Environmental Pollution 236 (2018) 100–110

Contents lists available at ScienceDirect

Environmental Pollution

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

Performance evaluation of *Bacteroidales* genetic markers for human and animal microbial source tracking in tropical agricultural watersheds☆

Pornjira Somnark [a], Natcha Chyerochana [b], Skorn Mongkolsuk [b, c, d], Kwanrawee Sirikanchana [b, d, *]

[a] Applied Biological Sciences, Chulabhorn Graduate Institute, Chulabhorn Royal Academy, Lak Si, Bangkok, Thailand
[b] Research Laboratory of Biotechnology, Chulabhorn Research Institute, Lak Si, Bangkok, Thailand
[c] Department of Biotechnology, Faculty of Science, Mahidol University, Bangkok, Thailand
[d] Center of Excellence on Environmental Health and Toxicology (EHT), Ministry of Education, Bangkok, Thailand

ARTICLE INFO

Article history:
Received 31 July 2017
Received in revised form
21 December 2017
Accepted 17 January 2018

Keywords:
Microbial source tracking
PCR
Bacteroidales
Tropical watershed
Southeast Asia

ABSTRACT

Microbial source tracking (MST) DNA-based assays have been used to successfully solve fecal pollution problems in many countries, particularly in developed nations. However, their application in developing countries has been limited but continues to increase. In this study, sixteen endpoint and quantitative PCR (qPCR) assays targeting universal and human-, swine-, and cattle-specific *Bacteroidales* gene markers were modified for endpoint PCR, evaluated for their performance with sewage and fecal samples from the Tha Chin watershed and subsequently validated with samples from the Chao Phraya watershed, Thailand. Sample sizes of 81 composite samples (from over 1620 individual samples) of farm animals of each type as well as 19 human sewage samples from the Tha Chin watershed were calculated using a stratified random sampling design to achieve a 90% confidence interval and an expected prevalence (i.e., desired assay's sensitivity) of 0.80. The best universal and human-, swine-, and cattle-specific fecal markers were BacUni EP, HF183/BFDrev EP, Pig-2-Bac EP, and Bac3 assays, respectively. The detection limits for these assays ranged from 30 to 3000 plasmid copies per PCR. The positive predictive values were high in universal and swine- and cattle-specific markers (85–100%), while the positive predictive value of the human-specific assay was 52.2%. The negative predictive values in all assays were relatively high (90.8–100%). A suite of PCR assays in Thailand was established for potential MST use in environmental waters, which supports the worldwide applicability of *Bacteroidales* gene markers. This study also emphasizes the importance of using a proper sample size in assessing the performance of MST markers in a new geographic region.

1. Introduction

Microbial source tracking (MST) is an approach used to identify source markers that indicate fecal hosts by taking advantage of genotypic and phenotypic traits of microorganisms in animal gastrointestinal tracts that are specific to their particular host species, for example, pigs, cows, chicken, goat, and sheep (Field and Samadpour, 2007; Sargent et al., 2011; Tran et al., 2015). The presence and distribution of these source markers in water resources can indicate contamination from particular hosts or groups. Fecal marker microorganisms belong to bacterial orders or higher (Bernhard and Field, 2000a; Kildare et al., 2007; Ohad et al., 2016; Shanks et al., 2008; Siefring et al., 2008), bacterial genera and species (Bernhard and Field, 2000a, 2000b; Haugland et al., 2010; Ohad et al., 2016; Weidhaas et al., 2010), or bacterial strains (Jofre et al., 2014; Leknoi et al., 2017; Sirikanchana et al., 2014; Wangkahad et al., 2017). The bacterial class Bacteroidetes and, in particular, the bacterial order *Bacteroidales* have been most frequently applied in MST (Ahmed et al., 2016; Harwood et al., 2014; Tran et al., 2015). Molecular techniques, specifically

☆ This paper has been recommended for acceptance by Dr. Harmon Sarah Michele.

* Corresponding author. Research Laboratory of Biotechnology, Chulabhorn Research Institute, Lak Si, Bangkok, Thailand.
E-mail address: kwanrawee@cri.or.th (K. Sirikanchana).

https://doi.org/10.1016/j.envpol.2018.01.052
0269-7491/

endpoint and quantitative PCR (qPCR) methods, have been designed to detect the order *Bacteroidales* to indicate the specific origins of fecal sources, such as human (Bernhard and Field, 2000a; Kildare et al., 2007; Siefring et al., 2008), swine (Dick et al., 2005; Mieszkin et al., 2009), cattle (Bernhard and Field, 2000a; Layton et al., 2006; Shanks et al., 2006b), chicken (Kobayashi et al., 2013; Lu et al., 2007), and universal (general) (Bernhard and Field, 2000b; Kildare et al., 2007; Siefring et al., 2008). An advantage of using molecular methods is that the samples are stable for long periods with proper storage and thus can be reanalyzed with assays that are developed or improved after the samples are stored.

