Science of the Total Environment 829 (2022) 154257

Contents lists available at ScienceDirect

Science of the Total Environment

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

Tracking bacterial pollution at a marine wastewater outfall site – A case study from Norway

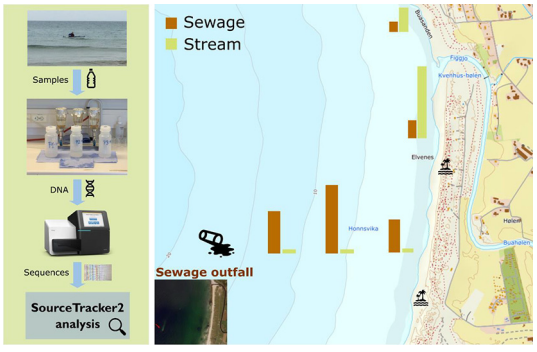
Andrea Bagi [a,*], Geir Skogerbø [b]

[a] NORCE Norwegian Research Centre, Marine Ecology, Mekjarvik 12, 4070 Randaberg, Norway
[b] IVAR IKS, Breiflåtveien 16, 4018 Stavanger, Norway

HIGHLIGHTS

- Whole community-based microbial source tracking provides valuable insights.
- SourceTracker revealed spreading of sewage effluent towards the beach.
- Both the sewage effluent and the stream contribute to fecal pollution at Bore beach.
- Low level presence of antimicrobial genes in treated sewage and seawater
- Overlap between source communities complicates source identification.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 30 November 2021
Received in revised form 9 February 2022
Accepted 27 February 2022
Available online 2 March 2022

Editor: Warish Ahmed

Keywords:
Metabarcoding
Metagenomics
Marine sewage outfall
Fecal pollution
Microbial source tracking

ABSTRACT

Coastal marine environments are increasingly affected by anthropogenic impacts, such as the release of sewage at outfall sites and agricultural run-off. Fecal pollution introduced to the sea through these activities poses risks of spreading microbial diseases and disseminating antibiotic resistant bacteria and their genes. The study area of this research, Bore beach, is situated between two such point sources, an outfall site where treated sewage is released 1 km off the coast and a stream that carries run-off from an agricultural area to the northern end of the beach. In order to investigate whether and to what extent fecal contamination from the sewage outfall reached the beach, we used microbial source tracking, based on whole community analysis. Samples were collected from sea water at varying distances from the sewage outfall site and along the beach, as well as from the sewage effluent and the stream. Amplicon sequencing of 16S rRNA genes from all the collected samples was carried out at two time points (June and September). In addition, the seawater at the sewage outfall site and the sewage effluent were subject to shotgun metagenomics. To estimate the contribution of the sewage effluent and the stream to the microbial communities at Bore beach, we employed SourceTracker2, a program that uses a Bayesian algorithm to perform such quantification. The SourceTracker2 results suggested that the sewage effluent is likely to spread fecal contamination towards the beach to a greater extent than anticipated based on the prevailing sea current. The estimated mixing proportions of sewage at the near-beach site (P4) were 0.22 and 0.035% in June and September, respectively. This was somewhat below that stream's contribution in June (0.028%) and 10-fold higher than the stream's contribution in September (0.004%). Our analysis identified a sewage signal in all the tested seawater samples.

1. Background

The health of coastal ecosystems and the range of services they provide for our society are highly dependent on the local water management practices adjacent to shorelines (Jurelevicius et al., 2021). Among the many

* Corresponding author.
E-mail address: anba@norceresearch.no (A. Bagi).

http://dx.doi.org/10.1016/j.scitotenv.2022.154257

anthropogenic influences, wastewater treatment plant effluents (or untreated sewage effluents) and agricultural runoffs represent major sources of contamination introducing excess organic material, nutrients, pollutants, and potentially pathogenic microorganisms into the marine environment (Daby et al., 2002; Owili, 2003). Besides their overall effect on ocean health, sewage outfalls and agricultural runoffs are a concern also due to the health hazards they represent to recreational users of beaches. In addition, they may contribute to the spreading of antibiotic resistance genes (Griffin et al., 2020). Wastewater contains various pathogens with considerable risk to human health. These pathogens include enteric bacteria (*Salmonella*, *Brucella*, pathogenic *E. coli*), parasitic protists (*Giardia*, *Cryptosporidium*) and viruses (adenoviruses, hepatoviruses). Cases of gastrointestinal and respiratory illnesses caused by pathogens do occur upon exposure to seawater contaminated with fecal waste, around the world (Prüss, 1998; WHO, 2003).

In order to protect human health, beach waters are monitored for fecal indicator bacteria (FIB), *e.g.*, thermophile coliforms, enterococci, *Escherichia coli*, *Clostridium perfringens*, and authorities may advise against beach use if FIB counts exceed a risk threshold (EU, 2006; Statens Helsetilsyn, 1994). Despite its widespread use and being the sole method included in most regulations for water quality assessment, the adequacy of FIB counts for detecting fecal contamination has been debated for a long time (Korajkic et al., 2018). This is mainly due to the fact that FIB does not only originate from anthropogenic fecal waste. FIB can be detected in fecal material of farmed and wild animals as well as non-fecal environmental sources such as soil. Differential fate of FIB in different environments, *e.g.*, marine *versus* freshwater as well as other of site-specific factors (*e.g.*, hydrological characteristics) further complicate the situation. At the same time, determining the origin of FIB at beaches is crucial for designing appropriate mitigation strategies to protect human health, therefore methods appropriate for that are needed.

In response to this need, several microbial source tracking (MST) approaches emerged as novel means of identifying sources of fecal pollution, and indicators highly specific to human and animal feces are continuously being established (Ahmed et al., 2015; Boehm et al., 2013; Hagedorn et al., 2011; Mathai et al., 2020; Steinbacher et al., 2021). MST methods mostly rely on molecular techniques, such as quantitative PCR or sequencing to discern the identity and quantity of indicator and host-specific microorganisms. Abundant and unique members of the human gut microbiome, *e.g.*, *Bacteroides* and *Prevotella* species, as well as their phages provide a reliable source for human-specific biomarker design. The most commonly used host-specific qPCR assays (for example *Bacteroides* assays) are typically targeting the 16S rRNA gene. Other genes, for example Shiga-like toxin 1 (*stx1*) and 2 (*stx2*) genes, the intimin (*eae*) gene of *E. coli* O157:H7 and non-virulence related alpha-1-6, mannanase gene of *Bacteroides thetaiotaomicron* have also been shown to work well to identify the host-source of detected fecal pollution, although they were employed only in a handful of studies (Aslan and Rose, 2013; Ibekwe et al., 2002; Vadde et al., 2019). Undoubtedly these methods are powerful tools to identify the host species (animal *versus* human) causing the fecal pollution. However, additional methods may be necessary when the potential sources of fecal bacteria (such as sewage effluent and river input) both contain human-derived pathogens, *i.e.*, the fecal bacteria are mainly associated with the same host (human). For resolving the environmental source of fecal pollution, comparative analysis of whole microbial communities has recently been developed, thanks to the emergence of cost-efficient next-generation sequencing technologies (Wani et al., 2021). Whole community analysis through amplicon sequencing of 16S rRNA genes provides insight into the presence of all bacteria simultaneously, while the above mentioned targeted methods, assess the presence of single microbial taxa (Unno et al., 2018). Once the community profiles are obtained through sequencing, there are several ways to assess the source of fecal pollution. Most recently, machine learning approaches able to either classify samples into predefined groups (random forest algorithms) or to calculate the contribution of potential sources (SourceTracker), have been developed (Knights et al., 2011; Roguet et al., 2018). The latter approach has proved to be a highly accurate method for MST (Dubinsky et al., 2016). It has also been further improved in order to be applicable for not only amplicon sequencing data but also to shotgun metagenomics datasets (McGhee et al., 2020).

