitivity (r) and specificity (s) were defined as r=a/(a+c) and s=d/(b+d), where *a*, when a faecal DNA sample is positive for the PCR marker of its own species (true positive); *b*, when a faecal DNA sample is positive for a PCR marker of another species (false positive); *c*, when a faecal DNA sample is negative for a PCR marker of its own species (false negative); *d*, when a faecal DNA sample is negative for a PCR marker of another species (true negative)(Fisher & Belle, 1993). In order to obtain a 95% confidence interval for the estimates of sensitivity and specificity 25 faecal samples from bovine and human origin were analyzed. In addition 10 pig and chicken faecal samples were analysed. The Fisher Exact Test as implemented in the statistical package SPSS 13 (SPSS Inc), was used to test the hypothesis that the sensitivity and specificity were the same in each country. The significance level was set at 5%.

## 3. Results

### 3.1. Detection limit

To ensure that all the laboratories had implemented the method successfully, a series of ten samples were sent to each laboratory. The contents of the samples were not disclosed to the participants. These were used to assess their proficiency in performing the analysis and to assess the limit of detection of the PCR assay. All laboratories correctly identified the positive and negative samples supplied. PCR reaction mixtures containing plasmid DNA harbouring the target sequence for *Bacteriodales* species specific for either human or bovine hosts that could be amplified with respectively the HF183F- Bac708R and the CF128F-Bac708R primer pairs were used as template. The limit of detection for all laboratories was 0.1 pg of template DNA ($2.5x10^4$ gene copies), which is in broad agreement with earlier reported detection limits (1 pg) for these primer pairs (Bernhard & Field, 2000a). Assuming that the *Bacteriodales* species specific for either human or bovine hosts contains 5 copies of the 16S rRNA gene (rRNA Operon Copy Number Collection; http://rrndb.cme.msu.edu), then the theoretical detection limit of this method is 5000 cells. Using dilutions of cow faeces in water, Bacteroidales species could be detected using general primers to well below 8 faecal coliforms per 100 ml, which corresponds to $10^{-4}$ g/l wet faeces. A similar detection limit was reported previously (Field *et al.*, 2003).

**3.2. Performance of the *Bacteriodales* 16S rRNA gene as marker for human faeces**

DNA was extracted from faecal samples from human volunteers in the four participating Atlantic rim countries and subjected to PCR using the HF183F-Bac708R primer pair. The sensitivity of this primer pair to detect human faeces was high in all countries, although regional variations were detected. These varied from 75.9% in the United Kingdom to 100% in France, with an average sensitivity of 85.6% (Table 1). In order to determine the

specificity of the marker, DNA was isolated from bovine, pig and chicken faeces and used as template in a PCR reaction with the HF183F-Bac708R primer pair. These animals represent the majority of livestock present in the catchment areas of the regions were the analysis took place. The average specificity of this primer pair to differentiate between faeces from human and these animals was 96.7%, with regional variations between 91% (France) and 100% (United Kingdom, Ireland; Table 1). Interestingly, the average specificity of this primer pair for pig faeces (100% in all countries) was higher than that for bovine faeces (88-100%), although the pig intestinal track is more similar to humans than that of cattle.

### 3.3. Performance of the *Bacteriodales* 16S rRNA gene as marker for bovine faeces

DNA extracted from bovine faecal samples was subjected to PCR amplification using the CF128F-Bac708R primer combination. The sensitivity of this primer combination to detect *Bacteriodales* sp. specific for bovine faeces was 100% in all countries (Table 2). Although the average specificity of this marker was high (72%), its performance in terms of specificity differed significantly between countries ($p<0.001$), ranging from 41% in Portugal to 96% in Ireland. Analysis of the data shows that the specificity of this marker fell into three groups. High specificities were detected in Ireland for all three non-bovine faecal samples (96% average). Although high specificities were detected for chicken (100%) and human (average 96%) for faecal samples collected in France and the United Kingdom, the marker failed to differentiate between bovine and pig faeces in these countries. The specificity of this marker was poor in Portugal with values ranging from 10 to 48% (Table 2). The analyses of blind samples at the onset of this project was used to ensure that uniform procedures for DNA extraction, amplification and analyses of

results were in place. To confirm the poor specificity of the CF128F-Bac708R marker in Portugal, the samples were subsequently independently analysed in the laboratories of the three other participating countries, giving the same result. The low specificity of the CF128F-Bac708R primer combination observed in Portugal in contrast to that of the other three countries was therefore not due to differences in experimental procedures or interpretation of data.

