Article:
Karunakaran, E., Battarbee, R., Tait, S. et al. (6 more authors) (2024) Integrating molecular microbial methods to improve faecal pollution management in rivers with designated bathing waters. Science of The Total Environment, 912. 168565. ISSN 0048-9697

https://doi.org/10.1016/j.scitotenv.2023.168565

Science of the Total Environment 912 (2024) 168565

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Integrating molecular microbial methods to improve faecal pollution management in rivers with designated bathing waters

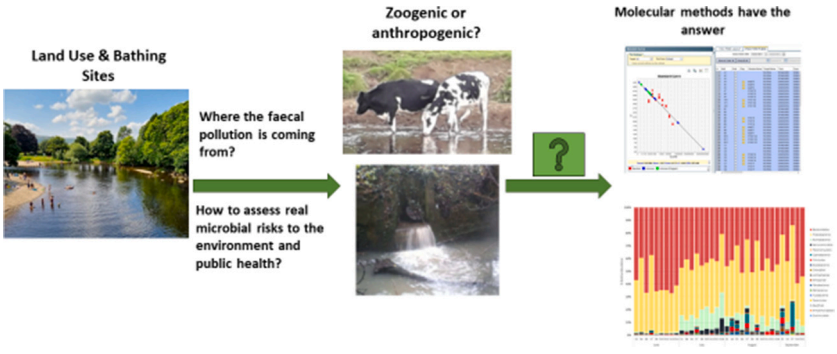
Esther Karunakaran [a], Rick Battarbee [b,c], Simon Tait [d], Bruno Melo Brentan [e], Cathal Berney [c], James Grinham [a], Maria Angeles Herrero [d], Ronex Omolo [d], Isabel Douterelo [d,*]

[a] Department of Chemical and Biological Engineering, University of Sheffield, Sheffield S1 3JD, UK
[b] Environmental Change Research Centre, University College London, London WC1E 6BT, UK
[c] Addingham Environment Group, Addingham, West Yorkshire LS29 0PD, UK
[d] Department of Civil and Structural Engineering, University of Sheffield, Sheffield S1 3JD, UK
[e] Hydraulic Engineering and Water Resource Department, Federal University of Minas Gerais, Belo Horizonte 31270, Brazil

HIGHLIGHTS

- Traditional FIB alone did not accurately describe pathogen presence.
- Different sources of faecal pollution shaped bacterial community structure.
- Land-use had a clear role in pathogens and opportunistic pathogens occurrence.
- Sites affected by faecal pollution were enriched in *Mycobacterium* spp, and *Aeromonas* spp.
- After a heavy rainfall event, an increase in bacteria from agricultural land sources was observed.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Warish Ahmed

Keywords:
Bathing water
Bacterial communities
Faecal pollution
Microbial source tracking

ABSTRACT

Rivers are at risk from a variety of pollution sources. Faecal pollution is of particular concern since it disperses pathogenic microorganisms in the aquatic environment. Currently, faecal pollution levels in rivers is monitored using faecal indicator bacteria (FIB) that do not offer information about pollution sources and associated risks. This study used a combined molecular approach, along with measurements of water quality, to gain information on pollution sources, and risk levels, in a newly designated recreational bathing site in the River Wharfe (UK). Physico-chemical parameters were monitored *in situ*, with water quality multiparameter monitoring sondes installed during the 2021 bathing season. The molecular approach was based on quantitative PCR (qPCR)-aided Microbial Source Tracking (MST) and 16S rRNA gene metabarcoding to obtain a fingerprint of bacterial communities and identify potential bioindicators.

The analysis from the water quality sondes showed that ammonium was the main parameter determining the distribution of FIB values. Lower faecal pollution levels were detected in the main river when compared to tributaries, except for samples in the river located downstream of a wastewater treatment plant. The faecal pollution type (anthropogenic vs. zoogenic) changed the diversity and the structure of bacterial communities,

* Corresponding author.
E-mail address: i.douterelo@sheffield.ac.uk (I. Douterelo).

https://doi.org/10.1016/j.scitotenv.2023.168565
Received 13 August 2023; Received in revised form 23 October 2023; Accepted 11 November 2023
Available online 17 November 2023

giving a distinctive fingerprint that can be used to inform source. DNA-based methods showed that the presence of human-derived bacteria was associated with *Escherichia coli* spikes, coinciding with higher bacterial diversity and the presence of potential pathogenic bacteria mainly of the genus *Mycobacterium*, *Aeromonas* and *Clostridium*. Samples collected after a heavy rainfall event were associated with an increase in *Bacteroidales*, which are markers of faecal pollution, including *Bacteroides graminisolvens*, a ruminant marker associated with surface runoff from agricultural sources. The combined use of qPCR and 16S rRNA sequencing was able to identify pollution sources, and novel bacterial indicators, thereby aiding decision-making and management strategies in recreational bathing rivers.

1. Introduction

Water bodies around the world are under pressure, and one of the main causes of water quality deterioration is faecal pollution, raised by increasing levels of population growth, rapid urban development and climate change (Miller and Hutchins, 2017). In the United Kingdom (UK), around 75 % of rivers in England and Wales do not meet good ecological status (Environment Agency, 2020). Urban influences are key in determining the overall condition of UK rivers, and significant effort has been made in reducing agricultural and industrial pollution in surface waters, yet urban diffuse pollution is responsible for 49 % of failures to meet water quality targets (Defra, 2012). Although water utilities employ wastewater treatment and enhance sewerage infrastructure to manage contaminants (Miller and Hutchins, 2017), the presence of faecal contamination in surface waters can still occur. These sources include discharges from wastewater treatment plants (WWTPs), combined sewer overflows (CSOs), misconnections in pipes, septic tanks, and runoff from manure-fertilised areas, as well as faeces from livestock or wildlife (van Heijnsbergen et al., 2022). Faecal pollution can compromise human health and river ecosystems by enriching the presence of pathogenic microorganisms, such as bacteria, fungi, viruses, and protozoa in the environment (Gerba, 2015; Korajkic et al., 2018). The use of surface water as a drinking water source and for recreation has increased with urbanisation, and pathogens can be transmitted to those who use surface water for recreational uses such as swimming (Xie et al., 2022). Wild swimming and other outdoor recreational activities, in natural bodies of water, saw a significant rise in interest across European countries during the COVID-19 pandemic (Niebaum et al., 2023). This surge in interest in activities like wild swimming can be attributed to lockdowns and restrictions during the pandemic, limiting people's ability to engage in indoor activities. As a result, people turned to outdoor activities as a safe and accessible way to stay active, connect with nature and improve mental well-being. Recreational and wild swimming have been associated with the risk of gastrointestinal, respiratory, ear and eye infections, and dermatological conditions caused by virus, bacteria and protozoa (Chamberlain et al., 2019). The parameters used in the European Union to assess water quality, and to inform risks in designated bathing areas, are legislated in the EU Bathing Water Directive (Directive 2006/7/EC). In the Directive, faecal pollution in bathing sites is monitored for concentrations of indicator bacteria (i.e., *Escherichia coli* and intestinal enterococci) (European Environment Agency, 2023). Following Brexit, the EU directive has been transposed into UK law by Defra, and in the UK the Environment Act, 2021 requires sewerage undertakers to continuously monitor the quality of the receiving water upstream and downstream of their assets (Environment Act, 2021). However, the FIB technique relies on cultivation-based approaches, built on letting *E. coli* and intestinal enterococci (IE), within a discrete river sample, grow in favourable/selective and controlled laboratory conditions, and then counting the number of colonies forming units (CFU) per mL of water. The use of FIB to estimate faecal pollution and associated faecal pathogens in water is useful for stakeholders, and the water industry, to evaluate compliance with legislation. However, the method is more than 150 years old (Holcomb and Stewart, 2020), and has several limitations when compared with molecular methods. Firstly, the method is based on selective quantification of a few indicator species and does not reflect the diversity of pathogenic species in a contaminated environmental sample. FIB typically exhibit a weak correlation with the presence of pathogens in aquatic ecosystems (Saingam et al., 2020). Therefore, there are many microbial risks, which are not identified and measured through the quantification of FIB (Holcomb and Stewart, 2020). Secondly, FIB counts offer no help in tackling pollution, as they do not reveal whether the source of contamination is human-derived or from agricultural runoff. Therefore, there is an urgent need for improved methods and new indicators, to quantify the extent of the different pollution sources and risks, associated with contamination in bathing waters, in order to manage such pollution and better protect public health. A better understanding of the characteristics of microbial communities (i.e., distribution, abundance, and structure) is essential for gaining insight into the potential microbial risks to public health, and to inform bathing water management. The use of molecular approaches can enhance traditional surveillance methods by providing additional information on faecal pollution markers, and other associated - risks including pathogens such as *Salmonella* spp., *Shigella* spp., and *Vibrio* spp. (Stec et al., 2022). Microbial Source Tracking (MST) methods can be used to discriminate between human-derived and non-human sources of faecal contamination. One of the most widely used MST approaches utilises qPCR to quantify host-specific microorganisms (Shanks et al., 2009: Sagova-Mareckova et al., 2021), and *Bacteroides* gene markers are commonly used to discriminate between human and other animal faecal sources (e.g. pig, cattle, dog, etc.) (Mieszkin et al., 2010; Harwood et al., 2014). In addition, sequencing of the 16S rRNA gene can provide a signature/fingerprint of the bacterial community structure that can enable the identification of pollution sources, and specific type of pollution (Douterelo et al., 2014).

This research will focus on a stretch of the River Wharfe in Ilkley that in 2020 was added to the list of bathing sites in England, becoming the first river designated as a bathing site in the UK (Defra, 2020). According to the Ilkley Clean River Group (a group of Ilkley residents), during the summer of 2019 almost 2000 people were observed to use the designated bathing site in Ilkley. The presence of pollution in such a popular recreational site reinforces the importance of addressing environmental concerns to protect public health. In this study, we used a combination of two molecular approaches to: (i) identify the sources and diversity of faecal bacteria at sites upstream and downstream of the designated bathing site; (ii) investigate the bacterial community structure of river and stream samples; and (iii) identify potential pathogens associated with exposure of humans to faecal pollution. The investigation of land uses and main sources of microbial pollutants in this newly designated bathing site will aid to better monitor the river and help the local organisations with statutory duties to protect the river from contamination to ultimately protect the environment and public health. This investigation reinforces the urgent need of initiatives such as the “One planet: one health” call. This initiative is supported by renowned environmental scientists, advocating for the establishment of a global science–policy body dedicated to addressing pollution and its negative effects on both humans and the environment at a global scale (Brack et al., 2022).

