View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by NERC Open Research Archive

Water Research 126 (2017) 101–110

Contents lists available at ScienceDirect

Water Research

journal homepage: www.elsevier.com/locate/watres

A highly specific *Escherichia coli* qPCR and its comparison with existing methods for environmental waters

David I. Walker [a, *], Jonathan McQuillan [b], Michael Taiwo [a], Rachel Parks [a], Craig A. Stenton [a], Hywel Morgan [c], Matthew C. Mowlem [b], David N. Lees [a]

[a] Centre for Environment, Fisheries and Aquaculture Science, Weymouth Laboratory, Weymouth, Dorset, UK
[b] National Oceanography Centre, Southampton, Hampshire, UK
[c] School of Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, UK

ARTICLE INFO

Article history:
Received 10 May 2017
Received in revised form
7 August 2017
Accepted 14 August 2017
Available online 25 August 2017

Keywords:
Escherichia coli
Environmental waters
Real-time PCR
Real-time NASBA
Rapid methods
ybbW

ABSTRACT

The presence of *Escherichia coli* in environmental waters is considered as evidence of faecal contamination and is therefore commonly used as an indicator in both water quality and food safety analysis. The long period of time between sample collection and obtaining results from existing culture based methods means that contamination events may already impact public health by the time they are detected. The adoption of molecular based methods for *E. coli* could significantly reduce the time to detection. A new quantitative real-time PCR (qPCR) assay was developed to detect the *ybbW* gene sequence, which was found to be 100% exclusive and inclusive (specific and sensitive) for *E. coli* and directly compared for its ability to quantify *E. coli* in environmental waters against colony counts, quantitative real-time NASBA (qNASBA) targeting *clpB* and qPCR targeting *uidA*. Of the 87 *E. coli* strains tested, 100% were found to be *ybbW* positive, 94.2% were culture positive, 100% were *clpB* positive and 98.9% were *uidA* positive. The qPCR assays had a linear range of quantification over several orders of magnitude, and had high amplification efficiencies when using single isolates as a template. This compared favourably with qNASBA which showed poor linearity and amplification efficiency. When the assays were applied to environmental water samples, qNASBA was unable to reliably quantify *E. coli* while both qPCR assays were capable of predicting *E. coli* concentrations in environmental waters. This study highlights the inability of qNASBA targeting mRNA to quantify *E. coli* in environmental waters, and presents the first *E. coli* qPCR assay with 100% target exclusivity. The application of a highly exclusive and inclusive qPCR assay has the potential to allow water quality managers to reliably and rapidly detect and quantify *E. coli* and therefore take appropriate measures to reduce the risk to public health posed by faecal contamination.

Crown Copyright © 2017 Published by Elsevier Ltd. This is an open access article under the Open Government License (OGL) (http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/).

1. Introduction

The commensal bacterium, *Escherichia coli* is found ubiquitously in the intestines of warm-blooded vertebrates including humans and can enter the environment at high concentrations from faecal matter. Accordingly, the presence of *E. coli* is considered as probable evidence of faecal contamination in water (Edberg et al., 2000). Detection and quantification of *E. coli* and other faecal coliform bacteria is used globally as an indicator of the risk to human health posed by co-occurring enteric pathogens including virulent strains of *E. coli* itself. Faecal indicator bacterial standards have been set for managing faecal pollution exposure risks from bathing waters and from shellfish consumption. Many enteric pathogens pose a significant threat to human health, even at low concentration, and so methods for the enumeration of indicator *E. coli* must be sufficiently sensitive to ensure that low cell numbers are detected and quantified accurately.

Existing legislative methods to monitor *E. coli* in water supplies and food-stuffs are generally based on bacterial cultivation in

* Corresponding author. Centre for Environment Fisheries and Aquaculture Science, Barrack Road, Weymouth, Dorset, DT4 8UB, UK.
E-mail address: david.walker@cefas.co.uk (D.I. Walker).

http://dx.doi.org/10.1016/j.watres.2017.08.032
0043-1354/Crown Copyright © 2017 Published by Elsevier Ltd. This is an open access article under the Open Government License (OGL) (http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/).

selective medium (e.g. TBX, MLGA) with or without prior enrichment on a selective enrichment medium (e.g. MMGB, MLSB). However, a principal limitation of traditional culture-based analysis techniques is the time delay of at least 18 h before useful results can be obtained and used to enact an intervention, leading to increased public health risk, especially during stochastic contamination events. Real time measurement of *E. coli* culture kinetics in which bacterial replication is measured by proxy using, for example, electrical impedance or gas pressure measurements can reduce the time to positivity, but may still require long cultivation periods when starting with low cell numbers. To address this problem, a number of new approaches have been devised to reduce the delay between sampling and intervention, including (i) the development of portable or deployable *in situ* analytical technologies which obviate the need to return samples to a centralised lab (McQuillan and Robidart, 2017), and (ii) rapid molecular bio-analytical methods which can provide meaningful data in the order of minutes to hours (Heijnen and Medema, 2009; Maheux et al., 2009).

Nucleic acid amplification is a culture-independent technique for the rapid and sensitive detection and quantification of potentially any species in an environmental sample based upon their unique genetic markers. For example, quantitative PCR (qPCR) has been used to measure *E. coli* based on the amplification of sequences from specific genomic loci including *uidA* (Frahm and Obst, 2003; Silkie et al., 2008) and *tuf* (Maheux et al., 2011). However, the amplification of genomic DNA can give an overestimation of viable *E. coli* cells because it may persist in an environment following cell death. Nucleic Acid Sequence Based Amplification (NASBA) is an alternative genetic amplification method which works directly with mRNA, a highly labile intermediate between gene and protein, which is considered to degrade rapidly following cell death, and is therefore a suitable proxy for detection of viable cells. Targeting a heat-inducible gene transcript, *clpB*, Heijnen and Medema (2009) were able to enhance the sensitivity of NASBA detection for water-borne *E. coli* by gently heating the sample prior to RNA extraction. In each case, the target sequence determines the exclusivity and inclusivity (specificity and sensitivity) of the analysis. The *Escherichia* genus is extraordinarily diverse and multiple sequence alignment analysis of *E. coli* genomes reveals that the number of gene sequences common to each member of this species is considerably less than the average genome size. This makes the selection of truly ubiquitous, and specific genetic targets in *E. coli* challenging and primer sequences must be carefully selected to avoid co-amplification of material extracted from closely related organisms, not derived from faecal contamination. Indeed, the *E. coli* PCR and NASBA assays developed to date have not demonstrated 100% exclusivity to *E. coli* and amplify at least some non-*E. coli* species (e.g. *Shigella* spp (Maheux et al., 2009)).

Despite these challenges, the use of molecular methods for the enumeration of *E. coli* and other bacteria in food and water samples could significantly reduce the delay associated with cell culture, specifically as an “early warning” of risk. In this study, the potential of both DNA- and RNA-based molecular quantification of *E. coli* in environmental water samples was evaluated using qPCR and quantitative NASBA (qNASBA) respectively. The water samples were analysed using membrane filtration followed by cell culture, as well as by nucleic acid extraction followed by two qPCR assays, including a novel assay developed in this work, and a heat-inducible qNASBA assay. The relative merits and limitations of each method are discussed with a view to the implementation of molecular analytics to complement the existing, statutory monitoring of *E. coli* in food and water supplies.

