



# Humans and Hoofed Livestock Are the Main Sources of Fecal Contamination of Rivers Used for Crop Irrigation: A Microbial Source Tracking Approach

Constanza Díaz-Gavidia<sup>1,2</sup>, Carla Barría<sup>1,2</sup>, Daniel L. Weller<sup>3,4</sup>, Marília Salgado-Caxito<sup>2,5</sup>, Erika M. Estrada<sup>6</sup>, Aníbal Araya<sup>2,7</sup>, Leonardo Vera<sup>8</sup>, Woutrina Smith<sup>9</sup>, Minji Kim<sup>10</sup>, Andrea I. Moreno-Switt<sup>2,5</sup>, Jorge Olivares-Pacheco<sup>2,7</sup> and Aiko D. Adell<sup>1,2\*</sup>

## OPEN ACCESS

### Edited by:

Gerson Nakazato,  
State University of Londrina, Brazil

### Reviewed by:

Hamed Mirjalali,  
Shahid Beheshti University of Medical  
Sciences, Iran  
Tatiana Amabile De Campos,  
University of Brasília,  
Brazil

### \*Correspondence:

Aiko D. Adell  
aiko.adell@unab.cl

### Specialty section:

This article was submitted to  
Antimicrobials, Resistance and  
Chemotherapy,  
a section of the journal  
Frontiers in Microbiology

Received: 31 August 2021

Accepted: 19 May 2022

Published: 30 June 2022

### Citation:

Díaz-Gavidia C, Barría C, Weller DL,  
Salgado-Caxito M, Estrada EM,  
Araya A, Vera L, Smith W, Kim M,  
Moreno-Switt AI,  
Olivares-Pacheco J and  
Adell AD (2022) Humans and Hoofed  
Livestock Are the Main Sources of  
Fecal Contamination of Rivers Used  
for Crop Irrigation: A Microbial Source  
Tracking Approach.  
Front. Microbiol. 13:768527.  
doi: 10.3389/fmicb.2022.768527

<sup>1</sup>Escuela de Medicina Veterinaria, Facultad de Ciencias de la Vida, Universidad Andres Bello, Santiago, Chile,

<sup>2</sup>Millennium Initiative for Collaborative Research on Bacterial Resistance (MICROB-R), Santiago, Chile, <sup>3</sup>Department of Environmental and Forest Biology, College of Environmental Science and Forestry, State University of New York, Syracuse, NY, United States, <sup>4</sup>Department of Biostatistics and Computational Biology, University of Rochester Medical Center, Rochester, NY, United States, <sup>5</sup>Escuela de Medicina Veterinaria, Facultad de Agronomía e Ingeniería Forestal, Facultad de Ciencias Biológicas y Facultad de Medicina, Pontificia Universidad Católica de Chile, Santiago, Chile,

<sup>6</sup>Department of Food Science and Technology, Eastern Shore Agricultural Research and Extension Center, Virginia Tech, Painter, Virginia, <sup>7</sup>Grupo de Resistencia Antimicrobiana en Bacterias Patógenas y Ambientales (GRABPA), Instituto de Biología, Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile, <sup>8</sup>Escuela Ingeniería Ambiental, Facultad de Ciencias de la Vida, Universidad Andres Bello, Santiago, Chile, <sup>9</sup>One Health Institute, School of Veterinary Medicine, University of California, Davis, Davis, CA, United States, <sup>10</sup>Department of Civil and Environmental Engineering, University of California, Davis, Davis, CA, United States

Freshwater bodies receive waste, feces, and fecal microorganisms from agricultural, urban, and natural activities. In this study, the probable sources of fecal contamination were determined. Also, antibiotic resistant bacteria (ARB) were detected in the two main rivers of central Chile. Surface water samples were collected from 12 sampling sites in the Maipo ( $n=8$ ) and Maule Rivers ( $n=4$ ) every 3 months, from August 2017 until April 2019. To determine the fecal contamination level, fecal coliforms were quantified using the most probable number (MPN) method and the source of fecal contamination was determined by Microbial Source Tracking (MST) using the *Cryptosporidium* and *Giardia* genotyping method. Separately, to determine if antimicrobial resistance bacteria (AMB) were present in the rivers, *Escherichia coli* and environmental bacteria were isolated, and the antibiotic susceptibility profile was determined. Fecal coliform levels in the Maule and Maipo Rivers ranged between 1 and 130 MPN/100-ml, and 2 and 30,000 MPN/100-ml, respectively. Based on the MST results using *Cryptosporidium* and *Giardia* host-specific species, human, cattle, birds, and/or dogs hosts were the probable sources of fecal contamination in both rivers, with human and cattle host-specific species being more frequently detected. Conditional tree analysis indicated that coliform levels were significantly associated with the river system (Maipo versus Maule), land use, and season. Fecal coliform levels were significantly ( $p < 0.006$ ) higher at urban and agricultural sites than at sites immediately downstream of treatment centers, livestock areas, or natural areas.

Three out of eight (37.5%) *E. coli* isolates presented a multidrug-resistance (MDR) phenotype. Similarly, 6.6% (117/1768) and 5.1% (44/863) of environmental isolates, in Maipo and Maule River showed and MDR phenotype. Efforts to reduce fecal discharge into these rivers should thus focus on agriculture and urban land uses as these areas were contributing the most and more frequently to fecal contamination into the rivers, while human and cattle fecal discharges were identified as the most likely source of this fecal contamination by the MST approach. This information can be used to design better mitigation strategies, thereby reducing the burden of waterborne diseases and AMR in Central Chile.

**Keywords:** microbial source tracking, water quality, waterborne pathogens, *Cryptosporidium*, *Giardia*, fecal coliforms, antimicrobial resistance

## INTRODUCTION

Fecal pollution in water is associated with negative health impacts, economic losses (including increased health costs), and pathogen contamination of raw foods (Boelee et al., 2019). The World Health Organization reports that diarrheal disease is the second leading cause of death in children under 5-year old, with approximately 1.7 billion cases of childhood diarrheal disease and 525,000 deaths of children per year (WHO, World Health Organization, 2017a). The diarrheal disease mostly results from contaminated food and water sources (WHO, World Health Organization, 2017b). This is worrisome considering that 780 million individuals worldwide lack access to improved drinking-water and 2.5 billion lack improved sanitation (WHO and UNCF, World Health Organization and United Nations Children's Fund, 2017).

The quality of many water sources has been greatly affected by fecal pathogens from anthropogenic activity such as urban and animal protein production (Pandey et al., 2014). In general, fecal contamination of surface water could be mainly attributed to human activities, such as poor management of fecal waste on farms leading to contaminated runoff into water (Fenton et al., 2021); the excessive use of fertilizers in agriculture (Khan et al., 2018), and the discharge of water treatment plants, since many microorganisms can colonize infrastructure and resist treatment (Newton and McClary, 2019). Some studies have investigated the effect of human activities on diverse pathogen distribution through the water; for example, in a 19-year meta-analysis review, the distribution of some parasites such as *Cryptosporidium* spp. and *Giardia* spp. can be enhanced by climate conditions, water availability, increasing illegal usage of raw wastewater, night soil, and manure amendment for crop soils (Javanmard et al., 2020). For example, the authors indicated that the transmission of parasites is probably limited to rural areas in industrialized districts or those cities where tourism is the main business, while in those regions where traditional animal husbandry is the main business, rain increases the surface water that carries (oo)cysts of parasites dispersed by herds to downstream farms traditionally irrigated with surface waters (Javanmard et al., 2020). Lewis et al. (2005) observed high levels of inputs of fecal coliforms (over  $10^8$  fecal coliforms/hectare) in discharges from areas with large cattle herds during precipitation events.

Similarly, Wilkes et al. (2011) reported that the likelihood of detecting *E. coli* O157:H7 was associated with upstream livestock pasture density, with 20% of the detections located where cattle had access to the waterway. Land use is a classification that provides information on the type and intensity of human activities occurring in a given geographical area (LaGro, 2005). Among the land-use classification systems, the most general classification and commonly used for regions and other large-scale applications consist of broad land use categories, such as “agriculture” or “urban and built-up” land use (LaGro, 2005). Therefore, different land uses drive different amounts of fecal material to rivers depending on factors such as human activities, livestock activities, and agricultural activities (Wen et al., 2017; Camara et al., 2019; Weller et al., 2020a) either by direct contact or in indirect ways such as dragged by rain (Ding et al., 2015).

In addition, several studies have reported antimicrobial-resistant *Enterobacteriaceae* in surface water (van Hoek et al., 2015; Guyomard-Rabenirina et al., 2017; Azuma and Hayashi, 2021). Therefore, it is important to consider that bacteria that reach water sources through feces may be carriers of antibiotic resistance genes (ARGs), abetting the spread of antimicrobial resistance (AMR; Pilmis et al., 2020). For instance, multi-drug resistance (MDR) *Salmonella Enteritidis* and Typhimurium, were detected in rivers and irrigation canals, with the higher detection rates observed in rural areas compared to urban and peri-urban areas (Martínez et al., 2017); therefore, land use is an important factor to be considered when evaluating the health risk to human that is in contact with surface water sources. These fecal and antibiotic resistance bacteria (ARB) discharge adversely affect water quality in rivers which is highly relevant to human health, quality of water used for crop irrigation, recreational activities, animal health (both livestock and wildlife) and the environment (Sbodio et al., 2013).

Many of the programs monitoring water quality only focus on measuring Fecal Indicator Bacteria (FIB; Holcomb and Stewart, 2020). These FIB are used because they are commonly found in human and animal feces and are known commensal organisms. Therefore, their presence in water may indicate the presence of pathogens (Korajkic et al., 2018). Among the FIB, the coliform group, consisting of the bacteria *Enterobacter*, *Klebsiella*, *E. coli*, and *Citrobacter*, are used as indicators of fecal pollution of surface water worldwide (Holcomb and

Stewart, 2020; Goshu et al., 2021). The presence of high levels of fecal coliforms is a frequently used indicator of the presence of fecally-associated bacteria, including enteric pathogens. These fecal pathogens, specifically protozoa and bacteria, can be transmitted and dispersed in many aquatic environments such as oceans, rivers, creeks, and lakes (Lusk et al., 2017; Devane et al., 2018; Shapiro et al., 2018). Despite this, the major limitation is that they do not provide information regarding the fecal source of contamination as any animal, including humans, can excrete this FIB through their feces (Field and Samadpour, 2007). Research related to determining the level of microbial contamination caused by inland surface waters has not included fecal source contamination. To overcome this limitation of the FIB, Microbial Source Tracking (MST), an approach to determine the contributions of specific hosts of fecal contamination to water sources, has been increasingly used (Raza et al., 2021). Among the MST tools reported in the scientific literature are the protozoan parasites *Cryptosporidium* spp. and *Giardia* spp., as both provide knowledge of host specificity (Ryu et al., 2011; Prystajecy et al., 2014; Xiao et al., 2018). For instance, the genus *Cryptosporidium* is composed of genotypes that are host specific and genotypes that infect various hosts (Xiao et al., 2000), a situation that also occurs with *Giardia* genotypes and assemblages (Feng and Xiao, 2011). *Giardia* is divided into six species based on morphological characteristics, among which *Giardia duodenalis* (synonyms *Giardia lamblia* and *Giardia intestinalis*) infects a variety of mammals, including humans (Cacciò et al., 2005) and will be referred to as *G. duodenalis* throughout this manuscript. Besides, *G. duodenalis* have genetic groupings (henceforth assemblages) of *G. duodenalis* and species, each with specific hosts. Therefore, identifying the genotype of *Cryptosporidium* spp. and assemblage for *Giardia duodenalis* can provide information on the host source (animal or human) of origin of the fecal contamination and its zoonotic potential or if they are infective to productive animals such as cows.

Chile provides an ideal condition to study the water quality of rivers because of the geography it presents. In addition, it is a narrow country, allowing an entire river that crossed different land use zones to be sampled on the same day. Water quality is particularly important for Chile as the rivers are the main source of irrigation and drinking water source. Therefore, the objective of this study was to determine the fecal contamination levels using fecal coliforms as indicators, identify the genotypes/assemblages of *Cryptosporidium* and *Giardia*, respectively, to determine the probable source of fecal contamination and detect the presence of antibiotic-resistant bacteria (ARB) in two main rivers of central Chile based on land use (urban, agriculture, livestock, and natural). Finally, we evaluate the association between the presence of fecal coliforms, *Cryptosporidium* spp., *Giardia duodenalis*, and ARB with environmental factors. This study provides information regarding the land uses and sources that contribute the most to rivers' biological contamination with fecal material and ARB, consequently providing information to create strategies to manage and reduce loads of this fecal contamination to the rivers.

