

1 **Metagenomic analysis of viruses, bacteria and protozoa in irrigation water**

2

3 Marta Rusiñol<sup>1</sup>, Sandra Martínez-Puchol<sup>1</sup>, Natalia Timoneda<sup>1,2</sup>, Xavier Fernández-Cassi<sup>1</sup>, Alba Pérez-  
4 Cataluña<sup>3</sup>, Ana Fernández-Bravo<sup>3</sup>, Laura Moreno-Mesonero<sup>4</sup>, Yolanda Moreno<sup>4</sup>, Jose Luís Alonso<sup>4</sup>,  
5 Maria José Figueras<sup>3</sup>, Josep Francesc Abril<sup>2</sup>, Sílvia Bofill-Mas<sup>1</sup> and Rosina Girones<sup>1</sup>.

6

7 1 Laboratory of Viruses Contaminants of Water and Food. Department of Genetics, Microbiology and  
8 Statistics. Faculty of Biology, University of Barcelona, Catalonia, Spain.

9 2 Computational Genomics Lab, Department of Genetics, Microbiology and Statistics. Faculty of  
10 Biology and Institute of Biomedicine (IBUB), University of Barcelona, Catalonia, Spain.

11 3 Microbiology Unit, Faculty of Medicine and Health Sciences, University Rovira and Virgili. Reus,  
12 Catalonia, Spain.

13 4 Instituto de Ingeniería del Agua y Medio Ambiente, Universitat Politècnica de València. Valencia,  
14 Spain.

15

16 **Keywords:** Irrigation water, Metagenomics, Virus, Bacteria, Protozoa

17

18 **Abstract**

19 Viruses (e.g., noroviruses and hepatitis A and E virus), bacteria (e.g., *Salmonella spp.* and pathogenic  
20 *Escherichia coli*) and protozoa (e.g., *Cryptosporidium parvum* and *Giardia intestinalis*) are well-  
21 known contributors to food-borne illnesses linked to contaminated fresh produce. As agricultural  
22 irrigation increases the total amount of water used annually, reclaimed water is a good alternative to  
23 reduce dependency on conventional irrigation water sources. European guidelines have established  
24 acceptable concentrations of certain pathogens and/or indicators in irrigation water, depending on the  
25 irrigation system used and the irrigated crop. However, the incidences of food-borne infections are  
26 known to be underestimated and all the different pathogens contributing to these infections are not  
27 known. Next-generation sequencing (NGS) enables the determination of the viral, bacterial and

28 protozoan populations present in a water sample, providing an opportunity to detect emerging  
29 pathogens and develop improved tools for monitoring the quality of irrigation water. This is a  
30 descriptive study of the virome, bacteriome and parasitome present in different irrigation water  
31 sources. We applied the same concentration method for all the studied samples and specific  
32 metagenomic approaches to characterize both DNA and RNA viruses, bacteria and protozoa.

33 In general, most of the known viral species corresponded to plant viruses and bacteriophages. Viral  
34 diversity in river water varied over the year, with higher bacteriophage prevalences during the autumn  
35 and winter. Reservoir water contained *Enterobacter cloacae*, an opportunistic human pathogen and an  
36 indicator of fecal contamination, as well as *Naegleria australiensis* and *Naegleria clarki*. Hepatitis E  
37 virus and *Naegleria fowleri*, emerging human pathogens, were detected in groundwater. Reclaimed  
38 water produced in a constructed wetland system presented a virome and bacteriome that resembled  
39 those of freshwater samples (river and reservoir water). Viral, bacterial and protozoan pathogens were  
40 occasionally detected in the different irrigation water sources included in this study, justifying the use  
41 of improved NGS techniques to get a comprehensive evaluation of microbial species and potential  
42 environmental health hazards associated to irrigation water.

## 43 **1. Introduction**

44 Agricultural irrigation accounts for 36% of the total annual water usage in Europe, reaching up to  
45 80% in some parts of the Mediterranean region (European Environment Agency, 2012). In Spain 79%  
46 of the total irrigated area (3Mha) is irrigated with surface water and only 21% with groundwater  
47 (Drewes et al., 2017). River flow rate fluctuation and the overexploitation of groundwater resources  
48 are existing problems in Europe, where integrated water resource plans are currently being  
49 implemented. Moreover, climate change is expected to intensify problems of water scarcity and affect  
50 irrigation requirements in the Mediterranean region (Collins et al., 2009). To reduce dependency on  
51 freshwater, reclaimed water can be used for irrigation, provided that the risk associated with pathogen  
52 contamination can be minimised. Reclaimed water can be relatively nutrient rich and may reduce the  
53 need for additional applications of inorganic fertilizers (Parsons et al., 2001). In Catalonia, around 300

54 wastewater treatment plants (WWTP) treat 700 hm<sup>3</sup> of water every year to produce 204 hm<sup>3</sup> of  
55 reclaimed water (ACA, 2016; Pérez et al., 2011). The use of reclaimed water sources is regulated by  
56 different national regulations and European guidelines, which stipulate the acceptable microbial  
57 concentrations in irrigation water based on the irrigation system being used and the irrigated crop (RD  
58 1620, 2007; WHO, 2006). In 2017, the European Commission published the minimum quality  
59 requirements for water reuse in agricultural irrigation and aquifer recharge, producing a guidance  
60 document addressing microbiological risks in the primary production of fresh fruits and vegetables  
61 (Alcalde-Sanz and Gawlik, 2017; EU C163, 2017).

62

63 The number of reported illnesses linked to contaminated fresh produce has increased in both  
64 developed and developing countries, with the majority of reported outbreaks being in Europe  
65 (Chatziprodromidou et al., 2018). Viruses (e.g., noroviruses and hepatitis A and E virus), bacteria  
66 (e.g., *Salmonella spp.* and pathogenic *Escherichia coli*) and protozoa (e.g., *Cryptosporidium parvum*  
67 and *Giardia intestinalis*) are well-known contributors to these reported food-borne diseases; however,  
68 the number of food-borne infections is known to be underestimated and all the different pathogens,  
69 including new and emerging strains, contributing to these infections are not known. The risk  
70 associated with the consumption of fresh produce irrigated with water containing pathogens depends  
71 on many variables, the most important being the concentration of pathogens in the irrigation water  
72 (Gonzales-Gustavson et al., 2019). The levels of exposure to solar radiation, the irrigation system, the  
73 time from irrigation to consumption and the type of crop are other important factors (Pachepsky et al.,  
74 2011; Thebo et al., 2017). If crops are irrigated with reclaimed water, the concentration loads of the  
75 pathogens and their specific persistence in the environment are also important factors.

76 Current regulations and monitoring programs involve the measurements of selected microbial  
77 indicators, which do not necessarily correlate with the reference pathogens (Drewes et al., 2017).  
78 Therefore, more resistant pathogens, zoonotic microorganisms excreted by livestock or even  
79 uncharacterized pathogenic strains may also be present or underestimated in irrigation water.

80

81 While pathogen-specific tests for water quality assessment are appropriate for estimating irrigation  
82 water risks, next-generation sequencing (NGS) techniques can detect a wide range of microbial  
83 populations in a given sample. Using NGS technologies, most studies have focused on bacterial  
84 communities instead of viral or protozoan populations in different environments. Several studies have  
85 demonstrated that sewage reflects the bacteriome (Newton et al., 2015), virome (Cantalupo et al.,  
86 2011; Fernandez-Cassi et al., 2018; Ng et al., 2012) and parasitome (Bradley et al., 2016) of human  
87 populations. Different studies have also characterized virus, bacteria or protozoa diversity in  
88 groundwater (D'Auria et al., 2018), river water (Uyaguari-Diaz et al., 2016), and reclaimed water  
89 (Rosario et al., 2009). Metagenomic methods have been used to tackle viruses and bacteria found in  
90 fresh fruits and vegetables surfaces, highlighting the potential role of irrigation water as vector of  
91 waterborne pathogens (Aw et al., 2016; Fernandez-Cassi et al., 2017; Jackson et al., 2013; Leff and  
92 Fierer, 2013). Viruses lack a common marker for their taxonomical classification and identification,  
93 making difficult to monitor the entire population in a single assay. Uyaguari-Diaz and colleagues  
94 selected the gene 23 (encoding the major capsid protein of T4 bacteriophages) and the gene encoding  
95 the RNA-dependent RNA polymerase (present in most RNA viruses) as the target amplicons for viral  
96 metagenomics (Shi et al., 2016; Uyaguari-Diaz et al., 2016). As with the use of 16S or 18S ribosomal  
97 RNA analysis to identify bacteria or protozoa, respectively, the use of a common marker for  
98 identifying viruses might result in an imprecise taxonomical classification due to the poor resolution  
99 caused by using a single gene (Poretsky et al., 2014). Viruses show great diversity in their genome  
100 structure and organization. Thus, important waterborne pathogenic DNA viruses (e.g. adenovirus)  
101 lacking an RNA-dependent RNA polymerase might be overlooked when using this marker to identify  
102 viruses. In this study, we applied a metagenomic approach to characterize the viral populations in  
103 different water sources used for irrigation, including both DNA and RNA viruses. To our knowledge,  
104 this is the first study to assess the virome, bacteriome and parasitome of the same irrigation water  
105 samples using metagenomics after a single concentration method for all the studied microorganisms.  
106  
107  
108

## 109 2. Materials and methods

### 110 2.1. Water samples and concentration of microorganisms using the skimmed milk flocculation 111 method (SMF).

112 Four conventional irrigation water sources (drinking water, reservoir water, groundwater and river  
113 water) and one reclaimed water source, currently used for irrigation, were selected for the study. Raw  
114 sewage was included as a reference of the microbial contamination circulating within the studied area.  
115 Samples were collected every month over a period of one year, from April 2016 to March 2017.

116 Drinking water was sampled directly from water tanks across the water distribution network of  
117 Barcelona. This water source is frequently used in urban community orchards. Reservoir water, a  
118 common source of irrigation water in the Catalan territory, was collected from the Foix river  
119 reservoir. Groundwater, which is a conventional irrigation water source in the Mediterranean region,  
120 was sampled from the north-east of Catalonia, a region with intensive farming and agricultural  
121 activities. River water samples were collected from the Fluvià river, which receives effluents from 24  
122 small WWTPs treating up to 100,000 Hab/Eq. The orchard sites surrounding the river are supplied  
123 directly with water diverted or pumped from the river.

