E. Manville,[1] M. S. Bhullar,[2] L. Nwadike,[3,4] A. Mustapha,[4] and V. Trinetta[1*]

[1]Food Science Institute, Kansas State University, 216 Call Hall, Manhattan, KS 66503, USA
[2]Dept. of Horticulture and Natural Resources, Kansas State University, 22201 W. Innovation Drive, Olathe, KS 66061, USA
[3]Kansas State University and the University of Missouri, K-State Research and Extension, 22201 W. Innovation Drive, Olathe, KS 66061, USA
[4]Dept. of Food Science, University of Missouri, 220 Eckles Hall, Columbia, MO 65211, USA

PEER-REVIEWED ARTICLE

Food Protection Trends, Vol 43, No. 4, p. 329–342
https://doi.org/10.4315/FPT-22-038

Characterization of *Escherichia coli* Isolates from Agricultural Water on Kansas and Missouri Fresh Produce Farms by Whole-Genome Sequencing

ABSTRACT

Contaminated agricultural water has been a source of pathogenic *Escherichia coli* in recent produce-related outbreaks. The purpose of this study was to characterize *E. coli* isolates from agricultural water sources by using whole-genome sequencing (WGS) to better understand contamination routes. Groundwater and surface water samples were collected quarterly from five farms in Missouri and Kansas over a 1-year period. Samples were tested for generic *E. coli* by using U.S. Environmental Protection Agency Method 1603, and presumptive *E. coli* colonies were isolated. In total, 570 isolates were analyzed by PCR, with 191 of these isolates confirmed as *E. coli*. WGS was completed using an Illumina MiSeq system. The de novo genome assemblies were obtained with Shovill pipeline version 0.9. The NCBI Pathogen Detection system was used to identify antimicrobial resistance (AMR) genes. The prevalence of *E. coli* was higher during spring and summer than winter. A diverse serotype pool was observed where more than 53% of isolates could be linked to a bovine source as the potential animal host. An AMR analysis showed that 100% of isolates carried at least two antimicrobial resistance genes. Recognizing the diversity of *E. coli* may help guide agricultural water assessments as proposed in the new agricultural water rule Food Safety Modernization Act Produce Safety Rule.

INTRODUCTION

Escherichia coli is a bacterial species commonly found in the intestinal tracts of humans and animals *(78)*. Not all *E. coli* strains are pathogenic. Nevertheless, over the years, numerous foodborne outbreaks linked to pathogenic *E. coli* have been traced back to agricultural water systems *(17, 28, 47)*. Shiga toxin–producing *E. coli* (STEC) are an important cause of foodborne disease that, upon ingestion, can result in severe gastrointestinal diseases *(66)*. Other strains of concern that have been associated with waterborne outbreaks include enterotoxigenic *E. coli*, enteropathogenic *E. coli* (EPEC), enterohemorrhagic *E. coli* (EHEC), and enteroinvasive *E. coli* *(31)*. These pathogens have been found to survive in various environments with a high evolutionary capacity *(70)*. With the increasing awareness of the survivability of *E. coli* in water

*Author for correspondence: Phone: +1 785.532.1667; Fax: +1 785.532.5861; Email: vtrinetta@ksu.edu

and the advancements in molecular methodologies, research efforts seek an understanding of pathogen transmission routes into agricultural water systems and their impact on public health.

The Food Safety Modernization Act (FSMA) Produce Safety Rule (PSR) established requirements for produce farmers to assess the quality of agricultural water sources because water sources have been found to be one of the most important pathways of produce contamination *(67)*. Because generic *E. coli* is recognized as an indicator of fecal contamination *(58)*, its presence and quantity can help guide growers' actions to improve agricultural water quality. Finding *E. coli* in water does not directly indicate the presence of pathogenic microorganisms; however, it does indicate that there is an increased risk of the presence of fecal-borne bacteria *(58)*.

Recent multistate produce outbreak investigations have highlighted the need for implementing control and analysis programs in accordance with new regulatory frameworks *(68)*. It is important for produce growers to understand potential sources of contamination in agricultural water systems and implement control measures to reduce the likelihood of such contamination. The importance of using rapid methods to trace contaminations in the food supply chain, compared with more traditional and time-consuming tests, will improve outbreak investigations *(14)*. Genomic techniques, such as whole-genome sequencing (WGS), allow for rapid identification of pathogens and have led to the development of databases (e.g., GenomeTrakr) with information available to assist microbial monitoring, surveillance, and source identification across the food industry.

The produce industry is rapidly expanding in Kansas and Missouri, with a noticeable increase in the production of specialty crops. Based on the 2017 U.S. Census of Agriculture (http://quickstats.nass.usda.gov/?source_desc=CENSUS) estimates, there are 1,449 vegetable farms in Missouri and 498 such farms in Kansas. A recent report *(28)* indicates that agricultural water is one of the least understood topics of the FSMA PSR in midwestern states and that chemical or physical methods for water treatments are not commonly used, especially on small- and medium-sized produce farms (reality of Kansas and Missouri). Many of these produce farms are excluded or exempted from the FSMA PSR. Nevertheless, it is important to support the growing industry in these states and ensure produce safety, because each year, Kansas and Missouri growers produce approximately $26 million and $81 million of fruits and vegetables, respectively *(28)*.

In a previous study conducted from 2018 to 2020 *(28)*, a Kansas State University–University of Missouri Extension produce safety team sampled 426 agricultural water sources in Kansas and Missouri produce farms for a comparative assessment of microbial quality. In the current study, we evaluated and compared the prevalence of microbial contamination in water sources collected on Kansas and Missouri produce farms. No difference in prevalence of *E. coli* was observed between the water sources from the two states, and overall, the number of *E. coli* reported in surface water (exposed to the environment) sources was statistically greater than that in groundwater (below the Earth's surface) sources *(28)*. Five produce farms were identified as hot spots, having high most probable number (MPN)/100 ml values of generic *E. coli*. The microbial testing threshold value was set at >2,419.6 MPN/100 ml. To better identify possible risk factors on produce farms and educate growers on safe management practices, the present research aimed to further characterize *E. coli* isolates from agricultural water sources collected on selected Kansas and Missouri fresh produce farms by using WGS and trace isolates back to the source that likely contaminated the water.

MATERIALS AND METHODS

Sample collection

Based on a previous study *(28)*, where produce farms in Missouri and Kansas received free water analysis, five farms with water sources of high MPN values (MPN/100 ml) of generic *E. coli* were selected for further investigation. In total, nine water sources on these five produce farms were sampled quarterly over a 1-year period (2020 to 2021) during fall, winter, spring, and summer. Weather reports available at The Weather Channel (www.weather.com) were used to record temperatures on the day of sampling. All pond water sources (named with capital letters) were collected on farms with domesticated animals, including ovine, porcine, bovine, canine (dogs), and poultry. All ponds, excluding pond A, were surrounded by a fence, and many were inaccessible to domesticated animals. *Table 1* gives the list of the water sources sampled in this study by location (state) and produce farm ID. Sources included groundwater and surface water. All groundwater samples were collected from a well, whereas the surface water samples were collected from ponds and a cistern from rainwater catchment. Subsamples at various depths and positions across the body of water were taken for each water source and combined to create one sample. Pond water sources had a 500-ml composite sample of five 100-ml samples randomly collected around the circumference of the pond. For cistern samples, 300-ml composite samples were collected, with three 100-ml samples taken at varying depths in the cistern. Well samples were collected through an attached water pump with two 100-ml samples taken after allowing water to run for 30 seconds, making a 200-ml composite sample. Samples were collected in a sanitary manner in sterilized 120-ml polyethylene bottles containing sodium thiosulfate (IDEXX, Westbrook, MN, USA). For the surface water sampling, collection bottles were fixed to a 3.5-m-long pole in a way that allowed the bottle to enter the water inverted just below the surface of the water, be

TABLE 1. Water sources sampled in this study from the five selected produce farms in Kansas and Missouri

Farm ID	Location (state)	Water source	Sample ID
1	Missouri	Pond	A
		Pond	B
2	Missouri	Pond	C
		Pond	D
3	Missouri	Pond	E
		Well	M
4	Kansas	Pond	F
5	Kansas	Cistern	G
		Pond	H

rotated to collect the water sample, and turned upright to be removed from the water *(Fig. 1)*. Upon collection, retrieved bottles were capped and put on ice to be transported for further processing within 12 h of the sampling period.

Figure 1. Depiction of method used to collect samples from surface water sources by using a collection bottle fixed to a long pole in a way that allowed the bottle to enter the water inverted just below the surface of the water, be rotated to collect the water sample, and turned upright to be removed from the water.

Isolation of *E. coli*

The U.S. Environmental Protection Agency Method 1603 was followed for water collection, filtration, and colony isolation *(65)*. In brief, 100 ml was taken from the composite sample to be filtered through a 0.45-µm membrane filter. Sample dilutions were completed for spring and summer samples to 1/10 and 1/100, respectively, whereas no dilution was performed for fall and winter samples. This difference was due to the overwhelming quantity of isolates collected through filtration during the spring and summer, making colony isolation difficult. After filtration, membrane filters were placed on modified membrane-thermotolerant *E. coli* agar (mTEC; BD Difco, Sparks, MD, USA). Presumptive *E. coli* colonies (magenta and/or red) were counted and reported as *E. coli* CFU/100 ml. Up to five presumptive positive colonies per water sample were streaked for isolation on tryptic soy agar (TSA; BD Difco). From each TSA plate, a single colony was selected and transferred to tryptic soy broth (BD Difco). Cells were preserved using CryoCare bacteria preservers (Key Scientific Products, Inc., Stamford, TX, USA) and stored at −80°C until further use.

