Curr Microbiol (2011) 63:439–449
DOI 10.1007/s00284-011-0002-6

Escherichia coli Diversity in Livestock Manures and Agriculturally Impacted Stream Waters

Kimberly L. Cook · Carl H. Bolster · Kati A. Ayers · Dale N. Reynolds

Received: 9 May 2011 / Accepted: 7 August 2011 / Published online: 28 August 2011
© Springer Science+Business Media, LLC (outside the USA) 2011

Abstract *Escherichia coli* (*E. coli*) isolate diversity enhances the likelihood of survival, spread, and/or transmission of the organism among environments. Understanding the ecology of this important organism is requisite for development of more accurate protocols for monitoring and regulatory purposes. In this study, *E. coli* diversity, gene profiles and transport properties of isolates from different livestock and water sources were evaluated. Strain diversity was evaluated by BOX-PCR, phylotyping, and profiling for 15 genes associated with adhesion, toxin production, iron acquisition or capsular synthesis. Attachment efficiencies were calculated for 17 isolates following transport through saturated porous media. Richness of genotype profiles for livestock isolates was relatively low (25, 12, and 11 for swine, poultry and dairy, respectively) compared to those from stream-water (115 and 126 from dry or wet weather events, respectively). Attachment efficiencies varied by an order of magnitude (0.039–0.44) and the isolate with the highest attachment efficiency possessed the largest suite of targeted genes including those for adherence (*iha*, *agn43*, and *fimH*), surface exclusion (*traT*) and the siderophore $iroN_{E.coli}$. Variation in *E. coli* isolates based on temporal and ecological source was found to translate to equally broad ranges in transport efficiency underscoring the large degree of genotypic and phenotypic variation that exists among *E. coli* isolates. The impact of this diversity on genetic exchange and the concomitant effect on the organisms' fate and transport under in situ environmental conditions warrant further investigation. These factors also require careful consideration for purposes of modeling, source tracking, and risk assessment.

K. L. Cook (✉) · C. H. Bolster
USDA Agricultural Research Service, AWMRU,
230 Bennett Lane, Bowling Green, KY 42104, USA
e-mail: kim.cook@ars.usda.gov

K. A. Ayers
Western Kentucky University, 1906 College Heights Blvd,
Bowling Green, KY 42101, USA

D. N. Reynolds
Kentucky Division of Water, 1508 Westen Street,
Bowling Green, KY 42104, USA

Introduction

Escherichia coli (*E. coli*) is a dominant intestinal commensal organism, an important fecal indicator bacterium (FIB), a potential pathogen and it has been used for microbial source tracking (MST) purposes. Recent studies have highlighted the extremely high diversity of *E. coli* isolates in water, human, and livestock sources [8, 16, 18, 21, 28]. These strain-level differences in the genotype and phenotype of *E. coli* isolates likely influence not only pathogenicity but also fitness in secondary habitats [32].

The versatility of *E. coli* is due to its acquisition of traits (fitness genes, plasmids, etc.) which aid in its survival and maintenance in diverse environments [27]. Type I fimbriae, the adhesion antigen 43 (Ag43) and the outer membrane protein Iha are each associated with adhesion, cell aggregation and biofilm formation, properties that may influence the fate and transport of *E. coli* isolates in transit from one host to another [27, 29, 30, 32]. Siderophores ($iroN_{E.coli}$) and genes involved in heme transport (*chuA*) are often associated with virulence, but may also confer a selective advantage in iron-limited secondary habitats. Few studies have evaluated how genes associated with adhesion, toxin production, iron acquisition or capsular synthesis correlate with transport of *E. coli* isolates [11, 24, 32].

Both environmental and cellular characteristics determine the fate and transport of microbial cells following deposition in soil or waters. Accordingly, studies have found that the distribution of *E. coli* genotypes is influenced by stream order, hydrologic conditions, spatial and temporal distribution, and land use patterns [1, 18, 24, 28]. Genotypic diversity (cell surface characteristics and genetic composition) of *E. coli* isolates has also been shown to affect transport through saturated porous media and biofilm formation even when strains are taken from the same source [2, 3, 10, 32].

The importance of strain diversity as it relates to FIB and MST lies in the effect it has on expressed traits (particularly adhesion) of the organism. Isolates that will be transported to water and sediments are the subset of the population that will be used for monitoring purposes. The goals of this study were to (1) determine the diversity of *E. coli* in manures from diverse livestock species and stream-water samples taken following dry and wet weather events; (2) determine the profile of adhesion, toxin production, iron acquisition, or capsular synthesis genes in representative isolates; and (3) evaluate the effect of strain level differences on the attachment and transport of *E. coli* through water-saturated porous media. Knowledge about these processes should aid in predicting its behavior in the environment, improve the design of transport models and ultimately strengthen monitoring and oversight processes.

Materials and Methods

Escherichia coli Isolation from Source Samples

Five 1-L samples of liquid swine or dairy effluent were collected from a farm in western KY. The samples were collected at five randomly chosen locations within the lagoon at a depth of ~1 meter using 1-L sterilized Nalgene bottles attached to an adjustable sampling pole. De-caked poultry litter was collected from five random locations inside a commercial broiler house in KY prior to the fifth flock. Samples were returned to the laboratory and processed immediately. Slurry and litter samples (10 g or 10 ml) were serially diluted in phosphate buffer, plated onto Modified m-Tech agar (BD-Diagnostic Systems, Franklin Lakes, NJ) and incubated at 37°C for 24 h. Stream-water samples ($n = 17$) were obtained from Bacon Creek located in western KY. Bacon Creek is on the EPA 303(d) list of impaired streams for pathogen presence. The predominant land use within the 90.5 square mile watershed is agricultural, but there are also surrounding rural communities with straight-pipes or septic systems. Samples were collected from the same locations (except for one site which was unavailable following the wet weather event) following one dry weather event ($n = 9$; 72 h without any rain) and one wet weather event ($n = 8$; 72 h of no rain followed by enough rainfall to cause runoff that reaches the stream). Stream-water samples (0.5, 5.0, and 10.0 ml) were filtered in triplicate onto 0.45 µm nitrocellulose membranes (Millipore, Billerica, MA) which were placed onto mFC agar for fecal coliforms (BD-Diagnostic Systems, Franklin Lakes, NJ) and incubated at 37°C for 24 h. Following incubation, each presumptive *E. coli* colony was plated onto EMB (BD-Diagnostic Systems, Franklin Lakes, NJ) and mFC (BD-Diagnostic Systems, Franklin Lakes, NJ) agar. Presumptive *E. coli* ($N = 1346$) were inoculated into 96 well microtiter plates (Fisher Scientific, St. Louis, MO), grown overnight at 37°C in 150 µl of Luria–Bertani (LB) liquid and stored at −80°C with 15% glycerol until use. Isolates were confirmed as *E. coli* by PCR analysis of the *uidA* gene by a modification of the method of Frahm and Obst [12]. The reaction targeted an 82 bp fragment of the of the *E. coli uidA* gene sequence using primers 784F 5′-GTG TGA TAT CTA CCC GCT TCG C-3′ and 866R 5′-AGA ACG GTT TGT GGT TAA TCA GGA-3′ [12]. The amplification mixture contained 1× EconoTaq PLUS master mix (Lucigen, Middleton, WI), 3.5 mM $MgCl_2$, and 600 nM each primer. The cycle conditions were as follows: 15 min at 95°C, 39 cycles at 95°C for 15 s and 60°C for 60 s, and 72°C for 45 s. PCR mixtures (15 µl) were electrophoresed on 2% agarose gel and ethidium bromide stained gel images were visualized with a FOTO/Analyst Investigator/Eclipse system (Fotodyne Inc., Hartland, WI).