The study was not intended to apportion sources for the faecal bacteria at the new bathing water site, nor to provide a definitive assessment of the health risks of bathing at the site. More data from samples collected under different weather patterns and flow conditions would be

needed to provide a precise pollution source apportionment in the area. Our objective is to illustrate a range of molecular methods (e.g., qPCR and sequencing) that could be used at this, and other sites in future, to provide a more accurate and comprehensive understanding of health and environmental risks than those currently mandated.

2. Materials and methods

2.1. Study sites and sampling

Fieldwork was conducted on the River Wharfe at Ilkley (West Yorkshire, UK) upstream and downstream of a newly designated bathing water site (Fig. 1, and Table 1, site code S10). In total, 14 sampling sites (Fig. 1, Table 1) along a 16 km stretch of the main river, and on a number of tributary streams, were selected based on FIB data reported in a previous independent study (Battarbee et al., 2020). These were selected to represent different sources of contaminants and optimise the molecular approach for the studied sites. Sites were located in tributaries (S2, S3, S4, S5, S7 and S9) and along the main river, upstream (S1, S6, S8) and downstream (S11, S12, S13 and S14) of the bathing water site (S10). To show overall land uses in the catchment area, sampling sites were visualised by QGIS software (http://www.qgis.org). Open Rivers GIS data website (www.ordnancesurvey.co.uk) and IGISMAP (https://www.igismap.com/) were used to extract River Wharfe and UK administrative boundary shapefiles. The land cover map was obtained from the UK Centre for Ecology & Hydrology website (https://eidc.ac.uk/) (Fig. 1). Potential primary sources of faecal bacteria along the main river include agricultural livestock, a small village Sewage Treatment Works (STW) and septic tanks upstream of S1, and various storm overflows and agricultural sources between S1 and S10.

Sites S11 to S14 occur mainly in an urban area where sources are potentially dominated by discharges from STWs. Six sites were located on tributaries, two of these were upstream (S2) and downstream (S3) of a small STW and one (S4) on a small beck, at its confluence with the Wharfe. Other tributaries included sampling site S5, known to have a high concentration of *E. coli* due to the presence of a large septic tank close to the confluence, site S7 which is draining a sub-catchment with land-use dominated by livestock agriculture, and S9 that is a tributary known to have periodically high concentrations of *E. coli*, thought to be due to a number of misconnections in its urban catchment (Battarbee et al., 2020; Battarbee and Secrett, 2021a, 2021b). The sampling campaign was supported by citizen scientists from the Ilkley Clean River Group, and the Addingham Environment Group, who aided monitoring

Table 1
Description of sampling sites and catchment type.

Sample code	Catchment type	Description
S1	Main river_GW	Agricultural land and small villages
S2	Tributary_GW	Upstream of small STW
S3	Tributary_GW	Downstream of small STW
S4	Tributary_GW	Confluence with R. Wharfe
S5	Tributary_GW	Confluence with R. Wharfe
S6	Main river_U	Upstream of Sewage Pumping Station (SPS)
S7	Tributary_GW	Rural catchment, livestock and septic tanks
S8	Main river_U	Upstream of CSOs
S9	Tributary_U	Downstream of CSO
S10	Main river_U	Main bathing site, upstream of STW
S11	Main river_U	Downstream of STW
S12	Main river_U	Downstream of STW
S13	Main river_U	Downstream of STW and CSOs
S14	Main river GW/U	Downstream of STWs

GW – Grassland/Woodland, U – Urban/Suburban.
STW-Sewage Treatment Works.
CSO-Combined Sewer Overflow.

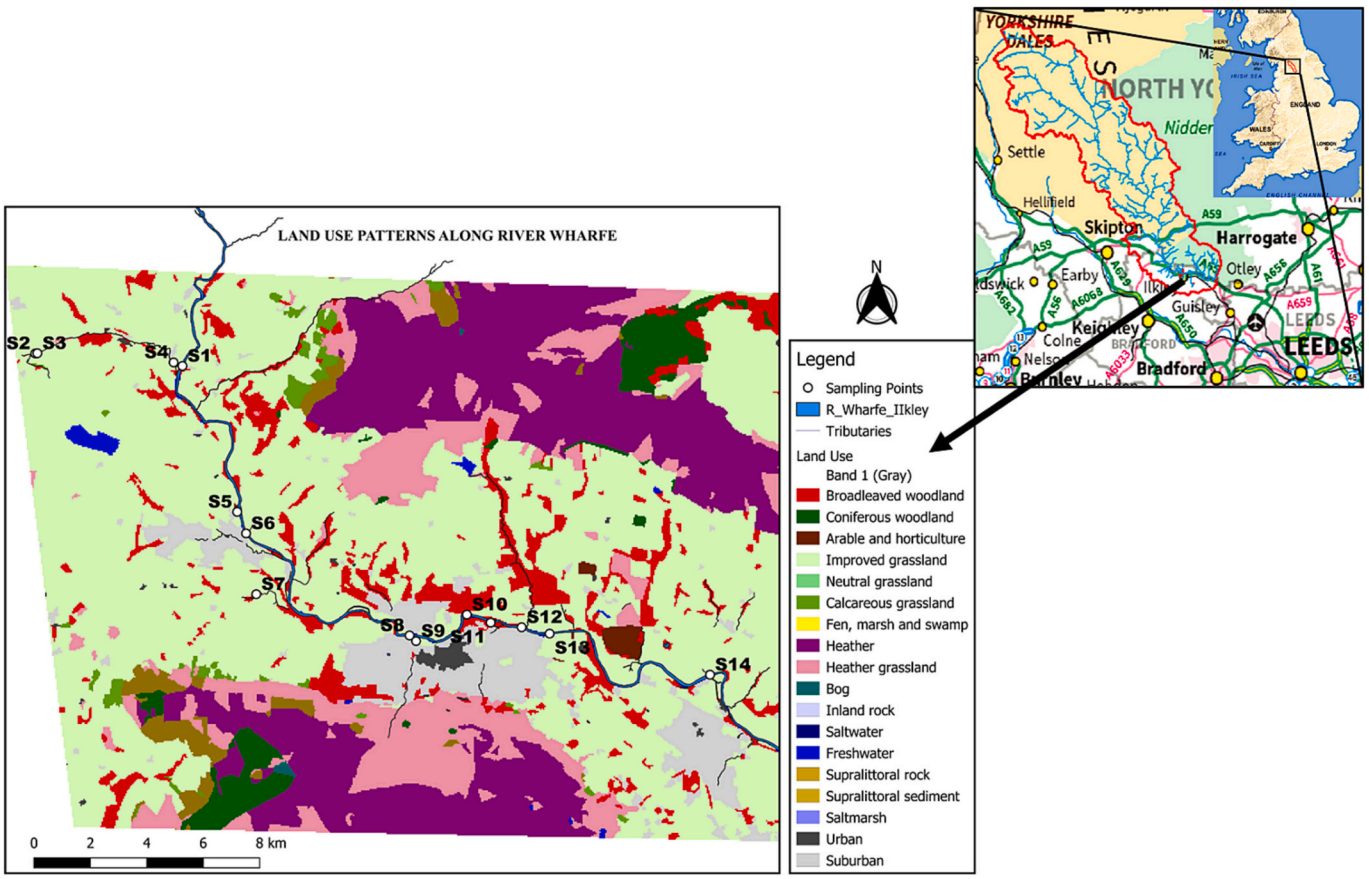


Fig. 1. Map showing sampling sites and land uses along the stretch of the River Wharfe sampled in the 2021 bathing season. The bathing site is S10.

the river for *E. coli* during the bathing season of 2021.

Samples were collected from sites S1, S4, S6, S7, S8, S10, S11, S13 and S14 in June, July and August of 2021. However, the August sampling included two extra sampling sites (S5 and S9) since high levels of pollution were monitored by measuring *E. coli* in previous months, and we were interested in determining the bacterial structure and composition at these sites. In September, due to limited resources, only 5 sites were sampled (S2, S3, S7 S10 and S12), targeting only places associated with high faecal pollution. To summarise, a total of 35 water samples were collected from mid-stream flow in the river or tributary, without disturbing the sediment, from 14 different sampling sites during the bathing season of 2021. For each sampling event, samples were collected by filling 5 L Nalgene plastic bottles previously disinfected with Virkon (SLS, UK) and rinsed afterwards to sterilise deionised water.

The samples were then transported on ice, to the University of Sheffield, and filtered for subsequent molecular analysis within 24 h of collection. From each bulk 5 L water sample, three replicates were obtained by filtering aliquots of 500–1000 mL through 0.2 μm sterile nitrocellulose membrane filters (Sartorius, UK). The total volume filtered depended on the concentration of suspended solids present in the samples, and the subsequent level of saturation of the filters. Please note that after heavy rainfall events, the maximum amount of water filtered was 500 mL due to the saturation of the filters. A total of 105 filters were obtained and preserved in sterile bags at −80 °C for subsequent DNA extraction. Simultaneously, water samples were collected in sterile containers for *E. coli* analysis by members of the Ilkley Clean River Group, Addingham Environment Group and Yorkshire Dales Rivers Trust (YDRT), and sent to an UKAS accredited laboratory for analysis (ISO/IEC 17025), ALS Environmental Coventry (UK).

2.2. Water quality monitoring sondes and culture-based FIB

In situ water quality multiparametric monitoring sondes were installed by Yorkshire Water Ltd. (YW Ltd.) in sites S4, S5, S6, S7, S9, S10, S11 during the bathing season of 2021. The multiparametric water quality sondes Xylem EX02 (Xylem Analytics, UK) were measuring continously and in real-time the levels of dissolved oxygen (DO), temperature, pH, conductivity, and ammonium. Simultaneously, YW Ltd. collected spot samples through the bathing season, at the sites where the sondes were installed, to measure *E. coli* and intestinal enterococci using current standard methods (ISO9308-1 and ISO7899-2, respectively). These gave two sets of independent measurements of FIB: 1) samples collected by citizen scientists at exactly the same time and site that samples were collected for DNA analysis and 2) water samples collected by YW Ltd. at those sites where the multiparametric sondes were installed. Principal Component Analysis (PCA) was used to reveal the key physico-chemical parameters determining the FIB levels measured by YW Ltd. at the sites where the sondes were installed. The PCA was performed with the Statistics and Machine Learning Toolbox of MATLAB 9.13 (https://www.mathworks.com).

2.3. Water filtration and DNA extractions

In total, DNA from 105 filters was extracted using a CTAB-chloroform-based method as previously described (Karunakaran et al., 2016). The quantity and purity of the extracted DNA (260/280 ratio) were measured using Nanodrop ND-1000 spectrophotometer (NanoDrop, Wilmington, DE). Based on the 260/280 ratio, those samples showing higher levels of inhibitors (260/280 $\leq$ 1.6) due to for example high levels of organics, were purified using a Wizard® DNA Clean-Up System (Promega UK).