## MATERIALS AND METHODS

### Study Sites

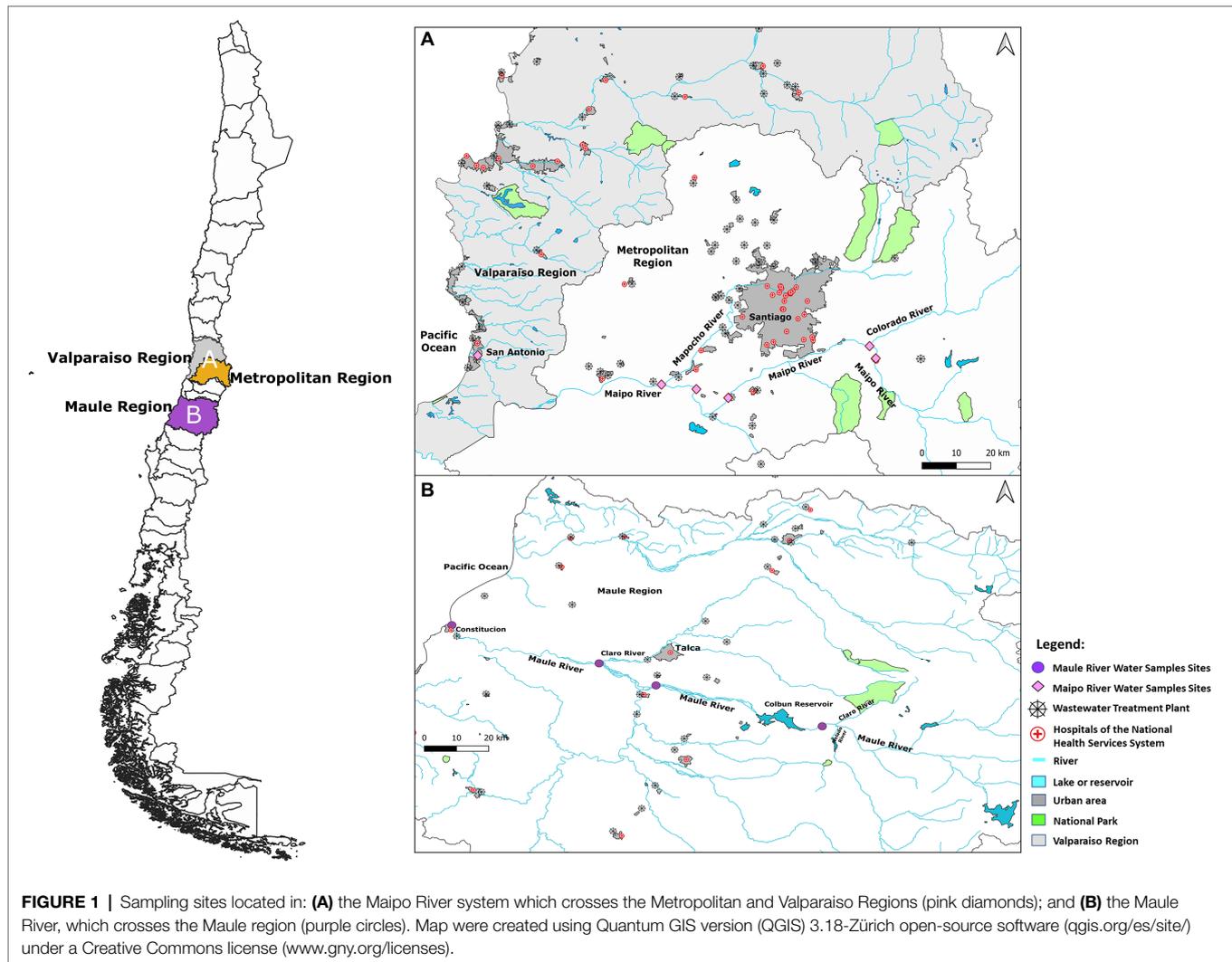
Sampling sites corresponded to sites representing different land-use areas in the Maipo and Mapocho Rivers, both located in the central macrozone of Chile. The Maipo crosses the Metropolitan and Valparaíso Regions. The Maule crosses the Maule Region of the country. The Maipo is the primary source of drinking water for the Metropolitan region and supplies approximately 70% of the drinking water and 90% of the irrigation demands (DGA, Dirección General de Aguas, 2004a). Similarly, the Maule provides irrigation to approximately 118,263 hectares and drinking water to approximately 99,6% of the urban population (DGA, Dirección General de Aguas, 2004b). Fecal coliforms levels in the Maipo and Maule Rivers have exceeded levels permitted by the Chilean water quality standard (INN, Instituto Nacional de Normalización, 1978). For example, fecal coliforms in the Maipo River ranged between  $<2$  and  $8.50 \times 10^6$  fecal coliforms/100 ml in past reports (DGA, Dirección General de Aguas, 2004a), while the Maule River between 540 and  $1.1 \times 10^4$  fecal coliforms/100 ml (DGA, Dirección General de Aguas, 2004b).

Land use throughout each river system was categorized using information from the Land Use Atlas of the Chilean Military Geographic Institute (Instituto Geográfico Militar de Chile, 2005). The land-use categories considered were as: (i) natural, (ii) agriculture, (iii) urban, and (iv) livestock and/or forestry. The Maipo and Maule Rivers share similar land use classifications, allowing a comparison between land use zones and both river courses. The natural zone represents an area with low influence of urban, agricultural, and/or livestock activity. At the same time, the other sites have large urban, agricultural, or livestock/forestry activities, respectively. Both the Maipo River's natural and agricultural land use areas receive the discharge of two wastewater treatment plants (WWTP), causing these zones to receive urban discharges.

For this reason, river water samples were collected before and after WWTP effluent discharge to determine whether this action influences the microbial source tracking by introducing *Cryptosporidium* spp. and *Giardia* spp. species that are human host specific into agricultural or natural areas. The selected urban sites corresponded to the discharge of rivers that have an important influence of major cities; for instance, the Mapocho River crosses the Chilean capital and major metropolis (Santiago) and discharges in the studied site in the Maipo River, and the Claro River receives the urban influence of the capital and major city in the Maule region (Talca) and discharges in the area sampled in the Maule River (Figure 1).

### Water Samples

Water samples were collected from each river every 3 months between June 2017 and April 2019. The timing of sampling was selected to ensure sampling during each season: summer (December–March), fall (March–June), winter (June–September), and spring (September–December). Fall and winter sampling events correspond to the rain event, while spring and summer sampling events correspond to the dry season. All sampling sites were georeferenced using Google Maps. Water was collected



in a sterile plastic bottle at each sampling site, which was rinsed three times in the river before collecting the sample. The sample was collected by submerging the bottle at least 10 cm below the surface. Immediately after sample collection, bottles were placed in a cooler and transported on ice. Samples were all processed within 24 h. Physical parameters of pH, water temperature ( $^{\circ}\text{C}$ ), conductivity ( $\mu\text{S}$  or  $\text{mS}$ ), salinity (ppm or ppt), and total dissolved solvents (ppm or ppt) were measured *in-situ* in each site using the Yalitech AM 006 Waterproof Multiparameter Meter Combo 6. Weather conditions (sunny, cloudy, or raining) were also recorded at the time of sampling.

### Fecal Coliforms Quantification

Water samples (15 ml) were used for fecal coliform enumeration following the Chilean Regulation NCh2313/22:1995 protocol for the detection of fecal coliforms in wastewater using EC medium [Instituto Nacional de Normalización (INN), 1995] to determine if each sample was compliant with Chilean water quality standards (INN, Instituto Nacional de Normalización, 1978) and whether the river complied with the water quality

standards for surface waters. Specifically, based on this regulation, we considered that a sample was compliant with the standard if fecal coliforms levels were below 1,000 fecal coliforms /100 ml of water; compliance is based on the assumption that fecal coliform levels above 1,000 fecal coliforms/100 ml is indicative of fecal contamination. The most probable number (MPN) of fecal coliforms in 100 ml of water was conducted following the instructions of the Chilean Regulation NCh2313/22:1995 protocol.

### Protozoa Quantification and Sequencing

To identify possible sources of fecal contamination, 20 L water samples were concentrated to 200 ml using Hollow-Fiber Ultrafiltration and B Braun DIACAP HIPS 20 (catalog code: 10-10-149144-5) hemodialysis filters following the methodology described by Hill et al. (2005) and Bambic et al. (2011). *Cryptosporidium* spp. oocysts and *Giardia duodenalis* cysts in the ultra-filtrated water were quantified according to the United States Environmental Protection Agency (EPA) method 1623 (Miller et al., 2005). *Cryptosporidium* spp. oocysts and

*Giardia duodenalis* cysts were concentrated using Dynabeads *Cryptosporidium*/*Giardia* Combo kit (Applied Biosystems, IDEXX Laboratories Inc., Maine, United States) immunomagnetic separation (IMS). IMS was conducted following Miller et al. (2005) with some modifications. Since metallic sediments interfered with the IMS step, a sediment wash was performed before adding the 100 µl IMS Dynal *Cryptosporidium* spp. and *Giardia duodenalis* beads to remove metallic sediments. Levels of each target were stained using Waterborne kit reagents (Aqua-Glo G/C Direct; Waterborne Inc., New Orleans, LA) per the manufacturer's instructions and enumerated according to Environmental Protection Agency (EPA) method 1623 (Miller et al., 2005). Briefly, 3-well slides (well diameter 14mm) were stained with fluorescein isothiocyanate (FITC), and DAPI (4,6-diamidino-2-phenylindole) and the entire well was examined for protozoa using an Olympus epifluorescence microscope. Organisms were visualized at ×200 magnification, and identifications were confirmed at ×400 (Miller et al., 2005). *Cryptosporidium parvum* and *C. hominis* oocysts were identifiable as 5–7 µm-diameter spheres with apple green outlines, often with a midline seam. *Cryptosporidium* spp. oocysts were identified as 5–7 µm-diameter spheres outlined in apple green and often with a midline seam, whereas *Cryptosporidium andersoni*/*C. muris* like organisms were identified as 5-by-7-µm elliptical forms (Miller et al., 2005). *Giardia* cysts were also apple green but oval and 9–14 µm long (Miller et al., 2005). The same experienced microscopist read all slides. As described below, samples were then subjected to DNA extraction and conventional PCR for confirmation and sequencing of *Cryptosporidium* spp. and *Giardia duodenalis*.

## Protozoa Molecular Analyses

A modified version of a previously published method (Miller et al., 2005) was used for extracting DNA from all the slides. After protozoa quantification the slide well for each sample was scraped with a scalpel blade and washed with sterile PBS into a separate microcentrifuge tube. Proteinase K (40 µl) was added to the sample and kept at 56°C overnight for sample digestion. DNA was extracted using Qiagen DNA Minikit (Qiagen Inc., Valencia, CA), and subsequently purified using a QIAamp column. Extracted DNA was stored at –20°C until PCR analysis.

The *Cryptosporidium* PCR was based on molecular characterization of the 18S rRNA gene using nested primers that amplify ~825 bp product (Xiao et al., 1999, 2000), and using conventional primers that amplified a 298-bp DNA fragment (Morgan et al., 1997). Amplification conditions were described by Adell et al. (2014). The *Giardia* PCR was based on molecular characterization of the glutamate dehydrogenase (GDH) and the SSU rRNA genes. GDH confirmation amplifies a ~432-bp fragment using a semi-nested-PCR protocol and primers (Read et al., 2004), while a ~292 bp final fragment of the SSU rRNA gene was amplified using a nested PCR protocol and primers (Appelbee et al., 2003). All primers are listed in **Supplementary Table S1**. All DNA was purified and submitted to Macrogen (Macrogen, Inc. Korea) for sequencing.

The protozoan parasites *Cryptosporidium* spp. and *Giardia* spp. have been used as an MST tool because both provide knowledge of host specificity (Ryu et al., 2011; Prystajec et al.,

2014; Xiao et al., 2018). For instance, the genus *Cryptosporidium* is composed of genotypes that are host specific and genotypes that infect various hosts (Xiao et al., 2000). For example, the major hosts for *C. parvum* are cattle, other ruminants, and humans; for *C. andersoni* cattle and bactrian camels; for *C. baileyi*, poultry and birds; for *C. meleagridis*, humans, turkeys, and birds; for *C. canis*, dogs; and other 10 species that have specific hosts (Xiao et al., 2004; Cacciò et al., 2005).

A situation also occurs with *Giardia* genotypes and assemblages (Feng and Xiao, 2011). *Giardia* spp. are divided into six species based on morphological characteristics, among which *Giardia duodenalis* (synonyms *G. lamblia* and *G. intestinalis*) infects a variety of mammals, including humans and livestock (Cacciò et al., 2005) and will be referred to as *G. duodenalis* throughout the manuscript. *G. duodenalis* also have several genetic groupings (henceforth assemblages), each of them having specific hosts. Thus, assemblage A can infect humans, other primates, livestock, dogs, cats, rodents, and other wild animals; similarly assemblage B is specific to humans and other primates, dogs, some species of wild mammals, assemblage C and D (also reported as *Giardia canis*) are specific to dogs and other canids; assemblage E (also reported as *G. bovis*) to cattle and other hoofed livestock; assemblage F (also reported as *G. cati*) to cats; and assemblage G (or *G. simondi*) to rats (Cacciò et al., 2005; Monis et al., 2009).

Therefore, in our study the identifying the genotype of *Cryptosporidium* and assemblage for *Giardia duodenalis* will be used to provide information on the host source (animal or human) of origin of the fecal contamination, its infective potential to productive animals (cattle), and its zoonotic potential.