BOX-PCR and Computer-Assisted DNA Fingerprint Analysis

Stored *E. coli* culture (10 µl) was transferred to 140 µl fresh LB broth and incubated overnight as described above. Isolates were pelleted by centrifugation (Eppendorf 5810R, Hauppauge, NY) at 3220×*g* for 45 min and 100 µl of the supernatant was discarded. Pellets were washed with 150 µl 0.1× PBS, pelleted as before and diluted by discarding 150 µl of the supernatant and adding 100 µl sterile water. Rep-PCR using 0.5 µM of the BOX AIR primer (5′-CTACGGCAAGGCGACGCTGACG-3′) (BOX-PCR) was carried out in a 25 µl reaction containing 1× EconoTaq PLUS master mix (Lucigen, Middleton, WI) and 2 µl of diluted cell suspension [7]. The PCR was conducted on a PTC-200 model thermal cycler (MJ Research/BioRad, Hercules, CA) as previously described [32] except the initial denaturation step was extended to 12 min. PCR mixtures (10 µl) were electrophoresed on 1% agarose gel supplemented with 40% synergel (Diversified Biotech, Boston, MA) for 4 h at 75 V. Ethidium bromide stained gel images were captured in quadrants with a FOTO/Analyst Investigator/Eclipse system (Fotodyne Inc., Hartland, WI). Gel images were captured, bands identified and the data were

statistically analyzed using Fingerprinting II Software Version 3.0 (BioRad, Hercules, CA). DNA fingerprints were compared using the Jaccard method for calculating similarity coefficients with optimization and position tolerance settings of 1.81 and 1.06, respectively. Cluster analyses of similarity matrices were performed by an unweighted pair group method with arithmetic mean (UPGMA). Isolates were grouped into unique BOX-PCR profiles based on 80% similarity in BOX-PCR fingerprint pattern. Rarefaction analysis and clone library coverage of *E. coli* isolates were computed using RarefactWin (version 1.3; S.M. Holland, Univ. GA, Athens, GA, http://www.uga.edu/strata/software/). Coverage was calculated by Good's method using the formula $[1 - n/N] \times 100$, where n is the number of fingerprint profiles in a sample represented by one isolate and N is the total number of isolates in that sample. Shannon-Wiener diversity index measures the rarity and commonness of species in a community. The Shannon-Wiener diversity index was determined using an on-line calculator (http://www.changbioscience.com/genetics/shannon.html) which uses the formula $H' = -\sum p_i \ln(pi)$ to calculate the diversity index; were p_i is the proportion of total number of species made up of the ith species. Source sub-group isolates were selected from the major BOX-PCR fingerprint groups to represent at least 80% of strains from each source (stream-water samples were not represented at this level due to the high degree of diversity).

Isolate Typing and Genetic Profiling

Escherichia coli isolates from the source sub-groups ($n = 221$) were used for more complete genotypic and phenotypic analyses as described below. Isolates were placed in one of four phylogenetic groups (A, B1, B2, or D) using the triplex PCR method of Clermont et al. [6] as modified by Higgins et al. [14]. Using this method, two genes (*chuA* and *yjaA*) and a non-coding region of the genome (TSPE4.C2) are amplified in one PCR reaction. Multiplex PCR was conducted using the PCR method of Chapman et al. [5] for pool I–VI to target genes involved in adhesion, toxin production, iron acquisition, or capsular synthesis (Table 1). PCR analyses for the presence of *agn43* were carried out as previously described [32] except the initial denaturation step was extended to 15 min. STEC were assessed using the multiplex PCR assay developed by Patton [23] for the simultaneous detection of stx_1, stx_2, and *eaeA* (Table 1). Assays were carried out in Qiagen HotStart Taq Master Mix (Qiagen, Valencia, CA) in a total volume of 25 µl (except the STEC multiplex which was in 50 µl) with 2 µl of genomic DNA extract and 0.5 µM primer in a PTC-200 model thermal cycler (MJ Research/BioRad, Hercules, CA). PCR mixtures were electrophoresed on 2% agarose gel 1 h at 75 V. Ethidium bromide stained gel images were captured with a FOTO/Analyst Investigator/Eclipse system (Fotodyne Inc., Hartland, WI).

Transport Experiments

Transport experiments were conducted using 17 of the *E. coli* isolates using methods previously described Bolster et al. [3]. In brief, a fresh culture of the isolate was grown overnight in LB broth for 3.5 h, centrifuged at $3{,}700 \times g$ and washed in 1 mM KCl. The bacterial solution was diluted to an optical density 546 nm of 0.19–0.21 (BioSpec-mini, Shimadzu, Kyoto, Japan) to obtain a cell concentration of $\sim 1.1 \times 10^8$ cells. Transport experiments were conducted through 10-cm long water-saturated columns packed with clean quartz sand (Unimin™, New Caanan, CT) ranging in size from 710 to 850 µm in diameter. Columns were operated at a flow rate of 0.67 ml min^{-1} for a Darcian velocity of ~ 0.36 cm min^{-1}. Bacterial solutions were injected into each column for 38 min (~ 1.2 pore volumes) followed by bacteria-free electrolyte solution. Effluent was collected every 3 min using a Spectra/Chrom CF-1 fraction collector (Spectrum Chromatography, Houston, TX) fraction collector. Effluent concentrations of *E. coli* were determined by measuring the optical density of the samples at a wavelength of 546 nm.

The breakthrough curves (BTCs) for each column experiment were fit with a first-order deposition model using weighted least squares regression to obtain a deposition coefficient (see Bolster et al. [3] for details). Bacterial attachment efficiencies (α), which describe a cell's ability to become associated with the sand surface, was calculated from the fitted values of the bacterial deposition rate, k, using the hemispheres-in-cell model derived by Ma et al. [19]:

$$\alpha = \frac{2}{3}\frac{d_c k}{(1-\varepsilon)v\eta}\left[\frac{3-\varepsilon}{3-3\varepsilon} - \frac{2(3-\varepsilon)}{\pi(3-3\varepsilon)}\cos^{-1}\left(\frac{3-3\varepsilon}{3-\varepsilon}\right)^{1/2} + \frac{2}{\pi}\sqrt{2\left(\frac{3-\varepsilon}{3-3\varepsilon}\right)^{-1/2} - 1}\right]^{-1}$$

where d_c is the sand grain diameter (cm), ε is porosity, and η is the single-collector efficiency calculated using the correlation equation of Ma et al. [19] using a bacterial density of 1055 kg m^{-3} and a Hamaker constant of 6.5×10^{-21} J [25].