2.4. qPCR with host specific markers: Bacteroidales general faecal indicators, human and ruminant

For the qPCR analysis, 75 samples (including replicates) were analysed to obtain representative samples of sites affected by faecal pollution from unknown origin, yet central to managing the designated bathing site for the water utility. These 75 DNA samples were from sites S1, S4, S5, S6, S7, S8, S10 and S13 from June, July, August and S7, S10 and S12 in September. The samples were pooled according to site and date and diluted 1/10 in molecular biology grade water (Thermofisher, UK). All the samples were run in triplicate, including a positive control and a negative control (PCR reaction without DNA) to verify the lack of contamination in the reagents. The dilutions and filtered sample volumes were considered when calculating the copy numbers per litre. The qPCR reactions were performed using an Applied Biosystems 7500 Real-Time PCR System (Thermofisher Scientific, UK). All qPCR reactions (20 μL total volume) were performed in triplicate and the reaction mixture for all Taqman chemistry-based qPCR assays included 10 μL of TaqMan Environmental Master Mix 2.0 (Applied Biosystems, Foster City, CA, USA), 2 μL of the probe/primer set with a final concentration of 2 μM probe and 10 μM each primer, and 5 μL of the 10-fold diluted target DNA template.

The primers and Taqman probes and the conditions used for the qPCR reactions are provided in Supplementary Material Table S1. Quantification of samples involved the use of standard curves prepared by using serial dilutions of each target gene (Osborn and Smith, 2005) amplified from the following sources: 1) for general *Bacteroidales*, DNA from *Bacteroides dorei*, *Bacteroides vulgaris* and *Bacteroides uniformis* was amplified and pooled (DMSZ, Germany); 2) for human derived *Bacteroidales*, DNA was extracted and amplified from samples obtained from an effluent from a local YW Ltd. wastewater treatment plant, and 3) for ruminant specific *Bacteroidales*, DNA was extracted, amplified and pooled from faeces from sheep, cow, pig, goat and horse obtained from Graves Park Animal Farm (Sheffield, UK). All the amplicons were purified using a QIAquick PCR Purification Kit (Quiagen, UK) and then quantified using Qubit Fluorometric quantification (Thermofisher, UK).

2.5. 16S rRNA gene sequencing and analysis

To obtain representative samples to analyse bacterial diversity through sequencing of the 16S rRNA gene, 3 sample replicates from the same site and date were pooled using equal amounts of DNA, to obtain a representative sample. In total 35 pooled samples were obtained, quantified and sent for sequencing to Mr. DNA (Shallowater, TX, USA). Sequencing was performed on a MiSeq platform following the manufacturer's guidelines. The 16S rRNA gene was sequenced by using primers 28F and 519R spanning the V1 to V3 hypervariable regions. These primers were used in a PCR reaction using the HotStarTaq Plus Master Mix Kit (Qiagen, USA) under the following conditions: 94 °C for 3 min, followed by 30–35 cycles of 94 °C for 30 s, 53 °C for 40 s and 72 °C for 1 min, after which a final elongation step at 72 °C for 5 min was performed. The purified amplicons were used to prepare the Illumina DNA library for sequencing.

In summary, an initial quality control of the sequencing raw data was carried out to remove sequencing errors, and to filter sequences with <150 pb. Then, pair-end sequences were joined and dereplicated, and chimeras were removed from the analysis using UCHIME2 (Edgar, 2018). Final Operational Taxonomic Units (OTUs) were taxonomically classified using BLASTn against a curated database derived from Ribosome Database Project II (RDPII) and The National Center for Biotechnology Information (NCBI) (Cole et al., 2013). Alpha-diversity was calculated as a measurement of Chao 1 index (richness estimator), Simpson index (dominance), and Shannon index (diversity) using unrarefied OTU tables at 97 % sequences similarities cut-off (Calero Preciado et al., 2022). The data were transformed by square root calculations, and Bray–Curtis similarity matrices were generated using the software Primer v7 (PRIMER-E, Plymouth, UK) and visualised using a non-metric multidimensional scaling (nMDS) diagram. Analysis of similarity statistics (ANOSIM) was calculated using the same Bray–Curtis distance matrix to test the significance of differences

between samples based on month of sampling.

3. Results

3.1. Water quality monitoring (multiparametric sondes and FIB)

In situ water quality measurements (Supplementary Material, Table S2) from the water quality monitoring sondes installed by YW Ltd. showed that temperature during the bathing season ranged from a minimum of 12.5 °C in S9 in August to a maximum of 19.5 °C in S7 and S11 in July. The lowest DO concentration 8.77 ± 0.81 mg/L was recorded at S10 in July. The recorded pH was stable across the monitoring sites and ranged between 7.9 and 8.9 for the whole bathing season. The concentration of ammonium ranged between a minimum of 0.06 mg/L in S10 in August and September and a maximum of 0.28 mg/L in S9 in September, where high levels of FIB were also detected by YW Ltd. The lowest values of conductivity were observed in sites S9 and S10 (always <200 μS/cm) for the duration of the bathing season. The highest conductivity values were obtained in S4 ≥ 455 μS/cm in July and August. The highest counts of *E. coli* provided in the YW Ltd. analysis of water samples in the places where the multiparametric sondes were installed (average 8425 CFU/100 mL) and IE (average 9475 CFU/100 mL) were found in S7.

Fig. 2 shows the results of the Principal Component Analysis (PCA) carried out with the water physico-chemical parameters measured by the *in-situ* sondes and the *E. coli* and IE values provided by YW Ltd. The principal component Axis 1 accounted for 73.5 % of the variability and principal component Axis 2 for 16.3 %. The results from the PCA showed that ammonium was the main parameter influencing the distribution of FIB levels in the river basin, believed to be contributed from agricultural activities, livestock, and sewage discharges.

3.2. Faecal indicators results in discrete water samples in the sites used for subsequent molecular analysis

Higher concentrations of *E. coli* counts (Supplementary Material, Table S3) were observed in samples collected from site S11, reflecting the influence of a nearby STW. For this site, *E. coli* levels were particularly high for August (55,000 CFU/100 mL). This result is associated with a period of sustained rainfall the day before sampling and may have included the influence of both untreated storm overflow discharges as well as higher volumes of treated effluent. Several sampling locations on tributaries including S5 (15,000 CFU/100 mL) and S4 (9000 CFU/100 mL) showed high levels of *E. coli* in the August samples (after a heavy rainfall event) when compared with other months. Overall, during the bathing season, values from the main river upstream STWs were less than those from the tributaries.

3.3. Quantification of faecal markers by qPCR

Table 2 shows the results of qPCR analysis as number of target genes in the samples (copies/L). For each qPCR assay, the quantification limit was determined from the standard curve obtained from each target gene. The lowest concentration of standard gene copies that could be confidently quantified was considered as the quantification limit. All the qPCR results were normalised to gene copies/L of water filtered and the samples considered positive if the concentrations were above the quantification limit. General *Bacteroidales* (GenBac) were quantified in all the samples analysed, yet a higher copy number of this gene was consistently found in the main river samples from S13 ($\geq 2.9 \bullet 10^7$ gene copies/L) and in S12 in September ($1.3 \cdot 10^8$ gene copies/L), downstream of a STW. In August, most of the samples showed an increase in copy numbers, coincident with a heavy rainfall event occurring the day before the sampling. The qPCR for the human specific *Bacteroidales* gene (HF183), suggests higher levels of this gene in samples from sites S4 in July and August ($1.9 \cdot 10^5$ to $2.4 \cdot 10^5$ gene copies/L), S13 in June, July and August (2.6 to $4.3 \cdot 10^5$ gene copies/L) and in S12 in September ($1.1 \cdot 10^6$ gene copies/L). The amplification of ruminant specific *Bacteroidales* (RumBacB2), showed that this gene was present but Below the Quantification Limit (B.Q.L) for all the samples in June. In July, only in samples from the main river S1 and S13 was it possible to quantify the presence of this gene (gene copies $>4 \cdot 10^5$copies/L). However, in August, ruminant specific *Bacteroidales* were quantified in several samples, including samples from the main river (e.g., S1, S6, S10 and S13), indicating the potential influence of agricultural run off after a heavy rainfall event. In the samples analysed in September, ruminant specific

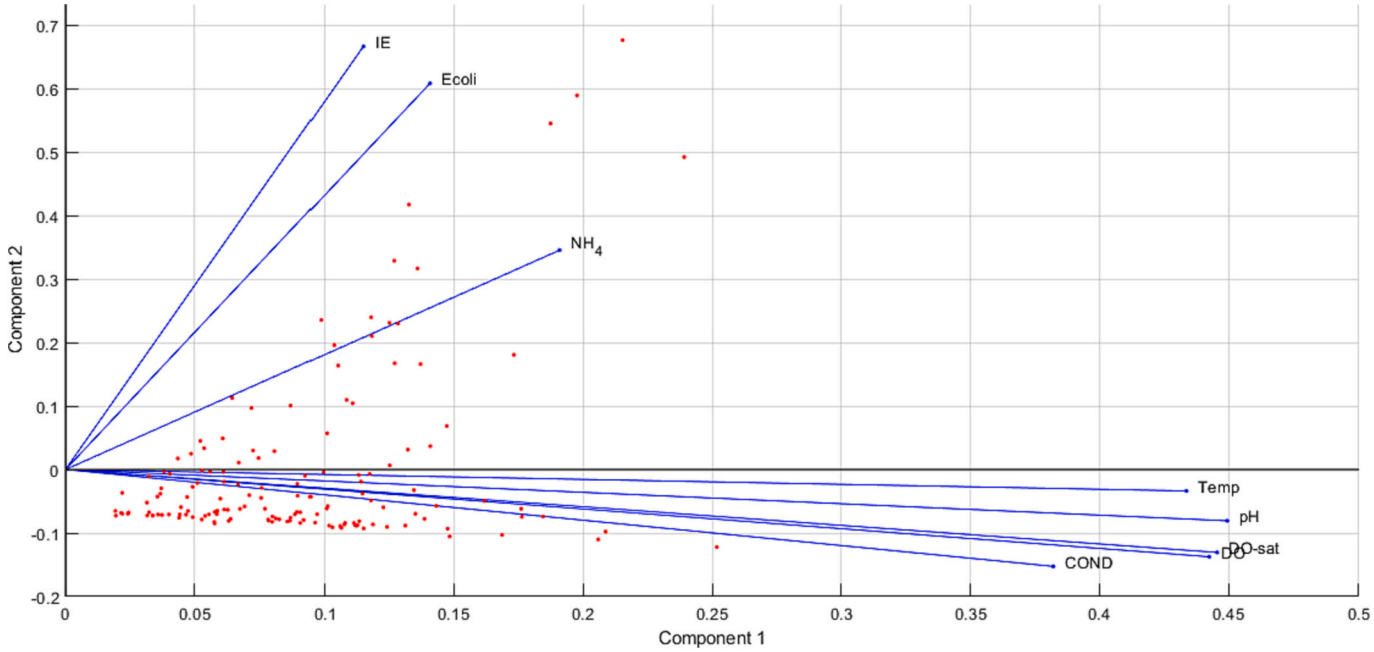


Fig. 2. PCA of physico-chemical water quality measurements from the multiparametric sondes and faecal indicators (*E. coli* and IE) measured by YW Ltd. at those sampling sites where the sondes were installed.

