**Identification of *Escherichia coli* from groups A, B1, B2 and D in drinking water in Brazil**

Renato H. Orsi, Nancy C. Stoppe, Maria Inês Z. Sato and Laura M. M. Ottoboni

**ABSTRACT**

The presence of *Escherichia coli* in drinking water is an indication of fecal contamination and can represent a risk of waterborne diseases. Forty-nine *E. coli* strains isolated from different sources of drinking water (distribution system, well, spring and mineral water) were placed into the phylogenetic groups A (15 strains), B1 (19 strains), B2 (2 strains) and D (13 strains). Approximately 30% of the strains analyzed belonged to groups B2 and D, which usually include potentially extraintestinal pathogenic strains. Moreover, the assignment of the strains to different phylogenetic groups indicates that different contamination events occurred in these waters. These results were compared with the distribution of *E. coli* strains isolated from two rivers and two dams into the phylogenetic groups. A significant difference was observed when the distribution of drinking water strains into the phylogenetic groups was compared to the results obtained from the Guarapiranga Dam and the Jaguari and Sorocaba Rivers. The results obtained in this work suggest that PCR-based methods can be used for a rapid assessment of potentially pathogenic *E. coli* strains in water samples.

**Key words** | *Escherichia coli*, drinking water, phylogenetic groups

**INTRODUCTION**

*Escherichia coli* is a normal inhabitant of the gastrointestinal tract of mammals. The *E. coli* strains are usually referred to as commensal, intestinal pathogenic or extraintestinal pathogenic (Russo & Johnson 2000). The pathogenic strains have been associated with several diseases including diarrhea, urinary tract infections and meningitis (Russo & Johnson 2005). In developing countries, diarrheal diseases are often associated with infant and child deaths (Murray & Lopez 1996; Sobel et al. 2004).

The presence of *E. coli* in water is an indication of fecal contamination and represents a risk of disease (Leclerc et al. 2001). According to Hunter et al. (2001) a fraction of the world’s population (~20%) has no access to safe drinking water. This fact, in conjunction with inadequate sanitation, leads to millions of deaths every year (Hunter et al. 2001). In developing countries, infantile diarrhea is associated with typical *E. coli* EPEC strains (Trabulsi et al. 2002). Since humans are considered to be reservoirs for these microorganisms (Trabulsi et al. 2002), infantile diarrhea can be associated with inappropriate management of domestic sewage and water quality (Ashbolt 2004).

*E. coli* strains can be separated into 4 main phylogenetic groups: A, B1, B2 and D (Selander et al. 1986; Herzer et al. 1990). Usually, the commensal strains are placed into the phylogenetic groups A and B1 (Johnson et al. 2001) and the extraintestinal pathogenic strains into group B2 and, to a lesser extent, group D (Picard et al. 1999; Johnson & Stell 2000). The intestinal pathogenic strains are usually assigned to groups A, B1 and D (Pupo et al. 1997). Clermont et al. (2000) described a simple PCR-based method that uses a combination of the chuA and yjaA genes and the DNA fragment TSPE4.C2 to assign *E. coli* strains to the phylogenetic groups A, B1, B2 and D. This methodology has been used, with different purposes, by authors interested...
in the assignment of *E. coli* strains to the phylogenetic groups (Gordon & Cowling 2003; Dixit et al. 2004).

The aim of the present work was to investigate the distribution of *E. coli* strains, isolated from different sources of drinking water, into the phylogenetic groups A, B1, B2 and D, and by doing this, assess the presence of potential pathogenic strains in drinking water. The results were compared with the data gathered by the analysis of *E. coli* strains isolated from two rivers and two dams.

**MATERIALS AND METHODS**

**Escherichia coli** strains

Forty nine strains of *E. coli* were isolated from different sources of drinking water routinely analyzed for potability by CETESB (Environment Protection Agency of São Paulo State, Brazil). The number of strains, the water source and the range of sample contamination are shown in Table 1. From raw waters (rivers and dams), a total of 261 *E. coli* strains were isolated (60 from the Jaguari River, 68 from the Sorocaba River, 65 from the Billings Dam and 68 from the Guarapiranga Dam).

**Strains isolation**

Water samples were collected in sterile disposable bottles according to Standard Methods 20th Edition (APHA 1998). The samples were chilled for transportation and examined within 24 hours. Samples were analyzed by the membrane filter technique using m-Endo agar LES (APHA 1998). Typical colonies (dark-red color with metallic sheen) were transferred to EC broth and incubated at 44.5 ± 0.2°C for 24 hours. Positive tubes were streaked on EMB agar (Merck). Isolated, typical *E. coli* colonies (metallic green colonies) were tested for lactose fermentation, oxidase test, citrate utilization, L-lysine decarboxylase, motility, glucose and sucrose fermentation, indole production, tryptophan desamination, hydrogen sulfide production and urea hydrolysis. Isolates that exhibited an *E. coli* biochemical profile were grown in LB broth, isolated in LB agar and stored in a 25% glycerol solution at −70°C. The 49 strains isolated from drinking water were isolated from 49 different samples, and only one strain of each EC positive sample was selected for PCR analysis. *Escherichia coli* ATCC 25922 was used as the control strain.