124 The outlet water from a sustainable wetland system treating the urban WWTP effluents, was selected  
125 as the reclaimed water source. The WWTP uses mechanical treatment to remove sand and large  
126 particles and a conventional activated sludge to digest the organic matter. About 70% of the  
127 secondary effluent is chlorinated and directly discharged into the River Congost, while the remaining  
128 30% is discharged into the constructed wetland system where its nitrogen and phosphorus contents are  
129 reduced by natural plant absorption. This system produces 120,000 m<sup>3</sup> of recycled water every year,  
130 which nowadays is mainly used to irrigate green areas. Raw sewage was directly collected from the  
131 urban WWTP.

132 All microorganisms were concentrated using a standardized SMF protocol (Calgua et al., 2013a,  
133 2013b; Gonzales-Gustavson et al., 2017). This protocol achieves high recovery efficiencies, with low  
134 variability for selected fecal indicators and a wide range of bacterial, viral and parasitic pathogens

135 (Gonzales-Gustavson et al., 2017). Briefly, irrigation water samples (10L), as well as raw sewage  
136 (500ml), were acidified to a pH 3.5, using 1 N HCl. Conductivity was also measured and adjusted  
137 with artificial sea salt (Sigma, Aldrich Chemie GMBH, Steinheim, Germany) to achieve a minimum  
138 conductivity of 1.5 mS/cm<sup>2</sup>. A pre-flocculated skimmed milk solution (PSM) was prepared by  
139 dissolving 10 g of skimmed milk powder (Difco-France) in 1 L of artificial seawater (33,33 gr of sea  
140 salts), adjusting the pH to 3.5. The PSM was then added to the samples to obtain a final concentration  
141 of 0.01 % of skimmed milk. The samples were stirred for 8 h at room temperature, and the flocs were  
142 allowed to settle by gravity for another 8 h. The supernatants were then removed, and the sediment  
143 was collected and transferred to 500-mL centrifuge containers and centrifuged at 8,000 g for 30 min at  
144 4 °C. Pellets were suspended in 5 mL of 0.2 M phosphate buffer, pH 7.5 (1:2, v/v of 0.2 M Na<sub>2</sub>HPO<sub>4</sub>  
145 and 0.2 M NaH<sub>2</sub>PO<sub>4</sub>). The samples were pooled according to the season they were collected in,  
146 distributed in refrigerated boxes among the collaborating laboratories and stored at -20 °C.

#### 147 *2.2. Viral DNA and RNA pre-NGS amplification and bioinformatics analysis*

148 Season-pooled concentrates (900 µL) were filtered through 0.45 µm Sterivex filters (Millipore). Free  
149 viral DNA and non-viral DNA was removed using TURBO DNA-free Kit (Ambion). The nucleic  
150 acids (NA) of viral origin were then extracted with the Viral RNA Mini Kit (Qiagen), without the  
151 carrier RNA. Sequence-independent, single-primer amplification (SISPA) was applied, as previously  
152 described (Fernandez-Cassi et al., 2018). Briefly, NAs were retrotranscribed using SuperScript III  
153 (Life Technologies) and tagged with random Primer A for the detection of both RNA and DNA  
154 viruses (Table 1). The second cDNA strand was constructed using Sequenase version 2.0 Kit  
155 (USB/Affymetrix). PCR amplification (10 min at 95°C, 25 cycles of 30 s at 94°C, 30 s at 40°C, and  
156 30 s at 50 °C, and a final step of 1 min at 70°C) was performed using the AmpliTaq Gold Master Mix  
157 (Life Technologies) and Primer B to obtain enough dsDNA for library preparation (Table 1). PCR  
158 products (100 µL) were cleaned and concentrated into smaller volumes (15 µL) with the DNA Clean  
159 & Concentrator Kit (Zymo research). Amplified DNA samples were quantified using Qubit 2.0  
160 fluorometer (Life Technologies, Oregon, USA) and libraries were constructed using a Nextera XT  
161 DNA Sample Preparation kit (Illumina Inc.) following the manufacturer's instructions. Samples were

162 sequenced on an Illumina MiSeq platform,  $2 \times 300$  bp, producing paired-end reads.

163 The quality of the raw and clean sequences was assessed using the FASTX-Toolkit version 0.0.14  
164 software (Hannon Lab; <http://www.hannonlab.org>). The sequences were cleaned using Trimmomatic  
165 version 0.32 (Bolger et al., 2014), with the trimming of low-quality reads presenting a Phred score  
166 above Q15 over a running window of 4 nucleotides. Duplicated reads were removed in a subsequent  
167 step to accelerate the downstream assembly. Viral reads were assembled into contigs, based on 90%  
168 identity over a minimum of 50% of the read length (coverage), using CLC Genomics Workbench 4.4  
169 (CLC bio USA). The resulting contigs were queried for sequence similarity, using NCBI-BLASTN  
170 against the NCBI Viral Genomes database (Brister et al., 2015) and the viral sequences from the  
171 GenBank database (Benson et al., 2015). On the other hand, NCBI-BLASTX was also run to compare  
172 the translated nucleotide contig sequences against the viral protein sequences from the UniProt  
173 database (UniProt Consortium, 2015). Based on the best BLAST result from those three comparisons  
174 on each contig, considering a minimum 90% identity cut-off and an E-value lower than  $10^{-5}$ , each  
175 sequence was classified into its likely taxonomic group.

### 176 *2.3. NGS of bacterial populations*

177 Total DNA was extracted from 300  $\mu$ l of the SMF concentrates using the DNeasy PowerSoil kit  
178 (Qiagen), following the manufacturer's instructions and was measured using Qubit 2.0 (Life  
179 Technologies, Invitrogen, Oregon, USA). The metagenomic analysis targeted the V3-V4 region of the  
180 16S rRNA gene with previously described primers (Claesson et al., 2009; Herlemann et al., 2011) and  
181 the overhang adaptor sequences from Illumina were added to the locus-specific primer for the targeted  
182 region (Table 1). Paired-end libraries were built using the Nextera XT DNA sample preparation kit  
183 (Illumina Inc.) and sequenced on an Illumina MiSeq platform,  $2 \times 250$  bp. End reads with a quality  
184 Phred score above Q30 were trimmed with QIIME 1.9.1. (Caporaso et al., 2010) and the cleaned reads  
185 were merged with the FLASH software (Magoč and Salzberg, 2011). All sequences with ambiguous  
186 base calls were also discarded. Taxonomic classification was performed using the SILVA database  
187 (Quast et al., 2012), with a 95% cut-off for sequence identity.

188        *2.4. NGS of protozoan populations*

189    A volume of 300 µl of each SMF concentrate was lysed using the FastPrep®-24 instrument (MP  
190    Biomedicals). DNA was extracted with the FastDNA® SPIN Kit (MP Biomedicals) for soil, according  
191    to the manufacturer's instructions. A modified protocol for the lysing step was followed, as described  
192    previously (Shields et al., 2013). Samples were first homogenized for 60 s. After the bead beating  
193    step, samples were placed on ice for 1 min and then homogenized for another 60 s. The final DNA  
194    products were eluted in a final volume of 50 µL. All the DNA extractions were performed in  
195    duplicate. Afterwards, the duplicate samples were mixed and purified with the OneStep™ PCR  
196    Inhibitor Removal Kit (Zymo Research) to remove inhibitors that could have affected downstream  
197    enzymatic reactions such as the PCR. Genomic DNA concentrations were measured with the Qubit®  
198    dsDNA BR Assay Kit (Thermo Fisher Scientific). The 260:280 ratio was measured using the  
199    NanoDrop ND-1000 UV/Vis spectrophotometer (Thermo Fisher Scientific). The amplicon sequencing  
200    protocol for protozoa, targeting the hypervariable V4 region of the 18S rRNA gene, was performed as  
201    previously described (Moreno et al., 2018), using the primers described in Table 1. Briefly, DNA  
202    amplicon libraries were built using the Nextera XT DNA Sample Preparation Kit (Illumina Inc.) and  
203    sequenced on an Illumina MiSeq platform. The obtained reads were analyzed using QIIME 1.9.1  
204    (Caporaso et al., 2010), applying additional scripts that are available in the Microbiome Helper  
205    Virtual Box (Comeau et al., 2017). OTUs were defined at the 97% similarity cut-off. The Protist  
206    Ribosomal Reference database pr2\_gb203\_version\_4.5 fasta (Guillou et al., 2013) was used as the  
207    reference.

208        *2.5. Statistics and heatmaps*

209    The richness and diversity indices were calculated using the Catchall version 4.0 software (Bunge et  
210    al., 2012), which included the species richness estimate (Chao1) and the Shannon diversity index as  
211    measures of the number of different species detected and the relative abundances of the species  
212    present in a sample, respectively. Heatmaps were generated using the R library ggplot2 graphics  
213    (Kolde, 2015).

214

### 215 3. Results

#### 216 3.1. Output of the metagenomic libraries

217 In this study, the viral, bacterial and protozoan communities present in the different irrigation water  
218 sources were evaluated using a metagenomic approach. The total number of reads obtained as well as  
219 the richness and diversity indices are summarized in Table 2. In total, 95.13% of the Illumina viral  
220 reads were assembled into contigs. The highest viral richness and diversity were observed in raw  
221 sewage samples, particularly those collected in winter. Drinking water samples presented the lowest  
222 number of viral reads compared to the other water sources. Regarding bacterial populations, reservoir  
223 water samples as well as river water samples collected in the summer and spring showed the highest  
224 richness and diversity values. For the reservoir, groundwater and river water samples, the highest  
225 Chao values and Shannon indices for bacteria were obtained in the samples collected during summer  
226 (Table 2). After quality filtering, trimming and detection of PCR-chimeras, the highest number of raw  
227 reads (110.766) obtained from the analysis of protozoa mock community was obtained for the pooled  
228 samples of reclaimed water collected in the summer season. Due to the wide range of raw reads,  
229 obtained from the 18S rRNA amplicon-based analysis, and the relatively low abundances of  
230 pathogenic protozoa, it was not possible to calculate richness and diversity values.