Table 2
Results of qPCR analysis using 3 different genetic markers: general *Bacteroidales* (GenBac). Human specific *Bacteroidales* (HF183), and ruminant specific *Bacteroidales* (RumBac2). Results are expressed as gene copy number/L.

	Site	GenBac	HF183	RumBacB2
June	S1	8.7E+06 ± 1.2E+06	5.1E+04 ± 1.6E+04	B.Q.L
	S4	1.3E+07 ± 1.7E+06	7.4E+04 ± 1.4E+03	B.Q.L
	S6	6.8E+06 ± 1.8E+06	3.4E+04 ± 4.3E+03	B.Q.L
	S7	9.8E+06 ± 9.5E+05	4.6E+03 ± 1.1E+03	B.Q.L
	S8	7.1E+06 ± 2.7E+06	3.1E+04 ± 2.8E+03	B.Q.L
	S10	8.6E+06 ± 7.9E+05	5.2E+04 ± 1.0E+04	B.Q.L
	S13	2.9E+07 ± 4.4E+05	4.1E+05 ± 1.4E+04	B.Q.L
July	S1	1.3E+07 ± 2.3E+06	3.5E+03 ± 1.3E+03	4.2E+05 ± 8.7E+02
	S4	4.8E+07 ± 3.2E+06	2.4E+05 ± 5.5E+03	B.Q.L
	S6	4.2E+06 ± 9.6E+05	B.Q.L	B.Q.L
	S7	9.1E+06 ± 2.1E+05	2.1E+04 ± 1.8E+03	B.Q.L
	S8	7.1E+06 ± 7.1E+05	3.9E+03 ± 1.4E+03	B.Q.L
	S10	8.6E+06 ± 7.9E+05	4.0E+04 ± 2.1E+04	B.Q.L
	S13	5.4E+07 ± 7.1E+06	2.6E+05 ± 1.4E+04	4.7E+05 ± 2.2E+04
August	S1	1.6E+07 ± 9.9E+05	9.9E+03 ± 5.8E+03	9.5E05 ± 0
	S4	9.1E+07 ± 1.4E+07	1.9E+05 ± 1.1E+04	B.Q.L
	S6	1.7E+07 ± 1.1E+05	3.0E+04 ± 6.4E+03	9.6E+05 ± 0
	S7	4.0E+07 ± 1.4E+06	2.4E+04 ± 4.1E+02	B.Q.L
	S8	2.1E+07 ± 1.5E+06	4.3E+04 ± 7.0E+03	B.Q.L
	S10	2.9E+07 ± 1.6E+06	1.4E+05 ± 7.0E+03	9.0E+05 ± 9.9E+04
	S13	5.4E+07 ± 2.6E+06	4.3E+05 ± 3.4E+04	8.8E+05 ± 4.6E+04
September	S7	5.0E+06 ± 1.5E+05	2.5E+04 ± 4.0E+03	B.Q.L
	S10	1.4E+07 ± 5.5E+05	2.3E+04 ± 1.4E+03	4.1E+04 ± 0
	S12	1.3E+08 ± 2.8E+07	1.1E+06 ± 1.4E+05	B.Q.L

B.Q.L. = Below Quantification Limit.

Bacteroidales were quantified in the bathing site (S10).

3.4. Metabarcoding of the 16S rRNA gene sequencing results

3.4.1. Relative abundance of bacterial groups at phylum level

Fig. 3 shows that the most abundant phylum in all the samples was *Proteobacteria* (>28 % in all samples) and *Bacteroidetes* (>20 % in all samples). *Bacteroidetes* were predominant in samples from the main river (e.g. S10, S13, S14). *Actinobacteria* was also present in all samples (6–20 %) over the monitored period but increased in abundance in all the samples after July. *Verrucomicrobia* (0.5–11 %) clearly increased in July, reaching 20 % of the total relative abundance in samples from S14. Regarding phyla commonly associated with animals' intestinal tracts, the main phyla detected through DNA sequencing were: i) *Bacteroidetes*, ii) *Firmicutes* and iii) *Fibrobacteres*. *Firmicutes* was consistently present at low percentages in all samples, but it was higher (>3 %) in two tributaries, S4 and S7, and increased its abundance in the main river in S11 and S13 in August samples (>2 %). The relative abundance of the phylum *Fibrobacteres*, was low in all samples (<5 %), but was higher in several tributaries, particularly in S4, S5, S7 in August (0.2–3 %) and in S7 in September (1 %), when compared with samples from the main river. Several tributaries showed higher levels of *Cyanobacteria* in August (S4, S5 and S7) and in S7 in September (7 %) when compared with other sampling months. These bacteria are traditional indicators of eutrophication in freshwater ecosystems.

3.4.2. Diversity of faecal indicators (human vs. non-human) in the bacterial communities

Several indicators of faecal pollution were found in the samples (Fig. 4A). Overall, the relative abundance of these indicator bacteria in the whole bacterial community in the samples was low, representing <2 % of the total bacterial diversity in a given sample. Traditional faecal indicators including *E. coli* and *Enterococcus* spp. were not detected in the samples by sequencing the 16S rRNA gene. This can be explained by the low level of representation that these bacteria have in the total microbial community when using DNA-based analysis. *E. coli* belongs to the *Enterobacteriales* order, and this order represented <0.5 % of the whole bacterial community in all the samples analysed. General indicators of faecal pollution are *Bacteroides* spp. Of this group, several species are human specific including: *Bacteroides dorei*, *Bacteroides uniformis* and *Bacteroides vulgatus*. These three species combined were present mainly in July in S4 (0.13 %) and S7 (0.09 %) (both tributaries) and in August in S9 (0.73 %) and in S11 (0.44 %) (Fig. 4A). Several species also associated with human-derived faecal pollution such as *Faecalibacterium* sp. and *Faecalibacterium prausnitzii* (Fitzgerald et al., 2018: Martín et al., 2017), were present in samples from tributaries including S4, S5 and S7. The abundance of *Bacteroides graminisolvens* (<0.2 %) a host-specific bacterium of ruminants (Nishiyama et al., 2009), was higher in tributary samples such as S7 in June (0.18 %) and S4 and S7 (0.08–0.17 %) in August. Several samples in the main river, S13 and S14 in August and S10 in September, showed very low counts of *B. graminisolvens* (<0.05 %) (Fig. 4A).

3.4.3. Diversity of pathogens and opportunistic pathogens (OPs) in the bacterial communities

Several species of potential pathogens and OPs were found in the samples (Fig. 4B), representing in several samples (S4, S7 and S11 in June, July and August, and in S2, S3, and S12 in September) a high percentage of the relative abundance of the whole bacterial community. The percentage of these type of bacteria tended to be higher in tributary samples including S4 in June (>27 %) and S7 in July (>11 %), and in sample S3 in September (>12 %), likely associated to its location downstream of a small STW. For all these samples, the main potential pathogenic bacterial genera were *Pseudomonas* spp. and *Mycobacterium* spp. Overall, *Pseudomonas* spp. were particularly abundant in the S4 sample in June (>25 %), in the S4 and S7 samples in July (>7 %), S7 and S11 in August (>8 %) and S3 in September (>10 %).

Pseudomonas spp. are ubiquitous in the environment and only a few species, including *Pseudomonas putida*, are pathogenic. *P. putida* occurred mainly in sample S4 in June (0.63 %) and S1 in August (0.54 %), and the relative abundance was very low for all other samples. Several species of *Aeromonas* spp. were also detected, representing, in S11 in August, 1.82 % of the total bacterial abundance and present in the same sampling site in other months but at a very low percentage <0.5 %. The presence of other potential pathogenic species such as *Clostridium* spp. was low, with a limited presence in samples, including one of the tributaries (S7, 0.6 %).

3.4.4. Alpha diversity indices: richness, diversity and dominance of bacterial species in the samples

Over the sampling period, richness (Chao >345 OTUs) was higher in samples from the tributaries S4 and S7 (Fig. 5A). Overall, samples collected from the main river showed less richness, with the exception of those downstream of a WWTP (S11 to S14). In August, a clear increase in