## Protozoa Phylogenetic Analysis

To allow for the identification of *G. duodenalis* and *Cryptosporidium* subtypes associated with specific hosts and, therefore, fecal sources, three custom databases (*Giardia*\_SSU, *Giardia*\_GDH, and *Crypto*\_18S) were built by searching the NCBI-nr database and publicly available literature. DNA sequences were aligned using ClustalW-2.1 (GNU Lesser GPL, 2010), and phylogenies were built in Molecular Evolutionary Genetics Analysis (MEGA) version 10.2.5 software (Kumar et al., 2018) using maximum likelihood and Tamura 3. Phylogenetic tree topological reliability was determined using the bootstrap method, with 1,000 replicates and bootstrap lower than 70% were deleted from the tree. Accession numbers with the nucleotide sequence data for PCR products used for phylogenetic classification obtained during this study are available in **Supplementary Table S2**.

## Statistical Analysis

Conditional inference trees (CTree) are a type of modified classification and regression tree (CART) to address fundamental limitations of traditional CART algorithms (e.g., bias toward using categorical variables with many levels/continuous variables for splits over categorical variables with fewer levels). It is important to note that this is a multivariable and not a multivariate analysis and is a standard approach used in environmental microbiology literature, including by water-focused

studies throughout the past decade (Bae et al., 2010; Strawn et al., 2013; Verhoughstraete et al., 2015; Weller et al., 2016, 2020a; Green et al., 2021; Guo and Lee, 2021; Hannan and Anmala, 2021). Ctrees and CARTs are ideal tools for capturing complex (e.g., hierarchical) relationships within data and are thus, well suited for observational studies in non-controlled environments (Kuhn and Johnson, 2016). For example, identifying combinations of factors (specific scenarios) associated with an increased or decreased likelihood of detecting each microbial target, which is something regression-based approaches cannot do (Kuhn and Johnson, 2016). Tree-based approaches also create easy-to-understand visuals and do not rely on significance testing, thereby avoiding limitations associated with multiple comparison testing (Kuhn and Johnson, 2016). Moreover, tree-based analyses do not generate effect estimates or odds ratios to quantify the strength of associations, which is worth mentioning as other studies have reported substantial variability in water quality within a waterway over time; therefore, the duration of our study was insufficient to generate reliable effect estimates (Kuhn and Johnson, 2016; Weller et al., 2020b,c).

Conditional tree analysis was used to characterize the relationship between environmental factors and (i) fecal coliform levels log transformed ( $\log_{10}$  fecal coliforms/100-ml), (ii) whether fecal coliform levels exceeded the Chilean water quality standard of 1,000 fecal coliforms/100-ml (Yes/No), and (iii) if *Giardia* spp. were detected by PCR (Yes/No); tree analyses were implemented using the *mlr* and *party* packages and visualized using the *partykit* as previously described (Strobl et al., 2007a,b, 2009; Boulesteix et al., 2015; Weller et al., 2020b,c). Briefly, hyperparameters were tuned, models trained, and performance assessed using three-fold cross-validation repeated 10 times; cross-validation was performed to optimize R-squared (for continuous outcomes) and area under the curve (AUC; for binary outcomes). For imbalanced, binary outcomes (i.e., *Giardia* detection versus non-detection), SMOTE resampling was performed as part of tuning (Kuhn and Johnson, 2016; Weller et al., 2021).

To minimize the potential for overfitting we: (i) tuned the *maxdepth* parameter, (ii) limited the upper bound of *maxdepth* to 6, and (iii) used a *mincriterion* of 0.95. Conditional forests analysis was used since it is robust to correlation and missingness in explanatory factors (by enabling surrogate splits), can handle both categorical and continuous outcomes, and can capture complex, hierarchical relationships (e.g., interactions) between explanatory factors. The environmental features considered were the river system (i.e., Maule versus Maipo); predominant land use at the site [i.e., plant-based agriculture (henceforth agricultural), livestock operations, natural, urban sites not at a wastewater discharge (henceforth urban) and urban sites with a wastewater discharge (henceforth treated urban)]; season (i.e., wet versus dry); if feces, garbage, or livestock were present at the sampling site; pH; water temperature; total dissolved solids; salinity; if it rained during the 24h before sampling; and weather conditions at the time of sampling (i.e., rainy, sunny, partly cloudy, or cloudy). For *Giardia*, coliform levels and if coliform levels met the Chilean water standard (i.e., were above or below 1,000 fecal coliforms/100-ml) were also considered explanatory factors. Separately, nonparametric

Spearman's one-tailed tests were conducted to assess the correlation between continuous outcome variables (i.e., MPN of fecal coliforms/100 ml, *Giardia* cysts/100 ml, and *Cryptosporidium* oocysts/100 ml) to enable comparison of the findings reported here with previous studies (e.g., Wu et al., 2011; Korajkic et al., 2018). A generalized linear model (GLM) was implemented to determine whether compliance with the Chilean water quality standard (i.e., fecal coliforms above versus below 1,000 fecal coliforms/100-ml) was associated with protozoa detection (i.e., if *Cryptosporidium*, *Giardia*, or no protozoa were detected). In both analyses (Spearman's correlation and GLM), differences with  $p < 0.05$  were statistically significant. All statistical analyses were performed using RStudio v1.1.456 or R v3.6.3 (R Core Team, 2020; RStudio Team, 2020).

## Detection of Antimicrobial Resistant *Escherichia coli*

To determine if multi-drug resistant *E. coli* were present in water samples, a subset of samples (from the first four of the eight total samplings) were tested ( $n=48$  total samples were tested). Briefly, 1 ml ultra-filtered samples were enriched in 5 ml of Buffer Peptone Water and incubated at 37°C for 24 h. Following incubation, 100  $\mu$ l of the samples were streaked onto plates with MacConkey agar supplemented with Ciprofloxacin (1 mg/L), Cefotaxime (1 mg/L), and Tetracycline (4 mg/L) and incubated at 37°C for 24 h. Colonies were selected according to morphology using a magnifying glass. This selection was made according to standard patterns of color (pink), as described previously (Jung and Hoilat, 2020). All isolates were identified using MALDI-TOF (BioMérieux, Marcy l'Étoile, France) and only the isolates identified as *E. coli* were tested against a panel of 15 antibiotics using the disk diffusion method to assess their antimicrobial susceptibility profile following CLSI guidelines (Clinical and Laboratory Standards Institute, 2018a). The antibiotic used were: Ampicillin (AMP, 10  $\mu$ g); Cefazolin (CFZ, 30  $\mu$ g); Ceftriaxone (CRO, 30  $\mu$ g); Cefepime (FEP, 30  $\mu$ g); Imipenem (IPM, 10  $\mu$ g); Ciprofloxacin (CIP, 5  $\mu$ g); Amikacin (AMK, 30  $\mu$ g); Gentamicin (GEN, 10  $\mu$ g); Fosfomycin/Trometamol (FOF, 200  $\mu$ g); Amoxicillin-Clavulanate (AMC, 30  $\mu$ g); Cefuroxime (CXM, 30  $\mu$ g); Aztreonam (ATM, 30  $\mu$ g); Chloramphenicol (CHL, 30  $\mu$ g); Tetracycline (TET, 30  $\mu$ g); and Trimethoprim/Sulfamethoxazole (SXT, 1.25/23.75  $\mu$ g). All antibiotics were supplied by OXOID (Hampshire, England). Isolates resistant to three or more antimicrobial classes were cataloged as MDR (Multi-Drug-Resistant Bacteria) following previously standardized criteria (Magiorakos et al., 2012).

## Detection of Environmental Resistant Bacteria

To further characterize environmentally resistant bacteria, 50 ml of water samples from Maipo and Maule Rivers, were passed through a 0.22- $\mu$ m filter. This filter was resuspended in 0.85% saline solution, and 200  $\mu$ l were seeded in TSB and R2A agar media in 120 mm square plates. These plates were incubated at 37°C, 25°C and 15°C, and the colonies were classified according to the standard pattern: shape, color, texture, and shape of the colony border and designated as morphotypes (Higuera-Llantén et al., 2018). A total of 2631 morphotypes

were isolated at all sampling points, media, and temperatures. Minimum Inhibitory concentrations (MICs) for Ampicillin (AMP), Erythromycin (ERY), Ceftazidime (CAZ), Imipenem (IMI), Ciprofloxacin (CIP), Gentamicin (GEN) and Tetracycline (TET) were determined by agar dilution methods as recommended by CLSI (Clinical and Laboratory Standards Institute, 2018b). Briefly, serial two-fold dilutions ranging from 0.1 to 2,048 µg/ml of any antibiotic were added into Mueller-Hinton agar (Difco). After overnight growth, all isolates were resuspended in 0.85% saline solution and inoculated in plates using a 96-needle replicator. The isolates that grew at 37°C were incubated for 24h, while those that grew at 25°C and 15°C were incubated for 48h. A Mueller-Hinton agar plate without antibiotics was used for growth and possible contamination control. These experiments were performed in triplicate. MIC was defined as the lowest concentration of antibiotic that inhibits the growth completely at least in two of three plates seeded at any temperature. As a reference, we used the strain *Escherichia coli* ATCC 25922, as recommended by CLSI (Clinical and Laboratory Standards Institute, 2018b).

## RESULTS

### Coliform Detection and Evaluation of Their Association to Environmental Factors

Fecal coliform levels in the Maule River ranged between 1 and 130 MPN/100-ml, not exceeding the Chilean water quality standard of 1,000 fecal coliforms/100-ml (Table 1). Conversely, fecal coliform levels in the Maipo ranged between 2 and 30,000 MPN/100-ml in 11 samples exceeding the water quality standard. Most of the Maipo River samples that exceeded the standard were mainly collected from urban sites. In both rivers, fecal coliform levels varied based on season, with higher levels during the dry season (mid-September to mid-March) compared to the wet season (mid-March to mid-September; Table 1). Fecal coliform levels also appeared to vary based on land use (Table 1).

The samples with the highest coliform levels were collected from urban sites in both rivers, while the samples with the lowest coliform levels were collected from natural sites (Table 1). While the range in fecal coliforms levels from treated urban sites (i.e., sites in the river at a wastewater discharge; <2–8,000 CFU/100-ml) overlapped the range for urban sites (i.e., sites not at a discharge; 34–30,000 CFU/100-ml), the mean level in treated sites (6,887 CFU/100-ml) was considerably lower than the mean in urban sites (32,705 CFU/100-ml); this difference was even more pronounced when only Maipo River sites were considered.

According to conditional tree analysis, fecal coliform levels were significantly lower in the Maule compared to the Maipo River ( $p < 0.001$ ; Figures 2A,B). Specifically, fecal coliforms were lowest in Maule River sites and highest in samples from Maipo River sites where the predominant land use was agricultural (i.e., plant-based production) or urban (i.e., not at a wastewater discharge) as opposed to livestock, natural, or treated urban (i.e., urban sites at a wastewater discharge; Figure 2A). All samples collected from the Maule River met the Chilean water quality standard, and this river was significantly associated with if a water sample was compliant with the standard ( $p < 0.01$ ; Figure 2B). For Maipo River, samples were most likely to be non-compliant if collected from sites where the predominant land use was agricultural, urban, or treated urban during the dry season. Samples were most likely to be compliant if collected from sites where the land use was natural or livestock-based production (Figure 2B). For urban, treated urban and agricultural sites, samples were more likely to meet the standard if collected during the wet season ( $p = 0.002$ ; Figure 2B).