Identification and confirmation of *E. coli* by PCR

A modified PCR protocol based on Walker et al. *(75)* was used. Primers targeting *clpB*, *uidA*, and *ybbW* (Biosearch Technologies, Petaluma, CA, USA) genes for *E. coli* spp. isolates are listed in *Table 2*. All primer stocks were prepared in 1× Tris-EDTA buffer (Integrated DNA Technologies, Coralville, IA, USA) to achieve a 100 µM stock concentration. Equal volumes and concentrations of primers were mixed with the addition of nuclease-free water to reach a working solution of 10 µM. Each reaction contained a combined total volume of 20 µl that included 1 µl of the three-target primer mix at 10 µM solution concentration, 10 µl of iQ Multiplex Powermix (Bio-Rad, Hercules, CA, USA), 5 µl of template DNA, and 4 µl of nuclease-free water. The parameters for the thermal cycler were set as follows: denaturation at 94°C for 5 minutes, followed by 29 cycles of 94°C for 30 seconds, 62°C for 30 seconds, and 68°C for 75 seconds, and a final extension step of 7 minutes at 68°C. All reactions were completed using a CFX96 touch real-time PCR detection system (Bio-Rad).

To separate amplified DNA fragments, PCR products were run on a 2200 TapeStation (Agilent, Waldronn, Germany). Positive *E. coli* samples showed tight bands with amplicons of

TABLE 2. Primer design for *E. coli* PCR identification and confirmation

Target	Sequence (5′–3′)
clpB	Forward: CATACGAATGCTGGATGCTG
	Reverse: TTTGAAGAACGTTTAAAAGGCG
uidA	Forward: ACCACGGTGATATCGTCCAC
	Reverse: TACAAGAAAGCCGGGCAAT
ybbW	Forward: AATCTGGCCGGGATTTTT
	Reverse: 5TGGCTCCGGCAATAATACAT

449, 454, and/or 447 bp, and to further confirm the presence of *E. coli*, positive samples were streaked on MacConkey agar (BD Difco).

DNA preparation

DNA extraction, library preparation, sequencing, and *in silico* analysis were performed as described by Domesle et al. *(20)*, with some modifications. Genomic DNA was extracted using DNeasy blood & tissue kits (Qiagen, Hilden, Germany). DNA concentrations were measured with a Qubit fluorometer 3.0 by using the dsDNA HS assay kit, according to the manufacturer's instructions (Thermo Fisher). DNA extracts were stored at −20°C until WGS analysis.

WGS characterization

The Nextera XT library preparation kit (Illumina, San Diego, CA, USA) was used for preparing paired-end libraries, and WGS was carried on a MiSeq sequencer using a 2 × 300-cycle MiSeq reagent V3 kit (Illumina). Trimming and *de novo* assemblies were obtained with Shovill version 0.9 (https://github.com/tseemann/shovill), available in the GalaxyTrakr pipeline (http://galaxytrakr.org/) *(2)*. Default parameters were used in all analyses *(80)*. The NCBI Prokaryotic Genomes Automatic Annotation pipeline (https://www.ncbi.nlm.nih.gov/genome/annotation_prok/) was used to annotate draft genomes of each isolate. Characterization was performed using an *in silico* multilocus sequence typing (MLST) for *E. coli*, based on genome sequences of seven housekeeping genes, and core genome MLST analysis by using EnteroBase version 1.1.2 (https://enterobase.warwick.ac.uk/).

Isolate characterization

The genetic relatedness of the isolates sequenced was compared with sequences of more than 200,000 *E. coli* isolates from various environmental and clinical sources available in the NCBI Pathogen Detection database. Through this database, the antimicrobial resistance (AMR) profiles of each isolate were identified through the AMRFinderPlus 3.10 tool (https://www.ncbi.nlm.nih.gov/pathogens/antimicrobial-resistance/AMRFinder/). Further characterization of *E. coli* isolates was conducted *in silico* through the Center for Genomic Epidemiology services. Virulence factors were predicted through the alignment of draft genome assembles per isolate against the database through VirulenceFinder 2.0 (version 2.0.3) (http://cge.cbs.dtu.dk/services/VirulenceFinder/) https://bio.tools/virulencefinder. Virulent factors that passed the default threshold of ≥ 90% nucleotide identity and ≥ 60% coverage were accepted as present in the isolate. Rapid serotype identification was completed using SerotypeFinder 2.0 (version 2.0.1), with accepted serotype genes falling within the default threshold of ≥ 85% nucleotide identity and ≥ 60% coverage per isolate. A phylogenetic tree was constructed based on the core genome MLST as defined by Ridom SeqSphere (version 8.4) (Ridom GmbH, Münster, Germany), where genomes were grouped based on serotype identification with distance centered on the seven *E. coli* Warwick genes. A retrospective literature study was used to further group isolates centered around animal reservoirs associated with specific serotypes. Articles obtained as suitable for this study were peer-reviewed original research articles.

Accession numbers

The sequences of all isolates used in this research are available under the BioProjects 357722 and 832234 on the NCBI website: https://www.ncbi.nlm.nih.gov/bioproject/357722 and https://www.ncbi.nlm.nih.gov/bioproject/832234).

RESULTS AND DISCUSSION

Prevalence of presumptive *E. coli* in sampled water sources

In total, 570 presumptive *E. coli* isolates were collected during the 1-year sampling period from nine different agricultural water sources. Pond D had overall the highest counts ($P < 0.05$) among sampling sites and seasons, with

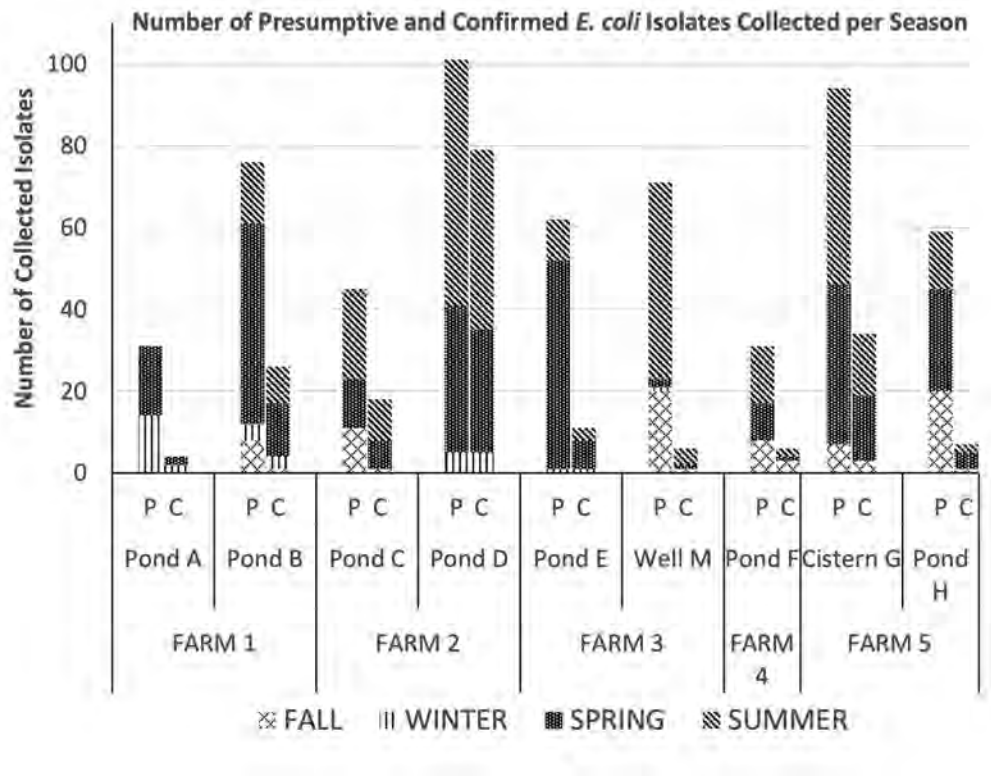

Figure 2. Total number of presumptive (P) and confirmed (C) *E. coli* isolates collected quarterly (fall, winter, spring, summer) over a 1-year period (2020 to 2021) shown per water source.

300 CFU/100 ml of water. Presumptive isolates were all tested by PCR: 33.5% (191/570) of samples were confirmed *E. coli*. The mTEC agar method is the current recommended protocol used to measure the presence of *E. coli* as a fecal indicator in water samples. This method requires specific multiple-colony verification protocols using taxonomic testing and highly trained staff. This could be a reasonable explanation for the variability between presumptive and confirmed isolates that we report in our study.

Most samples were recovered during summer (47%), followed by spring (42%), fall (6%), and winter (5%) *(Fig. 2)*. As expected, a seasonal effect was observed: in warmer seasons, a higher number of microorganisms was present in the different water sources (i.e., spring and summer, with temperatures reaching >20°C) compared with the winter sampling period, where the recorded temperatures were as low as 0°C. The seasonal impacts on *E. coli* concentrations in water are important to understand when implementing methods that prevent contaminants from entering the food supply. Confounding views on the seasonal variation of *E. coli* concentrations in water sources can be found throughout the scientific literature. Limitations, such as a lack of monitoring uniformity, climatic zone differences, climate change, and sampling size, make it difficult to generalize the seasonal patterns of *E. coli (36)*. Nevertheless, we observed that the summer held the highest concentration of *E. coli* isolates. Byappanahalli et al. *(15)* found high concentrations of *E. coli* in fresh water near Lake Michigan during warmer sampling periods, with concentrations peaking during the late summer. This is also supported by research from Oliver and Page *(51)* and Vivar and Fuentes *(73)*: warmer seasons showed increased numbers of excreted bacteria from animals, leading to water contamination. Combined with weather patterns and high temperatures, this tends to favor high reproductive and strong survivability environments for *E. coli*. Higher air temperatures have been shown to increase microbial concentrations in water, although the relationship of other meteorological factors also greatly influences microbial load *(76)*.