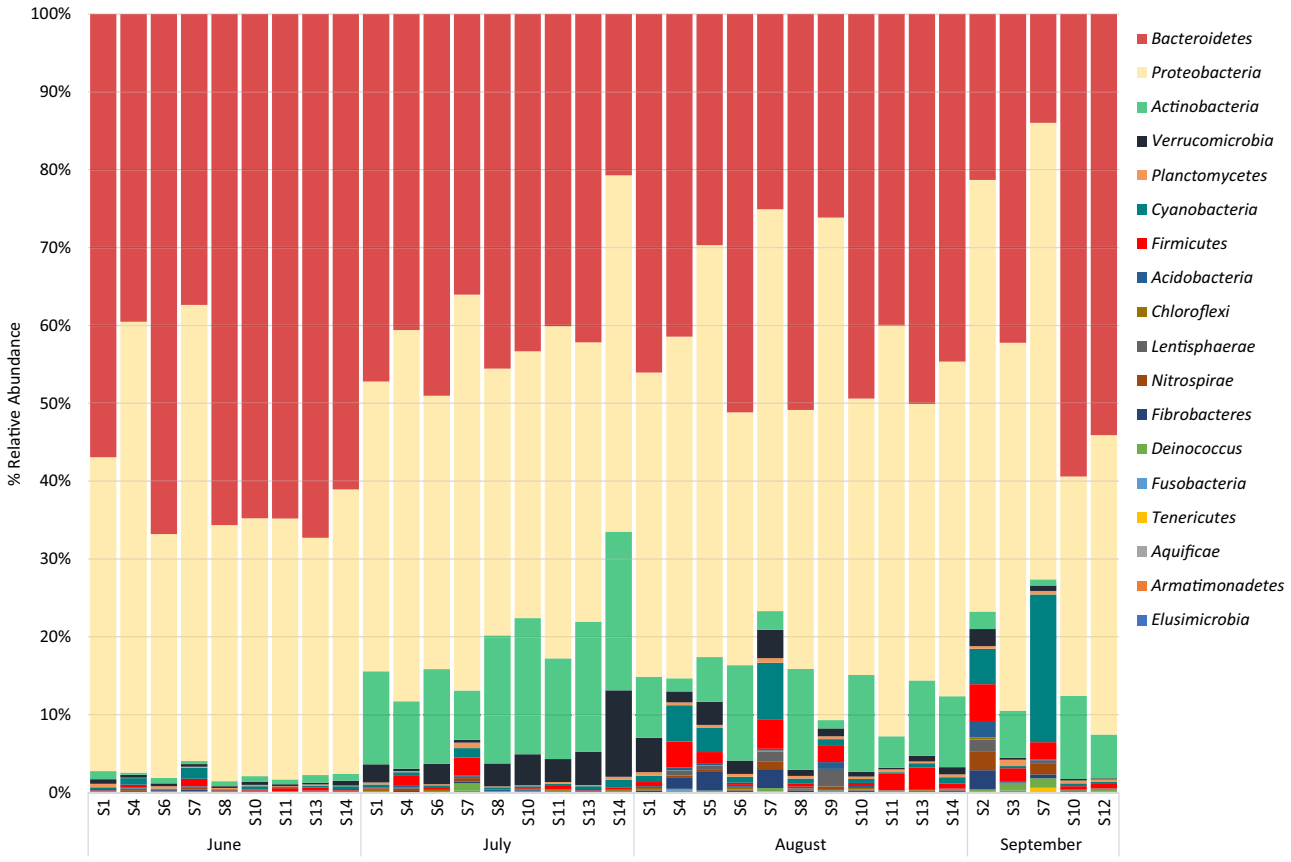


Fig. 3. Comparison of the relative abundance of the different phylum found in water samples over the sampled months.

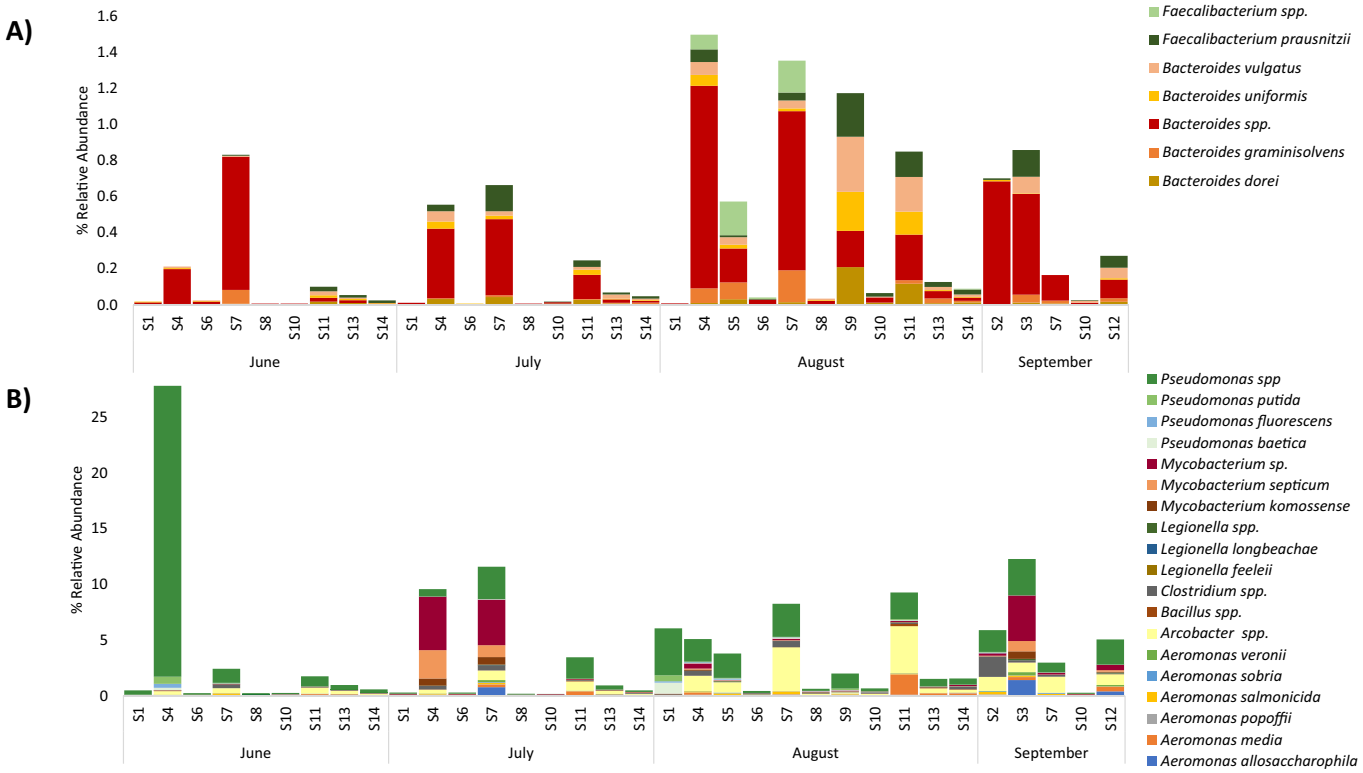


Fig. 4. A) Relative abundance of bacteria associated with faecal pollution; B) relative abundance of pathogens and opportunistic pathogens.

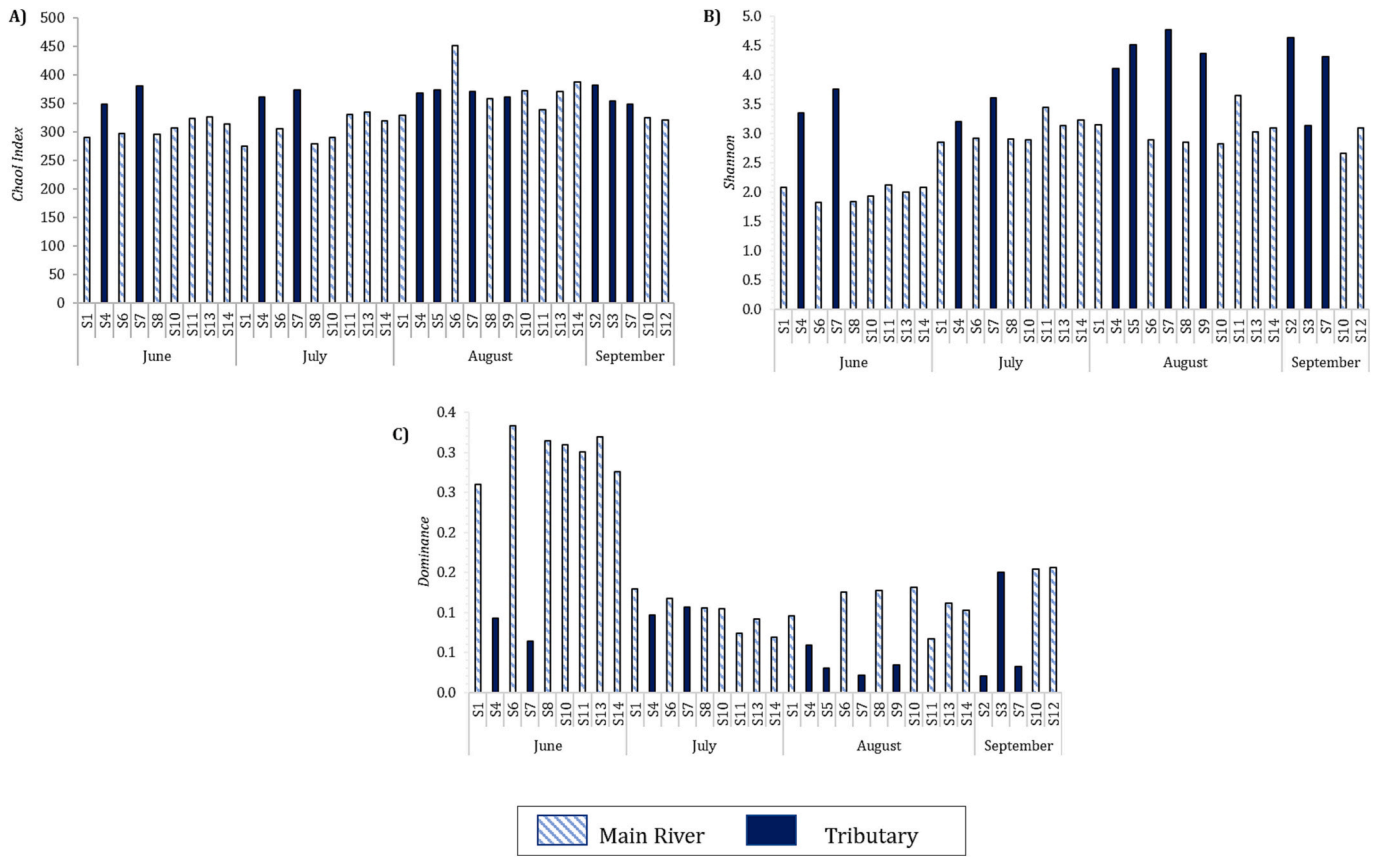


Fig. 5. Indices of A) richness using Chao I, B) Shannon diversity and C) dominance indices based on relative abundance of OTUs.

richness was observed in most of the samples, particularly in S6 (>400 OTUs). Shannon's diversity index (H′) (Fig. 5B), includes a measure of the number of OTUs in a community, taking into account the abundance of each OTUs (Willis, 2019). The samples with the lower diversity were those from S6 and S8 in June (H′ = 1.8 to 2.2). Higher Shannon (>4) values were found in August in several samples from tributaries S5 and S7, and in S7 in September. Dominance index results (Fig. 5C), indicate the dominance of one or few OTUs in the entire bacterial community. Generally, dominance was higher in the samples in June except for S4 and S7 and clearly decreased in August samples, particularly in S5 and

