DEEP-AMPLICON SEQUENCING (DAS) ANALYSIS TO DETERMINE THE PRESENCE OF PATHOGENIC HELICOBACTER SPECIES IN WASTEWATER REUSED FOR IRRIGATION

Irene Hortelanoa1, Yolanda Morenoa*, Laura Moreno-Mesoneroa, María Antonia Ferrúsb.

aResearch Institute of Water and Environmental Engineering (IIAMA), Universitat Politècnica de València, 46022 Valencia, Spain
bBiotechnology Department, Universitat Politècnica de València, 46022 Valencia, Spain

*Corresponding author: Yolanda Moreno. Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia, SPAIN; Tel/fax: +34963877090, Email: ymoren@upvnet.upv.es

1 These authors contributed equally to this work.

Abstract

Wastewater has become one of the most important and least expensive water for the agriculture sector, as well as an alternative to the overexploitation of water resources. However, inappropriate treatment before its reuse can result in a negative impact on the environment, such as the presence of pathogens. This poses an increased risk for environmental safety, which can subsequently lead to an increased risk for human health. Among all the emerging wastewater pathogens, bacteria of the genus Helicobacter are some of the most disturbing ones, since they are directly related to gastric illness and hepatobiliary and gastric cancer. Therefore, the aim of this study was to determine the presence of potentially pathogenic Helicobacter spp. in treated wastewater intended for irrigation. We used a next generation sequencing approach, based on Illumina sequencing in combination with culture and other molecular techniques (qPCR, FISH and DVC-FISH), to analyze 16 wastewater samples, with and without an enrichment step. By culture, one of the direct samples was positive for H. pylori. FISH and DVC-FISH techniques allowed for detecting viable Helicobacter spp., including H. pylori, in seven out of eight samples of wastewater from the tertiary effluents, while qPCR analysis yielded only three positive results. When wastewater microbiome was analyzed, Helicobacter genus was detected in 7 samples. The different molecular techniques used in the present study provided evidence, for the first time, of the presence of species belonging to the genus Helicobacter such as H.
pylori, H. hepaticus, H. pullorum and H. suis in wastewater samples, even after disinfection treatment.

Keywords

Wastewater, Helicobacter spp, molecular techniques, metagenomics, pathogens.

1. Introduction

Water is a scarce resource for a large part of the world population. Around 70% of the world water supplies are used for agriculture (Eslamian, 2016), being the economic sector with the greatest demand of water. The recovery and reuse of treated wastewater is an alternative to the overexploitation of natural water resources. For this reason, wastewater has become one of the most important and less expensive non-conventional water sources for agriculture (Drechsel et al., 2015; Eslamian, 2016).

In countries such as France, Italy, Spain, Cyprus, Israel, Jordan or USA, wastewater is a continuous source for agricultural irrigation (Aquarec 2006; Ndèye et al., 2008; Pedrero et al., 2010; EPA 2012; Kalavrouziotis et al., 2013), and has become a valuable resource. The use of treated wastewater for agriculture irrigation brings several advantages, such as less dependence on natural water sources (Parsons et al., 2010), presence of nutrients that reduce the use of artificial fertilizers (Pedrero et al., 2013b; Sánchez et al., 2014; Vivaldi et al., 2015; Lyu et al., 2016) and higher yield in irrigation (Vivaldi et al., 2015).

However, poor management prior to water reuse may cause negative impacts on the environment, such as the presence of pathogens, which can reach irrigated foods and affect human health (Lazarova et al., 2013).

Despite important progress in wastewater treatment technologies, waterborne human pathogens (viruses, bacteria and protozoa) can be present in the final effluent (Lazarova et al., 2013). Currently, determination of the microbiological quality of water is based on the presence of bacterial indicators, such as coliform bacteria and Escherichia coli. However, it has been reported the existence of emerging pathogens
that are more resistant to water treatments and disinfection processes than indicators.

This is the case of *Helicobacter pylori*, which is more resistant to chlorine than *E. coli* (Johnson et al., 1997), what can explain its survival in wastewater treatment plants (Moreno et al., 2012).

Among all the emerging wastewater pathogens, bacteria of the genus *Helicobacter* are some of the most disturbing ones, since they are directly related to gastric illness and hepatobiliary and gastric cancer. A relevant fact is that species belonging to *Helicobacter* genus can coexist with other pathogenic proteobacteria such as *Campylobacter* and *Arcobacter* in the same sample (Petersen et al., 2007).