Results

Escherichia coli Isolate Rep-PCR Profiling, Coverage, and Diversity Analysis

BOX-PCR fingerprint profiles for *E. coli* isolates from each source were clustered separately using UPGMA

Table 1 Oligonucleotide list of genes targeted in *E. coli* isolates

Gene target	Gene label	DNA sequence (5′ to 3′)	Size (bp)	Description/function	Reference
Single[a]					
agn43	Agn43F	GACTATGACCGGATTSTGGCAGGCT	498	Adhesion involved in diffuse adherence, cell aggregation	Yang et al. [32]
	Agn43R	GTGGCTCCAGCATCAGRTTGTCAG			
Multiplex 1[b]					
afa/draBC	Afa-F	GGCAGAGGGCCGGCAACAGGC	559	Adhesion; Central region of Dr antigen-specific fimbrial and afimbrial adhesion operon	Chapman et al. [5]
	Afa-R	CCCGTAACGCGCCAGCATCTC			
fimH	508Fim-H-F	TGCAGAACGGATAAGCCGTGG	508	D-Mannose-specific adhesin; type 1 fimbriae	Chapman et al. [5]
	508Fim-H-R	GCAGTCACCTGCCCTCCGGTA			
iha	IHA-F	CTGGCGGAGGCTCTGAGATCA	827	Novel nonhemagglutinin adhesin (0157:H7 and CFT073)	Chapman et al. [5]
	IHA-R	TCCTTAAGCTCCCGCGGCTGA			
hlyA	hly-F	AACAAGGATAAGCACTGTTCTGGCT	1177	Alpha-hemolysin	Chapman et al. [5]
	hly-R	ACCATATAAGCGGTCATTCCC			
kpsMT II	kpsMTII-F	GTCA GCGCATTTGCTGATACTGTTG	272	Grouping II capsular polysaccharide synthesis (K1, K5, and K12)	Chapman et al. [5]
	kpsMTII-R	CATCCAGACGATAAGCATGAGC			
$iroN_{E.coli}$	IRONEC-F	AAGTCAAAGCAGGGGTTGCCCG	665	Novel catecholate siderophore	Chapman et al. [5]
	IRONEC-R	GACGCCGACATTAAGACGCAG			
traT	Trat-F	GGTGTGGTGCGATGAGCACAG	290	Surface exclusion serum survival	Chapman et al. [5]
	Trat-R	CACGGTTCAGCCATCCCTGAG			
Multiplex 2[b]					
chuA	ChuA.1	GACGAACCAACGGTCAGGAT	279	Heme transport in enterohemorrhagic *E. coli* 0157:H7	Higgins et al. [14]
	ChuA.2	TGCCGCCAGTACCAAAGACA			
TSPE4C2	TspE4C2.1	GAGTAATGTCGGGGCATTCA	152	Anonymous DNA fragment	
	TspE4C2.2	CGCGCCAACAAAGTATTACG			
yjaA	YjaA.1	TGAAGTGTCAGGAGACGCTG	211	Identified in *E. coli* K12, function currently unknown	
	YjaA.2	ATGGAGAATGCGTTCCTCAAC			
Multiplex 3[b]					
eaeA	eae 1F	GACCCGGCACAAGCATAAGC	384	Adhesion; outer membrane protein required for attachment to epithelial cells	Patton and Patton [23]
	eae 2R	CCACCTGCAGCAACAAGAGG			
stx_1	stx 1F	ATAAATCGCCATTCGTTGACTAC	180	Toxin	
	stx 1R	AGAACGCCCACTGAGATCATC			
stx_2	stx 2F	GGCACTGTCTGAAACTGCTCC	255	Toxin	
	stx 2R	TCGCCAGTTATCTGACATTCTG			

[a] Single reaction refers to PCR run with only the listed primer set as described in the named reference

[b] Multiplex refers to multiplex PCR run with the primer sets listed as described in the named reference

analysis ($n = 1346$). Isolates from individual sources were placed into BOX-PCR fingerprint groups based on an 80% similarity criterion (Table 2). Richness of genotype profiles in livestock samples was relatively low (25, 12, and 11 for swine, poultry, and dairy, respectively) compared to that of *E. coli* isolates from stream-water following dry or wet weather events (115 and 126, respectively; Table 2). The diversity of isolates from the stream-water samples was also extremely high as compared to that of isolates from livestock sources. The majority of *E. coli* isolates from swine and dairy sources clustered into one dominant group (52.8 and 47.2%, respectively). Poultry isolates belonged to three dominant groups representing 30.6, 24.7, and 22.9% of all isolates. No *E. coli* genotype represented more than 7% of isolates from stream-water samples taken after wet or dry events. As a result, coverage of isolates from stream-water was very low (Table 2) and rarefaction analysis of BOX-PCR patterns for *E. coli* isolated from stream-water shows that curves never reached a horizontal asymptote, indicating that more isolates would be required to capture all of the diversity within these samples (representative curves for livestock (poultry) and stream-water samples (wet event) are shown in Fig. 1).

Table 2 Rep-PCR Genotype analysis of source sub-group *E. coli* isolates

Source	Total no. of isolates	Genotype richness	Coverage (%)	Shannon div. index	Source sub-group
Swine	339	25	99.1	1.97	49
Poultry	271	12	99.6	1.78	45
Dairy	339	11	99.4	1.59	42
Water-DE	185	115	59.5	4.54	42
Water-WE	212	126	62.3	4.64	43
Total	1346				221

Source indicates the sample type from which the *E. coli* isolates were obtained

Genotype richness refers to the number of unique BOX-PCR fingerprint types (based on 80% similarity)

Coverage $= 1 - (n/N)$ where n is the number of clones that occur only once in the library and N is the total number of clones

Shannon-Weiner diversity index was calculated using the formula $H = -S\ pi\ \ln\ (pi)$; where pi is the proportion of total number of species made up of the ith species

Source sub-group refers to the total number of isolates from genotype groups that were selected for 557 further analyses