2. Methods

2.1. Bacterial strains and culture

In total, 87 strains of *E. coli* were used in this study including the *E. coli* Collection of Reference Strains (ECOR; a collection of 72 *E. coli* isolates from a range of geographic and animal backgrounds obtained from the STEC Centre, Michigan State University, MI, USA), 10 Environmental isolates were collected as part of the routine bivalve shellfish hygiene monitoring programme operated by the Centre for Environment, Fisheries and Aquaculture Sciences (Cefas, UK), five reference strains were purchased from the UK National Collection of Type Cultures (NCTC) and one reference stain was purchased from the German Collection of Microorganisms and Cell Cultures (DSMZ). *E. coli* strains were used to determine the inclusivity of the qNASBA and qPCR methods for *E. coli*. Additionally, 23 non-*E. coli* species were used to test assay exclusivity. These included species closely related to *E. coli* (*E. fergusonii, E. albertii, E. vulneris, E. hermanii, Shimwellia blattae, Shigella sonnei, Shigella boydii and Shigella flexneri*) and more distantly related bacterial species (*Aeromonas caviae, Citrobacter freundii, Citrobacter koseri, Enterobacter aerogenes, Enterococcus faecalis, Enterococcus faecium, Klebsiella pneumoniae, Listeria monocytogenes, Pantoea agglomerans, Pseudomonas aeruginosa, Salmonella bongori, Salmonella Nottingham, Vibrio cholera, Vibrio parahaemolyticus and Vibrio vulnificus*). All cultures were grown overnight on either nutrient agar, LB agar, or TCBS agar at 37 °C or 30 °C according to the growth requirements of individual species. All culture media were purchased from Oxoid Ltd (Thermo-Fisher, UK).

2.2. Field sampling

Water samples were collected monthly from four different locations in the South West of England: The Axe estuary in the county of Devon, and effluent from a tertiary treated sewage works before and after UV treatment, effluent from a secondary treated sewage treatment works and the Bowleaze Cove bathing water sampling point; all situated in the county of Dorset.

With the exception of the tertiary treated sewage works, which was sampled on four occasions, samples were taken monthly from each site on five separate occasions between 21st January and 9th June 2016. Two litres of water were collected from each site using a sterile polypropylene bottle (Fisher Scientific). The water samples were transported on ice back to the laboratory where they were stored at 4 °C and used within 24 h of collection.

To allow comparisons of the effect of the sample matrix on qPCR and qNASBA results, a subsample was taken of those samples that were expected to have *E. coli* concentrations below the limit of quantification for the qPCR and qNASBA assays. These subsamples were spiked with approximately 10,000 *E. coli* (NCTC 9001) CFU/100 ml and then further analysed in the same way as all other samples. Culture-based Enumeration of *E. coli*.

E. coli concentrations were measured in triplicate using the single step TBX membrane filtration method (SCA, 2016). Briefly, between 100 and 10 ml of water was filtered onto a 45 mm diameter, 0.45 μm pore-size cellulose nitrate membrane disc (Fisher Scientific) and placed directly onto TBX medium (Oxoid). Samples expected to contain high levels of *E. coli* were diluted in maximum recovery diluent (MRD) to a 10^{-5} relative dilution, using a 10-fold dilution series. The filtered cells were recovered for 4 h at 30 °C and incubated for 18 h at 44 °C. *E. coli* were counted as blue/green colonies.

2.3. Nucleic acid extraction

Nucleic acids were extracted using either the GenElute Bacterial DNA kit (Sigma Aldrich) as per the manufacturers recommendations or the Nuclisens easyMag Extraction kit (BioMerieux). The Nuclisens method was used to extract nucleic acids from both laboratory isolates for inclusivity/exclusivity testing and environmental water samples. For isolates, a single colony from an agar plate culture was suspended in 150 μl of phosphate buffered saline (PBS, Fisher Scientific), which was added to 2 ml of Nuclisens Lysis buffer (BioMerieux), briefly vortexed to mix and left at room temperature for 10 min to lyse the cells. For environmental samples, suspended cells were collected by filtration prior to analysis; 100 ml from each sample was filtered onto a 47 mm diameter 0.45 μm pore sized Supor® filter membrane (Pall) with the exception those collected from secondary and tertiary sewage treatment works, which were highly turbid and for which a maximum of 50 ml could be filtered before the apparatus became blocked. To induce mRNA production, filters were place onto a pad soaked with MRD and recovered at 37 °C for 30 min before being heat shocked at 45 °C for 5 min. The membrane was then removed from the MRD soaked pad and placed into a 50 mm diameter Petri dish and covered with 2.2 ml of Nuclisens Lysis buffer and mixed on a Stuart SSL4 see-saw rocker at 30 oscillations/minute for 10 min at room temperature.

Following lysis, 2 ml of lysate was added to 50 μl of magnetic beads, vortexed briefly and left for 10 min to bind. Samples were then centrifuged at 1500 g for 2 min and the supernatant discarded. The beads were washed twice with 400 μl of wash buffer 1, twice with 500 μl of wash buffer 2 and once with 500 μl of wash buffer 3 on a MiniMag stand (Promega). Nucleic acids were eluted in 50 μl of elution buffer on a Thermomixer Compact (Eppendorf) at 60 °C and 1400 rpm for 5 min. Nucleic acid eluates were removed from the beads on a static magnetic stand (Promega) and split into two aliquots. For each sample, DNA concentration was measured using the Quantifluor fluorometer and ONE dsDNA reagents (Promega), RNA concentration was measured using the Qubit 2 fluorimeter with the high sensitivity RNA reagents (Invitrogen), and RNA integrity was measured using the high sensitivity RNA kit with a Bioanalyzer (Agilent).

2.4. Standard curves

DNA copy number standards were created for qPCR using the *E. coli* type strain (NCTC 9001). DNA was extracted from a single colony picked from a streak plate of *E. coli* using the GenElute Bacterial Genomic DNA kit (Sigma Aldrich) and quantified using the Quantifluor fluorimeter and ONE dsDNA reagents (Promega), according to the manufacturers recommended protocol. The *E. coli* DNA was serially diluted to between 2×10^6 and 0.2 copies per μl to create a standard curve.

To create RNA copy number standards for qNASBA, 2.5 μl of *E. coli* (NCTC 9001) DNA template was amplified using the GoTaq PCR kit (Promega). Each 50 μl reaction contained PCR buffer, 2.5 mM $MgCl_2$, 1 μM each of primers ColNasF1 and *ClpB*-R (Table 1), 0.25 mM of each dNTP and GoTaq DNA polymerase (Promega). PCR was performed using the following thermal cycling conditions: 5 min at 94 °C, followed by 40 cycles of 1 min at 94 °C, 1 min at 55 °C, 1 min at 72 °C, and a final extension step of 5 min at 72 °C. Amplification was confirmed by agarose gel electrophoresis, and the PCR amplicon was purified using the ChargeSwitch Pro PCR clean-up kit (Invitrogen). The PCR product was quantified using the Quantifluor fluorimeter and ONE dsDNA reagents (Promega). RNA transcripts were created using the Riboprobe T7 system (Promega) according to the manufacturers recommendations and purified by Direct-Zol spin columns with DNase I treatment (Zymo Research) and eluted into 50 μL of Tris-HCl EDTA (TE) buffer. Purified transcripts were quantified using the using Qubit HS RNA kit (Invitrogen), and the transcript length was confirmed using the high sensitivity RNA kit with a Bioanalyzer 2100 (Agilent). The transcripts were serially diluted to between 4×10^6 and 0.4 copies per μl to create a standard curve. Transcripts were checked for DNA contamination by null-PCR, using a Taq polymerase-based reaction.

