Genetic diversity of *Escherichia coli* isolates from surface water and groundwater in a rural environment

Maria Laura Gambero, Monica Blarasin, Susana Bettera
and Jesica Giuliano Albo

ABSTRACT

The genetic characteristics among *Escherichia coli* strains can be grouped by origin of isolation. Then, it is possible to use the genotypes as a tool to determine the source of water contamination. The aim of this study was to define water aptitude for human consumption in a rural basin and to assess the diversity of *E. coli* water populations. Thus, it was possible to identify the main sources of fecal contamination and to explore linkages with the hydrogeological environment and land uses. The bacteriological analysis showed that more than 50% of samples were unfit for human consumption. DNA fingerprinting analysis by BOX-PCR indicated low genotypic diversity of *E. coli* isolates taken from surface water and groundwater. The results suggested the presence of a dominant source of fecal contamination. The relationship between low genotypic diversity and land use would prove that water contamination comes from livestock. The genetic diversity of *E. coli* isolated from surface water was less than that identified in groundwater because of the different hydraulic features of both environments. Furthermore, each one of the two big strain groups identified in this basin is located in different sub-basins, showing that hydrological dynamics exerts selective pressure on bacteria DNA.

Key words | BOX-PCR, *E. coli*, groundwater contamination, rural environment

Maria Laura Gambero (corresponding author)
Susana Bettera
Laboratory of Food Microbiology,
    Section Microbiology,
National University of Río Cuarto,
Ruta Nac. 36 - Km. 601,
Río Cuarto,
Córdoba,
Argentina
E-mail: *mgambero@exa.unrc.edu.ar*

Monica Blarasin
Jesica Giuliano Albo
Laboratory Hydrogeology, Section Geology,
National University of Río Cuarto,
Ruta Nac. 36 - Km. 601,
Río Cuarto,
Córdoba,
Argentina

INTRODUCTION

The water quality is strongly influenced by urbanization, livestock and agricultural activity in agroecosystems. Increasing public concern over access to clean water has led to improve water source protection and management activities. Rural environments present unique challenges because they serve as water sources for a variety of activities (e.g. drinking water for humans and livestock, irrigation and recreation). Furthermore, both point and non-point sources of microbial pollution may be found (e.g. livestock operations, household septic systems and wildlife) (Chin 2006; Sinclaira *et al.* 2009).

Identifying sources of fecal pollution in waters is necessary to reduce the potential for human contact with enteric pathogens. Bacterial indicator organisms are commonly used to assess the presence of enteric pathogens in the water. The most common indicator microorganisms are fecal coliforms (FC) and its subgroup *Escherichia coli*. These microorganisms share a common feature: they all can inhabit the intestines of warm-blooded animals, including wildlife, livestock and humans, and therefore they can be excreted in the feces (Ibekwe *et al.* 2011; Carlos *et al.* 2012). However, *E. coli* presence in contaminated water will not provide information as to the actual source of contamination. This information is important, as fecal contamination resulting from human sources will establish a high public health risk because of the possible presence of pathogenic organisms. Additionally, if the fecal source is known, suitable management actions can be implemented to prevent further contamination and to mitigate the health risks (Harwood *et al.* 2000).

While *E. coli* has diverse genotypic and phenotypic characteristics, some are shared among strains exposed to similar environments due to selection pressure. If some of the characteristics among *E. coli* strains can be grouped by origin of isolation, then it is possible to use these genotypes as a tool to determine the source of unknown bacteria. The level of selective pressure exerted in environment may be a useful criterion for identifying the fecal contamination sources in water (Unno *et al.* 2010; Ibekwe *et al.* 2011).

Methods to determine sources of fecal pollution have usually included phenotypic and genetic characterization of fecal indicator bacteria isolate from surface waters. The study of the genetic diversity of *E. coli* isolates from water may inform on the diverse origins of contaminant sources. DNA fingerprinting techniques, such as pulsed-field gel electrophoresis (PFGE), ribotyping and repetitive element palindromic polymerase chain reaction (Rep-PCR), are reported to be an effective way to track the source of microbial pollution without many of the complications of antibiotic resistance analysis. (McLellan 2004; Mohapatra *et al.* 2007; Bonnet *et al.* 2009; Cuevas López *et al.* 2009; Lyautey *et al.* 2010; Unno *et al.* 2010). The Rep-PCR method uses naturally occurring conserved intergenic palindromic sequences for amplification by PCR. The PCR amplification uses primers, including the BOX A1R primer, complimentary to the palindromic sequences to produce unique DNA banding patterns (fingerprints) that are strain-specific. The Rep-PCR provides high taxonomic resolution and may act as a rapid detector of diversity and evolution of the microbial genomes being studied (Versalovic *et al.* 1994; Dombek *et al.* 2000).