The best-known species of *Helicobacter* genus, *H. pylori* is associated with gastrointestinal disorders which include gastritis, peptic ulcer, duodenal ulcer, gastric cancers and lymphoma associated with mucosal lymphoid tissue (MALT) (FDA 2014). Non- *H. pylori* Helicobacter species (NPH) such as *H. hepaticus*, *H. pullorum*, *H. suis* or *H. bilis* have been related to enteritis, ulcerative colitis, autoimmune hepatitis, hepatocarcinoma, chronic liver disease and autoimmune hepatobiliary disease. They are also associated to hepatobiliary and pancreatic cancers. These types of cancer are highly lethal and difficult to diagnose (Casswasll et al., 2010, Boutine et al., Mateos-Muñoz et al., 2013).

Different strategies, such as growing in biofilms or entering into the Viable But Non-Culturable (VBNC) state, have been suggested to be involved in *H. pylori* long-term survival in the environment (Cellini et al., 2008; Santiago et al., 2015). Different studies using molecular techniques have confirmed the presence of *H. pylori* in various aquatic environments, suggesting that contaminated water with human feces may be a reservoir for the pathogen (Twing et al., 2011; Moreno et al., 2012; Bahrami et al., 2013). It has been also demonstrated its survival in chlorinated water (Moreno et al., 2007).

Waterborne transmission has also been proposed for other *Helicobacter* species, including *H. mustelae*, *H. muridarum*, *H. felis*, *H. canadensis*, *H. pullorum*, *H. canis* (Azevedo et al., 2008) and *H. cetorum* (Goldman et al., 2009). Thus, for accurate determination of their epidemiology, it is crucial to determine the occurrence of
Helicobacter spp. in water distribution systems, as well as the role of water in its transmission. However, until now, the presence of NHPH species in wastewater has not been studied.

Cultivation and isolation of Helicobacter spp., especially NHPH, from environmental samples is extremely difficult (Øverby et al., 2016). Molecular techniques such as Polymerase Chain Reaction (PCR), and Fluorescence in situ Hybridization (FISH) have been used to detect Helicobacter spp. from treated wastewater and drinking, coastal, river, lake, recreational and glacial waters (Cunachi et al., 2016; Fernández-Delgado et al., 2016). Despite its high sensitivity, these techniques present a disadvantage, since they are not able to differentiate between viable and nonviable cells, since DNA from living and dead cells as well as extracellular DNA can be amplified (Wang and Levin, 2006; Pathak et al., 2012). Currently, some modifications of molecular conventional techniques, such as PMA q-PCR or Direct Viable Count method combined with FISH (DVC-FISH) (Santiago et al., 2015; Moreno-Mesonero et al., 2016) can be used for discriminating between viable and non-viable Helicobacter cells.

All these molecular assays are based on the detection or quantification of a single pathogen in a sample. Therefore, they do not provide information about the abundance and diversity of significant microorganisms present in the total microbial community (Ye and Zhang 2011). Thus, for knowing the ecology of pathogen in environment, it is very useful to simultaneously study all the microbiota of a sample.

Massive sequencing of high performance (Next Generation sequencing, NGS) is the only technique that allows this goal. Through the selection of specific primers, NGS can analyse phylogenetic groups or species (Qichao et al., 2014). It also allows non-cultivable microorganism’s detection and typing, since it discriminates even single nucleotide polymorphisms. This technology is currently being used to investigate the microbial diversity in aquatic environments, such a wastewater (Ye and Zhang, 2011; Cai and Zhang 2013; Lu et al., 2015,) drinking water (Vierheilig et al., 2015) and watersheds (Ibekwe et al., 2013).

Ilumina high-throughput sequencing technique (Logares et al., 2014) employs surface-attached automated dye sequencing, allowing the identification and characterization of pathogens at genus or even species level. This methodology can be extremely useful
for studying the diversity of pathogens in complex environmental samples and, consequently, to detect potential pathogens, which cannot be detected by culture techniques or other molecular techniques. 

Thus, the objective of the present work was to determine the presence of potentially pathogenic *Helicobacter* spp. in treated wastewater intended for irrigation. To reach this goal, Illumina-based NGS approach, specifically directed to *Helicobacter* genus, in combination with culture and molecular techniques, has been used.

2. Materials and methods

2.1. Sampling

Wastewater samples were collected from an activated sludge wastewater treatment plant in a Mediterranean region of Spain from February 2016 to May 2017. The urban wastewater treatment plant serves a total population of 155,674 equivalent inhabitants (36,625 m³/day). The final effluent is used for irrigation.

A total of 16 water samples were aseptically collected, 8 of them after biological secondary treatment (“S” samples), and 8 after disinfection treatment with chlorine and UV treatment (“T” samples). Samples were refrigerated and analyzed within 2 hours.

