Available online at www.sciencedirect.com
ScienceDirect
Journal of Microbiological Methods 70 (2007) 227–235
Journal of Microbiological Methods
www.elsevier.com/locate/jmicmeth

Genotyping of *Escherichia coli* from environmental and animal samples

James Higgins [a,*], Christina Hohn [b], Sally Hornor [c], Mark Frana [d], Mary Denver [e], Rolf Joerger [f]

[a] *USDA-ARS, Beltsville, MD 20705, USA*
[b] *NCSU College of Veterinary Medicine, Raleigh, NC, USA*
[c] *Biology Department, Anne Arundel Community College, Arnold, MD 21012, USA*
[d] *Department Biological Sciences, Salisbury University, Salisbury, MD 21801, USA*
[e] *The Maryland Zoo in Baltimore, Baltimore, MD 21217, USA*
[f] *Department Animal and Food Sciences, University of Delaware, Newark, DE 19716, USA*

Received 6 March 2007; received in revised form 17 April 2007; accepted 17 April 2007
Available online 24 April 2007

Abstract

The triplex PCR of Clermont et al. [Clermont, O., Bonacorsi, S., Bingen, E., 2000. Rapid and simple determination of the *Escherichia coli* phylogenetic groups. Appl. Environ. Microbiol. 66, 4555–4558.] was used to genotype *E. coli* isolates from the Mid-Atlantic region of the USA, obtained from freshwater, animal internal organs, and feces. Of 445 isolates subjected to genotyping, 118 isolates (26%) were genotype A, 111 (25%) genotype D, 140 (31%) genotype B1, and 76 (17%) genotype B2. All four genotypes were present in three sets of freshwater stream samples. When isolates from chicken cecal ingesta, cecal mucosa, and tracheal mucosa were screened, there was selective distribution of genotypes in these organs. Genotype D was rarely encountered in feces, milk, and intestinal tissues of dairy cows, while all four genotypes were represented in goose feces. Isolates from the feces of zoo animals reared in the US demonstrated a predominance of genotype B1. Thirty-six of the A isolates in our overall collection were subgenotype A_0, in which none of the three amplicons are observed; confirmation that these isolates were *E. coli* was done using an ancillary *lacZ* PCR assay. We conclude that the genotyping triplex PCR assay, used in combination with traditional culture methods, can be useful in categorizing *E. coli* from environmental and veterinary sources in the Mid-Atlantic region of the USA.
Published by Elsevier B.V.

Keywords: *E. coli*; Genotyping; *lacZ*; Water; Feces; Extraintestinal

1. Introduction

The Gram-negative bacterium *Escherichia coli* is a prominent member of the bacterial microbiota of the environment and in the feces of many species of birds and mammals. The distribution of this organism has historically been used as an index of water quality and in the past two decades, it has also been used as an indicator organism for source tracking purposes (Scott et al., 2002). There is some disagreement among microbial ecologists as to the ultimate utility of this approach. A number of investigators have reported some degree of success with experimental approaches using *E. coli* for source tracking (Ahmed et al., 2005; Fogarty et al., 2003; Hamilton et al., 2006; Hassan et al., 2005; Johnson et al., 2004; Leung et al., 2004; Seurinck et al., 2003; Stoeckel et al., 2004). However, Gordon and colleagues (Gordon, 2001; Gordon et al., 2002; Gordon and Cowling, 2003) argue that the genetic and phenotypic variability inherent in *E. coli* isolates from the same host and/or environmental niche precludes its use for source tracking, a viewpoint shared by Lasalde et al. (2005) based on their observations in work with restriction fragment length polymorphisms/PFGE analyses of *E. coli* isolates.

Many genotyping studies done on clinical isolates, as well as environmental and veterinary isolates, use assemblages of *E. coli* previously identified via biochemical and/or culture-based approaches (see for example Gordon et al. (2002), or Fogarty et al. (2003)). These isolates may be subjected to PCR for various pathogenicity genes; genotyping methods involving REP-PCR, ERIC-PCR, and PFGE; or sequencing of regions of genes of interest. The correct interpretation of data generated using these methods obviously is reliant to some degree on the fidelity of the method used to identify the *E. coli* in the first

* Corresponding author. USDA-ARS, Rm 202, Bldg 173, 10300 Baltimore Blvd, Beltsville 20705, MD, USA. Tel.: +1 301 504 6443; fax: +1 504 6608.
E-mail address: tarbandu12@juno.com (J. Higgins).

0167-7012/$ - see front matter. Published by Elsevier B.V.
doi:10.1016/j.mimet.2007.04.009

place. However, it is clear from an examination of the literature that there is some variation in the use of methods to assign a definitive identity to putative isolates of *E. coli*. Most assays described in the literature tend to focus on disease-causing strains of the organism; for example, a *gadA*/*gadB* PCR can successfully detect pathogenic strains of *E. coli*, with no cross-reactivity with non-*E. coli* coliforms (Grant et al., 2001).

One protocol that arguably comes closest to being a 'universal' method for *E. coli* identification is a triplex PCR assay that assigns the bacteria to one of four main phylogenetic groups (Clermont et al., 2000). This genotyping triplex assay relies on the amplification of a 279 bp fragment of the *chuA* gene; a 211 bp fragment of the *yjaA* gene; and a 152 bp fragment of TSPE4.C2, a noncoding region of the genome. The presence or absence of combinations of these three amplicons is used to assign the *E. coli* to a given genotype: A, B1, B2, or D. The utility of this method for analysis of *E. coli* population structure in environmental, veterinary, and clinical samples is well-documented (Dixit et al., 2004; Duriez et al., 2001; Escobar-Paramo et al., 2004b; Girardeau et al., 2005; Gordon et al., 2002; Gordon and Cowling, 2003; Johnson and Stell, 2000; Johnson et al., 2005a,b; Nougayrede et al., 2006; Power et al., 2005; Zhang et al., 2002).