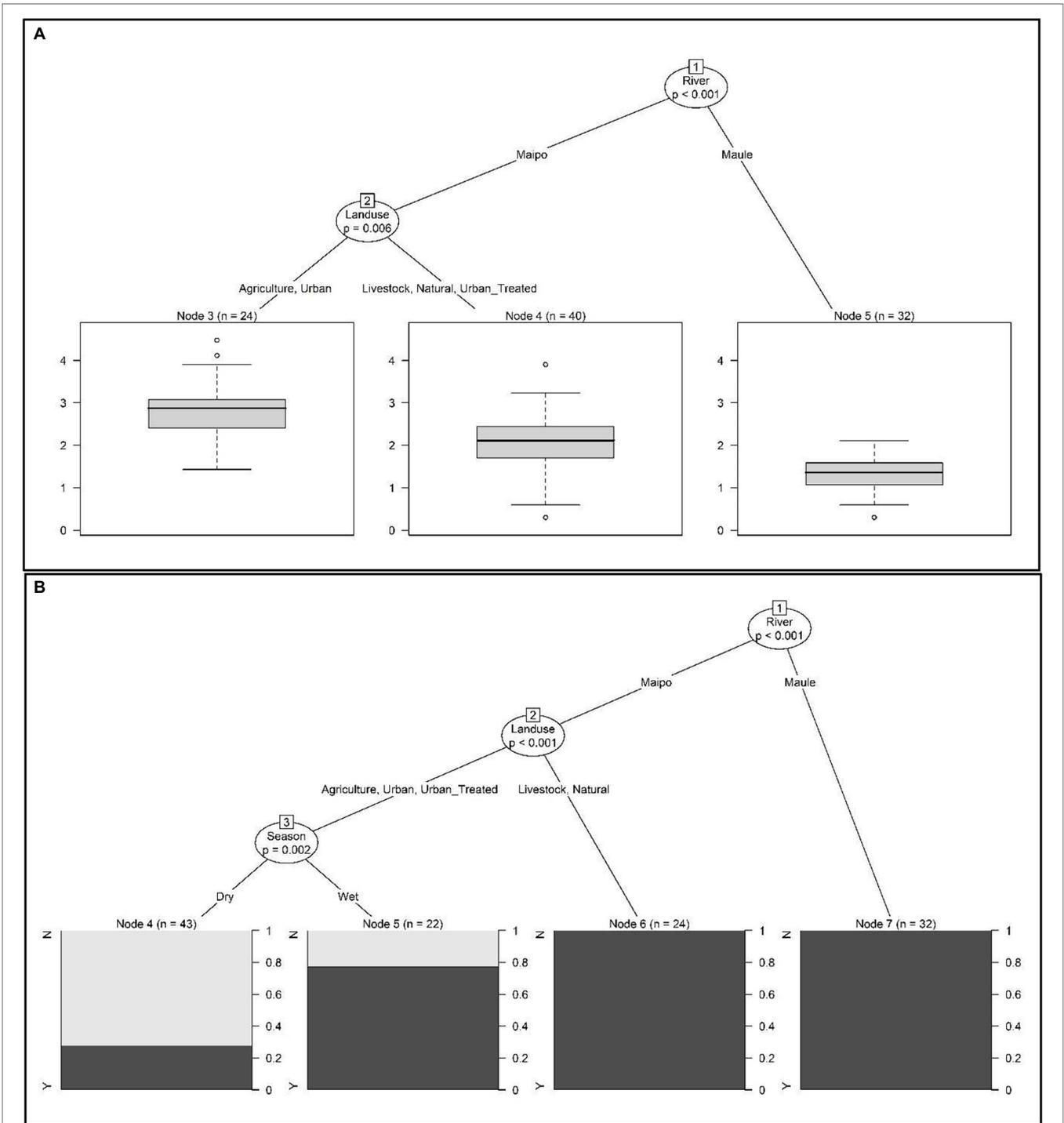
### Evaluation of *Cryptosporidium* spp. and *Giardia duodenalis* as Microbial Source Tracking Tools in Maipo and Maule Rivers and Their Associations With Environmental Factors

Higher *Cryptosporidium* spp. counts were found in water samples collected from the Maipo River (range=0–60 oocysts/10L;

**TABLE 1** | Fecal coliform levels (MPN/100ml) for each sample collected in the Maule and Maipo Rivers between 2017 and 2019.

River	Site [predominant land use(s)]	Fecal coliform counts (MNP/100ml) by season and year							
		2017			2018		2019		
		Winter	Spring	Summer	Fall	Winter	Spring	Summer	Fall
Maule	Site 1 (natural)	<2	13	<2	13	11	<2	23	2
	Site 2 (agricultural)	4	13	130	50	4	23	80	23
	Site 3 (urban)	50	70	30	30	110	130	17	23
	Site 4 (livestock/forestry)	22	2	30	30	30	17	22	50
Maipo	Site 5 (natural)	17	27	300	80	70	80	170	50
	Site 6 (treated urban)	200	4	30	4	8,000 <sup>1</sup>	80	1,700 <sup>1</sup>	<2
	Site 7 (natural)	50	27	130	50	230	70	33	80
	Site 8 (agricultural)	220	130	700	230	800	5,000 <sup>1</sup>	1,300 <sup>1</sup>	130
	Site 9 (treated urban)	80	27	800	130	230	1,100 <sup>1</sup>	300	300
	Site 10 (agricultural)	27	800	1,300 <sup>1</sup>	800	700	1,100 <sup>1</sup>	500	700
	Site 11 (urban)	34	280	30,000 <sup>1</sup>	800	1,100 <sup>1</sup>	700	13,000 <sup>1</sup>	8,000 <sup>1</sup>
	Site 12 (livestock)	280	280	130	220	300	220	130	500

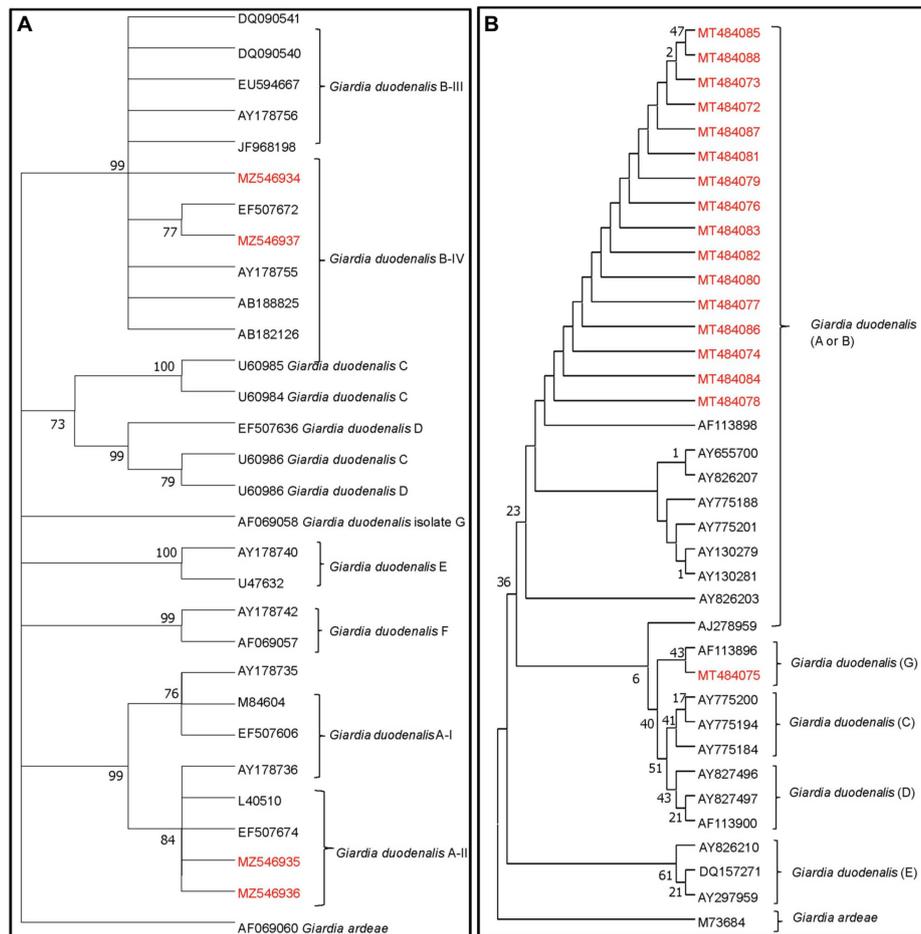
<sup>1</sup>Fecal coliform level exceeded the maximum limit of 1,000 fecal coliforms/ml established by the Chilean water quality standards (INN, Instituto Nacional de Normalización, 1978).



**FIGURE 2 |** Conditional inference trees were used to visualize hierarchical relationships between environmental factors and **(A)** fecal coliform levels ( $\log_{10}$  MPN/100-ml) in water samples, and **(B)** if fecal coliform levels were above or below the Chilean water quality standard (i.e., above or below 1,000 CFU/100-ml). In **(A)** the boxplot shows the distribution of  $\log_{10}$  fecal coliform levels in samples that met the given condition. In **(B)** the black bar in each plot shows the probability of the water sample being non-compliant with the Chilean water quality standard when the given conditions were met. For example, fecal coliform levels were highest in Maipo River samples collected from site with predominantly agricultural and urban land uses (that were not at a wastewater discharge site).

mean = 3.18 oocysts/10L) compared to the Maule River (range = 0–31 oocysts/10L; mean = 2 oocysts/10L; **Supplementary Figure S1**). Higher *Giardia* spp. counts were also found in the Maipo

(range = 0–171 cysts/10L, and a mean of 7 cysts/10L) compared to the Maule River (range = 0–4 cysts/10L, and a mean of <1 cysts/10L; **Supplementary Figure S1**). In Maipo River samples,



**FIGURE 3 |** Maximum likelihood phylogenetic analysis of *G. duodenalis* detected from river water samples. The analysis was constructed by using the Tamura 3 parameter model with MEGA 7.0 and the bootstrap values were calculated with 1,000 replicates. The analysis is based on (A) the GDH gene and (B) ssuRNA. The phylogenetic tree was rooted to *G. ardeae*. Depending on the assemblages of *G. duodenalis* determined by the different clades of the phylogenetic tree, the possible host source of fecal contamination could be inferred. The numbers at branch nodes represent bootstrap values greater than 70. Reference sequences included in the analysis are shown with their respective GenBank accession numbers. *G. duodenalis* strains characterized in this study are shown in red text.

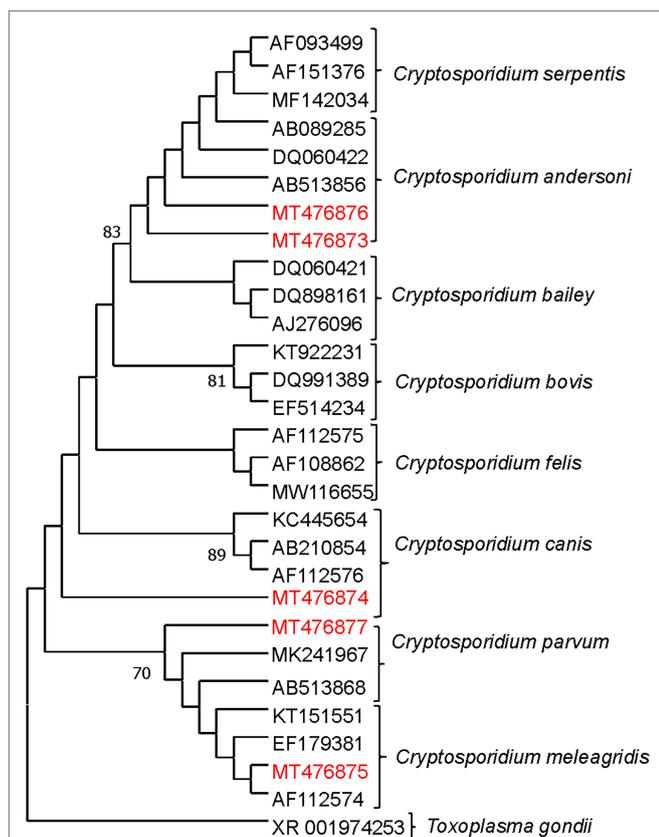
two assemblages of *G. duodenalis* were identified, including assemblages A or B (Figure 3; Supplementary Tables S2, S3). Only two *G. duodenalis* were detected in water samples collected from the Maule River. One of them corresponds to assembly A or B, and the other is assembly G. Also, four *Cryptosporidium* spp. were identified in the Maipo River, including *C. parvum*, *C. andersoni*, *C. meleagridis*, and *C. canis* (Figure 4; Supplementary Tables S2, S3). Sequencing analysis did not identify any *Cryptosporidium* spp. in the Maule River (Figure 4; Supplementary Tables S2, S3).

Most *Cryptosporidium*-positive samples were collected during the wet season. Whereas *Giardia*-positive samples were collected in both seasons, more positives were collected during the dry season. While fecal coliform levels were highest in samples collected from agricultural and urban sites, *G. duodenalis* assemblages A or B, and *C. meleagridis* were only detected in samples collected from agricultural sites (i.e., sites with plant-based production; Table 1; Figure 3; Supplementary

Tables S2, S3). While all, except one sample, collected from sites upstream of wastewater discharges (i.e., natural or agricultural sites) were *Giardia*-negative, multiple samples ( $n = 8$  samples) collected downstream of wastewater discharges (i.e., treated urban sites) were positive for *G. duodenalis* assemblages A or B (Supplementary Table S3). Indeed, conditional tree analysis indicated that the likelihood of detecting *Giardia* by PCR was greatest in samples collected downstream of wastewater discharges (i.e., treated urban sites;  $p < 0.001$ ) and lowest in samples collected from all other sites during the wet season (Figure 5).