In this study, we tested whole community analysis-based MST for fecal pollution source attribution through a case study at Bore beach (Borestranden) in Norway. Bore beach is located in the Jæren coastline (Rogaland, Norway), which is acknowledged as an international Hope Spot by the Mission Blue alliance, defining it as a place critical to the health of the oceans (https://mission-blue.org/hope-spots/). The beaches of Jæren have been protected since 1977, and in 2003 the Jærstrendene landscape conservation area was established (Miljøverndepartementet, 2003). Bore beach is considered one of the most beautiful beaches in Norway and a great surfing spot. It attracts around 100.000 visitors each year. The area surrounding this 3 km long beach consists of a camping site, holiday homes and agricultural land used for crop and livestock farming. From these areas, catchment water, and to some extent also wastewater, are directly collected by a small stream, Figgjo (also called Sele), which has its outlet at the northernmost end of the beach. In addition, there is a municipal wastewater treatment facility (WWTP) nearby, treating domestic wastewater and releasing its effluent offshore Bore beach. A recipient investigation is performed every four years along the entire Jæren coastline to monitor the impact of sewage release practices. During the last investigation, FIB detected near the wastewater outfall site raised the question whether FIB released at the sewage outfall might reach the beach. Increased recreational usage of the beach (bathing and surfing) in recent years has accentuated the importance of addressing this question. Therefore, our main goal was to identify whether and to what extent any FIBs at the beach might originate from the wastewater outfall or from the Figgjo stream. In addition, some of the samples were further explored to obtain insights into the presence of other fecal pollution indicators and into their antibiotic resistance profile.

2. Materials and methods

2.1. Study site

The WWTP facility in question performs chemical flocculation and subsequent separation of the suspended solids by filtration. The treated water (approximately 2000–10,000 m^3/day depending on the season and weather conditions) is collected in a reservoir and released through a pipeline to the sea periodically, in pulse-like manner (personal communication with WWTP operators). The prevailing northerly current is expected to carry the effluent plume away from the shore, however, this has not been documented. The wastewater outfall point is located approximately 1 km off the beach at 15 m depth. The tidal variation in the area is not large (diurnal tidal variation is of the order of 50 cm) and the recorded sea level difference at the two sampling points was within 20 cm. The stream draining into the beach is reported to have an average discharge of 10,5 m^3/s (https://en.wikipedia.org/wiki/Figgjoelva). There is at present no flow measurements on the river, but there are historical data (up to 2004, Norwegian Water Resources and Energy Directorate, https://sildre.nve.no/station/28.5.0) regarding precipitation and water level for a point 10 km upstream. Our calculations based on this historical data and recorded precipitation preceding the sampling days suggests that the river discharge would be on the order of 9.9 and 14.1 m^3/s on June 26 and September 11, respectively.

2.2. Sampling

The sampling was generally carried out in “good weather” conditions (Table 1), both for safety reasons, and also, since this is a recreational area, we assumed this would be the condition for which the investigation was most relevant. Seawater, stream water and treated wastewater samples were collected at three time points, 25 June (T0), 11 July (T1) and 10 September (T2) of 2019. The weather conditions at each sampling time are summarized in Table 1. The sample collection procedure was as follows. Seawater was sampled along the coast at Bore beach near a wastewater effluent outlet (P1–P7), water was taken from the Figgjo stream (P8) and from the effluent of the wastewater treatment plant at Bore (P9), with

Table 1
Weather conditions at Bore beach on the sampling days.

	June	July	September
Air temperature	17.6 °C	11.9 °C	12.7 °C
Wind speed	2.5 m/s	2.2 m/s	1.0 m/s
Wind direction	From Southeast	From North	From North
Rainfall	0 mm	0 mm	0 mm
Humidity	69%	98%	95%

locations as outlined in Fig. 1. Triplicate samples (1 L) were collected in sterile flasks for subsequent filtration and DNA extraction from all the 9 sampling points while a single 250 ml sample was collected for thermophile coliform analysis. Seawater samples were collected from a kayak (sampling points located using a handheld GPS device), samples P6 and P8 were collected by wading into the water.

Samples were sent to Eurofins Environment Testing Norway AS (Klepp, Norway) for the thermophile coliform analysis, which was according to the Norwegian Standard method NS4790 (Standard, 1990). Briefly, this method includes filtering of the water and incubation of the filters on a selective medium at 44.5 °C for 24 h. Thermophile coliforms are identified as dark blue colonies, which obtain their color from the pH indicator (methylene blue) turning blue in the presence of acid (lactic acid produced from lactose).

2.3. DNA extraction and sequencing

Water samples were brought to the laboratory within 3 h of sample collection and kept cold in the meantime. Filtration was carried out immediately upon arrival using 0.22 μm pore-size membrane filters (45 mm diameter, Millipore, USA). The volume of the filtered waters varied to some extent between the sampling points, with the majority being ~1 L while samples P6, P7 and P8 were ~500 ml or less (Supplementary Table S1). Of the treated wastewater (P9) only 20 ml could be passed through the filter. All filters were then either cut into small pieces and placed directly into a bead-tube prior to freezing or were placed in 50 ml Falcon tubes and frozen immediately. Filters stored intact were cut prior to the extraction step using sterile forceps and scissors. DNA was extracted using the DNA Power Lyser kit according to the manufacturer's instructions, including a 20-minute lysis step of vigorous mixing on a Vortex Genie at maximum speed (Qiagen). Quality and quantity of the total DNA was checked using NanoDrop and Qubit.

Triplicate raw DNA samples from two time points, T0 and T2, were sent to Novogene Europe (Cambridge, UK) for library preparation and 16S rRNA gene sequencing. The library preparation (PCR step) was carried out using the primers 341F (5′-CCTAYGGGRBGCASCAG-3′) and 806R (5′-GGACTACNNGGGTATCTAAT-3′) targeting the V3-V4 region (~466 bp). For sequencing, the Illumina Paired End 250 bp protocol was employed (expected throughput of 100,000 raw sequences per sample). In addition, triplicate wastewater (P9) and recipient seawater (P1) DNA samples were pooled from each time point and subjected to library preparation and shotgun sequencing at Novogene Europe (Cambridge, UK) using the Illumina Paired End 150 bp protocol. The number of samples sent for amplicon sequencing was restricted due to cost constraints. For the same reason, only samples from two sampling locations (triplicates pooled) were subjected to shotgun metagenomics.

Fig. 1. Overview of the sampling points at sea and at the stream (P1–P8) at Bore beach (Rogaland, Norway). The location of the wastewater treatment plant (P9) is not shown. Credit: norgeskart.no.

2.4. Amplicon data processing

Initial processing of the amplicons was performed by Novogene Europe (Cambridge, UK). Briefly, paired-end reads were demultiplexed, truncated to remove barcode and primer sequences and then merged using FLASH (V1.2.7). Quality filtering of the raw tags were performed to obtain high-quality (Q > 30) clean reads using a QIIME procedure (1.7.0). A reference-based chimera removal was then performed against the JGI Gold database with UCHIME. Finally, the reads were clustered into OTUs at 97% similarity and classified against the SILVA database (v132). Alpha diversity measures were calculated using the *alpha-phylogenetic* function in the *diversity* plugin of QIIME2 (Bolyen et al., 2019; Faith, 2007). Weighted UniFrac distances were calculated using the *distance* function (method = "wunifrac") in the *phyloseq* package (Lozupone et al., 2011; McMurdie and Holmes, 2013). SourceTracker2 was employed to analyse the relative contribution of the wastewater effluent and the stream to the composition of all other microbial communities (Knights et al., 2011). The command line SourceTracker2 tool was installed and used with default settings according to the instructions on the software's Github page (https://github.com/biota/sourcetracker2). As recommended by Ahmed et al. (2015), the sink samples were not rarefied, in order to include rare OTUs in the analysis. Raw sequences have been deposited to NCBI Sequence Read Archive (SRA) under the project number PRJNA752974. Accession numbers are specified in Supplementary Table S2A.

2.5. Metagenomics data processing

Processing of raw data including quality control, filtering, trimming, assembly, gene prediction, dereplication and the mapping of reads back onto the dereplicated genes have been carried out by the service provider (Novogene Europe, Cambridge, UK). These steps were performed as follows. Low quality bases (Q ≤ 38) exceeding 40 bp, reads with more than 10 consecutive N nucleotides and reads which overlapped with adapter sequences over 15 bp were trimmed. Clean_Q20 values (the percentage of bases whose quality score is greater than 20, *i.e.*, error rate < 0.01) were summarized. Samples passing QC were initially assembled using a SOAPdenovo protocol. The resulting scaffolds were trimmed at "N" to obtain fragments without "N". These scaftigs (*i.e.*, continuous sequences within scaffolds without N) were used as fragments to map clean reads back onto using Soap 2.21. Unutilized PE reads were collected for all samples after the first round of assembly and a mixed assembly was then conducted on the unutilized reads with the same assembly parameters. The "NOVO_MIX" genes represent the assembled genes from this second mixed assembly, while all other genes are named after the sample from which their scaftigs originated from. The scaftigs of each sample and mixed assembled scaftigs < 500 bp were trimmed and remaining scaftigs were used for further analysis and gene prediction. Scaftigs (≥500 bp) were used for ORF (Open Reading Frame) prediction by MetaGeneMark. ORFs ≤ 100 nt were trimmed and dereplicated by CD-HIT to generate gene catalogues. Dereplicating was performed with default parameters: identity of 95%, coverage of 90% (CD-HIT parameters: -c 0.95, -G 0, -aS 0.9, -g 1, -d 0). The longest gene was chosen as the representative gene (termed 'unigene'). Clean reads were mapped to the gene catalogue using SoapAligner with parameters: -m 200, -x 400, identity ≥ 95%. The resulting read count table was converted into relative counts for each gene (G_k) using the formula: $G_k = \frac{r_k}{L_k} \cdot \frac{1}{\sum \frac{r_i}{L_i}}$, where r_k is the read count of gene k and L_k is the length of gene k, r_i is the read count of the ith gene and L_i is the length if the ith gene. Read count table, relative count table, dereplicated nucleotide and amino acid sequences were used in subsequent analysis. KEGG pathway mapping and annotation of antibiotic resistance genes using the CARD database were performed by the service provider (Novogene Europe, Cambridge, UK). The provided assignments and read count tables were used for downstream statistical analysis as obtained. Raw sequences have been deposited to NCBI Sequence Read Archive (SRA) under the project number PRJNA752974. Accession numbers are specified in Supplementary Table S2B.