2.2. Wastewater sample analysis.

One liter of each wastewater sample was centrifuged at 3,220 g for 20 min. The supernatant was removed and centrifuged again at 8,000 g for 8 min at 4°C to ensure the elimination of all residual water. The resulted pellet was suspended in 10 mL of phosphate-buffered saline (PBS 1X: 130 mM sodium chloride, 10 mM sodium phosphate, pH 7.2). From this, 200 μL were taken for *Helicobacter* spp. culture, 1 mL for qPCR analysis, and another 1 mL for DVC-FISH analysis (“D” samples). An additional aliquot of 5 mL from each sample was incubated in 10 mL of Brucella broth enrichment media (BBL™ (Becton Dickinson, USA), 10% (v/v) fetal bovine serum (Fisher, USA) and Dent selective supplement (Oxoid, UK)) under microaerophilic conditions at 37°C for
24 hours. After the incubation period, the presence of *Helicobacter* spp. was also analyzed by culture, qPCR and FISH ("E" samples).

For Illumina-based NGS analysis, 1 mL of the 8 samples from tertiary effluent (after disinfection), both directly and after the enrichment step, were processed as described below.

### 2.3. Culture of *Helicobacter* spp.

Aliquots of 100 µL of all the samples were spread, before and after enrichment, on Dent selective agar (Blood Agar Base (Pronadisa, Spain), 10% (v/v) defibrinated horse blood (Oxoid, UK), 0.025% (v/v) sodium pyruvate (Fisher, USA), DENT supplement, (Oxoid)) and on Pylori® Agar (PA) (bioMérieux SA, Spain). A filter technique previously described by Steele and McDermot (1984) and modified by Moreno and Ferrús (2012) was used. Briefly, a sterile 0.65 µm pore size membrane (Whatman, Maidstone, England) was placed onto agar plates, and then 100 µL of each sample were placed on the filters. Plates were incubated under microaerophilic conditions at 37°C. After 24 h of incubation, the membranes were removed from the plates, which were incubated for 14 days.

After the incubation period, agar plates were examined for the presence of *Helicobacter* spp. Suspicious colonies were preliminary identified by Gram stain and subsequently analyzed by specific *Helicobacter* spp. rDNA PCR and *H. pylori* VacA qPCR as described below.

### 2.4. Detection of *Helicobacter* spp. and *H. pylori* by FISH and DVC-FISH

FISH and DVC-FISH analysis were performed according to the studies carried out by Moreno *et al.* (2003) and Piqueres *et al.* (2006), respectively.

For FISH, one mL of wastewater sample was centrifuged at 8,000 r.p.m at 4 ºC for 8 min. The supernatant was removed, and the pellet was resuspended in 1 mL of PBS 1X, washed and fixed with three volumes of 4 % paraformaldehyde for 1.5 h at 4ºC. Afterwards, samples were centrifuged and washed with PBS1X buffer and finally resuspended in 1:1 PBS/ethanol (v/v). Samples were stored at -20ºC until their hybridization.
For DVC-FISH analysis, another mL of each sample was added to 9 mL of DVC broth (BBL™ Brucella broth (BAS, Becton Dickinson, USA) supplemented with 5% fetal bovine serum (Fisher, USA) and 0.5 mg/L of novobiocin) and incubated for 24 hours at 37°C under specific microaerophilic conditions. After incubation, DVC broth was centrifuged at 8,000 rpm at 4°C for 8 min and fixed with paraformaldehyde as above described for FISH. Samples were stored at -20°C until their hybridization.

In this work, two different probes, one for Helicobacter spp. and another for H. pylori, were used:

A specific 16S rRNA probe (target position 717-737) designed by Chan et al., (2005) (HEL717-FAM-AGGTCGCCTCGCAATGAGTA) was used to identify Helicobacter spp. among the fixed samples. Previously, the specificity of HEL717-FAM was determinate in silico by “Probe match” online tool against RDP’S collection of 16S rRNA sequences (https://rdp.cme.msu.edu/probematch/search.jsp). The specificity was also determined in vitro by whole-cell hybridization with different reference strains belonging to Helicobacter genus (H. pylori DSM 10242, H. hepaticus DSM 22904, H. fennelliae DSM 7491, H. cinaedi DSM 5359 and H. suis DSM 19735) and other phylogenetic related species, belonging to the genus Arcobacter (A. butzleri NCTC 12481, A. skirrowii NCTC 12713, A. cryarophilus NCTC 11885) and Campylobacter (C. jejuni NCTC 11168). Reference strains were obtained from Leibnuz-Institute DSMZ-German Collection of Microorganisms and Cell Cultures and the National Collection of Types Cultures (NCTC, United Kingdom).