## 4. Discussion

*Bacteriodales* are a predominant species in the bacterial population of faeces in humans and other warm blooded animals as assessed by culturing (Holdeman *et al.*, 1976), fluorescent in situ hybridisation (Lay *et al.*, 2005) and PCR based techniques (Wang *et al.*, 1996). The fact that *Bacteriodales* species appears to be ubiquitous amongst animals, yet displays host dependent species diversity, lead to their development as markers for determining the origin of faecal pollution (Allsop & Stickler, 1985; Bernhard & Field, 2000b; Kreader, 1995). The aim of the current paper was to validate the use of the HF183F and CF128F primers that in conjunction with the general *Bacteriodales* primer Bac708R amplify a segment of the 16S rRNA gene of *Bacteriodales* species specific for either human or bovine faeces.

The HF183F marker displayed high sensitivity and specificity in all four Atlantic Rim countries. Previously, this marker was shown to have 100% specificity and 84% sensitivity when tested in Oregon, USA (Bernhard & Field, 2000a). A real-time PCR assay allowing quantification of the HF183F marker was recently developed, which showed that this marker is present in 5 out of 7 Belgian individuals tested (Seurinck *et al.*, 2005). The *Bacteriodales* species that carries the HF183F marker therefore appears to be

widespread in the human population both in the United States and in the European Union. It was recently shown that there is little variation in microbial populations in human faeces in five northern European countries, including France and the United Kingdom, that participated in the current study (Lay *et al.*, 2005). The human HF183F marker is therefore likely to be of general use throughout the European Union.

The CF128F marker, which is indicative for bovine feces, displayed 100% sensitivity in all Atlantic Rim countries tested, indicating that, like its HF183 counterpart, the *Bacteriodales* species carrying this marker also is widely distributed in both the United States and the European Union. A sensitivity of 100% was previously reported by Bernhard and Field (Bernhard & Field, 2000a). Their studies showed that this marker is also present in faeces of other ruminants (deer, elk, goat, llama and sheep) but not in non-ruminant animals (cat, dog, duck, seagull and pig), and therefore concluded that the CF128F is a marker for ruminant faeces rather than bovine specific. The data presented here indicate a difference in host specificity of the CF128F marker; in contrast to the United States and Ireland, the CF128F marker is present in pig faeces in France, the United Kingdom and Portugal. In the latter country it is also present in human and chicken faeces. The reason for this may be that a single *Bacteriodales* species carrying the CF128F marker occurs in both the human and animal population or that the Portuguese human and animal population harbours different yet related *Bacteriodales* species each carrying the CF128F marker.

## 5. Conclusions

The aim of this paper was to determine whether the amplification of 16S rRNA genes of *Bacteriodales* species, which in the United States have been shown to be associated with