### Correlation Between Fecal Coliforms and Protozoa Detected in River Water Samples

Spearman's correlation analysis indicated that coliforms counts (MPN/100 ml) significantly, and positively correlated with *Giardia* counts (cysts/100 ml; correlation coefficient = 0.31;  $p = 0.002$ ) but



**FIGURE 4 |** Maximum likelihood phylogenetic analysis of *Cryptosporidium* detected from river water samples. The analysis was constructed by using Tamura 3 parameter model with MEGA 7.0 and the bootstrap values were calculated with 1,000 replicates. The analysis was based on the SSU gene. The phylogenetic tree was rooted to *Toxoplasma gondii*. The numbers at branch nodes represent bootstrap values greater than 70. Reference sequences included in the analysis are shown with their respective GenBank accession numbers. *Cryptosporidium* strains characterized in this study are shown in red text. Based on the *Cryptosporidium* species determined by the different clades of the phylogenetic tree, the possible host source of fecal contamination could be inferred. \*Indicates that the sequence was inferred (or confirmed) using blast on the NCBI platform.

were not significantly correlated with *Cryptosporidium* counts (oocysts/100 ml; correlation coefficient = 0.03;  $p = 0.75$ ). However, the lack of a correlation between coliform and *Cryptosporidium* spp. levels may be due to the large number of *Cryptosporidium*-negative samples. According to GLM analysis, neither *G. duodenalis* (OR = 0.84;  $p = 0.896$ ; 95% Confidence Interval = 0.20, 6.33) or *Cryptosporidium* spp. (OR = 0.48;  $p = 0.531$ ; 95% Confidence Interval = 0.06, 9.95) detection were associated with fecal coliforms levels being compliant with the Chilean water quality standard of 1,000 CFU of fecal coliforms/100-ml.

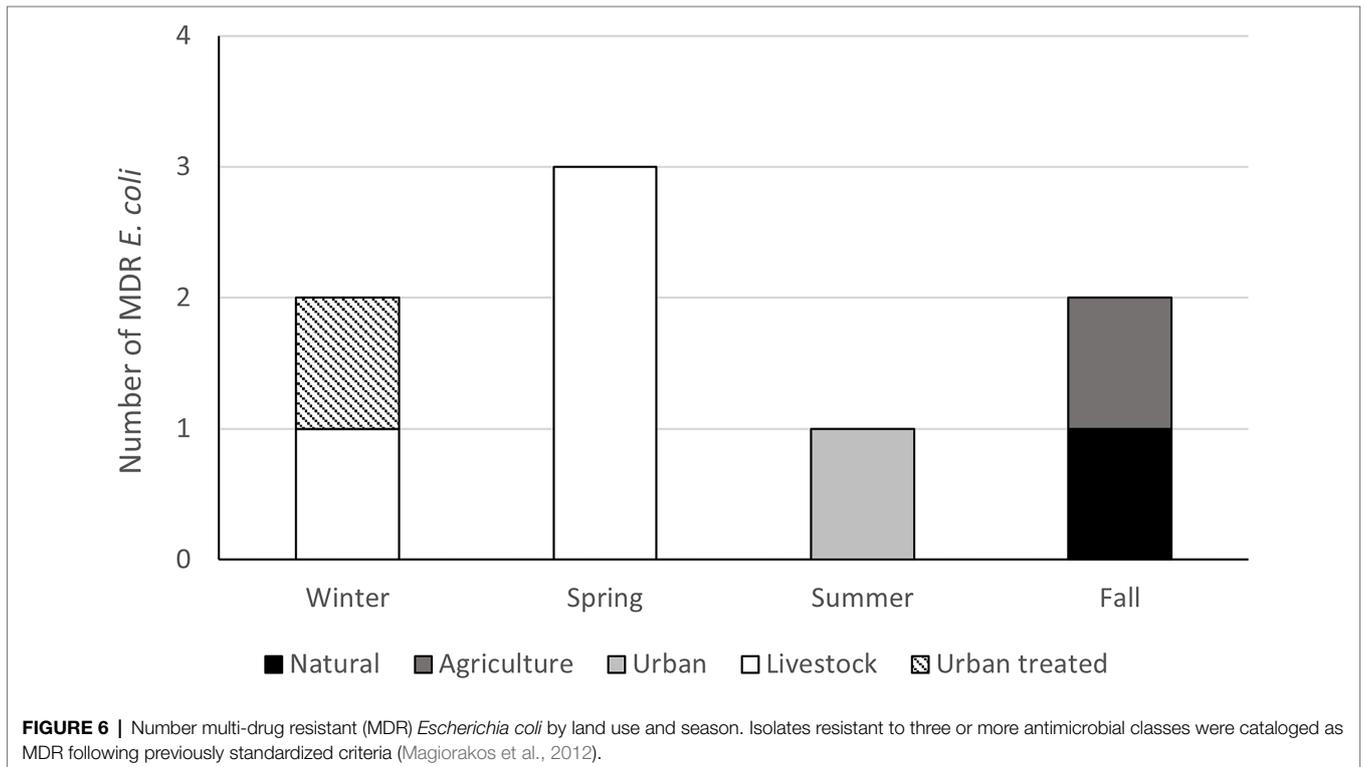
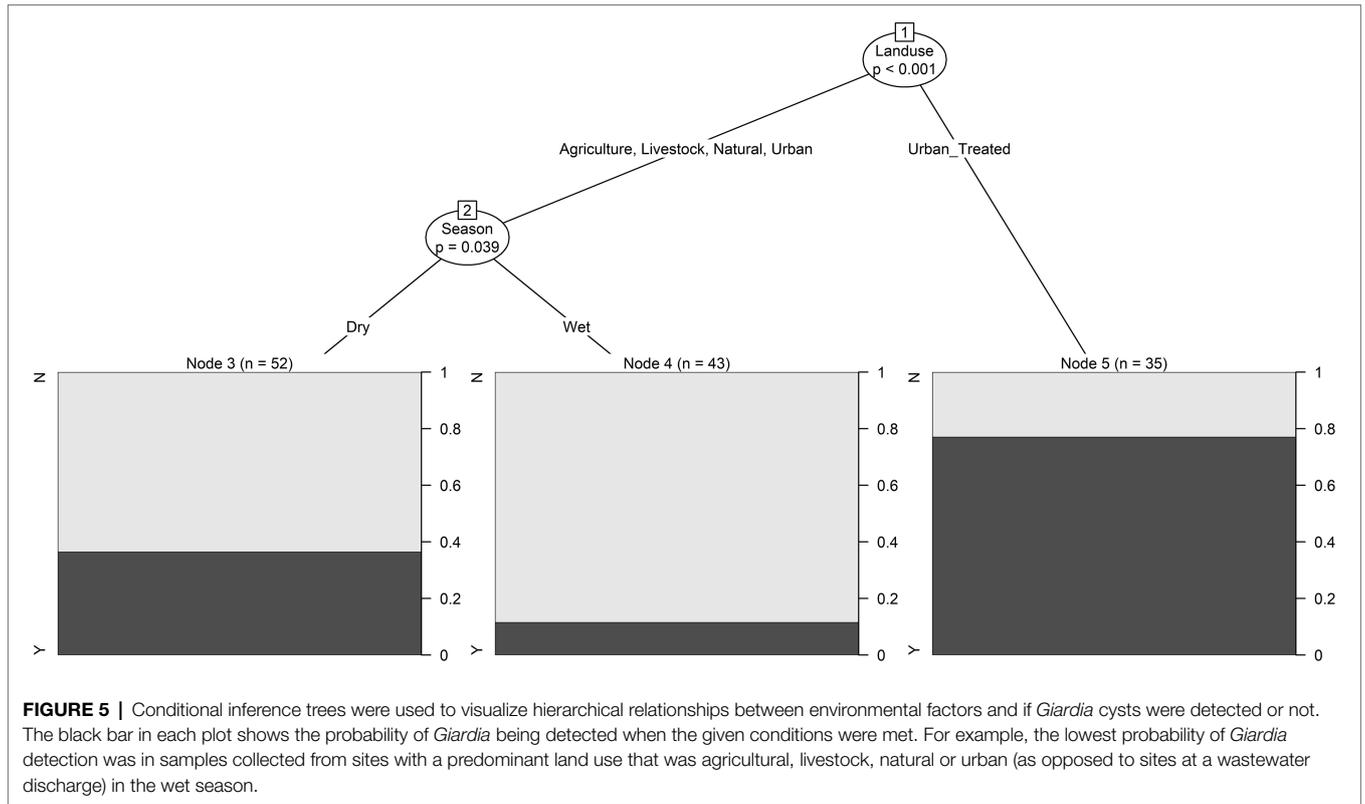
### Antimicrobial Resistant *Escherichia coli* and Environmental Resistant Bacteria

To evaluate antimicrobial resistance in enteric bacteria, *E. coli* were isolated, and their resistance profile was characterized. Twenty-eight out of 48 samples analyzed (58.3%) contained

Gram-negative bacteria resistant to at least one of the antibiotics tested (Ciprofloxacin, Cefotaxime, and Tetracycline). We obtained 66 isolates; 26 (lactose positives) isolates were characterized at the species level by MALDI-TOF MS. Finally, we identified eight *E. coli* isolates and three of them showed a multi-drug resistant profile (Figure 6; Supplementary Table S4). Regarding the environmental bacteria, 2,631 morphotypes were isolated at different temperatures (37°C, 25°C, and 15°C) and media. The major number of colonies were obtained at 25°C and in the R2A medium, demonstrating that the optimal temperature for environmental bacterial growth is 25°C. To determine the AMR profile, the MICs for Ampicillin (AMP), Erythromycin (ERY), Cefazidime (CAZ), Imipenem (IMI), Ciprofloxacin (CIP), Gentamicin (GEN), and Tetracycline (TET) were determined. Bacteria that showed resistance to at least three antibiotics of different families were classified as MDR bacteria. The percentage of MDR-bacteria was evaluated by season and land use and summer was the season with higher values in most of the land use analyzed (>2.5%), followed by the winter season where a high percentage of resistance was observed in both urban and livestock land uses (>2%; Figure 7).

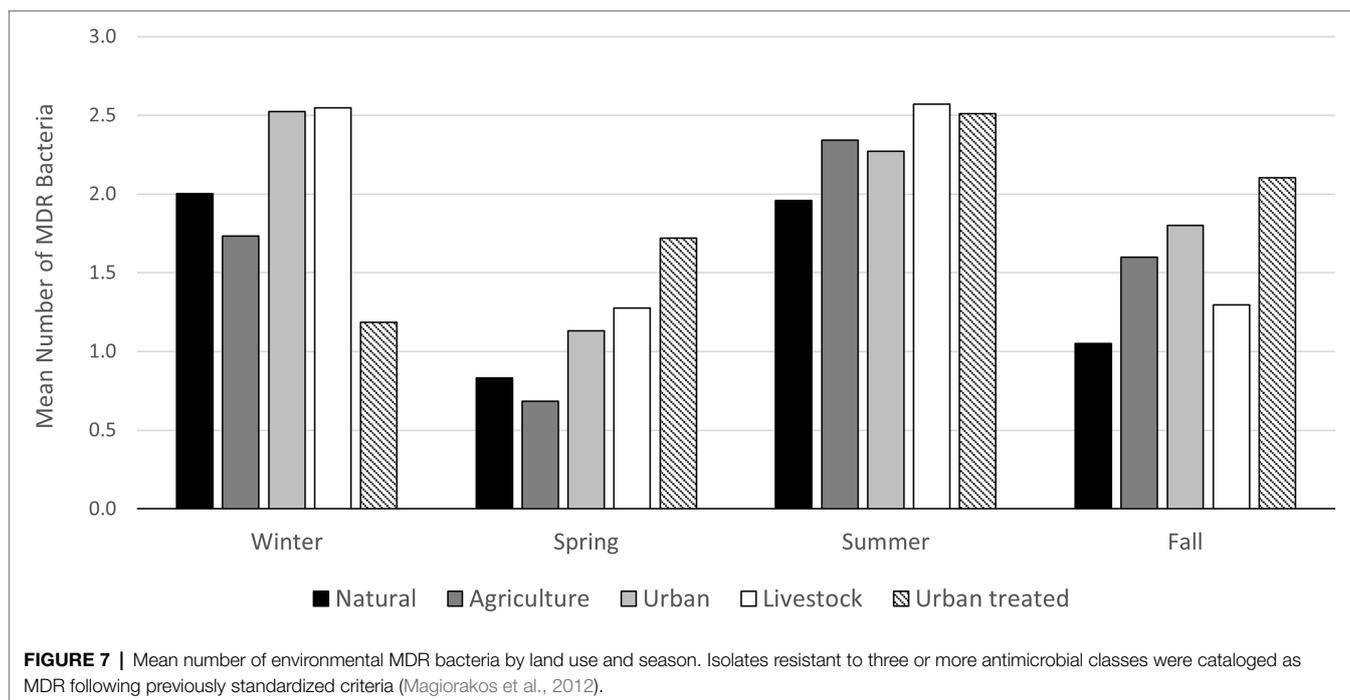
## DISCUSSION

The contamination of water with fecal material represents a problem for public health. Determining the source of fecal material could be useful to reduce the load of fecal material in rivers to designing mitigation measures that provide water with adequate quality for agriculture and livestock. In this study, fecal contamination levels were determined, and the probable sources of contamination were identified through an MST analysis, together with antimicrobial resistance, in two main rivers of central Chile based on land use. The main finding of our study suggests that: (i) microbiological water quality is affected by environmental conditions, particularly land use and season, and (ii) identification of *Cryptosporidium* and *Giardia* species provides information about the possible host source of the fecal contamination. Microbial water quality in the Maipo River was worse than in the Maule River; this was expected due to the larger human population in the Maipo (approx. 9 million) compared to the Maule River (approx. 1 million) basin (INE, Instituto Nacional de Estadísticas, 2017). Indeed, past studies have linked impaired water quality to increased anthropogenic disturbance (Lenart-Boroń et al., 2017; Cui et al., 2019; Yuan et al., 2019). For example, Yuan et al. (2019) examined water quality in canals in Suzhou, China, and found that greater urbanization was associated with higher microbial loads, including fecal markers and human pathogens. Similarly, Cui et al. (2019) examined 41 water samples of 15 rivers urban rivers located in Changzhou City, China and reported that the samples harbored diverse enteric pathogens, which were highly correlated to the human fecal marker abundances. Correspondingly, Lenart-Boroń et al. (2017) evaluated the water quality at 21 sites in major rivers of the Podhale region (southern Poland) for a year. Their results suggest that an increased share of built-up areas and arable



land results in significant deterioration of water quality and concluded that point sources of pollution (e.g., effluents from treatment plants or discharge of untreated sewage from

households) were more important than non-point sources (e.g., surface runoff). These findings are supported by the results of the conditional trees in the present study; specifically, high



fecal coliform levels and a greater likelihood of detecting *Giardia* were observed for samples collected from sites with anthropogenic influence (e.g., urban and agricultural). Based on our findings efforts to mitigate fecal contamination in Central Chile should focus on the Maipo River, and specifically on sources of contamination in urban and agricultural regions. Further studies are needed to identify specific sources of contamination in urban and agricultural reaches of the Maipo River.