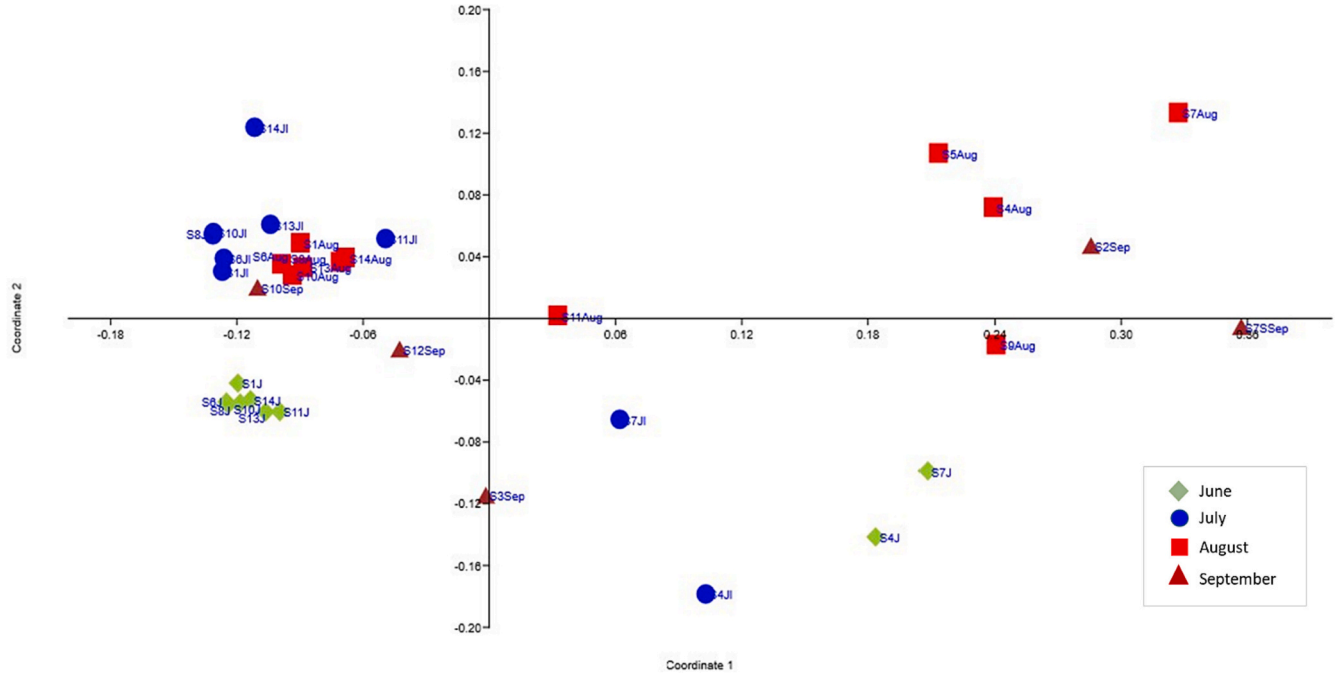


Fig. 6. Two-dimensional non-metric MDS ordination plot of samples analysed by 16S rRNA sequencing during the bathing season of 2021.

S7.

3.4.5. Non-metric Multidimensional Scaling (nMDS) analysis of bacterial communities at species level

The nMDS analysis (Fig. 6) was used to obtain clustering of samples at 97 % sequence similarity cut-off. Variability in bacterial community distribution between sample points was observed showing a temporal pattern with samples tending to cluster according to date of sampling. For example, most of the samples in June clustered together except for S7 and S4, both samples from tributaries. The nMDS analysis also showed a tendency for samples to separate spatially, with samples obtained from tributaries, with higher levels of pollution, tending to group together and separate from the main river samples. Samples from several tributaries including S4, S5, S7 and S9 in August and September tended to separate from the main cluster of samples of the months analysed. The separation of samples according to month was supported by the statistical ANOSIM analysis for June and July months ($R = 0.35$ and $p = 0.002$) and June and August ($R = 0.32$, $p = 0.005$).

4. Discussion

4.1. Water quality monitoring and faecal pollution indicators

The analysis of continous and real-time data from YW Ltd. water quality multiparametric sondes showed that the primary factor determining the distribution of faecal indicators (*E. coli* and IE) in the samples was ammonium. Ammonium is an important parameter monitored as part of the Water Framework Directive and has been previously recognised as a sentinel parameter of the microbiological pollution load in rivers (Cabral and Marques, 2006). Several studies in catchments around the world (Vadeboncoeur et al., 2018; Reynolds et al., 2021) have shown that ammonium sources are mainly manure, agricultural runoff, sewage and municipal effluent discharges, showing its relevance as a water quality parameter suitable for assessment of faecal pollution in rivers. In this study, ammonium was higher in S4 (0.18–0.20 mg/L) and S9 (0.26–0.28 mg/L), coincident with higher relative abundance of faecal pollution markers in the samples (*Bacteroides* spp. and *Faecalibacterium* spp.) when compared with other samples (Fig. 4). The sampling point S4, is located downstream of a STW and the site S9 has been associated with sewer misconnections. At the bathing site (S10), levels of ammonium were always low and ranged between 0.06 and 0.07 mg/L, confirming that anthropic sources are not the major cause of pollution at this site. Faecal pollution in the studied area was mainly associated with discharges from a small STW (S3 and S4 samples), potential leaks from septic tanks and surface water runoff (S5 and S7), and undetected sewer misconnections (S9). However, in agricultural/farming areas (Fig. 1 Table 1), faecal bacteria from ruminant sources were found including *Bacteroides graminisolvens* detected by metabarcoding and RumBacB2 gene copy numbers quantified by qPCR. At these sites, the detection of RumBacB2 genes and sequences belonging to *B. graminisolvens* was low, and when the levels of detection were met in the qPCR analysis, the ruminant marker levels tended to be higher when compared with the human markers (HF183). *Bacteroides graminisolvens*, has been previously found in methanogenic reactors of cattle farms (Abe et al., 2012; Nishiyama et al., 2009), showing potential as a MST marker from farming effluents. During dry weather conditions (e.g., June/July), faecal bacteria from ruminant sources were below the limit of detection when analysed by qPCR. However, in August, in samples collected shortly after a heavy rainfall event, the abundance of bacteria from ruminant sources increased (Table 2). Similar findings regarding the impact of intense rainfall on faecal pollution were observed by Peed et al. (2011). The authors reported a positive significant correlation between abundance of human associated markers and septic systems following a wet weather event. Overall, the lower concentrations of faecal pollution found in the main river upstream of the primary urban area, when compared to the tributaries, can be attributed to the dilution effect caused by the significantly larger volume of water in the main river channel. However, the main river samples downstream of a STW (from S11 to S14) showed high levels of human-faecal indicators (quantified mainly as GenBac and HF183 copy numbers). Samples in August, collected after a wet weather event, showed an increase in faecal markers (both human and ruminant). Heavy rainfall surface run-off processes, and CSOs losses that are released directly into a water body can result in water quality changes in receiving waters, showing the importance of incorporating wet weather monitoring in surveillance campaigns (Munro et al., 2019; Tian et al., 2022). Further research is required in agricultural catchments at different times of the year and under different flow conditions to better apportion sources between humans and ruminants.

4.2. Microbial community composition, structure and indicators of faecal pollution

The utilisation of DNA sequencing to analyse bacterial community composition provides a valuable tool for monitoring influences of pollution and environmental changes. This approach enables the identification of specific microorganisms that can serve as bioindicators for assessing water quality and identifying pollution events (Yergeau et al., 2012). In the present study, the bacterial community structure exhibited spatial and temporal variability across the sampling sites, reflecting influences from urban and agricultural activities within the catchment area (Fig. 6). The impact of faecal pollution on bacterial community composition was clearly demonstrated through DNA sequencing results, which revealed distinct bacterial fingerprints in sites most affected by faecal pollution (Fig. 3). Overall, bacterial communities associated with polluted sites exhibited higher abundances of *Bacteroides* spp., *Firmicutes*, pathogenic organisms, and opportunistic pathogens (Fig. 4), along with reduced community diversity (Fig. 5). Previous research has indicated that STWs (influent and untreated waters) predominantly contain microbiomes represented by *Firmicutes* (McLellan et al., 2010; Shanks et al., 2013) and *Bacteroidetes* (Newton et al., 2015). *Bacteroides* spp. are known to be closely associated with faecal matter, as they are strict anaerobic bacteria with limited survival rates outside of a host in the environment (Eklas, 2021). In this study, *Bacteroides uniformis* was found to be more prevalent in samples affected by wastewater treatment plants downstream from the bathing site (e.g., S11 to S14). *B. uniformis* is a resident of the human intestines and has been proposed as a potential candidate indicator for faecal pollution (Morita et al., 2021). Overall, *Bacteroides* spp. have been identified as suitable indicator organisms for CSO contamination in urban rivers (Ekhlas et al., 2021). Additionally, the combined quantification of *Bacteroides* spp. and *E. coli*, as suggested by Mulugeta et al. (2012), can enhance water quality assessment in aquatic environments. Another bacterial group frequently present in polluted samples in this study was *Firmicutes*. Bacteria belonging to the phylum *Firmicutes* are important inhabitants of the human gastrointestinal microbiome (Siezen and Kleerebezem, 2011), and are among the major bacterial groups found in river catchments influenced by human faecal sources (Raza et al., 2021). Some bacteria belonging to *Firmicutes* such as *Bacillus* and *Enterococcus* can produce resistant endospores that can persist in the environment (Browne et al., 2021). These microorganisms have been found in waters that have been contaminated over extended periods, making them excellent bioindicators of long-term faecal pollution (Edberg et al., 2000; Cabral, 2010). *Faecalibacterium prausnitzii*, a member of the *Firmicutes* phylum identified in this study, is one of the most abundant bacteria in the healthy human microbiota (Martín et al., 2017). Similar findings were recently reported by Sun et al. (2017), who observed elevated levels of *Firmicutes* in the Yangtze River, attributing them to faecal bacteria discharged from a nearby wastewater treatment plant. The findings from this study suggest that bacteria belonging to *Bacteroides* and *Firmicutes* can serve as reliable indicators of sewage pollution in surface waters. These bacteria, typically found in human faeces or sewage, but absent or present in minimal

amounts in surface freshwaters, hold potential as indicators for assessing sewage contamination. Furthermore, with advancements in high-throughput sequencing technology such as the portable sequencer called MinION (Werner et al., 2022) surveillance of pollution using sequencing may one day be routinely employed to monitor the overall health- status of rivers.

4.3. Risk from pathogens and opportunistic pathogens (OPs)

The relative abundance of pathogen-like sequences at the genera and species level varied based on land use and the abundance of FIB. The presence of human-derived bacteria coincided with higher bacterial diversity and the occurrence of potential pathogens, particularly bacteria belonging to the genera *Mycobacterium*, *Aeromonas*, *Arcobacter*, and *Clostridium*. Among *Mycobacterium* species, NonTuberculous *Mycobacteria* (NTM) were present in tributary samples, particularly in sites S1 and S2 in July. NTM is a subgroup within the *Mycobacterium* genus and includes numerous pathogenic species that are commonly found in aquatic environments (Delghandi et al., 2020). Similarly, *Aeromonas* which are considered opportunistic pathogens, are frequently isolated from freshwater ecosystems with varying degrees of faecal pollution and they can impact immunocompromised individuals who have had contact with contaminated water (Pettibone, 1998). *Arcobacter*, is an emerging waterborne pathogen and its presence has been linked to faecal pollution in rivers in Spain (Collado et al., 2008), Brazil (Godoy et al., 2020), and Japan (Ekhlas et al., 2021). Species of the genus *Clostridium* are widespread in nature and commonly inhabit the gastrointestinal tracts of humans and animals (Guo et al., 2020). Higher levels of *Pseudomonas* were also present in polluted sites in this study, most of which are not pathogenic and are commonly found in freshwater ecosystems. However, some *Pseudomonas* strains, such as *Pseudomonas putida*, can cause infections (Fernández et al., 2015).