either human or bovine faeces, could be used as marker in source tracing of faecal pollution in the Atlantic Rim countries of the European Union. The data showed that the HF183F marker was generally applicable in all countries. Although the sensitivity of the marker is not 100%, human faecal pollution is generally due to release of sewage from wastewater treatment plants or septic tanks, and therefore represents faecal contamination not by individuals but by populations. Bearing this in mind, the HF183F marker should detect faecal contamination from a human population with 100% sensitivity. Although the CF128F marker displayed 100% sensitivity, the low specificity observed in Portugal and to a lesser extent in France and the United Kingdom suggests a limited usefulness of this marker, and underscores the need for thorough local validation before this methodology is applied. Although the human and bovine markers have been detected in water samples during field studies (Bernhard *et al.*, 2003; Seurinck *et al.*, 2006), relatively little is known regarding the behaviour of these markers following discharge by the hosts, although it has been established that environmental parameters such as temperature have a major impact on their persistance *in situ* (Kreader, 1998; Seurinck *et al.*, 2005). Further studies in the Atlantic Rim countries to establish the behaviour of these faecal markers following discharge are required to increase the confidence in the ability of this method to determine the source of faecal contamination. Bearing these current limitations in mind, the method appears useful in identifying and distinguishing human and non-human faecal pollution in the Atlantic Rim area, particularly in the context of water quality management required by the new Bathing Water Directive.

**Acknowledgements**

238 This research (ICREW, Improving Coastal and Recreational Waters) was funded by the
239 European Regional Development Fund through the INTERREG-IIIB programme for the
240 Atlantic Area of Europe. The authors thank Katherine Field for making plasmids
241 containing the target sequences of the CF128 and HF183 markers available.

242

243 References
244

245 **Allsop, K. & Stickler, D. J. (1985).** An assessment of *Bacteroides fragilis* group
246 organisms as indicators of human faecal pollution. *J Appl Bacteriol* **58,** 95-99.

247 **Bernhard, A. E. & Field, K. G. (2000a).** A PCR assay to discriminate human and
248 ruminant feces on the basis of host differences in *Bacteroides-Prevotella* genes
249 encoding 16S rRNA. *Appl Environ Microbiol* **66,** 4571-4574.

250 **Bernhard, A. E. & Field, K. G. (2000b).** Identification of nonpoint sources of fecal
251 pollution in coastal waters by using host-specific 16S ribosomal DNA genetic
252 markers from fecal anaerobes. *Appl Environ Microbiol* **66,** 1587-1594.

253 **Bernhard, A. E., Goyard, T., Simonich, M. T., & Field, K. G. (2003).** Application of
254 a rapid method for identifying fecal pollution sources in a multi-use estuary.
255 *Water Res* **37,** 909-913.

256 **Boehm, A. B., Fuhrman, J. A., Mrse, R. D., & Grant, S. B. (2003).** Tiered approach
257 for identification of a human fecal pollution source at a recreational beach: case
258 study at Avalon Bay, Catalina Island, California. *Environ Sci Technol* **37,** 673-680.

259 **Dick, L. K., Bernhard, A. E., Brodeur, T. J., Santo Domingo, J. W., Simpson, J. M.,
260 Walters, S. P., & Field, K. G. (2005).** Host distributions of uncultivated fecal
261 Bacteroidales bacteria reveal genetic markers for fecal source identification. *Appl
262 Environ Microbiol* **71,** 3184-3191.

263 **Field, K. G., Bernhard, A. E., & Brodeur, T. J. (2003).** Molecular approaches to
264 microbiological monitoring: fecal source detection. *Environ Monit Assess* **81,**
265 313-326.

266 **Fisher, L. D. & Belle, G. (1993).** *Biostatistics - a methodology for health sciences.,* New
267 York: John Wiley and Sons

268 **Holdeman, L. V., Good, I. J., & Moore, W. E. (1976).** Human fecal flora: variation in
269 bacterial composition within individuals and a possible effect of emotional stress.
270 *Appl Environ Microbiol* **31,** 359-375.

**Kreader, C. A. (1995).** Design and evaluation of Bacteroides DNA probes for the specific detection of human fecal pollution. *Appl Environ Microbiol* **61,** 1171-1179.

**Kreader, C. A. (1996).** Relief of amplification inhibition in PCR with bovine serum albumin or T4 gene 32 protein. *Appl Environ Microbiol* **62,** 1102-1106.

**Kreader, C. A. (1998).** Persistence of PCR-detectable *Bacteroides distasonis* from human feces in river water. *Appl Environ Microbiol* **64,** 4103-4105.