Provided the limited number of groundwater sources collected within the parameters of this study, it is difficult to understand whether surface and groundwater sources carry equal risks when characterizing *E. coli*. It can however be noted that the cistern and well were observed to be safeguarded water sources inaccessible to domestic and wild animals, but still resulted in *E. coli* contamination. The cistern sampled had a large stone lid covering the reservoir's opening, but was not consistently covered and was housed in a shed near the main entrance of the building where people were frequently passing. Cisterns can be contaminated with *E. coli* through organic debris

or associated animal feces, whether that be through roof runoff or dust kicked up by foot traffic. This is supported by Lévesque et al. *(41)*, who found that soil and dust carried by wind can be washed into freshwater containers after analyzing rainwater collection tanks in Bermuda, with 66% of the 102 tanks sampled containing *E. coli*. Flood and heavy rains have also been known to wash debris and fecal pollutants into rain catchment systems and cisterns *(18)*. It is important to note that maintaining the water quality of a cistern can be difficult, with numerous routes of microbial contamination, and should be regularly monitored *(18)*. The well was the only groundwater source and had a covered opening with a hand pump to collect water. The well was next to a pond (pond E), and even though the well was not exposed to the environment, large quantities of *E. coli* could be found. Based on our observations, cross-contamination from the pond could have been possible. Andrade et al. *(4)* addresses three ways *E. coli* could contaminate well water: (i) directly through well opening, (ii) water recharge/deep drainage (geological pathways including shifts in soil), and (iii) direct migration (contaminated groundwater mixing with noncontaminated groundwater). Any of these factors could explain the transfer of *E. coli* contaminates into the well sampled during this study. Sasakova et al. *(59)* found contaminated surface water could eventually lead to the contamination of groundwater sources and that an aquifer environment could facilitate pathogen survival. Many factors can facilitate the microbial quality of groundwater sources, so it is important to note that these sources are still susceptible to contaminants.

Isolate characterization

In total, 99 different serotypes of *E. coli* could be identified *(Table 3)*. Reports of *E. coli* in water have exposed high discrepancies of serotypes with both pathogenic and nonpathogenic isolates. Maloo et al. *(45)* and Ramteke and Tewari *(54)* also noted the remarkable diversity found among the serotypes collected from both recreational water and drinking water. The most prevalent serotype identified in our study was *E. coli* O65:H38, with 12 identified isolates, all of which were collected during spring. Of the total *E. coli* isolates, only two were found to be a part of the "Big 6," including O26 (SAMN28816604) and O45 (SAMN23828792) *(10)*. Each isolate was however sourced from two different ponds during different seasons: O26 was isolated during spring and O45 during summer.

In *Table 3* and *Figure 3*, isolates are grouped based on the same flagellar H-antigens. The genetic relations inferred from phylogenetic relatedness *(Fig. 3)* can be used in source attribution and might indicate the same contamination source. Therefore, to quantify the relative importance of animal hosts as pathogen reservoirs, we performed a literature review to understand which serotypes could be historically linked to a potential host and therefore contamination source. The link is only theoretical, nevertheless it is useful to initially understand the potential risks for these water sources. Papers published earlier than 2010 were excluded from our search, and only publications linked to water source illness and *E. coli* contamination were considered. Literature supporting serotype-based identification shows isolates deriving from more than one potential host. Bovine sources were found to be the most common source of likely contamination, encompassing 53% of the *E. coli* serotypes. Other isolates could likely be traced back to human (38%), ovine (17%), poultry (11%), porcine (8%), and deer (2%) sources.

Observations collected during the sampling periods support the likely causes of contamination linked through serotype identification. As mentioned, all water samples were collected on farms with domesticated animals. Even if ponds were protected by a fence, the presence of *E. coli* in our samples indicates that the pathogen could be spread regardless of the confinement. Osman et al. *(52)* found domestic calves, sheep, and goats to harbor many of the same *E. coli* serovars as confirmed in this research. This finding supports the hypothesis that there could be a direct or indirect animal-to-human or animal-to-animal transmission of *E. coli*, creating overlap in serotype profiles *(33)*. Half of the isolates characterized in this study could be traced back to bovine origination. *E. coli* is a natural part of the ruminant microbiota; nevertheless, it has been suggested that even wildlife could be an asymptomatic reservoir *(22, 61)*. A deer running through a cow field could carry *E. coli* from that field on its hooves and eventually contaminate a water source. *E. coli* can be carried on animals' fur, hooves, paws, skin, feathers, and feces, thereby spreading *E. coli* to the areas where they roam *(16)*. *E. coli* contamination in agricultural water as observed is difficult to track, with numerous direct and indirect routes of transmission.

At least two antimicrobial resistance genes were expressed in all the isolates collected *(Table 3)*. More than 15 different AMR genes were reported through *in silico* analysis, covering many of the antibiotic classes. Each antibiotic class was grouped based on chemical structure *(77)*. The most prevalent AMR gene was *blaEC*, expressing resistance to β-lactam, spanning across 98% of the samples. The next most prevalent Amr gene was *mdtM* (94%), a gene expressing a multidrug-resistant protein including resistance to nucleoside, phenicol, lincosamide, and fluoroquinolone antibiotics through an efflux pump complex *(3)*. Fifty-seven percent of the isolates had the AMR gene *acrF* (57%), encoding an efflux pump resistance mechanism to fluoroquinolone *(3)*. Finally, the other most frequent genes expressed resistance to tetracyclines (*tetA* [14%] and *tetB* [10%]) and aminoglycosides (aph(3")-Ib [15%] and aph(6)-Id [15%]). Coinciding with the results of our study, Liao et al. *(42)* found all collected *E. coli* isolates from an aquafarm to have at least two drug resistance genes, indicating a multidrug-resistant sampling pool. One study suggests bodies of water are ideal

TABLE 3. Most frequent antimicrobial (AMR) and virulent genes found from *E. coli* isolates and potential source identification of isolates based on a literature review of where *E. coli* serotypes can be found *(1, 5-8, 11-13, 17, 19, 21, 23, 25, 26, 30, 32, 34-35, 37-38, 40, 44, 46-47, 50, 53, 55-57, 60, 62-63, 71, 74, 79, 81)*

				AMR							VIRULENCE							POTENTIAL ANIMAL HOST							
Biosample ID	Location	Water Source	Serotype	*arcF*	*blaEC*	*mdtM*	*aph(3'')-Ib*	*aph(6)-Id*	*tet(A)*	*tet(B)*	*gad*	*lpfA*	*terC*	*ompT*	*iss*	*traT*	*stx2*	Human	Bovine	Ovine	Caprine	Porcine	Poultry	Domestic Pet	Deer
SAMN28571028	MO	Pond D	O74:H8	+	+	+					+		+	+	+	+			+						
SAMN23828784	MO	Pond B	O8:H8	+	+	+					+		+	+	+				+	+					
SAMN23828785	MO	Pond B	O8:H8	+	+	+							+	+	+				+	+					
SAMN23828788	MO	Pond B	O8:H8	+	+	+							+		+				+	+					
SAMN23828789	MO	Pond B	O8:H8	+	+	+							+		+				+	+					
SAMN23828790	MO	Pond B	O8:H8	+	+	+							+	+	+				+	+					
SAMN23828791	MO	Pond B	O8:H8	+	+	+							+	+	+				+	+					
SAMN23828811	MO	Pond D	O74:H8	+	+	+					+		+	+	+	+			+						
SAMN23828820	MO	Pond D	O74:H8	+	+	+					+		+	+	+	+			+						
SAMN23894232	MO	Pond B	O8:H8	+	+	+							+	+	+					+					
SAMN23894240	MO	Pond D	O74:H8	+	+	+							+	+	+	+			+						
SAMN23894243	MO	Pond D	O74:H8	+	+	+							+	+	+	+			+						
SAMN23894250	MO	Pond E	O8:H8	+	+	+							+	+	+					+					
SAMN28544968	MO	Pond C	:H16	+	+	+												+	+			+	+		
SAMN28816569	MO	Pond B	O175:H16		+	+						+								+					
SAMN28816586	MO	Pond B	O175:H16		+	+						+								+					
SAMN28061266	KS	Cistern G	:H16	+	+	+							+					+	+			+	+		
SAMN28816574	MO	Pond B	O175:H16		+	+						+								+					
SAMN28061261	KS	Cistern G	:H16	+	+	+							+					+	+			+	+		
SAMN23828823	MO	Pond D	O175:H16		+	+						+													
SAMN23894246	MO	Pond D	O140:H36	+	+	+					+	+													
SAMN23894234	MO	Pond C	:H7	+	+	+					+	+	+	+	+			+	+						
SAMN28571031	MO	Pond D	O8:H7	+	+	+					+	+	+					+							
SAMN23894238	MO	Pond D	O8:H7	+	+	+					+	+	+					+							
SAMN23828800	MO	Pond D	O22:H8	+	+	+					+	+		+			+		+					+	
SAMN23828816	MO	Pond D	O22:H8	+	+	+					+	+		+			+		+		+			+	
SAMN23828817	MO	Pond D	O22:H8	+	+	+					+	+		+			+		+		+			+	
SAMN23894239	MO	Pond D	O22:H8	+	+	+					+	+		+			+		+		+			+	
SAMN28816588	MO	Pond E	O88:H7	+	+	+				+	+	+	+	+		+			+						
SAMN23828819	MO	Pond D	:H7	+	+	+					+	+	+			+									
SAMN23894248	MO	Pond E	O147:H7	+	+	+					+	+	+	+	+			+							
SAMN24300382	KS	Cistern G	O8:H8	+	+	+					+		+							+					
SAMN28816589	MO	Pond A	O6:H49	+	+	+					+	+	+							+					
SAMN23894227	MO	Pond B	O6:H49	+	+	+					+	+	+							+					
SAMN23894231	MO	Pond D	O148:H28	+	+	+						+	+					+							
SAMN28816579	MO	Pond B	O8:H7	+	+	+					+	+	+		+			+							
SAMN23894260	KS	Cistern G	:H20	+	+	+					+	+	+	+				+							
SAMN28816597	KS	Cistern G	O144:H7	+	+	+					+	+	+												
SAMN23828796	MO	Pond C	O178:H7	+	+	+					+	+	+						+						+
SAMN23828818	MO	Pond D	O178:H7	+	+	+					+	+	+						+						+
SAMN28816618	MO	Pond B	O22:H7	+	+	+					+	+	+						+						
SAMN28816582	MO	Pond B	:H7	+	+	+					+	+	+					+	+						
SAMN28816572	MO	Pond B	O22:H7	+	+	+					+	+	+						+						
SAMN23828806	MO	Pond D	O22:H7	+	+	+					+	+	+						+						
SAMN28544844	MO	Pond D	:H34	+	+	+						+		+				+	+						
SAMN28816575	MO	Well M	O39:H34	+	+	+					+	+													
SAMN28816583	MO	Pond D	:H2	+	+	+					+			+	+			+	+						
SAMN23828804	MO	Pond D	:H2	+	+	+					+			+	+										