**Phylogenetic group determination**

The phylogenetic group of each strain was determined by using triplex PCR as described by Clermont et al. (2000). For this, the genomic DNA of each strain was isolated using the GenomicPrep Cells and Tissue DNA Isolation Kit (Amersham Biosciences), following the specifications of the manufacturer. The isolated DNA was used in PCR amplifications that were carried out in a PT-100 thermal cycler (MJ Research) to identify the *chuA* (279 bp) and the *yjaA* (211 bp) genes and the DNA fragment TSPE4.C2 (152 bp). Each reaction contained 2 μl of 10X Taq polymerase buffer, 1.5 mM MgCl2, 20 pmol of each primer (Clermont et al. 2000), 2 μM of each dNTP, 2.5 units of Taq polymerase (Amersham Biosciences) and 200 ng of genomic DNA, in a 20 μl mixture. The amplification conditions were: initial denaturation at 94°C for 5 min followed by 30 cycles of 30 s at 94°C, 30 s at 55°C, 30 s at 72°C, and a final extension of 7 min at 72°C. The amplification products were separated by electrophoresis in a 2% agarose gel containing ethidium bromide. After electrophoresis, the gel was photographed and the strains were assigned to phylogenetic group B2 (*chuA + , yjaA + ) or D (*chuA + , yjaA − ) or B1 (*chuA − , TSPE4.C2 + ) or A (*chuA − , TSPE4.C2 − ). A negative control (reaction lacking the template DNA) was included in all amplifications performed. One *E. coli* strain from each group (A, B1, B2 and D) was used as positive control in the amplifications.

<table>
<thead>
<tr>
<th>Source</th>
<th>Number of strains</th>
<th>Total coliform (TC) densities (CFU/100 ml)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distribution system (Tap water and cistern)</td>
<td>10</td>
<td>1–100</td>
</tr>
<tr>
<td>Wells</td>
<td>18</td>
<td>3–1,500</td>
</tr>
<tr>
<td>Spring</td>
<td>14</td>
<td>3–1,500</td>
</tr>
<tr>
<td>Mineral water</td>
<td>7</td>
<td>1–200</td>
</tr>
</tbody>
</table>
Statistical analysis

The associations between phylogenetic groups and water source (river, dam, distribution system, well, spring and mineral water) were assessed by means of contingency Chi-squares performed with the Biostat v. 2.0 software (Ayres et al. 2000).

RESULTS AND DISCUSSION

*Escherichia coli* are abundant in human and nearly all warm-blooded animal feces \((10^7 - 10^9/g)\). Therefore, this organism is a reliable index of recent fecal contamination and indicates a risk of waterborne diseases (Szewzyk et al. 2000; Leclerc et al. 2001). These bacteria have been used worldwide to measure the microbial water quality, and the national and international regulations establish the absence of this indicator microorganism in any 100 ml sample of drinking water (WHO 1996).

In this work, 49 *E. coli* strains were isolated from EC positive drinking water samples. As shown in Table 1, the level of contamination of these samples ranged from low polluted (1 total coliform/100 ml) to high polluted (1,500 total coliforms/100 ml). Since contaminated water represents a risk for human health, we decided to gain more information about the *E. coli* strains isolated from these water samples by assigning them to the phylogenetic groups A, B1, B2 and D, using the PCR based technique described by Clermont et al. (2000).

As shown in Table 2, only 2 strains, 1 isolated from a drinking-water distribution system and the other from spring water, were assigned to phylogenetic group B2. *E. coli* strains belonging to group B2 are highly pathogenic and frequently responsible for extraintestinal infections in humans (Lecointre et al. 1998; Duriez et al. 2001). Therefore, the presence of these strains, especially in the distribution system water sample, deserves attention.

The remaining strains isolated from the drinking-water distribution system were assigned to group A (3 strains), B1 (4 strains) and D (2 strains). The presence of *E. coli* in the water distribution system could be an indication of inadequate decontamination or recontamination of the water in the distribution system (Mossel 1982).

Other 13 strains isolated from spring water were assigned to groups A (2 strains), B1 (4 strains) and D (7 strains). It is interesting to note that the highest percentage of strains belonging to group D was found in the spring water samples. *E. coli* strains from group D have fewer virulence determinants than strains from group B2 (Lecointre et al. 1998). Extraintestinal pathogenic *E. coli* can be found in group D (Picard et al. 1999) and according to Clermont et al. (2000), *E. coli* O157:H7 could belong to this phylogenetic group. Domestic and wildlife animal feces may be the source of pristine water contamination by these bacteria (Leclerc et al. 2002) that have been pointed out as the principal cause of diarrheal diseases in North America (Kassenborg et al. 2004).