3. Results and discussions

In order to identify the source of fecal contamination at Bore beach, we performed amplicon sequencing (16S rRNA gene) of seawater samples along a line between the wastewater outfall and the beach, together with samples along the beach roughly 200 m away from the shore in June and September 2019. The microbial communities of the two potential sources (the Figgjo stream and the wastewater effluent) were analysed in the same manner. Thermophile coliform counts were determined according to standard techniques. In addition, we performed shotgun metagenomics on the collected wastewater samples and the seawater sample at the marine outfall site for an in-depth characterization of alternative fecal pollution markers and antibiotic resistance genes.

Sampling point P5 was chosen as reference site for this study due to the prevailing northerly current, presumably transporting the wastewater plume in the direction of sampling point P7. We considered this site (P5) to represent a near-beach community of microbes with no obvious point source contamination present, but likely influenced by diffuse terrestrial inputs. The sampling point P4 was chosen to represent the beach as this location should be most affected by the wastewater plume in the absence of any coastal currents.

3.1. Water quality analysis – fecal indicator bacteria (FIB)

The thermophile (fecal) coliform analysis showed that except for the reference site (P5) all other samples contained detectable levels of coliforms at least at one sampling time. The stream samples (P8) and the seawater at the stream outlet (P6) consistently showed the presence of coliforms. The beach-near seawater samples (P4, P5 and P7) had at all times coliform levels below 100 CFU/ml indicating "good" water quality according to the Norwegian guidelines, which employ the term "good water quality" when CFU/ml of thermophile coliform bacteria is below 100 (Statens Helsetilsyn, 1994) (Table 2).

The samples closest to the sewage outfall (P2–P4) varied between "not detected (<10 CFU/ml)" and 120 CFU/ml.

3.2. Overview of the microbial community profiles

Alpha diversity patterns varied between the two sampling times, with most samples increasing in diversity from June to September. The P2, P3 and P4 and the wastewater communities showed modest changes, while the rest of the samples exhibiting larger increases in phylogenetic diversity (Fig. 2) (Faith, 2007). Other measures of alpha diversity showed similar trends (Supplementary Table S3). As the Simpson indices suggested, most communities were species rich and the species were evenly distributed, with a small increase from June to September. The wastewater sample showed the lowest value in both seasons, with a somewhat lower index in September compared to June.

Amplicon sequencing of the June and September samples revealed seasonal changes in the resident marine microbial communities and a clear influence of the wastewater at the marine outfall site (P1) in September

Table 2
Thermophile coliform results determined by cultivation-based method (Eurofins AS). Dates 26.06.2019, 11.07.2019 and 11.09.2019 correspond to T0, T1 and T2, respectively.

Sampling point	Thermophile coliforms (CFU/100 ml)		
	26.06.2019	11.07.2019	11.09.2019
P1	<10	30	>1500
P2	<10	20	20
P3	<10	90	120
P4	<10	10	20
P5	<10	<10	<10
P6	40	670	250
P7	20	20	<10
P8	70	160	430

(Fig. 3). The most dominant genera in the P1-P5 samples included several typical temperate pelagic community members, *e.g.*, *Planktomarina*, *Synechococcus*, *Candidatus* Pelagibacter, OM60/NOR5 clade, *Candidatus* Puniceispirillum and unidentified SAR116 clade bacterium (Flombaum et al., 2013; Giebel et al., 2011, 2019; Oh et al., 2010; Yan et al., 2009). The major difference between the two seasons for these seawater samples were a higher relative abundance of *Flavobacteriaceae* in June compared to September. This family showed an increase in relative abundance from the P1 site towards P4 with a concurrent decrease of SAR116 clade and Family I of unidentified *Cyanobacteria*. In June, both the P6 and P7 samples resembled the stream community with a higher relative abundance of *Comamonadaceae*, *Moraxellaceae* and *Cytophagaceae* in comparison to the other seawater samples, while in September, only P6 sample was similar to the stream community. Based on the family level composition, the main difference between P8 and P9 samples was apparent from the relative abundances of *Campylobacteraceae*, *Moraxellaceae*, *Cytophagaceae*, *Comamonadaceae*, and *Sporichthyaceae*. Members of the family *Comamonadaceae* (a major group of *β*-proteobacteria) have been isolated from both freshwater environments as well as from activated sludge (Khan et al., 2002; Nuy et al., 2020). The OTUs classified as *Comamonadaceae* in the stream samples were mainly associated with *Limnohabitans* (reaching up to 21% relative abundance), while in the sewage samples they were mainly associated with *Acidovorax* and *Comamonas denitrificans* (reaching up to 2 and 1% relative abundance, respectively). Besides *Comamonadaceae*, the stream samples were characterized by a high relative abundance of *Cytophagaceae* (predominantly *Pseudarcicella*) and *Sporichthyaceae* (predominantly hgcl clade actinobacteria), while wastewater samples were clearly dominated by *Campylobacteraceae*, almost exclusively represented by genus *Arcobacter* (Fig. 3). The presence of hgcl clade actinobacteria in the stream sample is interesting as this group has been well-known as adapted to nutrient-poor conditions (Farkas et al., 2020). Considering the surrounding agricultural land, one would expect the Figgjo stream to be rather nutrient rich, favoring organisms typical in mesotrophic freshwater environments.

In order to further explore the differences between the two pollutant-source communities, we performed differential abundance analysis. DESeq2 identified 111 and 236 OTUs as significantly more abundant in the P8 and P9 samples, respectively (log2 fold change > 2 and adjusted *p*-value < 0.05, Supplementary Table S4). These differentially abundant OTUs belonged to typical freshwater environment bacteria (*e.g.*, *Limnohabitans*, *Sporichthyaceae*, *Polynucleobacter*) in case of the stream, while mainly to gut microbiota members (*e.g.*, Bacteroides, *Firmicutes*, *Prevotella*, *Lactococcus*, *Ruminococcus*) in addition to *Arcobacter*, in case of the wastewater.

Grouping of samples on the non-metric multidimensional scaling (NMDS) plot reflected a weak season-dependent pattern, which was confirmed by PERMANOVA (the adonis function in the vegan package, 9999 permutations, $R^2 = 0.147, p = 0.0054$) (Fig. 4). The two sources, P8 and P9, were clearly distinct from the rest of the samples (and each other) in June, while in September, P9 grouped close to the seawater sample at the outfall site (P1) and P8 to the seawater at the stream outlet site (P6). The major difference between the two sampling times was the apparent swap between P7 and P6 in terms of being the most similar to the stream sample. Most likely this indicated a change in the current nearest to the shore (turning southward instead of the typical northern direction). In addition, the grouping of two P5 (reference site) replicates closer to P6 and the stream in September was also apparent. This suggested a higher load of terrestrial input at P5 in September. Although there is no obvious source of freshwater input to sea near the reference site, it is not unlikely that run-off from the nearby agricultural lands reaches the shore *via* other routes. This is especially likely following heavy rains such as that was the case in September. In addition, we cannot rule out the possibility of groundwater seepage in the form of submarine groundwater discharge (Burnett et al., 2003; Luijendijk et al., 2020). The beta diversity analysis also suggested that the P1 sample is different from the nearest seawater samples (P2, P3 and P4) even when the wastewater plume was not observable on the surface. Samples P2, P3 and P4 grouped tightly together at both sampling times, while the P1 samples remained outside of this cluster.