A cell number standard was created from *E. coli* (NCTC 9001) that was grown to log phase in nutrient broth. The *E. coli* concentration was estimated by optical density and confirmed by culture. The broth culture was serially and the nucleic acids were extracted using the magnetic bead extraction method described above to give a set of standards with an equivalent of 2×10^5 to 0.2 CFU/μl. Copy number and cell number standard curves generated using the *ybbW* qPCR described below assay and were compared to determine the relationship between genome copy number and equivalent CFU counts.

2.5. Real-time NASBA

Real-time NASBA was carried out in triplicate using the Nuclisens EasyQ basic kit (BioMerieux). Primers and molecular beacons were synthesised by Sigma Genosys (UK) and Eurogentec (UK) respectively. A reagent mixture with a volume of 7.5 μL containing rehydrated NASBA reagents in reagent diluent, 120 mM KCl, 250 nM of forward primer (ColNasF1; Table 1), 250 nM of reverse primer (ColNasR1), 125 nM of molecular beacon (ColBeac) and 2.5 μL of template, was added to 0.2 mL PCR tubes (Thermo-Fisher, UK). Then, 2.5 μl of enzyme mixture was added to the inside of optical grade PCR tube caps (Thermo-Fisher, UK) and placed carefully on top of the tubes. The samples were denatured for 2 min at 65 °C, and cooled to 41 °C with the heat-sensitive enzymes protected from the high temperatures in the tube caps. The tubes were briefly centrifuged to introduce the enzyme mixture to the reaction and then placed in a Stratagene© MxPro - Mx3000P system (Agilent) preheated to 41 °C. Molecular beacon fluorescence was measured

Table 1
Primers and probes.

Name	Target	Sequence	Original reference
ColNasF1	*clpB*	5′-AATTCTAATACGACTCACTATAGGGAGAAGGCTGGACGGCGACKATCCGGTCTTCA-3′	(Heijnen and Medema, 2009)
ColNasR1	*clpB*	5′-AAATCCACATTTCTGACGAGG-3′	(Heijnen and Medema, 2009)
ColBeac	*clpB*	5′-FAM-CGATCGGGGTAAAGTRATTCGCCTGGAACGATCG-BHQ1-3′	(Heijnen and Medema, 2009)
ClpB-R	*clpB*	5′-GCACCATCGCGTAATTGGTC-3′	This study
*uidA*_FPrimer	*uidA*	5′-CGGAAGCAACGCGTAAACTC-3′	(Silkie et al., 2008)
*uidA*_RPrimer	*uidA*	5′-TGAGCGTCGCAGAACATTACA-3′	(Silkie et al., 2008)
*uidA*_probe	*uidA*	5′-FAM-CGCGTCCGATCACCTGCGTC-TAMRA-3′	(Silkie et al., 2008)
401 F	*ybbW*	5′-TGATTGGCAAAATCTGGCCG-3′	This study
611 R	*ybbW*	5′-GAAATCGCCCAAATCGCCAT-3′	This study

at 30 s intervals over 120 min. Each sample was prepared in three replicate reactions; standards were prepared in 5 replicate reactions. Due to a high level of variability found between qNASBA runs, each NASBA run included triplicate samples containing 10^6 *E. coli* type strain cells to allow for normalisation of results between runs. Normalisation was performed using the following equation: $R_N = \frac{R_R}{\overline{X}/\overline{Y}}$ where R_N is the normalised TtP, R_R is the raw TtP, $\overline{X}$ is the arithmetic mean of TtP for all 10^6 cell standards in a specific qNASBA run and $\overline{Y}$ is the arithmetic mean of TtP for all 10^6 cell standards in all qNASBA runs.

2.6. Real-time PCR

A novel real-time PCR assay for the detection and quantification of *E. coli* was developed to target a fragment of the *ybbW* gene, encoding a putative Allantoin transporter. The *ybbW* gene sequence is part of the *E. coli* 'core genome', in which each gene exists in >95% of all sequenced strains. A consensus of the *ybbW* coding sequence was queried against the non-redundant nucleotide database held by the National Centre for Biotechnology Information (NCBI) using the Basic Logical Alignment Search Tool (BLAST). At the time of study the *ybbW* consensus sequence was found to have no significant matches to database sequences excluding those associated with *E. coli*. It was therefore selected for the development of a new assay based upon its inclusivity and exclusivity within *E. coli*. Multiple sets of primers were selected from conserved regions within an alignment of 60 annotated *E. coli* genome sequences (see supporting information Fig. S1), using the primer design function of Geneious R8 (Version 8.1.7). Primers were purchased from MWG Eurofins (Germany), and prepared in ultrapure water at a working concentration of 10 μM. Each primer set was evaluated by performing real-time PCR to amplify the *ybbW* target fragment from a stock of *E. coli* DNA, prepared as described above from strain NCTC 9001. Each reaction contained 12.5 μl of IQ SYBR Supermix containing reaction buffer, dNTPs, Taq polymerase and SYBR Green II DNA binding dye (Biorad), 9.5 μl of RT-PCR grade water (Agilent), 1 μl of each primer (final concentration 5 mM) and 1 μl of template DNA at 100 ng/μl; the final volume was 25 μl. The thermal cycling parameters were 95 °C for 2 min followed by 40 cycles of 95 °C for 15 s, 55–70 °C for 20 s and 72 °C for 45 s. A final extension step of 72 °C for 2 min was followed by a high resolution melting curve analysis to ensure product specificity. All reactions were carried out using a LightCycler 2.0 real-time PCR machine (Roche). The optimal primer set, and optimal annealing temperature, was selected by comparing amplification rate, total fluorescence and product specificity. Primers 401 F and 611 R (Table 1) were used for all subsequent reactions, with a primer annealing temperature of 68 °C; all other parameters were as described above.

An existing qPCR method developed by Silkie et al. (2008) for the detection and enumeration of *E. coli* targeting the *uidA* gene was also used for this study. Primers and probe were synthesised by Sigma Genosys (UK) and Eurogentec (UK) respectively. Each 25 μl PCR reaction contained 1 × Universal Taqman PCR Mastermix (Applied Biosystems), 300 nM of primer *uidA*_FPrimer, 900 nM of primer *uidA*_RPrimer, 300 nM of probe *uidA*_probe and 5 μl template DNA. The reactions were carried out on a Stratagene© MxPro - Mx3000P system (Agilent) with the following cycling parameters: 1 cycle at 95 °C for 10 min followed by 50 cycles at 95 °C for 20 s and 60 °C for 60 s. All qPCR reactions (including the *ybbW* and *uidA* assays) were carried out in triplicate for each sample.

3. Results

In this work, molecular methods for the detection and enumeration of *E. coli* in water samples were compared against a current, statutory culture based method; growth on selective and differential TBX medium. The molecular methods included a quantitative NASBA (qNASBA) assay targeting the *clpB* gene transcript, which can be induced by gently heating the sample prior to RNA extraction, a pre-existing qPCR assay targeting the *uidA* gene and a novel qPCR assay, developed in this study, which targets the *ybbW* gene, encoding a putative Allantoin transporter.