Source Sub-Group Typing

To better characterize *E. coli* isolates from source samples, sub-groups of isolates were chosen ($n = 221$; Table 2). These source sub-group isolates represented 81, 90, and 87% of genotype patterns for swine, poultry and dairy isolates, respectively. However, due to the high level of diversity in stream-water samples, sub-group isolates only represented 37 and 48% of genotype patterns for stream-water isolates from the wet and dry weather events, respectively. All of the 221 *E. coli* isolates from source sub-groups were subjected to another BOX-PCR followed by UPGMA cluster analysis of fingerprint profiles. Genotype richness for source sub-groups were consistent with those from the original source samples; livestock source sub-groups for swine, poultry and dairy were clustered into 22, 14, and 9 genotypes, respectively (Fig. 2). The number of clusters within the sub-groups was similar to those of the original isolates (Table 2) which suggests that sub-group isolates were representative of the original isolates from each livestock source. Genotype profiles for *E. coli* isolates from individual animal sources did not cluster together, but were inter-mixed among isolates from other livestock and from stream-water samples (Fig. 2). *E. coli* sub-group isolates from stream-water clustered into 14 and 21 genotype profiles for dry and wet weather events, respectively. This value was significantly lower than for the original stream-water isolates (Table 2), suggesting that the *E. coli* sub-group isolates from these samples were not representative of the entire *E. coli* population. This was expected given the incredible diversity found in the original stream-water samples. Genotype profiles for *E. coli* isolates from stream-water clustered with swine isolates (11%) as well as dairy and poultry isolates (both 23%). Around 10% of isolates from stream-water clustered into groups with

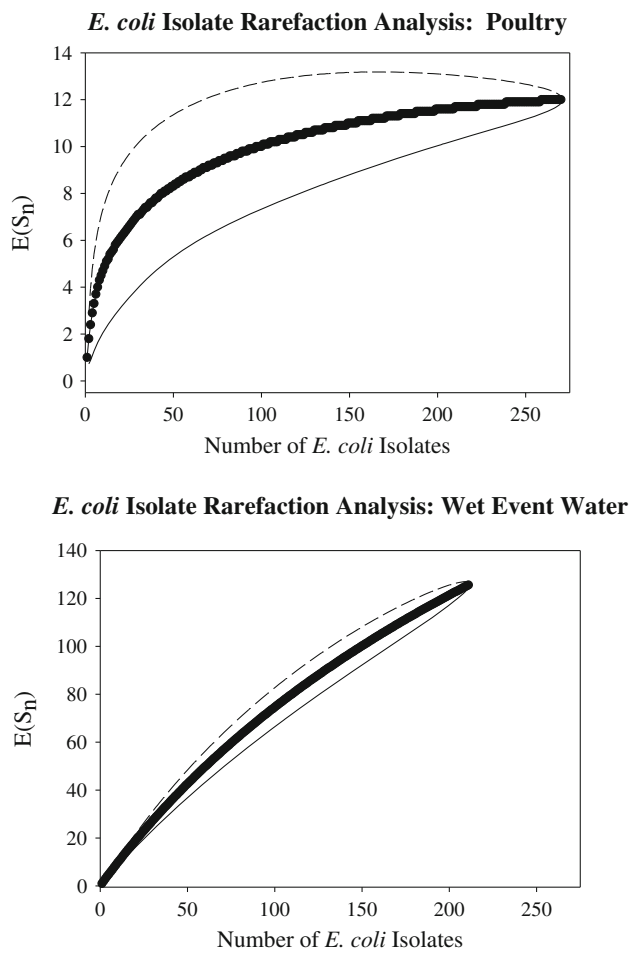


Fig. 1 Rarefaction analysis of *E. coli* isolate diversity in **a** poultry and **b** stream-water wet weather samples was carried out using the software, RarefactWin 1.3. The expected number of isolates [$E(S_n)$] was calculated from the number of isolates occurring in OTU's grouped at the 80% BOX-PCR profile similarity level. *Dashed lines* represent 99% confidence intervals

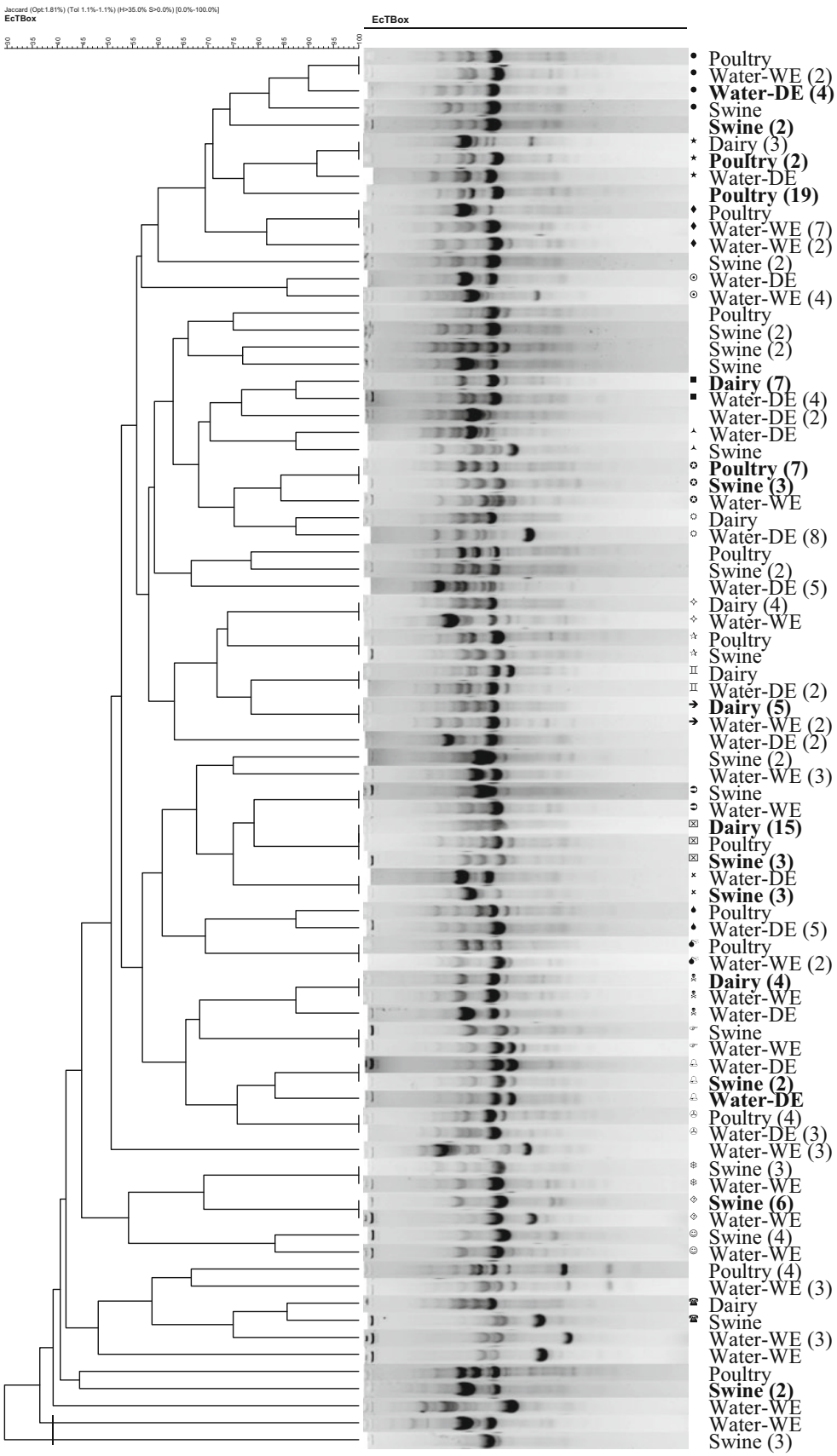


Fig. 2 The dendrogram shows the percent similarity of BOX-PCR fingerprint banding patterns based on UPGMA cluster analysis. Fingerprint patterns from 221 *E. coli* sub-group isolates from dairy, poultry, swine, stream-water after dry event (*DE*), or stream-water after wet event (*WE*) were clustered. *Number* in parentheses shows the number of *E. coli* isolates represented by that source. Groups with the same *symbol* are at least 80% similar. Sources in *bold* contain isolates that were used for column transport studies

