Elucidation of fecal inputs into the River Tagus catchment (Portugal) using source-specific mitochondrial DNA, HAdV, and phage markers

S. Monteiro[a,b,*], J. Ebdon[a], R. Santos[b], H. Taylor[a]

[a] *School of Environment and Technology, University of Brighton, Brighton, UK*

[b] *Laboratorio Analises, Instituto Superior Tecnico, Lisbon, Portugal*

Keywords: Water quality; remediation; fecal indicators; source tracking; mitochondrial markers; viruses

[*] Corresponding author at: Laboratorio Analises, Instituto Superior Tecnico, Avenida Rovisco Pais, 1049-001 Lisbon, Portugal.
E-mail address: silvia.monteiro@tecnico.ulisboa.pt

ABSTRACT

Determining the source of fecal contamination in a water body is important for the application of appropriate remediation measures. However, it has been suggested in the extant literature that this can best be achieved using a 'toolbox' of molecular- and culture-based methods. In response, this study deployed three indicators (*Escherichia coli* (EC), intestinal enterococci (IE) and somatic coliphages (SC)), one culture-dependent human marker (*Bacteroides* (GB-124) bacteriophage) and five culture-independent markers (human adenovirus (HAdV), human (HMMit), cattle (CWMit), pig (PGMit) and poultry (PLMit) mitochondrial DNA markers (mtDNA)) within the River Tagus catchment (n = 105). Water samples were collected monthly over a 13-month sampling campaign at four sites (impacted by significant specific human and non-human inputs and influenced by differing degrees of marine and freshwater mixing) to determine the dominant fecal inputs and assess geographical, temporal, and meteorological (precipitation, UV, temperature) fluctuations. Our results revealed that all sampling sites were not only highly impacted by fecal contamination but that this contamination originated from human and from a range of agricultural animal sources. HMMit was present in a higher percentage (83%) and concentration (4.20 log GC/100 mL) than HAdV (32%, 2.23 log GC/100 mL) and GB-124 bacteriophage with the latter being detected once. Animal mtDNA markers were detected, with CWMit found in 73% of samples with mean concentration of 3.74 log GC/100 mL. Correlation was found between concentrations of fecal indicators (EC, IE and SC), CWMit and season. Levels of CWMit were found to be related to physico-chemical parameters, such as temperature and UV radiation, possibly as a result of the increasing presence of livestock outside in warmer months. This study provides the first evaluation of such a source-associated 'toolbox' for monitoring surface water in Portugal, and the

conclusions may inform future implementation of surveillance and remediation strategies for improving water quality.

1. Introduction

Whilst water is undoubtedly one of the most important resources on the planet, it is also a major vehicle for disease transmission, continuing to cause 1.7 billion cases of childhood diarrheal disease and 525,000 preventable deaths/year (mainly in infants under 5 years of age) (WHO, 2017).

In addition, according to the existing climate change scenario, by 2030, water scarcity in arid and semi-arid places will force the reallocation of 24 to 700 million people (World Water Assessment Programme, 2009; Hameeteman, 2013).

It is, therefore, crucial to ensure freshwater sources are kept free of fecal contamination, especially human sources as they are more likely to contain microorganisms adapted to infect human hosts. As such, water contaminated with human fecal pollution may present significant risks to human health as they may contain pathogens, such as enteric viruses (Noroviruses, Hepatitis A viruses, and Enteroviruses), enteric bacteria (*Salmonella enterica*, *Shigella* spp., *Campylobacter* spp.) and/or protozoa (*Cryptosporidium* spp. and *Giardia lamblia*) (Seidel *et al*., 2016; Haramoto *et al*., 2018; Holcomb *et al*., 2020). Nonetheless, non-human animal sources may also act as reservoirs for certain 'zoonotic' pathogens, including *Escherichia coli* O157 H7, Hepatitis E virus and *Cryptosporidium* (Cotruvo *et al*., 2004).

Traditional fecal indicator bacteria (FIB), e.g. *Escherichia coli* (EC) and intestinal enterococci (IE), have been used during the last century to monitor microbial water quality (Ashbolt *et*., 2001). However, it is well known that most of these microorganisms are not host-associated (i.e. not limited to humans) but also exist in the intestines of other warm-blooded animals (Scott *et al*., 2002). Therefore, such organisms cannot be used to predict the source of fecal contamination and to accurately understand which pathogens may be present, or where they originate from. Epidemiological studies, when contamination is mainly from non-point sources, have also failed to correlate risks to human health and FIB (Dwight *et al*., 2004; Colford *et al*., 2007). This may, in part, be attributed to the fact that whilst EC and IE are able to inhabit different warm-blooded animals, many pathogens are limited to human hosts, especially enteric viruses (Harwood *et al*., 1999). Additionally, several strains of FIB have been shown to be able to become saprophytes, persisting in certain habitats, including soils, aquatic sediments, and vegetation (Ishii *et al*., 2006; Ksoll *et al*., 2007; Badgley *et al*., 2011). Therefore, the role of FIB in determining effective management of water quality can significantly improve by employing new methodologies that allow for the discrimination of the sources of these microorganisms. Source tracking (ST) methods have been shown to provide rapid and efficient fecal source discrimination in order to expedite cost-effective remediation actions (Ahmed *et al*., 2019; Kongprajug *et al*., 2019). Several ST methods have been tested previously including anaerobic bacteria (*Bifidobacterium* and *Bacteroides*), enteric viruses and bacteriophages and mitochondrial DNA (mtDNA). For instance, human adenovirus (HAdV), a member of the enteric viruses family, has been previously shown to have the potential to determine human sources of fecal contamination in different settings. HAdV is highly specific to human fecal contamination and is detected throughout the year at high

86 concentrations. On the other hand, bacteriophages infecting a specific host strain
87 (such as *Bacteroides fragilis* GB-124) have been described as potential markers to
88 evaluate fecal contamination from human origin (Payan *et al*., 2005; Ebdon *et al*.,
89 2012; McMinn *et al*., 2014). Such bacteriophages have been detected throughout the
90 world in relatively high concentrations in municipal wastewaters worldwide and their
91 temporal and geographic stability has already been addressed in several studies
92 (Ebdon *et al*., 2012; Jofre *et al*., 2014; McMinn *et al*., 2014). Mitochondrial DNA
93 markers have been used to a less extent to determine the sources of fecal
94 contamination although their potential as markers of fecal contamination in
95 environmental waters is considerable. Feces contain substantial amounts of exfoliated
96 epithelial cells and mtDNA is present in many copies per cell (Iyengar *et al*., 1991;
97 Gerber *et al*., 2001; Andreasson *et al*., 2002). Such feature ensures a robust PCR
98 signal in accordance with 16S rRNA genes. Targeting directly mtDNA implies
99 identifying the animal species directly rather than microbial species they host.
100 Nonetheless, further studies are necessary to understand the relevance of carry-over
101 on the environmental presence of a certain mtDNA marker. Carry-over was
102 determined by Caldwell *et al.* (2007) in half of the participants of the study who
103 consumed beef, the concentration was 1-2 orders below than that of the human target,
104 Martellini *et al*. (2005) were unable to detected animal signals from the feces of the
105 human volunteer who had eaten meat products the previous day. Therefore, such
106 results seem to indicate that the impact of carry-over in environmental waters where
107 the dilution effect is an important factor may be negligible.
108 Information on the environmental persistence of mtDNA markers is also sparse. The
109 persistence of these markers in was determined previously but in a microcosmos
110 setting, not evaluating directly in environmental waters where interactions with the

111 different physico-chemical and biological processes found in the environment are not
112 evaluated (Martellini *et al*., 2005; He *et al*., 2016).
113 In this study, we describe a new ST 'toolbox' of methods comprising both culture- and
114 molecular-dependent approaches and its application to identify sources within the R.
115 Tagus catchment. The objectives of the study were: (i) to compare the performance
116 (for the first time) of animal-associated mitochondrial DNA markers (cattle (CWMit),
117 pig (PGMit) and poultry (PLMit) fecal contamination) and human-associated markers,
118 including a mitochondrial marker (HMMit), HAdV and GB-124 bacteriophage; (ii) to
119 evaluate the seasonality of each marker; (iii) to study, in a real catchment, the
120 relationship of each MST marker to several environmental conditions, including water
121 and environmental temperature, salinity, and UV radiation, to have a better
122 understanding of the persistence of these markers in environmental freshwaters,
123 particularly those impacted by marine waters, and (iv) to determine the influence of
124 the occurrence of rain and precipitation levels on the presence of these markers in
125 the R. Tagus catchment. This study was design to provide new knowledge on a set of
126 less common MST markers and their potential usefulness in a "toolbox"-like approach
127 for pollution surveillance and aiding in the development of strategies to enhance the
128 water quality of the R. Tagus.
129