3.1. Inclusivity and exclusivity tests

The inclusivity and exclusivity of each method for *E. coli* was evaluated using a panel of 87 unique *E. coli* strains and 23 non-*E. coli* bacterial strains, including closely related members of the *Escherichia* genus. The results are summarised in Table 2. The qNASBA assay and the *ybbW* qPCR assay had 100% inclusivity for the *E. coli* strains tested, whereas the *uidA* qPCR assay and the TBX culture-based assay were 98.9% and 94.3% inclusive, respectively. The TBX culture, which is used as a statutory test for *E. coli* in water samples, did not detect 3 of the 72 ECOR strains and 2 with a O157:H7 serotype (NCTC 12900 & DSMZ 19206), which grew as cream-coloured colonies. The qNASBA assay and the *uidA* qPCR assay gave positive results for all *Shigella* spp. tested. The qNASBA was also positive for two of the non-*E. coli Escherichia* spp.; *E. albertii* and *E. fergusonii*. In contrast, the *ybbW* qPCR assay was 100% exclusive, giving no positive results for the non-*E. coli* species tested.

3.2. Standard curves: linearity and limit of detection

Standard curves were generated from a 10-fold dilution series of target sequence for each assay, shown in Fig. 1.

The qNASBA assay had an estimated detection limit of <10 copies, however the time taken for the reaction to amplify the 10 target copies to detectable levels (the Time To Positivity; TTP) was highly variable; the mean TTP was 59.7 min with a standard deviation of 15.6 min. An estimated single copy of the qNASBA target (*clpB* gene transcript) could be amplified in some instances, albeit this was never achieved consistently between replicate reactions. At an estimated 10^2 *clpB* transcript copies, the correlation coefficient was linear ($R^2 = 0.936$) but the amplification efficiency was low (61.7%). Amplification of between 10^4 and 10^7 estimated target copies produced a linear correlation ($R^2 = 0.962$) and the amplification efficiency was calculated as 103.4%.

The *ybbW* qPCR assay had an estimated detection limit of <10 copies. Fewer than 10 copies were never detected in our experiments. When the reactions contained between 10 and 10^6 estimated copies, the standard deviation between repeats was <0.30 cycles. Within this range there was a linear correlation ($R^2 = 0.999$) and the amplification efficiency was calculated as 94.0%.

The *uidA* qPCR assay also had a detection limit of <10 copies, with no positive amplifications below this level. When amplifying between 10 and 10^6 copies, the standard deviation between replicate reactions was <0.32 cycles, there was a linear correlation ($R^2 = 0.998$) and the amplification efficiency was calculated as 89.8%.

3.3. Field sample testing

Water samples were collected from several sites in the south west of England, at monthly intervals. The number of viable *E. coli* cells in each water sample was initially measured using the culture-based methodology, and is summarised in Fig. 2. Statistical ANOVA tests showed that *E. coli* concentration differed significantly between the samples ($p < 0.001$). Post ANOVA Tukey tests showed that the *E. coli* cell number at the secondary sewage works effluent

Table 2
Inclusivity and exclusivity tests of NASBA and two qPCR assays.

Species	Collection	Strain	TBX culture	*clpB* qNASBA	*uidA* qPCR	*ybbW* qPCR
Escherichia coli	ECOR	1–72	+(69)–(3)	+(72)	+(72)	+(72)
Escherichia coli	NCTC	9001	+	+	+	+
Escherichia coli	NCTC	12241	+	+	+	+
Escherichia coli	NCTC	13216	+	+	+	+
Escherichia coli	NCTC	12900	–	+	+	+
Escherichia coli	DSMZ	19206	–	+	+	+
Escherichia coli	Cefas	15/282	+	+	+	+
Escherichia coli	Cefas	15/291	+	+	+	+
Escherichia coli	Cefas	15/292	+	+	+	+
Escherichia coli	Cefas	15/297A	+	+	+	+
Escherichia coli	Cefas	15/300A	+	+	+	+
Escherichia coli	Cefas	15/300C	+	+	–	+
Escherichia coli	Cefas	15/415	+	+	+	+
Escherichia coli	Cefas	15/416	+	+	+	+
Escherichia coli	Cefas	15/424	+	+	+	+
Escherichia coli	Cefas	15/435A	+	+	+	+
Aeromonas caviae	NCTC	10852	–	–	–	–
Citrobacter freundii	NCTC	9750	–	–	–	–
Citrobacter koseri	NCTC	10786	–	–	–	–
Enterobacter aerogenes	NCTC	10006	–	–	–	–
Enterococcus faecalis	NCTC	775	–	–	–	–
Enterococcus faecium	NCTC	7171	–	–	–	–
Escherichia albertii	NCTC	17582	–	+	–	–
Escherichia fergusonii	NCTC	12128	–	+	–	–
Escherichia hermanii	NCTC	12129	–	–	–	–
Escherichia vulneris	NCTC	12130	–	–	–	–
Klebsiella pneumoniae	NCTC	9633	–	–	–	–
Listeria monocytogenes	NCTC	11994	–	–	–	–
Pantoea agglomerans	NCTC	9381	–	–	–	–
Pseudomonas aeruginosa	NCTC	10332	–	–	–	–
Salmonella bongori	DSMZ	13772	–	–	–	–
Salmonella Nottingham	NCTC	7832	–	–	–	–
Shigella boydii	DSMZ	7532	–	+	+	–
Shigella flexneri	DSMZ	4782	–	+	+	–
Shigella sonnei	DSMZ	5570	+	+	+	–
Shimwellia blattae	NCTC	12127	–	–	–	–
Vibrio cholera	NCTC	8042	–	–	–	–
Vibrio parahaemolyticus	NCTC	10885	–	–	–	–
Vibrio vulnificus	NCTC	11067	–	–	–	–
No template control			–	–	–	–

and the tertiary sewage treatment works effluent were similar, but the number of viable *E. coli* cells that could be detected following UV treatment was significantly lower. At the estuary site, there was a high level of variation between the monthly samples. This was due to an ebb tide during three of the samplings bringing contamination from up-estuary and a flood tide during two of the samplings bringing relatively clean water from the English Channel.

In tandem with the culture based analysis, nucleic acids were extracted from all of the water samples and characterised as shown in Table 3. There were relatively low concentrations of RNA and DNA in the estuarine and bathing water samples compared with the sewage effluent samples. The RNA concentrations were too low in the estuarine and bathing water samples to enable a RNA integrity number (RIN) to be calculated.

To determine whether copy numbers can reliable predict CFU in environmental samples, the relationship between RNA or DNA copy number (measured using either the *uidA* qPCR, *ybbW* qPCR or *clpB* qNASBA assay) and the number of *E. coli* CFU (measured by culture) is summarised in Fig. 3. For the *uidA* qPCR assay there was a log-linear relationship between target copy number and the number of CFU ($R^2 = 0.639$) when there was greater than 100 CFU/PCR reaction. For the *ybbW* qPCR assay, which used 1 μl of template DNA, there was a log-linear relationship between target copy number and the number of CFU ($R^2 = 0.673$) where there were greater than 10 CFU/PCR. There was no clear relationship between RNA copy number and CFU in the *clpB* qNASBA assay.