Based on pairwise comparison of weighted UniFrac distances between P4 and all other samples in June, the composition of the prokaryotic community nearest the beach (P4) resembled most the reference site, followed by the P3 sample, the seawater at the stream outlet site and the P2 sampling point (Fig. 5). The situation changed somewhat in September, when the reference site and the seawater at the stream outlet site showed greater dissimilarities from P4 with a simultaneously higher similarity between the P4, P3, P2 and P7 samples. The lower similarity of P4 to the stream microbial community again suggested a southward spreading of the stream water in September. At both sampling times, P4 had an intermediate similarity to the wastewater outfall site with a greater difference in September, when the P1 sample was clearly dominated by wastewater microorganisms.

3.3. SourceTracker2 analysis suggests sewage being transported towards the shore

Machine learning approaches have been shown to accurately identify sources of fecal pollution, particularly with the Bayesian algorithm-based program, SourceTracker (Henry et al., 2016; Mathai et al., 2020), which was able to identify human fecal pollution along a Russian river with great specificity, using microbial community data obtained directly from fecal material of different animal species (Dubinsky et al., 2016). Nevertheless, one caveat of SourceTracker analysis is important to acknowledge, namely the potential bias (inflation of false positive results) from overlapping bacterial composition of source

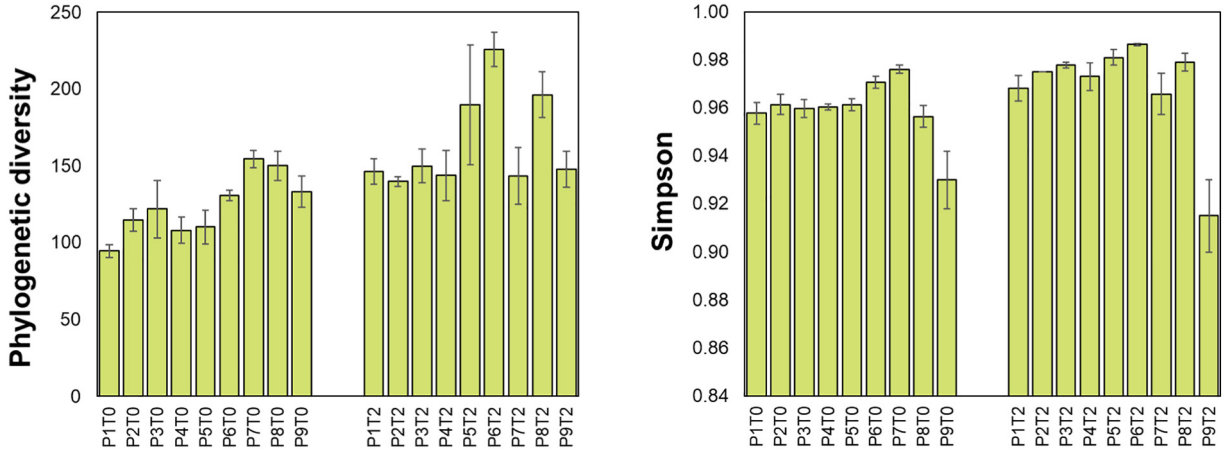


Fig. 2. Alpha diversity metrics: Faith's phylogenetic diversity and Simpson index of the 9 sampling points in June (T0) and September (T2).

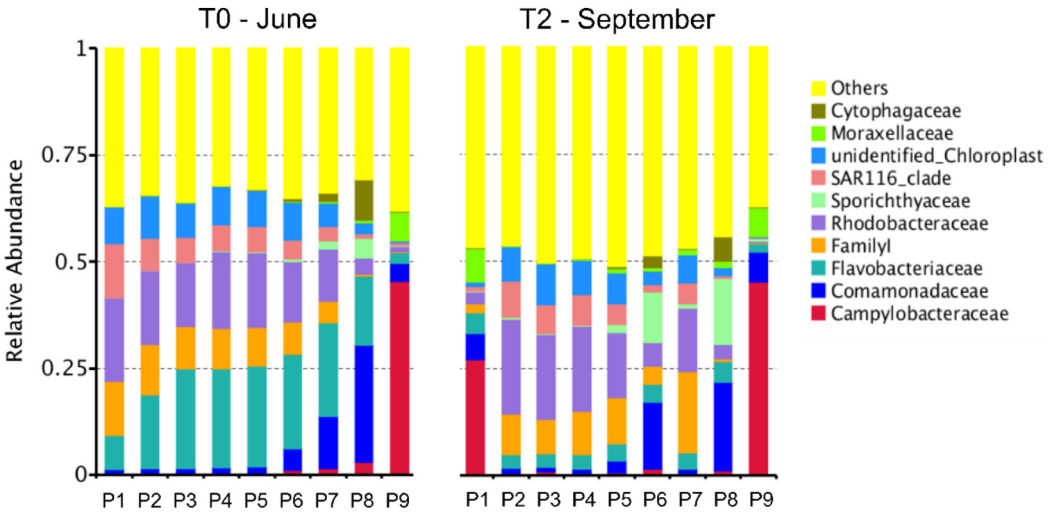


Fig. 3. Family level composition of the 18 microbial communities from the 9 sampling points at two sampling times. Data from triplicates was pooled and only the top 10 families are shown.

communities. Studies performed in the context of fecal pollution tracking in marine and freshwater environments found that taking the local environmental context into account can curb this problem and is in fact essential for accurate source predictions (Hägglund et al., 2018; Mathai et al., 2020). Following the recommendation from Hägglund et al. (2018) for ensuring the highest achievable accuracy for our analysis, we included the reference site (P5) as a third source in the SourceTracker analysis. This site was expected to represent a typical beach-near seawater community in the area, with no obvious point source pollution and thermophile coliform numbers below detection limit.

Our SourceTracker2 results indicated that the stream had low mixing proportions at P1, P2, P3 and P4 sampling points at both sampling times, and in September at P7 (0.004–0.28%, Supplementary Table S6, Fig. 6A). The contribution of the stream was highest in the June P7 sample (33.99%, compared to 0.052% in September). Sewage mixing proportions at the sampling points closest to the sewage outfall site (P2, P3 and P4) were also low, ranging between 0.002 and 1.31% (Fig. 6C, Supplementary Table S6). The predicted sewage signal at these three points did not correlate well with the mixing proportions obtained at P1 at any given time. This is probably best explained by the pulse-like release of sewage effluent from the wastewater treatment plant. Thus, the results obtained in June may be explained by the sampling being done some time after a pulse of sewage effluent, resulting in no observable plume at P1 (reflected in a sewage signal of 0.002%), but detectable sewage signals closer to the shore as the sewage was transported away from the release site (Fig. 6C). In September, the sampling event coincided with a pulse, hence the observed high mixing proportions at P1 (and a visible plume). At the beach-near site (P4) the SourceTracker2 results indicated approximately equal mixing proportions from the stream and from the sewage in June (0.22% and 0.28% from sewage and stream, respectively, Fig. 6B). In September, the overall mixing proportions were lower than in June for both sewage and river, but the sewage mixing proportion was magnitude higher contribution than that of the stream (Fig. 6B).

Considering the low number of thermophile coliform counts detected in our study site, it was not surprising to find low mixing proportions for sewage and stream sources at P2, P3 and P4 locations. According to Staley et al. (2018) SourceTracker could accurately predict as low as 0.025 v/v% fecal material spiked into natural water samples. Thus, the SourceTracker predictions may be accurate despite the apparently low sewage (and stream) contributions in in several of the samples. Although Hägglund et al. (2018) demonstrated a strong positive correlation between coliform counts and