130 ## 2. Materials and methods

131 ### 2.1. Sample location

132 Water samples were collected at four sampling sites in the area of the R. Tagus
133 catchment close to the city of Lisbon (population of 2.27 x 10^6 inhabitants) (Fig. 1; INE,
134 2019). All samples were collected between February 2015 and March 2016 (13
135 months) at approximately one-month intervals (n= 105). The four sampling points

selected were: site (1) Vila Franca de Xira (VFX), site (2) Marina Parque das Nações (MNdP), site (3) Alcântara (Alc), and site (4) Belém (Bel) (Fig. 1). Sampling site 1 was selected because of its location at the upper most extent of the tidal limit. Water quality at this sampling site in particular is not impacted by waters from the Atlantic Ocean during high tide, a situation quite distinct from the other sampling sites. Moreover, site 1 is a well-known location for the rearing of livestock (mainly cattle, but also pigs) and for bulls used for bull fighting. Sampling site 2 was chosen because of its proximity to the discharge of two significant wastewater treatment plants (WWTP 3 and 4, Fig. 1). WWTP 4 receives wastewater from livestock farms and discharges into one of the R. Tagus tributaries (Rio Trancão). It is important to understand whether the Rio Trancão increases the contamination in R. Tagus and if so, which sources are the most important to control. This area is also highly urbanized and used for recreational purposes such as dog walking. Sampling site 3 was chosen because of its proximity to the discharge point of the largest WWTP in Lisbon (WWTP 6 Pop. Equiv = 800,000). This WWTP receives and treats effluents from the majority of the Lisbon metropolitan area and is therefore a main source of human fecal contamination within the catchment. This sampling site is also highly influenced by oceanic mixing with water from the Atlantic Ocean during high tide. Finally, sampling site 4 was selected because of its popularity as a recreational site and the fact that it is heavily influenced by the Atlantic Ocean. Sampling was conducted between February 2015 to March 2016, in order to understand the influence of seasonality on the presence and distribution of source-associated markers.

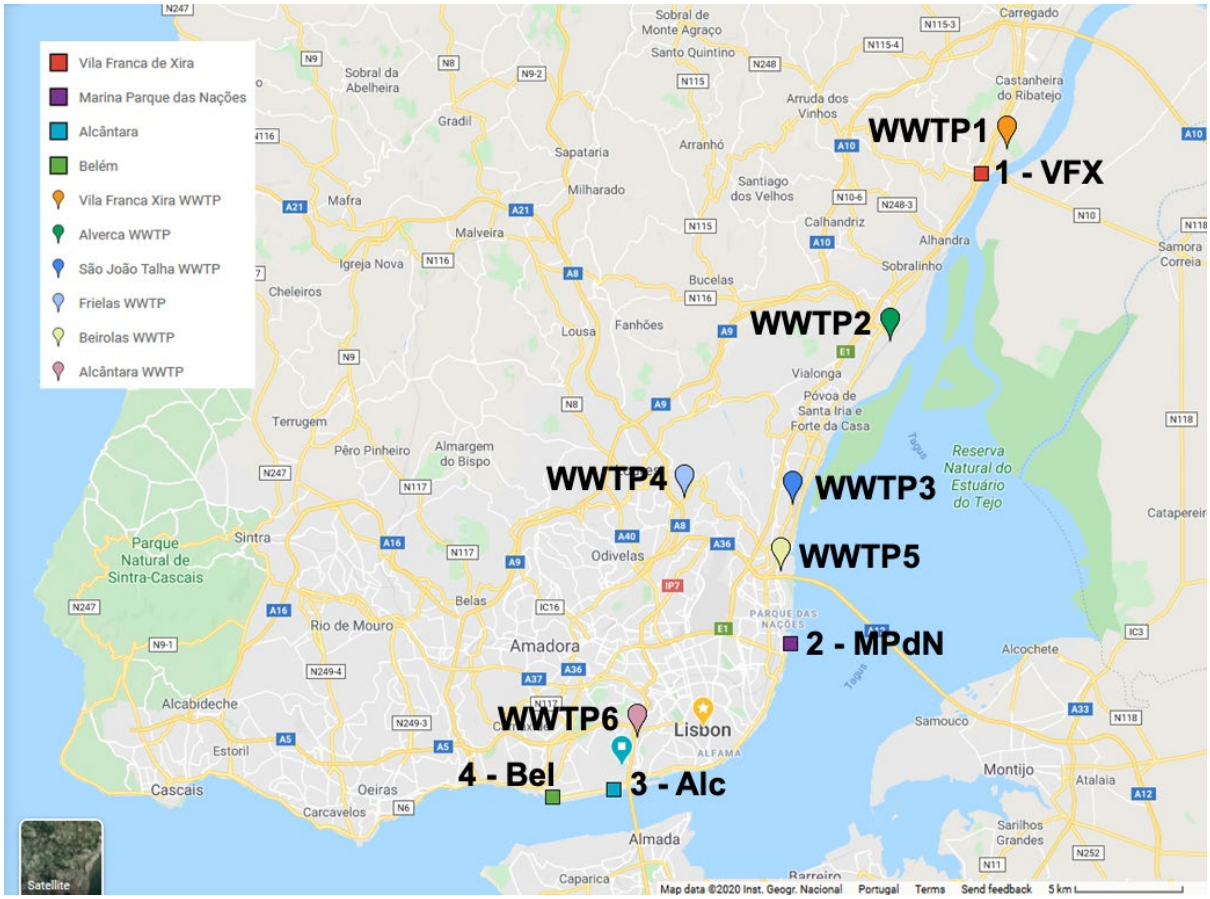


Fig. 1. The R. Tagus (Rio Tejo) catchment, sampling sites and location of wastewater treatment plants (WWTPs) (Google, 2020). Balloon symbols represent the six WWTPs impacting R. Tagus and the squares represent the four sampling locations.

2.2. Sample collection

In this study, one human-targeted culture-dependent assay (GB-124) phage, two human (HMMit and HAdV) and three animal qPCR-based assays (cattle - CWMit; pig – PGMit; and poultry – PLMit) were tested (Ebdon *et al*., 2007; Schill and Mathes, 2008; Rusiñol *et al*., 2014). Moreover, indicators of fecal contamination, EC, IE and SC were also determined. At each site, 2 L of river water was collected in 2 x 1 L sterile polyethylene bottles (Vidrolab 2, Portugal) for fecal indicator bacteria and bacteriophage detection and enumeration. On each occasion another 2 L sample was collected for the concentration of mtDNA markers, along with a 10 L grab sample collected in a 10 L sterile polyethylene bottle (Vidrolab 2, Portugal) for the analysis of

HAdV. All samples were then transported to the laboratory facilities at (5 ± 3) ºC within a period of 4 hours after collection and analyzed or treated immediately upon arrival to the laboratory.

Stool samples from animals were obtained from the Biblioteca/Bedeteca dos Olivais. One human volunteer at the laboratory provided stool samples. All fecal samples were frozen at -30 ºC. Stool samples were used to obtain the genomic material for the construction of standard curves for the quantification of each mtDNA marker and for the creation of positive controls for the qPCR reactions.

2.3. Fecal indicator bacteria

EC and IE were enumerated in river water samples by the Most Probable Number (MPN) method using the Colilert and Enterolert systems (IDEXX Laboratories, USA), according to standard methods (ASTM, 1999; UK Environment Agency, 2009). Briefly, each sample, or dilution thereof was poured into a 100 mL sterile plastic bottle and mixed with sterile H_2O up to a final volume of 100 mL. The appropriate substrate was added to the 100 mL bottle and the mixture was allowed to settle until dissolved. The contents were then poured into a Quanti-Tray® and sealed, then incubated at (37 ± 1) ºC for 18-22 h and at (41.0 ± 0.5) ºC for 18-22 h, for EC and IE, respectively. Positive and negative controls were run for all tests and samples were analyzed in duplicate. Following the incubation period, the trays were observed under UV light and wells showing fluorescence were counted.