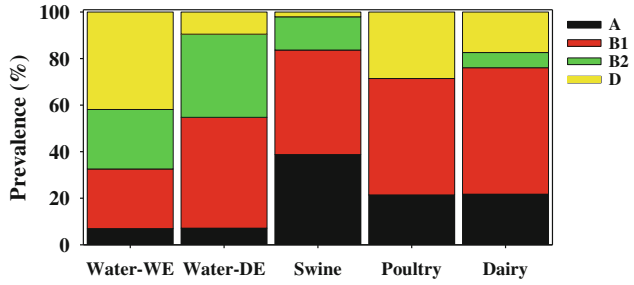


Fig. 3 Stacked bar chart showing the phylogroup distribution of *E. coli* isolates (221) from dairy, poultry, swine, stream-water after dry event (*DE*), or stream-water after wet event (*WE*). Prevalence represents the percent of isolates that are type A, B1, B2, or D

mixed livestock species. Over 34% of *E. coli* isolates from stream-water had genotype profiles that were distinct from those of the tested livestock species (Fig. 2). Furthermore, only 18% of the 84 *E. coli* isolates from the wet and dry event sub-groups clustered together.

Phylogenetic Grouping of *E. coli* Isolates

The triplex PCR method of Clermont et al. [6] was used to assign the 221 *E. coli* isolates to one of four phylogenetic groups (phylogroup; Fig. 3). Phylogroup B1 was the most dominant in livestock isolates (25–50%). *E. coli* isolates from stream-water samples had fewer phylogroup A isolates (7%) compared to livestock (21–39%), but greater percentages of B2 isolates (36 and 26% for dry and wet weather events, respectively) than livestock (14, 7, and 0% for swine, poultry and dairy, respectively). B1 isolates were also dominant in stream-water samples from the dry weather event (48%). However, phylogroup D was most dominant in isolates from the wet weather event (42% of isolates). Except for phylogroup A, the phylogroup profile for *E. coli* isolates from samples taken after the dry event were significantly different from those taken from the same locations after the wet weather event (Fig. 3).

Genetic Profiling of *E. coli* Isolates

Escherichia coli source sub-group isolates were evaluated for the presence of genes associated with adhesion (*afa/draBC*, *iha*, *agn*43, *eaeA*, and *fimH*), toxin production (*hlyA*, stx_1, and stx_2), capsular polysaccharide synthesis (*kpsMTII*) and siderophores ($iroN_{E.coli}$, *chuA*) (Table 3). The occurrence of these genes in *E. coli* isolates varied significantly among livestock species and even between stream-water samples taken after a wet or a dry event. No fimbrial dr-binding adhesion (*afa/draBC*) genes were present in any of the *E. coli* isolates and no single isolate was positive for all three (stx_1, stx_2, and *eaeA*) of the Shiga toxigenic *E. coli* (STEC) virulence genes. However, the three genes occurred separately in up to 8% of isolates and in all sample sources except for poultry (Table 3). The most commonly detected genes were the adhesions *fimH* (present in 80–95% of isolates) and *agn43* (present in 40–100% of isolates). The *traT* gene which encodes the outer membrane surface exclusion protein was present in 75% of all sub-group isolates. Another important adhesion factor *iha* was present in only 5% of isolates and none of the isolates from stream-water were positive. The siderophore $iroN_{E.coli}$ was present in only 24% of isolates. *E. coli* from the dry event associated stream-water samples had the highest percentage of positive *agn43* (100%) and *traT* (93%) isolates, while those associated with the wet event had the lowest percentage that were positive for those genes (40 and 62%, respectively).

Transport Studies

Transport of 17 of the *E. coli* isolates (5 poultry; 5 dairy; 4 stream-water; 3 from swine) through uncoated quartz sand was evaluated (Fig. 4). Bacterial attachment efficiencies among the 17 *E. coli* isolates varied by over an order of magnitude (attachment efficiency ranged from 0.039–0.44). The bacterial attachment efficiencies for these isolates were compared to those for the eight swine isolates evaluated in a previous study [3]; the isolates are shown in

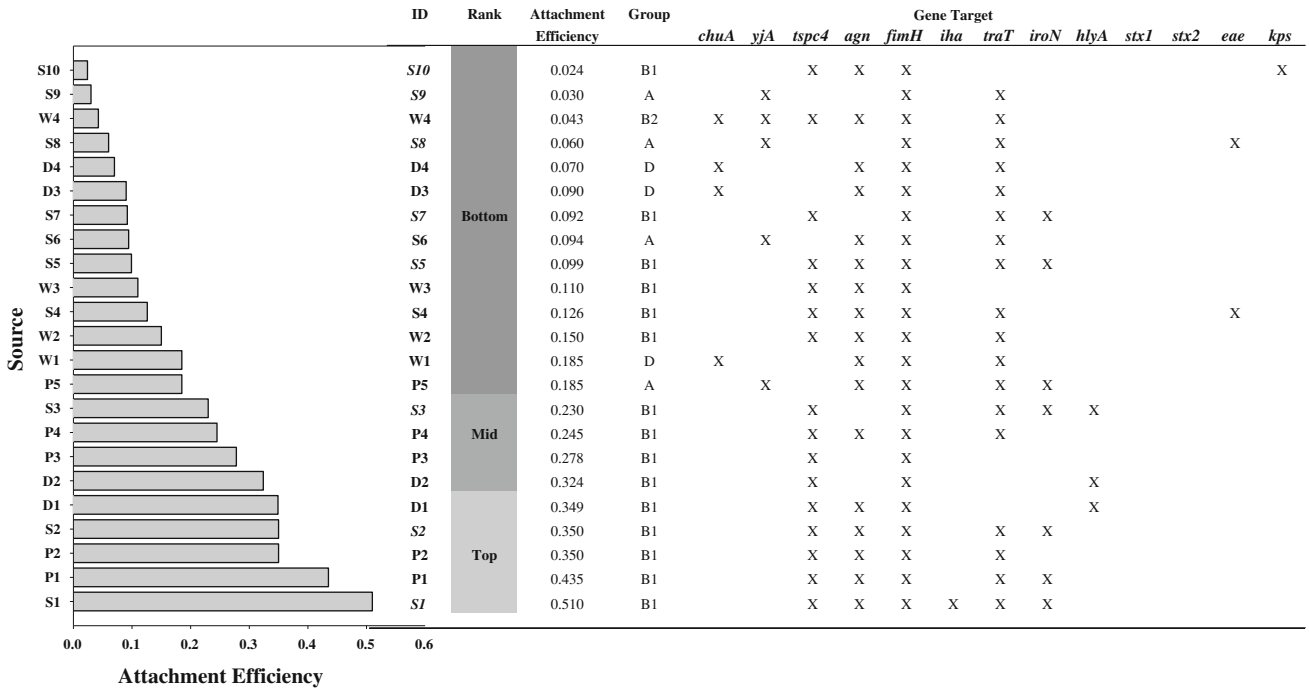
Table 3 Gene prevalence within *E. coli* isolates from each source sub-group