231

#### 232 3.2. Virome of the different irrigation water sources and raw sewage

233 The diversity of the viral species detected is graphically represented in Figure 1 according to the viral  
234 family and water source. Human viral pathogens were detected in all the different types of samples  
235 analyzed, except for drinking and reservoir water (Table 3). Over 67% of the viral metagenomic  
236 sequences showed significant similarity to the viral sequences stored in one or more of the three  
237 databases used (NCBI Viral Genomes database, GenBank and UniProt). Most of the known viral  
238 species (> 77%) detected in the conventional irrigation water sources were plant viruses or  
239 bacteriophages. Bacteriophages belonging to the *Siphoviridae* (dsDNA), *Myoviridae* (dsDNA),  
240 *Podoviridae* (dsDNA) and *Microviridae* (ssDNA) families were always the most represented in all the  
241 irrigation water sources.

242 The 95% of the viral reads from the drinking water samples corresponded to phages infecting  
243 different bacteria, including *Arthrobacter* spp., *Propionibacterium* spp. and *Pseudomonas* spp. Up to  
244 39 different species belonging to 6 phage families were detected in this type of water. NGS did not  
245 detect any waterborne pathogens in drinking water (Table 3), with only a few contigs identified that  
246 corresponded to the *Myoviridae* and *Microviridae* phage families and *Phycodnaviridae*.  
247 *Phycodnaviridae* (dsDNA), which includes viruses that infect freshwater eukaryotic algae and  
248 phytoplankton, was very prevalent in the reservoir, groundwater and river water samples. Viruses  
249 belonging to *Virgaviridae* (ssRNA+) and *Tombusviridae* (ssRNA+) were highly abundant in the river  
250 water samples collected in the autumn and winter, whereas a greater diversity of viruses from the  
251 *Phycodnaviridae* family were detected in groundwater in summer and river water in spring. A large  
252 number of viruses affecting insects, including members of the *Iridoviridae* (dsDNA) and *Parvoviridae*  
253 (ssDNA) families that infect cockroaches, flies and mosquitoes, were mainly detected in the river and  
254 reservoir water samples. The drinking water samples had higher proportions of known viral sequences  
255 (77%) than river (68%) or reclaimed water (67%) samples; however, a large fraction of the viral reads  
256 could not be assigned. Reclaimed water samples presented a mixture of bacteriophages, plant viruses  
257 and viruses infecting invertebrates (e.g., aphids and flies), but no human pathogens. Less than 1% of  
258 the viral reads corresponded to non-pathogenic viruses that infect humans. More than 30 different  
259 viral families were detected in the urban sewage samples, with the majority of the contigs belonging  
260 to the *Siphoviridae* and *Myoviridae* families, which showed the highest diversity in the samples  
261 collected during the summer (167 and 103 different species, respectively) and winter (172 and 144  
262 species, respectively). When analyzing the data obtained from the raw sewage samples pooled  
263 according to season (Figure 1), pathogenic viruses belonging to the *Caliciviridae* and *Astroviridae*  
264 families were present throughout the year, but the number of contigs showed seasonality, being higher  
265 during the autumn and winter (Table 3). A high diversity of human astrovirus contigs were detected in  
266 the raw sewage samples collected in the winter and autumn (maximum length, 2,483 bp). Human  
267 adenovirus 41 was detected in raw sewage, as well as in groundwater samples collected during the  
268 autumn. Norovirus GI and GII as well as sapovirus GI, GIV and GV, which are important human  
269 pathogens belonging to the *Caliciviridae* family, were found throughout the year in raw sewage

270 samples. However, the number of species, as well as the number of detected reads, was higher during  
271 the winter and autumn. Norovirus GII was also detected in the river water samples collected during  
272 the autumn, corresponding with the peak excretion season. A 396-bp contig belonging to the beta  
273 human papillomavirus 104 was found in the raw sewage sample collected during spring. Within the  
274 *Picornaviridae* family, sequences belonging to human enteroviruses A, B and C were detected in the  
275 raw sewage samples throughout the year. Remarkably, a 653- and an 867-bp contig were detected that  
276 corresponded to EV-71, showing 86% and 94.8% identity over 99% of the sequence coverage,  
277 respectively. Among the viruses belonging to the *Picornaviridae* family that could cause  
278 gastroenteritis, the Aichi virus presented greater abundance during the summer, whereas salivirus was  
279 more prevalent during the winter. Interestingly, hepatitis E virus (HEV) was detected in the raw  
280 sewage and groundwater samples collected in the summer.

281

### 282 *3.3. Bacteriome of the different irrigation water sources and raw sewage*

283 The diversity of the bacteriome varied among the samples. Drinking water as well as samples of  
284 groundwater, river water and reclaimed water collected in the winter and autumn presented the lowest  
285 diversity values (Table 2). Raw sewage diversities did not vary across seasons. Figure 2 shows the  
286 relative abundances of several bacteria that could be related with human infections in each water  
287 sample.

288 The most abundant bacterial genus in drinking water, accounting for 33.2% of the bacterial reads, was  
289 *Pseudomonas* (Figure 2). The genera *Cellvibrio* (27.0%), *Anoxybacillus* (5%) and *Caulobacter* (5%),  
290 which are associated with plant rhizospheres, streams and fresh water, were also prevalent in drinking  
291 water.

292 Most of the bacterial reads (18%) for reservoir water samples corresponded to genera representing  
293 less than 1% abundance in the samples, thus demonstrating high bacterial diversity and richness. Only  
294 the *Bacillus* (17%) and *Enterobacter* (6%) genera presented relatively high abundances in this type of  
295 irrigation water source (Figure 2). The other taxa represented by 4 to 6% of the reads were *Raoultella*,  
296 *Actinobacter* and non-culturable bacteria.

297 Regarding waterborne human pathogens, groundwater samples collected in the autumn presented the  
298 highest abundance of *Enterobacter* (34.9%) (Figure 2), while the genera *Geobacillus* (20%) and  
299 *Thermus* (18.1%) were abundant in the samples collected in the spring. The genus *Pantoea* showed a  
300 relative abundance of 38.1% in the samples collected in winter, but less than 1% abundance in the  
301 samples from the other seasons. Different bacterial genera with relative abundances ranging from 9%  
302 to 14% were observed in the samples collected in summer, which showed the highest values for  
303 bacterial richness and diversity compared to the samples collected in the other seasons (Table 2). The  
304 percentage of non-culturable bacteria ranged from 0.65% in winter to 4.8% in spring.

305 River water samples also presented variations in bacterial richness and diversity between the seasons.  
306 *Limnohabitans*, common planktonic bacteria occurring in river streams, was detected in low  
307 percentages across the seasons (3.9%), but showed increased prevalence in the autumn, with over a  
308 third of the total number of reads corresponding to this bacterial genus (32.5%). The human-  
309 associated *Bacillus*, *Pseudomonas* and *Yersinia* genera showed low abundances throughout the year,  
310 but increased prevalence in the winter (maximum abundances of 23.9%, 21.6% and 12.4%,  
311 respectively) (Figure 2). Similarly, *Flavobacterium* presented increased relative abundance during the  
312 autumn (21.1%). *Enterobacter* and *Serratia* were detected in all the river water samples, but at lower  
313 percentages (Figure 2).

314 The most prevalent bacteria in reclaimed water was *Actinobacteria* (10.8% in spring to 2.5% in  
315 autumn), with bacteria belonging to the *Limnohabitans* genus being relatively abundant throughout  
316 the year. The number of reads corresponding to Rickettsiales bacteria increased only in the winter, but  
317 none of these were associated with pathogenic species. The most prevalent human pathogens were  
318 *Flavobacterium* and *Serratia* bacteria, showing peak abundances during the winter (Figure 2).

319 The number of human-associated bacterial genera ranged from 10 in the winter to 6-7 in the other  
320 seasons for raw sewage samples. Among these genera, *Arcobacter* was the most abundant, with  
321 percentages ranging from 14.9% in the autumn to 34.9% in the spring. However, *Arcobacter* showed  
322 a generally high prevalence throughout the year, being more prevalent in the spring and summer and  
323 showing a small decrease in the colder seasons. *Pseudomonas* was more prevalent in the summer and  
324 winter, decreasing to 1% abundance in the spring. The other less abundant genera were *Acinetobacter*,

325 *Aeromonas*, *Bacteroides* and *Streptococcus*, with some of these containing bacteria that are  
326 pathogenic to humans. Differences within a genus were observed between the seasons. For example,  
327 the genus *Yersinia* showed an abundance of 7.4% in the winter, but an abundance of less than 0.1% in  
328 the other seasons (Figure 2).

329

### 330 3.4. Parasitome of the different irrigation water sources

331 The waterborne protozoan pathogens detected in raw sewage and the different irrigation water sources  
332 are summarized in Table 4. *Blastocystis* sp. was the most prevalent protozoa, while different free-  
333 living amoeba (FLA) such as *Acanthamoeba* spp, *Entamoeba coli*, *E. dispar*, *E. moshkovskii*,  
334 *Naegleria fowleri*, *N. australiensis*, *N. clarki* and members of the *Hartmannellidae* family (including  
335 *Vermamoeba vermiformis*) were also detected. *Naegleria* spp. were most frequently detected in  
336 reservoir water, groundwater, river water and reclaimed water samples collected during the summer.  
337 *N. fowleri*, the causative agent of primary amoebic meningoencephalitis (PAM), was detected in  
338 groundwater. *Cryptosporidium* spp. was detected in reclaimed water samples collected during the  
339 summer at a low relative abundance (< 0.05%). *Giardia* was not detected in any of the samples  
340 analyzed.

341

## 342 4. Discussion

343

344 The metagenomic approach enables the description of viral, bacterial and protozoan populations in a  
345 water sample, providing an opportunity to detect emerging pathogens and develop improved tools for  
346 monitoring the quality of irrigation water. However, it is not an exhaustive detection method, since  
347 pathogens detected by other molecular tools (e.g., q(RT)-PCR) can be missed (Fernandez-Cassi et al.,  
348 2018). Moreover, the detection of pathogens using metagenomics may be hampered by low target  
349 abundances and/or by higher microbial concentrations in the sample. The quantification of bacterial,  
350 protozoan or viral abundances based on the number of reads is in fact a relative number because the  
351 resulting data only indicate values in relation to the total number of microorganisms in a specific  
352 sample. Since metagenomics assays are not suitable for quantification, these results are intended to

353 indicate the global distribution of families/species and the relationship between diversity and  
354 seasonality.