**Lay, C., Rigottier-Gois, L., Holmstrom, K., Rajilic, M., Vaughan, E. E., de Vos, W. M., Collins, M. D., Thiel, R., Namsolleck, P., Blaut, M., & Dore, J. (2005).** Colonic microbiota signatures across five northern European countries. *Appl Environ Microbiol* **71,** 4153-4155.

**Layton, A., McKay, L., Williams, D., Garrett, V., Gentry, R., & 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,** 4214-4224.

**Sambrook, J. & Russell, D. W. (2001).** *Molecular cloning: a laboratory manual*, 3rd. Cold Spring Harbor, N.Y: Cold Spring Harbour Laboratory Press

**Scott, T. M., Rose, J. B., Jenkins, T. M., Farrah, S. R., & Lukasik, J. (2002).** Microbial source tracking: current methodology and future directions. *Appl Environ Microbiol* **68,** 5796-5803.

**Seurinck, S., Defoirdt, T., Verstraete, W., & 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,** 249-259.

**Seurinck, S., Verdievel, M., Verstraete, W., & Siciliano, S. D. (2006).** Identification of human fecal pollution sources in a coastal area: a case study at Oostende (Belgium). *J Water Health* **4,** 167-175.

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

| Country | All | | | France | | | Ireland | | | Portugal | | | UK | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | s (%) | r (%) | n | s (%) | r (%) | n | s (%) | r (%) | n | s (%) | r (%) | n | s (%) | r (%) | n |
| *Source of faeces* | | | | | | | | | | | | | | | |
| **Human** | - | **86** | **104** | - | **100** | **25** | - | **88** | **25** | - | **76** | **29** | - | **80** | **25** |
| **Non-human** | **97** | - | **180** | **91** | - | **45** | **100** | - | **45** | **96** | - | **43** | **100** | - | **45** |
| Bovine | 95 | - | 100 | 88 | - | 25 | 100 | - | 25 | 92 | - | 25 | 100 | - | 25 |
| Pig | 100 | - | 40 | 100 | - | 10 | 100 | - | 10 | 100 | - | 10 | 100 | - | 10 |
| Chicken | 98 | - | 40 | 90 | - | 10 | 100 | - | 10 | 100 | - | 10 | 100 | - | 10 |

**Table 1:** Determination of the sensitivity (r) and specificity (s) of the HF183F-Bac708R primer pairs for detecting human and non-human faeces. The column 'all' contains the average of all four countries. Non-human represents the average value of bovine, pig and chicken specificity. n total number of samples tested. General *Bacteriodales* 16S rDNA was detected in all faecal samples tested.

| Country | All | | | France | | | Ireland | | | Portugal | | | UK | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | s (%) | r (%) | n | s (%) | r (%) | n | s (%) | r (%) | n | s (%) | r (%) | n | s (%) | r (%) | n |
| *Source of faeces* | | | | | | | | | | | | | | | |
| **Bovine** | - | **100** | **100** | - | **100** | **25** | - | **100** | **25** | - | **100** | **25** | - | **100** | **25** |
| **Non-bovine** | **72** | - | **184** | **76** | - | **45** | **96** | - | **45** | **41** | - | **49** | **80** | - | **45** |
| Human | 84 | - | 104 | 92 | - | 25 | 100 | - | 25 | 48 | - | 29 | 100 | - | 25 |
| Pig | 30 | - | 40 | 10 | - | 10 | 80 | - | 10 | 10 | - | 10 | 10 | - | 10 |
| Chicken | 88 | - | 40 | 100 | - | 10 | 100 | - | 10 | 50 | - | 10 | 100 | - | 10 |

**Table 2:** Determination of the sensitivity (r) and specificity (s) of the CF128F-Bac708R primer pairs for detecting bovine and non-bovine. The column 'all' contains the average of all four countries. Non-bovine represents the average value of human, pig and chicken specificity. n total number of samples tested. General *Bacteriodales* 16S rDNA was detected in all faecal samples tested.