We also used a 16S rRNA oligonucleotide probe (HPY-CTGGAGAGACTAAGCCCTCC) which has proven to be specific for H. pylori detection in environmental samples (Moreno-Mesonero et al., 2020; Vesga et al., 2018)

For both assays, FISH and DVC-FISH, aliquots of 5 µL of each fixed sample were hybridized on diagnostic slides, pretreated with 0.1% gelatin. Slides with fixed samples were dehydrated by serial immersions in 50%, 80% and 100% ethanol for 3 min each. Formamide at a final concentration of 40% and 50 ng of each probe were included in the hybridization buffer (0.9 mL/L NaCl, 0.01 % SDS, 20mm/L Tris-HCl [pH 7.6]). The reaction was completed under dark conditions at 46°C for 1.5 h. Afterwards, the slides
were incubated with 50 mL of washing solution (0.10 M NaCl, 0.02 M HCl-Tris, 0.01 % SDS and 0.005 M EDTA) under dark conditions at 48°C for 15 min. Finally, slides were washed with distilled water and air-dried.

Slides were mounted with VECTASHIELD® Antifade Mounting Medium (VECTOR laboratories, United Kingdom) and visualized using epifluorescence Olympus BX 50 with U-MWIB, U-MWIB and U-MIWG exciter filters. Pictures were taken with the camera Olympus DP-12 camera. For all analysis, a fixed pure culture of the *H. pylori* NCTC 11637 strain was used as a positive control.

### 2.5. Detection of *Helicobacter pylori* by qPCR

DNA extraction from samples was carried out using the GeneJet™ genomic DNA purification kit (ThermoScientific, Germany). Specific *H. pylori* qPCR based on SYBR green I fluorescence was conducted using VacA primers to amplify a 372 bp fragment according to Vesga *et al.* (2018). Reactions were performed in a final volume of 20 μl containing 2 μl of Light-Cycler® FastStart DNA SYBR Green I (Roche Applied Science, Spain), 1.6 μl of MgCl₂ (50 mM), 0.5 μl of each primer (20 mM) and 2 μl of DNA template. The amplification consisted of an initial DNA denaturation step at 95 °C for 10 min, followed by 40 cycles of: 95 °C for 10 s, 62 °C for 5 s and 72 °C for 16 s; and finally, one cycle at 72 °C for 15 s and another at 40 °C for 30 s. Amplifications were conducted by duplicate. DNA from *H. pylori* strain NCTC 11637 was used as a positive control and qPCR mix without DNA served as negative control in all the qPCR analysis.

### 2.6. Amplicon-based metagenomics analysis for detection of *Helicobacter* spp.

#### Evaluation of primers

The primers HS-Forward (5´-CTAACATGCAAGTCGACGA-3) and HS2-Reverse (5´-GTGCTTATTCGTTAGATACCGTCAT-3´) were selected to amplify the V3-V4 region of the 16S rRNA gene of Helicobacteraceae family. The primer HS was designed in this study and HS2 was previously used by Huang *et al.* (2002).
The Oligo Analyzer 3.1 tool (Integrated DNA Technology) was used to determine the HS and HS2 oligonucleotide properties. TestPrime tool (https://arbsilva.de/search/testprime) was used to calculate the coverage of the primer pairs, using Silva SSU r132 database; and the sequence collection RefNR, allowing 0, 1 and 2 mismatches in the union of primer pairs (Klindworth et al., 2013).

Specificity was also evaluated in vitro by conventional PCR as described below. DNA from *H. pylori* DSM 10242, *H. hepaticus* DSM 22904, *H. fennelliae* DSM 7491, *H. cinaedi* DSM 5359 and *H. suis* DSM 19735, as well as reference strains of *Arcobacter* (A. butzleri NCTC 12481, A. skirrowii NCTC 12713, A. cryarophilus NCTC 11885, A. cibarius (environmental isolate)) and *Campylobacter* (*C. jejuni* NCTC 11168) were used to determine the specificity of the pair of primers.

**PCR reaction**

DNA extraction from samples was carried out using the GeneJet™ genomic DNA purification kit (ThermoScientific, Germany). The amplification was performed in a final reaction volume of 25 µL which contained 2.5 µL 10X Buffer, 1.30 µL of MgCl2 (50Mm), 0.75 µL of each deoxynucleotide (10mM), 0.5 µL of each primer (10 mM), 5 U of *Taq* polymerase (IBIAN TECHNOLOGIES S.L, Zaragoza, Spain) and 4 µL of DNA template. Cycling conditions were 95°C for 5 min, followed by a 28-cycle of amplification (denaturation at 95°C for 45s; annealing at an optimum temperature of 59°C for 45s and extension at 72°C for 45 s) and additional extension step for 5 min at 72°C. Annealing temperature range between 55°C to 66°C was tested.