The correlation between fecal coliforms and parasites was evaluated. While we found evidence of a significant correlation between fecal coliform and *Giardia* levels, there was correlation between fecal coliform and *Cryptosporidium* levels, which may be driven by the small number of *Cryptosporidium*-positive samples in the present study. Indeed, a review that assessed relationships between FIBs and pathogens in water hypothesized that the failure to find associations between FIBs and the target was due to insufficient data (Wu et al., 2011). However, it is also important to recognize, that failure to find evidence of a relationship may also be due to the absence of a relationship. Indeed, several recent studies have shown that pathogen-FIB relationships are dependent on the waterway and/or environmental context of the sampling site and sampling event (McEgan et al., 2013; Weller et al., 2020b). Overall, our finding of an association between FIB levels and *Giardia* levels is consistent with past studies. Many were analyzed in review studies on the relationship between FIBs and pathogens and parasites in recreational water (Korajkic et al., 2018). Specifically, this review found that almost 50% of the 73 studies considered performed statistical analysis to determine if there was a significant association between FIB and other pathogens (including parasites). Only considering freshwater studies ( $n=41$ ),

a positive association with pathogens was reported in two-thirds of them ( $n=23$ ), and the remaining studies did not support statistical analysis (Korajkic et al., 2018).

*Cryptosporidium* spp. and *Giardia* spp. have been extensively used to perform MST analysis (Prystajecy et al., 2014; dela Peña et al., 2021; González-Fernández et al., 2021). During this study, *Cryptosporidium* spp. and *G. duodenalis* were determined in water samples. Most *Cryptosporidium* spp. and *G. duodenalis* assemblages corresponded to subtypes or species that are not host-specific, such as *C. parvum* and *G. duodenalis* assemblages A and B, respectively (Xiao et al., 2000, 2004; Cacciò et al., 2005; Monis et al., 2009). Even though the amplicon used in our study generated sequences too short to allow differentiation between assemblages A or B, it does not affect our result interpretation as both assemblages are not host-specific. Furthermore, despite that *G. duodenalis* A and B are host-unspecific, it has been reported that humans and livestock are among the main hosts affected by these assemblages (Cacciò et al., 2005; Monis et al., 2009). Nevertheless, it was possible to identify a *C. canis* (in one sample from natural land use), *C. meleagridis* (in one sample from agricultural land use), and *C. andersoni* (in two samples from urban land use), which are specific to dogs, humans, and birds, and cattle, respectively (Xiao et al., 2000, 2004; Cacciò et al., 2005). *Cryptosporidium canis* has been reported to infect wild canines such as foxes (Zhang et al., 2016; Papini and Verin, 2019). Since various canines such as foxes (*Lycalopex culpaeus*) and free-ranging dogs, among others, inhabit the natural sections of the river systems studied (Cevitanes et al., 2020; Benavides et al., 2021), this may explain the detection of *C. canis* in the present study from samples collected from sites where the predominant land use is natural. *G. duodenalis* (assemblage

G) associated with rats and/or mice was detected in the livestock/forestry area in the Maule River (Cacciò et al., 2005; Monis et al., 2009). Therefore, based on our results, *Cryptosporidium* spp. and *G. duodenalis* provided information about the possible host source of fecal contamination in rivers. To our knowledge, a standard profile for *Cryptosporidium* spp. does not exist. Therefore, we cannot make comparisons. Our study had a low likelihood of detecting *Cryptosporidium*, with many samples presenting non-detection of this parasite. A possible explanation of this result is that metallic sediments present in the river water samples may negatively affect the performance of the IMS technique and, therefore, hamper the recovery of *Cryptosporidium* spp. Therefore, improvements must be conducted in the detection protocol as we presented complications when using the immunomagnetic separation technique (IMS) as the rivers in Chile naturally have high concentrations of metallic sediments (copper, among others; Ghorbani and Kuan, 2016), which might have interfered with the detection of these protozoa, causing a possible underestimation in the detection. This could explain the low levels and identification of *Cryptosporidium* spp. in river water samples in our studies. Further studies should be conducted to evaluate whether the metallic sediments of mining activities interfere with detection methods involving magnetic beads and how to overcome this problem.

In past studies, *Cryptosporidium* oocysts and *Giardia* cysts have been detected in treated effluents of WWTP (Kitajima et al., 2014; Domenech et al., 2018), and reports indicate that some treatments along the WWTP process (stabilization ponds, constructed wetland, ultrafiltration, and UV irradiation) are found to be effective for the removal of *Cryptosporidium* oocysts, while other treatments (activated sludge, high-rate filtration, and chlorine disinfection) are ineffective (Nasser, 2016). In Chile, the removal of protozoa from effluent and the use of UV light, ozone, or other treatments besides chlorine in the tertiary treatment is not mandatory for WWTP (Ministerio Secretaría General de la Presidencia, 2001); therefore, it is highly likely that protozoa are not completely removed from the treated effluent. We collected river water samples upstream and downstream of the discharge of treated effluent of WWTP to determine the effect on the number and species of protozoa released in the river water. It is worth mentioning that once the effluent is released from the treatment plant by an underground pipe, it flows through an open canal for approximately 30 m before reaching the river. Along the trajectory of the effluent, it crosses some households that have horses, chickens and/or dogs. Therefore, fecal or protozoan contamination may come from flowing through this canal rather than from the WWTP. This could explain the presence of *G. duodenalis* assemblages A or B without a specific host that is probably from human sources. Considering this finding, we suggest that the effluents from the WWTP should be discharged directly into the rivers through underground pipes rather than into open canals to evaluate whether these protozoa are coming from the effluent of wastewater treatment, therefore associated with humans, or are coming from other animal sources (Encyclopedia Britannica, 2021).

Aquatic ecosystems are considered the most important matrices for the release, mixing, persistence, and dissemination of ARB (Baquero et al., 2008; Taylor et al., 2011; Yang et al., 2018); especially freshwater bodies since mixing occurs with a high risk of genetic exchange between coliforms and environmental bacteria (Kraemer et al., 2019; Nnadozie and Odume, 2019). Determining the impact of the mixture between environmental bacteria and bacteria that arise from contamination by human activity, like coliforms, is one of the great challenges in the AMR phenomenon. In this study, we have shown that the greatest amount of environmental ARB and coliforms are isolated in the summer season, which could increase the probability of genetic exchange between these two microbiomes. The same effect could be observed when the variable land use is analyzed. In both cases, soils classified as urban presented the highest levels of fecal coliforms and environmental ARB, which further enables genetic exchange between these two bacterial groups. Understanding the factors that favor the mobilization of antibiotic resistance determinants between different microbiomes is pivotal in the resistance phenomenon. In this study, we have determined that the summer season and the type of urban land use are the factors where the greatest levels of coliforms were counted and where the largest number of environmental ARBs were isolated; this could favor the genetic exchange between these two microbiomes, generating a major problem in the AMR phenomenon.

## CONCLUSION

Our results indicate that river water used as irrigation water has high levels of fecal coliforms, with many sites evaluated having noncompliance with the Chilean water quality standard, protozoa, and AMR. Fecal contamination (level or based on the compliance of a water quality regulation), protozoa levels, and ARB presence in river water depended on the interaction of environmental factors such as river, land use, and season. In our study, the Maipo River had higher levels of coliforms (and noncompliance with the Chilean water quality criteria), protozoa detection, and ARB than the Maule River. The land use that contributed most to the fecal and protozoa contamination of river water was after the discharge of wastewater treatment plants (urban treated), urban land use, and agriculture. Therefore, the environmental factors of land use, river, and season should be considered when designing studies to evaluate water contamination as the microbial contamination varies according to the human activity in the area of interest.

The presence of different species of *Cryptosporidium* spp. and assemblages of *Giardia duodenalis* indicate that humans, cattle, birds, rats/mice, and dogs are the most probable source of origin of the fecal contamination. Furthermore, these protozoa are known to be zoonotic and therefore may pose a potential public health threat if this water is directly or indirectly consumed. Therefore, based on our results, we can conclude that *Cryptosporidium* spp. and *G. duodenalis* detection and land use can be used as MST tools to provide information on the possible source of fecal contamination. Finally, we can say that the highest levels of fecal coliforms were detected in both the summer season and the land

use classified as urban, and the largest amount of environmental ARB was isolated, increasing the probabilities of mobilization and dispersion of genetic elements involved in AMR. These data could aid in implementing measures to mitigate the dispersion of coliforms and ARB, thus reducing risks to public health.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

## AUTHOR CONTRIBUTIONS

CD-G, CB, LV, MK, AM-S, JO-P, and ADA designed the study, planned the field work logistics, and wrote the manuscript. DW and AA ran the statistical analysis, while CD-G constructed the phylogenetic trees. CD-G, CB, MS-C, EE, AA, MK, and ADA conducted experiments. DW, MS-C, EE, WS, MK, AM-S, JO-P, and ADA designed experiments and critically reviewed the manuscript. All authors contributed to the article and approved the submitted version.

## FUNDING

We acknowledge the financial support received from the FONDECYT 11160116 and 1221536 to ADA, FONDECYT

1181167 to AM-S, and ANID Millennium Science Initiative/Millennium Initiative for Collaborative Research on Bacterial Resistance, MICROB-R, NCN17\_081 to ADA, AM-S, and JO-P. Data analysis was supported by the National Institute of Environmental Health Sciences of the National Institutes of Health (NIH) under award number T32Es007271.

## ACKNOWLEDGMENTS

The authors thank field teams who supported collections of the river water samples. The collection of river water samples from the San Antonio Sites was facilitated through the School of Military Engineers of San Antonio, the authors specifically acknowledge Captain Ricardo Pasten and Cape Pablo Tapia. The authors also acknowledge the Wastewater Treatment Plant for helping us in finding the pipe that discharged treated water into the river. The authors thank Miles Daniels from the University of California, Santa Cruz for coming to Chile to train us in the installation and use of the hollow ultrafiltration system. The authors are also grateful to the students that assisted in the processing of water sample collection and processing as part of their internships and course training.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2022.768527/full#supplementary-material>