(Continued on the next page)

TABLE 3. Most frequent antimicrobial (AMR) and virulent genes found from *E. coli* isolates and potential source identification of isolates based on a literature review of where *E. coli* serotypes can be found *(1, 5-8, 11-13, 17, 19, 21, 23, 25, 26, 30, 32, 34-35, 37-38, 40, 44, 46-47, 50, 53, 55-57, 60, 62-63, 71, 74, 79, 81)* (cont.)

				AMR							VIRULENCE							POTENTIAL ANIMAL HOST							
Biosample ID	Location	Water Source	Serotype	*arcF*	*blaEC*	*mdtM*	*aph(3'')- Ib*	*aph(6)-Id*	*tet(A)*	*tet(B)*	*gad*	*lpfA*	*terC*	*ompT*	*iss*	*traT*	*stx2*	Human	Bovine	Ovine	Caprine	Porcine	Poultry	Domestic Pet	Deer
SAMN28816619	MO	Pond D	O43:H2	+	+	+	+	+		+	+	+							+						
SAMN23828803	MO	Pond D	O43:H2	+	+	+	+	+		+	+	+							+						
SAMN23828805	MO	Pond D	O43:H2	+	+	+	+	+		+	+	+							+						
SAMN23828807	MO	Pond D	O43:H2	+	+	+	+	+		+	+	+							+						
SAMN23828808	MO	Pond D	O43:H2	+	+	+	+	+		+	+	+							+						
SAMN23828814	MO	Pond D	O43:H2	+	+	+	+	+		+	+	+							+						
SAMN23828824	MO	Pond D	O43:H2	+	+	+	+	+		+	+	+							+						
SAMN23894237	MO	Pond D	:H2	+	+	+	+	+		+	+	+						+	+						
SAMN23894244	MO	Pond D	O43:H2	+	+	+	+	+		+	+	+							+						
SAMN24300391	MO	Pond D	O43:H2	+	+	+	+	+		+	+	+							+						
SAMN28816593	KS	Pond H	O136:H12	+	+	+					+			+					+						
SAMN24300379	KS	Pond H	O147:H21	+	+	+					+	+	+			+		+							
SAMN28544843	MO	Pond D	O65:H38	+	+	+					+	+	+	+						+					
SAMN28545094	MO	Pond D	O65:H38	+	+	+					+	+	+	+						+					
SAMN28544964	MO	Pond D	O65:H38	+	+	+					+	+	+	+						+					
SAMN28545100	MO	Pond D	O65:H38	+	+	+					+	+	+	+						+					
SAMN28544963	MO	Pond D	O65:H38	+	+	+					+	+	+	+						+					
SAMN28544841	MO	Pond D	O65:H38	+	+	+					+	+	+	+						+					
SAMN28571030	MO	Pond D	O65:H38	+	+	+					+	+	+	+						+					
SAMN28571026	MO	Pond D	O65:H38	+	+	+					+	+	+	+						+					
SAMN28816613	MO	Pond D	O65:H38	+	+	+					+	+	+	+						+					
SAMN28571033	MO	Pond D	:H38	+	+	+						+	+	+				+					+		
SAMN28571036	MO	Pond D	O65:H38	+	+	+					+	+	+	+						+					
SAMN28571039	MO	Pond D	O65:H38	+	+	+					+	+	+	+						+					
SAMN28571035	MO	Pond D	O65:H38	+	+	+					+	+	+	+						+					
SAMN23894235	MO	Pond D	:H21	+	+	+												+	+						
SAMN23894252	KS	Pond F	O8:H7	+	+	+					+	+	+		+			+							
SAMN28816600	KS	Pond H	O8:H19	+	+	+	+	+	+	+	+		+	+					+	+				+	
SAMN28816578	MO	Pond E	:H21	+	+	+					+	+	+					+	+						
SAMN28545097	MO	Pond D	O1:H7	+	+	+					+	+				+		+	+						
SAMN28545096	MO	Pond D	O1:H7	+	+	+			+		+	+			+	+		+	+						
SAMN28816558	MO	Pond D	O1:H7	+	+	+			+		+	+			+	+		+	+						
SAMN28544966	MO	Pond D	O1:H7	+	+	+			+		+	+			+	+		+	+						
SAMN28545098	MO	Pond D	O1:H7	+	+	+			+		+	+			+	+		+	+						
SAMN28571027	MO	Pond D	O1:H7	+	+	+					+	+			+	+		+	+						
SAMN28571037	MO	Pond D	O1:H7	+	+	+					+	+			+			+	+						
SAMN28571029	MO	Pond D	O1:H7	+	+	+			+		+	+			+	+		+	+						
SAMN23828792	MO	Pond C	O45:H8	+	+	+					+	+			+	+			+						
SAMN23828809	MO	Pond D	O3:H8	+	+	+					+	+		+		+						+			
SAMN23894228	MO	Pond B	O149:H45		+	+					+	+		+		+		+							
SAMN23894236	MO	Pond D	:H9	+	+	+					+	+			+			+	+						
SAMN28061271	KS	Cistern G	:H7	+	+	+					+	+	+	+				+	+						
SAMN28816576	MO	Well M	:H21	+	+	+					+	+			+			+	+						
SAMN23894221	KS	Cistern G	:H21	+	+	+						+		+	+			+	+						
SAMN23894222	KS	Cistern G	:H21	+	+	+						+		+	+			+	+						
SAMN23894225	MO	Pond A	:H21	+	+	+					+	+		+	+			+	+						
SAMN23828793	MO	Pond C	:H21	+	+	+					+	+		+	+			+	+						
SAMN28816605	KS	Pond H	:H2/H47	+	+	+					+	+						+	+						

(Continued on the next page)

TABLE 3. Most frequent antimicrobial (AMR) and virulent genes found from *E. coli* isolates and potential source identification of isolates based on a literature review of where *E. coli* serotypes can be found *(1, 5-8, 11-13, 17, 19, 21, 23, 25, 26, 30, 32, 34-35, 37-38, 40, 44, 46-47, 50, 53, 55-57, 60, 62-63, 71, 74, 79, 81)* (cont.)