The results obtained from molecular methods indicate that human sources of faecal pollution significantly contribute to overall pollution within the study area. In samples taken within 48 h following a heavy rainfall event in August, the concentration levels of faecal markers and the presence of potential OPs, particularly *Mycobacterium* spp., increased in all samples. This suggests an input of new faecal pollution from different sources due to surface-water runoff. Similar findings have been observed in other studies where rainfall has been identified as the main driver of water quality variations and an elevated microbial risk (Tornevi et al., 2014; García-Aljaro et al., 2017). These human sources continuously introduce microbial inputs to the river, exerting a selective pressure on the natural microbial communities. This pressure can lead to the enrichment of certain microbial groups, as observed in this study, with an increase in the dominance of potential pathogens like *Mycobacterium* and *Aeromonas* (Fig. 4). DNA sequencing profiling provided a powerful tool for understanding the dynamics of potential pathogens present in the samples. The understanding of the distribution patterns of these adverse microbes is crucial for catchment managers in making informed decisions regarding the implementation of appropriate practices, to mitigate and reduce potential risks. This might involve implementing better water treatment technologies, improving sanitation practices, or implementing targeted interventions in specific areas with higher pathogen loads. The results presented in this study contributed to research carried out to improve surveillance and mitigation efforts carried out by the local water company. However, from a regulatory perspective, epidemiological studies are necessary to develop risk assessment models and evaluate the use of molecular methods in setting regulatory thresholds and determining the association between the detection of certain indicators and illness. By examining the relationships between various factors, including exposure to faecal pollution, environmental conditions (e.g., heavy storms, heatwaves, floods, and droughts), demographic data, and health outcomes, these models can quantify the likelihood of illness associated with specific exposure levels (Holcomb and Stewart, 2020).

5. Conclusion

Assessing contamination sources in recreational waters is of utmost importance for directing outreach and mitigation efforts. This study highlights the effective application of microbial genetic analysis, in identifying fingerprints and alternative indicators of faecal pollution. These methods offer a fast and efficient way to quantify new water quality indicators and assess water quality. However, further research is needed to determine if these indicator microorganisms vary across different locations, and if they can accurately assess human health risks, especially when the primary source of faecal contamination is not human-related. Understanding the potential risks associated with faecal pollution is crucial, and it should be noted that the presence of pathogenic taxa based on biomarker DNA alone may not always correlate with actual public health risks, as for example the presence of virulence genes should also be considered. However, shotgun metagenomics can help address this concern by providing a more comprehensive understanding of the genetic content of an environmental sample and relevant genes associated with health risks.

This study highlights the significance of quantifying pollution sources, both point and nonpoint sources, to effectively manage river water quality and establish appropriate environmental governance policies. A comprehensive understanding of the current state and long-term trends in water quality is pivotal in enabling policymakers to propose adaptive and sustainable management solutions. Continued research, data sharing, and advancements in sequencing technologies will all contribute to the development of robust water quality monitoring practices and informed decision-making processes in the field of water resource management.

CRediT authorship contribution statement

E. Karunakaran: Methodology, Analysis, Writing–review and editing. **R. Battarbee**: Conceptualization, Methodology, Funding acquisition, Writing–original draft. **S. Tait:** Writing– original draft, Funding acquisition. **B.M. Brentan**: Methodology, Analysis, **C. Berney**, methodology, writing-review, and editing. **J. Grinham**: Methodology. **M.A. Herrero:** Methodology. **R. Omolo:** Methodology, Analysis. **I. Douterelo**: Conceptualization, Methodology, Resources, Investigation, Writing – original draft, writing review and editing, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Isabel Douterelo reports financial support and writing assistance were provided by Yorkshire Water Ltd.

Data availability

Data will be made available on request.

Acknowledgements

This research has been supported by Yorkshire Water Ltd. and by in kind contributions from local citizen science groups including the Ilkley Clean River Group and Addingham Environment Group. The *Escherichia coli* analysis was funded by the Yorkshire Dales River Trust and the Yorkshire Dales National Park Authority. We want to thank Graves Park Animal Farm in Sheffield for providing samples to aid the qPCR analysis.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2023.168565.

References

Abe, K., Ueki, A., Ohtaki, Y., Kaku, N., Watanabe, K., Ueki, K., 2012. Anaerocella delicata gen. nov., sp. nov., a strictly anaerobic bacterium in the phylum Bacteroidetes isolated from a methanogenic reactor of cattle farms. J. Gen. Appl. Microbiol. 58 (6), 405–412. https://doi.org/10.2323/jgam.58.405 (PMID: 23337575).

Battarbee, R.W., Secrett, M., 2021a. Improving Water Quality on the River Wharfe from Oughtershaw to the Ouse: A Citizen Science Project. Faecal bacteria data from tributary becks between Bolton Abbey and Ilkley on the 23rd August 2021. https://www.ydrt.org.uk/wp-content/uploads/2023/01/Wharfe_iWHARFE21_Report_Tribs_Final-.pdf.

Battarbee, R.W., Secrett, M., 2021b. iWharfe, Improving Water Quality on the River Wharfe from Oughtershaw to the Ouse: A Citizen Science Project. Faecal Bacteria Data from Samples Collected from Recreational Sites on the 23rd August 2021. https://www.ydrt.org.uk/wp-content/uploads/2022/02/Wharfe_iWHARFE21_Report_21v2.pdf.

Battarbee, R.W., Secrett, M., Malby, R., Shackleton, K., Taylor, M., Simons, C., 2020. iWharfe, Improving Water Quality on the River Wharfe from Oughtershaw to the Ouse: A Citizens Science Project. Faecal Bacteria Data from Samples Collected on the 24th August 2020. https://www.ydrt.org.uk/wpcontent/uploads/2021/03/Wharfe_iWHARFE_MicrobiologyReport_Small.pdf.

Brack, W., Bennett, D., Breitholtz, M., et al., 2022. One planet, one health: a call to support the initiative on a global science–policy body on chemicals and waste. Environ. Sci. Eur. 34 (1), 1–10.

Browne, H.P., Almeida, A., Kumar, N., Vervier, K., Adoum, A.T., Viciani, E., Dawson, N.J. R., Forster, S.C., Cormie, C., Goulding, D., Lawley, T.D., 2021. Host adaptation in gut Firmicutes is associated with sporulation loss and altered transmission cycle. Genome Biol. 22 (1), 204.

Cabral, J.P., 2010. Water microbiology. Bacterial pathogens and water. Int. J. Environ. Res. Public Health 7 (10), 3657–3703.

Cabral, J.P., Marques, C., 2006. Faecal coliform bacteria in Febros River (Northwest Portugal): temporal variation, correlation with water parameters, and species identification. Environ. Monit. Assess. 118 (1), 21–36.

Calero Preciado, C., Soria-Carrasco, V., Boxall, J., Douterelo, I., 2022. Climate change and management of biofilms within drinking water distribution systems. Frontiers in Environmental. Science 10.

Chamberlain, M., Marshall, A.N., Keeler, S., 2019. Open water swimming: medical and water quality considerations. Curr. Sports Med. Rep. 18, 121–128.

Cole, J.R., Wang, Q., Fish, J.A., Chai, B., McGarrell, D.M., Sun, Y., Brown, C.T., Porras-Alfaro, A., Kuske, C.R., Tiedje, J.M., 2013. Ribosomal Database Project: data and tools for high throughput rRNA analysis. Nucleic Acids Res. 42 (D1), D633–D642.

Collado, L., Inza, I., Guarro, J., Figueras, M.J., 2008. Presence of Arcobacter spp. in environmental waters correlates with high levels of fecal pollution. Environ. Microbiol. 10 (6), 1635–1640.

Defra, 2012. Tackling Water Pollution from the Urban Environment. Consultation on a Strategy to Address Diffuse Water Pollution from the Built Environment. November 2012. Retrived on 7th of August 2023 from. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/82602/consult-udwp-doc-20121120.pdf.

Defra, 2020. Ilkley Gets Green Light for First River Bathing Site in England. Press release, Retrived on 7th of August 2023 from. https://www.gov.uk/government/news/ilkley-gets-green-light-for-first-river-bathing-site-in-england.

Delghandi, M.R., Waldner, K., El-Matbouli, M., Menanteau-Ledouble, S., 2020. Identification Mycobacterium spp. in the natural water of two Austrian Rivers. Microorganisms 8 (9).

Douterelo, I., Boxall, J.B., Deines, P., Sekar, R., Fish, K.E., Biggs, C.A., 2014. Methodological approaches for studying the microbial ecology of drinking water distribution systems. Water Res. 65, 134–156.

Edberg, S.C., Rice, E.W., Karlin, R.J., Allen, M.J., 2000. Escherichia coli: the best biological indicator of faecal pollution? J. Appl. Microbiol. 88 (Supplement 1), S66–S72.

Edgar, R.C., 2018. UCHIME2: improved error correction for Illumina amplicon sequencing. Bioinformatics 34 (23), 4114–4115.

Ekhlas, D., Kurisu, F., Kasuga, I., Cernava, T., Berg, G., Liu, M., Furumai, H., 2021. Identification of new eligible indicator organisms for combined sewer overflow via 16S rRNA gene amplicon sequencing in Kanda River. Tokyo. Journal of Environmental Management 284, 112059.

Eklas, S., 2021. Bacteroides spp. as indicators of fecal pollution in aquatic environments: a review. Water Sci. Technol. 83 (8), 1851–1866.

Environment Act, 2021. c. 30, s. 82. Retrived on 1st October 2023 from. https://www.legislation.gov.uk/ukpga/2021/30/section/82/enacted.

Environment Agency, 2020. The State of the Environment: Water Quality in Rivers and Lakes. Retrieved 7th August 2023 from. https://www.gov.uk/government/organisations/environment-agency.

European Environment Agency, 2023. State of Bathing Waters in 2023. Retrieved 7th August 2023 from. https://www.eea.europa.eu/themes/water/europes-seas-and-coasts/assessments/state-of-bathing-water/state-of-bathing-waters-in-2023.

Fernández, M., Porcel, M., de la Torre, J., Molina-Henares, M.A., Daddaoua, A., Llamas, M.A., Roca, A., Carriel, V., Garzón, I., Ramos, J.L., Alaminos, M., Duque, E., 2015. Analysis of the pathogenic potential of nosocomial Pseudomonas putida strains. Front. Microbiol. 6.