Comparisons by paired T tests showed that while there was a significant difference in *E. coli* concentration between samples collected before and after UV treatment as measured by membrane filtration ($p < 0.001$), there were no significant differences in copy numbers for the *uidA* qPCR assay ($p = 0.585$), *ybbW* qPCR assay ($p = 0.289$) or *clpB* qNASBA assay ($p = 0.115$).

Table 4 shows the percentage of reactions that had no amplification at different *E. coli* concentration ranges for all three assays tested. Reactions containing fewer than an equivalent of 10 *E. coli* CFU did not reliably amplify for any of the assays. The qNASBA assay had a high likelihood of amplification failure even at the highest *E. coli* concentrations. For the *uidA* qPCR assay, a single extraction replicate of a spiked seawater sample containing 660 *E. coli* CFU/PCR did not amplify in all three PCR replicates, but the other two extraction replicates for the same sample amplified in all PCR replicates.

3.4. Calculated CFU counts by qPCR

The relationship between DNA copy numbers per PCR reaction (C_r) and equivalent CFU counts per PCR reaction (E_r) was calculated as $logE_r = 0.9logC_r + 0.1883$. The calculated *E. coli* concentrations (E_c) in field samples was calculated as $E_c = 100\left(\frac{E_r}{P_e}/\frac{P_l}{V_f}\right)$ where P_e is the proportion of DNA eluate used in a PCR reaction, P_l is the proportion of cell lysate that was purified and V_f is the volume of sample that was filtered. Fig. 4 shows the calculated *E. coli*

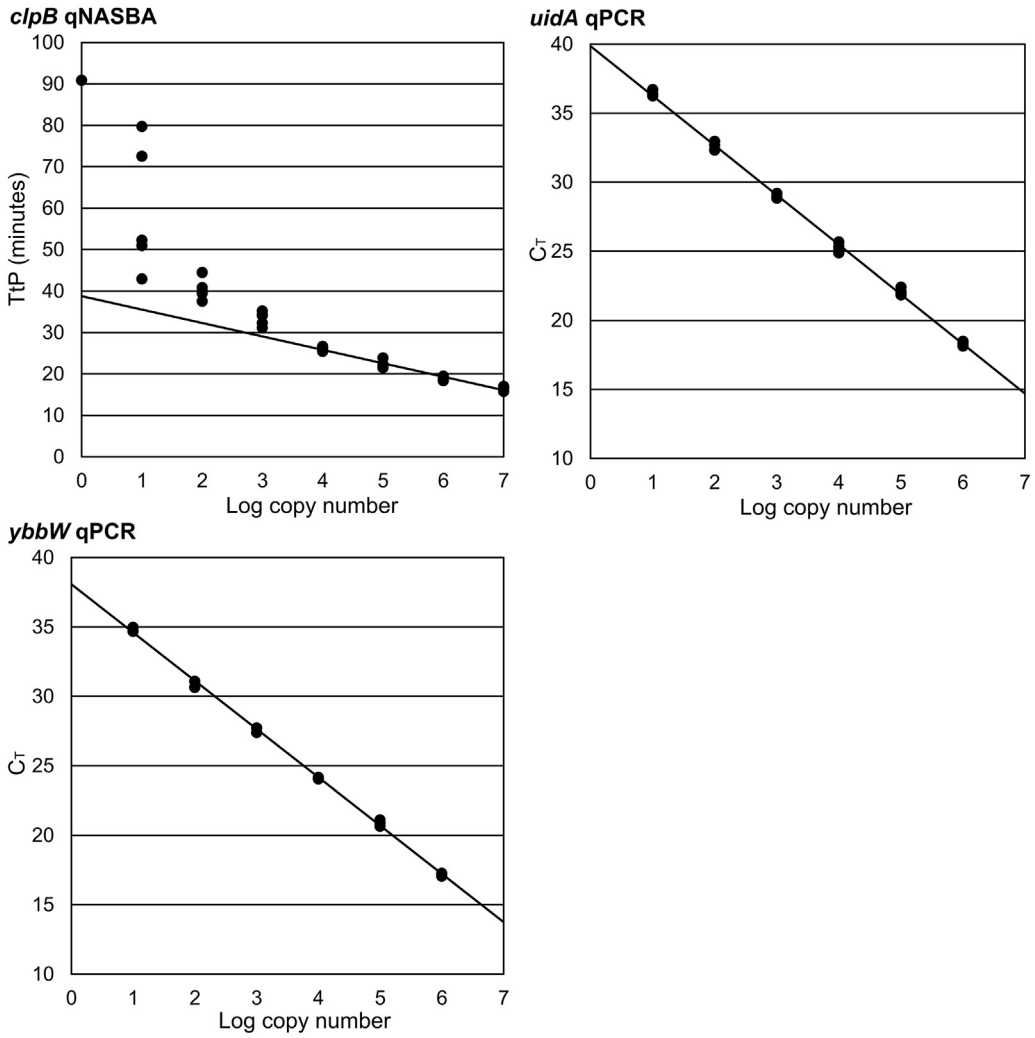

Fig. 1. Standards curves for the *clpB* qNASBA using RNA transcripts derived from *E. coli* (NCTC 9001) and *uidA* qPCR and *ybbW* qPCR using *E. coli* (NCTC 9001) genomes.

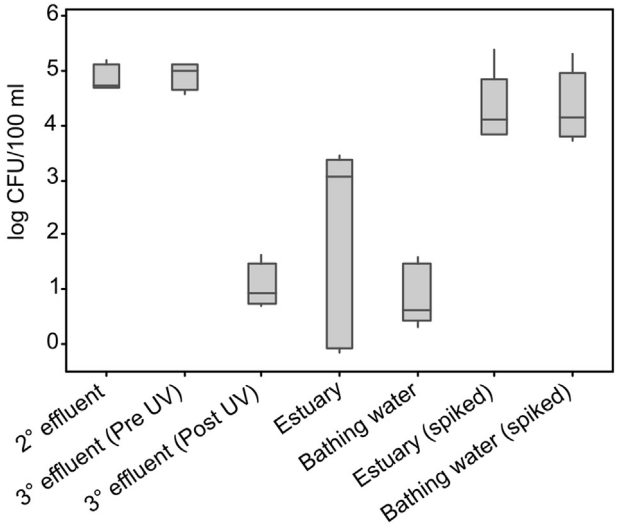

Fig. 2. Summary of *E. coli* concentrations in water samples taken from several sites in the south west of England. For two of sites (estuary and bathing water), samples were further quantified following spiking with a culture of *E. coli* (NCTC 9001).

concentrations against the measured *E. coli* concentrations for those samples field samples that had *E. coli* concentrations of >1000 CFU/100 ml and therefore fell within the linear range for quantification by *uidA* and *ybbW* qPCR as shown above.

There was a linear relationship between calculated and measured *E. coli* concentrations for both *uidA* and *ybbW* qPCR assays ($R^2 = 0.717$ and 0.748 respectively). The average proportion of log calculated to log measured *E. coli* concentration was 0.98 and 0.95 for the *uidA* and *ybbW* qPCR assays respectively. One-way ANOVA tests showed that there were no significant differences in the proportion of log calculated to log measured *E. coli* concentration between matrices for the *ybbW* ($P = 0.0.77$) or *uidA* ($P = 0.091$) qPCR assays.