2.4. Bacteriophages

Somatic coliphages (SC) were chosen for their potential as viral indicators of fecal pollution, whereas *B. fragilis* (strain GB-124) bacteriophages were tested as potential

198 human-associated markers, as they have shown to be limited to human sources in the
199 UK and US (Payan *et al*., 2005; Ebdon *et al*., 2007; McMinn *et al*., 2014, 2017). The
200 detection and enumeration of bacteriophages was performed in duplicate using the
201 standardized double-layer methods specific for the detection of SC and GB-124
202 phages (ISO 2000, 2001).
203 Water samples were analyzed directly for SC (as they are more abundant than source-
204 associated markers) whereas 2 L filtrations were performed for GB-124 phages
205 (Mendez *et al*., 2004). For the determination of GB-124 phages, river samples were
206 amended to a final concentration of 0.05M $MgCl_2$, filtered through a 47 mm 0.22 µm
207 mixed cellulose esters membrane filter (Whatman, GE Healthcare, US), after which
208 the membranes were cut into eight fragments and placed in a flask containing 5 mL of
209 elution buffer (3% w/v beef extract, 3% v/v Tween 80, and 0.5M NaCl, final pH 9.0).
210 The eluate was placed in an ultrasonic bath for 5 min. The membrane was removed
211 and the eluate was filtered through a 0.22 µm polyvinylidene fluoride (PVDF) filter
212 (PALL, UK) with low protein binding (Tartera *et al*., 1992). This constitutes the
213 concentrated sample. One mL of the *Bacteroides* GB-124 inoculum culture was added
214 to 2.5 mL of semi-solid agar medium (*Bacteroides* Phage Recovery Medium agar
215 (BPRM); 1% agar). One mL of concentrated sample was added subsequently to the
216 mixture. The latter was slowly vortexed in order to avoid the formation of air bubbles,
217 poured over BPRM with 2% agar, and left to set. The plates were then inverted and
218 placed in anaerobic jars containing an anaerobic sachet (Oxoid, UK).
219 SC were determined by adding 1 mL of EC inoculum culture to 2.5 mL of semi-solid
220 agar medium (Modified Scholtan's Agar (MSA)) and 1 mL of samples pre-filtered using
221 a 0.22 µm PVDF filter, vortexed, and poured over a layer of solid MSA and left to set.
222 For both phages the plates were incubated at (36 ± 2) ºC for 18 (± 2) h.

2.5. Molecular-based assays

For the concentration of mtDNA, 2 L river water samples were centrifuged at 9000 xg for 15 min and the supernatant carefully removed until all the volume was centrifuged. The pellet was then resuspended in 10 mL of sample (Martellini *et al*., 2005).

To concentrate HAdV from 10 L river water, a method based on direct organic flocculation was used, with the adhesion of viral particles to pre-flocculated skimmed milk (Calgua *et al*., 2008). In brief, the pH of the samples was pre-regulated to pH 3.5 by adding 1 N HCl. Five mL of pre-flocculated 1 % (w/v) skimmed milk was then added to the sample, so that the final concentration of skimmed milk was 0.01 % (w/v). The samples were stirred at room temperature for 8 h. The material in suspension was allowed to settle for a further 8 h and the supernatant was carefully removed. The flocculated sample was then centrifuged at 12 ºC for 30 min at 7 000 xg. Following centrifugation, the supernatant was carefully removed and the sediment was resuspended in 8 mL of 0.2 M phosphate buffer (0.2 M of sodium monohydrogen phosphate (Na_2HPO_4) and a solution of sodium phosphate (NaH_2PO_4) at a proportion of 1:1; pH 7.5). The volume of the resuspended viral particles was finally adjusted to 10 mL. The eluates for mtDNA and HAdV analysis were kept at -80 ºC until analysis.

2.6. Nucleic acid extraction

DNA from the concentrated river water samples for analysis of HAdV was extracted using QIAamp Viral RNA Mini Kit (Qiagen, Germany), DNA from river samples for the detection of mtDNA was extracted using QIAamp DNA Blood Mini kit (Qiagen, Germany), and DNA to generate the positive controls and standard curves for mtDNA were extracted from stools using QIAamp Fast DNA Stool mini kit (Qiagen, Germany)

248 according to the manufacturer's instructions. Final eluted volume was 200 µL for river
249 for mtDNA extraction (in the river and stool samples) and 80 µL for HAdV extraction.
250 The samples were kept at (-30 ± 5) ºC until further analysis, within 6 months.
251
252 **2.7. Real-time PCR**

253 Primers and probe sequences of the human, bovine, pig and poultry markers used are
254 shown in Table S1. The amplifications were performed in 25 µL of reaction mixture
255 using the TaqMan® Universal Master Mix II (Applied Biosystems, US). A volume of
256 12.5 µL of master mix was mixed with the respective concentration of primers and
257 probes for each assay (forward and reverse primers – 800 nM, probes – 200 nM). The
258 volume was adjusted to 20 µL with sterile DNA and RNA-free water. In addition, 10-
259 fold and 100-fold dilutions of every DNA extraction were also assayed. The PCR
260 reactions were carried out in a 7300 Real-Time PCR System (Applied Biosystems,
261 US) according to the original manuscript (Hernroth *et al*., 2002; Schill and Mathes,
262 2008). For HAdV, the first step occurs at 50 ºC for 2 min, followed by an activation
263 step at 95 ºC for 10 min, 40 cycles of denaturation at 95 ºC for 15 sec and annealing
264 at 60 ºC for 1 min. For all mtDNA assays, the first step consisted of an activation step
265 at 94 ºC for 3 min, followed by 40 cycles of denaturation at 94 ºC for 3 sec and
266 annealing at 60 ºC for the human, bovine and poultry assays and 57 ºC for pig assay.
267 Quantification was performed using specific standard curves built of seven points and
268 in duplicate for each MST marker. A standard curve was performed with each qPCR
269 run. Quantification for HAdV was constructed using the Amplirun® Adenovirus DNA
270 control (Vircell, Spain) with values ranging between 1.36 and 1.36 x 10^4 gene copies
271 (GC)/µL. For the quantification of mtDNA markers, the standard curves were obtained
272 by amplification of each target using conventional PCR from extracted stool samples.

After amplification, the amplicons were run in a 4% electrophoresis agarose gel. Upon visualization, each band was excised from the gel and DNA was purified using the illustra GFX PCR DNA and Gel band purification kit according to the manufacturer's instructions (GE Healthcare Life Sciences, US). Following purification of a DNA stock for each quantitative parameter, the concentration of each stock was quantified using the NanoDrop ND-100 spectrophotometer (ThermoScientifics, DE). Serial dilutions of each stock were prepared in DNA/RNA-free water in order to prepare the standard curve for qPCR. Standard curve for HMMit ranged between 5.87 and 5.87 x 10^6 GC/uL, for CWMit ranged from 5.5 and 5.5 x 10^6 GC/uL, for PGMit the standard curve varied between 5.71 and 5.71 x 10^6 GC/uL and for PLMit between 4.1 and 4.1 x 10^6 GC/uL. To ensure that no cross-contamination existed during the complete protocol several quality control steps were added. A process control accompanying the process of concentration, extraction and qPCR was added each time samples were analyzed. A negative extraction control followed the process of extraction. In addition, positive and a non-template controls were tested also with each qPCR run. All results were expressed in terms of GC/100 mL of sample.

2.8. Physico-chemical parameters

Different physico-chemical parameters may influence the behavior of microorganisms in the environment, including temperature (atmospheric and water), UV radiation, salinity, and rainfall (during 24 h previous to sampling). Data for physico-chemical parameters such as atmospheric temperature, rainfall, and UV radiation were obtained from the Instituto Português do Mar e da Atmosfera (IPMA, 2016). Salinity and water temperature data were obtained from the Marine, Environment and Technology Center (MARETEC, 2016).