The amplified products were analyzed by 1% agarose gel containing RedSafe™ and observed under UV light.

**High-throughput 16S rRNA Sequencing**

A mock sample was created in order to evaluate both the optimal conditions of the *Helicobacter* spp. 16S rRNA fragment amplification and the bioinformatics pipeline after Illumina sequencing. The mock solution was formed by the mix of different volumes of DNA of each reference strain listed above, in order to give a concentration of 2 ng each into a single tube.
In order to test the suitability of the amplicon-sequencing conditions in the complex sample community, 3 μl of a water sample from the effluent after tertiary treatment (sample W1TD) was inoculated with 7 μl of the mock solution.

Finally, DNA from all the direct and enriched samples taken after disinfection treatment was processed by Illumina amplicon-based sequencing.

All sequencing analysis (mock, sample W1TD inoculated with the mock and wastewater samples) were carried out by using Illumina MiSeq platform through a 2 × 300bp paired-end run at FISABIO sequencing service (Valencia, Spain). The amplicon sequencing protocol targets V3-V4 region of the 16S rRNA gene, using the primers previously described. DNA amplicon libraries were created as explained by Illuminaguide (http://www.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf) using the enzyme KAPA HiFi HotStart (KAPABIOSYSTEMS, USA). PCR conditions were: 95ºC for 3 min, following by 28-cycles of amplification (denaturation at 95ºC for 30 s; annealing at 55ºC for 30 s and extension at 72ºC for 30 s) and an additional extension step for 5 min at 72ºC.

**Bioinformatics analysis.**

QIIME 1.9.1 (Caporaso et al., 2010) was used to analyze the raw sequences obtained from the Illumina sequencing platform. For that, the corresponding scripts used in Microbiome Helper virtual box (Comeau et al., 2017) were applied. Initially, PEAR v0.9.19 (Zhang et al., 2014) was used to join forward and reverse sequences. Moreover, to prove that reads were rightly stitched, FastQC (Andrews, 2010) tool was employed. Afterward, FASTX-Toolkit v0.014 (Gordon 2009) was applied to filter the stitched reads by length and quality score (minimum of Q30 over at least 90% of the read or less than 200 bp and reads with any ambiguous base (“N”) were removed). Following this, FastQC tool was used again to remove the merged sequences with low-quality tails.

The tool VSEARCH v1.11.1 (Rognes et al., 2016) was used to filter the potential chimeric sequences. SortMeRNA v2.0 (Kopylova et al., 2012) and SUMACLUST v1.0.00
(Mercier et al., 2013) methods were used to process the remaining sequences to form clusters using the open QIIME sequences. Operational Taxonomic Units (OUTs) were defined at the 97% genetic similarity cut-off. The Silva v132 high quality ribosomal RNA database was used as the reference (Quast et al., 2013).

OTUs’ sequences which were not identified at species level by applying Silva v132 rRNA database were aligned against NCBI database by using the BLAST online tool (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

3. Results and Discussion

Various routes of transmission for Helicobacter spp. have been proposed, including the oral-oral and fecal-oral route, since this microorganism has been isolated from saliva and feces (Anand et al., 2014; Safaei et al., 2011). Additionally, many studies have confirmed the presence of Helicobacter spp. by molecular techniques in different aquatic environments. In this report, the presence of Helicobacter species was evidenced in wastewater samples after secondary and tertiary treatment by molecular techniques, including NGS, and culture.

3.1. Culture of Helicobacter spp. and H. pylori

All the enriched samples were negative for Helicobacter growth in any of the culture media used. However, characteristic colonies were observed in three plates from direct samples, although it was not possible to isolate the colonies, due to the presence of a mass of non-specific bacterial growth. Thus, to demonstrate the presence of culturable Helicobacter cells, all the growth from these plates where suspicious colonies were observed was recovered, and DNA was extracted and analyzed by Helicobacter spp. PCR and H. pylori qPCR amplicons were subsequently sequenced.