				AMR							VIRULENCE							POTENTIAL ANIMAL HOST							
Biosample ID	Location	Water Source	Serotype	*arcF*	*blaEC*	*mdtM*	*aph(3'')- Ib*	*aph(6)-Id*	*tet(A)*	*tet(B)*	*gad*	*IpfA*	*terC*	*ompT*	*iss*	*traT*	*stx2*	Human	Bovine	Ovine	Caprine	Porcine	Poultry	Domestic Pet	Deer
SAMN23894229	MO	Pond B	:H7	+	+	+				+	+	+	+	+	+	+		+	+						
SAMN28816615	MO	Pond E	:H10	+	+	+						+				+		+	+			+	+		
SAMN28816557	MO	Pond E	:H10	+	+	+					+	+				+		+	+			+	+		
SAMN23828822	MO	Pond D	:H11/H54	+	+	+				+	+	+						+	+						
SAMN23828787	MO	Pond B	O120:H10	+	+	+					+	+		+				+	+						
SAMN23828810	MO	Pond D	O4:H10	+	+	+					+	+							+						
SAMN28544842	KS	Cistern G	O23:H16	+	+	+	+	+	+		+												+		
SAMN28061098	KS	Cistern G	:H16	+	+	+	+	+	+		+							+	+			+	+		
SAMN28544972	KS	Cistern G	O23:H16	+	+	+	+	+	+							+							+		
SAMN24300383	KS	Cistern G	O23:H16	+	+	+	+	+	+			+											+		
SAMN24300384	KS	Cistern G	O23:H16	+	+	+	+	+	+		+	+											+		
SAMN24300386	KS	Cistern G	O23:H16	+	+	+	+	+	+		+	+				+							+		
SAMN24300387	KS	Cistern G	O23:H16	+	+	+	+	+	+														+		
SAMN24300388	KS	Cistern G	O23:H16	+	+	+	+	+	+			+											+		
SAMN23894216	MO	Pond C	O9:H10	+	+						+							+							
SAMN23894259	KS	Cistern G	O9:H30	+	+	+	+	+			+											+			
SAMN23894262	KS	Cistern G	O9:H30	+	+	+	+	+			+											+			
SAMN23828798	MO	Pond C	O155/08:H20	+	+	+					+														
SAMN28816603	MO	Pond B	O141:H16	+	+	+					+	+		+	+				+						
SAMN23828813	MO	Pond D	O141:H16	+	+	+					+	+		+	+				+						
SAMN23828821	MO	Pond D	O141/08:H16	+	+	+					+	+		+	+				+						
SAMN23828799	MO	Pond D	O8:H28	+	+	+					+	+													
SAMN23828815	MO	Pond D	O8:H28	+	+	+					+														
SAMN23894220	KS	Pond F	O7:H16	+	+	+			+		+	+	+			+			+						
SAMN28545093	MO	Pond D	O35:H21	+							+	+	+			+									
SAMN28571034	MO	Pond D	:H21	+							+	+	+			+		+	+						
SAMN23894219	KS	Pond F	O30:H25	+	+	+				+	+	+	+												
SAMN23894247	MO	Pond D	O3:H21	+	+	+					+	+	+	+							+				
SAMN23894257	KS	Cistern G	:H21	+	+	+			+		+	+	+					+	+						
SAMN24300380	KS	Pond H	O22:H21	+	+	+					+	+	+						+						
SAMN28571038	MO	Pond D	O159:H21	+	+	+					+	+	+						+				+		
SAMN23894223	KS	Cistern G	O88:H21	+	+	+	+	+	+		+	+	+	+		+			+						
SAMN23894224	KS	Pond H	O88:H21	+	+	+	+	+	+		+	+	+	+		+			+						
SAMN23894256	KS	Cistern G	:H21	+	+	+					+	+	+					+	+						
SAMN28816550	MO	Well M	O156:H12	+	+	+					+	+	+												
SAMN23894217	MO	Well M	O59:H28	+	+	+					+	+	+												
SAMN23894249	MO	Pond E	O12:H16	+	+	+					+														
SAMN23894230	MO	Pond D	O8:H19	+	+	+					+	+							+						
SAMN23894255	KS	Cistern G	O8:H19	+	+	+					+	+							+	+				+	
SAMN24300377	MO	Pond D	O8:H19	+	+	+						+							+	+				+	
SAMN28816551	KS	Cistern G	:H38	+	+	+					+							+					+		
SAMN28816617	KS	Cistern G	O9:H12	+	+	+							+						+						
SAMN28816602	MO	Pond B	:H26	+	+								+					+							
SAMN23828794	MO	Pond C	:H26	+	+	+					+	+	+					+							
SAMN28816591	MO	Pond A	O96:H20		+	+					+	+	+							+					
SAMN28544965	MO	Pond D	O150:H45	+	+	+					+														
SAMN28816604	MO	Pond E	O26:H11	+	+	+			+		+	+	+	+	+				+						
SAMN28571032	MO	Pond D	O1:H19	+	+	+						+	+			+			+			+			

(Continued on the next page)

TABLE 3. Most frequent antimicrobial (AMR) and virulent genes found from *E. coli* isolates and potential source identification of isolates based on a literature review of where *E. coli* serotypes can be found *(1, 5-8, 11-13, 17, 19, 21, 23, 25, 26, 30, 32, 34-35, 37-38, 40, 44, 46-47, 50, 53, 55-57, 60, 62-63, 71, 74, 79, 81)* (cont.)

				AMR							VIRULENCE							POTENTIAL ANIMAL HOST							
Biosample ID	Location	Water Source	Serotype	*arcF*	*blaEC*	*mdtM*	*aph(3'')- Ib*	*aph(6)-Id*	*tet(A)*	*tet(B)*	*gad*	*lpfA*	*terC*	*ompT*	*iss*	*traT*	*stx2*	Human	Bovine	Ovine	Caprine	Porcine	Poultry	Domestic Pet	Deer
SAMN23894261	KS	Cistern G	:H31	+	+	+			+		+	+	+					+	+						
SAMN23828786	MO	Pond B	:H19		+	+					+	+	+	+	+								+		
SAMN23894233	MO	Pond C	O9:H14	+	+	+					+	+						+							
SAMN28816596	MO	Pond B	:H42	+	+	+							+					+	+						
SAMN28816611	MO	Pond B	:H42	+	+	+					+		+					+	+						
SAMN28816548	MO	Pond B	:H42	+	+	+					+		+					+	+						
SAMN24300381	KS	Pond F	O149:H2	+		+					+	+		+	+				+	+					
SAMN28061102	MO	Pond E	O84:H2	+	+	+					+	+		+					+						
SAMN23894263	KS	Cistern G	O174:H4	+	+	+						+										+			
SAMN28544970	KS	Cistern G	:H29		+	+			+		+		+						+						
SAMN23894254	KS	Cistern G	O176:H4	+	+	+	+	+	+				+								+				
SAMN28816584	MO	Pond B	:H32	+	+	+					+		+					+							
SAMN28816553	MO	Pond E	O106:H31		+	+					+		+					+							
SAMN28061108	KS	Cistern G	O21:H12	+	+	+	+	+		+	+		+					+							
SAMN28816577	KS	Cistern G	:H12	+	+	+					+		+					+				+	+		
SAMN23828797	MO	Pond C	O180:H10	+	+	+					+		+					+							
SAMN23894253	KS	Pond F	O81:H27	+	+	+					+		+					+							
SAMN23894258	KS	Cistern G	OX13:H10	+	+	+	+	+		+	+		+			+			+							
SAMN24300389	KS	Cistern G	O89:H4	+	+	+			+		+		+												
SAMN23828795	MO	Pond C	O107:H27	+	+	+					+		+					+							
SAMN23894264	KS	Cistern G	:H51	+	+						+							+							
SAMN23894245	MO	Pond D	:H20	+	+	+	+	+	+											+					
SAMN24300390	MO	Pond C	O9:H30		+	+																+			
SAMN28061105	MO	Pond D	:H46	+	+	+	+	+	+					+					+						
SAMN28545095	MO	Pond D	:H34	+	+	+								+				+	+						
SAMN23894241	MO	Pond D	O75:H31	+		+																		+	
SAMN24300376	MO	Pond D	O133:H29	+	+									+		+									
SAMN23828825	MO	Pond D	O138:H28	+	+																				
SAMN28816598	MO	Pond D	O17:H45	+	+	+				+						+							+		
SAMN23828801	MO	Pond D	O73:H45	+	+	+										+									
SAMN23894242	MO	Pond D	O73:H45	+	+	+				+						+									
SAMN23894215	MO	Pond B	O146:H28	+	+								+	+	+										+
SAMN23894226	MO	Pond A	O81:H1	+	+								+	+											
SAMN28816580	MO	Pond E	O113:H5	+	+									+									+		
SAMN23828802	MO	Pond D	O17:H18	+	+	+					+	+		+	+			+					+		
SAMN23894218	KS	Pond F	O15:H18	+	+						+	+		+	+			+							
SAMN24300385	KS	Cistern G	:H45	+	+	+			+					+	+	+							+		
SAMN28816568	MO	Well M	O8:H10	+	+									+								+			+
SAMN28816556	MO	Well M	O148:H8		+	+													+						
SAMN28544969	MO	Pond C	:H2	+	+	+					+	+	+					+	+						
SAMN28545101	MO	Pond C	:H2	+	+	+					+	+	+					+	+						
SAMN28544967	MO	Pond C	:H2	+	+	+					+	+	+			+		+	+						
SAMN28544971	MO	Pond C	:H2	+	+	+					+	+	+					+	+						
SAMN24300378	KS	Pond H	:H2	+	+	+					+	+	+					+	+						
SAMN23828812	MO	Pond D	O39:H7	+	+	+					+	+		+				+	+						
SAMN28544840	MO	Pond C	:H8	+	+	+							+					+	+			+			
SAMN28545099	MO	Pond C	:H8	+	+	+							+					+	+			+			