Fitzgerald, G.F., et al., 2018. Faecalibacterium prausnitzii: a window into the microbiome and disease. Front. Microbiol. 9, 2114.

García-Aljaro, C., Martín-Díaz, J., Viñas-Balada, E., Calero-Cáceres, W., Lucena, F., Blanch, A.R., 2017. Mobilisation of microbial indicators, microbial source tracking markers and pathogens after rainfall events. Water Res. 112, 248–253. https://doi.org/10.1016/j.watres.2017.02.003, 2017 Apr 1. Epub 2017 Feb 2. PMID: 28171819.

Gerba, C.P., 2015. Chapter 22 - environmentally transmitted pathogens. In: Pepper, I.L., Gerba, C.P., Gentry, T.J. (Eds.), Environmental Microbiology, Third edition. Academic Press, pp. 509–550.

Godoy, R.G., Marcondes, M.A., Pessôa, R., Nascimento, A., Victor, J.R., Duarte, A.J.D.S., Clissa, P.B., Sanabani, S.S., 2020. Bacterial community composition and potential pathogens along the Pinheiros River in the southeast of Brazil. Sci. Rep. 10 (1), 1–9.

Guo, P., Zhang, K., Ma, X., He, P., 2020. Clostridium species as probiotics: potentials and challenges. Journal of Animal Science and Biotechnology 11 (1), 24.

Harwood, V.J., Staley, C., Badgley, B.D., Borges, K., Korajkic, A., 2014. Microbial source tracking markers for detection of fecal contamination in environmental waters: relationships between pathogens and human health outcomes. FEMS Microbiol. Rev. 38 (1), 1–40.

Holcomb, D.A., Stewart, J.R., 2020. Microbial indicators of fecal pollution: recent progress and challenges in assessing water quality. Curr Environ Health Rep 7 (3), 311–324.

Karunakaran, E., Vernon, D., Biggs, C.A., Saul, A., Crawford, D., Jensen, H., 2016. Enumeration of sulphate-reducing bacteria for assessing potential for hydrogen sulphide production in urban drainage systems. Water Sci. Technol. 73 (12), 3087–3094.

Korajkic, A., McMinn, B.R., Harwood, V.J., 2018. Relationships between microbial indicators and pathogens in recreational water settings. Int. J. Environ. Res. Public Health 15 (12), 2842.

Martín, R., Miquel, S., Benevides, L., Bridonneau, C., Robert, V., Hudault, S., Chain, F., Berteau, O., Azevedo, V., Chatel, J.M., Sokol, H., Bermúdez-Humarán, L.G., Thomas, M., Langella, P., 2017. Functional characterization of novel Faecalibacterium prausnitzii strains isolated from healthy volunteers: a step forward in the use of F. prausnitzii as a next-generation probiotic. Front. Microbiol. 8, 1226.

McLellan, S.L., Huse, S.M., Mueller-Spitz, S.R., Andreishcheva, E.N., Sogin, M.L., 2010. Diversity and population structure of sewage-derived microorganisms in wastewater treatment plant influent. Environ. Microbiol. 12 (2), 378–392.

Mieszkin, S., Yala, J.F., Joubrel, R., Gourmelon, M., 2010. Phylogenetic analysis of Bacteroidales 16S rRNA gene sequences from human and animal effluents and assessment of ruminant faecal pollution by real-time PCR. J. Appl. Microbiol. 108 (3), 974–984.

Miller, J.D., Hutchins, M., 2017. The impacts of urbanisation and climate change on urban flooding and urban water quality: a review of the evidence concerning the United Kingdom. Journal of Hydrology: Regional Studies 12, 345–362.

Morita, A.K.M., et al., 2021. Pollution threat to water and soil quality by dumpsites and non-sanitary landfills in Brazil: a review. Waste Manag. 131, 163–176.

Mulugeta, S., et al., 2012. Contamination level and location of recreational freshwater influence the ability to predict Escherichia coli concentration by qPCR targeting Bacteroides. J. Environ. Manage. 103, 95–101.

Munro, K., Martins, C.P.B., Loewenthal, M., Comber, S., Cowan, D.A., Pereira, L., 2019. Evaluation of combined sewer overflow impacts on short-term pharmaceutical and illicit drug occurrence in a heavily urbanised tidal river catchment (London, UK). Science of the Total Environment 657, 1099–1111.

Newton, R.J., McLellan, S.L., Dila, D.K., Vineis, J.H., Morrison, H.G., Eren, A.M., Sogin, M.L., 2015. Sewage reflects the microbiomes of human populations. MBio 6, e02574s.

Niebaum, G., Berlekamp, J., Schmitt, H., Lämmchen, V., Klasmeier, J., 2023. Geo-referenced simulations of E. coli in a sub-catchment of the Vecht River using a probabilistic approach. Sci. Total Environ. 868, 161627.

Nishiyama, T., Ueki, A., Kaku, N., Watanabe, K., Ueki, K., 2009. Bacteroides graminisolvens sp. nov., a xylanolytic anaerobe isolated from a methanogenic reactor treating cattle waste. Int. J. Syst. Evol. Microbiol. 59 (Pt 8), 1901–1907. https://doi.org/10.1099/ijs.0.008268-0, 2009 Aug. Epub 2009 Jun 30. PMID: 19567576.

Osborn, A.M., Smith, C.J. (Eds.), 2005. Molecular Microbial Ecology. Garland Science.

Peed, L.A., Nietch, C.T., Kelty, C.A., Meckes, M., Mooney, T., Sivaganesan, M., Shanks, O. C., 2011. Combining land use information and small stream sampling with PCR-based methods for better characterization of diffuse sources of human fecal pollution. Environ. Sci. Technol. 45 (13), 5652–5659.

Pettibone, G.W., 1998. Population dynamics of Aeromonas spp. in an urban river watershed. J. Appl. Microbiol. 85 (4), 723–730.

Raza, S., Kim, J., Sadowsky, M.J., Unno, T., 2021. Microbial source tracking using metagenomics and other new technologies. J. Microbiol. 59 (3), 259–269.

Reynolds, L.J., Martin, N.A., Sala-Comorera, L., Callanan, K., Doyle, P., O'Leary, C., Buggy, P., Nolan, T.M., O'Hare, G.M.P., O'Sullivan, J.J., Meijer, W.G., 2021. Identifying sources of faecal contamination in a small urban stream catchment: a multiparametric approach [original research]. Front. Microbiol. 12.

Sagova-Mareckova, M., Alonso, L., Vautrin, F., Moenne-Loccoz, Y., Müller, E., Ploy, M.C., Moreira, D., 2021. Impact of environmental factors on the abundance and diversity of antibiotic-resistant bacteria in freshwater ecosystems. Environ. Sci. Technol. 55 (24), 16724–16734.

Saingam, P., Li, B., Yan, T., 2020. Fecal indicator bacteria, direct pathogen detection, and microbial community analysis provide different microbiological water quality assessment of a tropical urban marine estuary. Water Res. 185, 116280.

Shanks, O.C., Kelty, C.A., Sivaganesan, M., Varma, M., Haugland, R.A., 2009. Quantitative PCR for genetic markers of human fecal pollution. Appl. Environ. Microbiol. 75 (17), 5507–5513.

Shanks, O.C., et al., 2013. Comparison of the microbial community structures of untreated wastewaters from different geographic locales. Appl. Environ. Microbiol. 79 (9), 2906–2913.

Siezen, R.J., Kleerebezem, M., 2011. The human gut microbiome: are we our enterotypes? J. Microbial. Biotechnol. 4 (5), 550.
Stec, J., Kosikowska, U., Mendrycka, M., Stępień-Pyśniak, D., Niedźwiedzka-Rystwej, P., Bębnowska, D., Hrynkiewicz, R., Ziętara-Wysocka, J., Grywalska, E., 2022. Opportunistic pathogens of recreational waters with emphasis on antimicrobial resistance—a possible subject of human health concern. Int. J. Environ. Res. Public Health 19 (12), 7308.
Sun, Z., Zhu, X., Pan, Y., Zhang, J., 2017. Assessing terrestrial water storage and flood potential using GRACE data in the Yangtze River Basin, China. Remote Sensing 9 (10), 1011.
Tian, Z., Yu, Z., Li, Y., Ke, Q., Liu, J., Luo, H., et al., 2022. Frontiers in Environmental Science, Vol. 10. https://doi.org/10.3389/fenvs.2022.887446.
Tornevi, A., Bergstedt, O., Forsberg, B., 2014. Precipitation effects on microbial pollution in a river: lag structures and seasonal effect modification. Plos One 9 (5), e98546.
Vadeboncoeur, Y., et al., 2018. The role of ammonium in mediating climate change effects on carbon and nitrogen fluxes in agricultural catchments. Biogeochemistry 139 (1), 1–16.
van Heijnsbergen, E., et al., 2022. (Antibiotic-resistant) *E. coli* in the Dutch–German Vecht catchment—monitoring and modeling. Environ. Sci. Technol. 56 (21), 15064–15073.
Werner, D., Acharya, K., Blackburn, A., Zan, R., Plaimart, J., Allen, B., Mgana, S.M., Sabai, S.M., Halla, F.F., Massawa, S.M., Haile, A.T., Hiruy, A.M., Mohammed, J., Vinitnantharat, S., Thongsamer, T., Pantha, K., Mota Filho, C.R., Lopes, B.C., 2022. MinION nanopore sequencing accelerates progress towards ubiquitous genetics in water research. Water 14, 2491.
Willis, A.D., 2019. Rarefaction, alpha diversity, and statistics. Front. Microbiol. 10.
Xie, Y., Liu, X., Wei, H., Chen, X., Gong, N., Ahmad, S., Lee, T., Ismail, S., Ni, S.-Q., 2022. Insight into impact of sewage discharge on microbial dynamics and pathogenicity in river ecosystem. Sci. Rep. 12 (1), 6894.
Yergeau, E., Lawrence, J.R., Sanschagrin, S., Waiser, M.J., Korber, D.R., Greer, C.W., 2012. Next-generation sequencing of microbial communities in the Athabasca River and its tributaries in relation to oil sands mining activities. Appl. Environ. Microbiol. 78, 7626–7637.

Further reading

Haugland, R.A., Varma, M., Sivaganesan, M., Kelty, C., Peed, L., Shanks, O.C., 2010. Evaluation of genetic markers from the 16S rRNA gene V2 region for use in quantitative detection of selected Bacteroidales species and human fecal waste by qPCR. Syst. Appl. Microbiol. 33 (6), 348–357.
Siefring, S., Varma, M., Atikovic, E., Wymer, L., Haugland, R.A., 2008. Improved real-time PCR assays for the detection of fecal indicator bacteria in surface waters with different instrument and reagent systems. J. Water Health 6 (2), 225–237.