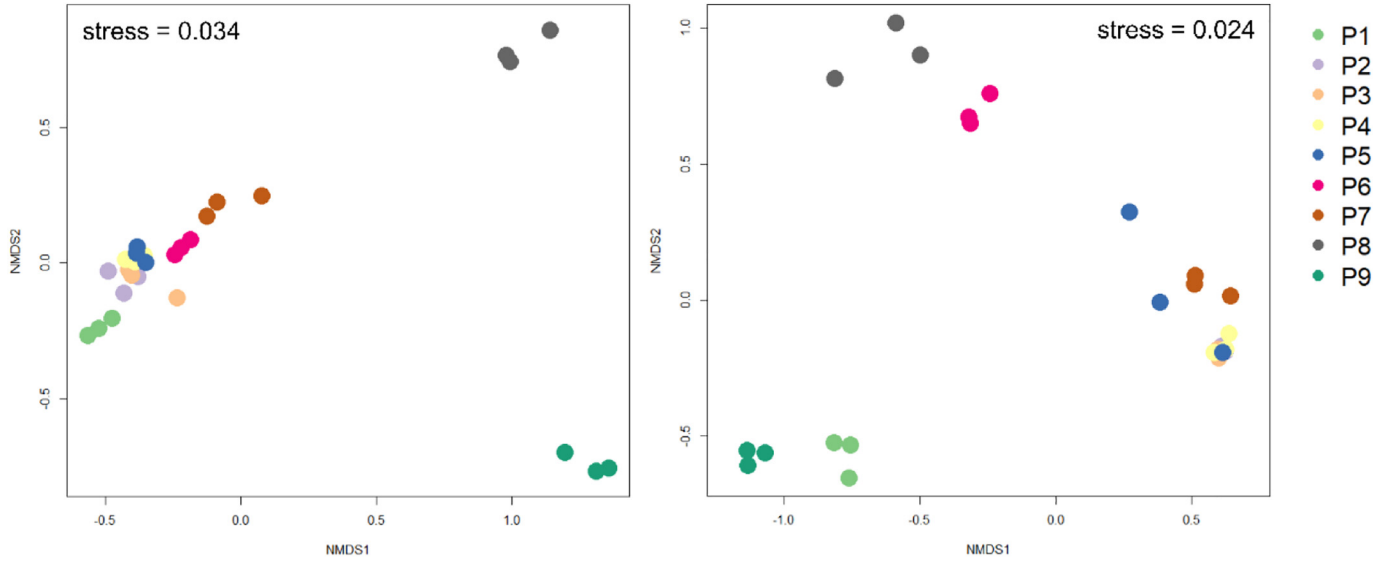


Fig. 4. Non-metric multidimensional scaling (NMDS) representation of community distances (distance measure = Bray-Curtis). June (A) and September (B) samples.

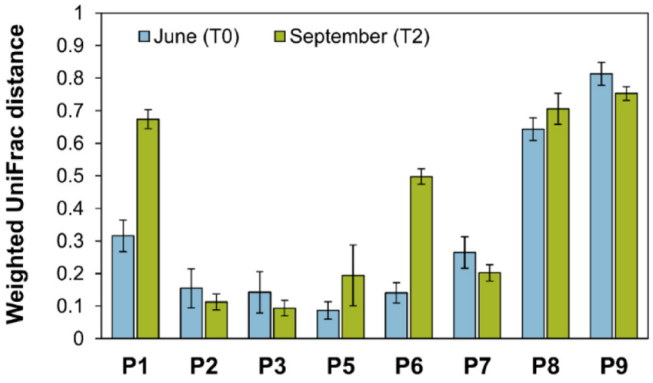


Fig. 5. Weighted UniFrac distances between P4 samples (closest to the beach) and all other sampling points. Completely distinct = 1 and identical = 0, *i.e.*, the higher the bar the less similar the sample is to the seasonally corresponding P4 sample.

the mixing proportions assigned to a pooled fecal contamination source (sum of several potential fecal materials contributing to the pollution), in our case, there was no such correlation between the two parameters. In June, there were no detectable coliforms in P1, P2, P3, P4 and P5 (<10 CFU/100 ml) despite SourceTracker2 indicating a sewage signal in P2, P3 and P4 (0.66% at P4).

The SourceTracker2 sewage mixing proportions at P2, P3 and P4 were approximately one order of magnitude higher in June than in September, despite coliforms only being detected in September. We considered whether the lower mixing proportions estimated by SourceTracker2 for September might be the consequence of alterations in the microbial community at the reference point (P5) between June and September. The reference point was chosen to represent a beach-near seawater community that was not much influenced by the two main sources of potential pollution (the river and the sewage outfall) but still so that it might be influenced by diffuse seepage or seasonal run-off from the adjacent land area. The community composition at P5 was very similar to P2, P3 and P4 in June, while in September, the P5 community appeared to be more similar to the river. A decrease in weighted UniFrac distances between P8 and P5, and P9 and P5 from June to September corroborated this (distance between P8 and P5 decreased from 0.63 to 0.56 and between P9 and P5 it decreased from 0.81 to 0.67, from June to September, respectively). Accumulated precipitation over the week preceding the sampling dates was six times higher in September than in June (7 and 42 mm in June and September, respectively), and may have caused a stronger "unspecific" run-off from the land in September leading to an altered microbial composition at P5. It appears reasonable that an altered microbial community at P5 in September would influence the SourceTracker2 estimates of the stream and sewage mixing proportions. Thus, we performed additional SourceTracker2 runs on the September data to look at the possible effect of this, first using the June reference community (P5T0) as the third source, and second, removing P5 as a third source altogether (Supplementary Table S7). Re-running the SourceTracker2 analysis of the September samples with these alterations increased both the sewage and the stream mixing proportions at P2, P3 and P4, 5–20-fold. While these tests in no way constitute proof of the mixing proportions in September, they do indicate the possibility that the admixture of pollutants along the beach may have been higher than the initial SourceTracker2 analysis indicated.

Taken together, our results highlight that overlap between the source communities poses a great challenge which needs to be addressed in the context of complex source tracking problems. Despite this challenge, the different approaches tested here all pointed to the conclusion that both the sewage effluent and the stream water spreads near the beach in ways we did not anticipate. In order to be able to interpret our observations, investigating the local mixing patterns of these two freshwater inputs (sewage and stream) under various weather and tidal conditions would be vital. The sampling in this study was carried out under calm wind conditions, which will probably reflect those of traditional recreational usage of the beach (bathing), but not necessarily more modern forms (surfing and wind surfing) that require rougher weather conditions. As prevailing winds in the area are westerly, it is entirely conceivable that more contamination from the sewage and/or the stream may be transported towards or along the beach, increasing the potential health risk for bathers, in particular, surfers. As the usage of the site for recreational purpose is likely increase, it seems reasonable to recommend further investigations, particular under rougher weather conditions. In addition, laboratory experiments could help validating the accuracy of SourceTracker predictions, demonstrating its ability to assign correct mixing proportions to similar contamination sources.

3.4. Presence of fecal pollution indicators: Bacteroides, Arcobacter and crAssphage

3.4.1. Bacteroides

Bacteroides species are common and abundant members of the gut microbiota of several warm-blooded animals and are obligate anaerobes exhibiting limited aerotolerance. Their presence in aerobe aquatic environments is therefore not expected and most likely signals fecal pollution. In our amplicon dataset, 44 OTUs belonged to the genus *Bacteroides*, of which 16 could only be assigned to the genus level (*Bacteroides*) while the remaining 28 OTUs were classified into 24 different *Bacteroides* species.

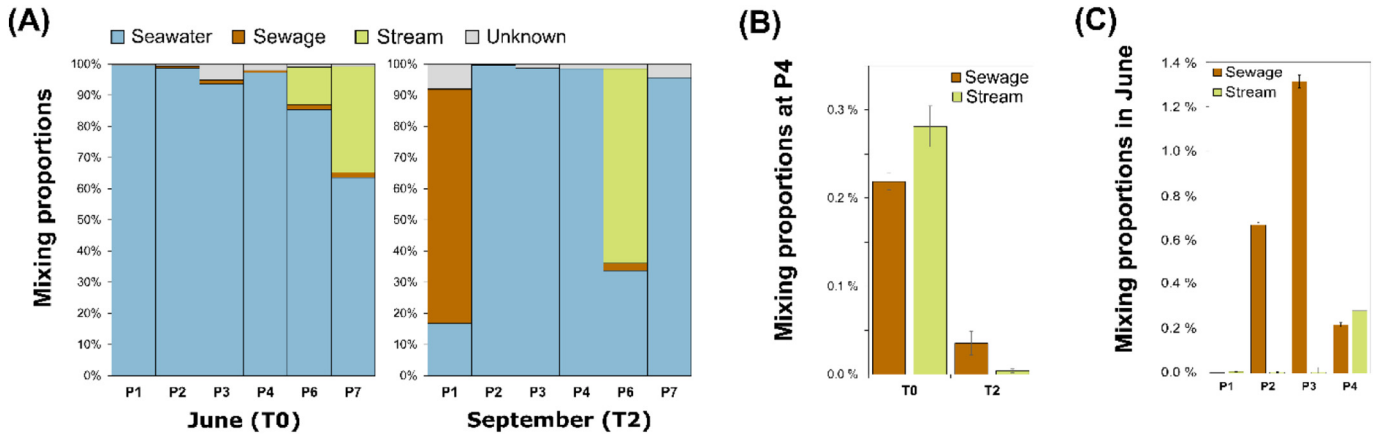


Fig. 6. Mixing proportions estimated by SourceTracker2 with 3 sources specified: P5 = seawater, P8 = stream and P9 = sewage. Mixing proportions in all samples (A), mixing proportions of sewage and stream in the P4 samples (B) and mixing proportions of sewage and stream in June between the outfall site (P1) and the near-beach sample (P4) (C).

