Evaluation of a human-associated genetic marker for *Escherichia coli* (H8) for fecal source tracking in Thailand

Pimchanok Nopprapun, Suwanna Kitpati Boontanon, Hidenori Harada IWA, Nawatch Surinkul and Shigeo Fujii IWA

ABSTRACT

High levels of microbial fecal pollution are a major concern in many countries. A human-associated genetic marker for *Escherichia coli* (H8) has recently been developed for fecal source tracking. The assessment of the H8 marker performance is crucial before it can be applied as a suitable method for fecal source tracking in each country. The performance (specificity and sensitivity) of the H8 marker was evaluated by using non-target host groups (cattle, buffalo, chicken, duck, and pig feces) and target host groups (influent and effluent from a wastewater treatment plant and septages). SYBR based real-time PCR (polymerase chain reaction) was done on 400 *E. coli* isolates from non-target and target host groups after *E. coli* isolation. It was found that the specificity from animal feces samples collected in Thailand was 96%. Moreover, influent, effluent, and septage samples showed the values of the sensitivity at 18, 12, and 36%, respectively. All of the non-target host groups were found to be significantly different with positive proportions from the target host group (septage samples) ($p \leq 0.01$). Based on the results, this marker is recommended for use as a human-associated *E. coli* marker for identifying sources of fecal pollution in Thailand.

Key words | *E. coli*, fecal source tracking, human genetic marker, Thailand, wastewater

Pimchanok Nopprapun
Suwanna Kitpati Boontanon (corresponding author)
Nawatch Surinkul
Civil and Environmental Engineering Department,
Mahidol University,
Salaya, Nakhon Pathom, 73170,
Thailand
E-mail: *suwanna.boo@mahidol.ac.th*

Suwanna Kitpati Boontanon
Hidenori Harada IWA
Shigeo Fujii IWA
Graduate School of Global Environmental Studies,
Kyoto University,
Yoshida-Honmachi, Sakyo-ku, Kyoto, 606-8501,
Japan

Hidenori Harada
Graduate School of Asian and African Area Studies,
Kyoto University,
46 Yoshida-Shimoadachi, Sakyo-ku, Kyoto,
606-8501,
Japan

HIGHLIGHTS

- A human-associated genetic marker for *E. coli* (H8) was evaluated for its performance in Thailand.
- The specificity and sensitivity of the H8 marker for detecting human-associated *E. coli* were 96 and 36% respectively.
- Host-specificity performance was high and comparable to the performance found in Japan and Australia.
- The sensitivity was low because target samples were mixed with nonpoint sources from a combined sewerage system

INTRODUCTION

Fecal contamination by water discharged from point sources and nonpoint sources causes waterborne diseases. This is a major concern in many countries, especially in developing countries (Pandey *et al.* 2014). All of the waterborne pathogens are currently unrealistic for monitoring in the aquatic environment because of the diversity of pathogens, including bacteria, viruses, and protozoa (Harwood *et al.* 2014). Thus, fecal indicator bacteria (FIB) have been selected for monitoring due to being easy to measure and their high levels in sewage and feces (Griffith *et al.* 2009). For example, *Escherichia coli* has been commonly used as an indicator organism of fecal pollution and freshwater for many years (Anderson *et al.* 2005). Furthermore, FIB have been studied for water quality for decades, but the limited information

doi: 10.2166/wst.2020.525

cannot provide the data about the original source in wastewater due to their nature (found in warm- and cold-blooded animal feces) (McLellan 2004; Field & Samadpour 2007).