Many investigators have evaluated the use of Rep-PCR method to determine sources of environmental fecal bacteria in waters. Therefore, Dombek *et al.* (2000) reported the use of Rep-PCR DNA fingerprint to differentiate *E. coli* strains isolated from human and animal sources. Somarelli *et al.* (2007) reported molecular fingerprinting (BOX-PCR) of fecal microorganism isolated from the Conesus lake. Ma *et al.* (2011) used BOX-PCR to differentiate sources of environmental *E. coli* and Carlos *et al.* (2012) to discriminate *E. coli* strains from animal and waste sources. The investigators showed that this technique has greater discriminatory power. Studies on the genetic characterization and diversity of *E. coli* isolated from groundwater in order to assess the fecal contamination source are not fully developed in general. To our knowledge, our investigation group has made the first studies in Argentina using BOX-PCR DNA fingerprinting to evaluate genetic diversity of *E. coli* isolated groundwater in order to assess fecal contamination source. Gambero *et al.* (2016) have reported genetic and antibiotic studies of *E. coli* isolated from a sedimentary aquifer in the rural environment in Cordoba, Argentina.

In this framework, the aim of this study was to define water aptitude for human consumption considering general bacteriological characteristics and to elucidate the diversity of *E. coli* water populations, identifying the main sources of fecal contamination and exploring linkages with the hydrogeological environment and land uses.

STUDY AREA

Climate, land uses and hydrogeology

The Barranquita-Knutzen basin covers an area of 292 km^2. The basin is located in the Chacopampeana plain, in the south of Córdoba province, Argentina (Figure 1). The climate is dry sub-humid with little water excess. The average annual precipitation is 732 mm, divided into two periods, the wettest from November to March with 70% of the annual total precipitation. Water for human consumption and livestock activities is mostly supplied by the unconfined aquifer. The land is used mainly for farming activity, which includes soybean, wheat and corn crops, and makes use of pesticides and fertilizers (phosphates, sulphates and urea). On the other hand, livestock activity includes breeding of cattle, pigs, sheep, horses and poultry, being worth mentioning that frequently, the streams are used as drinking water for the cattle.

The unconfined aquifer is made up of loessical type sediments (mainly very fine sands and silts) interlayered with paleochannels that lie at different depths (sands and gravels with high hydraulic conductivity). Groundwater flows in a general almost WE direction from the piedmont to the low land areas (Figure 2), but clearly shows different local flow direction and the depth of the water table is variable, ranging from 2.5 m to 24.0 m. Both characteristics, groundwater flow and water table depth, are mainly controlled by the relief
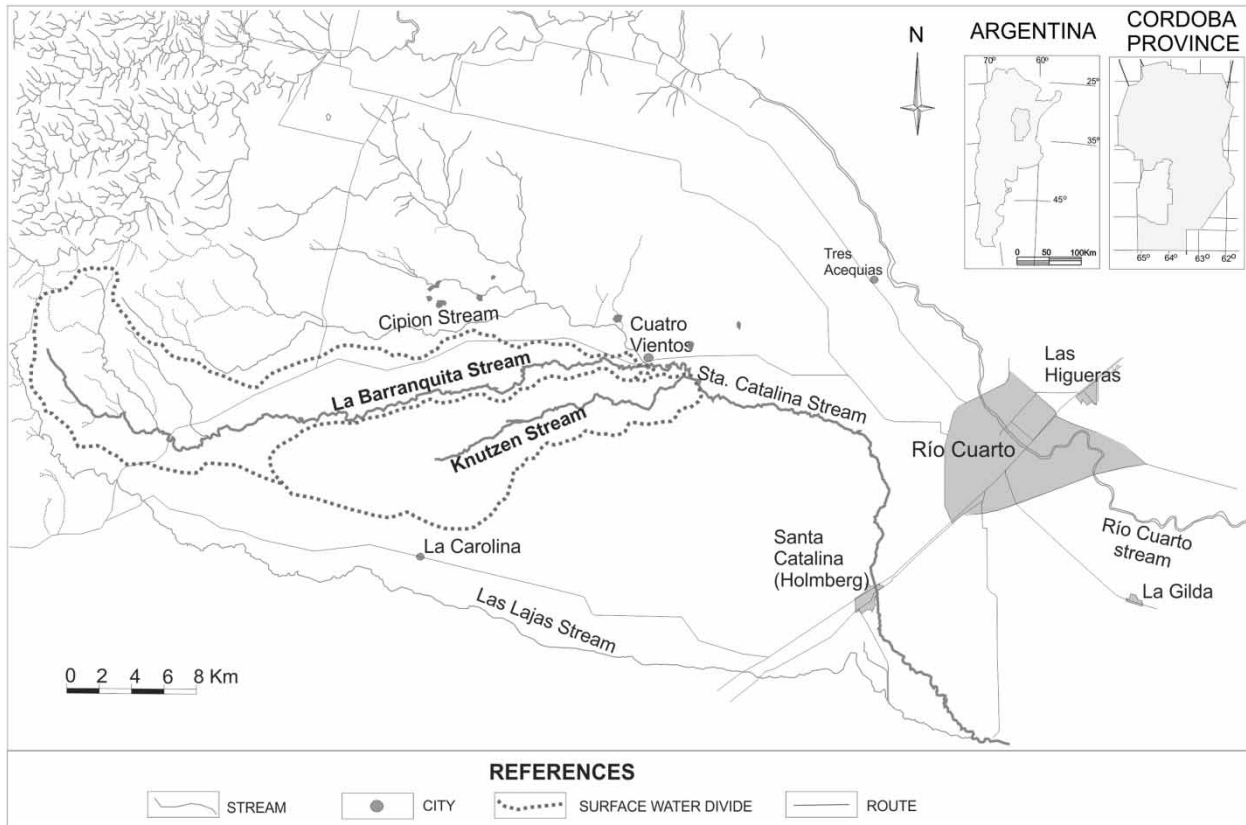
Figure 1 │ Location of study area in Córdoba, Argentina.