Most of these species were absent from the P1 sample in June and from the P7 sample in September. Moreover, they remained below 0.1% in all other seawater samples, except for the P5 and P6 locations in September. The stream contained on average 0.17% and 0.40% of these *Bacteroides* species, while their relative abundance was highest in the sewage with 1.55 and 4.50% in June and September, respectively. Several species-specific 16S rRNA gene-based markers have been developed to distinguish human from animal fecal pollution. This approach relies on the findings that some *Bacteroides* species are more frequently found in human fecal material than in animal. Most of the Bacteroides species at Bore were present in at least 10 samples (out of 18), *B. cellulosilyticus* being found in all samples. There was no apparent difference between the *Bacteroides* species of the sewage and the stream, suggesting that the fecal pollution in the stream may be mostly from human origin.

3.4.2. Arcobacter

Arcobacter dominated the wastewater effluent sample in both the 16S rRNA amplicon and the shotgun metagenomic data, with relative abundances of 26–45% in the amplicon data and 12–17% in the shotgun metagenomic data. High relative abundance of *Arcobacter* in a wastewater microbial community is not unprecedented despite *Arcobacter* making up only a small fraction of the human gut microbiome in most healthy people (Fisher et al., 2014). *Arcobacter* is thus commonly found in domestic wastewater (Merga et al., 2014) as demonstrated by several studies (Fisher et al., 2014; Kristensen et al., 2020; Lu et al., 2015). Moreover, the MiDAS database (McIlroy et al., 2017) has 81 entries of *Arcobacter* in wastewater influent, for which the average and maximum contents are 4 and 21%, respectively. In our amplicon dataset, 17 OTUs were classified as *Arcobacter*, more specifically as *Arcobacter sp.*, *Arcobacter cryaerophilus* and Arcobacter_sp._L. The highest relative abundance of *Arcobacter* (sum of all *Arcobacter* OTUs) was observed in the wastewater effluent sample in both seasons (45 ± 4%) followed by the outfall site in September (26 ± 4%) and the stream sample in June (3 ± 0.2%). In general, the genus *Arcobacter* appears to be a group with several human and animal pathogen members as well as species not known to be pathogenic (Ferreira et al., 2016; Figueras et al., 2014). A triad of *Arcobacter* species: *Arcobacter butzleri*, *Arcobacter cryaerophilus*, and *Arcobacter skirrowii* have been recognized as emerging foodborne pathogens with antimicrobial resistance potential (Nguyen et al., 2020). Both *A. cryaerophilus* and *A. skirrowii* have been associated with cases of persistent diarrhea while *A. butzleri* has been defined by the International Commission on Microbiological Specifications for Foods as an important zoonotic agent and a serious hazard to human health through seafood contamination (Figueras et al., 2014; Levican et al., 2016; Wybo et al., 2004). In our amplicon dataset, only *A. cryaerophilus* was detected, while all three species were present in the shotgun metagenomics data (Fig. 7 and Table 3). Approximately one third of the *Arcobacter* OTUs in the sewage outfall was classified as *A. cryaerophilus*, making up 13–18% of the sewage effluent community. The relative abundance of *A. cryaerophilus* in the seawater samples (with the exception of P1T0) and the stream remained below 0.5% and was below 0.01% in the P1T0 and P7T2 samples.

Previous studies of Danish wastewater have shown that pathogenic and other bacteria are removed in treatment plants at efficiencies of 99% or more. On the other hand, *Arcobacter* was significantly more abundant in the supernatant than in activated sludge, suggesting that *Arcobacter* cells did not attach well to the activated sludge flocs. Thus, if we consider that the wastewater treatment at Bore consists only of chemical flocculation (using a polymer), we may have a situation in which most other bacteria are precipitated along with the suspended matter, leaving *Arcobacter* as a major constituent of the remaining bacterial community.

3.4.3. crAssphage

Besides fecal indicator bacteria, the *Bacteroides* phage, *crAssphage*, characteristic of human feces, was shown to provide a sensitive novel biomarker of fecal pollution of human origin (Farkas et al., 2019; Karkman et al., 2019). Although *crAssphage* has also been detected in animal feces (less frequently than in human), Ahmed et al. (2018) found that all tested

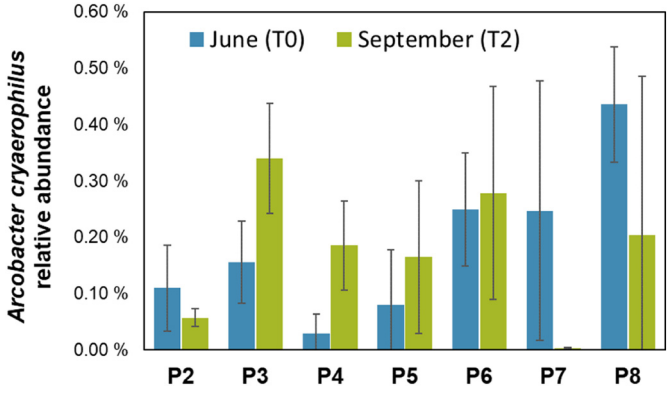


Fig. 7. Relative abundances of *Arcobacter cryaerophilus* in the 16S rRNA amplicon dataset (P1 site not shown due to high relative abundance at T2).

crAssphage assays had >90% specificity to human feces. We investigated whether this marker was present in our dataset and indeed found sequences classified as *crAssphage* in the wastewater effluent and P1 samples. Read numbers assigned to *crAssphage* genes were significantly lower at P1 when the wastewater plume was not observed at the surface (T0 and T1; 37 and 72 reads, respectively, ~0.0004–0.0008% relative abundance) than in the wastewater (P9) or at P1 when the wastewater plume was evident (T2; 8466 sequence reads, ~0.02–0.07% relative abundance). Unfortunately, the stream sample could not be included in the shotgun metagenomic analysis, hence we do not have information regarding the presence of this marker there.

3.5. Antibiotic resistance profiles

Excessive use of antibiotics has launched a phenomenon that some researchers early on predicted might become the next global pandemic: the pandemic of antibiotic resistance (Nadimpalli et al., 2021). Despite the early recognition of antimicrobial resistance transferability and warnings regarding (1) the potential effects of antimicrobial overuse, (2) antibiotic resistance, its spreading, and (3) effects on human health, it has only relatively recently become a hot topic (Mercer et al., 1971; Prestinaci et al., 2015). Among the many sources of antibiotic resistance genes, wastewater effluents are of particular interest because they distribute both antibiotic residues, their metabolic products, and resistant bacteria into aquatic environments. Often these aquatic environments are then used as drinking water supply, or for recreational purposes, or as sites for aquaculture industry. ARBs in sewage effluents mixing with aquatic microorganisms can exchange resistance genes through horizontal gene transfer (Tomova et al., 2015). Here, we used the shotgun metagenomic data from the sewage effluent and seawater from the outfall site at Bore beach to gain insight into the local antibiotic resistance gene profile. To our knowledge, this is the first study mapping ARGs in the area.

In total, 1997 genes (0.09–0.23% of all genes) matched antibiotic resistance related genes of the Comprehensive Antibiotic Resistance Database (CARD; https://card.mcmaster.ca/), representing 231 resistance types. The most prevalent and abundant ones are shown in Table 4. Proteobacteria and Actinobacteria ARGs were largely of the type RND antibiotic efflux pumps, while Firmicutes harbored mainly the *vanR* type ARGs and Bacteroidetes the *vanR* and *dfr* type resistance genes. Examples of antibiotic resistance-related

Table 3
Relative abundances of *Arcobacter cryaerophilus*, *Arcobacter butzleri* and *Arcobacter skirrowii* in the metagenomic dataset.

	P1T0	P1T1	P1T2	P9T0	P9T1	P9T2
A. cryaerophilus	0.013%	0.013%	1.0%	2.7%	2.4%	1.6%
A. butzleri	0.0023%	0.0022%	0.21%	0.48%	0.43%	0.36%
A. skirrowii	0.0054%	0.0028%	0.099%	0.25%	0.23%	0.15%

genes from the abundant *Arcobacter* genus included the *acrB* (multidrug efflux pump), RND antibiotic efflux pumps, multidrug ABC transporter permease, multidrug efflux SMR transporter, multidrug resistance (MDR) efflux transporter family protein and the bleomycin resistance family protein. Isolates of the *Arcobacter* species discussed above (*A. cryaerophilus*, *A. butzleri*, *A. skirrowii*) obtained from environmental samples have been shown to exhibit resistance to several antibiotics, including tetracycline (Sciortino et al., 2021). Nevertheless, in our dataset, only a single *tetA* gene (coding for *tetracycline resistance protein*) was found among the *Arcobacter* genes.