4. Discussion

In this study the potential of nucleic acid amplification for the rapid detection and quantification of *E. coli* in environmental waters was evaluated and compared with the TBX colourimetric culture based assay. The TBX assay, which is currently used for statutory testing of water samples, failed to identify 5 of 87 (5.7%) *E. coli* strains from a range of animal hosts and geographic locations. In contrast, the nucleic acid amplification-based assays were more inclusive; a *uidA*-targeted qPCR assay amplified 98.8% of the strains

Table 3
RNA concentrations, RNA integrity numbers (RIN) and DNA concentrations for water samples. Standard deviations are in brackets. *Several samples had RNA concentrations that were below the limit of quantification (LOQ) for measurement of RIN.

Sample type	Average RNA concentration (ng/μl)	Average RIN	Average DNA concentration (ng/μl)
2° treated effluent	16.3 (4.7)	7.3 (0.5)	3.9 (1.5)
3° treated effluent (pre UV)	21.7 (10.5)	8.2 (0.5)	3.1 (1.4)
3° treated effluent (post UV)	15.6 (6.6)	7.7 (0.4)	2.9 (1.0)
Estuary	4.3 (2.1)	<LOQ*	1.3 (0.8)
Bathing water	5 (2.5)	<LOQ*	0.8 (0.5)
Estuary (spiked)	5.8 (2.5)	<LOQ*	1.3 (0.8)
Bathing water (spiked)	5.6 (2.4)	<LOQ*	0.9 (0.7)

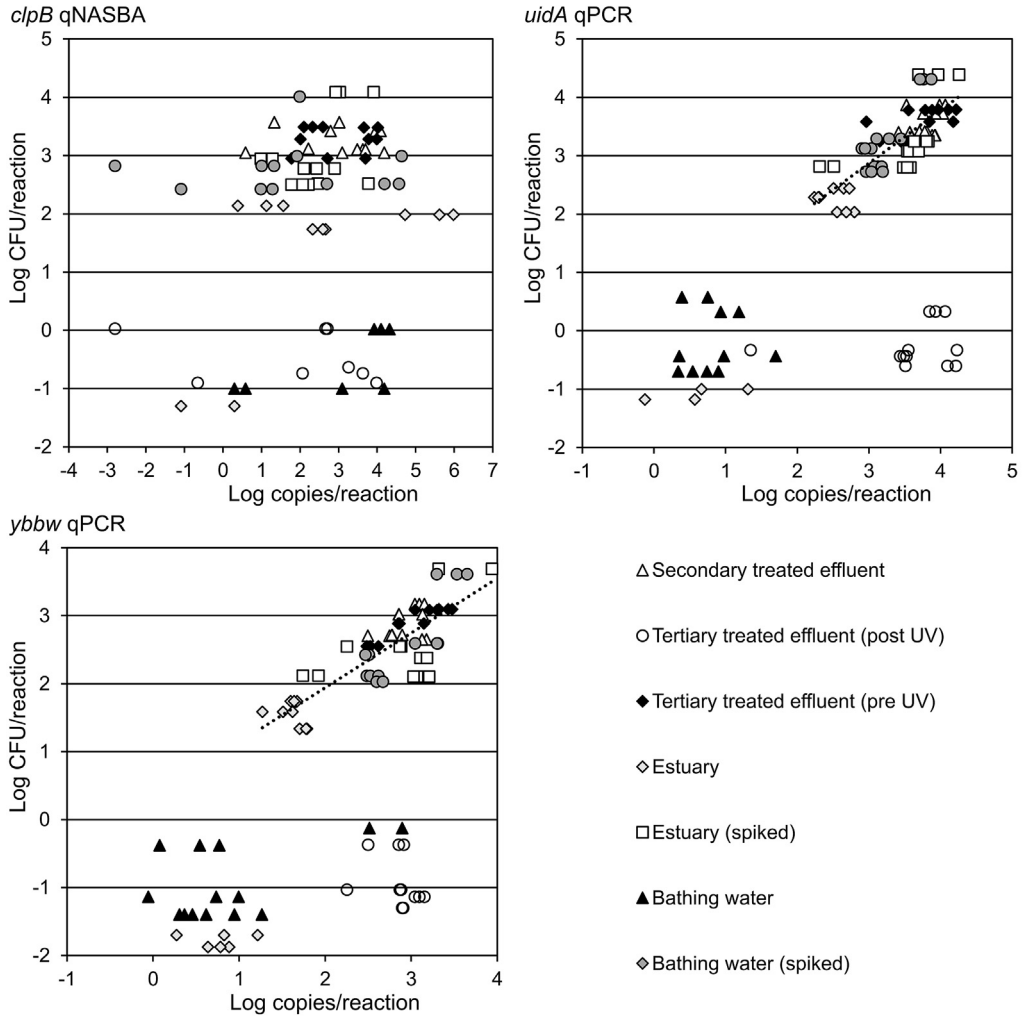

Fig. 3. Comparisons of copy numbers from *uidA* qPCR, *ybbW* qPCR and clpB qNASBA against colony counts for *E. coli* in environmental water samples. The dotted lines represent linear regressions in qPCRs with >100 CFU/PCR for the *uidA* assay ($R^2 = 0.639$) and >10 CFU/PCR for the *ybbW* assay ($R^2 = 0.673$).

Table 4
Comparison of the percentage of negative amplifications in three molecular assays over several *E. coli* concentration ranges.

CFU/reaction	% reactions with no amplification (n)		
	clpB qNASBA	*uidA* qPCR	*ybbW* qPCR
<1	51.3 (78)	29.5 (78)	21.9 (105)
1–10	29.6 (27)	25.9 (27)	n/a (0)
10–100	0.0 (18)	n/a (0)	0.0 (27)
100–1000	18.9 (90)	4.8 (63)	2.5 (120)
1000–10000	17.3 (75)	0.0 (120)	5.0 (60)
>10000	11.1 (27)	0.0 (27)	0.0 (3)

tested, whilst a *ybbW*-targeted qPCR assay and a *clpB* mRNA-targeted qNASBA assay detected all of the *E. coli strains* tested. This high level of inclusivity has also been demonstrated for other molecular assays for *E. coli* (Maheux et al., 2009). However, to date no single *E. coli* assay has been described which shows both complete inclusivity and complete exclusivity; for example all of the pre-existing assays for *E. coli* also detect *Shigella* spp. Here we have shown that the *ybbW* assay does not detect any of the tested *Shigella* spp. or other non-*E. coli* species and so is the most exclusive *E. coli* assay developed to date. Additionally, while the TBX culture, *clpB* NASBA and *uidA* qPCR assays also gave false positive results for

