Title: Validation of a five-colony pool analysis using Multiplex Real Time PCR for detection of diarrheagenic E. coli

Running title: Five-colony Real-time multiplex PCR for diarrheagenic E. coli

Barletta F\textsuperscript{1}, Ochoa TJ\textsuperscript{1,2}, Ecker L\textsuperscript{3}, Gil AI\textsuperscript{3}, Lanata CF\textsuperscript{3}, Cleary TG\textsuperscript{2}

\textsuperscript{1}Instituto de Medicina Tropical, Universidad Peruana Cayetano Heredia, Lima, Perú.
\textsuperscript{2}Center for Infectious Diseases, University of Texas School of Public Health, TX.
\textsuperscript{3}Instituto de Investigación Nutricional, Lima, Perú.

* Corresponding author:
Francesca Barletta Solari, MSc.
Instituto de Medicina Tropical
Universidad Peruana Cayetano Heredia
Av. Honorio Delgado 340
San Martin de Porras, Lima 33, Perú
Phone 51-1-482-3910
Fax: 51-1-482-3404
E-mail: francescabarletta@yahoo.es
ABBREVIATIONS

PCR  polymerase chain reaction, EPEC enteropathogenic  *E. coli*, ETEC enterotoxigenic *E. coli*, STEC Shiga toxin producing *E. coli*, EAEC enteroaggregative *E. coli*, EIEC enteroinvasive *E. coli*, DAEC diffusely adherent *E. coli*, Tm DNA melting temperature
ABSTRACT

Five *E. coli* colonies/patient were studied to evaluate the reliability of multiplex real time PCR assay for detection diarrheagenic *E. coli* groups, using a pool of 5 colonies rather than individual colonies. Sensitivity and specificity were 98% and 100%, respectively at a fifth of the cost of the individual colony analysis.

Diarrheal disease remains a leading cause of mortality and morbidity in the world (1,2), particularly in developing countries. In Peru, enteric infections represent the third most important cause of mortality in the first five years of life (6). The diarrheagenic *Escherichia coli* (*E. coli*) are important causes of diarrhea in children in the developing world (9) and to a lesser extent the developed world (4). *E. coli* associated with diarrhea have been classified into six groups based on clinical, epidemiological and molecular criteria (Table 1). Although these *E. coli* commonly cause gastroenteritis in children living in developing countries, they are not routinely sought in clinical laboratories. A multiplex real time PCR to simultaneously identify 8 virulence genes associated with the six classes of diarrheagenic *E. coli* has been previously developed (5). To make this assay less expensive for use in developing countries, we conducted this study to determine the sensitivity and specificity of the multiplex real time PCR analysis using a pool the five colonies rather than analyzing individual colonies.

Stool specimens were collected from 1025 children 2m to 12m of age from a diarrhea surveillance study conducted in low socioeconomic communities of Lima, Perú during 2006 and 2007 (11), and were cultured on MacConkey agar plates. After overnight incubation at 37°C, five lactose positive *E. coli* colonies/patient were sub-cultured on a new
MacConkey agar plate to assure pure individual colonies. Prototypical strains (DAEC-5019, STEC-91/8626, ETEC-H10407, EPEC-2348/69, EIEC-213 and EAEC-042) were used as positive controls, and *E. coli* C600 was used as a negative control.

DNA was extracted by boiling either a single colony in 50 µL or a pool of 5 colonies in 100 µL of PCR water for 5min followed by 15min at room temperature and a centrifugation at 13000rpm for 15min. 2 µL of this crude lysate was used as template in a 25 µL total PCR reaction volume.

PCR was performed using a PTC-200 thermal cycler and real time fluorescence monitoring by a Chromo 4 optical detector (Biorad). Each multiplex PCR assay contained the sets of primers previously validated (5), 0.5U of Phusion polymerase (Finnzyme) in high fidelity Phusion buffer with a final concentration of 200µM dNTPs and 4mM MgCl$_2$. SYBR Green I (Cambrex Bio Science) was diluted as recommended by the manufacturer and a hot-start of 98°C for 30sec was used to prevent non specific amplification. The protocol consisted of incubation a 98°C for 20sec, 60°C for 20sec, 72°C for 30sec, and 75°C for 10sec. After 25 cycles a melting curve was determined with a ramp speed of 2.5°C/sec between 73°C and 95°C with a reading every 0.2°C. Melting points were automatically calculated by the Opticon Monitor software which plotted the negative derivative of fluorescence with respect to temperature (-d(F)/dT versus T).

We analyzed 1158 pools from children (<1 yr of age) with (858) and without (292) diarrhea, who were enrolled in the cohort study. 27.5% (319/1158) of pools were positive for diarrheagenic *E. coli* and 72.5% (839/1158) were negative. The most common pathogens were EAEC (12.5% diarrhea vs. 16.1% controls), EPEC (5.8% diarrhea vs. 7.9%...
controls), DAEC (3.7% diarrhea vs. 2.1% controls) and ETEC (2.5% diarrhea vs. 1.0% controls). EIEC were not detected in these samples.