2.9. Data analysis

All data analysis was performed with either Microsoft Excel 2016 or IBM SPSS Statistics 25 (IBM, NY). The limit of detection (LoD) was calculated by making serial dilutions of the standards used for each marker and determining the lowest quantity of the marker with a confidence level of 99%. All data were converted into a logarithmic format (log_{10}). Samples with negative results for each parameter were transformed and used as follows for the statistical analysis:

$$\frac{Method\ Limit\ of\ detection}{2}$$

To determine the influence of sampling location and season in the occurrence of each marker, the analysis of variances of two-way factors without replication was used. The evaluation of the impact of each physico-chemical parameters in the occurrence of each marker was carried out using Spearman's rank order correlation. One-way ANOVA was performed to determine the influence of the occurrence of precipitation on each marker and Spearman's rank order correlation was used for calculation of correlation coefficients between the levels of precipitation (mm), water temperature and salinity values and the concentration of each marker. Spearman's rank order correlation was used for calculation of correlation coefficients between parameters.

3. Results

3.1. Indicators of fecal contamination

A total of 105 samples were collected from the four different sampling points within the R. Tagus catchment. EC and IE were present in the majority of river water samples (99% and 98%, respectively) with median concentrations of 2.68 log EC MPN/100mL

and 2.02 log IE MPN/100mL (Table 1). SC were also present in the vast majority of the river water samples tested (99%) with a median concentration of 2.60 log PFU/100mL. These results show a high level of fecal contamination at all the sampling points tested.

3.2. Source Tracking markers

The LoD for assays HAdV, HMMit, PGMit, CWMit, and PLMit were 20.0, 8.0, 8.0, 5.0, and 5.0 GC per reaction, respectively. The human-associated markers of fecal contamination (GB-124 phages, HAdV, and HMMit) displayed notably different distributions (Table 1). HMMit was the most frequently detected marker (present in 83% of samples) and was also present at the highest concentration (4.20 log GC/100mL). HAdV were found in 32% of samples with a median concentration of 2.23 log GC/100mL. Conversely GB-124 phages were detected in just a single river water sample and at low concentration (1 PFU/100mL).

From the non-human markers (cattle (CWMit), pig (PGMit) and poultry (PLMit)), CWMit was the most frequently detected marker (present in 73% of samples) and was also found at the highest concentration (median 3.74 log GC/100mL) (Table 1). PGMit was detected in over 50% of the samples at a median of 2.99 log GC/100mL.

Table 1. Quantitative results and % of positive samples for FIB, non-specific phages (SC), and human (GB124, HMMit, HAdV), cattle (CWMit), pig (PGMit) and poultry (PLMit) fecal markers in R. Tagus catchment

		Concentration	
Parameter	*n*	Median (log *units/100mL)	% positives
EC	105	2.68	99
IE	105	2.02	98
SC	105	2.60	99
GB-124	105	0.00	1
HAdV	105	2.23	32
HMMit	105	4.20	83
CWMit	105	3.74	73
PGMit	105	2.99	56

PLMit	105	2.33	39

*units = CFU (EC, IE); PFU (SC, GB-124); GC (HAdV, HMMit, CWMit, PGMit, PLMit)

Finally, PLMit was the least frequently detected mtDNA marker, present in 39% of samples and at the lowest concentration (2.33 log GC/100mL).

3.3. Geographical and temporal distribution

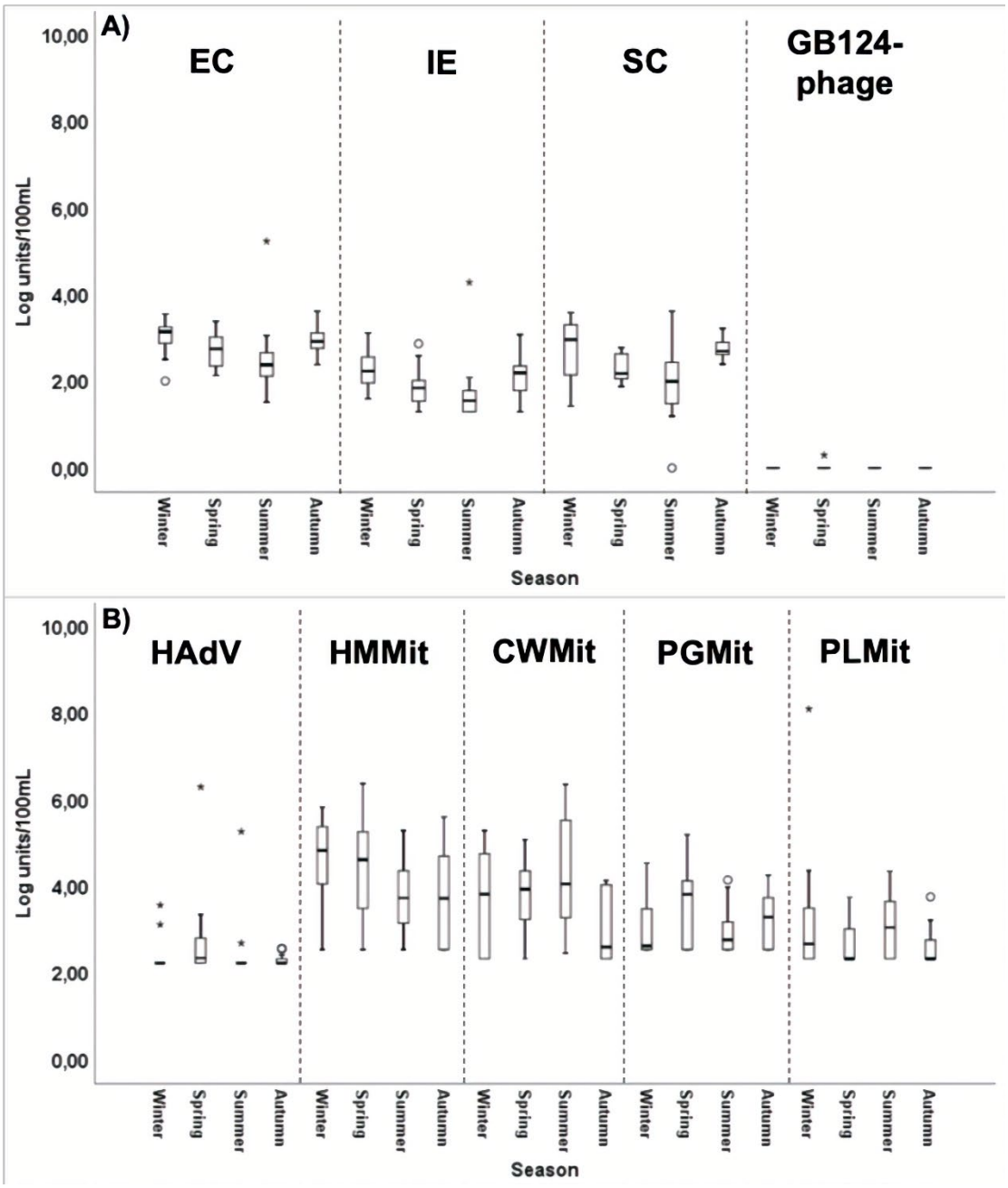
EC and IE were detected in all of the samples tested during winter and autumn, whereas during spring and summer, they were detected in approximately 90% of samples from the sampling points. SC were detected in all samples tested, with the exception of two samples collected during the summer at sampling site 4 (Bel). The median concentration and frequency of detection of each parameter are summarized in Table 2. Human fecal contamination (as determined by the presence of HAdV and HMMit) was detected all year around. HAdV was less prevalent during summer months (22% of samples positive), whereas HMMit was found to be less prevalent during autumn with lower concentrations during summer and autumn months. CWMit marker was also detected frequently (regardless of time of year) but was more abundant during the spring and summer months. PGMit was also consistently detected at similar concentrations, irrespective of season. PLMit was more common during the summer with positive results found at all sampling sites (percentage of positive between 40% and 75% positive).