4. Conclusions

Taken together, our findings confirm that there is a low level persistent fecal pollution at Bore beach, caused by the combination of the catchment water delivered by the stream and the sewage outfall site. Potentially pathogenic species of *Bacteroides* and *Arcobacter* were found in all samples. The machine learning approach (SourceTracker2) also indicated a stronger spread of the sewage effluent towards the beach than might have been expected based on the general knowledge concerning the prevailing ocean current off the beach. However, a more detailed tracing the transport trajectory of the sewage plume would be necessary to reach a more definite conclusion with respect to whether and to what extent sewage contaminants reach the bathing and surfing areas. Further investigations into the actual route of the sewage plume along the coast would be advantageous in this respect, particularly under rougher weather conditions. In addition, examinations of the microbial communities of the sediment in the area could reveal whether there is a reservoir of sewage-related microbes that may become resuspended under certain circumstances. Furthermore, in order to fully protect the health of recreational users at Bore, the contribution of the stream should not be neglected, as this source delivers large quantities of water closer to the sites of usage. The published dataset represents a modest yet pioneering recollection of antibiotic resistance genes from the area, providing a baseline for future research. Our results suggest a low level of antimicrobial gene presence in both the sewage and the seawater, composed mostly of efflux systems. These systems are likely to be relevant to bacterial adaptation to the gut environment, rather than to synthetic antibiotics, suggesting a relatively low risk in terms of ARG dissemination from the sewage outfall.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2022.154257.

CRediT authorship contribution statement

Andrea Bagi: Conceptualization, Investigation, Formal analysis, Visualization, Writing – original draft, Project administration. **Geir Skogerbø:** Conceptualization, Investigation, Writing – review & editing, Supervision, Funding acquisition, Project administration.

Table 4
The most prevalent and abundant antibiotic resistance genes and their characteristics.

Resistance gene	Description
*mtr*A	Transcriptional activator of the *Mtr*CDE multidrug efflux pump of *Neisseria gonorrhoeae*.
*arl*R	Response regulator that binds to the *nor*A promoter to activate expression.
*rpo*B2	Rifampin-sensitive beta-subunit of RNA polymerase (*rpo*B)/rifampin-resistant beta-subunit of RNA polymerase (*rpo*B2) genes.
ugd	Required for the synthesis and transfer of 4-amino-4-deoxy-L-arabinose (*Ara*4N) to Lipid A, which allows gram-negative bacteria to resist the antimicrobial activity of cationic antimicrobial peptides and antibiotics such as polymyxin.
*bcr*A	ABC transporter found in *Bacillus licheniformis* that confers bacitracin resistance.
*van*RF	*van*R variant found in the *van*F gene cluster
*par*Y mutant	Point mutation on the *Streptomyces rishiriensis par*Y resulting in aminocoumarin resistance.
*sul*4	Dihydropteroate synthase gene and mobile sulfonamide resistance gene shown to confer resistance when expressed in *E. coli*.
*kdp*E	Transcriptional activator that is part of the two-component system *Kdp*D/*Kdp*E that is studied for its regulatory role in potassium transport and has been identified as an adaptive regulator involved in the virulence and intracellular survival of pathogenic bacteria.

Declaration of competing interest

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

Acknowledgements

Funding: This work was supported by the Regional Research Fund (Regionalt forskningsfond Vestlandet, RFF Vest), grant number: 299709.

References

Ahmed, W., Staley, C., Sadowsky, M.J., Gyawali, P., Sidhu, J.P.S., Palmer, A., et al., 2015. Toolbox approaches using molecular markers and 16S rRNA gene amplicon data sets for identification of fecal pollution in surface water. Appl. Environ. Microbiol. 81, 7067–7077.
Ahmed, W., Payyappat, S., Cassidy, M., Besley, C., Power, K., 2018. Novel crAssphage marker genes ascertain sewage pollution in a recreational lake receiving urban stormwater runoff. Water Res. 145, 769–778.
Aslan, A., Rose, J.B., 2013. Evaluation of the host specificity of Bacteroides thetaiotaomicron alpha-1-6, mannanase gene as a sewage marker. Lett. Appl. Microbiol. 56, 51–56.
Boehm, A.B., Van De Werfhorst, L.C., Griffith, J.F., Holden, P.A., Jay, J.A., Shanks, O.C., et al., 2013. Performance of forty-one microbial source tracking methods: a twenty-seven lab evaluation study. Water Res. 47, 6812–6828.
Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., et al., 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat. Biotechnol. 37, 852–857.
Burnett, W.C., Bokuniewicz, H., Huettel, M., Moore, W.S., Taniguchi, M., 2003. Groundwater and pore water inputs to the coastal zone. Biogeochemistry 66, 3–33.
Daby, D., Turner, J., Jago, C., 2002. Microbial and nutrient pollution of coastal bathing waters in Mauritius. Environ. Int. 27, 555–566.
Dubinsky, E.A., Butkus, S.R., Andersen, G.L., 2016. Microbial source tracking in impaired watersheds using PhyloChip and machine-learning classification. Water Res. 105, 56–64.
EU, 2006. Directive 2006/7/EC of the European Parliament and of the Council of 15 February 2006 concerning the management of bathing water quality and repealing Directive 76/160/EEC.
Faith, D.P., 2007. The role of the phylogenetic diversity measure, PD, in bio-informatics: getting the definition right. Evol. Bioinform. Online 2, 277–283.
Farkas, K., Adriaenssens, E.M., Walker, D.I., McDonald, J.E., Malham, S.K., Jones, D.L., 2019. Critical evaluation of CrAssphage as a molecular marker for human-derived wastewater contamination in the aquatic environment. Food Environ. Virol. 11, 113–119.
Farkas, M., Kaszab, E., Radó, J., Háhn, J., Tóth, G., Harkai, P., et al., 2020. Planktonic and benthic bacterial communities of the largest central European shallow lake, Lake Balaton and its main inflow Zala River. Curr. Microbiol. 77, 4016–4028.
Ferreira, S., Queiroz, J.A., Oleastro, M., Domingues, F.C., 2016. Insights in the pathogenesis and resistance of Arcobacter: a review. Crit. Rev. Microbiol. 42, 364–383.
Figueras, M.J., Levican, A., Pujol, I., Ballester, F., Rabada Quilez, M.J., Gomez-Bertomeu, F., 2014. A severe case of persistent diarrhoea associated with Arcobacter cryaerophilus but attributed to Campylobacter sp. and a review of the clinical incidence of Arcobacter spp. New Microbes New Infect. 2, 31–37.
Fisher, J.C., Levican, A., Figueras, M.J., McLellan, S.L., 2014. Population dynamics and ecology of Arcobacter in sewage. Front. Microbiol. 5, 525.
Flombaum, P., Gallegos, J.L., Gordillo, R.A., Rincón, J., Zabala, L.L., Jiao, N., et al., 2013. Present and future global distributions of the marine Cyanobacteria Prochlorococcus and Synechococcus. Proc. Natl. Acad. Sci. 110, 9824–9829.
Giebel, H.-A., Kalhoefer, D., Lemke, A., Thole, S., Gahl-Janssen, R., Simon, M., et al., 2011. Distribution of Roseobacter RCA and SAR11 lineages in the North Sea and characteristics of an abundant RCA isolate. ISME J. 5, 8–19.
Giebel, H.-A., Wolterink, M., Brinkhoff, T., Simon, M., 2019. Complementary energy acquisition via aerobic anoxygenic photosynthesis and carbon monoxide oxidation by Planktomarina temperata of the Roseobacter group. FEMS Microbiol. Ecol. 95.
Griffin, D.W., Banks, K., Gregg, K., Shedler, S., Walker, B.K., 2020. Antibiotic resistance in marine microbial communities proximal to a Florida sewage outfall system. Antibiotics 9.
Hagedorn, C., Blanch, A.R., Harwood, V.J., 2011. Microbial Source Tracking: Methods, Applications, And Case Studies. Springer Science & Business Media.
Hägglund, M., Bäckman, S., Macellaro, A., Lindgren, P., Borgmästars, E., Jacobsson, K., et al., 2018. Accounting for bacterial overlap between raw water communities and contaminating sources improves the accuracy of signature-based microbial source tracking. Front. Microbiol. 9.
Henry, R., Schang, C., Coutts, S., Kolotelo, P., Prosser, T., Crosbie, N., et al., 2016. Into the deep: evaluation of SourceTracker for assessment of faecal contamination of coastal waters. Water Res. 93, 242–253.
Ibekwe, A.M., Watt, P.M., Grieve, C.M., Sharma, V.K., Lyons, S.R., 2002. Multiplex fluorogenic real-time PCR for detection and quantification of Escherichia coli O157:H7 in dairy wastewater wetlands. Appl. Environ. Microbiol. 68, 4853–4862.
Jurelevicius, D., Cotta, S.R., Montezzi, L.F., Dias, A.C.F., Mason, O.U., Picão, R.C., et al., 2021. Enrichment of potential pathogens in marine microbiomes with different degrees of anthropogenic activity. Environ. Pollut. 268, 115757.