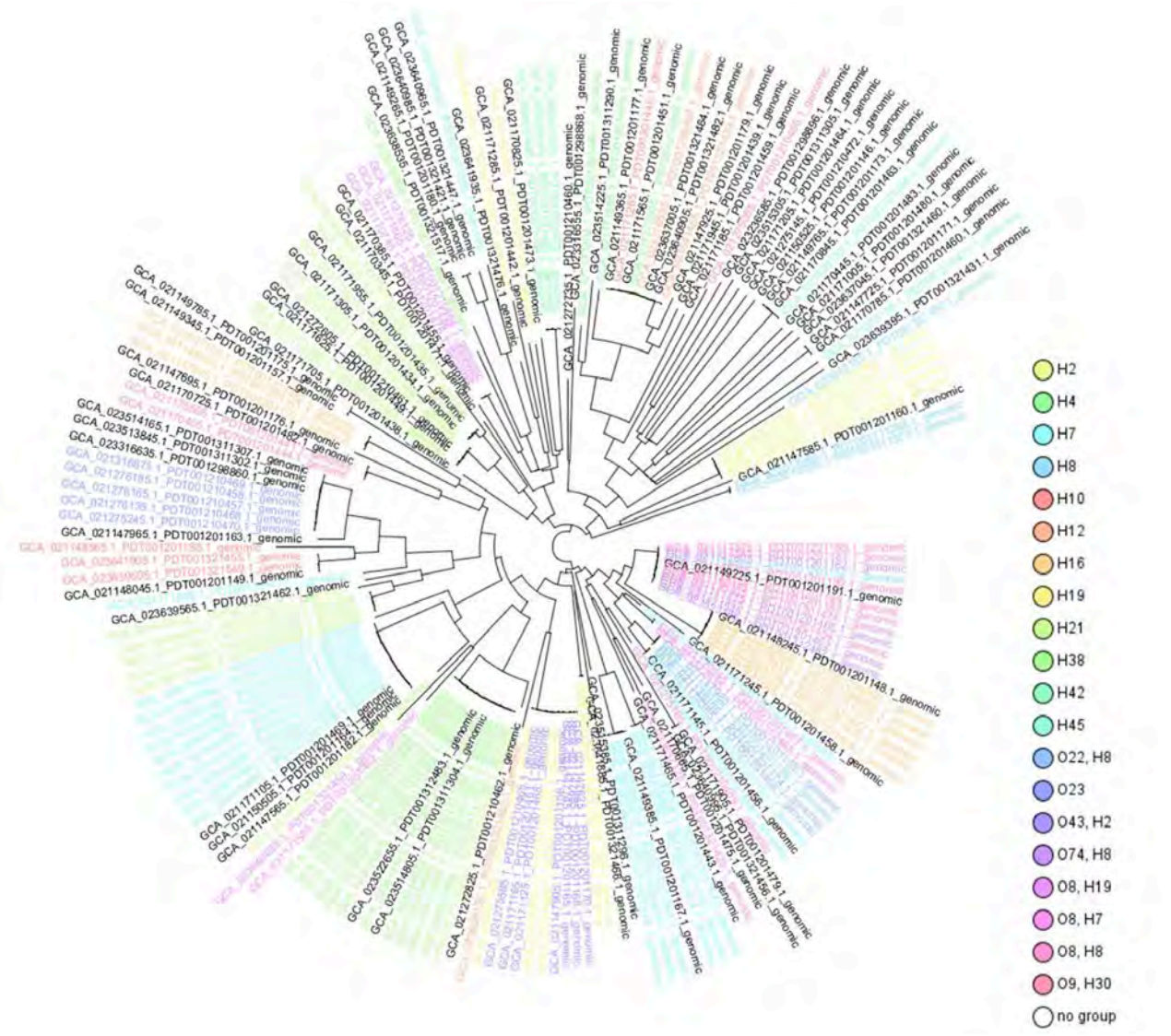

Figure 3. Example of a phylogenic tree of the 191 isolates, grouped based on H-flagella antigens. Distances were based on seven Warwick genes.

places for bacterial AMR evolution to occur naturally due to dense bacterial populations and pollution forcing bacteria to adapt *(43)*. Water sources are opportunistic environments for bacterial AMR gene acquisitions *(42)*. Furthermore, Massé et al. *(48)* found residual antibiotics in the feces of livestock. A few of the most prevalent antibiotics used in livestock include tetracyclines, fluoroquinolones, and β-lactams, all of which appear as AMR genes within the collected isolates.

Virulence genes are also shown in *Table 3*, with more than 45 different genes identified through *in silico* analysis. Four isolates, all serotyped to be O22:H8 (SAMN23828800, SAMN23828816, SAMN23828817, and SAMN23894239), could be confirmed as Shiga toxin–producing (stx_2) isolates, one of the most puissant toxins that is associated with the manifestation of hemolytic uremic syndrome *(49)*. This toxin, when ingested, can cause severe illness and even death and is thus a threat to public health. This is pressing information because many STEC virulent markers are on mobile genetic elements and can be easily transferred to other isolates *(24)*. These STEC isolates were each collected from a pond source during summer.

E. coli isolates collected from water have shown high rates of virulent factors, as reported in numerous surveillance studies, including an investigation of the U.S. Salish Sea, by Vingino et al. *(72)*, and a study of the U.S. Great Lakes by Hamelin et al. *(29)*. Herein, the most frequently expressed virulence gene found was *gad* (73%), a highly specific gene to the *E. coli* species commonly used as a prescreening

marker for pathogenic groups *(27)*. This gene allows isolates to tolerate highly acidic conditions when enduring routes through the gastrointestinal tract by the decarboxylation of glutamate where cells can repel incoming protons *(9)*. The virulence gene *ipfA* was found in 64% of the isolates, whereas the gene *iss* was found in 23%. These two genes, generally found in EHEC or EPEC pathotypes, encode for fimbriae that help cellular adhesion and increase serum survival, respectively *(69)*. Virulence gene *terC* was found in more than half of the isolates (51%) and represents the key protein needed for tellurite resistance *(64)*. Of the total samples, only seven *E. coli* isolates did not show any virulence genes.

Overall *E. coli* has a great degree of assortment among virulence genes with rapid adaptation resulting in capricious pathogenicity *(39)*. Understanding virulence gene profiles may contribute to knowledge on emergent *E. coli* pathotypes.

CONCLUSIONS

This study highlights the importance of agricultural water sources and seasonality when assessing produce safety. Cisterns and wells, as observed in this study, can harbor high concentrations of *E. coli* despite these sources not being exposed to the surrounding environment. Produce growers should consider this potential when selecting a water source to irrigate fresh produce. Furthermore, our research shows the need for a more expansive assessment of agricultural water within Kansas and Missouri to understand the true diversity of *E. coli* in the waters within those states and the importance of characterizing *E. coli* serotypes to help guide agricultural water assessments as proposed in the new agricultural water rule. It can also be concluded that there are numerous direct and indirect routes that *E. coli* can use to contaminate both surface and groundwater sources. Produce growers may be able to use the information collected in this study to make informed decisions on how to better prevent *E. coli* contamination routes. As shown in this study, *E. coli* can be carried to a water source without an animal host having access to the water source.

REFERENCES

1. Acosta-Dibarrat, J., E. Enriquez-Gómez, M. Talavera-Rojas, E. Soriano-Vargas, A. Navarro, and R. Morales-Espinosa. 2021. Characterization of commensal *Escherichia coli* isolates from slaughtered sheep in Mexico. *J. Infect. Dev. Ctry.* 15:1755–1760.
2. Afgan, E., D. Baker, B. Batut, M. van den Beek, D. Bouvier, M. Čech, J. Chilton, D. Clements, N. Coraor, B. A. Grüning, A. Guerler, J. Hillman-Jackson, S. Hiltemann, V. Jalili, H. Rasche, N. Soranzo, J. Goecks, J. Taylor, A. Nekrutenko, and D. Blankenberg. 2018. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.* 46:W537–W544.
3. Alcock B. P., A. R. Raphenya, T. T. Y. Lau, K. K. Tsang, M. Bouchard, A. Edalatmand, W. Huynh, A. V. Nguyen, A. A. Cheng, S. Liu, S. Y. Min, A. Miroshnichenko, H. K. Tran, R. E. Werfalli, J. A. Nasir, M. Oloni, D. J. Speicher, A. Florescu, B. Singh, M. Faltyn, A. Hernandez-Koutoucheva, A. N. Sharma, E. Bordeleau, A. C. Pawlowski, H. L. Zubyk, D. Dooley, E. Griffiths, F. Maguire, G. L. Winsor, R. G. Beiko, F. S. L. Brinkman, W. W. L. Hsiao, G. V. Domselaar, and A. G. McArthur. 2020. CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res.* 48:D517–D525.
4. Andrade, L., J. O'Dwyer, E. O'Neill, and P. Hynds. 2018. Surface water flooding, groundwater contamination, and enteric disease in developed countries: a scoping review of connections and consequences. *Environ. Pollut.* 236:540–549.
5. Awosile, B., J. Reyes-Velez, Y. Cuesta-Astroz, J. C. Rodríguez-Lecompte, M. E. Saab, L. C. Heider, G. Keefe, J. Sánchez, and J. T. McClure. 2020. Short communication: whole-genome sequence analysis of 4 fecal blaCMY-2-producing *Escherichia coli* isolates from Holstein dairy calves. *J. Dairy Sci.* 103:877–883.
6. Barlow, R. S., K. S. Gobius, and P. M. Desmarchelier. 2006. Shiga toxin-producing *Escherichia coli* in ground beef and lamb cuts: results of a one-year study. *Int. J. Food Microbiol.* 111:1–5.
7. Bentancor, A., M. V. Rumi, C. Carbonari, E. Gerhardt, M. Larzábal, D. A. Vilte, V. Pistone-Creydt, I. Chinen, C. Ibarra, A. Cataldi, and E. C. Mercado. 2012. Profile of Shiga toxin-producing *Escherichia coli* strains isolated from dogs and cats and genetic relationships with isolates from cattle, meat and humans. *Vet. Microbiol.* 156:336–342.
8. Berenger, B. M., L. Chui, C. Ferrato, T. Lloyd, V. Li, and D. R. Pillai. 2022. Performance of four commercial real-time PCR assays for the detection of bacterial enteric pathogens in clinical samples. *Int. J. Infect. Dis.*114:195–201.
9. Bergholz, T. M., C. L. Tarr, L. M. Christensen, D. J. Betting, and T. S. Whittam. 2007. Recent gene conversions between duplicated glutamate decarboxylase genes (gadA and gadB) in pathogenic *Escherichia coli*. *Mol. Biol. Evol.* 24:2323–2333.
10. Bertoldi, B., S. Richardson, R. G. Schneider, P. Kurdmongkoltham, and K. R. Schneider. 2017. Preventing foodborne illness: *E. coli* the big six. Available at: https://www.nifa.usda.gov/sites/default/files/resource/Preventing-Foodborne-Illness-E-coli-the-big-six.pdf. Accessed 15 January 2023.
11. Bettelheim, K. A. 1978. The sources of 'OH' serotypes of *Escherichia coli*. *J. Hyg.* 80:83–113.
12. Bettelheim, K. A., A. Kuzevski, R. A. Gilbert, D. O. Krause, and C. S. McSweeney. 2005. The diversity of *Escherichia coli* serotypes and biotypes in cattle faeces. *J. Appl. Microbiol.* 98:699–709.
13. Bosilevac, J. M., and M. Koohmaraie. 2011. Prevalence and characterization of non-O157 Shiga toxin-producing *Escherichia coli* isolates from commercial ground beef in the United States. *Appl. Environ. Microbiol.* 77:2103–2112.
14. Brown, E., U. Dessai, S. McGarry, and P. Gerner-Smidt. 2019. Use of whole-genome sequencing for food safety and public health in the United States. *Foodborne Pathog. Dis.* 16:441–450.
15. Byappanahalli, M., M. Fowler, D. Shively, and R. Whitman. 2003. Ubiquity and persistence of *Escherichia coli* in a midwestern coastal stream. *Appl. Environ. Microbiol.* 69:4549–4555.
16. Centers for Disease Control and Prevention. 2019. *E. coli* infection. Available at: https://www.cdc.gov/healthypets/diseases/ecoli.html. Accessed 11 October 2021.
17. Center for Disease Control and Prevention. 2019. Outbreak of *E. coli* linked to romaine lettuce. Available at: https://www.cdc.gov/ecoli/2018/o157h7-11-18/index.html. Accessed 11 October 2021.
18. Centers for Disease Control and Prevention. 2022. Cistern and other rain catchment systems. Available at: https://www.cdc.gov/healthywater/drinking/ private/rainwater-collection.html. Accessed 2 November 2022.