(Giuliano Albo & Blarasin 2014). The lithology of the unsaturated zone (USZ) and water table depth are key factors that control the aquifer recharge and the aquifer vulnerability to contamination (Giuliano Albo & Blarasin 2014).

The surface water and groundwater salinity values are between 508–2,050 μS/cm and 654–1,078 μS/cm, respectively (Figure 2). The water geochemical features are variable, from calcium–sodium bicarbonate to sodium sulphate water types, which depend on the geochemical processes described for this aquifer in Giuliano Albo & Blarasin (2014).

METHODS

Water sampling

Environmental water samples ($n = 41$) were collected in the wet season (spring). Thirty-six samples were collected from groundwater and five from surface water. To ensure the

representativeness of groundwater samples, each well was emptied three times. The water samples (500 mL) were taken for bacteriological indicator analysis according to Argentine Food Code (AFC 2012). All samples were collected aseptically in sterile bottles, stored at 4°C and analyzed within 24 h of collection in the laboratory of the National University of Rio Cuarto (Argentina).

Bacteriological analysis

The determination of Heterotrophic Plate Counts (HPC) was carried in nutrient agar, incubated at 35°C during 24 h. The total coliforms (TC) and FC were determined through the multiple-tube fermentation (MTF) technique. Probability tables (McCrady tables) were used to determine the Most Probable Number (MPN) and to estimate the number of coliform organisms per 100 ml of water. TC were incubated in MacConkey broth at 35°C during 24–48 h and FC in Brilla broth at 44.5°C during 24 h. The presence of *E. coli* was
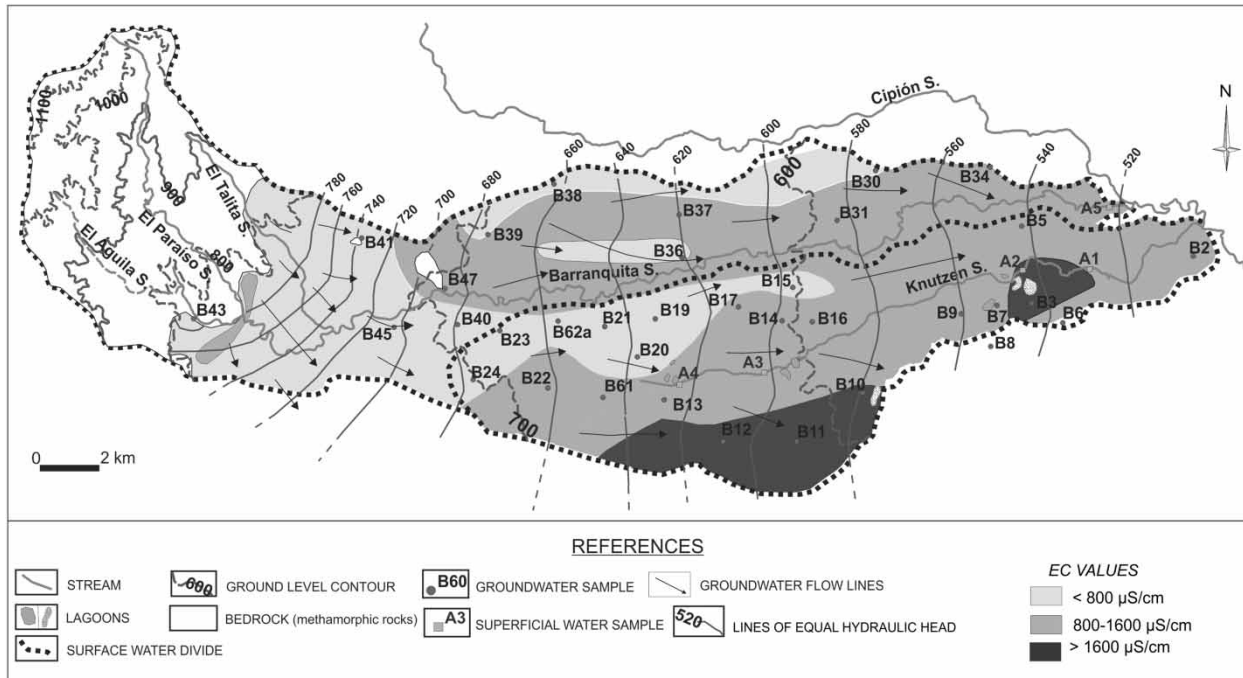
Figure 2 │ Location map of study area showing sampling sites, groundwater direction and salinity in the Barranquita-Knutzen basin, Córdoba, Argentina.