qPCR can quantitatively measure fluorescently labeled amplification products using a hydrolysis probe or non-specific intercalating dye, which in principle can provide more sensitive detection than endpoint PCR (Botes et al., 2013; Smith and Osborn, 2009). qPCR can also deliver more accurate measurement because it avoids amplification bias that may occur in later cycles by measuring fluorescent signal in real-time during an exponential amplification phase (Smith and Osborn, 2009). Although a number of studies have demonstrated good performance of general and source-specific qPCR assays for MST application (Kildare et al., 2007; Layton et al., 2006; Mieszkin et al., 2009; Shanks et al., 2008), challenge of implementing qPCR methods in developing countries remains an obstacle. Endpoint PCR, in contrast, involves lower instrumental and yearly maintenance costs, lower costs for each PCR (i.e., consumables, reagents, etc.), and less technical skill (Riedel et al., 2014), all of which suit the application for use in developing countries. In fact, promotion of PCR laboratories in developing countries has also been increasingly emphasized (Ragheb and Jimenez, 2014; World Health Organization, 2016). Consequently, there is a pressing need to investigate endpoint PCR assays for both general and source-specific fecal markers for use in developing countries. Nonetheless, for some fecal sources, no suitable endpoint PCR assays have been established (Ballesté et al., 2010; Bernhard and Field, 2000a; Gourmelon et al., 2007; Shanks et al., 2010; Toledo-Hernandez et al., 2013). Therefore, it is intriguing to adapt the widely used qPCR assays to endpoint platforms and evaluate the performance of the modified endpoint PCR counterparts. Apart from challenges in adapting the assay from qPCR to PCR platforms, including probable loss of detection sensitivity and accuracy (Botes et al., 2013; Smith and Osborn, 2009), variability in PCR performance could further be impacted by laboratory protocols and reagents used and therefore could vary from laboratory to laboratory (Al-soud and Radstrom, 1998; Riedel et al., 2014; Rissanen et al., 2010; Saeidi et al., 2017; Silva and Domingues, 2015). In practice, improved qPCR performance over PCR counterparts with similar primers and target amplicons appeared to be assay-specific, as no improved sensitivity was shown in one study, while 10 times improvement was shown in another study (Dagher et al., 2004; Riedel et al., 2014).

Moreover, although fecal markers have been identified and developed for source tracking applications in many countries, the microorganism communities in gastrointestinal tracts can be dissimilar for each host species in different geographical areas, which may be due to climate, food, antibiotics, and other region-specific factors (Adami and Cavazzoni, 1999; Looft et al., 2014; Lu et al., 2013; Mah et al., 2008; Shanks et al., 2011). This potential variation can lead to variability in the performance of fecal markers among regions. Sets of performance criteria facilitate the selection of the best assay when method evaluation is critical. Sensitivity, specificity, accuracy, positive predictive value, and negative predictive value have been used as criteria in the evaluation of MST assays and are determined by measuring the intrinsic properties of MST markers in regions of interest (Ballesté et al., 2010; Cao et al., 2015; Fremaux et al., 2009; Gómez-Doñate et al., 2015; Green et al., 2012; Harwood et al., 2013; Li et al., 2015, 2016; Odagiri et al., 2015; Schriewer et al., 2013; Wangkahad et al., 2017). Although no strict guidelines are available, suitable values of sensitivity and specificity have been recommended to be 0.80 and 0.90, respectively (Ahmed et al., 2016; Boehm et al., 2013; U.S. Environmental Protection Agency, 2005). It is also clear that these performance characteristics depend on the number of samples collected for testing markers; however, a lack of sample size guidelines has been acknowledged (Ahmed et al., 2016; Harwood et al., 2014; Santo Domingo et al., 2007). In epidemiology, a minimum sample size can be computed to ensure that samples represent the prevalence of markers or diseases in a population within a specified confidence level and acceptable error (Thompson, 2012). When evaluating MST assay performance, such a minimum sample size can be calculated based on the expected marker prevalence, as defined by the assay's sensitivity. Calculations of sample size vary with the method of sampling. Three common statistical sampling designs used in animal populations are simple random sampling, cluster or two-stage sampling, and stratified random sampling (Leon et al., 2011). Simple random sampling, or random sampling without replacement, is the simplest sampling design and involves randomly selecting samples with equal probabilities of selection from a population (Thompson, 2012). Cluster or two-stage sampling is a type of sampling method in which a whole population is divided into clusters, and certain clusters are randomly selected for individual sampling. In contrast, stratified random sampling classifies a population into strata of similar traits; samples are then collected from each stratum by simple random sampling. The stratified random selection method is generally used when strata are classified by geographical region or other factors. Because one goal of this study was to establish a set of MST markers that could be used throughout the Tha Chin watershed, stratified random sampling was most appropriate for calculating the sample size of target hosts to achieve the expected sensitivity.

The ultimate goal of this study was to establish a set of endpoint PCR assays for tracking universal and human-, swine-, and cattle-specific fecal markers in agricultural watersheds in Thailand. The specific objectives were to 1) perform method modification of the widely used endpoint and qPCR primers to endpoint PCR assays for universal (general) and human-, swine-, and cattle-specific fecal origins; 2) evaluate performance of the modified endpoint PCR assays on sewage and fecal samples of known origins from Tha Chin watershed and validate in Chao Phraya watershed; and 3) select the best modified endpoint PCR assays and define their corresponding detection limits.

2. Materials and methods

2.1. Study area

The Tha Chin watershed is located in the central part of Thailand (Fig. 1). It covers 13,477.16 km^2 comprising 13 provinces. The Tha Chin River is 325 km long, receives water from the Chao Phraya River in Chai Nat Province and is connected to the Gulf of Thailand in Samut Sakhon Province. The four main provinces (Chai Nat, Suphan Buri, Nakhon Pathom, and Samut Sakhon) through which the Tha Chin River passes were selected as sample collection regions in this study (Hydro and Agro Informatics Institute, 2012a). For a decade, the Tha Chin River had one of the poorest water qualities among the 65 monitored freshwater sources in Thailand (Pollution Control Department (PCD), 2015). The parameters that pose water quality problems for the Tha Chin River include dissolved oxygen (DO), biochemical oxygen demand (BOD), total coliform bacteria (TCB), fecal coliform bacteria (FCB), and ammonia-nitrogen (NH_3-N) (Pollution Control Department (PCD),