Karkman, A., Pärnänen, K., Larsson, D.G.J., 2019. Fecal pollution can explain antibiotic resistance gene abundances in anthropogenically impacted environments. Nat. Commun. 10, 80.
Khan, S.T., Horiba, Y., Yamamoto, M., Hiraishi, A., 2002. Members of the family Comamonadaceae as primary poly(3-hydroxybutyrate-co-3-hydroxyvalerate)-degrading denitrifiers in activated sludge as revealed by a polyphasic approach. Appl. Environ. Microbiol. 68, 3206–3214.
Knights, D., Kuczynski, J., Charlson, E.S., Zaneveld, J., Mozer, M.C., Collman, R.G., et al., 2011a. Bayesian community-wide culture-independent microbial source tracking. Nat. Methods 8, 761–763.
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, 2842.
Kristensen, J.M., Nierychlo, M., Albertsen, M., Nielsen, P.H., 2020. Bacteria from the genus Arcobacter are abundant in effluent from wastewater treatment plants. Appl. Environ. Microbiol. 86.
Levican, A., Collado, L., Figueras, M.J., 2016. The use of two culturing methods in parallel reveals a high prevalence and diversity of Arcobacter spp. in a wastewater treatment plant. Biomed. Res. Int. 2016, 8132058.
Lozupone, C., Lladser, M.E., Knights, D., Stombaugh, J., Knight, R., 2011. UniFrac: an effective distance metric for microbial community comparison. ISME J. 5, 169–172.
Lu, X., Zhang, X.X., Wang, Z., Huang, K., Wang, Y., Liang, W., et al., 2015. Bacterial pathogens and community composition in advanced sewage treatment systems revealed by metagenomics analysis based on high-throughput sequencing. PLoS One 10, e0125549.
Luijendijk, E., Gleeson, T., Moosdorf, N., 2020. Fresh groundwater discharge insignificant for the world's oceans but important for coastal ecosystems. Nat. Commun. 11, 1260.
Mathai, P.P., Staley, C., Sadowsky, M.J., 2020. Sequence-enabled community-based microbial source tracking in surface waters using machine learning classification: a review. J. Microbiol. Methods 177, 106050.
McGhee, J.J., Rawson, N., Bailey, B.A., Fernandez-Guerra, A., Sisk-Hackworth, L., Kelley, S.T., 2020. Meta-SourceTracker: application of Bayesian source tracking to shotgun metagenomics. PeerJ 8, e8783.
McIlroy, S.J., Kirkegaard, R.H., McIlroy, B., Nierychlo, M., Kristensen, J.M., Karst, S.M., 2017. MiDAS 2.0: an ecosystem-specific taxonomy and online database for the organisms of wastewater treatment systems expanded for anaerobic digester groups. Database 2017, bax016.
McMurdie, P.J., Holmes, S., 2013. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One 8, e61217.
Mercer, H.D., Pocurull, D., Gaines, S., Wilson, S., Bennett, J.V., 1971. Characteristics of antimicrobial resistance of Escherichia coli from animals: relationship to veterinary and management uses of antimicrobial agents. Appl. Microbiol. 22, 700–705.
Merga, J.Y., Royden, A., Pandey, A.K., Williams, N.J., 2014. Arcobacter spp. isolated from untreated domestic effluent. Lett. Appl. Microbiol. 59, 122–126.
Miljøverndepartementet, 2003. In: Miljøverndepartementet (Ed.), Forskrift om vern av Jærstrendene landskapsvernområde med biotopfredingar og naturminne i Randaberg, Sola, Klepp og Hå kommunar, Rogaland.
Nadimpalli, M.L., Chan, C.W., Doron, S., 2021. Antibiotic resistance: a call to action to prevent the next epidemic of inequality. Nat. Med. 27, 187–188.
Nguyen, P.T., Juarez, O., Restaino, L., 2020. A new method for detection of Arcobacter butzleri, Arcobacter cryaerophilus, and Arcobacter skirrowii using a novel chromogenic agar. J. Food Prot. 84, 160–168.
Nuy, J.K., Hoetzinger, M., Hahn, M.W., Beisser, D., Boenigk, J., 2020. Ecological differentiation in two major freshwater bacterial taxa along environmental gradients. Front. Microbiol. 11.
Oh, H.-M., Kwon Kae, K., Kang, I., Kang Sung, G., Lee, J.-H., Kim, S.-J., et al., 2010. Complete genome sequence of "Candidatus puniceispirillum marinum" IMCC1322, a representative of the SAR116 clade in the Alphaproteobacteria. J. Bacteriol. 192, 3240–3241.
Owili, M.A., 2003. Assessment of impact of sewage effluents on coastal water quality in Hafnarfjordur, Iceland. The United Nations Fishery Training Program, Final Report.
Prestinaci, F., Pezzotti, P., Pantosti, A., 2015. Antimicrobial resistance: a global multifaceted phenomenon. Pathogens Glob. Health 109, 309–318.
Prüss, A., 1998. Review of epidemiological studies on health effects from exposure to recreational water. Int. J. Epidemiol. 27, 1–9.
Roguet, A., Eren, A.M., Newton, R.J., McLellan, S.L., 2018. Fecal source identification using random forest. Microbiome 6, 185.
Sciortino, S., Arculeo, P., Alio, V., Cardamone, C., Nicastro, L., Arculeo, M., 2021. Occurrence and antimicrobial resistance of Arcobacter spp. recovered from aquatic environments. Antibiotics (Basel, Switzerland) 288, 10.
Staley, C., Kaiser, T., Lobos, A., Ahmed, W., Harwood, V.J., Brown, C.M., et al., 2018. Application of SourceTracker for accurate identification of fecal pollution in recreational freshwater: a double-blinded study. Environ. Sci. Technol. 52, 4207–4217.
Standard N, 1990. Vannundersøkelse - Termotolerante koliforme bakterier og presumptiv E. coli - Membranfiltermetode. Standard Norway.
Statens Helsetilsyn, 1994. Vannkvalitetsnormer for friluftsbad.
Steinbacher, S.D., Savio, D., Demeter, K., Karl, M., Kandler, W., Kirschner, A.K.T., et al., 2021. Genetic microbial faecal source tracking: rising technology to support future water quality testing and safety management. Österreichische Wasser- und Abfallwirtschaft 73, 468–481.
Tomova, A., Ivanova, L., Buschmann, A.H., Rioseco, M.L., Kalsi, R.K., Godfrey, H.P., et al., 2015. Antimicrobial resistance genes in marine bacteria and human uropathogenic Escherichia coli from a region of intensive aquaculture. Environ. Microbiol. Rep. 7, 803–809.
Unno, T., Staley, C., Brown, C.M., Han, D., Sadowsky, M.J., Hur, H.-G., 2018. Fecal pollution: new trends and challenges in microbial source tracking using next-generation sequencing. Environ. Microbiol. 20, 3132–3140.
Vadde, K.K., McCarthy, A.J., Rong, R., Sekar, R., 2019. Quantification of microbial source tracking and pathogenic bacterial markers in water and sediments of Tiaoxi River (Taihu Watershed). Front. Microbiol. 10.
Wani, G.A., Khan, M.A., Dar, M.A., Shah, M.A., Reshi, Z.A., 2021. Next generation high throughput sequencing to assess microbial communities: an application based on water quality. Bull. Environ. Contam. Toxicol. 106, 727–733.
WHO, 2003. World Health Organization - Guidelines for Safe Recreational Water Geneva, Switzerland.
Wybo, I., Breynaert, J., Lauwers, S., Lindenburg, F., Houf, K., 2004. Isolation of Arcobacter skirrowii from a patient with chronic diarrhea. J. Clin. Microbiol. 42, 1851–1852.
Yan, S., Fuchs, B.M., Lenk, S., Harder, J., Wulf, J., Jiao, N.Z., et al., 2009. Biogeography and phylogeny of the NOR5/OM60 clade of Gammaproteobacteria. Syst. Appl. Microbiol. 32, 124–139.