Somewhat disappointingly, adoption of the genotyping approach of Clermont et al. by North American microbial ecologists has been slow. The benefits of using a universal, established protocol for cataloging *E. coli* diversity have been amply demonstrated in clinical microbiology scenarios, so there is reason to believe genotyping can be useful in other situations as well. Accordingly, we were interested in testing the utility of the genotyping triplex of Clermont et al. as a means of identifying *E. coli* recovered from animal feces and water in the eastern USA. The specificity of these assays for other enteric bacteria encountered in these environments was also examined, as well as the utility of a PCR targeting the *lacZ* gene as an ancillary method of confirmation of *E. coli*. This paper reports conclusions drawn from these efforts.

2. Materials and methods

2.1. DNA extraction

Bacterial colonies from agar plates were cored using Pasteur pipettes and deposited into 0.2 ml microfuge tubes containing 100 µl of Instagene™ matrix (Bio-Rad, Hercules, California). Broth cultures (1.5 ml volume) were pelleted by centrifugation and the pellet resuspended in 100–200 µl of Instagene™ matrix. Both colony plugs and broth pellets were subjected to a 15 min heating step at 56 °C, followed by heating at 100 °C for 8 min; tubes were centrifuged to pellet the Instagene resin, and 3 µl to 5 µl volumes of supernatant were used as template for PCR.

2.2. Triplex PCR for genotyping

In our hands the reaction conditions reported by Clermont et al. (2000), consisting of 30 cycles of a 5 s denaturation step, and a 10 s annealing/extension step, failed to yield consistent results with our positive control *E. coli* ATCC 25922, which, as genotype B2, should amplify all three products. We subsequently found that thermal cycling conditions of: 35 cycles of 94 °C for 15 s, 60 °C for 30 s, and 72 °C for 45 s, provided reproducible amplification of the three bands in *E. coli* 25922 when used in conjunction with 1.5 mM $MgCl_2$, 20 pmol each of the TSPE4.C2 and *chuA* primers, 80 pmol each of the *yjaA* primers, and 5 µl 10× PCR reaction buffer in a total reaction volume of 50 µl, including 5 µl DNA template (all reagents from Invitrogen, Gaithersburg, Maryland). PCR was conducted on a PTC-200 model thermal cycler (MJ Research/BioRad, Hercules, CA). *E. coli* strain ATCC 25922 was used as a positive control for all genotyping PCRs. For those rare (i.e., <2%) genotyping PCR assays which failed to generate all three amplicons from the positive control, the assay was repeated until the positive control provided satisfactory results. Isolates refractory to this approach were subjected to *lacZ* PCR (below).

2.3. lacZ and 16S rRNA PCR

One feature of the genotyping triplex PCR is that bacteria are assigned to genotype A based on either the presence of only the *yjaA* amplicon, or the absence of all three amplicons (Clermont et al., 2000). Isolates failing to amplify any of the three amplicons are referred to as subgenotype "A_0" (Escobar-Paramo et al., 2004a,b). To confirm that our A_0 isolates were *E. coli*, we subjected them to PCR for a 365-bp fragment of the *E. coli* beta-galactosidase gene (hereafter referred to as *lacZ*) using the "Big Z forward" primer: 5′ gca gcg ttg ttg cag tgc 3′ and the "Big Z reverse" primer: 5′ gtc ccg cag cgc aga c 3′; corresponding to nucleotides 2458–2475 and 2807–2822, respectively, of the *E. coli* beta-galactosidase gene (Genbank Accession No. V00296; Zell and Fritz, 1987). Reagents for PCR included 1 *U Taq* polymerase, 1.5 mM $MgCl_2$, 50 pmol each primer, 5 µl 10× PCR buffer, and 5 µl DNA template, in a total reaction volume of 50 µl; thermal cycling conditions were: 35 cycles of 15 s at 94 °C, 30 s at 60 °C, 45 s at 72 °C, with a final extension of 7 min at 72 °C, on a PTC-200 model thermal cycler.

For amplification and sequencing of the 16S rRNA gene, we used the primers of Kazor et al. (2003), with 1 U *Taq* polymerase, 1.5 mM $MgCl_2$, 5 µl 10× PCR buffer, 50 pmol each primer, and 5 µl DNA template in a 50 µl total reaction volume; thermal cycling conditions were: 35 cycles of 15 s at 94 °C, 1:30 min at 58 °C, 2 min at 72 °C, with a final extension of 7 min at 72 °C, on a PTC-200 model thermal cycler.

For sequencing of PCR products, 5 µl aliquots of product were treated with 2 µl ExoSAP-IT® reagent (USB Corp., Cleveland, OH) according to the manufacturer's recommendations, and 2 µl portions of the ExoSAP-IT® treated PCR product were used as template for dye-terminator cycle sequencing with the Big Dye™ 3.1 kit (PE Biosystems, Foster City, CA). Sequencing reactions were electrophoresed on an Applied Biosystems ABI 3100 model automated fluorescent sequencer (Foster City, California). Forward and reverse reads were used to create consensus sequences which were analyzed using the clustalw alignment and tree-drawing components of the Lasergene™ v. 6 software package (DNAStar, Madison, Wisconsin).