## REFERENCES

- Adell, A. D., Smith, W. A., Shapiro, K., Melli, A., and Conrad, P. A. (2014). Molecular epidemiology of *Cryptosporidium* spp. and *Giardia* spp. in mussels (*Mytilus californianus*) and California Sea lions (*Zalophus californianus*) from Central California. *Appl. Environ. Microbiol.* 80, 7732–7740. doi: 10.1128/AEM.02922-14
- Appelbee, A. J., Frederick, L. M., Heitman, T. L., and Olson, M. E. (2003). Prevalence and genotyping of *Giardia duodenalis* from beef calves in Alberta, Canada. *Vet. Parasitol.* 112, 289–294. doi: 10.1016/S0304-4017(02)00422-3
- Azuma, T., and Hayashi, T. (2021). Effects of natural sunlight on antimicrobial-resistant bacteria (AMRB) and antimicrobial-susceptible bacteria (AMSB) in wastewater and river water. *Sci. Total Environ.* 766:142568. doi: 10.1016/j.scitotenv.2020.142568
- Bae, H.-K., Olson, B. H., Hsu, K.-L., and Sorooshian, S. (2010). Classification and regression tree (CART) analysis for indicator bacterial concentration prediction for a Californian coastal area. *Water Sci. Technol.* 61, 545–553. doi: 10.2166/wst.2010.842
- Bambic, D., McBride, G., Miller, W., Stott, R., and Wuertz, S. (2011). *Quantification of Pathogens and Sources of Microbial Indicators for QMRA in Recreational Waters. Vol. 10*. IWA Publishing.
- Baquero, F., Martínez, J. L., and Cantón, R. (2008). Antibiotics and antibiotic resistance in water environments. *Curr. Opin. Biotechnol.* 19, 260–265. doi: 10.1016/j.COPBIO.2008.05.006
- Benavides, J. A., Salgado-Caxito, M., Opazo-Capurro, A., González Muñoz, P., Piñero, A., Otto Medina, M., et al. (2021). ESBL-producing *Escherichia coli* carrying CTX-M genes circulating among livestock, dogs, and wild mammals in small-scale farms of Central Chile. *Antibiotics* 10:510. doi: 10.3390/antibiotics10050510
- Boelee, E., Geerling, G., van der Zaan, B., Blauw, A., and Vethaak, A. D. (2019). Water and health: From environmental pressures to integrated responses. *Acta Trop.* 193, 217–226. doi: 10.1016/j.actatropica.2019.03.011
- Boulesteix, A. L., Janitza, S., Hapfelmeier, A., Van Steen, K., and Strobl, C. (2015). Letter to the editor: On the term “interaction” and related phrases in the literature on random forests. *Brief. Bioinform.* 16, 338–345. doi: 10.1093/bib/bbu012
- Cacciò, M. S., Thompson, R. C. A., McLauchlin, J., and Smith, V. H. (2005). Unravelling *Cryptosporidium* and *Giardia* epidemiology. *Trends Parasitol.* 21, 430–437. doi: 10.1016/J.PT.2005.06.013
- Camara, M., Jamil, N. R., and Abdullah, A. F. Bin (2019). Impact of land uses on water quality in Malaysia: a review. *Ecol. Process.* 8, 1–10. doi: 10.1186/S13717-019-0164-X
- Cevadanes, A., Esperón, F., Di Cataldo, S., Neves, E., Sallaberry-Pincheira, N., and Millán, J. (2020). Antimicrobial resistance genes in Andean foxes inhabiting anthropized landscapes in Central Chile. *Sci. Total Environ.* 724:138247. doi: 10.1016/j.scitotenv.2020.138247
- Clinical and Laboratory Standards Institute (2018a). *Performance Standards for Antimicrobial Susceptibility Testing. 28th Edn*. Wayne, PA: Clinical and Laboratory Standards Institute.
- Clinical and Laboratory Standards Institute (2018b). CLSI Document M07Ed11E. Methods for Dilution Antimicrobial Susceptibility Tests for bacteria that grow Aerobically; Approved Standard—Eleven edition Guidelines. Available at: <https://clsi.org/standards/products/microbiology/documents/m07/> (Accessed June 1, 2022).
- Cui, Q., Huang, Y., Wang, H., and Fang, T. (2019). Diversity and abundance of bacterial pathogens in urban rivers impacted by domestic sewage. *Environ. Pollut.* 249, 24–35. doi: 10.1016/j.envpol.2019.02.094

- dela Peña, L. B. R. O., Vejano, M. R. A., and Rivera, W. L. (2021). Molecular surveillance of *Cryptosporidium* spp. for microbial source tracking of fecal contamination in Laguna Lake, Philippines. *J. Water Health* 19, 534–544. doi: 10.2166/wh.2021.059
- Devane, M. L., Weaver, L., Singh, S. K., and Gilpin, B. J. (2018). Fecal source tracking methods to elucidate critical sources of pathogens and contaminant microbial transport through New Zealand agricultural watersheds—A review. *J. Environ. Manag.* 222, 293–303. doi: 10.1016/j.jenvman.2018.05.033
- DGA, Dirección General de Aguas (2004a). Diagnóstico y clasificación de los cursos y cuerpos de agua Según objetivos de calidad. Cuenca del Río Maipo. Dir. Gen. aguas, 1–105. Available at: <https://mma.gob.cl/wp-content/uploads/2017/12/Maipo.pdf> (Accessed August 30, 2021).
- DGA, Dirección General de Aguas (2004b). Diagnóstico y clasificación de los cursos y cuerpos de agua según objetivos de calidad. Cuenca del Río Maule. 105. Available at: <http://portal.mma.gob.cl/wp-content/uploads/2017/12/Lluta.pdf> (Accessed August 30, 2021).
- Ding, J., Jiang, Y., Fu, L., Liu, Q., Peng, Q., and Kang, M. (2015). Impacts of land use on surface water quality in a Subtropical River basin: A case study of the Dongjiang River basin, southeastern China. *Water* 7, 4427–4445. doi: 10.3390/W7084427
- Domenech, E., Amorós, I., Moreno, Y., and Alonso, J. L. (2018). *Cryptosporidium* and *Giardia* safety margin increase in leafy green vegetables irrigated with treated wastewater. *Int. J. Hyg. Environ. Health* 221, 112–119. doi: 10.1016/j.ijheh.2017.10.009
- Encyclopedia Britannica (2021). Wastewater Treatment for Pollution Control. Water Cris. Available at: <https://www.britannica.com/explore/savingearth/wastewater-treatment> (Accessed August 30, 2021).
- Feng, Y., and Xiao, L. (2011). Zoonotic potential and molecular epidemiology of *Giardia* species and giardiasis. *Clin. Microbiol. Rev.* 24, 110–140. doi: 10.1128/CMR.00033-10
- Fenton, O., Tuohy, P., Daly, K., Moloney, T., Rice, P., and Murnane, J. G. (2021). A review of on-farm roadway runoff characterisation and potential management options for Ireland. *Water Air Soil Pollut.* 232, 1–22. doi: 10.1007/s11270-021-05027-0
- Field, K. G., and Samadpour, M. (2007). Fecal source tracking, the indicator paradigm, and managing water quality. *Water Res.* 41, 3517–3538. doi: 10.1016/j.watres.2007.06.056
- Ghorbani, Y., and Kuan, S. H. (2016). A review of sustainable development in the Chilean mining sector: past, present and future. *Int. J. Mining Reclam. Environ.* 31, 137–165. doi: 10.1080/17480930.2015.1128799
- González-Fernández, A., Symonds, E. M., Gallard-Gongora, J. F., Mull, B., Lukasik, J. O., Rivera Navarro, P., et al. (2021). Relationships among microbial indicators of fecal pollution, microbial source tracking markers, and pathogens in Costa Rican coastal waters. *Water Res.* 188:116507. doi: 10.1016/j.watres.2020.116507
- Goshu, G., Koelmans, A. A., and de Klein, J. J. M. (2021). Performance of faecal indicator bacteria, microbial source tracking, and pollution risk mapping in tropical water. *Environ. Pollut.* 276:116693. doi: 10.1016/j.envpol.2021.116693
- Green, H., Wilder, M., Wiedmann, M., and Weller, D. (2021). Integrative survey of 68 non-overlapping upstate New York watersheds reveals stream features associated with aquatic fecal contamination. *Front. Microbiol.* 12:684533. doi: 10.3389/fmicb.2021.684533
- Guo, J., and Lee, J. H. W. (2021). Development of predictive models for “very poor” beach water quality gradings using class-imbalance learning. *Environ. Sci. Technol.* 55(21), 14990–15000. doi: 10.1021/acs.est.1c03350
- Guyomard-Rabenirina, S., Dartron, C., Falord, M., Sadikalay, S., Ducat, C., Richard, V., et al. (2017). Resistance to antimicrobial drugs in different surface waters and wastewaters of Guadeloupe. *PLoS One* 12:e0173155. doi: 10.1371/journal.pone.0173155
- Hannan, A., and Anmala, J. (2021). Classification and prediction of fecal coliform in stream waters using decision trees (DTs) for upper Green River watershed, Kentucky, USA. *Water* 13:2790. doi: 10.3390/w13192790
- Higuera-Llantén, S., Vásquez-Ponce, F., Barrientos-Espinoza, B., Mardones, F. O., Marshall, S. H., and Olivares-Pacheco, J. (2018). Extended antibiotic treatment in salmon farms select multiresistant gut bacteria with a high prevalence of antibiotic resistance genes. *PLoS One* 13:e0203641. doi: 10.1371/journal.pone.0203641
- Hill, V. R., Polaczyk, A. L., Hahn, D., Narayanan, J., Cromeans, T. L., Roberts, J. M., et al. (2005). Development of a rapid method for simultaneous recovery of diverse microbes in drinking water by ultrafiltration with sodium polyphosphate and surfactants. *Appl. Environ. Microbiol.* 71, 6878–6884. doi: 10.1128/AEM.71.11.6878-6884.2005
- Holcomb, D. A., and Stewart, J. R. (2020). Microbial indicators of fecal pollution: recent Progress and challenges in assessing water quality. *Curr. Environ. Heal. Reports.* 7, 311–324. doi: 10.1007/s40572-020-00278-1
- INE, Instituto Nacional de Estadísticas (2017). Resultado Censo 2017 para la región del Maule. Available at: <http://resultados.censo2017.cl/Region?R=R07> (Accessed 11 November 2020).
- INN, Instituto Nacional de Normalización (1978). Requisitos de calidad del agua para diferentes usos (NCh1333.Of78). Chile Available at: [https://ciperchile.cl/pdfs/11-2013/norovirus/NCh1333-1978\\_Mod-1987.pdf](https://ciperchile.cl/pdfs/11-2013/norovirus/NCh1333-1978_Mod-1987.pdf) (Accessed November 11, 2020).
- Instituto Geográfico Militar de Chile (2005). *Atlas de la República de Chile*. Santiago, Chile.
- Instituto Nacional de Normalización (INN) (1995). NCh2313/22: 1995. Industrial and sewage effluents—Test methods—Part 22: Determination of faecal coliform organisms in EC medium. Available at: <https://ecommerce.inn.cl/nch231322199542212> (Accessed: March 21, 2022)
- Javanmard, E., Mirsamadi, E. S., Olfatfar, M., Ghasemi, E., Saki, F., Mirjalali, H., et al. (2020). Prevalence of *Cryptosporidium* and *Giardia* in vegetables in Iran: a nineteen-years meta-analysis review. *J. Environ. Health Sci. Eng.* 18, 1629–1641. doi: 10.1007/s40201-020-00493-w
- Jung, B., and Hoilat, G. J. (2020). MacConkey medium. StatPearls Publishing Available at: <https://www.ncbi.nlm.nih.gov/books/NBK557394/> (Accessed August 30, 2021).
- Khan, M. N., Mobin, M., Abbas, Z. K., and Alamri, S. A. (2018). Fertilizers and their contaminants in soils, surface and groundwater. *Anthr.* 5, 225–240. doi: 10.1016/B978-0-12-409548-9.09888-2
- Kitajima, M., Haramoto, E., Iker, B. C., and Gerba, C. P. (2014). Occurrence of *Cryptosporidium*, *Giardia*, and *Cyclospora* in influent and effluent water at wastewater treatment plants in Arizona. *Sci. Total Environ.* 484, 129–136. doi: 10.1016/j.scitotenv.2014.03.036
- Korajkic, A., McMin, B. R., and Harwood, V. J. (2018). Relationships between microbial indicators and pathogens in recreational water settings. *Int. J. Environ. Res. Public Health* 15:2842. doi: 10.3390/ijerph15122842
- Kraemer, S. A., Ramachandran, A., and Perron, G. G. (2019). Antibiotic pollution in the environment: From microbial ecology to public policy. *Microorganisms* 7:180. doi: 10.3390/microorganisms7060180
- Kuhn, M., and Johnson, K. (2016). *Applied Predictive Modeling*. New York: Springer Nature.
- Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096
- LaGro, J. A. (2005). “Land-use classification,” in *Encyclopedia of Soils in the Environment*. ed. D. Hillel (Oxford: Elsevier), 321–328.
- Lenart-Boroń, A., Wolanin, A., Jelonekiewicz, E., and Żelazny, M. (2017). The effect of anthropogenic pressure shown by microbiological and chemical water quality indicators on the main rivers of Podhale, southern Poland. *Environ. Sci. Pollut. Res.* 24, 12938–12948. doi: 10.1007/s11356-017-8826-7
- Lewis, D. J., Atwill, E. R., Lennox, M. S., Hou, L., Karle, B., and Tate, K. W. (2005). Linking on-farm dairy management practices to storm-flow fecal coliform loading for California coastal watersheds. *Environ. Monit. Assess.* 107, 407–425. doi: 10.1007/s10661-005-3911-7
- Lusk, M. G., Toor, G. S., Yang, Y. Y., Mechtensimer, S., De, M., and Obreza, T. A. (2017). A review of the fate and transport of nitrogen, phosphorus, pathogens, and trace organic chemicals in septic systems. *Crit. Rev. Environ. Sci. Technol.* 47, 455–541. doi: 10.1080/10643389.2017.1327787
- Magiorakos, A. P., Srinivasan, A., Carey, R. B., Carmeli, Y., Falagas, M. E., Giske, C. G., et al. (2012). Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance. *Clin. Microbiol. Infect.* 18, 268–281. doi: 10.1111/j.1469-0691.2011.03570.x
- Martínez, M. C., Retamal, P., Rojas-Aedo, J. F., Fernández, J., Fernández, A., and Lapierre, L. (2017). Multidrug-resistant outbreak-associated *Salmonella* strains in irrigation water from the metropolitan region, Chile. *Zoonoses Public Health* 64, 299–304. doi: 10.1111/zph.12311
- McEgan, R., Mootian, G., Goodridge, L. D., Schaffner, D. W., and Danyluk, M. D. (2013). Predicting *Salmonella* populations from biological, chemical, and