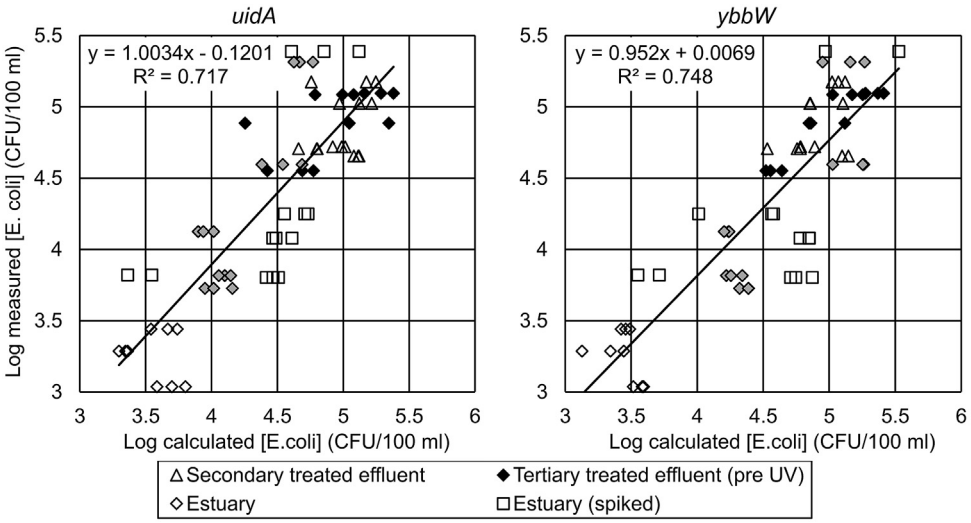

Fig. 4. Calculated *E. coli* concentrations from *uidA* and *ybbW* qPCR results against measured *E. coli* results for those samples with *E. coli* concentrations >1000 CFU/100 ml as measured by culture.

non-*E. coli Escherichia* species, the *ybbW* qPCR assay was shown to be exclusive to *E. coli*. The use of *E. coli* as an indicator of faecal contamination is a common practice internationally in food, recreational waters and potable waters (WHO, 2003, 2011; FAO, 2015), and so the increased level of exclusivity of the *ybbW* assay offers an advantage over other *E. coli* detection assays for public health applications.

The theoretical limit of quantification (LOQ) for the qNASBA was approximately 10^4 copies, higher than either qPCR assay studied. The study by Heijnen and Medema (2009), in which the qNASBA assay was first reported, did not attempt to characterise the assay in this way, but showed it was possible to detect *E. coli* from samples with 1 CFU/ml if the samples were heat shocked before RNA extraction. However, the relatively low amplification efficiency for the qNASBA assay for samples containing fewer than 10^4 copies found in this study indicates that this assay would not be suitable for quantification of *E. coli* in samples from most environmental samples. This is corroborated by the lack of any correlation between CFU counts and qNASBA results for environmental water samples collected from a range of sites across southwest England. While it might be expected that there would be no correlation between those samples with low *E. coli* concentrations such as bathing waters and estuarine waters, there was also no correlation with those samples with high *E. coli* levels such as sewage effluents. McLellan (2004) showed that there is a large diversity of *E. coli* genotypes associated with environmental waters and so it is likely that in the field samples collected in this study there were several different strains of *E. coli* present which may express genes at different levels (Vital et al., 2015). This would be reflected in a high degree of variability in qNASBA results as was found here. In this study, a subsample of bathing waters and estuarine waters which contained low concentrations of *E. coli* were spiked with a single strain (NCTC 9001) of *E. coli*, resulting in samples that were dominated by a single *E. coli* strain. The qNASBA results for these samples had large variability and so indicates that intra-strain gene expression variability is likely to cause as much variability in an mRNA based approach as would inter-strain gene expression variability. It is therefore probable that the lack of correlation between qNASBA and culture results was also due to the high LOQ for the qNASBA assay. The qNASBA assay had a high number of failed reactions for field samples, even for those samples with the highest CFU counts. For water samples collected at the tertiary sewage treatment works, those samples collected before UV treatment had the lowest failure rate and had the highest average RNA concentrations. After UV treatment RNA concentrations were lower and there was a small but significant decrease in RNA integrity number, which may account for some of the failed NASBA reactions in samples from this site.

It is also possible that failed qNASBA reactions were a result of the presence of inhibitory substances. However, Rutjes et al. (2006) showed that substances that often inhibit PCR reactions, had no effect on NASBA. The use of an internal amplification control (Rodríguez-Lázaro et al., 2004) in future studies would add certainty to any negative results obtained by both qNASBA and qPCR assays. The USA environmental protection agency (EPA) routinely uses salmon sperm DNA as an internal control in the qPCR assay for Enterococci in water (EPA, 2013) and could likely be adapted to work with the *E. coli* qPCR assays described in this study. However, this dsDNA internal amplification control would not be suitable for use with qNASBA because NASBA requires RNA as the starting material. An alternative, RNA based IAC would therefore be required in any further qNASBA method development.

A high degree of variability was found between qNASBA runs, which meant that all qNASBA data needed to be normalised across runs to allow for comparable data analysis. We suspect that the main contributing factor for this high inter-run variation was the large number of handling steps required when performing NASBA, as well and the requirement for multiple enzymes to work in tandem to achieve amplification. Unlike PCR which uses thermotolerant polymerases, the enzymes used in NASBA are not thermotolerant and must be added to the reaction mixture following denaturation of RNA at 65 °C. In this study, this was achieved by placing the enzymes into the caps of the reaction vessels and briefly centrifuging before the isothermal NASBA was carried out. This additional handling step introduced a period of time where the reaction was not held at a constant temperature, and may have resulted in some reformation of secondary RNA structure, rendering some of the target unavailable to the NASBA enzymes. During the course of this study, we found that it was possible to generate positive NASBA reactions with a single step reaction in which the enzymes were mixed with the rest of the reaction mixture, if denaturation was performed at 55 °C instead of 65 °C (data not shown). However this was not studied in detail and it is possible that not all templates would be equally denatured at

such low temperatures. Additionally, despite our best efforts to control the length of time taken to remove reaction plates, centrifuge and put the plates into the thermocycler for the isothermal reaction, there was some degree variability introduced at this point. The use of lab-on-a-chip (LOC) devices may help to resolve these issues, by allowing the separation of the enzymes from the reaction mixtures during denaturation, and then automatically introducing enzymes to the reaction under tightly controlled conditions. This would allow efficient NASBA reactions to be performed with a reduction in hands-on time.

The theoretical LOQ for both qPCR assays was less than 10 copies. No positive results were obtained for reactions containing fewer than 10 copies and so it was not possible to calculate the LOD. To accurately establish the theoretical limit of quantification and limit of detection for these assays, PCR containing between 1 and 10 copies would need to be performed and the results analysed according to published guidance (Wilrich and Wilrich, 2009). However, the data obtained in this study indicate that unlike qNASBA, both qPCR assays have similar detection limits to other *E. coli* qPCR methods (Maheux et al., 2009).

In contrast to qNASBA results for field samples, correlations were found between *E. coli* CFU counts and copy number results for both PCR assays. Both qPCR assays had comparable correlation coefficients within their linear ranges, but the *ybbW* assay performed slightly better than the *uidA* assay. Despite the good degree of correlation found for both qPCR assays using DNA from a single *E. coli* strain (Fig. 1), correlation coefficients fell to 0.639 and 0.673 for *uidA* and *ybbW* respectively when these assays were applied to environmental water samples. While it is likely that a large proportion of the decrease in correlation coefficient was due to error inherent in all microbiological methodologies, it is possible that variation was introduced by the presence of non-viable *E. coli* that were detected by PCR but not by the TBX culture method. The inability of PCR to distinguish between viable and non-viable bacterial cells is well documented, and is demonstrated in this study by the similarity in detected copy numbers for sewage effluent sampled before and after UV treatment, despite significant differences in CFU counts. The use of dyes such as propidium monoazide (PMA) in combination with PCR is well established as a means to discriminate viable from non-viable bacterial cells by PCR (Bae and Wuertz, 2009; Gensberger et al., 2014) and the adoption of these methods may serve to increase the reliability of these PCR assays when measuring viable *E. coli* in environmental waters.