Table 2. Quantitative results and percentage of positive samples for FIB, non-specific phages (SC), and human (GB124, HMMit, HAdV), cattle (CWMit), pig (PGMit) and poultry (PLMit) fecal markers in R. Tagus catchment by location and season

	Concentration							
	Median (*units/100mL)	**% positive**	**Median (*units/100mL)**	**% positive**	**Median (*units/100mL)**	**% positive**	**Median (*units/100mL)**	**% positive**
	Sampling site							
Parameter	**1. VFX**		**2. MPdN**		**3.Alc**		**4. Bel**	
EC	**2.96**	100	2.68	100	2.69	100	2.36	96
IE	1.79	96	**2.34**	100	2.02	100	1.91	96
SC	2.65	100	2.54	100	**2.70**	100	2.30	96
GB-124 phage	0.00	4	0.00	0	0.00	0	0.00	0
HAdV	2.23	24	2.23	33	**2.56**	52	2.23	32
HMMit	4.22	84	4.20	82	4.04	84	**4.30**	89
CWMit	3.73	76	**3.92**	85	3.49	68	3.68	75
PGMit	3.06	68	**3.21**	63	2.54	44	3.09	57
PLMit	2.33	36	**3.02**	59	2.33	28	2.33	32
	Season							
	Winter		**Spring**		**Summer**		**Autumn**	
EC	**3.15**	100	2.60	98	2.24	100	2.92	100
IE	**2.46**	100	1.91	95	1.60	100	2.20	100
SC	**3.22**	100	2.45	100	2.12	94	2.70	100
GB-124 phage	0.00	0	0.00	2	0.00	0	0.00	0
HAdV	2.23	29	2.23	43	2.23	22	2.23	42
HMMit	**4.53**	84	4.25	91	3.74	89	3.73	58
CWMit	2.80	52	4.00	93	**4.32**	94	2.60	50
PGMit	2.72	55	3.26	61	2.77	56	**3.30**	58
PLMit	2.33	45	2.33	30	**3.02**	61	2.33	25

*units = CFU (EC, IE); PFU (SC, GB-124); GC (HAdV, HMMit, CWMit, PGMit, PLMit); Bold denotes the highest reading for each parameter (with respect to sampling site and season)

370 The concentration of fecal indicators and ST markers differed considerably throughout
371 the year, with the HMMit and CWMit markers showing greater variance in their
372 seasonal values compared to the remaining markers (Fig. 2). The results showed that
373 the concentration of EC, IE and SC generally decreased during summer and spring
374 (Table 2, Fig. 2A).
375



376

377 **Fig. 2.** Concentration of the various fecal indicators and source tracking markers with respect to season
378 A) culture-based indicators/markers and B) molecular-based markers in the R. Tagus catchment.

379 In general, the presence of markers did not appear to be significantly influenced by
380 the sampling site location (Table 2, Fig. S1). The influence of sampling site and season
381 upon each parameter was conducted using the analysis of variance (two-way
382 ANOVA).
383 The concentration of the fecal indicators (EC, IE, SC) was highly affected by
384 seasonality ($p < 0.05$) and for EC the sampling site also played an important role ($p <$
385 0.05). The variance of the concentration of most source-associated markers was not
386 clearly attributed to either sampling site, or seasonal factors ($p > 0.05$; Fig. 2B), with
387 the exception of the CWMit that showed a statistically significant difference with
388 respect to season ($p < 0.05$; Fig. 2B). Further discrimination was performed to
389 understand the effect of season on the distribution of the indicators and markers (Table
390 S3). For EC, IE and SC there was a significant difference between winter
391 concentrations and concentrations in spring and summer ($p < 0.001$ and < 0.05),
392 whereas CWMit concentrations were lowest and highest during autumn and summer,
393 respectively. Nonetheless, CWMit concentrations also varied with respect to the other
394 seasons ($p < 0.05$). EC concentrations were furthermore affected by the sampling site,
395 with Site 4 (Belém) primarily responsible for the observed differences (Table S4).
396
397 **3.4. Correlation between physico-chemical parameters and non- and source-**
398 **associated markers**

399 Different physico-chemical parameters are known to influence the behavior of
400 microorganisms in the environment, including temperature (atmospheric and water),
401 UV radiation, salinity, and rainfall (IPMA, 2016; MARETEC, 2016). To understand
402 whether temperature and/or UV influenced the concentrations of each parameter,
403 Spearman's rank order correlation was carried out. Analysis of the effect of UV and
404 temperature on the indicators and markers (Table S5) showed that UV radiation and

405 atmospheric temperature appeared to have a combined impact on EC, IE, SC, and
406 CWMit. The remaining markers were not significantly affected by atmospheric
407 temperature or UV radiation levels. Another important factor potentially influencing
408 environmental water quality is precipitation, which was evaluated in two ways. Firstly,
409 a binary analysis was conducted (one-way ANOVA) on the influence of rainfall vs. no
410 rainfall (on day of sampling or 24 hrs before) on the presence of each fecal indicator
411 and marker. Secondly, the relationship between rainfall levels (mm) and the
412 concentration at which fecal indicators and source-specific markers were detected
413 was also determined using Spearman's rank order correlation. One-way ANOVA
414 showed that for the majority of the fecal indicators and markers analyzed, occurrence
415 of rainfall on the day of sampling (or 24 hrs prior), had little, or no appreciable influence
416 on which markers were present (Table S6). However, the one-way analysis of variance
417 identified significant correlation between IE, SC, and PLMit marker and presence of
418 rainfall on the day of sampling (or 24 hrs prior) ($p < 0.05$). Additionally, the influence
419 of rainfall was noticeable (though to a lesser extent) for SC and PLMit in samples from
420 site 3 (Alc) and site 4 (Bel), respectively (Table S6; Fig. S2). Although some
421 parameters were related to the occurrence of rainfall, the correlation between the
422 levels of rainfall and the concentration of other parameters showed a different pattern
423 (Table S7). Interestingly, the presence of rainfall appeared to be moderately correlated
424 with HAdV concentration (Spearman rank order correlation; r = 0.417, p-value < 0.05).
425 The level of rainfall was also negatively associated with the concentrations of IE and
426 SC but showed no correlation with the remaining parameters. The relationship
427 between water temperature/salinity and the concentration of the different parameters
428 was also evaluated at each site using Spearman's correlation (Table S8). The results
429 revealed that water temperature negatively affected concentrations of EC, IE, and SC

(r = -0.379, p < 0.01; r = -0.404, p < 0.01; r = -0.389, p < 0.01). Weak but positive correlation was found between water temperature and CWMit (r = 0.293, p < 0.01). The remaining parameters were not shown to correlate with water temperature. Water salinity only weakly correlated with the concentration of EC (r = -0.268, p < 0.01) and no correlation was found with IE. Although no correlation was found between other tested parameters and salinity, this parameter influenced most of them negatively (r below zero).

3.5. Relationship between source-associated markers and fecal indicators

To determine whether the presence and concentration of source-associated markers (HMMit, HAdV, GB-124, CWMit, PGMit, and PLMit) were quantitatively correlated to the presence and concentration of fecal indicators (FIB and SC), Spearman's correlation was performed. Among fecal indicators, the strongest correlation was found (Table 3) between EC and IE (r = 0.673, p < 0.05) and the weakest between IE and SC (r = 0.469, p < 0.05). Significant but weak correlations were also observed between EC and HAdV (p < 0.05), EC and PGMit (p < 0.05), IE and PGMit (p < 0.05), HAdV and CWMit (p < 0.05) and CWMit and PGMit (p < 0.05) (Table 3).

The concentration of human- and cattle-associated markers tended to decrease with respect to the concentration of fecal indicators (Table 3; Table S9). PGMit always presented the highest correlation coefficients with EC and SC.

Table 3. Spearman's correlation between the Log_{10} concentration of non-specific fecal markers and the different source-specific markers.

	EC	IE	SC	GB124 phages	HAdV	HMMit	CWMit	PGMit	PLMit
EC	-	**0.673**[a]	**0.511**[a]	0.055	0.204[b]	-0.091	-0.107	0.221[b]	0.093
IE		-	**0.469**[a]	-0.042	0.083	-0.059	-0.069	0.247[b]	0.193[b]
SC			-	-0.097	0.134	-0.079	-0.188	0.080	0.058
GB124 phages				-	0.202[b]	0.019	0.046	0.132	-0.075
HAdV					-	-0.012	0.223[b]	0.014	0.144
HMMit						-	0.061	0.022	-0.056
CWMit							-	0.255[a]	-0.010
PGMit								-	-0.043
PLMit									-

a Two tailed significance, $p < 0.01$.

b Two tailed significance, $p < 0.05$.

Conversely, the remaining markers showed low, or no significant correlations with the fecal indicators. To understand whether the source-specific MST markers correlated quantitatively to the non-specific fecal markers during rainfall events, correlation analysis using Spearman rank correlation coefficients were performed in this situation (Table S10). PGMit was moderately positively correlated with IE during rain events (r = 0.444, p-value = 0.05). Although not statistically correlated at a p-value of 0.05, PGMit presented the highest correlation coefficients with the remaining indicators of fecal contamination, EC and SC. In stark contrast, the remaining markers showed low, or no significant correlations with the non-specific markers.