Source	Total	chuA	yjaA	TSPE4. C2	agn	fimH	iha	traT	iroN	hlyA	stxl	stx2	eae	kps	afa
Swine	49	16	39	35	69	80	4	67	29	2	3	1	8	0	0
Chicken	45	24	18	49	87	89	16	80	56	0	0	0	0	0	0
Dairy	42	26	19	50	90	93	5	76	17	19	4	6	5	0	0
Water-dry event	42	50	40	83	100	90	0	93	19	12	1	2	3	0	0
Water-wet event	43	58	14	23	40	95	0	62	0	2	5	5	8	0	0

Values represent the percent of total isolates which were positive for the indicated gene

Source indicates the sample type from which the *E. coli* isolates were obtained

Total indicates the number of *E. coli* isolates tested for each gene target, except for stx_1, stx_2, and *eae* which were tested on all 1,346 *E. coli* isolates



ID	Rank	Attachment Efficiency	Group	Gene Target												
				chuA	*yjA*	*tspc4*	*agn*	*fimH*	*iha*	*traT*	*iroN*	*hlyA*	*stx1*	*stx2*	*eae*	*kps*
S10		0.024	B1			X	X	X								X
S9		0.030	A		X			X		X						
W4		0.043	B2	X	X	X	X	X		X						
S8		0.060	A		X			X		X					X	
D4		0.070	D	X			X	X		X						
D3		0.090	D	X			X	X		X						
S7	Bottom	0.092	B1			X		X		X	X					
S6		0.094	A		X		X	X		X						
S5		0.099	B1			X	X	X		X	X					
W3		0.110	B1			X	X	X								
S4		0.126	B1			X	X	X		X					X	
W2		0.150	B1			X	X	X		X						
W1		0.185	D	X			X	X		X						
P5		0.185	A		X		X	X		X	X					
S3		0.230	B1			X		X		X	X	X				
P4	Mid	0.245	B1			X	X	X		X						
P3		0.278	B1			X		X								
D2		0.324	B1			X		X				X				
D1		0.349	B1			X	X	X				X				
S2		0.350	B1			X	X	X		X	X					
P2	Top	0.350	B1			X	X	X		X						
P1		0.435	B1			X	X	X		X	X					
S1		0.510	B1			X	X	X	X	X	X					

Fig. 4 Graphical representation of attachment efficiencies of isolates used for transport studies as well as attachment efficiency values, phylogroup assignment, and presence/absence of target gene sequences. Swine isolates used in column studies reported by Bolster et al. [2] are *italicized*. Rank indicates the location of the isolate in the top, bottom or mid-level of attachment efficiency values. Group indicates the phylogroup designation of the isolate based on the PCR typing method of Clermont et al. [6]. Gene target positive response is indicated by "*X*". *S* Swine, *D* Dairy, *P* Poultry, *W* Stream-water

italics in Fig. 4. The distribution of phylogroups among the isolates used for transport studies was dominated by group B1 followed by group A, D, and B2 (65, 17, 13, and 4%, respectively). The isolate with the highest attachment efficiency possessed the largest suite of targeted genes including those for adherence (*iha*, *agn43*, and *fimH*), surface exclusion (*traT*) and the siderophore $iroN_{E.coli}$. The five *E. coli* isolates with the highest attachments efficiencies were all positive for *agn43* and *fimH*. A higher portion of these isolates were also positive for $iroN_{E.coli}$ (65% in the top, 25% in the mid and 21% in the lowest group).

Discussion

In this study, the diversity of *E. coli* was evaluated in three livestock source samples and in stream-water samples taken from the same location following a wet and a dry weather event. BOX-PCR followed by rarefaction and coverage analysis of fingerprint profiles from between 271 and 339 isolates showed that the coverage of *E. coli* in livestock sources was over 99% and was well represented by the selected isolates. Alternatively, rarefaction analysis, genotype richness and coverage all suggest that the diversity of *E. coli* within water source samples was much higher and was not represented by the number of isolates obtained. Further analyses showed that the isolates did not cluster according to livestock species and over 34% of isolates from stream-water did not cluster with any livestock genotype. Several other studies have also found differences between strains present in fecal material and those found in soils and water sources [16, 28].

Only 18% of isolates taken from the same locations along a stream following wet and dry events clustered together, suggesting that the temporal diversity in these samples was also great. Temporal variability has also been seen in *E. coli* strains taken at different time points from dairy [28], swine [8] and water [4] sources. Others have found that specific *E. coli* genotypes in soil and water sources were present in multiple sample locations and sample periods [15, 18, 21]. Researchers have suggested that *E. coli* populations may be selected in the environment and these adapted strains become permanent, naturalized members of the community [1, 15]. While 34% of our *E. coli* isolates from stream-water were not represented by the livestock isolates, the large diversity and lack of clonal groups suggest that these isolates may be transient rather than naturalized *E. coli* strains.

While BOX-PCR provides a useful method for separation of large groups of isolates, in this study, that profile did not appear to correlate to other genetic or phenotypic factors. For example, 20 sub-group isolates were selected

to represent a BOX-PCR group that encompassed 53% of the 339 total *E. coli* isolates from swine. Based on BOX-PCR, these isolates were 81% similar. However, the 20 isolates fell into 3 phylogroups (8 type A, 8 type B1, and 4 type B2). Most were positive for *fimH* (90%) and *traT* (75%), however, only 15 were positive for *agn43*, 2 for *iha*, and 6 for $iroN_{E.coli}$. This supports the results of others who found that community fingerprints had no correlation to other characteristics such as phylotype, virulence, or antibiotic resistance [8, 17, 20, 32].

The phylogrouping method of Clermont et al. [6] showed that all four phylogroups (A, B1, B2, and D) were represented in isolates from stream-water, poultry and swine. Dairy lacked B2 isolates which were also found to be a minor group (4%) in the dairy isolates investigated by Son et al. [28]. The *E. coli* isolates taken following the wet weather event were dominated by phylogroup D (42%) while those from the dry event were dominated by phylogroup B1 (only 9.5% D). Significant shifts in phylogroup dominance were also seen in a study on a rural creek watershed conducted by Ratajczak et al. [24]. They suggest that the shift in phylogroups was due to run-off and subsequent input of fecal material from human sources and from dairy. In this study, dairy sources had the highest concentration of group D isolates (28.6%) and may have contributed to the high concentration of phylogroup D isolates following the wet weather event which resulted in run-off and a concomitant increase in total cells (two orders of magnitude; data not shown). However, it was not possible to assess the impact of input from rural septic tanks or straight line pipes on this watershed since no human sources were tested.

Recent studies have highlighted the potential for the environment to serve as a reservoir for commensal and pathogenic *E. coli* including many strains possessing genes associated with virulence or those important to adherence and transport [9, 11, 17, 20]. In this study, the most commonly detected genes were those related to adhesion (*agn43* and *fimH*) and the surface exclusion lipoprotein (*traT*). In a study of 286 bovine feedlot isolates Yang et al. [32] found that all were positive for *fimH*, but only 48% possessed *agn43*. In our study, most of the 221 sub-isolates were positive for *fimH* and similar numbers were positive for *agn43*. Only samples obtained following the wet weather event had low levels of *agn43* positive *E. coli* isolates (39%). Alternatively, *E. coli* isolates in samples obtained following the dry weather event were 100% positive for *agn43*. The two *E. coli* populations also differed significantly in isolates positive for *traT*, $iroN_{E.coli}$, *hlyA*, and *iha*. This suggests that there were significant differences in the source of input between the two events. In a study of the influence of hydrologic conditions on *E. coli* in a rural watershed in France, Ratajczak et al. [24] also found that populations were influenced by dry versus wet events.

There were also differences in the occurrence of genes associated with adhesion, toxin production, iron acquisition or capsular synthesis among livestock isolates. In addition to genotype and phylotype data, the diversity of gene profiles from the livestock isolates, suggests that *E. coli* populations differ significantly among these sources. Our results support those of others who suggest that genes associated with adhesion, toxin production, iron acquisition, or capsular synthesis are endemic to *E. coli* strains in the environment [17, 20] and among livestock species [5, 31]. Pathogenicity may not be correlated with the occurrence of these traits in diverse *E. coli* isolates, but it certainly increases the potential for transfer among strains.

Escherichia coli isolates harboring genes linked to enterohemorrhagic strains (particularly *E. coli* O157:H7) are of particular importance due to their link to serious human infections [23]. In our study, the only samples with no positive stx_1, stx_2, and/or *eaeA* isolates were from chicken. In fact, the percentage of isolates positive for stx_1 and stx_2 were similar to those for *eaeA* (varying between 1 and 8%). Similarly, Maldonado et al. [20] found that ~5% of 1,698 environmental isolates of *E. coli* were positive for *eaeA*, stx_1, stx_2, or *hlyA*. Isolates from swine and the wet event stream-water samples had the highest proportion of *eaeA* positives (8% each). Positive isolates from stream-water obtained following the wet event (4–8%) were consistently higher than those from the dry event (1–3%). Hamilton et al. [13] also found uneven distribution of intimin positive isolates based on both sample location and time of sampling. These results along with those for the other genes and the genotyping data suggest that the temporal distribution of *E. coli* strains is extremely dynamic.

Seventeen isolates were chosen for transport based on isolation source, phylotype and gene profile. Attachment efficiencies, phylogenetic typing, and gene profiles were compared for the selected isolates and for eight swine isolates used for previous transport studies [3]. Similar to findings from other studies, attachment efficiencies among isolates from this study varied by over an order of magnitude [10]. Furthermore, the range in attachment efficiencies observed in this study is within the range observed for eight swine isolates obtained in an earlier study under identical experimental conditions, suggesting that a finite range of attachment efficiencies may exist for *E. coli*, though further experiments are needed to confirm this finding [3].

Chapman [5] suggested that the presence of virulence genes may or may not pre-dispose the strain to pathogenicity however these characteristics can influence the fate and transport of the strain. The environmental fitness and adaptability of isolates carrying particular suites of genes may be influenced by acquired traits. For example, we found that isolate S1 (isolate SP3B03 from [3]) had the

highest attachment efficiency, possessed genes for type 1 fimbrae (*fimH*), the self-recognizing adhesion antigen 43 (*agn43*) and for surface exclusion serum survival (*traT*). It was the only strain to possess *iha* a gene which encodes an OMP that confers adherence to non-adherent laboratory strains [29]. Others have also found that genes associated with adherence are more common in *E. coli* strains that form biofilms [22] and have higher attachment efficiencies [11]. On the other hand, the isolate with the lowest attachment efficiency (S10; isolate SP4-H03 from [3]) was the only one to possess the *kpsMTII* gene which encodes Group II capsular polysaccharides (K1, K5, and K12). Schembri et al. [26] found that aggregation could be inhibited in *E. coli* isolates expressing these capsular polysaccharides. They suggest that capsules physically inhibit Ag43–Ag43 interaction thereby reducing adherence to colon cancer cells. This may explain why isolate S10, which possessed genes encoding the *kpsMTII* gene and *agn43* had the lowest attachment efficiency.

Although studies of genes associated with adhesion, toxin production, iron acquisition, or capsular synthesis in *E. coli* isolates from water sources have been conducted, these have not been correlated with transport characteristics. This is an important factor that warrants further research given the importance of *E. coli* as an indicator organism. It is possible that current monitoring criteria select for the sub-set of the *E. coli* population that is more likely to be transported (i.e., non-adherent). This could lead to biases in data interpretation if, for example, ruminants are more likely to have *E. coli* isolates with fewer genes important to adherence while poultry are dominated by isolates with high levels of adherence. This speaks to the ultimate goal of these studies of genotypic and phenotypic diversity of *E. coli*, which is to address the ecology of this important indicator organism and to identify factors that influence its fate and transport in the environment. The validity of the indicator paradigm and the feasibility of MST can not be fully evaluated without a better understanding of the ecology of the targeted populations.

Acknowledgments The authors wish to thanks Rohan Parekh and Stacy Antle for their technical assistance. This research was part of USDA-ARS National Program 206: Manure and By-product Utilization. Mention of a trademark or product anywhere in this paper is to describe experimental procedures and does not constitute a guarantee or warranty of the product by the USDA and does not imply its approval to the exclusion of other products or vendors that may also be suitable.

References

1. Badgley BD, Ferguson J, Heuvel AV, Kleinheinz GT, McDermott CM, Sandrin TR, Kinzelman J, Junion EA, Byappanahalli MN, Whitman RL, Sadowsky MJ (2011) Multi-scale temporal and spatial variation in genotypic composition of Cladophora-borne *Escherichia coli* populations in Lake Michigan. Water Res 45:721–731
2. Bolster CH, Hazendaroglu BZ, Walk ST (2009) Diversity in cell properties and transport behavior among 12 different environmental *Escherichia coli* isolates. J Environ Qual 38:465–472
3. Bolster CH, Cook KL, Marcus IM, Haznedaroglu BZ, Walker SL (2010) Correlating transport behavior with cell properties for eight porcine *Escherichia coli* isolates. Environ Sci Technol 44:5008–5014
4. Casarez EA, Pillai SD, Di Giovanni GD (2007) Genotype diversity of *Escherichia coli* isolates in natural waters determined by PFGE and ERIC-PCR. Water Res 41:3643–3648
5. Chapman T, Wu X-Y, Barchia I, Bettelheim KA, Driesen S, Trott D, Wilson M, Chin JJ-C (2006) Comparison of virulence gene profiles of *Escherichia coli* strains isolated from healthy and diarrheic swine. Appl Environ Microbiol 72:4782–4795
6. Clermont O, Bonacorsi S, Bingen E (2000) Rapid and simple determination of the *Escherichia coli* phylogenetic group. Appl Environ Microbiol 66:4555–4558
7. Dombek PE, Johnson LK, Zimmerley ST, Sadowski MJ (2000) Use of repetitive DNA sequences and the PCR to differentiate *Escherichia coli* isolates from human and animal sources. Appl Environ Microbiol 66:2572–2577
8. Duriez P, Topp E (2007) Temporal dynamics and impact of manure storage on antibiotic resistance patterns and population structure of *Escherichia coli* isolates from a commercial swine farm. Appl Environ Microbiol 73:5486–5493
9. Ewers C, Antao E-M, Diehl I, Philipp H-C, Wieler LH (2009) Intestine and environment of the chicken as reservoirs for extraintestinal pathogenic *Escherichia coli* strains with zoonotic potential. Appl Environ Microbiol 75:184–192
10. Foppen JW, Liem Y, Schijven J (2008) Effect of humic acid on the attachment of *Escherichia coli* in columns of goethite-coated sand. Water Res 42:211–219
11. Foppen JW, Lutterodt G, Röling WFM, Uhlenbrook S (2010) Towards understanding inter-strain attachment variations of *Escherichia coli* during transport in saturated quartz sand. Water Res 44:1202–1212
12. Frahm E, Obst U (2003) Application of fluorogen probe technique (Taqman PCR) to the detection of *Enterococcus* spp. and *Escherichia coli* in water samples. J Microbiol Met 52:123–131
13. Hamilton MJ, Hadi AZ, Griffith JF, Ishii S, Sadowski MJ (2010) Large scale analysis of virulence genes in *Escherichia coli* strains isolates from Avalon Bay, CA. Water Res 44:5463–5473
14. Higgins J, Hohn C, Hornor S, Frana M, Denver M, Joerger R (2007) Genotyping of *Escherichia coli* from environmental and animal samples. J Microbiol Methods 70:227–235
15. Ishii S, Ksoll WB, Hicks RE, Sadowsky MJ (2006) Presence and growth of naturalized Escherichia coli in temperate soils from Lake Superior watersheds. Appl Environ Microbiol 72:612–621
16. Lu Z, Lapen D, Scott A, Dang A, Topp E (2005) Identifying host sources of fecal pollution: diversity of *Escherichia coli* in confined dairy and swine production systems. Appl Environ Microbiol 71:5992–5998
17. Luna GM, Vignaroli C, Rinaldi C, Pusceddu A, Nicoletti L, Gabellini M, Danovaro R, Biavasco F (2010) Extraintestinal *Escherichia coli* carrying virulence genes in coastal sediments. Appl Environ Microbiol 76:5659–5668
18. Lyautey E, Lu Z, Lapen DR, Wilkes G, Scott A, Berkers T, Edge TA, Topp E (2010) Distribution and diversity of *Escherichia coli* populations in the South Nation River Drainage basin, eastern Ontario Canada. Appl Environ Microbiol 76:1486–1496
19. Ma H, Pedel J, Fife P, Johnson WP (2009) Hemispheres-in-cell geometry to predict colloid deposition in porous media. Environ Sci Technol 43:8573–8579

20. Maldonado Y, Fiser JC, Nakatsu CH, Bhunia AK (2005) Cytotoxicity potential and genotypic characterization of *Escherichia coli* isolates from environmental and food sources. Appl Environ Microbiol 71:1890–1898
21. McLellan SL (2004) Genetic diversity of *Escherichia coli* isolated from urban rivers and beach water. Appl Environ Microbiol 70:4658–4665
22. Naves P, del Prado G, Huelves L, Gracia M, Ruiz V, Blanco J, Dahbi G, Blanco M, del Carmen Ponte M, Soriano F (2008) Correlation between virulence factors and in vitro biofilm formation by *Escherichia coli* strains. Microbial Path 45:86–91
23. Patton A, Patton J (1998) Detection and characterization of shiga toxigenic *Escherichia coli* by using multiplex PCR assays fro *stx1*, *stx2*, *eaeA*, Enterohemorrhagic *E. coli hlyA*, *rfb*0111, and *rfb*0157. J Clinical Microbiol 36:598–602
24. Ratajczak M, Laroche E, Berthe T, Clermont O, Pawlak B, Denamur E (2010) Influence of hydrological conditions on the *Escherichia coli* population structure in the water of a creek on a rural watershed. BMC Microbiol 10:1–10
25. Redman JA, Walker SL, Elimelech M (2004) Bacterial adhesion and transport in porous media: role of the secondary energy minimum. Environ Sci Technol 38:1777–1785
26. Schembri MA, Dalsgaard D, Klemm P (2004) Capsule shields the function of short bacterial adhesins. J Bacteriol 186:1249–1257
27. Sherlock O, Dobrindt U, Jensen JB, Vejborg RM, Klemm P (2006) Glycosylation of the self-recognizing *Escherichia coli* Ag43 autotransporter protein. Appl Environ Microbiol 188:1798–1807
28. Son I, Van Kessel JS, Karns JS (2009) Genotypic diversity of *Escherichia coli* in a dairy farm. Foodborne Pathog Dis 6:837–847
29. Tarr PI, Bilge SS, Vary JCJ, Jelacic S, Habeeb RL, Ward TR, Baylor MR, Besser TE (2000) Iha: a novel *Escherichia coli* O157:H7 adherence-conferring molecule encoded on a recently acquired chromosomal island of conserved structure. Infect Immun 68:1400–1407
30. Van Houdt R, Michiels CW (2005) Role of bacterial cell surface structures in *Escherichia coli* biofilm formation. Res Microbiol 156:626–633
31. Wu X-Y, Chapman T, Trott D, Bettelheim KA, Do TN, Driesen S, Walker MJ, Chin J (2007) Comparative analysis of virulence genes, genetic diversity and phylogeny of commensal and enterotoxigenic *Escherichia coli* isolates from weaned pigs. Appl Environ Microbiol 73:83–91
32. Yang H–H, Vinopal RT, Grasson D, Smets BF (2004) High diversity among environmental *Escherichia coli* isolates from a bovine feedlot. Appl Environ Microbiol 70:1528–1536