We also analyzed 2065 individual colonies (1595 colonies from positive pools and 500 colonies from 100 randomly selected negative pools). There was a trend, although not statistically significant, for a higher number of positive colonies per sample in the diarrhea group than control for some pathogens (i.e. DAEC 3.04 ± 1.64 in diarrhea vs. 1.94 ± 1.47 in controls) in the diarrhea group. For purpose of the pool versus individual colony analysis diarrhea and control samples were grouped together.

On a pool-colony analysis, the specificity was 100% for each diarrheagenic *E. coli* group, and the sensitivity ranged from 96% to 100% (**Table 1**). There were no false positives; in no case (0/319) did a peak occur at a melting point inconsistent with the prototype strain genotype. There were 5 false negative pools (5/100) as indicated by individual colony analysis which were positive for DAEC, EPEC, ETEC and EAEC (1 strain each). The sensitivity was 98%; when false negative pool samples were run a second time the sensitivity improved to 100%.

Among positive samples, mixed infections could also be detected (**Figure 1**). The most common mixed infections were EAEC + EPEC (9/20), EAEC + ETEC (3/20), EPEC + ETEC (3/20). There was no competition between different set of primers used to detect the genes associated with diarrheagenic *E. coli* virulence.

Conventional PCR for diagnosis of diarrheagenic *E. coli* has limitations in the developing world. The cost, the time delay required to do gel analysis of PCR products, and the inability to analyze large numbers of strains, are major impediments to this approach. For this reason, we developed and validated a multiplex real-time PCR for detection of
diarrheagenic *E. coli* groups (5). The “gold standard” is the individual colony analysis, but the cost limits its use as a routine test for use in clinical laboratories.

In this study, we were able to save time and effort involved in testing for virulence factors, reducing the number of gene detection assays by use of a sensitive and specific colony pool multiplex real-time PCR. An advantage of this technique is that there is no need to run an electrophoresis gel to determine the presence of the amplicons because each has a characteristic melting temperature (Tm) that is detected in the denaturation curve. Indeed, this approach also allows identification of mixed infections involving two or three diarrheagenic *E. coli*. Individual positive colonies can be defined in a subsequent PCR if required for epidemiologic or antibiotic susceptibility studies. The rapid detection of diarrheagenic *E. coli* has important treatment implications. Some diarrheagenic *E. coli* (e.g. ETEC) respond to antimicrobial agents while for others (e.g. STEC) antibiotics should be avoided. Furthermore, diarrheagenic *E. coli* are commonly resistant to commonly used antibiotics (10) so that optimal treatment depends on rapid detection of the specific pathogen.

The choice of five colonies for analysis represents a compromise between the cost of analysis and the need to detect infection.

In summary, the analysis of a pool of 5 colonies for the detection of diarrheagenic *E. coli* by multiplex real-time PCR is a dramatically more cost effective, sensitive, and specific technique when compared to the individual colony analysis.

ACKNOWLEDGMENTS
Theresa Ochoa is supported by PHS - FIC 1K01TW007405 and Tom Cleary is supported by the PHS - NICHD R01- HD051716. This study was partially funded by Dr. Claudio Lanata’s Institutional Research Fund. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Child Health and Human Development or the National Institutes of Health.

No conflict of interest for all authors.

REFERENCES


FIGURE LEGENDS

Figure 1: Mix infection detected in a pool analysis colony showing both an EPEC \([eae^+]\), and an ETEC \([st^+lt^-]\) (A), and confirmed by individual colony analysis corresponding to an ETEC \([st^-lt^+]\) colony (B) and an EPEC \([eae^-]\) colony (C).
## Table 1: Sensitivity and Specificity of the pool analysis

<table>
<thead>
<tr>
<th>Diarrheagenic E. coli</th>
<th>Individual Colony n (%)</th>
<th>Pool Analysis n (%)</th>
<th>Sensitivity of the Pool Analysis</th>
<th>Specificity of the Pool Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>EAEC (aagR)</td>
<td>155 (13.4)</td>
<td>154 (13.3)</td>
<td>99%</td>
<td>100%</td>
</tr>
<tr>
<td>EPEC (eaeA)</td>
<td>74 (6.4)</td>
<td>73 (6.3)</td>
<td>99%</td>
<td>100%</td>
</tr>
<tr>
<td>DAEC (daaD)</td>
<td>40 (3.5)</td>
<td>38 (3.3)</td>
<td>95%</td>
<td>100%</td>
</tr>
<tr>
<td>ETEC (st, lt)</td>
<td>25 (2.2)</td>
<td>24 (2.1)</td>
<td>96%</td>
<td>100%</td>
</tr>
<tr>
<td>STEC/EHEC (eaeA, stx1, stx2)</td>
<td>6 (0.5)</td>
<td>6 (0.5)</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>ALL</td>
<td>300</td>
<td>295</td>
<td>98%</td>
<td>100%</td>
</tr>
</tbody>
</table>

EPEC = Enteropathogenic E. coli, EAEC = Enteroaggregative E. coli, DAEC = Diffusele adherent E. coli, ETEC = Enterotoxigenic E. coli, STEC = Shigatoxin-like E. coli
Figure 1

A

B

C

Temperature (°C)