# Prevalence of season-specific *Escherichia coli* strains in the Yeongsan River Basin of South Korea

Jeonghwan Jang,[1] Tatsuya Unno,[1] Seung Won Lee,[1] Kyung Hwa Cho,[3] Michael J. Sadowsky,[4,5] GwangPyo Ko,[6] Joon Ha Kim[1]** and Hor-Gil Hur[1,2]*

[1]*School of Environmental Science and Engineering and*
[2]*International Environmental Research Center, Gwangju Institute of Science and Technology, Gwangju, Korea.*
[3]*Department of Civil and Environmental Engineering, University of Massachusetts, Amherst, Massachusetts, USA.*
[4]*Department of Soil, Water and Climate and*
[5]*BioTechnology Institute, University of Minnesota, St. Paul, Minnesota, USA.*
[6]*Department of Environmental Public Health, Seoul National University, Seoul, Korea.*

## Summary

Seasonal and spatial variation in the genotypic richness of 3480 *Escherichia coli* isolates obtained from the Yeongsan River basin in South Korea was investigated by using the horizontal fluorophore-enhanced rep-PCR (HFERP) DNA fingerprinting technique. The relationship between 60 *E. coli* isolates from each of 58 freshwater samples was determined by using multidimensional scaling (MDS) analysis and self-organized maps (SOMs). The MDS analysis, done based on HFERP DNA fingerprints, showed that *E. coli* isolates obtained in October through December clustered tightly, while those obtained in other sampling periods were more genetically diverse. However, site-specific *E. coli* genotypes were not observed. SOMs analysis, done using the 10 most frequently isolated *E. coli* genotypes, showed the occurrence of season-specific *E. coli* genotypes and the main SOMs clusters were most influenced by temperature, strain diversity and biochemical oxygen demand. Diversity among *E. coli* genotypes tended to decrease as water temperature decreased, and the numbers of *E. coli* genotypes observed in urban area were greater, more diverse and less dependent on water temperature than those obtained from agricultural areas. Taken together, our findings indicate that an ecological approach needs to be considered in order to obtain a better understanding of *E. coli* community dynamics in the environment and that SOMs analysis is useful to visualize the multidimensional dependent variables that are influencing the types and dynamics of specific *E. coli* genotypes in the environment.

## Introduction

While *Escherichia coli* is generally thought to be a commensal bacterium that inhabits the intestinal tract of warm-blooded animals, including humans, this bacterium often makes it way into soils and water resulting in faecal contamination of the environment. Recently, *E. coli* was adopted as one of the key indicator organisms to examine faecal contamination in freshwater environments (USEPA, 2006). However, the use of *E. coli* as a faecal indicator microorganism has recently been questioned since the bacterium can survive for long period of time outside of the intestinal tract and may reproduce in several secondary habitats (Byappanahalli and Fujioka, 1998; Byappanahalli *et al.*, 2006; Ishii *et al.*, 2006; Sadowsky and Whitman, 2010). Naturalized *E. coli* genotypes have been reported in sand, soil and sediments as well as in association with macrophytic algae (Byappanahalli *et al.*, 2006; 2007; Ishii *et al.*, 2006; Sadowsky and Whitman, 2010). Several studies have suggested that there is a relationship between the genotypes of *E. coli* found among specific animal hosts and the geographical locations where they were isolated (Byappanahalli *et al.*, 2006; Ishii *et al.*, 2006). Dombek and colleagues (2000) reported that surrounding environmental conditions may be one of the factors influencing the genotypic richness of *E. coli* in the extraintestinal environment. Several studies have now provided convincing evidence that specific *E. coli* genotypes may become naturalized to some environments (Byappanahalli *et al.*, 2006; Ishii *et al.*, 2006; Walk *et al.*, 2007).

Diversity among *E. coli* isolates can be determined by using several genotypic methods, including the horizontal fluorophore enhanced rep-PCR (HFERP) DNA fingerprinting technique (Johnson *et al.*, 2004; Ishii and Sadowsky, 2009). Complex DNA fingerprint data from a large number of isolates have been analysed by using Pearson's, curve-based, product-moment correlation coefficient,