determined in 100 ml of sample in MacConkey broth incubated at 35°C during 24–48 h. Then, an aliquot was spread onto EMB agar plates and incubated at 35°C for 24 h. Isolates were confirmed as *E. coli* by using a series of biochemical tests, including indole, Voges-Proskauer, methyl red tests and the inability to grow on citrate agar (IMViC). The presence of *Pseudomonas aeruginosa* was determined on a volume of 100 ml of sample in Crystal Violet broth incubated at 35°C during 24–48 h. The isolation was carried out in Cetrimide agar plates and colonies were confirmed by the following biochemical tests: oxidase, growth at 42°C, and pigments production in Agar P and F. The methodology was carried out according to *Standard Methods for the Examination of Water and Wastewater* (APHA 2005).

DNA isolation and BOX-PCR fingerprinting

Genomic DNA from individual pure cultures of *E. coli* isolates was extracted by the use of Real Genomic DNA Extraction kit (Real Biotech Corporation, RBC), according to the manufacturer indications. The extracted DNA was used as the template DNA for Rep-PCR fingerprinting. The Rep-PCR assay mixture was prepared as described by

Rademaker & Bruijn (1996) modified. The final reaction mix (25 μl) consisted of 1X PCR buffer (Promega, Biodynamics), 1.5 mM MgCl$_2$, 0.2 mM of each deoxynucleoside triphosphate (Promega, Biodynamics), 10 pmol/μl of the BOX A1R primer (5′-CTACGGCAAGGCGACGCTGACG-3′) (BOX-PCR), 0.04 U/μl of *Taq* polymerase (Promega) and 3 μl of dilution template DNA.

Amplification was performed in a MJ Research thermocycler as follows: after an initial denaturation at 94°C for 5 min, 35 cycles of denaturation (94°C, 1 min), annealing (40°C, 2 min) and extension (72°C, 2 min) were performed, followed by a final extension (72°C, 10 min). A negative (distilled water) control was included in the PCR experiments. The PCR products were separated by horizontal gel electrophoresis on a 1.5% agarose gel. Gels were stained in 0.5 μg/ml ethidium bromide solution and the gel images were taken in a MiniBisPRO in TIF Format. The Rep-PCR fingerprints of the isolates were normalized with a 1 kb DNA ladder.

Statistical analysis

The Rep-PCR fingerprints were subjected to cluster analysis using the SPSS, v.11.5. software. This analysis was also

used to evaluate the similarity of the strains. In this program, the obtained results were converted to a two dimensional rectangular matrix data of binary codes for all *E. coli* strains, [+] for presence and [–] for absence of the PCR product band. The relationship between the fingerprints was estimated by the proportion of shared bands using the Dice coefficient and dendrograms were generated using the UPGMA method (Unweighted Pair Group Method with Arithmetic-Averages), with the NTSYS v. 2.1 software. (Dombek *et al.* 2000; Baldy-Chudzik & Stosik .2005; Somarelli *et al.* 2007).

RESULTS

Groundwater and surface water quality

The results of bacteriological analysis of groundwater samples are shown in Table 1. It was seen that 53 per cent of samples were unfit for human consumption taking into account all the general bacteriological indicators. FC and *E. coli* were detected as evidence of fecal contamination. Thus, seven wells were positive for FC (range, 4–430 MPN/100 ml) and eight wells were positive for *E. coli*. B17, B31 and B43 wells were greatest affected by the bacteriological contamination. The results of bacteriology analysis of surface

water samples are shown in Table 2. It was detected that three samples were positive for FC (range, 75–2,100 MPN/100 ml) and all the samples were positive for *E. coli*.

E. coli recovery from surface water and groundwater

From the 41 water samples extracted, 17 strains were confirmed as *E. coli* by biochemical tests, from which eight were isolated from groundwater and nine from surface water. The land use activities neighboring the sites where the *E. coli* strains were isolated are listed in Table 3.

Analysis of BOX-PCR DNA fingerprinting patterns

BOX primers were used to generate PCR fingerprints for *E. coli* isolates obtained from surface water and groundwater to determine their diversity and the population structure. Strains with BOX-PCR fingerprint patterns with ≥85% similarity were considered clonal populations (Malathum *et al.* 1998; McLellan 2004). The optimized technique of BOX-PCR method in this study produced high visualization of DNA fragments and was replicable.

The total *E. coli* ($n = 17$) isolates produced high quality BOX-PCR fingerprinting and generated DNA fragments

Table 1 │ Microbiological groundwater quality

Bacteriological parameters	Limits according to AFC	Min-Max	Samples numbers above the limit of AFC
Groundwater ($n = 36$)			
HPC (cfu/mL)	>500 cfu/ml	$0–1.6 \times 10^3$ cfu/ml	3
TC (MPN/100 mL)	≤3 MPN/100 ml	0–480 MPN/100 ml	12
E. coli	Absence	–	8
Ps. aeruginosa	Absence	–	8

Table 2 │ Microbiological surface water quality

| Bacteriological parameters	Limits according to AFC	Min-Max	Samples numbers above the limit of AFC
Surface water ($n = 5$)			
HPC (cfu/mL)	>500 cfu/ml	$2 \times 10^3–7 \times 10^3$ cfu/ml	5
TC (MPN/100 mL)	≤3 MPN/100 ml	300–2,100 MPN/100 ml	5
E. coli	Absence	–	5
Ps. aeruginosa	Absence	–	1

Table 3 | Land use characteristic and strains of *E. coli* isolated in each site

Barranquita-Knutzen basin	Site	Land use	Strains (*n* = 17)
Groundwater	B6	Agriculture and livestock	M29(1)
			M29(2)
	B16	Agriculture. Livestock: pens sheep, horses.	M40
	B24	Agriculture and livestock	M49
	B31	Agriculture. Livestock: pigs and sheep pens	M182
	B43	Agriculture. Livestock: pigs and goat pens	M170
	B61	Agriculture. Livestock: cattle	M180
	B62a	Pens cattle and pigs. Feedlots	M184
Surface water	A1	Agricultural Extensive livestock	M23(1)
			M23(2)
			M23(3)
	A2		M27(1)
			M27(2)
	A3		M37(1)
			M37(2)
	A4		M51(1)
			M51(2)

from 250 to 2,500 bp. Products of 750, 400, 300 and 250-bp have commonly occurred. The 300-bp band was the only band common to all the isolates. The greatest diversity of DNA fragments were observed within the range of 1,000–2,500-bp.

Cluster analysis of *E. coli* BOX-PCR fingerprints

To determine the relatedness of strains, a dendrogram based on BOX-derived fingerprint data was constructed (Figure 3). The cluster highlights that the BOX-PCR results show a high
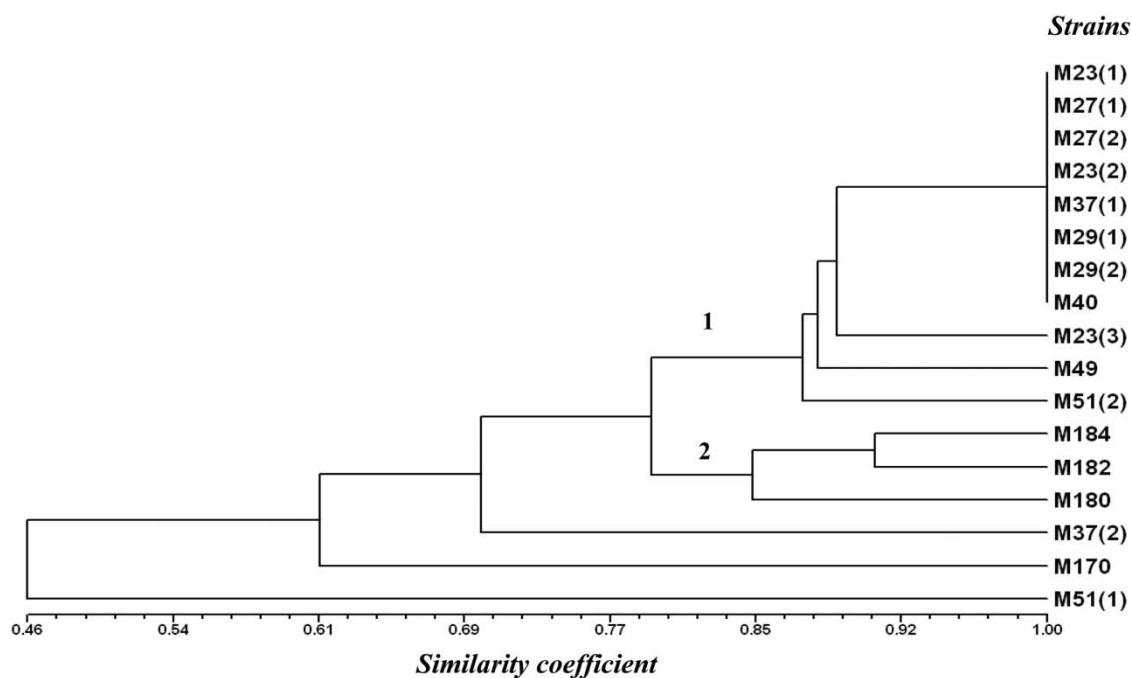
Figure 3 | Dendrogram based on UPGMA cluster analysis of BOX-PCR data from strains isolated from water samples from the Barranquita-Knutzen basin.

homogeneity within the group. From all the isolates, three represented individual genotypes M37(2), M170 and M51(1), and the other 14 were grouped in two different cluster. As shown in Figure 3, each group has a characteristic number of strains as well as separate inter- and intra-group similarity relations. Cluster 1 is formed by eleven isolates (64.7%) related to four different BOX-PCR patterns. Eight isolates had the same genotype. This cluster was made up of strains isolated from surface water and groundwater from the Knutzen basin. Cluster 2 was composed of three isolates, where two (M184 and M182) were grouped in a subcluster with high similarity (90%).

DISCUSSION

Bacteriological groundwater quality measurements were compared with the AFC water quality standards. The results indicated that 50% of the samples were unfit for human consumption. In addition, most of the sampling sites were affected by fecal contamination, because of the high levels of FC and the presence of *E. coli*. These values were related with land uses. Thus, it was found that most samples taken from wells located in the proximity of livestock had higher counts of total and FC and presence of *E. coli*. The hydrogeological setting plays an important role in groundwater contamination, specially the water table depth and USZ lithology, parameters that control the aquifer vulnerability to contamination. Thus, the USZ exerts a protection on the aquifer through physical (advection, dispersion and dilution), chemical (adsorption, oxidation, hydrolysis, redox reaction, dissolution) and biological (degradation or reduction) processes which produce the disappearance or attenuation of contaminants, influencing their arrival to the aquifer (Díaz Delgado *et al.* 2005). However, as was observed in this study, many contaminants persist through infiltration and reach the aquifer, in spite of the fact that the found USZ lithology was represented by fine sands and silts. In addition, the most contaminated wells showed shallow depths of the water table (2.58 to 13.2 m). This aspect is also greatly influenced by the contaminant load (duration, concentration, and so forth). In this case, the bacteriological analysis carried out in samples taken near livestock punctual sources, showed high concentration of total and FC and *E. coli*.

In the present study, the analysis of BOX-PCR fingerprinting revealed a low level of genetic diversity among *E. coli* water isolates, with two clusters, highlighting that cluster 1 has 100% of genetic similarity, gathering more than 60% of total strains. The whole dendrogram clearly indicates that this hydrological basin contains dominant strains with multiple isolates exhibiting identical (100%) or highly similar (greater than 85%) DNA fingerprints.

These results may be related to the arrival to surface water and groundwater of *E. coli* coming from the same contamination source. The contamination from a single fecal source would likely result in less diversity between isolates. Thus, different authors indicate that in water environments impacted by only one contamination source, the populations of *E. coli* were significantly less diverse than impacted by different sources (urban, industrial, sewage and agricultural activity) (Anjos Borges *et al.* 2003; McLellan 2004; Orsi *et al.* 2007; Lyautey *et al.* 2010). Thus, in Barranquita-Knutzen basin, livestock activity would probably be the main source of fecal contamination in surface and groundwater environments.

The multivariate statistical analysis of DNA fingerprinting allowed differentiating two clusters. Cluster 1 corresponds to all strains isolated from Knutzen sub-basin and cluster 2 corresponds to Barranquita sub-basin. This proves that hydrological factors influence the distribution of *E. coli* in the studied environment. Interestingly, these results have implications to manage the fecal contamination and the associated waterborne disease risks, in rural environment.

CONCLUSIONS

In the present study, the bacteriological analysis showed that more than 50% of water samples were unfit for human consumption.

The genetic diversity DNA fingerprinting analysis by BOX-PCR indicated low genotypic diversity of *E. coli* isolates from surface water and groundwater. The low genetic diversity suggested the presence of a dominant source of fecal contamination in the Barranquita-Knutzen basin.

The relationship between low genotypic diversity of *E. coli* and land use in the basin would prove that water contamination comes from livestock.

The genetic diversity of *E. coli* isolated from surface water was less than that identified in groundwater. This is influenced by the water body characteristics, considering that the stream samples are hydraulic linked whereas in the aquifer the samples are less hydraulic connected taking into account the different groundwater flow line directions. Moreover, the different local lithologies and geochemical water types in the aquifer may generate diverse microcosms in bacteria population.

The analysis of genetic diversity by BOX-PCR demonstrates that the hydrological dynamics have a great influence on microbial distribution. Accordingly, the water divide between the two sub-basins controls the two big different strains groups, showing that hydrological features can exert selective pressure on bacteria DNA.

ACKNOWLEDGEMENTS

REFERENCES

Anjos Borges, L. G., Vechia, V. & Corcao, G. 2003 Characterisation and genetic diversity via REP-PCR of *Escherichia coli* isolates from polluted waters in southern Brazil. *Microbiology Ecology* 45, 173–180.

APHA/AWWA/WEF 2005 *Standard Methods for the Examination of Water and Wastewater*, 21th edn. American Public Health Association/American Water Works Association/Water Environment Federation. Part. 9000. Microbiological examination. Washington, DC, USA.

Argentine Food Code (AFC) 2012 Law 18284/69. Chapter XII. Water drinks, water and sparkling water. Art 982 (Res MSyAS n° 494 del 7.07.94). Argentina.

Baldy-Chudzik, K. & Stosik, M. 2005 Specific genomic fingerprints of *Escherichia coli* strains with repetitive sequences and PCR as an effective tool for monitoring freshwater environments. *Polish Journal of Environmental Studies* 14 (5), 551–557.

Bonnet, C., Diarrassouba, F., Brousseau, R., Masson, L., Topp, E. & Diarra, M. S. 2009 Pathotype and antibiotic resistance gene distributions of *Escherichia coli* isolates from broiler chickens raised on antimicrobial-supplemented diets. *Applied and Environmental Microbiology* 75 (22), 6955–6962.

Carlos, C., Alexandrino, F., Stoppe, N., Sato, M. & Ottoboni, L. 2012 Use of *Escherichia coli* BOX-PCR fingerprints to identify sources of fecal contamination of water bodies in the State of São Paulo, Brazil. *Journal of Environmental Management* 93, 38–43.

Chin, D. 2006 *Water-Quality Engineering in Natural Systems.* John Wiley & Sons, Inc., Hoboken, NJ.

Cuevas López, O., León Félix, J., Edeza, M. & Quiroz, C. 2009 Detection and antibiotic resistance of *Escherichia coli* and *Salmonella* in water and agricultural soil. *Fitotec. Mex* 32 (2), 119–126.

Díaz Delgado, C., Alberich, M. & Lopez Vera, F. 2005 Hidrologia Subterránea. En: Recursos Hídricos. Conceptos básicos y estudios de casos en Iberoamérica. Piriguazú. Ediciones/ CIRA-UAEM. Pag: III-3 a III-101.

Dombek, P., Johnson, L., Zimmerley, S. & Sadowsky, M. 2000 Use of repetitive DNA sequences and the PCR to differentiate *Escherichia coli* isolates from human and animal sources. *Applied and Environmental Microbiology* 66 (6), 2572–2577.

Gambero, M. L., Blarasin, M. & Bettera, S. 2016 Tracking contamination: antibiotic resistance and molecular characteristics of *Escherichia coli* isolated from groundwater in an agroecosystem. Argentina Rend. Online Soc.Geol. It., Suppl. n. 1 al 39, 829. Società Geologica Italiana, Roma. DOI: 10.3301/ROL.2016.63.

Giuliano Albo, M. J. & Blarasin, y. M. 2014 Hidrogeoquímica y estimación del fondo natural de nitratos del agua subterránea en un agroecosistema del pedemonte de la sierra de Comechingones. *Revista de la Asociación Geológica Argentina* 71 (3), 378–392. http://ppct.caicyt.gov.ar/index. php/raga/article/view/3096/4349.

Harwood, V., Whitlock, J. & Withington, V. 2000 Classification of antibiotic resistance patterns of indicator bacteria by discriminant analysis: use in predicting the source of fecal contamination in subtropical waters. *Appl. Environ. Microbiol.* 66, 3698–3704.

Ibekwe, A., Murinda, S. & Graves, A. 2011 Genetic diversity and antimicrobial resistance of *Escherichia coli* from human and animal sources uncovers multiple resistances from human sources. *PLoS ONE* 6 (6), e20819. doi:10.1371/journal.pone. 0020819.

Lyautey, E., Lu, Z., Lapen, D. R., Wilkes, G., Scott, A., Berkers, T., Edge, T. A. & Topp, E. 2010 Distribution and diversity of *Escherichia coli* populations in the south nation river drainage basin, Eastern Ontario, Canada. *Applied and Environmental Microbiology 7* 6 (5), 1486–1496.

Ma, H., Fu, L. & Li, J. 2011 Differentiation of fecal *Escherichia coli* from human, livestock, and poultry sources by rep-PCR DNA fingerprinting on the shellfish culture area of East China Sea. *Curr. Microbiol.* 62, 1423–1430.

Malathum, K., Weinstock, G. & Murray, B. 1998 Repetitive sequence-based PCR versus pulsed-field gel electrophoresis for typing of *Enterococcus faecalis* at the subspecies level. *Journal of Clinical Microbiology* 36 (1), 211–215.

McLellan, S. L. 2004 Genetic diversity of *Escherichia coli* isolated from urban rivers and beach water. *Applied and Environmental Microbiology* 70 (8), 4658–4665.

Mohapatra, B., Broersma, K., Nordin, R. & Mazumder, A. 2007 Evaluation of repetitive extragenic palindromic – PCR of discrimination of fecal *Escherichia coli* from humans, and different domestic and wild animals. *Microbiology Inmunology* **51** (8), 733–740.

Orsi, R., Stoppe, N., Sato, M., Gomes, T., Prado, P., Manfio, G. & Ottoboni, L. 2007 Genetic variability and pathogenicity potential of *Escherichia coli* isolated from recreational water reservoirs. *Research in Microbiology* **158**, 420–427.

Rademaker, J. L. W. & de Bruijn, F. J. 1997 Characterization and classification of microbes by rep-PCR genomic fingerprinting and computer assisted pattern analysis. *DNA Markers: Protocols, Applications and Overviews* (G. Caetano-Anolles & P. M. Gresshoff, eds). John Wiley & Sons, New York, pp. 151–171.

Sinclair, A., Hebb, D., Jamieson, R., Gordon, R., Benedict, K., Fuller, K. & Stratton, G. 2009 Growing season surface water loading of fecal indicator organisms within a rural watershed. *Madanic Water Research* **43**, 1199–1206.

Somarelli, J., Makarewicz, J., Sia, R. & Simon, R. 2007 Wildlife identified as major source of *Escherichia coli* in agriculturally dominated watersheds by BOX A1R-derived genetic fingerprints. *Journal of Environmental Management* **82**, 60–65.

Unno, T., Dukki, H., Jang, J., Lee, S., Kim, J., Ko, G., Kim, B., Ahn, J., Sadowsky, M. & Hur, H. 2010 High diversity and abundance of antibiotic-resistant *Escherichia coli* isolated from humans and farm animal hosts in Jeonnam Province, South Korea. *Science of the Total Environment* **408**, 3499–3506.

Versalovic, J., Schneider, M., Bruijn, F. & Lupski, J. 1994 Genomic fingerprinting of bacteria using repetitive sequence based polymerase chain reaction. *Meth. Mol. Cell. Biol.* **40**, 5–25.