One of the direct samples taken after the tertiary (disinfection) treatment (W6TD) yielded the expected 16S rRNA fragment and VacA gene amplicons by PCR and, after alignment against NCBI database by BLAST tool, the sequence corresponded to the specie Helicobacter pylori, with a sequence identity of 100%.
Currently, no optimal culture medium is known to allow the isolation of *Helicobacter* spp. from feces and other environments such as food and water (Janet et al., 2003). Conventional culture analysis is extremely limited for recovering these bacteria from environmental samples, because of rapid overgrowth of accompanying microbiota. In this work, however, we have observed the occurrence of cultivable *H. pylori* in wastewater after disinfection treatment, what poses some questions about the actual risk for the consumption of vegetables irrigated with this type of water.

### 3.2. *Helicobacter* spp. and *H. pylori* detection by FISH and DVC-FISH.

*In silico* alignment by “Probe match” online tool against RDP’s collection of 16S rRNA sequences showed the specificity of HEL717-FAM probe for *Helicobacter* spp. under the specific conditions described above.

When *in vitro* assays were conducted, all *Helicobacter* strains included in this work showed positive hybridization with the probe. Species belonging to the Campylobacteriales order, such as *Campylobacter* and *Arcobacter*, were tested under the same conditions, yielding negative results.

Under rigorous conditions, the HEL717-FAM probe was able to demonstrate the presence of *Helicobacter* spp. in 11 out of the 16 (69%) enriched samples, 5 from the secondary effluent and 6 after tertiary treatment (Table 1). When DVC-FISH was used for detecting viable cells of *Helicobacter* spp. in direct samples with the same probe, positive results were obtained in 13 out of the 16 (81%) direct samples, 6 of them taken from secondary treatment and 7 after tertiary treatment.

When *H. pylori*-specific HPY probe was used, the organism was evidenced to be present in 10 out of the 16 (69%) enriched samples, 5 from the secondary effluent and 5 after tertiary treatment. DVC-FISH yielded viable elongated (viable) cells in 10 out of the 16 (63%) direct samples, 4 of them coming from secondary treatment and 6 after tertiary treatment.

*Helicobacter* spp. FISH assays yielded two negative samples after secondary treatment (W1SE and W4SE) which became positive after tertiary treatment (W1TE, W4TE). For
DVC-FISH, this fact was observed also in one sample (W1TD). With the use of specific H. pylori HPY probe, again, two FISH negative (W1SE and W4SE) and three DVC-FISH negative (W1SD, W4SD and W6SD) secondary samples were positive after tertiary treatment. This could be due, either because the massive presence of high amounts of microbiota in samples from the secondary, what hides the presence of fluorescent *Helicobacter* cells to the microscope operator, or to the fact that the samples were contaminated after going outside the secondary tank, during their passing through the tertiary equipment. In fact, *H. pylori* can maintain in equipment and tubes in biofilms, which may act as a reservoir and contamination source for this microorganism (García *et al.*, 2014).

Overall, FISH technique was able to detect *Helicobacter* spp., including *H. pylori*, in 7 out of 8 samples of wastewater from the tertiary effluents, all of them presenting viable, and thus, potentially infective cells. Our findings support that these techniques allow a rapid analysis to investigate the presence and epidemiology of *Helicobacter* spp., improving the safety and quality of reuse wastewater for agriculture purposes.

### 3.3. Detection of *Helicobacter pylori* by qPCR

Six samplings (75%) were positive for *H. pylori* by qPCR (Table 1), 2 of them both after secondary and tertiary treatments, two of them after secondary but not after tertiary and the other two only from tertiary treatment. These findings, in agreement with FISH results, show that the disinfection treatment of wastewater does not eliminate *H. pylori*.

In five samples in which hybridization or NGS showed the presence of *H. pylori* (W1TE, W2TD, W2TE, W4TE, W6TD), qPCR was negative, probably due to the presence of organic compounds in wastewater acting as PCR inhibitors (Li *et al.*., 2014). This is reinforced by the fact that in two of these samples (W1T and W4T) the qPCR was positive prior to enrichment and became negative after this step, which increases the amount of organic material in the template. Thus, although qPCR is a very specific technique, and has proven to be very sensitive for other types of samples (Botes *et al.*, 2013; Clavel *et al.*, 2016) it seems that it can yield false negative results when applied to complex matrices like wastewater.
3.4. Detection of *Helicobacter* spp. by Illumina sequencing

Amplicon-based sequencing is a culture-independent method which allows the simultaneous identification of microorganisms present in a sample, even at species level, and could be an accurate technique in order to detect low represented bacteria, such as pathogenic species, in complex environments (Cao et al., 2017). Moreover, the genus-specific NGS approach opens new perspectives in massive quantitative, specific and sensitive diagnostics (Pereira et al., 2017).