4. Discussion

This study was underpinned by a 13-month sampling campaign to determine the presence, concentration and behavior of the fecal indicators and novel source-associated markers, with respect to season and sampling site. This information was lacking from the extant literature for mitochondrial markers, so our findings represent an important gap in current knowledge that is crucial for establishing their suitability as source-associated markers for routine deployment in river catchments. Our study design not only encompassed seasonal fluctuations, but also investigated the influence of precipitation (presence and levels), since it has been demonstrated that more than 90% of the flux of fecal contamination from diffuse sources (as determined by fecal indicator concentrations) can be related to hydrological events (Reischer *et al*., 2008). The results of this catchment deployment have indicated that human fecal sources are the dominant source of contamination at all four sampling points monitored. Human fecal contamination was more prevalent than the remaining targeted sources of fecal contamination. Fecal contamination of cattle (CWMit) origin

481 also appeared to heavily influence water quality, as this marker was commonly
482 detected (in over 70% of samples) and at high concentrations (median concentration
483 3.74 log GC/100mL). Sampling sites 1 (VFX) and 2 (MNdP) were suspected to both
484 contain a significant component of contamination from cattle origins, since site 1 is
485 known for the rearing and exportation of cattle and site 2 receives the influent from
486 two WWTP known to receive waste from cattle farms, along with direct riverine inputs
487 (run-off) from the cattle farms on the south side of the river. Nonetheless, the other
488 animal sources were also frequently detected during this study, with more than half of
489 the samples testing positive for porcine contamination and about 40% positive for
490 poultry. Porcine fecal inputs were thought to be more likely at sampling site 2, due to
491 its proximity to pig farming operations on not only the north bank of the R. Tagus, but
492 also on the southern bank. Although found in higher concentration on sampling site 2,
493 porcine fecal contamination was also detected at all sampling locations, with sampling
494 site 1 having similar prevalence and concentration to sampling site 2. Lisbon
495 metropolitan area consists of 18 municipalities, many of which are predominantly
496 urban, with approximately three million inhabitants (INE, 2014). The city of Lisbon, with
497 a population of approximately 500,000 inhabitants (INE, 2014), is mainly an urban
498 area but many of the municipalities on the north and south bank of the River Tagus
499 are not only densely populated with human inhabitants but also characterized by high
500 numbers of agricultural animals, including cattle farms, pig farms and poultry abattoirs.
501 Despite being regarded as an urban area, in 2018 the Lisbon metropolitan area
502 contained 5% of the total number of cattle in Portugal (75,000 cattle units), 10% of the
503 total number of pigs (220,000 pigs) and more than 180,000 poultry (INE, 2019). In
504 addition to the large number of farm animals on the outskirts of the city, the farms and
505 abattoirs are often not equipped with adequate wastewater treatment processes, or

506 do not have any form of on-site treatment, meaning that fecal contamination generated
507 by these industries either goes mostly directly to the municipal wastewater, or is
508 discharged into nearby environmental waters. The WWTP featured in this study did
509 not (at the time the study was conducted) have sufficient treatment/disinfection in
510 place to substantially reduce the microbiological parameters (and also mitochondrial
511 cells) in the final effluent. These factors are likely to be the reason why a high
512 percentage of samples collected during this study contained high concentrations of
513 human and agricultural inputs. On the other hand, the human marker GB124
514 bacteriophage was detected in just one sample, at low concentration, suggesting that
515 this host appeared to be geographically restricted from the catchment and that other
516 *Bacteroides* hosts (e.g. GA-17) may be more appropriate in this region. The existence
517 of phages in feces and wastewater is highly dependent on the presence in the
518 mammalian gut of *Bacteroides* host strains, which are homogenous in terms of
519 receptors and modification-restriction enzymes resulting in capability of phage
520 replication (Puig *et al*., 1999; Ebdon *et al*., 2012; McMinn *et al*., 2014). Therefore, the
521 diversity of *Bacteroides* strains is the main factor responsible for the variations in the
522 levels of infectivity. *Bacteroides* strains have evolved over time in very well
523 compartmentalized environments and have been segregated according to regional
524 dietary regimens and other factors such as host age and climate (Benno *et al*., 1989;
525 Moore and Moore, 1995; Puig *et al*., 1999). *B. fragilis* strain GB124 was isolated in the
526 UK and the phages infecting this strain have been shown to be almost exclusively of
527 human origin. The diet in Southern Europe is somewhat different from that of the UK,
528 where climate is generally cooler and wetter, two factors that may influence the
529 different evolutionary paths in *Bacteroides* strains. Nonetheless, phages capable of
530 infecting *B. fragilis* strain GB124 have been found to be present at detectable levels

531 in studies carried out across the US and Brazil (McMinn *et al*., 2014; 2017; Prado *et*
532 *al*., 2018).
533 Fecal indicators generally displayed a marked seasonal pattern, with higher
534 concentrations during the colder months and lower concentrations during the warmer
535 months. This suggests that the decrease in the fecal indicators is correlated with the
536 increase in solar radiation (intensity and duration) and possible biological predation of
537 the indicators/markers within environmental waters (Kay *et al*., 2005; Schultz-
538 Fademrecht *et al*., 2008; Byappanahalli *et al*., 2012). Conversely, the concentration of
539 the source-associated markers was not impacted by seasonality, with the exception
540 of the cattle marker, which was present at higher concentration in warmer months,
541 particularly during summer. The increase in the concentration of this source-
542 associated marker is likely the result of a number of potential factors. First, cattle
543 numbers are far higher in the catchment during the warmer months (as animals
544 housed indoors over winter are moved outside); secondly, the chance of rainfall (storm
545 events) at this time of the year is more likely to mobilize diffuse fecal material into
546 surface waters; and thirdly, low-flow conditions mean that fecal inputs have a greater
547 impact on the water quality and source-associated markers are more abundant due to
548 reduced dilution. These findings are in accordance with the findings in other studies,
549 e.g., by Rusiñol *et al*., 2014, who detected contamination of animal origin at a higher
550 concentration during the summer in a study performed in five different countries
551 (Brazil, Greece, Hungary, Spain, and Sweden)
552 Our study also showed that IE, PLMit and HAdV were impacted by rainfall occurrence
553 and levels. Similar results were demonstrated in previous studies (Haramoto *et al*.,
554 2006; Hata *et al*., 2014; Rodrigues *et al*., 2015). Levels of the PLMit marker were
555 significantly lower in samples collected on dry days, compared to those collected

following precipitation (during previous 24hrs) and concur with previous findings (Barros *et al*., 2007). HMMit displayed no clear seasonality, being unaffected by atmospheric temperature and UV radiation levels. Surprisingly, little information exists on the influence of different environmental physicochemical parameters on the degradation of mitochondrial cells, or on the mitochondrial DNA itself. The detection of mitochondrial DNA from human, cattle, and porcine sources in raw wastewater, partially treated wastewater and disinfected effluent was studied previously (Martellini *et al*., 2005) and showed that following each treatment stage, only human mitochondrial DNA remained at detectable levels. He *et al*. (2016) determined that temperature and sunlight negatively correlated with the concentration of human and pig mtDNA markers in a field microcosm using dialysis tubes. Malla and Haramoto, 2020 have compiled the information on mtDNA MST markers and their persistence in the environment. However, the studies presented were performed in microcosms, rather than on samples collected directly from the river channel and were not tested alongside as larger number of physico-chemical parameters as featured in this study. The PGMit marker was weakly correlated to EC and IE during the entire sampling campaign but showed moderate correlation with IE during and following precipitation events. Therefore, this study demonstrated the capacity of certain markers, such as PGMit, to link source-associated markers concentration to fecal indicators. This should facilitate the estimation of the contribution from particular sources to the total load of fecal contamination assessed by conventional fecal indicators. The correlation found between PGMit and IE is even more interesting when considering the US EPA's revised recreational water quality criteria for IE, that aims to keep the risk of gastrointestinal illness in swimmers below 30 illnesses per 1000 swimmers (USEPA, 2012).