- physical indicators in Florida surface waters. *Appl. Environ. Microbiol.* 79, 4094–4105. doi: 10.1128/AEM.00777-13
- Miller, W. A., Atwill, E. R., Gardner, I. A., Miller, M. A., Fritz, H. M., Hedrick, R. P., et al. (2005). Clams (*Corbicula fluminea*) as bioindicators of fecal contamination with *Cryptosporidium* and *Giardia* spp. in freshwater ecosystems in California. *Int. J. Parasitol.* 35, 673–684. doi: 10.1016/j.ijpara.2005.01.002
- Ministerio Secretaría General de la Presidencia (2001) Decreto 90 establece norma de emisión para la regulación de contaminantes asociados a las descargas de residuos líquidos a aguas marinas y continentales superficiales. Available at: <https://www.bcn.cl/leychile/navegar?idNorma=182637> (Accessed: August 6, 2021).
- Monis, P. T., Caccio, S. M., and Thompson, R. C. A. (2009). Variation in *Giardia*: towards a taxonomic revision of the genus. *Trends Parasitol.* 25, 93–100. doi: 10.1016/j.pt.2008.11.006
- Morgan, U. M., Constantine, C. C., Forbes, D. A., and Thompson, R. C. (1997). Differentiation between human and animal isolates of *Cryptosporidium parvum* using rDNA sequencing and direct PCR analysis. *J. Parasitol.* 83, 825–830. doi: 10.2307/3284275
- Nasser, A. M. (2016). Removal of *Cryptosporidium* by wastewater treatment processes: a review. *J. Water Health* 14, 1–13. doi: 10.2166/wh.2015.131
- Newton, R. J., and McClary, J. S. (2019). The flux and impact of wastewater infrastructure microorganisms on human and ecosystem health. *Curr. Opin. Biotechnol.* 57, 145–150. doi: 10.1016/j.copbio.2019.03.015
- Nnadozie, C. E., and Odume, O. N. (2019). Freshwater environments as reservoirs of antibiotic resistant bacteria and their role in the dissemination of antibiotic resistance genes. *Environ. Pollut.* 254:113067. doi: 10.1016/j.envpol.2019.113067
- Pandey, P. K., Kass, P. H., Soupir, M. L., Biswas, S., and Singh, V. P. (2014). Contamination of water resources by pathogenic bacteria. *AMB Express* 4, 51–16. doi: 10.1186/s13568-014-0051-x
- Papini, R. A., and Verin, R. (2019). *Giardia* and *Cryptosporidium* in red foxes (*Vulpes vulpes*): screening for coproantigens in a population of Central Italy and mini-review of the literature. *Maced. Vet. Rev.* 42, 101–106. doi: 10.2478/macvetrev-2019-0013
- Pilmis, B., Monnier, A. Le, and Zahar, J.-R. (2020). Gut microbiota, antibiotic therapy and antimicrobial resistance: A narrative review. *Microorganisms* 8, 269. doi: 10.3390/MICROORGANISMS8020269.
- Prystajecy, N., Huck, P. M., Schreier, H., and Isaac-Renton, J. L. (2014). Assessment of *Giardia* and *Cryptosporidium* spp. as a microbial source tracking tool for surface water: application in a mixed-use watershed. *Appl. Environ. Microbiol.* 80, 2328–2336. doi: 10.1128/AEM.02037-13
- R Core Team (2020). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: <https://www.r-project.org/> (Accessed May 27, 2022).
- RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA. Available at: <http://www.rstudio.com/> (Accessed June 8, 2022).
- Raza, S., Kim, J., Sadowsky, M. J., and Unno, T. (2021). Microbial source tracking using metagenomics and other new technologies. *J. Microbiol.* 59, 259–269. doi: 10.1007/S12275-021-0668-9
- Read, C. M., Monis, P. T., and Thompson, R. C. A. (2004). Discrimination of all genotypes of *Giardia duodenalis* at the glutamate dehydrogenase locus using PCR-RFLP. *Infect. Genet. Evol.* 4, 125–130. doi: 10.1016/j.meegid.2004.02.001
- Ryu, H., Tran, H., Ware, M. W., Iker, B., Griffin, S., Egorov, A., et al. (2011). Application of leftover sample material from waterborne protozoa monitoring for the molecular detection of *Bacteroidales* and fecal source tracking markers. *J. Microbiol. Methods* 86, 337–343. doi: 10.1016/j.mimet.2011.06.001
- Sbodio, A., Maeda, S., Lopez-Velasco, G., and Suslow, T. V. (2013). Modified Moore swab optimization and validation in capturing *E. coli* O157: H7 and *Salmonella enterica* in large volume field samples of irrigation water. *Food Res. Int.* 51, 654–662. doi: 10.1016/j.foodres.2013.01.011
- Shapiro, K., Silver, M., Byrne, B. A., Berardi, T., Aguilar, B., Melli, A., et al. (2018). Fecal indicator bacteria and zoonotic pathogens in marine snow and California mussels (*Mytilus californianus*). *FEMS Microbiol. Ecol.* 94:fy172. doi: 10.1093/femsec/fy172
- Strawn, L. K., Fortes, E. D., Bihn, E. A., Nightingale, K. K., Gröhn, Y. T., Worobo, R. W., et al. (2013). Landscape and meteorological factors affecting prevalence of three food-borne pathogens in fruit and vegetable farms. *Appl. Environ. Microbiol.* 79, 588–600. doi: 10.1128/AEM.02491-12
- Strobl, C., Boulesteix, A.-L., Zeileis, A., and Hothorn, T. (2007a). Bias in random forest variable importance measures: illustrations, sources and a solution. *BMC Bioinformatics* 8, 25. doi: 10.1186/1471-2105-8-25
- Strobl, C., Boulesteix, A. L., Zeileis, A., and Hothorn, T. (2007b). Variable importance measures: illustrations, sources and a solution. *BMC Bioinformatics* 8, 8–25. doi: 10.1186/1471-2105-8-25
- Strobl, C., Hothorn, T., and Zeileis, A. (2009). Party on! A new, conditional variable importance measure for random forests available in the party package. *Contrib. Res. J.* 1, 14–17. doi: 10.32614/RJ-2009-013
- Taylor, N. G. H., Verner-Jeffreys, D. W., and Baker-Austin, C. (2011). Aquatic systems: maintaining, mixing and mobilising antimicrobial resistance? *Trends Ecol. Evolution* 26, 278–284. doi: 10.1016/j.tree.2011.03.004
- van Hoek, A. H. A. M., Veenman, C., van Overbeek, W. M., Lynch, G., de Roda Husman, A. M., and Blaak, H. (2015). Prevalence and characterization of ESBL- and AmpC-producing *Enterobacteriaceae* on retail vegetables. *Int. J. Food Microbiol.* 204, 1–8. doi: 10.1016/j.ijfoodmicro.2015.03.014
- Verhoughstraete, M. P., Martin, S. L., Kendall, A. D., Hyndman, D. W., and Rose, J. B. (2015). Linking fecal bacteria in rivers to landscape, geochemical, and hydrologic factors and sources at the basin scale. *Proc. Natl. Acad. Sci.* 112, 10419–10424. doi: 10.1073/pnas.1415836112
- Weller, D., Belias, A., Green, H., Roof, S., and Wiedmann, M. (2020a). Landscape, water quality, and weather factors associated with an increased likelihood of foodborne pathogen contamination of New York streams used to source water for produce production. *Front. Sustain. Food Syst.* 3:124. doi: 10.3389/fsufs.2019.00124
- Weller, D., Brassill, N., Rock, C., Ivanek, R., Mudrak, E., Roof, S., et al. (2020b). Complex interactions between weather, and microbial and physicochemical water quality impact the likelihood of detecting foodborne pathogens in agricultural water. *Front. Microbiol.* 11:134. doi: 10.3389/fmicb.2020.00134
- Weller, D. L., Love, T. M. T., Belias, A., and Wiedmann, M. (2020c). Predictive models may complement or provide an alternative to existing strategies for assessing the enteric pathogen contamination status of northeastern streams used to provide water for produce production. *Front. Sustain. Food Syst.* 4:561517. doi: 10.3389/fsufs.2020.561517
- Weller, D. L., Love, T. M. T., and Wiedmann, M. (2021). Comparison of resampling algorithms to address class imbalance when developing machine learning models to predict foodborne pathogen presence in agricultural water. *Front. Environ. Sci.* 9:701288. doi: 10.3389/fenvs.2021.701288
- Weller, D., Shiwakoti, S., Bergholz, P., Grohn, Y., Wiedmann, M., and Strawn, L. K. (2016). Validation of a previously developed geospatial model that predicts the prevalence of *Listeria monocytogenes* in New York state produce fields. *Appl. Environ. Microbiol.* 82, 797–807. doi: 10.1128/AEM.03088-15
- Wen, Y., Schoups, G., and Giesen, N. van de (2017). Organic pollution of rivers: combined threats of urbanization, livestock farming and global climate change. *Sci. Rep.* 7, 1–9. doi: 10.1038/srep43289.
- WHO and UNICEF, World Health Organization and United Nations Children's Fund (2017). Progress on drinking water, sanitation and hygiene. 114. Available at: <https://apps.who.int/iris/handle/10665/258617> (Accessed July 22, 2021).
- WHO, World Health Organization (2017a). Diarrhoeal disease, key facts. Key facts. Available at: <https://www.who.int/news-room/fact-sheets/detail/diarrhoeal-disease> (Accessed July 22, 2021).
- WHO, World Health Organization (2017b). Don't pollute my future! The impact of the environment on children's health. 33. Available at: <https://apps.who.int/iris/bitstream/handle/10665/254678/WHO-FWC-IHE-17.01-eng.pdf;jsessionid=A9C9B73B98D9B6B36C9C81ABA979E707> (Accessed July 22, 2021).
- Wilkes, G., Edge, T. A., Gannon, V. P. J., Jokinen, C., Lyautey, E., Neumann, N. F., et al. (2011). Associations among pathogenic bacteria, parasites, and environmental and land use factors in multiple mixed-use watersheds. *Water Res.* 45, 5807–5825. doi: 10.1016/j.watres.2011.06.021
- Wu, J., Long, S. C., Das, D., and Dorner, S. M. (2011). Are microbial indicators and pathogens correlated? A statistical analysis of 40 years of research. *J. Water Health* 9, 265–278. doi: 10.2166/wh.2011.117
- Xiao, L., Alderisio, K., Limor, J., Royer, M., and Lal, A. A. (2000). Identification of species and sources of *Cryptosporidium* oocysts in storm waters with a small-subunit rRNA-based diagnostic and genotyping tool. *Appl. Environ. Microbiol.* 66, 5492–5498. doi: 10.1128/AEM.66.12.5492-5498.2000

- Xiao, L., Lihua, R., Ryan, U., and Upton, S. J. (2004). *Cryptosporidium* taxonomy: recent advances and implications for public health. *Clin. Microbiol. Rev.* 17, 72–97. doi: 10.1128/CMR.17.1.72-97.2004
- Xiao, L., Morgan, U. M., Limor, J., Escalante, A., Arrowood, M., Shulanw, W., et al. (1999). Genetic diversity within *Cryptosporidium parvum* and related *Cryptosporidium* species. *Appl. Environ. Microbiol.* 65, 3386–3391. doi: 10.1128/AEM.65.8.3386-3391.1999
- Xiao, S., Zhang, Y., Zhao, X., Sun, L., and Hu, S. (2018). Presence and molecular characterization of *Cryptosporidium* and *Giardia* in recreational lake water in Tianjin, China: a preliminary study. *Sci. Rep.* 8, 2353–2358. doi: 10.1038/s41598-018-20902-3
- Yang, Y., Song, W., Lin, H., Wang, W., Du, L., and Xing, W. (2018). Antibiotics and antibiotic resistance genes in global lakes: A review and meta-analysis. *Environ. Int.* 116, 60–73. doi: 10.1016/j.envint.2018.04.011
- Yuan, T., Vadde, K. K., Tonkin, J. D., Wang, J., Lu, J., Zhang, Z., et al. (2019). Urbanization impacts the physicochemical characteristics and abundance of fecal markers and bacterial pathogens in surface water. *Int. J. Environ. Res. Public Health* 16, 1–19. doi: 10.3390/ijerph16101739
- Zhang, S., Tao, W., Liu, C., Jiang, Y., Wan, Q., Li, Q., et al. (2016). First report of *Cryptosporidium canis* in foxes (*Vulpes vulpes*) and raccoon dogs (*Nyctereutes procyonoides*) and identification of several novel subtype families for *Cryptosporidium* mink genotype in minks (*Mustela vison*) in China. *Infect. Genet. Evol.* 41, 21–25. doi: 10.1016/j.meegid.2016.03.016
- Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Díaz-Gavidia, Barría, Weller, Salgado-Caxito, Estrada, Araya, Vera, Smith, Kim, Moreno-Switt, Olivares-Pacheco and Adell. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.