#### 355 *4.1. Raw sewage, the main source of microbial contamination*

356 Season-pooled samples of raw sewage were analyzed as sewage is the most important source of  
357 environmental contamination and has been demonstrated to be a powerful tool for epidemiological  
358 surveillance and risk evaluation (Fernandez-Cassi et al., 2018; Newton et al., 2015). As previously  
359 reported, the majority of viruses identified in the raw sewage virome were bacteriophages (Bibby and  
360 Peccia, 2013; Cantalupo et al., 2011). The low abundance of known human pathogenic viral families  
361 (less than 1%) in raw sewage has also been reported in the literature (Cantalupo et al., 2011;  
362 Fernandez-Cassi et al., 2018).

363 Our metagenomic approach identified important human pathogens in raw sewage that are transmitted  
364 via the fecal-oral route. Human enteroviruses A, B and C, adenoviruses, astroviruses, caliciviruses  
365 and picornaviruses, which mainly cause gastroenteritis, were detected in raw sewage throughout the  
366 year. The higher abundance and diversity of astroviruses and caliciviruses during the colder seasons  
367 are consistent with their well-documented seasonal peak that has been demonstrated by other  
368 molecular tools (Bosch et al., 2014; Haramoto et al., 2005). HEV, an emerging zoonotic virus that  
369 causes acute hepatitis, was detected in raw sewage samples collected during the summer. Although it  
370 was occasionally detected, it is still important to report its presence as there is an increasing number  
371 of foodborne HEV cases being reported in Europe (EFSA BIOHAZ Panel et al., 2017).

372

373 The *Arcobacter* and *Pseudomonas* genera were the most abundant bacterial populations observed in  
374 raw sewage (Figure 2). Whereas *Pseudomonas* was more prevalent in the summer and winter,  
375 *Arcobacter* showed a higher prevalence throughout the year, as has been stated by other studies (Tang  
376 et al., 2016). Most of the bacterial reads corresponded to *A. cryaerophilus*, which is associated with  
377 human and animal disease and has been previously reported to be the predominant *Arcobacter* species  
378 occurring in raw sewage (Ferreira et al., 2015; Figueras et al., 2014; Fisher et al., 2014; Levican et al.,  
379 2016; Salas-Massó et al., 2018).

380 Other pathogenic bacteria such as *Bacteroides* and *Aeromonas* also occurred at high percentages in all  
381 the raw sewage samples analyzed, being more prevalent in those collected in the spring. Interestingly  
382 *Aeromonas* and *Arcobacter* have been found predominating in metagenomics analysis of wastewater  
383 and it have been indicated that this is due to the capacity of this bacteria to multiply in the sewerage  
384 system (McLellan et al., 2010; McLellan and Roguet, 2019).

385 *Blastocystis* spp. were the most prevalent protozoa detected in raw sewage although it should be noted  
386 that the rRNA copies of 18S used to identify protozoa vary in different types of waterborne protozoa  
387 (Moreno et al., 2018). The effects of *Blastocystis* on human health remain controversial (Greige et al.,  
388 2018). *Blastocystis* is a common protozoan parasite found in the human digestive tract and is the most  
389 common eukaryotic parasite of the human intestine (Greige et al., 2018; Nieves-Ramírez et al., 2018;  
390 Souppart et al., 2009). It has been linked to gastrointestinal symptoms, but has also been described as  
391 an asymptomatic colonizer of the normal intestinal microbiota (Nieves-Ramírez et al., 2018).

392 FLA are ubiquitous protozoa that can cause opportunistic and non-opportunistic infections in humans  
393 (Plutzer and Karanis, 2016; Serrano-Luna et al., 2013). FLA such as *Naegleria*, *Acanthamoeba* and  
394 *Vermamoeba* species are ubiquitously distributed worldwide in various warm aquatic environments  
395 (Panda et al., 2015). Waterborne transmission of pathogenic strains of *Acanthamoeba* spp. and  
396 *Naegleria* spp. is of great relevance (Plutzer and Karanis, 2016). In our study, *N. clarki* was the most  
397 frequently detected species in raw sewage.

398 Three *Entamoeba* spp. were identified in our raw sewage samples: *E. moshkovskii*, *E. coli* and *E.*  
399 *dispar*. *E. moshkovskii* frequently occurs in regions where amebiasis is highly prevalent and has been  
400 isolated from wastewater, rivers, lakes and brackish water (Serrano-Luna et al., 2013). *E. moshkovskii*  
401 is primarily considered an FLA, but it can infect humans (Ali et al., 2003; Mahmoudi, 2015). *E.*  
402 *dispar* has traditionally been considered non-pathogenic, but this view is being re-evaluated (Caliari et  
403 al., 2015). Current evidence suggests that the risk factors for acquiring *E. moshkovskii* infections are  
404 similar to those described for *E. histolytica* and *E. dispar* (Serrano-Luna et al., 2013).

405

406 4.2. Diversity of microorganisms in different irrigation water sources

407           4.2.1. *Drinking water*

408 Viruses infecting bacteria and algae were detected in drinking water. Previous studies with potable  
409 water have described latent forms of bacteriophages to be the most abundant viruses in this type of  
410 water source (Rosario et al., 2009). Although our results confirmed that bacteriophages were  
411 detectable after chlorination, there were no human viral pathogens in drinking water.

412 *Pseudomonas* (33.2%) and *Cellvibrio* (27.0%) bacteria were highly abundant in drinking water, as has  
413 been previously reported (Bertelli et al., 2018; El-Chakhtoura et al., 2018). Despite the low abundance  
414 of this species, it is still higher than that reported in other studies (Bertelli et al., 2018). Among the  
415 protozoa, we detected *Acanthamoeba* at low abundances in drinking water. This free-living  
416 opportunistic protozoan parasite has been detected in drinking water used for irrigation and has been  
417 also reported to be present in bottled water (Khan, 2003). Pathogenic protozoa have been detected in  
418 municipal drinking water systems, even in countries with highly regulated water infrastructures,  
419 causing waterborne and food-borne outbreaks (Kulinkina et al., 2016). This might be particularly  
420 alarming in countries where no residual disinfectant is applied to the water (Braeye et al., 2015). The  
421 low abundances of the two pathogenic microorganisms detected in the drinking water samples need to  
422 be analyzed further, given that NGS data present important limitations and that the detected  
423 microorganisms might be inactive.

424           4.2.2. *Reservoir water and river water*

425 Reservoir water and river water showed very similar distributions of the viral families, sharing 97 of  
426 the complete viral genomes. Although the two types of water sources were sampled from different  
427 river basins, it is well known that microbial communities are similar in these two ecosystems (Vaz-  
428 Moreira et al., 2014). The reduced runoff within the Mediterranean region that might be caused by  
429 climate change will affect the circulation patterns of river water into reservoirs and consequently the  
430 proportions of nutrients and organic carbon in these ecosystems (Imek et al., 2011). Thus, viral  
431 diversity may vary between intense rainfall events and long dry periods, when the demand for  
432 irrigation water increases and the limited water circulating into reservoirs increases anoxic conditions.

433 We observed higher viral diversity in the river water samples collected during autumn and winter.  
434 Except in the pooled samples collected in the summer, the most abundant *Microviridae* phage family  
435 detected was the recently described Eel River Basin Microviridae (ERBM), most frequently found in  
436 sediments (Bryson et al., 2015). It has been reported that the microbial community structure and  
437 associated metagenomes, especially for bacteriophage families, can change after rainfall events  
438 (Tseng et al., 2013). Thus, the seasonality in viral diversity for river water samples may be due to the  
439 fact that the river flow decreases during the summer months (from 8 m<sup>3</sup>/s to 1 m<sup>3</sup>/s) and  
440 bacteriophages are not resuspended from river sediments.

441 Regarding plant viruses, it is not clear how these viruses found in irrigation water could infect the  
442 plants being irrigated. Detailed studies are required to elucidate their potential impact on agriculture  
443 (Balique et al., 2015; Colson et al., 2010).

444 While pathogenic *Bacillus* species were not detected in reservoir water, 30% of the *Enterobacter*  
445 reads corresponded to *Enterobacter cloacae*, which is an opportunistic human pathogen and an  
446 indicator of fecal contamination (Chen et al., 2014). Thus, the presence of this bacterial species could  
447 indicate the existence of a source of human fecal contamination. Contigs corresponding to common  
448 fecal indicators, as crAssphage (1 in autumn) and Pepper mild mottle virus (PMMoV) (23 in winter  
449 and 29 in autumn), were detected in riverwater.

450 As stated above, *N. australiensis* and *N. clarki* were detected in reservoir water (0.13%) and river  
451 water (0.008%). *Naegleria* isolation has been reported to be higher during the warmer seasons in  
452 diverse aquatic habitats worldwide (De Jonckheere, 2004; Panda et al., 2015; Pernin et al., 1998). *N.*  
453 *australiensis* can be pathogenic to mice and could affect human health (Kao et al., 2012).

454 It is interesting to note that *Pseudomonas* and *Bacillus* bacteria were predominant in the pooled  
455 samples from winter and summer. Similar results of the relative abundances of the *Flavobacteriaceae*  
456 and *Comamonadaceae* families were also obtained from the nearby river Ter (Lekunberri et al.,  
457 2018).

#### 458 4.2.3. Groundwater

459 It has been reported that viruses can survive for a long time in groundwater and can pass through the  
460 aquifer pores that physically filter out larger pathogens (Hunt et al., 2010; Ogorzaly et al., 2010). In  
461 our study, waterborne excreted pathogens were not frequently detected in groundwater samples,  
462 whereas plant viruses and bacteriophages were widely prevalent. However, we did detect HEV in the  
463 pooled sample of groundwater collected in the summer and human adenovirus (HAdV) in the sample  
464 collected during autumn. HAdV, which indicates human fecal contamination, has been associated  
465 with precipitation events and can persist in the environment for a long time (up to 1,300 days)  
466 (Bradbury et al., 2013; Kauppinen et al., 2018). These results were confirmed by running specific RT-  
467 qPCR (Jothikumar et al., 2006) and sequencing (data not shown). Interestingly, we detected porcine  
468 adenovirus in the groundwater sample collected during the summer (at a concentration of 10E+03  
469 GC/L), confirming porcine fecal pollution (data not shown). HEV, a zoonotic pathogen that can cause  
470 self-limiting or fulminant hepatitis in humans, was detected in groundwater despite its low prevalence  
471 in the studied area. Intensive pig farming occurs in the surrounding areas from where the groundwater  
472 was extracted, suggesting that leakage from the slurry tanks could have contaminated the groundwater  
473 with the zoonotic pathogen.