581 This study demonstrated that the 13-month sampling campaign has contributed to an
582 improved understanding of the effect of seasonality on indicator and source-
583 associated marker dynamics and highlighted the importance of catchment and
584 precipitation event sampling for effective quantitative source tracking. Moreover, the
585 parametric analysis of precipitation events allowed for the quantitative assessment of
586 the porcine fecal contamination component. It must be emphasized here that the
587 potential dominant sources of fecal contamination in study sites on the R. Tagus were
588 either of human or cattle origin.
589 Whilst it is clear that cattle, pig and poultry rearing are common at certain points within
590 the study catchment, the prevalence and concentration of non-human markers during
591 the sampling campaign, highlight the potential impact that these non-point sources
592 can have on water quality within the R. Tagus catchment. Although human sources
593 are a major contributor to contamination within the R. Tagus (and as such represent a
594 risk to public health), fecal contamination from a range of non-human sources may
595 also present considerable potential risks to public health. For example, several
596 zoonotic pathogens are able to infect or exist in a range of different warm-blooded
597 mammals (including humans), such as *Cryptosporidium* and *Giardia*, and Hepatitis E
598 virus. In addition to the application of tailored models to the area of study by combining
599 local samples including also factors such as spatial and seasonal inactivation, it has
600 been suggested (Ballesté *et al*., 2020) that dilution and aging may help improve the
601 accuracy reducing the number of source-associated markers to be tested. To
602 conclude, this study provides new data concerning the temporal and spatial
603 distribution of fecal indicators and source-associated markers and correlations
604 between different variables, which could be extremely helpful in the management of

water quality and in the application of remediation measures to enhance and restore the R. Tagus catchment.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

References

Ahmed, W., Payyappat, S., Cassidy, M., Besley, C., 2019. Enhanced insights from human and animal host-associated molecular marker genes in a freshwater lake receiving wet weather overflows. Sci. Rep. 9(1): 12503.

Andreasson, H., Gyllensten, U., Allen, M., 2002. Real-time DNA quantification of nuclear and mitochondrial DNA in forensic analysis. Biotechniques, 33, 407-411.

Ashbolt, N.J., Grabow, W.O.K., Snozzi, Snozzi, M., 2001. Indicators of microbial water quality, in: Fewtrell, L., Bartram, J. (Eds.), World Health Organization, Water Quality: Guidelines, Standards and Health. IWA Publishing, London, pp. 289 – 316.

ASTM, 1999. D6503-99 Standard test method for Enterococci in water using Enterolert. West Conshohocken, United States.

Badgley, B., Thomas, F., Harwood, V., 2011. Quantifying environmental reservoirs of fecal indicator bacteria associated with sediment and submerged aquatic vegetation. Environ. Microbiol. 13(4), 932–942.

Ballesté, E., Belanche-Muñoz, L.A., Farnleitner, A.H., Linke, R., Sommer, R., Santos, R., Monteiro, S., Maunula, L., Oristo, S., Tiehm, A., Stange, C., Blanch, A., 2020. Improving the identification of the source of faecal pollution in water using a modelling approach: from multi-source to aged and diluted samples. Water Res. 171: 115392.

Barros, L.S., Amaral, L.A., Lorenzon, C.S., Junior, J.L., Machado Neto, J.G., 2007. Potential microbiological contamination of effluents in poultry and swine abattoirs. Epidemiol. Infect. 135, 505-518.

Benno, Y., Endo, K., Mizutani, T., Namba, Y., Komori, T., Mitsuoka, T., 1989. Comparison of fecal microflora of elderly persons in rural and urban areas of Japan. Appl. Environ. Microbiol. 55(5), 1100–1105.

638 Byappanahalli, M.N., Nevers, M.B., Korajkic, A., Staley, Z.R., Harwood, V.J., 2012.
639 Enterococci in the environment. Microbiol. Mol. Biol. Rev. 76(4), 685-706.
640 Caldwell, J.M., Raley, M.E., Levine, J.F., 2007. Mitochondrial multiplex real-time PCR
641 as a source tracking method in fecal-contaminated effluents. Environ. Sci. Technol.
642 41(9), 3277-3283.
643 Calgua, B., Mengewein, A., Grunert, A., Bofill-Mas, S., Clemente-Casares, P.,
644 Hundesa, A., Wyn-Jones, A.P., López-Pila, J.M., Gironés, R., 2008. Development and
645 application of a one-step low cost procedure to concentrate viruses from seawater
646 samples. J. Virol. Methods 153(2), 79–83.
647 Colford, J., Wade, T.J., Schiff, K.C., Wright, C.C., Griffith, J.F., Sandhu, S.K., Burns,
648 S., Sobsey, M., Lovelace, G., Weisberg, S.B., 2007. Water quality indicators and the
649 risk of illness at beaches with nonpoint sources of fecal contamination. Epidemiology
650 18(1), 27–35.
651 Cotruvo, J.A., Dufour, A., Rees, G., Bartram, J., Carr, R., Cliver, D.O., Craun, G.F.,
652 Fayer, R., Gannon, V.P.J., 2004. Waterborne Zoonoses. World Health Organization,
653 IWA Publishing, London.
654 Dwight, R.H., Baker, D.B., Semenza, J.C., Olson, B.H., 2004. Health effects
655 associated with recreational coastal water use: urban versus rural California. Am. J.
656 Public Health 94(4), 565–567.
657 Ebdon, J., Muniesa, M., Taylor, H., 2007. The application of a recently isolated strain
658 of Bacteroides (GB-124) to identify human sources of faecal pollution in a temperate
659 river catchment. Water Res. 41(16), 3683–3690.
660 Ebdon, J.E., Sellwood, J., Shore, J., Taylor, H.D., 2012. Phages of Bacteroides (GB-
661 124): a novel tool for viral waterborne disease control? Environ. Sci. Technol. 46(2),
662 1163–1169.

663 Gerber A.S., Loggins, R., Kumar, S., Dowling, T.E., 2001. Does nonneutral evolution
664 shape observed patterns of DNA variation in animal mitochondrial genomes? Annu.
665 Rev. Genet. 35, 539-566.
666 Google, 2020. Google Maps. https://www.google.pt/maps/@38.7376736,-185,-
667 9.1392886,15z?hl=en (accessed in February 2020).
668 Hameeteman, E., 2013. Future Water (In)security: facts, figures, and predictions.
669 Global Water Institute, 2013.
670 Haramoto, E., Katayama, H., Oguma, K., Koibuchi, Y., Furumai, H., Ohgaki, S., 2006.
671 Effects of rainfall on the occurrence of human adenovirus, total coliforms, and
672 Escherichia coli in seawater. Water Sci. Technol. 54(3), 225–230.
673 Haramoto, E., 2018. Detection of waterborne protozoa, viruses, and bacteria in
674 groundwater and other water samples in the Kathmandu Valley, Nepal. IOP Conf. Ser.:
675 Earth Environ. Sci. 120, 012004.
676 Harwood, V., Butler, J., Parrish, D., Wagner, V., 1999. Isolation of fecal coliform
677 bacteria from the diamondback terrapin (Malaclemys terrapin centrata). Appl. Environ.
678 Microbiol. 65(2), 865–867.
679 Hata, A., Katayama, H., Kojima, K., Sano, S., Kasuga, I., Kitajima, M., Furumai, H.,
680 2014. Effects of rainfall events on the occurrence and detection efficiency of viruses
681 in river water impacted by combined sewer overflows. Sci. Total Environ. 468-469,
682 757–763.
683 He, X., Liu, P., Zheng, G., Chen, H., Shi, W., Cui, Y., Ren, H., Zhang, X.-X., 2016-
684 Evaluation of five microbial and four mitochondrial DNA markers for tracking human
685 and pig fecal pollution in freshwater. Sci. Rep 6, 35311.
686 Hernroth, B.E., Conden-Hansson, A.-C., Rehnstam-Holm, A.-S., Gironés, R., Allard,
687 A.K., 2002. Environmental factors influencing human viral pathogens and their