19. Díaz-Jiménez, D., I. García-Meniño, J. Fernández, V. García, and A. Mora. 2020. Chicken and turkey meat: consumer exposure to multidrug-resistant *Enterobacteriaceae* including mcr-carriers, uropathogenic *E. coli* and high-risk lineages such as ST131. *Int. J. Food Microbiol.* 331:108750.

20. Domesle, K. J., S. R. Young, and B. Ge. 2021. Rapid screening for *Salmonella* in raw pet food by loop-mediated isothermal amplification. *J. Food Prot.* 84:399–407.

21. Enriquez-Gómez, E., M. Talavera-Rojas, E. Soriano-Vargas, A. Navarro-Ocaña, V. Vega-Sánchez, S. A.-M. de Oca, and J. Acosta-Dibarrat. 2017. Serotypes, virulence genes profiles and antimicrobial resistance patterns of *Escherichia coli* recovered from feces of healthy lambs in Mexico. *Small Rumin. Res.* 153:41–47.

22. Fischer, J. R., T. Zhao, M. P. Doyle, M. R. Goldberg, C. A. Brown, C. T. Sewell, D. M. Kavanaugh, and C. D. Bauman. 2001. Experimental and field studies of *Escherichia coli* O157:H7 in white-tailed deer. *Appl. Environ. Microbiol.* 67:1218–1224.

23. Flament-Simon, S.-C., M. de Toro, V. García, J. E. Blanco, M. Blanco, M. P. Alonso, A. Goicoa, J. Díaz-González, M.-H. Nicolas-Chanoine, and J. Blanco. 2020. Molecular characteristics of extraintestinal pathogenic *E. coli* (ExPEC), uropathogenic *E. coli* (UPEC), and multidrug resistant *E. coli* isolated from healthy dogs in Spain. Whole genome sequencing of canine ST372 isolates and comparison with human isolates causing extraintestinal infections. *Microorganisms* 8:1712.

24. Franz, E., M. M. Klerks, O. J. De Vos, A. J. Termorshuizen, and A. H. C. van Bruggen. 2007. Prevalence of Shiga toxin-producing *Escherichia coli stx1, stx2, eaeA,* and *rfbE* genes and survival of *E. coli* O157:H7 in manure from organic and low-input conventional dairy farms. *Appl. Environ. Microbiol.* 73:2180–2190.

25. Fukuyama, M., R. Yokoyama, S. Sakata, K. Furuhata, K. Oonaka, M. Hara, Y. Satoh, K. Tabuchi, T. Itoh, A. Kai, and M. Matsuda. 1999. Study on the verotoxin-producing *Escherichia coli. J. Jpn. Assoc. Infect. Dis.* 73:1140–1144.

26. Gannon, V. P., S. D'Souza, T. Graham, R. K. King, K. Rahn, and S. Read. 1997. Use of the flagellar H7 gene as a target in multiplex PCR assays and improved specificity in identification of enterohemorrhagic *Escherichia coli* strains. *J. Clin. Microbiol.* 35:656–662.

27. Grant, M. A., S. D. Weagant, and P. Feng. 2001. Glutamate decarboxylase genes as a prescreening marker for detection of pathogenic *Escherichia coli* groups. *Appl. Environ. Microbiol.* 67:3110–3114.

28. Haley, O. C., L. Nwadike, Y. Zhao, J. M. Maher, S. E. Gragg, V. Trinetta, and M. Bhullar. 2022. Comparative assessment of the microbial quality of agricultural water on Kansas and Missouri fresh produce farms. *Food Prot. Trends* 42:186–193.

29. Hamelin, K., G. Bruant, A. El-Shaarawi, S. Hill, T. A. Edge, S. Bekal, J. M. Fairbrother, J. Harel, C. Maynard, L. Masson, and R. Brousseau. 2006. A virulence and antimicrobial resistance DNA microarray detects a high frequency of virulence genes in *Escherichia coli* isolates from Greater Lakes recreational waters. *Appl. Environ. Microbiol.* 72:4200–4206.

30. Hu, Y., Q. Zhang, and J. C. Meitzler. 1999. Rapid and sensitive detection of *Escherichia coli* O157:H7 in bovine faeces by a multiplex PCR. *J. Appl. Microbiol.* 87:867–876.

31. Hunter, P. R. 2003. Drinking water and diarrhoeal disease due to *Escherichia coli. J. Water Health* 1:65–72.

32. Hussein, H. S. 2007. Prevalence and pathogenicity of Shiga toxin-producing *Escherichia coli* in beef cattle and their products1,2. *J. Anim. Sci.* 85:E63–E72.

33. Johnson, J. R., P. Delavari, A. L. Stell, T. S. Whittam, U. Carlino, and T. A. Russo. 2001. Molecular comparison of extraintestinal *Escherichia coli* isolates of the same electrophoretic lineages from humans and domestic animals. *J. Infect. Dis.* 183:154–159.

34. Johnson, R. P., R. C. Clarke, J. B. Wilson, S. C. Read, K. Rahn, S. A. Renwick, K. A. Sandhu, D. Alves, M. A. Karmali, H. Lior, S. A. McEwen, J. S. Spika, and C. L. Gyles. 1996. Growing concerns and recent outbreaks involving non-O157:H7 serotypes of verotoxigenic *Escherichia coli. J. Food Prot.* 59:1112–1122.

35. Karama, M., A. O. Mainga, B. T. Cenci-Goga, M. Malahlela, S. El-Ashram, and A. Kalake. 2019. Molecular profiling and antimicrobial resistance of Shiga toxin-producing *Escherichia coli* O26, O45, O103, O121, O145 and O157 isolates from cattle on cow-calf operations in South Africa. *Sci. Rep.* 9:11930.

36. Kostyla, C., R. Bain, R. Cronk, and J. Bartram. 2015. Seasonal variation of fecal contamination in drinking water sources in developing countries: a systematic review. *Sci. Total Environ.* 514:333–343.

37. Krause, G., S. Zimmermann, and L. Beutin. 2005. Investigation of domestic animals and pets as a reservoir for intimin- (eae) gene positive *Escherichia coli* types. *Vet. Microbiol.* 106:87–95.

38. Kudva, I. T., P. G. Hatfield, and C. J. Hovde. 1997. Characterization of *Escherichia coli* O157:H7 and other Shiga toxin-producing *E. coli* serotypes isolated from sheep. *J. Clin. Microbiol.* 35:892–899.

39. Kuhnert, P., P. Boerlin, and J. Frey. 2000. Target genes for virulence assessment of *Escherichia coli* isolates from water, food and the environment. *FEMS Microbiol. Rev.* 24:107–117.

40. Leotta, G. A., N. Deza, J. Origlia, C. Toma, I. Chinen, E. Miliwebsky, S. Iyoda, S. Sosa-Estani, and M. Rivas. 2006. Detection and characterization of Shiga toxin-producing *Escherichia coli* in captive non-domestic mammals. *Vet. Microbiol.* 118:151–157.

41. Lévesque, B., D. Pereg, E. Watkinson, J. S. Maguire, L. Bissonnette, S. Gingras, P. Rouja, M. G. Bergeron, and E. Dewailly. 2008. Assessment of microbiological quality of drinking water from household tanks in Bermuda. *Can. J. Microbiol.* 54:495–500.

42. Liao, C.-Y., B. Balasubramanian, J. J. Peng, S. R. Tao, W. C. Liu, and Y. Ma. 2021. Antimicrobial resistance of *Escherichia coli* from aquaculture farms and their environment in Zhanjiang, China. *Front. Vet. Sci.* 8:806653.

43. Lupo, A., S. Coyne, and T. U. Berendonk. 2012. Origin and evolution of antibiotic resistance: the common mechanisms of emergence and spread in water bodies. *Front. Microbiol.* 3:18.

44. Malahlela, M. N., B. T. Cenci-Goga, M. C. Marufu, T. Y. Fonkui, L. Grispoldi, E. Etter, A. Kalake, and M. Karama. 2022. Occurrence, serotypes and virulence characteristics of Shiga-toxin-producing *Escherichia coli* isolates from goats on communal rangeland in South Africa. *Toxins* 14:353.

45. Maloo, A., A. B. Fulke, N. Mulani, S. Sukumaran, and A. Ram. 2017. Pathogenic multiple antimicrobial resistant *Escherichia coli* serotypes in recreational waters of Mumbai, India: a potential public health risk. *Environ. Sci. Pollut. Res.* 24:11504–11517.