Fecal source tracking is a procedure to identify microbes or genetic markers specifically associated with their host sources, e.g., human, pig, cow (Johnston *et al.* 2013). This method has been successfully solving fecal pollution problems in many countries, especially in developed countries, although this method has been limited in developing countries but continues to increase (Somnark *et al.* 2018). In addition, control of fecal contamination becomes difficult to implement as a management plan without the identification of its source. Therefore, fecal source tracking has become an important method for understanding the source of fecal contamination and for decreasing the hazard of waterborne diseases that occur in the environment. This method will help to produce an efficient way for water quality management. Human-specific *Bacteroides* genetic markers have been used to detect fecal pollution in the environment (Bernhard & Field 2000; Seurinck *et al.* 2005; Haugland *et al.* 2010). However, the limitation of the usage of the *Bacteroides* genetic marker is that it persists for shorter periods of time (99% decay in <8 days) compared to the *E. coli* genetic marker (99% decay in >15 days) (Tambalo *et al.* 2012). While *E. coli* has been widely used as an indicator bacteria of fecal pollution, integrated ways to use the *Bacteroides* marker with *E. coli* concentration have not been well established yet. Recently, a human-associated genetic marker for *E. coli* (H8) has been developed and used in Japan, Australia, and Bangladesh (Gomi *et al.* 2014; Warish *et al.* 2015; Harada *et al.* 2018). The H8 marker has been validated with other human-associated *E. coli* genetic markers. It was found that the H8 marker had the highest performance in terms of the specificity and sensitivity to detect human fecal pollution (Warish *et al.* 2015). Moreover, these studies have successfully used the H8 marker to check the proportion of human-associated *E. coli* in wastewater and drinking water. However, this H8 marker method has not been widely applied to investigate the fecal pollution in a water environment for fecal source tracking in Thailand. Therefore, the H8 marker required further investigation into its applicability and performance in identifying fecal contamination sources as well as reducing the related waterborne diseases and better managing water quality in Thailand. This study was conducted with these objectives: (1) to evaluate the performance (specificity and sensitivity) of the H8 marker by testing with fecal and wastewater samples from eight non-target and target host groups in Thailand, and (2) to compare the performance of the H8 marker with that of other countries.

MATERIALS AND METHODS

Non-target and target host groups sampling

Animal feces samples including cattle, buffalo, chicken, duck, and pig (non-target host groups) were collected in sterilized plastic bags for calculating the specificity of the H8 marker. For avoiding an *E. coli* clone colony, 10 feces samples from each animal species were collected from different livestock farms in Nakhon Pathom, Samut Sakhon, and Ratchaburi provinces, Thailand. Cattle, buffalo, chicken, duck, and pig were chosen as animal species as they were the top five of the most prevalent livestock animals in Thailand, according to *The Number of Livestock in Thailand in 2017* (Department of Livestock Development 2017).

To determine the sensitivity of the H8 marker, 10 samples from each influent and effluent (target host groups) were collected in 100 mL sterilized bottles from a domestic wastewater treatment plant serving approximately 566,000 people in Bangkok, Thailand. In addition, septage samples (target host groups) from 10 individual households in Samut Sakhon province, Thailand were also collected in 100 mL sterilized bottles and analyzed for evaluation of the sensitivity of the H8 marker. Ten individual samples per animal and human sources were selected for analysis as advised by the US Environmental Protection Agency (EPA) microbial source tracking guide document (US EPA 2005). Figure 1 shows examples of septic tanks which were used for collecting the samples in this study. All the samples were kept in the dark and transported on ice to the laboratory, then stored at 4 °C and processed within 12 hours.

Escherichia coli isolation

For livestock feces, samples were directly streaked on HiCrome Chromogenic Coliform Agar (recommended for detection of *E. coli*, HIMEDIA, India) by using a sterilized loop. A loop was dipped into and picked up a small amount of feces, and then dragged across the surface of the agar. This agar was incubated at 37 °C for 22 hours.

Moreover, each of 10 samples from influent, effluent, and septage samples was mixed carefully and 30 μL of each sample was used to make a dilution in 30 mL of phosphate buffered saline. Then, the samples were filtrated by

Figure 1 | Septic tanks for sample collection.

using a 100 mL sterilized funnel and disposable filtration devices (0.45 μm, white MCE membrane, Merck, Germany). After filtration, the filter was placed on HiCrome Chromogenic Coliform Agar and it was incubated at 37 °C for 22 hours. For *E. coli* analysis, *E. coli* isolates (blue colonies) were picked up by using sterilized toothpicks and transferred into each well of a 96-well microtiter plate filled with 50 μL MilliQ water. These 96 well plates were stored and kept at −20 °C for a maximum of 24 hours prior to polymerase chain reaction (PCR) analysis.