688 potential indicator organisms in the blue mussel, Mytilus edulis: the first scandinavian
689 report. Appl. Environ. Microbiol. 68(9), 4523–4533.
690 Holcomb, D.A., Stewart, J.R., 2020. Microbial indicators of fecal pollution: recent
691 progress and challenges in assessing water quality. Curr. Envir. Health Rpt. 7, 311-
692 324. doi: 10.1007/s40572-020-00278-1.
693 INE, 2014. Tipologia socioeconómica das Áreas Metropolitanas de Lisboa e Porto –
694 2011. Lisbon, Portugal.
695 INE, 2019. Instituto Nacional de Estatística – Estatística Agrícolas 2018, Edição 2019.
696 IPMA, 2016. Instituto Português do Mar e da Atmosfera.
697 http://www.ipma.pt/pt/index.html. (accessed in September 2016).
698 Ishii, S., Ksoll, W.B., Hicks, R.E., Sadowsky, M.J., 2006. Presence and growth of
699 naturalized Escherichia coli in temperate soils from Lake Superior watersheds. Appl.
700 Environ. Microbiol. 72(1), 612–621.
701 ISO 10705-2, 2000. ISO 10705-2 Water quality. Detection and enumeration of
702 bacteriophages - Part 2: Enumeration of somatic coliphages. International
703 Organisation for Standardisation, Geneva, Europe.
704 ISO 10705-4, 2001. ISO 10705-4 Water quality - Detection and enumeration of
705 bacteriophages - Part 4: Enumeration of bacteriophages infecting Bacteroides fragilis.
706 International Organisation for Standardisation, Geneva, Europe.
707 Iyengar, V., Albaugh, G.P, Lohani, A., Nair, P.P., 1991. Human stools as a source of
708 viable colonic epithelial cells. FASEB J. 5, 2856-2859.
709 Jofre, J., Blanch, A., Lucena, F., Muniesa, M., 2014. Bacteriophages infecting
710 Bacteroides as a marker for microbial source tracking. Water Res. 55, 1-11.
711 Kay, D., Stapleton, C.M., Wyer, M.D., McDonald, A.T., Crowther, J., Paul, N., Jones,
712 K., Francis, C., Watkins, J., Wilkinson, J., Humphrey, N., Lin, B., Yang, L., Falconer,

713 R.A., Gardner, S., 2005. Decay of intestinal enterococci concentrations in high-energy
714 estuarine and coastal waters: towards real-time *T*90 values for modelling faecal
715 indicators in recreational waters. Water Res. 39(4), 655-667.
716 Kongprajug, A., Chyerochana, N., Somnark, P., Kampaengthong, P.L., Mongksolsuk,
717 S., Sirikanchana, K., 2019. Human and animal microbial source tracking in a tropical
718 river with multiple land use activities. Int. J. Hyg. Environ. Health 222, 645-654.
719 Ksoll, W., Ishii, S., Sadowsky, M., Hicks, R., 2007. Presence and sources of fecal
720 coliform bacteria in epilithic periphyton communities of Lake Superior. Appl. Environ.
721 Microbiol. 73(12), 3771–3778.
722 Malla, B., Haramoto, E., 2020. Host-specific mitochondrial DNA markers for tracking
723 the sources of fecal pollution. Curr. Opin. Environ. Sci. Health 16, 34-46.
724 MARETEC, 2016. Marine Environment and Technology Center - Instituto Superior
725 Técnico. http://forecast.maretec.org/maps_tagusestuary3D.asp (accessed in October
726 2016).
727 Martellini, A., Payment, P., Villemur, R., 2005. Use of eukaryotic mitochondrial DNA
728 to differentiate human, bovine, porcine and ovine sources in fecally contaminated
729 surface water. Water Res. 39(4), 541–548.
730 McMinn, B.R., Korajkic, A., Ashbolt, N.J., 2014. Evaluation of Bacteroides fragilis GB-
731 124 bacteriophages as novel human-associated faecal indicators in the United States.
732 Lett. Appl. Microbiol. 59(1), 115–121.
733 McMinn, B.R., Ashbolt, N.J., Korajkic, A., 2017. Bacteriophages as indicators of fecal
734 pollution and enteric virus removal. Lett. Appl. Microbiol. 65(1), 11-26.
735 Mendez, J., Audicana, A., Isern, A., Llaneza, J., Moreno, B., Tarancon, M. L., Jofre,
736 J., Lucena, F., 2004. Standardised evaluation of the performance of a simple

membrane filtration-elution method to concentrate bacteriophages from drinking water. J. Virol. Methods 117(1), 19-25.

Moore, W., Moore, L., 1995. Intestinal floras of populations that have a high risk of colon cancer. Appl. Environ. Microbiol. 61(9), 3202–3207.

Payan, A., Ebdon, J., Taylor, H., Gantzer, C., Ottoson, J., Papageorgiou, G.T., Blanch, A.R., Lucena, F., Jofre, J., Muniesa, M., 2005. Method for isolation of Bacteroides bacteriophage host strains suitable for tracking sources of fecal pollution in water. Appl. Environ. Microbiol. 1(9), 5659–5662.

Prado, T., Bruni, A.C., Barbosa, M.R.F., Bonanno, V.M.S., Garcia, S.C., Sato, M.I.Z., 2018. Distribution of human fecal marker GB-124 bacteriophages in urban sewage and reclaimed water of São Paulo city, Brazil. J. Water Health, 16(2), 289-299.

Puig, A., Queralt, N., Jofre, J., Araujo, R., 1999. Diversity of Bacteroides fragilis strains in their capacity to recover phages from human and animal wastes and from fecally polluted wastewater. Appl. Environ. Microbiol. 65(4), 1772–1776.

Reischer, G.H., Haider, J.M., Sommer, R., Stadler, H., Keiblinger, K.M., Hornek, R., Zerobin, W., Mach, R.L., Farnleitner, A.H., 2008. Quantitative microbial faecal source tracking with sampling guided by hydrological catchment dynamics. Environ. Microbiol. 10(10), 2598–2608.

Rodrigues, M.T., Henzel, A., Staggemeier, FR., de Quevedo, D.M., Rigotto, C., Heinzelmann, L., do Nascimento, C.A., Spilki, F.R., 2015. Human adenovirus spread, rainfalls, and the occurrence of gastroenteritis cases in a Brazilian basin. Environ. Monit. Assess. 187, 720

Rusiñol, M., Fernandez-Cassi, X., Hundesa, A., Vieira, C., Kern, A., Eriksson, I., Ziros, P., Kay, D., Miagosotvich, M., Vargha, M., Allard, A., Vantarakis, A., Wyn-Jones, P., Bofill-Mas, S., Girones, R., 2014. Application of human and animal viral microbial

762 source tracking tools in fresh and marine waters from five different geographical areas.
763 Water Res. 59, 119-129.
764 Schill, W.B., Mathes, M.V., 2008. Real-time PCR detection and quantification of nine
765 potential sources of fecal contamination by analysis of mitochondrial cytochrome b
766 targets. Environ. Sci. Technol. 42, 5229-5234.
767 Schultz-Fademrecht, C., Wichern, M., Horn, H., 2008. The impact of sunlight on
768 inactivation of indicator microorganisms both in river water and benthic biofilms. Water
769 Res. 42(19), 4771-4779.
770 Scott, T.M., Rose, J.B., Jenkins, T.M., Farrah, S.R., Lukasik, J., 2002. Microbial
771 Source Tracking – current methodology and future directions. Appl. Environ. Microbiol.
772 68, 5796-5803.
773 Seidel, M., Jurzik, L., Brettar, I., Höfle, M.G., Griebler, C., 2016. Microbial and viral
774 pathogens in freshwater: currect research aspects studies in Germany. Environ. Earth
775 Sci. 75, 1384. doi: 10.1007/s12665-016-6189-x
776 Tartera, C., Araujo, R., Michel, T., Jofre, J., 1992. Culture and decontamination
777 methods affecting enumeration of phages infecting Bacteroides fragilis in sewage.
778 Appl. Environ. Microbiol. 58(8), 2670–2673.
779 UK Environment Agency, 2009. The microbiology of drinking water (2009) - Part 4 -
780 methods for the isolation and enumeration of coliform bacteria and Escherichia coli
781 (including E. coli O157:H7).
782 US EPA, 2012. Recreational water quality criteria. Document number 820-F-12-058.
783 Office of Water, United States Environmental Protection Agency, Washington DC.
784 WHO, 2017. Diarrhoeal disease. https://www.who.int/news-room/fact-
785 sheets/detail/diarrhoeal-disease (accessed in October 2019).

World Water Assessment Programme, 2009. The United Nations World Water Development Report: water in a changing world. UNESCO: Paris, and Earthscan: London.