46. Manges, A. R. 2016. *Escherichia coli* and urinary tract infections: the role of poultry-meat. *Clin. Microbiol. Infect.* 22:122–129.

47. Marder, E. P., K. N. Garman, L. A. Ingram, and J. R. Dunn. 2014. Multistate outbreak of *Escherichia coli* O157:H7 associated with bagged salad. *Foodborne Pathog. Dis.* 11:593–595.

48. Massé, D., N. Saady, and Y. Gilbert. 2014. Potential of biological processes to eliminate antibiotics in livestock manure: an overview. *Animals* 4:146–163.

49. Melton-Celsa, A. R. 2014. Shiga toxin (*stx*) classification, structure, and function. *Microbiol. Spectr.* 2:2.4.06.

50. Nielsen, E. M., M. N. Skov, J. J. Madsen, J. Lodal, J. B. Jespersen, and D. L. Baggesen. 2004. Verocytotoxin-producing *Escherichia coli* in wild birds and rodents in close proximity to farms. *Appl. Environ. Microbiol.* 70:6944–6947.

51. Oliver, D. M., and T. Page. 2016. Effects of seasonal meteorological variables on *E. coli* persistence in livestock faeces and implications for environmental and human health. *Sci. Rep.* 6:37101.

52. Osman, K. M., A. M. Mustafa, M. Elhariri, and G. S. AbdElhamed. 2013. The distribution of *Escherichia coli* serovars, virulence genes, gene association and combinations and virulence genes encoding serotypes in pathogenic *E. coli* recovered from diarrhoeic calves, sheep and goat: the distribution of *E. coli* serovars. *Transbound. Emerg. Dis.* 60:69–78.

53. Parma, A. E. 2000. Virulence genotypes and serotypes of verotoxigenic *Escherichia coli* isolated from cattle and foods in Argentina. *Eur. J. Epidemiol.* 16:757–762.

54. Ramteke, P. W., and S. Tewari. 2007. Serogroups of *Escherichia coli* from drinking water. *Environ. Monit. Assess.* 130:215–220.

55. Rehman, M. A., H. Rempel, C. D. Carrillo, K. Ziebell, K. Allen, A. R. Manges, E. Topp, and M. S. Diarra. 2022. Virulence genotype and phenotype of multiple antimicrobial-resistant *Escherichia coli* isolates from broilers assessed from a "One-Health" perspective. *J. Food Prot.* 85:336–354.

56. Rey, J. 2003. Serotypes, phage types and virulence genes of Shiga-producing *Escherichia coli* isolated from sheep in Spain. *Vet. Microbiol.* 94:47–56.

57. Reyes-Rodriguez, N. E., J. Barba-León, A. Navarro-Ocaña, V. Vega-Sanchez, F. R. Gómez De Anda, J. M. Talavera-Gonzalez, and M. Talavera-Rojas. 2020. Serotypes and stx2 subtyping of Shiga toxin producing *Escherichia coli* isolates from cattle carcasses and feces. *Rev. Mex. Cienc. Pecu.* 11:1030–1044.

58. Rodrigues, C., A. L. B. R. da Silva, and L. L. Dunn. 2019. factors impacting the prevalence of foodborne pathogens in agricultural water sources in the southeastern United States. Water 12:51.

59. Sasakova, N., G. Gregova, D. Takacova, J. Mojzisova, I. Papajova, J. Venglovsky, T. Szaboova, and S. Kovacova. 2018. Pollution of surface and ground water by sources related to agricultural activities. *Front. Sustain. Food Syst.* 2:42.

60. Segura, A., P. Auffret, C. Klopp, Y. Bertin, and E. Forano. 2017. Draft genome sequence and characterization of commensal *Escherichia coli* strain BG1 isolated from bovine gastro-intestinal tract. *Stand. Genom. Sci.* 12:61.

61. Stein, R. A., and D. E. Katz. 2017. *Escherichia coli*, cattle and the propagation of disease. *FEMS Microbiol. Lett.* 364:fnx050.

62. Sun, H., Y. Wan, P. Du, D. Liu, R. Li, P. Zhang, Y. Wu, S. Fanning, Y. Wang, and L. Bai. 2021. Investigation of tigecycline resistant *Escherichia coli* from raw meat reveals potential transmission among food-producing animals. *Food Control* 121:107633.

63. Tanabe, R. H. S., M. A. Vieira, N. A. B. Mariano, R. C. B. Dias, R. V. da Silva, C. M. Castro, L. F. dos Santos, C. H. Camargo, R. S. Yamatogi, V. L. M. Rall, and R. T. Hernandes. 2019. Identification and characterization of atypical enteropathogenic and Shiga toxin-producing *Escherichia coli* isolated from ground beef and poultry breast purchased in Botucatu, Brazil. *Braz. J. Microbiol.* 50:1099–1103.

64. Turkovicova, L., R. Smidak, G. Jung, J. Turna, G. Lubec, and J. Aradska. 2016. Proteomic analysis of the TerC interactome: novel links to tellurite resistance and pathogenicity. J. *Proteom.* 136:167–173.

65. U.S. Environmental Protection Agency. 2014. Method 1603: *Escherichia coli (E. coli)* in water by membrane filtration using modified membrane-thermotolerant *Escherichia coli* agar (modified mTEC). Available at: https://www.epa.gov/sites/default/files/2015-08/documents/method_1603_2009.pdf. Accessed 13 January 2021.

66. U.S. Food and Drug Administration. 2019. *Escherichia coli (E. coli)*. Available at: https://www.fda.gov/food/foodborne-pathogens/escherichia-coli-e-coli. Accessed 11 October 2021.

67. U.S. Food and Drug Administration. 2022. FSMA proposed rule of agricultural water. Available at: https://www.fda.gov/food/food-safety-modernization-act-fsma/fsma-proposed-rule-agricultural-water. Accessed 11 October 2021.

68. U.S. Food and Drug Administration. 2022. Outbreak investigation of *E. coli* O157:H7-spinach. Available at: https://www.fda.gov/food/outbreaks-foodborne-illness/outbreak-investigation-e-coli-o157h7-spinach-november-2021. Accessed 11 October 2021.

69. Valat, C., F. Auvray, K. Forest, V. Métayer, E. Gay, C. Peytavin de Garam, J. Y. Madec, and M. Haenni. 2012. Phylogenetic grouping and virulence potential of extended-spectrum-β-lactamase-producing *Escherichia coli* strains in cattle. *Appl. Environ. Microbiol.* 78:4677–4682.

70. van Elsas, J. D., A. V. Semenov, R. Costa, and J. T. Trevors. 2011. Survival of *Escherichia coli* in the environment: fundamental and public health aspects. *ISME J.* 5:173–183.

71. Vettorato, M. P., A. F. P. de Castro, M. C. Cergole-Novella, F. L. L. Camargo, K. Irino, and B. E. C. Guth. 2009. Shiga toxin-producing *Escherichia coli* and atypical enteropathogenic *Escherichia coli* strains isolated from healthy sheep of different populations in São Paulo, Brazil. *Lett. Appl. Microbiol.* 49:53–59.

72. Vingino, A., M. C. Roberts, M. Wainstein, J. West, S. A. Norman, D. Lambourn, J. Lahti, R. Ruiz, M. D'Angeli, S. J. Weissman, and P. Rabinowitz. 2021. Surveillance for antibiotic-resistant *E. coli* in Salish Sea ecosystem. *Antibiotics* 10:1201.

73. Vivar, M., and M. Fuentes. 2016. Using solar disinfected water: on the bacterial regrowth over 1-week of water usage including direct intake after sun exposure and long-term dark storage. *Sol. Energy* 131:138–148.

74. Vu-Khac, H., E. Holoda, E. Pilipcinec, M. Blanco, J. E. Blanco, G. Dahbi, A. Mora, C. López, E. A. González, and J. Blanco. 2007. Serotypes, virulence genes, intimin types and PFGE profiles of *Escherichia coli* isolated from piglets with diarrhea in Slovakia. *Vet. J.* 174:176–187.

75. Walker, D. I., J. McQuillan, M. Taiwo, R. Parks, C. A. Stenton, H. Morgan, M. C. Mowlem, and D. N. Lees. 2017. A highly specific *Escherichia coli* qPCR and its comparison with existing methods for environmental waters. *Water Res.* 126:101–110.

76. Ward, M., R. Dhingra, J. V. Remais, H. H. Chang, L. M. Johnston, L. A. Jaykus, and J. Leon. 2015. Associations between weather and microbial load on fresh produce prior to harvest. *J. Food Prot*. 78:849–854.

77. Werth, B. J. 2022. Overview of antibiotics. Merck Manual. Available at: https://www.merckmanuals.com/home/infections/antibiotics/overview-of-antibiotics. Accessed 29 July 2022.

78. World Health Organization. 2018. *E. coli.* Available at: https://www.who.int/news-room/fact-sheets/detail/e-coli. Accessed 11 October 2021.

79. Younis, G., A. Awad, and N. Mohamed. 2017. Phenotypic and genotypic characterization of antimicrobial susceptibility of avian pathogenic *Escherichia coli* isolated from broiler chickens. *Vet. World* 10:1167–1172.

80. Zhang, S., Y. Yin, M. B. Jones, Z. Zhang, B. L. Deatherage Kaiser, B. A. Dinsmore, C. Fitzgerald, P. I. Fields, and X. Deng. 2015. *Salmonella* serotype determination utilizing high-throughput genome sequencing data. *J. Clin. Microbiol.* 53:1685–1692.

81. Zhi, S., G. S. Banting, N. J. Ruecker, and N. F. Neumann. 2017. Stress resistance in naturalized waste water *E. coli* strains. *J. Environ. Eng. Sci.* 12:42–50.