The majority of the strains isolated from wells and mineral water were allocated in groups A and B1. That is, among the 18 strains isolated from wells, 9 belonged to group B1, 6 to A and 3 to D. The 7 strains isolated from mineral water were assigned to group A (4 strains), B1 (2 strains) and group D (1 strain). The phylogenetic groups A and B1 usually include commensal *E. coli* strains (Duriez et al. 2001).

A significant association was observed between the groups and spring water \((\chi^2 = 6.50; \text{D.F.} = 1; \text{P} = 0.011)\) where strains from groups B2 and D were more prevalent than strains from groups A and B1. This result suggests that the spring water samples, besides being more polluted (Table 1), have a higher prevalence of strains that may cause extraintestinal diseases than the other sources of water studied herein. No other significant association was observed between a specific source of water and the phylogenetic group categories (commensal/intestinal strains from groups A and B1 or extraintestinal strains from groups B2 and D, data not shown).

### Table 2 | *Escherichia coli* strains distribution into the phylogenetic groups A, B1, B2 and D

<table>
<thead>
<tr>
<th>Drinking water source</th>
<th>A</th>
<th>B1</th>
<th>B2</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distribution system</td>
<td>3 (30%)</td>
<td>4 (40%)</td>
<td>1 (10%)</td>
<td>2 (20%)</td>
</tr>
<tr>
<td>Well</td>
<td>6 (33%)</td>
<td>9 (50%)</td>
<td>0</td>
<td>3 (17%)</td>
</tr>
<tr>
<td>Spring</td>
<td>2 (14%)</td>
<td>4 (29%)</td>
<td>1 (7%)</td>
<td>7 (50%)</td>
</tr>
<tr>
<td>Mineral</td>
<td>4 (57%)</td>
<td>2 (29%)</td>
<td>0</td>
<td>1 (14%)</td>
</tr>
</tbody>
</table>
The distribution of the drinking water *E. coli* strains into the phylogenetic groups was compared with the distribution of strains isolated from 2 dams (Billings and Guarapiranga) and 2 rivers (Jaguari and Sorocaba) located in São Paulo State, Brazil. The dams are important water reservoirs located in the Metropolitan São Paulo region and receive more than 40,000 visitors who use these waters for recreational activities, including swimming and sailing, during the weekends. Both dams receive a heavy pollution load, mainly from irregular domestic discharges and to decrease microbiological risk, 21 sites from these dams are assessed weekly to evaluate bathing water quality (CETESB 2004).

The Jaguari and Sorocaba Rivers are part of the São Paulo State surface inland waters monitoring program. This program evaluates the water quality using two indexes, one for water supply and the other for aquatic life protection. These rivers are located in urbanized and industrialized areas and the demand for hydro resources is strong. The levels of chemical and biological parameters indicate that the main source of pollution in these rivers is derived from domestic sewage (CETESB 2004).

The distribution of strains isolated from drinking water into the phylogenetic groups A, B1, B2 and D showed a significant difference from the distribution of strains isolated from the rivers. These differences were observed when the results of both rivers were analyzed together ($\chi^2 = 21.216$; D.F. = 2; $P < 0.0001$) and separated ($\chi^2 = 14.556$; D.F. = 2; $P = 0.0007$ for the Sorocaba River and $\chi^2 = 17.988$; D.F. = 2; $P = 0.0001$ for the Jaguari River). The difference between the distribution of drinking water and river strains into the phylogenetic groups may be attributed to different sources of contamination and/or different selective pressures in the drinking and river waters.

No significant difference was observed between the distribution of strains into the phylogenetic groups when the results obtained from the Billings Dam were compared with the ones obtained from drinking water ($\chi^2 = 2.56$; D.F. = 2; $P = 0.2780$). However, a significant difference was observed when the distribution of drinking water strains into the phylogenetic groups was compared with the distribution of strains isolated from the Guarapiranga Dam ($\chi^2 = 13.495$; D.F. = 2; $P = 0.0012$).

**CONCLUSIONS**

The results obtained in this work suggest that PCR-based methods, applied to identify the phylogenetic groups A, B1, B2 and D, can be used for a rapid assessment of potentially pathogenic *E. coli* in water samples. This would be helpful as an initial screening assay and could be used to complement more time consuming traditional tests that use serological and animal assays. In this study, it was observed that about 30% of the drinking water strains analyzed belonged to groups B2 and D, which usually includes potentially extraintestinal pathogenic strains. Moreover, the fact that the *E. coli* strains analyzed were assigned to different phylogenetic groups suggests that different contamination events occurred on these samples of drinking water. Finally, the spring water samples analyzed seemed to have a predominance of potentially extraintestinal pathogenic strains and, therefore, deserve special attention.

**ACKNOWLEDGEMENTS**

This work was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) grant (2000/05721-8). RHO received a fellowship from CAPES. LMMO had a research fellowship from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). We thank the Microbiology and Parasitology Section of CETESB for detection and isolation of *E. coli* strains.

**REFERENCES**