SYBR Green based real-time PCR assays

To check the performance of specificity and sensitivity, SYBR Green based real-time PCR assays were performed on each *E. coli* isolate with a human-associated *E. coli* genetic marker (H8). Table 1 shows *E. coli* isolates which were used in this study. Two hundred and fifty *E. coli* isolates from non-target host groups and 150 *E. coli* isolates from target host groups were used to perform the analysis. In addition, the primer and target DNA sequence of the H8 genetic primer sets were used for SYBR Green based PCR assays, and are shown in Table 2.

Table 1 | *E. coli* isolates in this study

Host groups	Number of *E. coli* isolates	Number of *E. coli* isolates for PCR
Non-target host groups		
Cattle	162	50
Buffalo	108	50
Chicken	77	50
Pig	56	50
Duck	85	50
Target host groups		
Influent	103	50
Effluent	102	50
Human (septage)	131	50

In this study, the colony PCR method was used as a powerful tool for rapid and easy screening of the potentially large numbers of colonies, without DNA extraction, with high sensitivity (Bergkessel & Guthrie 2013). Each isolate of *E. coli* was screened for the presence of H8. Selective agar was used to grow *E. coli*, and isolates were collected by picking up a colony identified as *E. coli* based on its color profile. The mechanism used to identify *E. coli* from various environmental sources was based on *β*-glucuronidase as the target enzyme due to very high sensitivity and specificity of this method (Rice *et al.* 1990). Later, the separated isolates were directly transferred to a PCR master mix.

For the real-time PCR assays, the PCR mixture (15 μL) was composed of 4.9 μL of MilliQ water, 7.5 μL of Quanti-Fast SYBR Green PCR (QIAGEN, Germany), 0.3 μL each of forward and reverse primers, and 2 μL of *E. coli* samples. All PCR reactions were performed in a 96-well plate using the CFX96 touch real-time PCR system (BIO RAD, Singapore). Positive (DNA from control strains) and negative (sterile water) controls were included for each PCR assay. The Real-time PCR conditions were set at 95 °C × 5 min + (95 °C × 10 sec + 60 °C × 30 sec) × 40 cycles + melting curve analysis.

To analyze the results from SYBR Green based PCR assay, the Ct value (cycle threshold) of the amplification curve was used to determine the presence or absence of PCR amplification from each sample. The H8 positivity control was observed by the Ct value of the amplification curve as below

Table 2 | Primer sets for SYBR Green based PCR assays

Gene	Name	Primer sequence	Target DNA sequence	Product size	Source
H8	H8-F H8-R	ACAGTCAGCGAGATTCTTC GAACGTCAGCACCACCAA	ACAGTCAGCGAGATTCTTC TTGGTGGTGCTGACGTTC	177 bp	Gomi *et al.* (2014)

25 cycles. Moreover, the Tm value (melting temperature) was also determined based on the results of dissociation curve. The Tm value of the dissociation curve is shown within 91 ± 1 °C (Tm value of positive control) (Harada *et al.* 2018).

DATA ANALYSIS

Finally, all the results were evaluated by using the specificity and the sensitivity test. Specificity was calculated by using true negatives and false positives of the non-target host samples. Sensitivity was calculated by using true positives and false negatives of the target host samples (Odagiri *et al.* 2015).

Specificity was calculated as the number of non-target host samples identified correctly as negatives, divided by the total number of non-target host samples tested as shown in Equation (1).

$$\text{Specificity} = \frac{\text{true negatives}}{(\text{true negatives} + \text{false positives})} \quad (1)$$

On the other hand, sensitivity was calculated as the number of target host samples identified correctly as positives, divided by the total number of target host samples tested as shown in Equation (2).

$$\text{Sensitivity} = \frac{\text{true positives}}{(\text{true positives} + \text{false negatives})} \quad (2)$$

To differentiate the positive proportion of the H8 marker between human and animal, a proportion test was used to confirm the positive samples between the target and the non-target host groups by using RStudio (R 3.4.0), a programming language for statistical analysis and data science.

RESULTS AND DISCUSSION

H8 positive proportion from non-target and target host samples

In total, 400 fecal and wastewater samples were collected from five non-targets (cattle, buffalo, chicken, duck, and pig) and three (influent, effluent, and septage) host groups in Thailand. H8 positive percentages of isolates from non-target and target host groups are summarized in Table 3. For the specificity and sensitivity analysis, we screened 50 *E. coli* isolates from each sample to avoid false positive

Table 3 | H8 positive percentages of isolates from non-target and target host groups

Host groups	*n*	Number of H8 positive samples	Human-associated *E. coli* (%) H8	*p*-value[a]
Non-target host groups				
Cattle	50	6	12	0.01
Buffalo	50	3	6	0.001
Chicken	50	1	2	<0.001
Pig	50	1	2	<0.001
Duck	50	0	0	<0.001
Target host groups				
Influent	50	9	18	–
Effluent	50	6	12	–
Human (septage)	50	18	36	–

[a]Proportion test compared with human (septage).

detection of the H8 marker. From all of 250 non-target host groups tested in this study, 239 (96%) *E. coli* isolates were PCR negative for the H8 marker. However, one chicken, one pig six cattle, and three buffalo fecal samples were PCR positive. In addition, it was found that in a total of 150 *E. coli* isolates from target host groups the human excreta samples from septage (18/50) had the highest positive percentage of 36% with the H8 marker. Moreover, influent (9/50) and effluent (6/50) samples from a wastewater treatment plant showed the positive percentage only at 18 and 12%, respectively. The percentage of positive proportion of the H8 marker from the wastewater treatment plant was lower than the human excreta samples from septic tanks because wastewater treatment plants in Thailand have been normally constructed to use a combined sewerage system (Boontanon & Buathong 2013). This may cause some nonpoint sources that can potentially mix with the target wastewater samples such as animal feces, rain water, and toxic chemicals from urban runoff. Based on the highest positive percentage from target host groups, the positive proportion from septage samples were selected to calculate the difference between non-target and target host groups. The results showed that the H8 marker was found to be significantly different in positive proportions from human to all animal samples consisting of cattle (6/50), buffalo (3/50), chicken (1/50), pig (1/50), and duck (0/50) ($p \leq 0.01$).

Among the non-target samples, cattle feces samples were found to have the highest number of H8 positive isolates at six positive samples (12%). Such data were projecting a similar trend to that from a previous study in Australia which found five wastewater samples (50%) from cattle were positive with the H8 marker from their non-target host groups (Warish *et al.* 2015). These results show that cattle are the most frequent animal species in which the H8 marker has been detected (from two studies). In addition, another study conducted in Japan showed that pig (6.7%), cow (3.1%), and chicken (11.1%) fecal samples indicated positive results with the H8 marker (Gomi *et al.* 2014). These results illustrate the variation of the H8 marker performance in the non-target host samples. This variation may be attributed to differences in livestock feeding or livestock living conditions.

Assay performance of the H8 marker

Specificity and sensitivity are commonly used to determine and validate the host-specificity of genetic markers (Bernhard & Field 2000). The performances by specificity and sensitivity of genetic markers are crucial characteristics for the fecal source tracking method because low specificity and sensitivity of genetic markers may result in false positives and false negatives in the identification of the origin of nonpoint source fecal pollution (Stoeckel & Harwood 2007). According to universal performance criteria, it has been recommended that a specificity percentage with values below 80% need to be considered carefully and may not be suitable as a fecal source tracking method. Therefore, it has been suggested that the specificity and sensitivity of a genetic marker be tested for its applicability and performance in identifying fecal contamination sources for environmental surface water, especially for new geographical studies (US EPA 2005).