For those samples with an *E. coli* concentration of >1000 *E. coli* CFU/100 ml (as measured by TBX culture) it was possible to calculate a predicted *E. coli* concentration based on gene copy numbers for both the *ybbW* and *uidA* qPCR assays. No significant differences in the proportion of calculated to measured *E. coli* concentration was found between sample types for either the *ybbW* and *uidA* qPCR assays. This suggests that qPCR, when used in combination with the magnetic bead nucleic acid extraction method, is able to predict CFU counts in samples taken from several water types ranging from relatively dirty sewage effluent to relatively clean recreational seawater.

5. Conclusions

This study has shown that qNASBA is not suitable for the quantification of *E. coli* in water, due to a lack of correlation between mRNA copy number and *E. coli* CFU. Nor is qNASBA suitable for *E. coli* detection in environmental samples due to a large proportion of negative amplification results even at high *E. coli* concentrations. In contrast, qPCR is capable of reliably detecting *E. coli* from environmental waters and the development of the *ybbW* qPCR assay will allow more unambiguous identification of *E. coli* than has previously been achievable using rapid test methods. Additionally, the adoption of internal amplification controls and viability dyes such as PMA will likely make qPCR more able to accurately quantify *E. coli*. This would be of great benefit to water quality and food safety testing laboratories due to the rapidity of PCR based methods when compared with culture based methods. However, despite the theoretical low detection levels achievable by qPCR, the main factor in limiting detection in real samples is often the concentration and isolation of *E. coli* DNA. This may be addressed in part by filtering larger volumes of test water, however this is likely to be impractical for routine sample analysis and may introduce larger concentrations of inhibitory substances. It is therefore recommended that future study of rapid, molecular based methods for microbial detection and quantification in environmental waters should focus on the improvement of DNA extraction method efficiency, rapidity and reliability.

Acknowledgements

This work was supported jointly by the Natural Environment Research Council and the Biotechnology and Biological Sciences Research Council [grant number BB/M025837/1]; and the Centre for Environment, Fisheries and Aquaculture Science [grant number SCN529].

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.watres.2017.08.032.

References

Bae, S., Wuertz, S., 2009. Rapid decay of host-specific fecal Bacteroidales cells in seawater as measured by quantitative PCR with propidium monoazide. Water Res. Engl. 43 (19), 4850–4859. http://dx.doi.org/10.1016/j.watres.2009.06.053.

Edberg, S.C., Rice, E.W., Karlin, R.J., Allen, M.J., 2000. Escherichia coli: the best biological drinking water indicator for public health protection. In: Symposium Series (Society for Applied Microbiology), vol. 29, pp. 106S–116S. England.

EPA, 2013. Method 1609: Enterococci in Water by TaqMan® Quantitative Polymerase Chain Reaction (qPCR) with Internal Amplification Control (IAC) Assay. US Environmental Protection Agency Office of Water, Washington DC.

FAO, 2015. Standard for Live and Raw Bivalve Molluscs. Codex International Food Standards.

Frahm, E., Obst, U., 2003. Application of the fluorogenic probe technique (TaqMan PCR) to the detection of Enterococcus spp. and Escherichia coli in water samples. J. Microbiol. Methods 52 (1), 123–131. http://dx.doi.org/10.1016/S0167-7012(02)00150-1.

Gensberger, E.T., Polt, M., Konrad-Koszler, M., Kinner, P., Sessitsch, A., Kostic, T., 2014. Evaluation of quantitative PCR combined with PMA treatment for molecular assessment of microbial water quality. Water Res. Engl. 67, 367–376. http://dx.doi.org/10.1016/j.watres.2014.09.022.

Heijnen, L., Medema, G., 2009. Method for rapid detection of viable Escherichia coli in water using real-time NASBA. Water Res. 43 (12), 3124–3132. http://dx.doi.org/10.1016/j.watres.2009.04.025, 2009/05/30.

Maheux, A.F., Bissonnette, L., Boissinot, M., Bernier, J.L.T., Huppé, V., Picard, F.J., Bérubé, È., Bergeron, M.G., 2011. Rapid concentration and molecular enrichment approach for sensitive detection of Escherichia coli and Shigella species in potable water samples. Appl. Environ. Microbiol. 77 (17), 6199–6207. http://dx.doi.org/10.1128/AEM.02337-10.

Maheux, A.F., Picard, F.J., Boissinot, M., Bissonnette, L., Paradis, S., Bergeron, M.G., 2009. Analytical comparison of nine PCR primer sets designed to detect the presence of Escherichia coli/Shigella in water samples. Water Res. 43 (12), 3019–3028. http://dx.doi.org/10.1016/j.watres.2009.04.017.

McLellan, S.L., 2004. Genetic diversity of Escherichia coli isolated from urban rivers and beach water. Appl. Environ. Microbiol. 70 (8), 4658–4665. http://dx.doi.org/10.1128/AEM.70.8.4658-4665.2004.

McQuillan, J.S., Robidart, J.C., 2017. Molecular-biological sensing in aquatic environments: recent developments and emerging capabilities. Curr. Opin. Biotechnol. 45, 43–50. https://doi.org/10.1016/j.copbio.2016.11.022.

Rodríguez-Lázaro, D., D'Agostino, M., Pla, M., Cook, N., 2004. Construction strategy for an internal amplification control for real-time diagnostic assays using nucleic acid sequence-based amplification: development and clinical application. J. Clin. Microbiol. 42 (12), 5832–5836. http://dx.doi.org/10.1128/JCM.42.12.5832-5836.2004.

Rutjes, S.A., van den Berg, H.H.J.L., Lodder, W.J., de Roda Husman, A.M., 2006. Real-

time detection of noroviruses in surface water by use of a broadly reactive nucleic acid sequence-based amplification assay. Appl. Environ. Microbiol. U. S. 72 (8), 5349–5358. http://dx.doi.org/10.1128/AEM.00751-06.

SCA, 2016. The microbiology of recreational and environmental waters (2014) - Part 3-Methods for the isolation and enumeration of Escherichia coli (including E. coli O157:H7). Methods Exam. Waters Assoc. Mater. 25–31. http://www.standingcommitteeofanalysts.co.uk/methods/Microbiology/rec.html.

Silkie, S.S., Tolcher, M.P., Nelson, K.L., 2008. Reagent decontamination to eliminate false-positives in Escherichia coli qPCR. J. Microbiol. Methods 72 (3), 275–282. http://dx.doi.org/10.1016/j.mimet.2007.12.011.

Vital, M., Chai, B., Ostman, B., Cole, J., Konstantinidis, K.T., Tiedje, J.M., 2015. Gene expression analysis of E. coli strains provides insights into the role of gene regulation in diversification. ISME J. Engl. 9 (5), 1130–1140. http://dx.doi.org/10.1038/ismej.2014.204.

WHO, 2003. Guidelines for Safe Recreational Water Environments. In: Coastal and Fresh Waters, vol. 1.

WHO, 2011. Guidelines for Drinking-water Quality, fourth ed.

Wilrich, C., Wilrich, P.-T., 2009. Estimation of the POD function and the LOD of a qualitative microbiological measurement method. J. AOAC Int. U. S. 92 (6), 1763–1772.