474 Pooled samples of groundwater showed different distributions of bacterial populations between the  
475 seasons. *Geobacillus* (20.0% in the spring to 0.9% in the summer), followed by *Thermus* and  
476 *Anoxybacillus*, all showed seasonal peaks and were frequently distributed in water springs. The high  
477 abundance of proteobacteria in groundwater has been also reported previously (Hong et al., 2013).

478 It is interesting to point out the relative abundance of *Vermamoeba vermiformis* (2.63%) in the  
479 summer groundwater sample. *V. vermiformis* colonizes water systems and is a reservoir of pathogenic  
480 bacteria such as *Legionella pneumophila* (Fouque et al., 2014). The detection of *N. fowleri* in  
481 groundwater is of particular concern for human health as it is the principal cause of PAM, a rare and  
482 severe disease that causes inflammation and destruction of the brain. The risk of infection depends on  
483 the relatively high concentrations of *N. fowleri* in water (Cabanés et al., 2001). The relative  
484 abundance of *N. fowleri* in the summer groundwater sample was 0.63%, but we were not able to  
485 determine the concentration of the amoeba with our method. *Naegleria* spp. present an additional

486 threat to human health because they can act as vehicles for the multiplication and dispersal of  
487 pathogenic bacteria (Huang and Hsu, 2010).

#### 488 4.2.4. Reclaimed water

489 Plant viruses, including *Virgaviridae*, *Tombusviridae* and *Phycodnaviridae*, and bacteriophages, such  
490 as *Siphoviridae*, *Myoviridae*, *Podoviridae* and *Microviridae*, accounted for more than 90% of the  
491 known viral reads obtained with our reclaimed water samples, representing over 80% of the known  
492 viral species detected (Figure 1). The wetland from where the reclaimed water originated appeared to  
493 restore the naturally occurring microbial communities, as the distribution of the viral families detected  
494 in the reclaimed water samples resembled that of freshwater samples (river and reservoir water). We  
495 identified different circoviruses and parvoviruses in reclaimed water, which agrees with the findings  
496 of other metagenomic studies performed on raw sewage and reclaimed water produced after the  
497 chlorination of treated secondary effluents (Blinkova et al., 2009; Rosario et al., 2009). Viruses  
498 belonging to these two families are highly stable in the environment and have a broad host range that  
499 includes mammals and invertebrates. Given the huge biodiversity in complex ecosystems such as  
500 wetlands, these viruses might not be of concern for human health. CrAssphage (1 in summer) and  
501 PMMoV contigs (5, 21 and 23 in summer, winter and autumn respectively), were detected at the  
502 wetland effluent samples.

503 According to the minimum quality requirements for water reuse in agricultural irrigation, wetland  
504 effluents may be optimal for crop irrigation as class B recycled water, given that it is rich in  
505 macronutrients like nitrogen and phosphorus (reported mean values over the year of sampling: 63  
506 mg/l for N and 10.7 mg/l for P (data not shown)) and no human pathogens were detected.

507 The *Actinobacteria* and *Limnohabitans* genera were predominant in all the reclaimed water samples.  
508 These bacteria inhabit a broad range of freshwater habitats and can constitute up to 30% of the free-  
509 living bacteria in freshwater systems (Zwart et al., 2002). As previously stated for viruses, the  
510 bacteriome of reclaimed water was similar to that of river water, demonstrating that the treatment  
511 used to produce the reclaimed water restores the bacterial populations observed in river water.

512 It should be noted that the prevalence of *Cryptosporidium* spp. and *Giardia intestinalis* might have  
513 been underestimated by our metagenomic approach. These parasites are the two most commonly  
514 occurring protozoa associated with waterborne disease outbreaks (Plutzer and Karanis, 2016) and are  
515 highly prevalent in sewage and reclaimed water. We detected *Cryptosporidium* spp. in the samples  
516 collected during summer. Although the SMF method has demonstrated to be useful to recover  
517 *Giardia* (Gonzales-Gustavson et al., 2017), this protozoan pathogen was not found in any of the  
518 samples analyzed. The high microbial concentrations could have masked the presence of *Giardia*. It is  
519 also interesting to note that *Entamoeba* was detected in reclaimed water. Although the pathogenicity  
520 of the detected *E. moshkovskii* remains unknown, the presence of this protozoa confirms previous  
521 claims that its prevalence might be underestimated (Shimokawa et al., 2012).

#### 522 4.3. Metagenomic approach as a public health tool

523 Our results give a global picture of the microbial communities present in different types of irrigation  
524 water sources produced by using a single concentration method and a sequencing platform. This  
525 approach has huge potential given that NGS techniques are expected to become more affordable and  
526 automatized. However, the use of metagenomics to assess water quality will not replace current water  
527 quality tests in the near future. Increasing the sampling locations and the number of samples  
528 sequenced will contribute to the monitoring of microbial trends and diversity shifts. Furthermore, the  
529 metagenomics approach can be used to detect viable but non-culturable bacteria/protozoa as well as  
530 emerging and known viral pathogens. As these new tools are still in development, several  
531 improvements are needed to avoid possible biases and uncertainties associated with the pre-  
532 amplification step and the loss of information during the filtering and processing of the huge amounts  
533 of data. Further improvements in NGS technologies might enable the study of all 16S/18S rRNA gene  
534 sequences to provide greater resolution for better taxonomical classification. The earlier we identify  
535 the origin of fecal contamination, the better our chances are to find and eliminate the source of  
536 microbial pathogens. Thus, it is critical to have a comprehensive understanding of all the viruses,  
537 bacteria and protozoa (including potential pathogens) circulating in different irrigation water sources,  
538 as this information helps assess water quality.

539 We detected only a few human pathogens in both conventional irrigation water sources (drinking,  
540 reservoir, groundwater and river water) and reclaimed water. Microbial abundances and diversities in  
541 the reclaimed water samples were very similar to those in the river water and reservoir water samples.  
542 In the Mediterranean region, increased temperatures and decreased precipitation will cause a general  
543 decrease in water availability. This may result in pathogens occurring at higher concentrations in  
544 rivers as the river flow rate will not be sufficient to dilute the effluents from WWTPs (Rusiñol et al.,  
545 2015). Seasonal effects have to be carefully evaluated, as one pooled sample per season might result  
546 in overestimation of the seasonal effects. As sequencing costs of the high-throughput platforms are  
547 being reduced, researchers may develop studies with large number of samples per season, which we  
548 consider as one caveat of this study.

549 The Urban Waste Water Treatment Directive (91/271/CEE, 1991) does not require disinfection prior  
550 to discharge despite the fact that WWTP effluents contain elevated levels of pathogens (Naughton and  
551 Rousselot, 2017; Rusiñol and Girones, 2017). Thus, sustainable and green water treatments to produce  
552 reclaimed water are not only a way of producing new irrigation water sources, but also a way of  
553 reducing pathogen levels in the receiving water bodies. Water reuse for irrigation purposes is  
554 consistent with population growth plans and is a chance to close the human water cycle.

555

## 556 5. Conclusions

- 557 • Viral diversity in river water varied between seasons, with bacteriophages becoming more  
558 prevalent during autumn and winter, when norovirus GII strains also become detectable.
- 559 • Reservoir water contained *Enterobacter cloacae*, an opportunistic human pathogen and indicator  
560 of fecal contamination, *Naegleria australiensis* and *Naegleria clarki*.
- 561 • The metagenomic approach revealed the presence of the human pathogens HEV, HAdV and  
562 *Naegleria fowleri* in groundwater samples collected during the summer.
- 563 • A constructed wetland used as a sustainable system to treat secondary effluents from an urban  
564 WWTP seemed to restore naturally-occurring microbial communities as the virome and  
565 bacteriome of the reclaimed water resembled those of freshwater (river and reservoir water).

566

567 **Data availability statement**

568 Raw data can be found in zenodo under the DOI number 10.5281/zenodo.2620427.

569

570 **Acknowledgements**

571

572 This work was supported through a grant funded by the Spanish Ministry of Economy and  
573 Competitiveness (MINECO) in the frame of the collaborative international consortium JPIW2013-  
574 095-C03-01, JPIW2013-095-C03-02 and JPIW2013-095-C03-03 of the Water Challenges for a  
575 Changing World Joint Programming Initiative (Water JPI) Pilot Call and partially by AGL2017-  
576 86797-C2-1-R. Sílvia Bofill-Mas is a Serra-Hunter fellow at the University of Barcelona.

577

578 **References**

579

580 91/271/CEE, 1991. Tratamiento de las aguas residuales.

581 ACA, 2016. Evolució anual del volum d'aigua reutilitzat per usos (hm<sup>3</sup>).

582 Alcalde-Sanz, L., Gawlik, B.M., 2017. Minimum quality requirements for water reuse in agricultural  
583 irrigation and aquifer recharge Towards a legal instrument on water reuse at EU level.

584 <https://doi.org/10.2760/887727>

585 Ali, I.K.M., Hossain, M.B., Roy, S., Ayeh-Kumi, P.F., Petri, W.A., Haque, R., Clark, C.G., Clark,

586 C.G., 2003. Entamoeba moshkovskii infections in children, Bangladesh. Emerg. Infect. Dis. 9,

587 580–4. <https://doi.org/10.3201/EID0905.020548>

588 Aw, T.G., Wengert, S., Rose, J.B., 2016. Metagenomic analysis of viruses associated with field-

589 grown and retail lettuce identifies human and animal viruses. Int. J. Food Microbiol. 223, 50–56.

590 <https://doi.org/10.1016/j.ijfoodmicro.2016.02.008>

591 Balique, F., Lecoq, H., Raoult, D., Colson, P., 2015. Can plant viruses cross the kingdom border and

592 be pathogenic to humans? Viruses 7, 2074–2098. <https://doi.org/10.3390/v7042074>

593 Benson, D.A., Clark, K., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., Sayers, E.W., 2015. GenBank.

594 Nucleic Acids Res. 43, D30-5. <https://doi.org/10.1093/nar/gku1216>

595 Bertelli, C., Courtois, S., Rosikiewicz, M., Piriou, P., Aeby, S., Robert, S., Loret, J.-F., Greub, G.,  
596 2018. Reduced Chlorine in Drinking Water Distribution Systems Impacts Bacterial Biodiversity  
597 in Biofilms. *Front. Microbiol.* 9, 2520. <https://doi.org/10.3389/fmicb.2018.02520>

598 Bibby, K., Peccia, J., 2013. Identification of Viral Pathogen Diversity in Sewage Sludge by  
599 Metagenome Analysis. *Env. Sci Technol* 47, 1945–1951.  
600 <https://doi.org/10.1021/es305181x>.Identification

601 Blinkova, O., Rosario, K., Li, L., Kapoor, A., Slikas, B., Bernardin, F., Breitbart, M., Delwart, E.,  
602 2009. Frequent detection of highly diverse variants of cardiovirus, cosavirus, bocavirus, and  
603 circovirus in sewage samples collected in the United States. *J. Clin. Microbiol.* 47, 3507–13.  
604 <https://doi.org/10.1128/JCM.01062-09>

605 Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence  
606 data. *Bioinformatics* 30, 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>

607 Bosch, A., Pinto, R.M., Guix, S., 2014. Human astroviruses. *Clin. Microbiol. Rev.* 27, 1048–1074.  
608 <https://doi.org/10.1128/CMR.00013-14>

609 Bradbury, K.R., Borchardt, M.A., Gotkowitz, M., Spencer, S.K., Zhu, J., Hunt, R.J., 2013. Source and  
610 transport of human enteric viruses in deep municipal water supply wells. *Environ. Sci. Technol.*  
611 47, 4096–103. <https://doi.org/10.1021/es400509b>

612 Bradley, I.M., Pinto, A.J., Guest, J.S., 2016. Design and evaluation of illumina MiSeq-compatible,  
613 18S rRNA gene-specific primers for improved characterization of mixed phototrophic  
614 communities. *Appl. Environ. Microbiol.* 82, 5878–5891. <https://doi.org/10.1128/AEM.01630-16>

615 Braeye, T., DE Schrijver, K., Wollants, E., van Ranst, M., Verhaegen, J., 2015. A large community  
616 outbreak of gastroenteritis associated with consumption of drinking water contaminated by river  
617 water, Belgium, 2010. *Epidemiol. Infect.* 143, 711–9.  
618 <https://doi.org/10.1017/S0950268814001629>

619 Brister, J.R., Ako-adjei, D., Bao, Y., Blinkova, O., 2015. NCBI Viral Genomes Resource. *Nucleic*  
620 *Acids Res.* 43, D571–D577. <https://doi.org/10.1093/nar/gku1207>

621 Bryson, S.J., Thurber, A.R., Correa, A.M.S., Orphan, V.J., Vega Thurber, R., 2015. A novel sister

622 clade to the enterobacteria microviruses (family Microviridae ) identified in methane seep  
623 sediments. *Environ. Microbiol.* 17, 3708–3721. <https://doi.org/10.1111/1462-2920.12758>

624 Bunge, J., Woodard, L., Bohning, D., Foster, J.A., Connolly, S., Allen, H.K., 2012. Estimating  
625 population diversity with CatchAll. *Bioinformatics* 28, 1045–1047.  
626 <https://doi.org/10.1093/bioinformatics/bts075>

627 Cabanes, P.-A., Wallet, F., Pringuez, E., Pernin, P., 2001. Assessing the Risk of Primary Amoebic  
628 Meningoencephalitis from Swimming in the Presence of Environmental *Naegleria fowleri*. *Appl.*  
629 *Environ. Microbiol.* 67, 2927–2931. <https://doi.org/10.1128/AEM.67.7.2927-2931.2001>

630 Calgua, B., Fumian, T., Rusiñol, M., Rodriguez-Manzano, J., Mbayed, V.V.A., Bofill-Mas, S.,  
631 Miagostovich, M., Girones, R., 2013a. Detection and quantification of classic and emerging  
632 viruses by skimmed-milk flocculation and PCR in river water from two geographical areas.  
633 *Water Res.* 47, 2797–810. <https://doi.org/10.1016/j.watres.2013.02.043>

634 Calgua, B., Rodriguez-Manzano, J., Hundesa, A., Suñen, E., Calvo, M., Bofill-Mas, S., Girones, R.,  
635 2013b. New methods for the concentration of viruses from urban sewage using quantitative  
636 PCR. *J. Virol. Methods* 187, 215–21. <https://doi.org/10.1016/j.jviromet.2012.10.012>

637 Caliari, M., Gomes, M., Neumann, E., Oliveira, F.S., 2015. *Entamoeba dispar*: Could it be pathogenic.  
638 *Trop. Parasitol.* 5, 9. <https://doi.org/10.4103/2229-5070.149887>

639 Cantalupo, P.G., Calgua, B., Zhao, G., Hundesa, A., Wier, A.D., Katz, J.P., Grabe, M., Hendrix,  
640 R.W., Girones, R., Wang, D., Pipas, J.M., 2011. Raw Sewage Harbors Diverse Viral  
641 Populations. *MBio* 2, e00180-11-e00180-11. <https://doi.org/10.1128/mBio.00180-11.Editor>

642 Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N.,  
643 Peña, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E.,  
644 Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky,  
645 J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenko, T., Zaneveld, J., Knight, R.,  
646 2010. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7,  
647 335–6. <https://doi.org/10.1038/nmeth.f.303>

648 Chatziprodromidou, I.P., Bellou, M., Vantarakis, G., Vantarakis, A., Ioanna, C., Chatziprodromidou,  
649 P., 2018. Viral outbreaks linked to fresh produce consumption: a systematic review. *J. Appl.*

650 Microbiol. 124, 932–942. <https://doi.org/10.1111/jam.13747>

651 Chen, H.-L., Lu, J.-H., Wang, H.-H., Chen, S.-J., Chen, C.-J., Wu, K.-G., Tang, R.-B., 2014. Clinical  
652 analysis of *Enterobacter* bacteremia in pediatric patients: A 10-year study. *J. Microbiol.*  
653 *Immunol. Infect.* 47, 381–386. <https://doi.org/10.1016/j.jmii.2013.03.016>

654 Claesson, M.J., O’Sullivan, O., Wang, Q., Nikkilä, J., Marchesi, J.R., Smidt, H., de Vos, W.M., Ross,  
655 R.P., O’Toole, P.W., 2009. Comparative analysis of pyrosequencing and a phylogenetic  
656 microarray for exploring microbial community structures in the human distal intestine. *PLoS*  
657 *One* 4, e6669. <https://doi.org/10.1371/journal.pone.0006669>

658 Collins, R., Kristensen, P., Thyssen, N., 2009. Water resources across Europe — confronting water  
659 scarcity and drought. EEA Report 2/2009. Eea 60. <https://doi.org/10.2800/16803>

660 Colson, P., Richet, H., Desnues, C., Balique, F., Moal, V., Grob, J.-J., Berbis, P., Lecoq, H., Harlé, J.-  
661 R., Berland, Y., Raoult, D., 2010. Pepper mild mottle virus, a plant virus associated with specific  
662 immune responses, Fever, abdominal pains, and pruritus in humans. *PLoS One* 5, e10041.  
663 <https://doi.org/10.1371/journal.pone.0010041>

664 Comeau, A.M., Douglas, G.M., Langille, M.G.I., 2017. Microbiome Helper: a Custom and  
665 Streamlined Workflow for Microbiome Research. *mSystems* 2, e00127-16.  
666 <https://doi.org/10.1128/mSystems.00127-16>

667 D’Auria, G., Artacho, A., Rojas, R., Bautista, J., Méndez, R., Gamboa, M., Gamboa, J., Gómez-Cruz,  
668 R., 2018. Metagenomics of Bacterial Diversity in Villa Luz Caves with Sulfur Water Springs.  
669 *Genes (Basel)*. 9, 55. <https://doi.org/10.3390/genes9010055>

670 De Jonckheere, J.F., 2004. Molecular Definition and the Ubiquity of Species in the Genus *Naegleria*.  
671 *Protist* 155, 89–103. <https://doi.org/10.1078/1434461000167>

672 Drewes, J., Hübner, U., Zhiteneva, V., Karakurt, S., 2017. Characterization of unplanned water reuse  
673 in the EU. Garching.

674 EFSA BIOHAZ Panel, Ricci, A., Allende, A., Bolton, D., Chemaly, M., Davies, R., Salvador, P.,  
675 Escamez, F., Herman, L., Koutsoumanis, K., Lindqvist, R., Nørrung, B., Robertson, L., Ru, G.,  
676 Sanaa, M., Simmons, M., Skandamis, P., Snary, E., Speybroeck, N., Kuile, B. Ter, Threlfall, J.,  
677 Bartolo, I. Di, Johne, R., Pavio, N., Rutjes, S., Poel, W. Van Der, Vasickova, P., Hempen, M.,

678 Messens, W., Rizzi, V., Latronico, F., Girones, R., 2017. Public health risks associated with  
679 hepatitis E virus ( HEV ) as a food-borne pathogen 15. <https://doi.org/10.2903/j.efsa.2017.4886>

680 El-Chakhtoura, J., Saikaly, P.E., van Loosdrecht, M.C.M., Vrouwenvelder, J.S., 2018. Impact of  
681 Distribution and Network Flushing on the Drinking Water Microbiome. *Front. Microbiol.* 9,  
682 2205. <https://doi.org/10.3389/fmicb.2018.02205>

683 EU C163, 2017. Guidance document on addressing microbiological risks in fresh fruits and  
684 vegetables at primary production through good hygiene.

685 European Environment Agency, 2012. Towards efficient use of water resources in Europe.  
686 <https://doi.org/10.2800/95096>

687 Fernandez-Cassi, X., Timoneda, N., Gonzales-Gustavson, E., Abril, J.F., Bofill-Mas, S., Girones, R.,  
688 2017. A metagenomic assessment of viral contamination on fresh parsley plants irrigated with  
689 fecally tainted river water. *Int. J. Food Microbiol.* 257, 80–90.  
690 <https://doi.org/10.1016/j.ijfoodmicro.2017.06.001>

691 Fernandez-Cassi, X., Timoneda, N., Martínez-Puchol, S., Rusiñol, M., Rodríguez-Manzano, J.,  
692 Figuerola, N., Bofill-Mas, S., Abril, J.F., Girones, R., 2018. Metagenomics for the study of  
693 viruses in urban sewage as a tool for public health surveillance. *Sci. Total Environ.* 618, 870–  
694 880. <https://doi.org/10.1016/j.scitotenv.2017.08.249>

695 Ferreira, S., Queiroz, J.A., Oleastro, M., Domingues, F.C., 2015. Insights in the pathogenesis and  
696 resistance of *Arcobacter* : A review. *Crit. Rev. Microbiol.* 42, 1–20.  
697 <https://doi.org/10.3109/1040841X.2014.954523>

698 Figueras, M.J., Levican, A., Pujol, I., Ballester, F., Rabada Quilez, M.J., Gomez-Bertomeu, F., 2014.  
699 A severe case of persistent diarrhoea associated with *Arcobacter cryaerophilus* but attributed to  
700 *Campylobacter* sp. and a review of the clinical incidence of *Arcobacter* spp. *New Microbe New*  
701 *Infect* 2, 31–27. <https://doi.org/10.1002/2052-2975.35>

702 Fisher, J.C., Levican, A., Figueras, M.J., McLellan, S.L., 2014. Population dynamics and ecology of  
703 *Arcobacter* in sewage. *Front. Microbiol.* 5, 525. <https://doi.org/10.3389/fmicb.2014.00525>

704 Fouque, E., Trouilhé, M.-C., Thomas, V., Humeau, P., Héchard, Y., 2014. Encystment of  
705 *Vermamoeba* (*Hartmannella*) *vermiformis*: Effects of environmental conditions and cell

706 concentration. *Exp. Parasitol.* 145, S62–S68. <https://doi.org/10.1016/j.exppara.2014.03.029>

707 Gonzales-Gustavson, E., Cárdenas-Youngs, Y., Calvo, M., da Silva, M.F.M., Hundesa, A., Amorós,  
708 I., Moreno, Y., Moreno-Mesonero, L., Rosell, R., Ganges, L., Araujo, R., Girones, R., 2017.  
709 Characterization of the efficiency and uncertainty of skimmed milk flocculation for the  
710 simultaneous concentration and quantification of water-borne viruses, bacteria and protozoa. *J.*  
711 *Microbiol. Methods* 134, 46–53. <https://doi.org/10.1016/j.mimet.2017.01.006>

712 Gonzales-Gustavson, E., Rusiñol, M., Medema, G., Calvo, M., Girones, R., 2019. Quantitative risk  
713 assessment of norovirus and adenovirus for the use of reclaimed water to irrigate lettuce in  
714 Catalonia. *Water Res.* 153, 91–99. <https://doi.org/10.1016/j.watres.2018.12.070>

715 Greige, S., El Safadi, D., Bécu, N., Gantois, N., Pereira, B., Chabé, M., Benamrouz-Vanneste, S.,  
716 Certad, G., El Hage, R., Chemaly, M., Hamze, M., Viscogliosi, E., 2018. Prevalence and  
717 subtype distribution of *Blastocystis* sp. isolates from poultry in Lebanon and evidence of  
718 zoonotic potential. *Parasit. Vectors* 11, 389. <https://doi.org/10.1186/s13071-018-2975-5>

719 Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., Boutte, C., Burgaud, G., de  
720 Vargas, C., Decelle, J., Del Campo, J., Dolan, J.R., Dunthorn, M., Edvardsen, B., Holzmann, M.,  
721 Kooistra, W.H.C.F., Lara, E., Le Bescot, N., Logares, R., Mahé, F., Massana, R., Montresor, M.,  
722 Morard, R., Not, F., Pawlowski, J., Probert, I., Sauvadet, A.-L., Siano, R., Stoeck, T., Vaultot, D.,  
723 Zimmermann, P., Christen, R., 2013. The Protist Ribosomal Reference database (PR2): a catalog  
724 of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. *Nucleic Acids*  
725 *Res.* 41, D597-604. <https://doi.org/10.1093/nar/gks1160>

726 Haramoto, E., Katayama, H., Oguma, K., Ohgaki, S., 2005. Application of cation-coated filter method  
727 to detection of noroviruses, enteroviruses, adenoviruses, and torque teno viruses in the  
728 Tamagawa River in Japan. *Appl. Environ. Microbiol.* 71, 2403–2411.  
729 <https://doi.org/10.1128/AEM.71.5.2403>

730 Herlemann, D.P., Labrenz, M., Jürgens, K., Bertilsson, S., Waniek, J.J., Andersson, A.F., 2011.  
731 Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *ISME*  
732 *J.* 5, 1571–1579. <https://doi.org/10.1038/ismej.2011.41>

733 Hong, P.-Y., Yannarell, A., Dai, Q., Ekizoglu, M., 2013. Monitoring the perturbation of soil and

734 groundwater microbial communities due to pig 1 production activities 2 Downloaded from.  
735 Appl. Environ. Microbiol. <https://doi.org/10.1128/AEM.03760-12>

736 Huang, S.-W., Hsu, B.-M., 2010. Survey of *Naegleria* and its resisting bacteria-*Legionella* in hot  
737 spring water of Taiwan using molecular method. *Parasitol. Res.* 106, 1395–1402.  
738 <https://doi.org/10.1007/s00436-010-1815-0>

739 Hunt, R.J., Borchardt, M.A., Richards, K.D., Spencer, S.K., 2010. Assessment of sewer source  
740 contamination of drinking water wells using tracers and human enteric viruses. *Environ. Sci.*  
741 *Technol.* 44, 7956–63. <https://doi.org/10.1021/es100698m>

742 Imek, K.Š., Comerma, M., García, J.-C., Nedoma, J., Marcé, R., Armengol, J., 2011. The Effect of  
743 River Water Circulation on the Distribution and Functioning of Reservoir Microbial  
744 Communities as Determined by a Relative Distance Approach. *Ecosystems* 14, 1–14.  
745 <https://doi.org/10.1007/s10021-010-9388-4>

746 Jackson, C.R., Randolph, K.C., Osborn, S.L., Tyler, H.L., 2013. Culture dependent and independent  
747 analysis of bacterial communities associated with commercial salad leaf vegetables. *BMC*  
748 *Microbiol.* 13. <https://doi.org/10.1186/1471-2180-13-274>

749 Jothikumar, N., Cromeans, T.L., Robertson, B.H., Meng, X.J., Hill, V.R., 2006. A broadly reactive  
750 one-step real-time RT-PCR assay for rapid and sensitive detection of hepatitis E virus. *J. Virol.*  
751 *Methods* 131, 65–71. <https://doi.org/10.1016/j.jviromet.2005.07.004>

752 Kao, P.-M., Hsu, B.-M., Chiu, Y.-C., Chen, N.-H., Huang, K.-H., Shen, S.-M., 2012. Identification of  
753 the *Naegleria* Species in Natural Watersheds Used for Drinking and Recreational Purposes in  
754 Taiwan. *J. Environ. Eng.* 138, 893–898. [https://doi.org/10.1061/\(ASCE\)EE.1943-7870.0000549](https://doi.org/10.1061/(ASCE)EE.1943-7870.0000549)

755 Kauppinen, A., Pitkänen, T., Miettinen, I.T., 2018. Persistent Norovirus Contamination of  
756 Groundwater Supplies in Two Waterborne Outbreaks. *Food Environ. Virol.* 10, 39–50.  
757 <https://doi.org/10.1007/s12560-017-9320-6>

758 Khan, N.A., 2003. Pathogenesis of *Acanthamoeba* infections. *Microb. Pathog.* 34, 277–85.  
759 [https://doi.org/10.1016/S0882-4010\(03\)00061-5](https://doi.org/10.1016/S0882-4010(03)00061-5)

760 Kolde, R., 2015. Pheatmap: Pretty Heatmaps.

761 Kulinkina, A. V, Shinee, E., Guzmán Herrador, B.R., Nygård, K., Schmoll, O., 2016. The situation of

762 water-related infectious diseases in the pan-european region. Copenhagen. <https://doi.org/ISBN>  
763 9 789289 052023

764 Leff, J.W., Fierer, N., 2013. Bacterial Communities Associated with the Surfaces of Fresh Fruits and  
765 Vegetables. *PLoS One* 8, e59310. <https://doi.org/10.1371/journal.pone.0059310>

766 Lekunberri, I., Balcázar, J.L., Borrego, C.M., 2018. Metagenomic exploration reveals a marked  
767 change in the river resistome and mobilome after treated wastewater discharges. *Environ. Pollut.*  
768 234, 538–542. <https://doi.org/10.1016/j.envpol.2017.12.001>

769 Levican, A., Collado, L., Figueras, M.J., 2016. The Use of Two Culturing Methods in Parallel  
770 Reveals a High Prevalence and Diversity of *Arcobacter* spp. in a Wastewater Treatment Plant.  
771 *Biomed Res. Int.* 2016, 1–9. <https://doi.org/10.1155/2016/8132058>

772 Magoč, T., Salzberg, S.L., 2011. FLASH: fast length adjustment of short reads to improve genome  
773 assemblies. *Bioinformatics* 27, 2957–63. <https://doi.org/10.1093/bioinformatics/btr507>

774 Mahmoudi, M.R., 2015. Genotyping of *Giardia lamblia* and *Entamoeba* spp from river waters in Iran.  
775 *Parasitol. Res.* <https://doi.org/10.1007/s00436-015-4702-x>

776 McLellan, S.L., Huse, S.M., Mueller-Spitz, S.R., Andreishcheva, E.N., Sogin, M.L., 2010. Diversity  
777 and population structure of sewage-derived microorganisms in wastewater treatment plant  
778 influent. *Environ. Microbiol.* 12, 378–92. <https://doi.org/10.1111/j.1462-2920.2009.02075.x>

779 McLellan, S.L., Roguet, A., 2019. The unexpected habitat in sewer pipes for the propagation of  
780 microbial communities and their imprint on urban waters. *Curr. Opin. Biotechnol.* 57, 34–41.  
781 <https://doi.org/10.1016/j.copbio.2018.12.010>

782 Moreno, Y., Moreno-Mesonero, L., Amorós, I., Pérez, R., Morillo, J.A., Alonso, J.L., 2018. Multiple  
783 identification of most important waterborne protozoa in surface water used for irrigation  
784 purposes by 18S rRNA amplicon-based metagenomics. *Int. J. Hyg. Environ. Health* 221, 102–  
785 111.

786 Naughton, C., Rousselot, O., 2017. management of risk from excreta and wastewater activated sludge,  
787 in: J.B. Rose and B. Jiménez-Cisneros (Ed.), *GLOBAL WATER PATHOGEN PROJECT*.

788 Newton, R.J., McLellan, S.L., Dila, D.K., Vineis, J.H., Morrison, H.G., Murat Eren, A., Sogin, M.L.,  
789 2015. Sewage reflects the microbiomes of human populations. *MBio* 6, 1–9.