In this study, the specificity and sensitivity of the H8 marker were evaluated by using 250 fecal samples from five non-target host groups and 150 fecal/wastewater samples from three target host groups. According to Equations (1) and (2), the specificity and sensitivity of the H8 marker to human-associated *E. coli* strains were 96 and 36%, respectively. The specificity of the H8 marker was found to be at a high-performance level for using this marker for a fecal source tracking method as suggested by the US EPA microbial source tracking guide document (US EPA 2005). However, the number of H8 positive samples between influent and effluent were low and not significantly different. This indicated that the survivability of *E. coli* in influent and effluent was not significantly different for H8 positive *E. coli* and negative *E. coli*. The samples were collected from a domestic wastewater treatment plant with activated sludge process and the effluent discharged from the plant without disinfection. In addition, the sensitivity of the H8 marker in this study was a quite low percentage. The reason for the low sensitivity of the H8 genetic marker was that the H8 sequence was quite sensitive to cattle. For example, in the living conditions of a community with close livestock animal contact it was found that the evident components of the gut microbiome were being shared between children and cattle (Mosites *et al.* 2017). In addition, the consumption of cow's milk products can transfer dairy cattle-associated bacteria to the human gut (Milani *et al.* 2019). On the other hand, *Yersinia enterocolitica* is often a contaminant in meat foods such as beef, lamb, pork, and animal products such as raw milk, which mainly causes gastrointestinal manifestations in humans (Rahman *et al.* 2011). The genera like *Yersinia* which are found in animal intestines, whose sequences are also assimilative by BLAST searches with the H8 maker and positive PCR amplifications, may originate from them (Gomi *et al.* 2014).

Furthermore, small livestock farms in Thailand have run their businesses ranging from backyard animal farming to

industrial livestock farming enterprises (FAO 2002). The large concentration of animals and large farm sizes are causing environmental stress. Livestock animals around the villages therefore have more opportunity to be exposed to human feces in such living conditions, which probably increases the false negative rate with the H8 marker.

Comparison of the H8 performance

The performance of the H8 marker was compared to the three previous studies carried out in Japan (sensitivity: 30%, specificity: 99%), Australia (sensitivity: 45% and specificity: 94%), and Bangladesh (sensitivity: 16%, specificity not confirmed) (Gomi *et al.* 2014; Warish *et al.* 2015; Harada *et al.* 2018). Table 4 shows comparison of the H8 performance from the four studies.

According to this study, this marker seems to have good host-specificity performance (more than 80%) and is comparable to the performance in Japan, Australia, and Thailand. However, there is no data of the specificity in Bangladesh and other regions of the developing world. The sensitivity of the H8 marker in this study was 36% (septage samples from septic tanks), which is comparable to other studies such as in Japan and Australia which used sanitary wastewater samples. However, the sensitivity of the H8 marker in Bangladesh, where *E. coli* isolates from wastewater samples from a community toilet were analyzed, was quite low. The reasons behind this could be that animal feces (e.g. mice and ruminants) might accidentally mix into the community toilet, and human distal gut microbiota in Bangladesh could be different from other countries (Harris *et al.* 2016; Harada *et al.* 2018). Therefore, the performance of the fecal source tracking method has been significantly affected by gut microbiomes (Wu *et al.* 2011) and different geography (Yatsunenko *et al.* 2012). From these results, the performance of the H8 marker in Thailand was projecting a similar trend as compared with other countries. Thus, there is a good possibility of the usefulness of a human-associated *E. coli* genetic marker for fecal source tracking in water environments in Thailand.

Table 4 | Comparison of the H8 performance from four studies

	Specificity (%)	Sensitivity (%)
Japan (Gomi *et al.* 2014)	99	30
Australia (Warish *et al.* 2015)	94	45
Bangladesh (Harada *et al.* 2018)	–	16
This Thailand study	96	36

CONCLUSIONS

The performance (specificity and sensitivity) of the H8 marker in Thailand was evaluated and compared with three previous studies in Japan, Australia, and Bangladesh. The further purpose was to find a suitable human-specific *E. coli* genetic marker for identifying a human source of fecal pollution in environmental water for fecal source tracking in Thailand. In the present study, the specificity and sensitivity of the H8 marker for detecting human-associated *E. coli* strains were about 96 and 36%, respectively. Among non-target and target host group samples, the H8 marker was found to have good host-specificity and sensitivity compared with the studies in the three other countries. However, the sensitivity of the H8 marker was quite low because target samples were mixed with non-point sources from a combined sewerage system, which is a common practice used in Thailand at wastewater treatment plants. Furthermore, cattle's gut microbiome and cow milk consumption could easily deliver cattle-associated bacteria to humans. In this study, the *E. coli* isolates were screened and selected from eight host groups (animal and human samples). Therefore, this approach for the assessment of the H8 genetic marker is recommended for various types of animal and human sources from new geographical areas, because of the similarity of H8 sequences with *Yersinia* genera and closeness to cattle. In addition, some other factors may be different in each area such as gut microbiomes, animal husbandry practices, and living conditions, causing possible detection of false positives of the H8 marker. Therefore, it is crucial to evaluate the performance and suitability of this *E. coli* genetic marker for identifying human sources of fecal pollution in surface water.

ACKNOWLEDGEMENTS

This study was supported by the Faculty of Graduate Studies, Mahidol University and the Graduate School of Global Environmental Studies, Kyoto University.

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

REFERENCES

Anderson, K., Whitlock, J. & Harwood, V. 2005 Persistence and differential survival of fecal indicator bacteria in subtropical waters and sediments. *Applied and Environmental Microbiology* **71** (6), 3041–3048.

Bergkessel, M. & Guthrie, C. 2013 Colony PCR. *Methods in Enzymology* **529**, 299–309.

Bernhard, A. & Field, K. 2000 A PCR assay to discriminate human and ruminant feces on the basis of host differences in *Bacteroides-Prevotella* genes encoding 16S rRNA. *Applied and Environmental Microbiology* **66** (10), 4571–4574.

Boontanon, S. K. & Buathong, T. 2013 *On-site Management for Domestic Wastewater in Thailand*. pp. 1–10. Available from: http://wepa-db.net/pdf/1403policy_brief/WEPA_PB3_2013.pdf (accessed 19 July 2019).

Department of Livestock Development 2017 *The Number of Livestock in Thailand*. Available from: http://www.dld.go.th/th/ (accessed 7 November 2018).

FAO 2002 *Some Issues Associated with the Livestock Industries of the Asia-Pacific Region*. Food and Agriculture Organization, Bangkok, Thailand.

Field, K. G. & Samadpour, M. 2007 Fecal source tracking, the indicator paradigm, and managing water quality. *Water Research* **41** (16), 3517–3538.

Gomi, R., Matsuda, T., Matsui, Y. & Yoneda, M. 2014 Fecal source tracking in water by next-generation sequencing technologies using host-specific *escherichia coli* genetic markers. *Environmental Science & Technology* **48** (16), 9616–9623.

Griffith, J., Cao, Y., McGee, C. & Weisberg, S. 2009 Evaluation of rapid methods and novel indicators for assessing microbiological beach water quality. *Water Research* **43** (19), 4900–4907.

Harada, H., Fujimori, Y., Gomi, R., Ahsan, M., Fujii, S., Sakai, A. & Matsuda, T. 2018 Pathotyping of *Escherichia coli* isolated from community toilet wastewater and stored drinking water in a slum in Bangladesh. *Letters in Applied Microbiology* **66** (6), 542–548.

Harris, A., Pickering, A., Harris, M., Doza, S., Islam, M., Unicomb, L., Luby, S., Davis, J. & Boehm, A. 2016 Ruminants contribute fecal contamination to the urban household environment in Dhaka, Bangladesh. *Environmental Science & Technology* **50** (9), 4642–4649.

Harwood, V., Staley, C., Badgley, B., Borges, K. & Korajkic, A. 2014 Microbial source tracking markers for detection of fecal contamination in environmental waters: relationships between pathogens and human health outcomes. *FEMS Microbiology Reviews* **38** (1), 1–40.

Haugland, R., Varma, M., Sivaganesan, M., Kelty, C., Peed, L. & Shanks, O. 2010 Evaluation of genetic markers from the 16S rRNA gene V2 region for use in quantitative detection of selected *bacteroidales* species and human fecal waste by qPCR. *Systematic and Applied Microbiology* **33** (6), 348–357.

Johnston, C., Byappanahalli, M., Gibson, J., Ufnar, J., Whitman, R. & Stewart, J. 2013 Probabilistic analysis showing that a combination of *Bacteroides* and *Methanobrevibacter* source tracking markers is effective for identifying waters contaminated by human fecal pollution. *Environmental Science & Technology* **47** (23), 13621–13628.

McLellan, S. 2004 Genetic diversity of *escherichia coli* isolated from urban rivers and beach water. *Applied and Environmental Microbiology* **70** (8), 4658–4665.

Milani, C., Duranti, S., Napoli, S., Alessandri, G., Mancabelli, L., Anzalone, R., Longhi, G., Viappiani, A., Mangifesta, M., Lugli, G., Bernasconi, S., Ossiprandi, M., Sinderen, D., Ventura, M. & Turroni, F. 2019 Colonization of the human gut by bovine bacteria present in Parmesan cheese. *Nature Communications* **10** (1), 1286.

Mosites, E., Sammons, M., Otiang, E., Eng, A., Noecker, C., Manor, O., Hilton, S., Thumbi, S. M., Onyango, C., Garland-Lewis, G., Call, D. R., Njenga, M. K., Wasserheit, J. N., Zambriski, J. A., Walson, J. L., Palmer, G. H., Montgomery, J., Borenstein, E., Omore, R. & Rabinowitz, P. M. 2017 Microbiome sharing between children, livestock and household surfaces in western Kenya. *PLOS ONE* **12** (2), 0171017.

Odagiri, M., Schriewer, A., Hanley, K., Wuertz, S., Misra, P., Panigrahi, P. & Jenkins, M. 2015 Validation of *Bacteroidales* quantitative PCR assays targeting human and animal fecal contamination in the public and domestic domains in India. *Science of The Total Environment* **502**, 462–470.

Pandey, P., Kass, P., Soupir, M., Biswas, S. & Singh, V. 2014 Contamination of water resources by pathogenic bacteria. *AMB Express* **4** (1), 51.

Rahman, A., Bonny, T., Stonsaovapak, S. & Ananchaipattana, C. 2011 *Yersinia enterocolitica*: epidemiological studies and outbreaks. *Journal of Pathogens* **2011**, 1–11.

Rice, E., Allen, M. & Edberg, S. 1990 Efficacy of beta-glucuronidase assay for identification of *Escherichia coli* by the defined-substrate technology. *Applied and Environmental Microbiology* **56** (5), 1203–1205.

Seurinck, S., Defoirdt, T., Verstraete, W. & Siciliano, S. 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. *Environmental Microbiology* **7** (2), 249–259.

Somnark, P., Chyerochana, N., Mongkolsuk, S. & Sirikanchana, K. 2018 Performance evaluation of *Bacteroidales* genetic markers for human and animal microbial source tracking in tropical agricultural watersheds. *Environmental Pollution* **236**, 100–110.

Stoeckel, D. & Harwood, V. 2007 Performance, design, and analysis in microbial source tracking studies. *Applied and Environmental Microbiology* **73** (8), 2405–2415.

Tambalo, D., Fremaux, B., Boa, T. & Yost, C. 2012 Persistence of host-associated *Bacteroidales* gene markers and their quantitative detection in an urban and agricultural mixed prairie watershed. *Water Research* **46** (9), 2891–2904.

US EPA 2005 *Microbial Source Tracking Guide Document*. US EPA, Office of Research and Development, Cincinnati, OH, USA.

Warish, A., Triplett, C., Gomi, R., Gyawali, P., Hodgers, L. & Toze, S. 2015 Assessment of genetic markers for tracking the sources of human wastewater associated *escherichia coli* in environmental waters. *Environmental Science & Technology* **49** (15), 9341–9346.

Wu, G. D., Chen, J., Hoffmann, C., Bittinger, K., Chen, Y. Y., Keilbaugh, S. A., Bewtra, M., Knights, D., Walters, W. A., Knight, R., Sinha, R., Gilroy, E., Gupta, K., Baldassano, R., Nessel, L., Li, H., Bushman, F. D. & Lewis, J. D. 2011 Linking long-term dietary patterns with gut microbial enterotypes. *Science* **334**, 105–108.

Yatsunenko, T., Rey, F. E., Manary, M. J., Trehan, I., Dominguez-Bello, M. G., Contreras, M., Magris, M., Hidalgo, G., Baldassano, R. N., Anokhin, A. P., Heath, A.C., Warner, B., Reeder, J., Kuczynski, J., Caporaso, J. G., Lozupone, C. A., Lauber, C., Clemente, J., Knights, D., Knight, R. & Gordon, J. I. 2012 Human gut microbiome viewed across age and geography. *Nature* **486**, 222–227.

First received 28 January 2020; accepted in revised form 21 October 2020. Available online 4 November 2020