In this work, a region of the V3-V4 segment of the 16S rRNA gene of *Helicobacter* spp. was selected for amplification and deep sequencing through Illumina MiSeq.

**Evaluation of primers**

The primers were analyzed *in silico*, to know their coverage, through the TestPrime program using the Silva SSUr132 database.

In all the cases, *in silico* analysis with the databank sequences demonstrated that the primers allowed for recovering mostly amplified sequences belonging to *Helicobacter* spp.. When no mismatch was allowed, primers aligned against *Helicobacter* spp. with 64.6% coverage. However, when 1 mismatch was allowed in the annealing of the primers, *Helicobacter* spp. showed a coverage of 90.1% and *Campylobacter* sequence coverage was 22.7%, while Epsilonproteobacteraota yielded a coverage of 14.1%.

When 2 mismatches were allowed, sequences coverage increased to 96.2% for *Helicobacter*, 31.3% for Campylobacteraceae, 2.6% for Arcobacteraceae and 19.1% for total Epsilonproteobacteraota. According to this *in silico* analysis, it was not possible to establish an optimal PCR reaction for the only detection of *Helicobacter* spp.

Amplification conditions were then tested *in vitro* by amplification of the DNA of all the reference strains included in the study. At an annealing temperature of 59°C all Helicobacter DNAs were specifically amplified. Nevertheless, the standard annealing temperature of 55°C recommended by Illumina protocol was used for DAS in environmental samples.
Optimization of 16S r RNA amplicon-based sequencing

The potential bias of the primers due to the sample properties and microbial composition was evaluated by comparing the sequences obtained after the amplicon-based sequencing analysis of a mock community, an environmental sample (W1TE) and the same environmental sample inoculated with the mock.

After quality filtering, trimming and elimination of chimeras, the analysis of the three samples yielded 19614 high-quality reads which were clustered in 233 prokaryotic OTUs with assigned taxonomy, assigned at 97% similarity against Silva v132 database.

All species included in the mock community were recovered after sequencing analysis. As expected, a bias towards *Helicobacter* amplification was detected, since the sequences of the other species were underrepresented (Table 2).

In the water sample inoculated with the mock (W1TD+mock), only *Helicobacter* sequences were recovered after sequencing. Amplicon-based sequencing does not seem to be affected by possible inhibitor substances present in the sample, since almost the same number of sequences of *Helicobacter* were recovered from both the inoculated sample and the mock (Table 2).

Some authors have proposed the use of mock communities when NGS is performed in environmental samples. Their inclusion can directly help post-sequencing analyses (Bokulich et al., 2013; Brooks et al., 2015; Pereira et al., 2017). It can also be used to calculate the chimerism rate, the sequencing error rate and the drift in the representation of a community structure (Schloss et al., 2011). In this work, despite the mock was inoculated with the same amount of DNA of each reference strain, the relative abundance of detected *Helicobacter* sequences was much higher than the other species. This fact is even more evident in the results obtained from the sample inoculated with the mock, in which only *Helicobacter* sequences were recovered after amplicon-based sequencing, and shows the important bias of the sequencing procedure with the used primers, designed for favoring *Helicobacter* detection.

However, even with this bias, after the amplicon sequencing analysis of the mock community and water sample, sequences belonging to other families and even other...
Phyla were also recovered, especially other phylogenetically close genus belonging to Campylobacterales order, such as *Arcobacter* or *Campylobacter*. This fact was expected, according to the previous *in silico* and *in vitro* evaluation of the primers, and taking into account the low stringent conditions of the amplification reaction, as described by Illumina, which can allow amplification of more abundant sequences.

**Helicobacter spp. microbiome in treated wastewater samples**

Wastewater samples after tertiary treatment were analyzed by amplicon-based Illumina MiSeq sequencing, both direct and after enrichment procedure. After NGS analysis, as expected by *in silico* and *in vitro* evaluation, it was observed that the primers also allow the identification of some sequences belonging to other phyla different to Proteobacteria.

A total of 746,198 raw sequences were obtained. After filtering by quality, joining paired-end reads and eliminating chimeras, a total of 673,902 high quality sequences remained, which were grouped into 3,078 OTUs (3% cutoff level).