790 <https://doi.org/10.1128/mBio.02574-14>

791 Ng, T.F.F., Marine, R., Wang, C., Simmonds, P., Kapusinszky, B., Bodhidatta, L., Oderinde, B.S.,  
792 Wommack, K.E., Delwart, E., 2012. High Variety of Known and New RNA and DNA Viruses  
793 of Diverse Origins in Untreated Sewage. *J. Virol.* 86, 12161–12175.  
794 <https://doi.org/10.1128/JVI.00869-12>

795 Nieves-Ramírez, M.E., Partida-Rodríguez, O., Laforest-Lapointe, I., Reynolds, L.A., Brown, E.M.,  
796 Valdez-Salazar, A., Morán-Silva, P., Rojas-Velázquez, L., Morien, E., Parfrey, L.W., Jin, M.,  
797 Walter, J., Torres, J., Arrieta, M.C., Ximénez-García, C., Finlay, B.B., 2018. Asymptomatic  
798 Intestinal Colonization with Protist *Blastocystis* Is Strongly Associated with Distinct  
799 Microbiome Ecological Patterns. *mSystems* 3. <https://doi.org/10.1128/mSystems.00007-18>

800 Ogorzaly, L., Bertrand, I., Paris, M., Maul, A., Gantzer, C., 2010. Occurrence, survival, and  
801 persistence of human adenoviruses and F-specific RNA phages in raw groundwater. *Appl.*  
802 *Environ. Microbiol.* 76, 8019–8025. <https://doi.org/10.1128/AEM.00917-10>

803 Pachepsky, Y., Shelton, D.R., Mclain, J.E.T.T., Patel, J., Mandrell, R.E., 2011. Irrigation Waters as a  
804 Source of Pathogenic Microorganisms in Produce. A Review, 1st ed, *Advances in Agronomy*.  
805 Elsevier Inc. <https://doi.org/10.1016/B978-0-12-386473-4.00007-5>

806 Panda, A., Khalil, S., Mirdha, B.R., Singh, Y., Kaushik, S., 2015. Prevalence of *Naegleria fowleri* in  
807 Environmental Samples from Northern Part of India. *PLoS One* 10, e0137736.  
808 <https://doi.org/10.1371/journal.pone.0137736>

809 Parsons, L.R., Wheaton, T.A., Castle, W.S., 2001. High application rates of reclaimed water benefit  
810 citrus tree growth and fruit production. *HortScience* 36, 1273–1277.

811 Pérez, S., Köck, Ma., Tong, L., GInebreda, A., López-Serna, R., Postigo, C., Brix, R., López de Alda,  
812 M., Petrovic, M., Wang, Y., Barceló, D., 2011. Wastewater Reuse in the Mediterranean Area of  
813 Catalonia, Spain: Case Study of Reuse of Tertiary Effluent from a Wastewater Treatment Plant  
814 at el Prat de Llobregat (Barcelona), in: Barceló, D., Petrović, M., Afferden, M. (Eds.), *Waste*  
815 *Water Treatment and Reuse in the Mediterranean Region*. Springer, pp. 249–294.

816 Pernin, P., Pélandakis, M., Rouby, Y., Faure, A., Siclet, F., 1998. Comparative recoveries of  
817 *Naegleria fowleri* amoebae from seeded river water by filtration and centrifugation. *Appl.*

818 Environ. Microbiol. 64, 955–9.

819 Plutzer, J., Karanis, P., 2016. Neglected waterborne parasitic protozoa and their detection in water.  
820 Water Res. 101, 318–332. <https://doi.org/10.1016/j.watres.2016.05.085>

821 Poretsky, R., Rodriguez-R, L.M., Luo, C., Tsementzi, D., Konstantinidis, K.T., 2014. Strengths and  
822 Limitations of 16S rRNA Gene Amplicon Sequencing in Revealing Temporal Microbial  
823 Community Dynamics. PLoS One 9, e93827. <https://doi.org/10.1371/journal.pone.0093827>

824 Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glöckner, F.O.,  
825 2012. The SILVA ribosomal RNA gene database project: improved data processing and web-  
826 based tools. Nucleic Acids Res. 41, D590–D596. <https://doi.org/10.1093/nar/gks1219>

827 RD 1620, 2007. Real Decreto 1620/2007. Régimen jurídico de la reutilización de las aguas  
828 regeneradas, 2007. Boletín Oficial del Estado 294-21092.

829 Rosario, K., Nilsson, C., Lim, Y.W., Ruan, Y., Breitbart, M., 2009. Metagenomic analysis of viruses  
830 in reclaimed water. Environ. Microbiol. 11, 2806–2820. <https://doi.org/10.1111/j.1462-2920.2009.01964.x>

831

832 Rusiñol, M., Fernandez-Cassi, X., Timoneda, N., Carratalà, A., Abril, J.F.J.F., Silvera, C., Figueras,  
833 M.J.M.J., Gelati, E., Rodó, X., Kay, D., Wyn-Jones, P., Bofill-Mas, S., Girones, R., 2015.  
834 Evidence of viral dissemination and seasonality in a Mediterranean river catchment:  
835 Implications for water pollution management. J. Environ. Manage. 159, 58–67.  
836 <https://doi.org/10.1016/j.jenvman.2015.05.019>

837 Rusiñol, M., Girones, R., 2017. Summary of Excreted and Waterborne Viruses, in: Rose, J.B.,  
838 Jiménez-Cisneros, B. (Eds.), Global Water Pathogens Project. Michigan State University, E.  
839 Lansing, MI, UNESCO, Michigan.

840 Salas-Massó, N., Figueras, M.J., Andree, K.B., Furones, M.D., 2018. Do the Escherichia coli  
841 European Union shellfish safety standards predict the presence of Arcobacter spp., a potential  
842 zoonotic pathogen? Sci. Total Environ. 624, 1171–1179.  
843 <https://doi.org/10.1016/j.scitotenv.2017.12.178>

844 Serrano-Luna, J., Piña-Vázquez, C., Reyes-López, M., Ortiz-Estrada, G., de la Garza, M., 2013.  
845 Proteases from Entamoeba spp. and Pathogenic Free-Living Amoebae as Virulence Factors. J.

846 Trop. Med. 2013, 890603. <https://doi.org/10.1155/2013/890603>

847 Shi, M., Lin, X.-D., Tian, J.-H., Chen, L.-J., Chen, X., Li, C.-X., Qin, X.-C., Li, J., Cao, J.-P., Eden,  
848 J.-S., Buchmann, J., Wang, W., Xu, J., Holmes, E.C., Zhang, Y.-Z., 2016. Redefining the  
849 invertebrate RNA virosphere. *Nature* 540, 539–543. <https://doi.org/10.1038/nature20167>

850 Shields, J.M., Joo, J., Kim, R., Murphy, H.R., 2013. Assessment of three commercial DNA extraction  
851 kits and a laboratory-developed method for detecting *Cryptosporidium* and *Cyclospora* in  
852 raspberry wash, basil wash and pesto. *J. Microbiol. Methods* 92, 51–58.  
853 <https://doi.org/10.1016/j.mimet.2012.11.001>

854 Shimokawa, C., Kabir, M., Taniuchi, M., Mondal, D., Kobayashi, S., Ali, I.K.M., Sobuz, S.U., Senba,  
855 M., Houpt, E., Haque, R., Petri, W.A., Jr, Hamano, S., 2012. *Entamoeba moshkovskii* Is  
856 Associated With Diarrhea in Infants and Causes Diarrhea and Colitis in Mice. *J. Infect. Dis.* 206,  
857 744. <https://doi.org/10.1093/INFDIS/JIS414>

858 Souppart, L., Sancier, G., Cian, A., Wawrzyniak, I., Delbac, F., Capron, M., Dei-Cas, E., Boorom, K.,  
859 Delhaes, L., Viscogliosi, E., 2009. Molecular epidemiology of human *Blastocystis* isolates in  
860 France. *Parasitol. Res.* 105, 413–421. <https://doi.org/10.1007/s00436-009-1398-9>

861 Tang, J., Bu, Y., Zhang, X.-X., Huang, K., He, X., Ye, L., Shan, Z., Ren, H., 2016. Metagenomic  
862 analysis of bacterial community composition and antibiotic resistance genes in a wastewater  
863 treatment plant and its receiving surface water. *Ecotoxicol. Environ. Saf.* 132, 260–269.  
864 <https://doi.org/10.1016/j.ecoenv.2016.06.016>

865 Thebo, A.L., Drechsel, P., Lambin, E.F., Nelson, K.L., 2017. A global, spatially-explicit assessment  
866 of irrigated croplands influenced by urban wastewater flows. *Environ. Res. Lett.* 12.  
867 <https://doi.org/10.1088/1748-9326/aa75d1>

868 Tseng, C.-H., Chiang, P.-W., Shiah, F.-K., Chen, Y.-L., Liou, J.-R., Hsu, T.-C., Maheswararajah, S.,  
869 Saeed, I., Halgamuge, S., Tang, S.-L., 2013. Microbial and viral metagenomes of a subtropical  
870 freshwater reservoir subject to climatic disturbances. *ISME J.* 7, 2374–86.  
871 <https://doi.org/10.1038/ismej.2013.118>

872 Uyaguari-Diaz, M.I., Chan, M., Chaban, B.L., Croxen, M.A., Finke, J.F., Hill, J.E., Peabody, M.A.,  
873 Van Rossum, T., Suttle, C.A., Brinkman, F.S.L., Isaac-Renton, J., Prystajeky, N.A., Tang, P.,

874 2016. A comprehensive method for amplicon-based and metagenomic characterization of  
875 viruses, bacteria, and eukaryotes in freshwater samples. *Microbiome* 4, 20.  
876 <https://doi.org/10.1186/s40168-016-0166-1>

877 Vaz-Moreira, I., Nunes, O.C., Elia, C., Manaia, M., 2014. Bacterial diversity and antibiotic resistance  
878 in water habitats: searching the links with the human microbiome. *FEMS Microbiol Rev* 38,  
879 761–718. <https://doi.org/10.1111/1574-6976.12062>

880 WHO, 2006. GUIDELINES FOR THE SAFE USE OF WASTEWATER , EXCRETA AND  
881 GREYWATER Volume 2 Wastewater use in agriculture.

882 Zwart, G., Crump, B., Kamst-van Agterveld, M., Hagen, F., Han, S., 2002. Typical freshwater  
883 bacteria: an analysis of available 16S rRNA gene sequences from plankton of lakes and rivers.  
884 *Aquat. Microb. Ecol.* 28, 141–155.

885