With regard to relative abundance, Proteobacteria phylum was the most dominant in all analyzed samples (31.214% of the total sequences recovered), as expected, followed by Bacteroidetes (23.7%) and Firmicutes (15.01%). This result is according to previous studies that demonstrated that this phylum is the most abundant in municipal wastewater treatment plants (Wagner and Loy, 2002; Wang *et al.*, 2012; Qiao *et al.*, 2015). These results are also consistent with other investigations in which wastewater samples were analyzed by high throughput pyrosequencing and Illumina analysis (Ye and Zhang 2013; Kumaraswamy *et al.*, 2014; Xie *et al.*, 2018). The Epsilonbacteraeota phylum (28.04%) was the most abundant in 5 out of 8 wastewater direct samples. All bacteria belonging to the Epsilonbacteraeota in the analyzed wastewater samples were assigned to the Campylobacterales order. Regarding bacterial genera, *Arcobacter* and *Helicobacter* represented 27.06% of the total reads.

*Helicobacter* spp. sequences were present in 7 out of 8 water samples (Table 3). As the assigned taxonomy at 97% was unable to discriminate between species, the sequences
of each OTU of Helicobacter spp., were aligned against NCBI database by the BLAST online tool.

According to NCBI database, Helicobacter spp. taxonomy was assigned at species level in 20 OTUs. Eight OTUs, corresponding to H. hepaticus, with a sequence identity higher than 97%, were recovered in all samples, except for WT1 and WT8 (Table 3). Sequences identified as H. pylori, were present in samplings W2, W4, W5, W6 and W7. Furthermore, 2 OTUs identified as H. pullorum with a sequence identity of 98% and 99%, were present in samples W4TDE and W7DE, while an only OTU of H. suis was present in sample W5TD. The remaining OTUs, recovered from samplings W1, W4, W5 and W6, were identified only at genus level with a sequence identity higher than 97%.

In two direct samples which were positive for Helicobacter spp. / H. pylori by other molecular methods (W1TD and W3TD), NGS yielded negative results.

Results showed that an enrichment step enhanced H. hepaticus sequences identification, since in 3 negative direct samples (W2TD, W3TD and W7TD), sequences assigned to this genus were recovered after enrichment (Table 3). On the contrary, enrichment seems to be a disadvantage for H. pylori detection by NGS, as observed in samples W6TE and W7TE.

Even though H. pylori is the most known pathogen belonging to the genus Helicobacter and there are numerous recent studies about its transmission and presence in wastewater (Bai et al., 2016; Vesga et al., 2018; Farhadkhani et al., 2019) in this work, the dominant Helicobacter species was H. hepaticus (6.16%), followed by H. pylori (0.19%). A research carried out by Hamada et al. (2009) suggested that viable H. hepaticus can infect the liver, the gallbladder epithelium or the intestine of humans. To our knowledge this is the first time that H. hepaticus is detected in environmental samples. Currently there is no investigation that reports the coexistence of both pathogens in aquatic environments.

Additionally, low relative abundances of the species H. pullorum and H. suis were found. H. pullorum has been associated with Crohn’s disease and cholelithiasis (Fukuda et al., 2002; Bohr et al., 2004; Chen et al., 2007; Karagin et al., 2010). H. suis is the
most predominant gastric NHPH in humans and involves 13.9% to 78.5% of human NHPH infections (Groote et al., 2005; Haesebrouck et al., 2009). The main reservoirs of those two potentially pathogenic NHPH species are pigs and chickens, respectively (Ménard and Smet, 2019). Thus, they could reach water through animal faces. However, our results are too preliminary to obtain conclusions about this subject. There is only a previous study completed by Fernandez-Delgado et al. (2016), in which the presence of both Helicobacter species in freshwater was detected.

Since the aim of this work was to optimize the detection of Helicobacter spp., in the present work we included an enrichment step with the purpose to increase sensitivity and selectivity in the detection of Helicobacter spp. in wastewater samples. Other authors reported the effectiveness of an enrichment process to increase the sensitivity of detection techniques based on PCR (Ahmed et al., 2009; Santiago et al., 2015). However, results have only showed the effectiveness of this procedure for H. hepaticus NGS detection, while for the other techniques there were no significant differences between enriched and direct sample. As a limited number of samples have been examined, these results should be confirmed by more wide studies.

4. Conclusion

To our knowledge this is the first study which reports the presence of Helicobacter spp. in wastewater samples from Spain, comprising disinfection effluents used for irrigation purposes, carried out by different molecular techniques, including 16S rRNA deep sequencing. These results demonstrate the presence of species belonging to the genus Helicobacter such as H. pylori, H. hepaticus, H. pullorum and H. suis in wastewater samples, even after disinfection treatment. This suggests the possible role of wastewater as a vehicle of transmission of pathogenic Helicobacter species to humans through irrigation. The amplicon-sequencing method optimized here showed to be a specific and sensitive method for the simultaneous detection of pathogenic Helicobacter species in environmental samples.
Bibliography:


