ARTICLE OPEN

Faecal pollution source tracking in the holy Bagmati River by portable 16S rRNA gene sequencing

Kalyan Pantha^{1,8}, Kishor Acharya^{2,8}✉, Sanjeeb Mohapatra^{3,8}, Santosh Khanal⁴, Niroj Amatya⁵, Carolina Ospina-Betancourth², Giacomo Butte^{1,2}, Suman Dhun Shrestha⁶, Prajwal Rajbhandari⁷ and David Werner²

A suitcase laboratory was used for 16S rRNA amplicon sequencing to assess microbial water quality in the holy Bagmati River, Kathmandu, Nepal. SourceTracker analysis and Volcano plots revealed that microbial communities in the downstream part of the river were mainly contributed by untreated sewage. Seasonal variability in the sewage microbiome was reflected in the downstream river water quality. The bacterial genera *Acidovorax*, *Geobacillus* and *Caulobacter* predominated in the upstream sites, while genera containing putative human pathogens and gut bacteria, such as *Clostridium*, *Prevotella*, *Arcobacter*, *Lactobacillus*, *Enterococcus* and *Streptococcus* become prominent in the downstream sites. Marker gene qPCR assays for total bacteria, total coliforms, Human *E. coli*, *Arcobacter butzleri* and *Vibrio cholerae* confirmed the sequencing data trends. Even though basic sanitation provision is nowadays near universal in Nepal, our findings show how inadequate wastewater management may turn an urban river into an open sewer, which poses a public health risk.

npj Clean Water (2021)4:12; <https://doi.org/10.1038/s41545-021-00099-1>

INTRODUCTION

Bagmati is the principal river of the Bagmati River Basin, one of the major basins of Nepal. It flows through the Kathmandu Valley, which includes Kathmandu, Lalitpur and Bhaktapur. The river basin is mainly fed by natural springs and monsoon rainfall and is considered as the source of Nepalese civilisation and urbanisation^{1,2}. The river flows along the famous Pashupatinath Temple, which holds great religious value for Hindu, and is also a World Heritage Site attracting more than a million tourists every year from across the globe³. Every day, hundreds and thousands of Hindu pilgrims perform bathing activities as a religious ritual in the Bagmati River. Water from the upstream part of this river is used for drinking, irrigation, livestock, industrial and domestic purposes by the people of the Kathmandu Valley^{3,4}. Every day, 30 million litres of water is sourced from the river and its tributaries for domestic purposes⁵. However, the river has been polluted due to rapid urbanisation, population growth, unmanaged sewage connection (from domestic, industrial and agricultural effluents) and solid waste disposal along the river banks^{4,6–9}. Approximately 21,000 kg of domestic sewage is discharged daily into the river from the Kathmandu Valley, including industrial BOD load discharges of 3151 kg per day¹. This pollution has not only affected the aesthetic value of the river, but has also caused high microbial and chemical loads, posing a considerable threat to public health and the ecosystem⁷. Furthermore, the river water pollution directly affects the surface, subsurface and groundwater quality, including the shallow wells in vicinity of the river, as they are directly interconnected^{5,10}. More than fifty percent of Kathmandu Valley residents rely on ground and subsurface water for various household activities including drinking¹¹. Various studies have reported frequent incidents of diarrhoeal diseases in Nepal, including in the Kathmandu Valley, which is mostly

caused by drinking water contaminated with human or animal faeces^{12–14}. Therefore, robust and frequent screening of river water microbiomes is essential to get an insight into faecal contamination sources and the pathways of waterborne disease transmission, which will ultimately help in designing effective interventions for safeguarding public and aquatic ecosystem health. Quantitative “Microbial Source Tracking” (MST) methods employing qPCR and next-generation sequencing (NGS) techniques are increasingly being used to identify the sources of faecal contamination in watersheds^{15,16}.

Most of the studies done on the Bagmati River have focused on spatiotemporal variation in organic matter, nutrients, major ions and trace elements^{4,6,17,18}. Only a few recent studies have addressed the microbial water quality and microbiome diversity of this river^{14,19,20}. Tandukar et al. quantified several human enteric viruses (i.e., adenoviruses, noroviruses and enteroviruses) and faecal indicator bacteria (i.e., total coliforms, *Escherichia coli*, *Enterococcus* spp and *Bacteroides*) present in the watersheds of the Bagmati River using real-time qPCR, and they also used fluorescence microscopy to quantify helminths, such as *Cryptosporidium* spp. and *Giardia* spp.²⁰. Likewise, Khanal et al. explored the helminths communities in the Bagmati River watersheds using Illumina NGS¹⁹. Shrestha et al. studied the Bagmati River microbiome and quantified *Arcobacter* using Illumina NGS and qPCR, respectively¹⁴. Although qPCR results can be obtained within 2 h, they provide limited insight into the composition of water microbial communities, as it can only identify the targeted pathogens that are specifically chosen for monitoring. In contrast, high-throughput molecular methods, such as Illumina NGS allow more comprehensive water microbiome characterisations, but there will be delay in the availability of data from centralised laboratory facilities, which can restrict rapid decision making²¹.

¹Group for Rural Infrastructure Development, Wise use House, Jwagal, Lalitpur, Nepal. ²School of Engineering, Newcastle University, Newcastle upon Tyne NE1 7RU, UK. ³Environmental Science and Engineering Department, Indian Institute of Technology Bombay, Mumbai MH-400076, India. ⁴Center for Pediatric Genomic Medicine, Children's Mercy Kansas City, Kansas City, MO 64110, USA. ⁵Department of Medical Microbiology, Nobel College, Pokhara University, Kathmandu, Nepal. ⁶School of Engineering, Kathmandu University, Kavre, Nepal. ⁷Department of Applied Microbiology and Food Technology, Research Institute for Bioscience and Biotechnology (RIBB), Kathmandu, Nepal. ⁸These authors contributed equally: Kalyan Pantha, Kishor Acharya, Sanjeeb Mohapatra. ✉email: kishor.acharya@ncl.ac.uk

This is especially a problem in resource-limited settings, like in developing countries²².

An exciting development for near-real-time microbial water quality surveying in developing countries is the recent release of the low-cost, field-deployable and memory-stick sized MinION sequencer from Oxford Nanopore Technologies (ONT) Ltd. Its portability offers the possibility to comprehensively and quickly survey microbial water quality in any kind of setting, including in resource-limited developing countries^{23–25}. In our previous work, we developed and validated a suitcase laboratory, which combined the MinION sequencer with other small equipment items to enable water microbiome characterisation by portable 16S rRNA gene amplicon sequencing^{24,25}. In the current study, we deployed this toolbox to survey the microbiomes of the Bagmati River for different reaches of the river and different seasons. Our previous work had focused on the method development and validation²⁴, and the assembly of small equipment items into a suitcase laboratory²⁵. The current study reports the innovative application of these tools in combination with SourceTracker (ST) and Volcano plot analysis to facilitate microbial water quality monitoring at the watershed-scale in a culturally, religiously and economically important river located in a developing country.

RESULTS

Microbial community analysis

An overall microbial community analysis is presented as PCA plots and a dendrogram in Fig. 1. The cluster analysis showed good agreement between sample replicates, which clustered most closely. WWTP influent (i.e., untreated sewage) collected in the post-monsoon season clustered most closely with water samples from S4, S5 and S6 collected at same time, while WWTP influent from the monsoon season clustered with water samples from S6 collected in the same season. The WWTP effluent from both monsoon and post-monsoon season clustered together. The PCA plot with data from all the sampling times (Fig. 1b) generally showed a separation of downstream and WWTP influent water samples from the upstream and WWTP effluent samples along principal component 1, with only a few exceptions. Genera mostly found in the human gut microbiome^{15,26} like *Streptococcus*, *Trichococcus*, *Lactobacillus*, *Enterococcus*, *Prevotella* and *Arcobacter*, were highly prevalent in downstream and WWTP influent water

samples, which separated these samples from the upstream water samples in the PCA. Among the three factors analysed (i.e., location, sampling time and water sample types), locations and sampling time had a significant effect on the similarity of the samples in the ANOSIM, although with relatively low *R* values (ANOSIM; Location: *R* = 0.29, *p* value = 0.001 and Sampling time: *R* = 0.16, *p* value = 0.01). ANOSIM further indicated no statistically significant differences between the microbial communities in water from locations S4, S5 and S6 and the wastewater influent (ANOSIM; (1) S4 and Inf: *R* = 0.0309, *p* value = 0.357; (2) S5 and Inf: *R* = 0.0617, *p* value = 0.369 and (3) S6 and Inf: *R* = 0.0123, *p* value = 0.3690).

An interesting picture emerged when a separate PCA (Fig. 1c, d) and cluster analysis (Supplementary Fig. 1) was conducted for water samples from the monsoon and the post-monsoon season. In both seasons, there were substantial, but seasonally distinct, contributions of genera found in the human gut microbiome to the variance among water samples along principal component 1: in the monsoon season *Arcobacter*, *Aeromonas*, *Streptococcus* and *Prevotella* had significant PC1 loadings; in the post-monsoon season *Enterococcus*, *Acinetobacter*, *Streptococcus* and *Trichococcus* had significant PC1 loadings. Separation of wastewater treatment plant effluent (WWTP Effluent) samples along PC1 away from the WWTP influent (WWTP Influent) samples in both sampling events signified the benefits of wastewater treatment, because human gut-associated genera became less predominant in treated wastewater microbiomes, as expected²⁷. Accordingly, there was a clear separation of the most upstream water samples from the most downstream water samples along PC1 in both events, with the downstream water samples becoming more similar to WWTP Influent (Fig. 1c, d). Evidently, as the Bagmati River flowed into more densely populated areas, the characteristics of its water microbiome changed from a composition more similar to treated, to a composition more similar to untreated urban sewage, but the composition of the urban sewage was variable for the monsoon and post-monsoon season.

Abundance of human gut and putative pathogenic bacteria in the water microbiomes

A more detailed breakdown of the microbial community composition in the Bagmati River for the monsoon and post-monsoon season is reported in Table 1, and Supplementary Tables

Fig. 1 Cluster and PCA analysis at rank genus for 16S rRNA gene sequencing reads. **a** Cluster analysis [all seasons], PCA plot **b** all seasons, **c** monsoon [June 2019] and **d** post monsoon [August 2019]. Arrows in the PCA plots indicate the ten variables with the highest loadings (vector lengths) in the PC1 and 2 space.

Table 1. Average total percentage of putative human gut bacteria and putative human pathogenic bacteria at both genus and species level in the Bagmati River, WWTP influent and WWTP effluent during the post monsoon (Post monsoon, August 2019).

	S1_P-mon	S2_P-mon	S3_P-mon	S4_P-mon	S5_P-mon	S6_P-mon	Inf_P-mon	Eff_P-mon
Human gut genera (%)	3.290 (0.007)	2.432 (0.117)	6.905 (0.030)	36.198 (0.509)	22.914 (0.237)	30.855 (0.381)	19.059 (0.351)	1.003 (0.029)
Human putative pathogen genera (%)	5.038 (0.078)	2.848 (0.110)	5.669 (0.022)	20.468 (0.266)	11.005 (0.075)	11.377 (0.102)	13.312 (0.244)	2.652 (0.032)
Human gut species (%)	0.521 (0.035)	0.321 (0.001)	1.600 (0.048)	11.367 (0.085)	8.262 (0.034)	13.899 (0.218)	5.011 (0.072)	0.128 (0.002)
Human putative pathogen species (%)	0.375 (0.016)	0.161 (0.005)	0.352 (0.037)	2.419 (0.043)	1.005 (0.016)	1.113 (0.027)	3.898 (0.012)	0.681 (0.016)

Inf and Eff, respectively, indicated WWTP influent and WWTP effluent, while P-mon indicates Post-monsoon (August 2019). The number in the parenthesis is the standard deviation of the duplicate analysis. Data from the other sampling events are available as supporting information, Tables S1 and S2.

1 and 2, which compare the total percentage relative abundance of putative human gut²⁸ and pathogenic²⁹ bacteria at genus and species level for different sampling sites in the Bagmati River, and the WWTP influent and effluent (Refer to Supplementary Tables 3–5 for more detailed lists of bacteria). Based on our previous findings²⁴, species identities are not always reliable due to the limited read accuracy of the MinION sequencing reads, but the overall trends are nonetheless indicative of changes in microbial composition. For all sampling events, the water collected at the most upstream site S1 and S2 showed the lowest relative abundance for both human gut and putative pathogenic bacteria, whereas the highest relative abundance was observed in the water collected at the most downstream sites S4–S6 (Table 1, Supplementary Tables 1 and 2). The microbial water quality of water samples collected at site S1 can be considered as baseline data, as this watershed is distant from the densely populated Kathmandu Valley and has the minimal influence of human and urbanisation activities. Figure 2, and additional figures in Supplementary Information (Supplementary Fig. 2–9) show how the abundance of human gut and putative pathogenic genera changed in space and time along the Bagmati River. As the river flowed downstream, the abundance of some of these groups of bacteria increased, and the most drastic and significant increase was observed at the sites S4, S5 and S6 downstream of the Pashupatinath Temple as compared to site S1 (Two-sample *t* test, *p* value < 0.05) (Table 1, Supplementary Table 1, Supplementary Table 2, Supplementary Figs. 2, 3, 6 and 7). In comparison with the WWTP influent and effluent samples, the relative abundance of human putative pathogenic species in the water at sites S1 and S2 was significantly lower than in WWTP effluent (Two-sample *t* test, *p* value < 0.05), [Table 1 and Supplementary Table 1]. However, the pathogen load in the WWTP effluent can always be further reduced by tertiary treatment processes, such as chlorination before it is discharged into the receiving water bodies³⁰. Downstream of Pashupatinath Temple, the abundance of some of the human gut and putative pathogenic genera were similar or even greater than observed in WWTP influent (Table 1, Supplementary Table 1, Supplementary Figs. 4, 5, 8 and 9). These more detailed results supported the observations from the PCA and showed that the microbial water quality of the Bagmati River deteriorated significantly downstream of Pashupatinath Temple as a result of faecal pollution, which was most likely due to the discharge of untreated sewage into the river, a significant public health hazard. Similar results have been reported in previous studies, although for a more limited number of sample sites^{14,20}.

Effect of untreated sewage discharge in the downstream receiving river

ST analysis was performed to better understand the impact of three sources on the microbial water quality of the Bagmati River: (1) WWTP influent as a proxy for untreated sewage discharges, (2)

WWTP effluent and (3) the most upstream river water, which provides a baseline. Each source was distinct in both seasons based on PCA analysis, showing that the leave-one-out source class prediction provided a reasonable reflection of sources. This then allowed us to proportionate source influences in sinks for two different seasonal events (i.e., monsoon [June 2019] and post monsoon [August 2019]; no wastewater treatment data was available for the July 2019). In the monsoon season, sequences of microbial communities in location S2 were mostly sourced from the upstream river ($63 \pm 1.19\%$), while WWTP effluent contributed $16.46 \pm 1.18\%$ (Fig. 3a). In contrast, at S3, S4 and S5, both the upstream river ($35.48 \pm 0.2\%$ for S3; $30.27 \pm 0.4\%$ for S4; and $32.73 \pm 1.12\%$ for S5) and WWTP influent [i.e., untreated sewage] ($26.90 \pm 0.75\%$ for S3; $48.95 \pm 0.12\%$ for S4; and $23.23 \pm 0.17\%$ for S5) had significant contributions (Fig. 3a). The sequences of microbial communities in location S6 were mostly dominated by sequences from the untreated sewage ($72.70 \pm 0.34\%$), and to lesser extent by the river upstream (S1) and WWTP effluent (Fig. 3a). Interestingly, at all five sites, unknown sources showed an influence (13.33–18.28%). Finally, in the post-monsoon season, the sequences of microbial communities in sites S4 ($74.66 \pm 0.21\%$), S5 ($62.73 \pm 0.29\%$) and S6 ($83.71 \pm 0.19\%$) were largely constituted by the sequences found in WWTP influent, while at sites S2 and S3, the upstream river (S1) contributed $77.91 \pm 0.23\%$ and $32.65 \pm 0.09\%$, respectively (Fig. 3b). Interestingly, the unknown source contributed 50% of the sequences in location S3, while in other sites the contribution from an unknown source ranged between 13.66 and 19.7%. Based on the recent study from ENPHO, in the Kathmandu valley, 76.63% of faecal waste is discharged into the water bodies without treatment³¹. These findings suggest that the whole watersheds network can be divided into two main microbial source communities (river upstream and WWTP influent), and their seasonally variable composition and blending explained where, how and why faecal contamination influenced the microbiomes of the Bagmati River. In addition, the water flowing in the downstream part of the Bagmati River had mostly the characteristics of untreated sewage discharged directly into the river.

Faecal markers in the river quantified with qPCR

The qPCR results on different marker genes, including, faecal marker genes for the water samples collected from different sites at three different sampling events, including WWTP influent and effluent are presented in Fig. 4. These results were from the further analysis of DNA samples from Nepal at Newcastle University in the UK. All the pathogen or faecal marker genes were detected in the WWTP influent and effluent, and the three sites downstream of Pashupatinath Temple (i.e., S4, S5 and S6) for all sampling events. The concentration of these genes in these three sites was significantly higher than in upstream sites and WWTP effluent (Two-sample *t* test, *p* value < 0.05), and comparable to what we observed in the WWTP influent, except for the

Fig. 2 Volcano plots comparing the abundance of specific genera in S6 relative to S1. Comparison of Human gut genera²⁸ for **a** monsoon and **c** post-monsoon seasons, human pathogenic genera²⁹ for **b** monsoon and **d** post-monsoon seasons. Green and blue dots indicate the genera in the S6 site that are at least twofold (x-axis) higher and lower, respectively, than detected in the S1 site, and high statistical significance ($-\log_{10} p$ value, y-axis). The dashed blue line shows where the p value = 0.05, with points above the line having the p value < 0.05 and points below the line having $p > 0.05$. The relative abundance of human gut and pathogenic genera on the left side of the solid blue line decreased, while on the right side it increased as compared to abundance in the site S1.

ciaB gene. In the other three sites, which were located upstream of Pashupatinath Temple, the *ciaB* gene was always detected and quantified in the water collected from site S3, but only detected in one water sample from site S2 in the post-monsoon season (August 2019). Except for the July 2019 water samples from site S1, *E. coli* from human origin were detected in water samples collected from all locations. Whereas the *ompW* gene, which is a specific marker for *Vibrio cholerae*, was only detected at sites S2 and S3 in the monsoon water samples (June 2019) (Fig. 4). Furthermore, a boxplot presented in Supplementary Fig. 10 revealed the range of concentration of marker genes in the studied seasons for the different sampling events. The observed trends for all the marker genes suggested that the concentration of all the studied marker genes increased as the river flowed downstream. Marker genes for total bacteria, Human *E. coli* and *Arcobacter butzleri* were significantly reduced in the WWTP effluent as compared to WWTP influent (Two-sample *t* test, p value < 0.05) (Supplementary Fig. 10).

Cross-comparison and validation of MinION data with other microbiology data

Our previous analysis of a MOCK community revealed that MinION sequenced reads might result in false positive outcomes,

especially at the species level, which indicated the need for validation with alternative methods, such as qPCR²⁴. To validate the MinION's result in this study we performed qPCR on extracted DNA for different marker genes at Newcastle University after returning to UK as presented in Supplementary Table 6. The marker genes used to quantify *Vibrio cholerae*, *Arcobacter butzleri* and Human *E. coli* were virulence genes and specifically present only in those microorganisms. Supplementary Fig. 11 shows the extent of correlation between different microbial water quality indicators determined by qPCR and NGS approaches. A significant Spearman rank correlation was observed between *Arcobacter butzleri*, total coliform and human *E. coli* quantified with qPCR, and other putative faecal indicator bacteria screened with the MinION (Supplementary Fig. 11). These correlations were well aligned with our earlier results^{24,25}, and substantiated the usefulness of the portable MinION platform for bacterial hazard screening in water samples via sequencing of 16S rRNA gene amplicons.

Basic microbiology for faecal indicator bacteria

Supplementary Table 7 shows the preliminary MPN index for coliform bacteria in different water samples collected from six sites in the monsoon and post-monsoon season. Coliform bacteria were present in water collected from all the sites. The confirmative

Fig. 3 Relative contribution of various sources to the sinks in different seasons. **a** Monsoon (June 2019) and **b** post monsoon (August 2019).

Fig. 4 qPCR results for marker genes in water and wastewater samples. Water for **a** monsoon (June 2019), **b** monsoon (July 2019) and **c** post monsoon (August 2019), and wastewater for **d** monsoon (June 2019) and post monsoon (August 2019). Data points are an average of duplicate samples and error bars indicate the standard deviation.

test for faecal indicator organisms in Supplementary Table 8 showed the growth of characteristic greenish metallic shine bacterial colonies on EMB agar at 44.5 °C in all the samples, which previously showed the presence of coliform bacteria except in the most upstream water samples from S1 and S2 collected in the monsoon season (June 2019). This further confirmed the presence of faecal indicator organism *E. coli* at the downstream sampling sites and provides further evidence for the faecal pollution at these sites.

DISCUSSION

In this study, we demonstrated the suitability of the portable MinION NGS platform for comprehensive water quality monitoring and faecal pollution source tracking in a watershed via an innovative combination of 16S rRNA amplicon sequencing, ST and Volcano plot analysis. This demonstrated the practical feasibility of microbial community analysis and bacterial hazard screening with this portable laboratory in a developing country, and is to the best of our knowledge the first report of water microbiomes sequenced in Nepal. Several 16S rRNA gene sequences retrieved from water samples were associated with bacterial groups with public health relevance, and qPCR assays performed later in the UK to target virulence genes confirmed the *Vibrio cholerae*, *Arcobacter butzleri* and human faeces derived *E. coli* hazards.

Putative faecal indicator bacteria were frequently detected in high abundances in the river water at sites downstream of Pashupatinath Temple, which is a World Heritage Site attracting more than one million tourists each year, and many of them take a bath in the river as a ritual. There was a significant increase of putative faecal indicator bacteria from the most upstream to the most downstream site, demonstrating the deleterious impact of poorly managed sanitation in the densely settled urban areas of the Kathmandu Valley on the Bagmati River water quality. During the post-monsoon season, the PCA and ST analysis elucidated that the microbial communities in the more downstream sites most strongly resembled the microbial communities of WWTP influent, which is untreated urban sewage. At the most downstream site (S6) of the Bagmati River in the Kathmandu Valley, the microbial communities of the water were always most similar to WWTP influent communities during both the monsoon and post-monsoon season. This must be attributed to the practice of discharging untreated sewage into the river, as soon as it passes the Pashupatinath Temple. These findings for putative faecal indicator bacteria in the NGS libraries were well aligned with the qPCR results for faecal marker genes. The lowest abundance of these microorganism was seen at S1, which is commonly used as a source of drinking water for Kathmandu City, due to little human interference^{14,20}. Poor microbial river water quality from site S4 in the downstream could be attributed to unsafe sanitation, and open defecation along in the river as has been reported elsewhere^{14,20,32}. Additional pollution contributed by the river tributaries also cannot be ignored³³. Overall, these results further verify deterioration of the river water quality as the river flows through Kathmandu city, which is of significant concern due to the possible exposure pathway for waterborne diseases. As the occurrence of enteric viruses strongly correlates with the faecal indicators²⁰, additional studies may be conducted to identify pathogenic human viruses in the river water.

People taking a bath in the Bagmati river (as a routine activity or for a holy dip) between sites S3 and S4 are vulnerable to faecally polluted water and associated waterborne diseases, such as typhoid, cholera, paratyphoid fever, dysentery, jaundice and amoebiasis^{34,35}. In addition, people residing nearby the Bagmati River in urban settings are expected to be exposed indirectly through consumption of water from shallow wells and vegetables grown with polluted water²⁴. Many research findings have already highlighted river pollution as one of the root cause of

groundwater pollution, which includes both pollution with anthropogenic chemicals^{36–38} and potential pathogenic microbes³⁹. Chronic exposure to such compounds and microbes may have severe consequences, including diseases conditions, multidrug resistance and endocrine disruption^{40,41}. In addition, consumption of vegetables grown through drip irrigation in this polluted watershed may contribute an additional disease risk as has been reported elsewhere^{42,43}. This study revealed how WWTP effluent had substantially better microbial water quality than the downstream Bagmati River, which demonstrates the urgent need for a less fragmented sanitation provision with proper sewage treatment. For instance, appropriate sewage collection strategies and subsequent treatment of sewage can be an appropriate intervention to prevent further degradation of river water quality.

At all the studied areas, the bacterial community structure, including the abundance of putative faecal indicator bacteria in the monsoon season was distinct and more diverse than in the post-monsoon season. Certain groups of faecal indicator bacteria prevailed in significantly higher abundance in the monsoon season, while separate groups were prevalent in the post-monsoon season. For instance, *Enterococcus*, *Streptococcus*, *Bacillus*, *Clostridium* and some genera from *Enterobacteriaceae* were significantly higher in downstream sites than in the WWTP influent during the post-monsoon season. Leight et al., detected higher abundance of organisms, such as *Streptococcus*, *Clostridium* and *Bacteroides* and some genera from the *Enterobacteriaceae* in the studied river during the wet season, and further suggested that some of these might be allochthonous organisms that might get washed in from the land²⁶. Such temporal variation in the abundance of the putative faecal indicator bacteria highlights the necessity of taking into consideration source variability when selecting the most appropriate indicator for MST. In this study of the Bagmati River, species from the genera *Enterococcus* and *Streptococcus* appeared to be suitable indicators for faecal pollution during the post-monsoon season, while species from the genera *Prevotella* and *Arcobacter* were better indicators during the monsoon season. Therefore, it is imperative to study pollution sources and river water in conjuncture, as was done in this study, to account for seasonal variability of indicator bacteria in the sources.

Urban water pollution, mostly due to the discharge of untreated sewage and agricultural or storm-water runoff into waterways from land, are a growing problem in developing countries⁴⁴. The observed microbial water quality at downstream sites in the Bagmati River illustrated this problem in a socioeconomically important watershed. This further implies that current sanitation provision in Kathmandu is insufficient for the protection of economies, public health and ecosystems, and thus implementation of safely managed sanitation system is urgent. In the absence of adequate wastewater treatment facilities and safe faecal sludge management, rivers become open sewers that pose a significant exposure risk to the public, not dissimilar to open defecation. The ongoing Asian Development Bank project [Kathmandu Valley Wastewater Management Project (ADB Loan No. 3000-NEP)] in the Kathmandu Valley, Nepal, has as its main objective to modernise and expand the wastewater network and treatment facilities in the Kathmandu Valley, which shall improve and help in managing the wastewater services in the Kathmandu Valley in the future⁴⁵. The project aims to rehabilitate the sewerage network and further expand the capacity of WWTPs to 90.5 million litres from around 16 million litres per day by constructing five new treatment plants⁴⁵. The data presented here provide the current status quo for water quality in the Bagmati River, which will be useful for documenting the environmental benefits of the new wastewater treatment infrastructures in the future.

In this study, we used a low-cost, portable sequencer, the MinION that was also used in our previous study²⁵, to characterise in Nepal the microbiome of the Bagmati River from its source to

the point where it leaves the Kathmandu Valley. Portable sequencing allowed to simultaneously screen, identify and monitor multiple indicator bacteria for faecal pollution, and produce multivariate data for faecal pollution source tracking, which is particularly valuable when there is temporal variability in the source composition, as was illustrated in this study. This technology can enable on-site monitoring, and make a vital contribution to the more rapid characterisation of water quality, which underpins the design of adequate treatment technologies and regulatory frameworks for mitigating the associated problems^{25,46}. Furthermore, such portable methods to comprehensively screen the microbiome of a biological system will play a major role in monitoring wastewater treatment systems and the protection of both ecosystem and human health. Overall, the application of portable sequencing technologies for faecal pollution source tracking will help achieve the UN Sustainable Development Goal (SDG) 6; clean water and sanitation⁴⁷. For example, faecal pollution source tracking is essential for SDG 6.3, which aims to protect both ecosystem and human health by eliminating, minimising and significantly reducing different streams of pollution into water bodies⁴⁷.

MATERIAL AND METHODS

Study site and sample collection

The study investigated a 25 km stretch of the Bagmati River ranging from Sundarijal (Upstream) to Chovar (Downstream). The Bagmati catchment drains an area of 595.4 km² with Chovar as its outlet. Several tributaries of different order feed the river, including Manahara khola, Hanumante khola, Dhobi khola, Tukucha khola, Bishnumati khola and Nakkhu khola. The altitude of the basin area varies from 1178 to 2723 m above the mean sea level.

Eight sampling sites as shown in Fig. 5 were selected based on the population density. The upstream region, which was located above the holy Pashupatinath Temple, contained sampling sites S1, S2 and S3, while the downstream region included sampling sites S4, S5 and S6 (Supplementary Table 9 in supporting information). The sites S4 and S5 were in the areas where the river flows through the settlements with the highest population density. The most downstream sampling site S6 (Chovar), was in the area where the river flows beyond the Valley into a region with a lower population density. In addition, the WWTP influent and effluent of the Guheshwori WWTP, located in between the sampling locations S3 and S4, were also sampled. From each

Fig. 5 Locations of the sampling sites in the Kathmandu Valley, Nepal. Six different sites in the Bagmati River were sampled, two times during the monsoon and once in the post monsoon, while WWTP influent and effluent were sampled only once in monsoon and post monsoon. “Base map and data from OpenStreetMap and OpenStreetMap Foundation”

sampling site, 2 L of water were collected in sterile bottles during the monsoon (7th June and 20th July, 2019) and post monsoon (30th August, 2019) season. The average weekly rainfall in the Kathmandu Valley was 42, 29.85 and 7.09 mm, respectively, in the week of June 7th, July 20th and August 30th. Daily rainfall data for the Kathmandu Valley during June, July and August are reported in Supplementary Table 10 in supporting information. Samples were transported immediately to the laboratory in cold boxes, containing ice cubes and packs, and processed within 2 h.

16S rRNA gene amplicon sequencing analysis of water quality with a suitcase laboratory

All the necessary portable equipment and consumables for the 16S rRNA gene amplicon sequencing analysis with the memory-stick sized MinION of ONT were packed into two check-in sized luggage items at Newcastle University and transported to Nepal by air and road. The molecular microbiology consumables including flow cells, PCR reagents and 16S rRNA gene amplicon sequencing kit required cold storage. Therefore, they were placed into a polystyrene box with the cool packs used by ONT to ship their flow cells to customers. This packaging could maintain the refrigerator temperature of about 4 °C for ~2 days. The portable laboratory²⁵ was set-up at a bench space provided by the Research Institute for Bioscience and Biotechnology, Kathmandu, Nepal.

Biomass for total DNA extraction from each sampling event was concentrated in duplicates by independently filtering 100 mL of water through a 0.22 µm membrane (Sartorius UK Limited, Surrey, UK), except for site S1, for which 250 mL were filtered, because a lower amount of biomass was anticipated. Membrane filters were folded aseptically, transferred individually in a 15 mL sterile falcon tubes, and stored at −20 °C until use. These filtrations were conducted in the CEMAT water laboratories (Kathmandu, Nepal).

Total DNA from the retained biomass was extracted using a PowerWater DNA Isolation Kit as per manufacturer's instruction (QIAGEN, Crawley, UK), while the isolated DNA was quantified using a Qubit dsDNA HS Assay Kit (Life Technologies, UK). The 16S rRNA gene amplicon library preparation, subsequent sequencing of the libraries, and data processing and statistical analysis were performed in accordance with the method described elsewhere²⁵.

Metadata collection

For context and validation of the sequencing data generated with the portable MinION toolbox in Nepal, the multiple tube method to assess the most probable number (MPN) of faecal indicator bacteria was performed for the monsoon (June 2019) and pre-monsoon season (August 2019) at CEMAT water laboratories (Kathmandu, Nepal), according to the procedure described elsewhere⁴⁸, and qPCR was performed at Newcastle University using surplus of the DNA extracted in Nepal to quantify different faecal marker genes. Briefly, the MPN index with a 95% confidence limit was worked out with the help of an MPN Chart. The presumptive test was performed using MacConkey broth. Further, a confirmatory test was conducted by inoculating one loopful of the sample from each presumptive positive tube on EMB agar and incubating it at 44.5 °C for 24 h. The greenish metallic shine colonies from EMB agar were then streaked on nutrient agar and incubated at 37 °C overnight. Standard IMViC tests were performed for isolated colonies from the nutrient agar for confirmation of *E. coli* and a further complete test was done on MacConkey broth⁴⁸.

qPCR assays for target genes were conducted on a Bio-Rad CFX 1000 system (Bio-Rad, Hercules, CA USA) by following the method described elsewhere²⁵. Briefly, primers mentioned in Supplementary Table 6 were used to quantify the target marker genes in a final reaction volume of 15 µL. The final reaction mix

consisted of 2 µL of template DNA, 7.5 µL of so Advanced Universal SYBR Green Supermix (Bio-Rad, Hercules, CA USA), 0.75 µL (500 nmol L^{−1}) of each forward and reverse primer, and 4 µL molecular grade H₂O (Invitrogen, Life Technologies, Paisley, UK) were mixed together. The qPCR programme for the quantification of each marker genes was 98 °C for 3 min (1×), then 98 °C for 15 s followed by primer annealing temperature (T_a) for 60 s (Supplementary Table 6) (40 cycles). At the end of each assay, the qPCR products were subjected to melt curve analysis with gradual increase of temperature (0.2 °C temperature increments every 10 s) from 65 to 95 °C. Standard curves were generated for every qPCR assay. All the samples, standards and negative controls (molecular grade water replaced the DNA template) were run in duplicates, while to minimise the effect of inhibitors on PCR reactions, the DNA samples were diluted to a working solution of 5 ng µL^{−1} using molecular grade water.

Data analysis

A correlation analysis (i.e., Spearman rank) between the water quality markers and indicators quantified, respectively, by qPCR and NGS approaches was performed in R-studio (Version; R version 3.6.2). This helped to validate and cross-compare the outcomes for different methods. Finally, comparison between the abundance of detected human gut and pathogenic genera for both monsoon (June 2019) and post monsoon (August 2019) in S1 (the most upstream and rural site) versus other sites were performed with a Volcano plot in R-studio. A similar comparison between the WWTP influent versus water from the other sites, including the WWTP effluent, were also performed with a Volcano plot in R-studio (Version; R version 3.6.2).

ST analysis as described elsewhere⁴⁹ was used to evaluate the relative contribution of microbial communities from different “sources” along the Bagmati river to the “sinks”, during both the monsoon (June 2019) and post-monsoon (August 2019) season. Sources consist of the most upstream site, S1 (*n* = 2), WWTP influent (*n* = 2) and WWTP effluent (*n* = 2). While the sinks consist of the five sites in the Bagmati River labelled as S2, S3, S4, S5 and S6. ST uses Gibb's sampling (Markov chain Monte Carlo-algorithm) to evaluate the relative contribution of sources, while all the undefined operational taxonomic units in the sinks will be assigned as from unknown sources. ST analysis was performed with a sequencing depth of 15,000 with 100 iterations, ten re-starts and the auto-tuning functionality.

DATA AVAILABILITY

Raw sequencing data from the 16S rRNA sequencing are registered on the NCBI biosample database with sequence read archive (SRA) accession number PRJNA690679. Additional data created during this research are openly available at <https://doi.org/10.25405/data.ncl.13160147.v1>. Please contact Newcastle Research Data Service at rdm@ncl.ac.uk for access instructions.

Received: 18 August 2020; Accepted: 12 January 2021;

Published online: 18 February 2021

REFERENCES

- Dahal, A., Khanal, M. & Ale, M. *Bagmati River Festival: Conservation of Degrading River* (Georgia Institute of Technology, 2011).
- Subedi, A., Poudel, P. & Acharya, T. *Temporal Shift of Bagmati River over 25 Years Using Landsat*, Vol. 42 (International archives of the photogrammetry, remote sensing and spatial information sciences, 2019).
- Bajracharya, S. B. et al. *Bagmati action plan (2009–2014)* (Government of Nepal, High Powered Committee for Integrated Development of the Bagmati Civilization and National Trust for Nature Conservation, 2009).
- Bhatt, M. P. et al. Chemistry of the heavily urbanized Bagmati River system in Kathmandu Valley, Nepal: export of organic matter, nutrients, major ions, silica, and metals. *Environ. Earth Sci.* **71**, 911–922 (2014).

5. Malla, R. et al. Physico-chemical and oxygen-hydrogen isotopic assessment of Bagmati and Bishnumati rivers and the shallow groundwater along the river corridors in Kathmandu Valley. *Nepal. J. Water Resour. Prot.* **7**, 1435 (2015).
6. Kannel, P. R., Lee, S. & Lee, Y.-S. Assessment of spatial-temporal patterns of surface and ground water qualities and factors influencing management strategy of groundwater system in an urban river corridor of Nepal. *J. Environ. Manag.* **86**, 595–604 (2008).
7. Deshar, B. D. Squatters problems along Bagmati Riverside in Nepal and its impact on environment and economy. *Int. J. Environ. Eng. Manag.* **4**, 127–142 (2013).
8. Gautam, R., Shrestha, J. K. & Shrestha, G. K. C. Assessment of river water intrusion at the periphery of Bagmati River in Kathmandu Valley. *Nepal J. Sci. Technol.* **14**, 137–146 (2013).
9. Mishra, B. K. et al. Assessment of Bagmati river pollution in Kathmandu Valley: Scenario-based modeling and analysis for sustainable urban development. *Sustainability Water Qual. Ecol.* **9**, 67–77 (2017).
10. Devkota, D. C. & Watanabe, K. Impact of solid waste on water quality of Bishnumati River and surrounding areas in Kathmandu. *Nepal J. Nepal. Geol. Soc.* **31**, 19–24 (2005).
11. Udmale, P. et al. The status of domestic water demand: supply deficit in the Kathmandu Valley, Nepal. *Water* **8**, 196 (2016).
12. Uprety, S. et al. The effect of the 2015 earthquake on the bacterial community compositions in water in Nepal. *Front. Microbiol.* **8**, 2380 (2017).
13. Uprety, S. et al. Assessment of microbial risks by characterization of *Escherichia coli* presence to analyze the public health risks from poor water quality in Nepal. *Int. J. Hyg. Environ. Health* **226**, 113484 (2020).
14. Shrestha, R. G. et al. Prevalence of Arcobacter and other pathogenic bacteria in river water in nepal. *Water* **11**, 1416 (2019).
15. Vadde, K. K. et al. Next-generation sequencing reveals fecal contamination and potentially pathogenic bacteria in a major inflow river of Taihu Lake. *Environ. Pollut.* **254**, 113108 (2019).
16. Green, H. et al. Microbial source-tracking reveals origins of fecal contamination in a recovering watershed. *Water* **11**, 2162 (2019).
17. Kannel, P. R. et al. Spatial-temporal variation and comparative assessment of water qualities of urban river system: a case study of the river Bagmati (Nepal). *Environ. Monit. Assess.* **129**, 433–459 (2007).
18. Bhatt, M. P. & McDowell, W. H. Evolution of chemistry along the Bagmati drainage network in Kathmandu Valley. *Water, Air, Soil Pollut.* **185**, 165–176 (2007).
19. Khanal, S., Gyawali, P. & Shrestha, B. Determining the Helminth Communities in Environmental Water of Kathmandu, Nepal by using NGS. In *Skill, Knowledge and Innovation Transfer to Nepal*, Vol. 77 (2014).
20. Tandukar, S. et al. Presence of human enteric viruses, protozoa, and indicators of pathogens in the Bagmati River, Nepal. *Pathogens* **7**, 38 (2018).
21. Peletz, R. et al. Why do water quality monitoring programs succeed or fail? A qualitative comparative analysis of regulated testing systems in sub-Saharan Africa. *Int. J. Hyg. Environ. Health* **221**, 907–920 (2018).
22. Quick, J. et al. Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. *Nat. Protoc.* **12**, 1261 (2017).
23. Urban, L. et al. Freshwater monitoring by nanopore sequencing. <https://www.biorxiv.org/content/10.1101/2020.02.06.936302v1>. (2020).
24. Acharya, K. et al. A comparative assessment of conventional and molecular methods, including MinION nanopore sequencing, for surveying water quality. *Sci. Rep.* **9**, 1–11 (2019).
25. Acharya, K. et al. Metagenomic water quality monitoring with a portable laboratory. *Water Res.* **184**, 116112 (2020).
26. Leight, A. K., Crump, B. C. & Hood, R. R. Assessment of fecal indicator bacteria and potential pathogen co-occurrence at a shellfish growing area. *Front. Microbiol.* **9**, 384 (2018).
27. Adrados, B. et al. Comparison of removal efficiency of pathogenic microbes in four types of wastewater treatment systems in Denmark. *Ecol. Eng.* **124**, 1–6 (2018).
28. King, C. H. et al. Baseline human gut microbiota profile in healthy people and standard reporting template. *PLoS ONE* **14**, 9 (2019).
29. HSE. *The Approved List of Biological Agents* (HSE, 2013).
30. Mohapatra, S. et al. The novel SARS-CoV-2 pandemic: possible environmental transmission, detection, persistence and fate during wastewater and water treatment. *Sci. Total Environ.* 142746 (2020).
31. ENPHO. *Shit Flow Diagram* (ENPHO, 2020).
32. Liyanage, C. P. & Yamada, K. Impact of population growth on the water quality of natural water bodies. *Sustainability* **9**, 1405 (2017).
33. Adhikari, M. P., Neupane, M. R. & Kafle, M. Physico-chemical parameterization and determination of effect of tributaries on enhancement of pollutants in Bagmati River. *J. Nepal Chem. Soc.* **40**, 36–43 (2019).
34. Nwidi, L. L. et al. Assessment of the water quality and prevalence of water borne diseases in Amassoma, Niger Delta, Nigeria. *African J. Biotechnol.* **7**, 17 (2008).
35. Cissé, G. Food-borne and water-borne diseases under climate change in low-and middle-income countries: further efforts needed for reducing environmental health exposure risks. *Acta Tropica.* **194**, 181–188 (2019).
36. Scheytt, T. et al. Pharmaceuticals in groundwater: clofibric acid beneath sewage farms south of Berlin, Germany (ACS Publications, 2001).
37. Gottschall, N. et al. Pharmaceutical and personal care products in groundwater, subsurface drainage, soil, and wheat grain, following a high single application of municipal biosolids to a field. *Chemosphere* **87**, 194–203 (2012).
38. Müller, B. et al. Pharmaceuticals as indicators of sewage-influenced groundwater. *Hydrogeol. J.* **20**, 1117–1129 (2012).
39. Pathmaja, S. et al. Risk of groundwater pollution in open dug wells in the Pungudutivu Island of the Jaffna Peninsula, Sri Lanka. *J. Geol. Soc. Sri Lanka* **20**, (2019).
40. Menon, N. G. et al. *Review on Occurrence and Toxicity of Pharmaceutical Contamination in Southeast Asia, in Emerging Issues in the Water Environment during Anthropocene*. 63–91 (Springer, 2020).
41. Janousková, J., Krcmery, V. & Kadlecová, O. Salmonellae with antibiotic resistance and R plasmids in Danube River. *Zentralbl. Bakteriol. Orig. A* **233**, 495–504 (1975).
42. Lonigro, A. et al. Faecal pollution on vegetables and soil drip irrigated with treated municipal wastewaters. *Agric. Water Manag.* **174**, 66–73 (2016).
43. Jongman, M. & Korsten, L. Irrigation water quality and microbial safety of leafy greens in different vegetable production systems: a review. *Food Rev. Int.* **34**(4), 308–328 (2018).
44. World Health Organization. *Sanitation Safety Planning: Manual for Safe Use and Disposal of Wastewater Greywater and Excreta* (World Health Organization, 2015).
45. KVMWP. *Environmental Compliance Monitoring Report* (ed KVMWP Project) (KVMWP, 2020).
46. Aarestrup, F. M. & Woolhouse, M. E. J. Using sewage for surveillance of antimicrobial resistance. *Science* **367**, 630–632 (2020).
47. UN. *Sustainable Development Goals (SDGs)* (UN, 2015). <https://sdgs.un.org/goals>.
48. APHA. *Standard Methods for the Examination of Water and Wastewater* (American Public Health Association, 1920).
49. Knights, D. et al. Bayesian community-wide culture-independent microbial source tracking. *Nat. Methods* **8**, 761–763 (2011).

ACKNOWLEDGEMENTS

The travel, subsistence and consumables funding for this work was provided by Newcastle University's Global Challenges Academy, via Research England GCRF QR funding. This work also received additional support and salary funding from the United Kingdom's Engineering and Physical Sciences Research Council (EPSRC) (grant number: EP/P028527/1) and Water Security and Sustainable Development Hub (funded by; the UK Research and Innovation's Global Challenges Research Fund (GCRF), via the Economic and Social Research Council (ESRC), grant number: ES/S008179/1).

AUTHOR CONTRIBUTIONS

KP: Conceptualisation, methodology, investigation and writing—original draft. KA: Funding acquisition, conceptualisation, methodology, formal analysis, investigation and writing—original draft. SM: Conceptualisation, methodology, writing—review and editing. SK: Software and formal analysis. NS: Conceptualisation and methodology. CO-B: Investigation, writing—review and editing. GB: Investigation, writing—review and editing. SDS: Writing—review and editing. PR: Resources and investigation. DW: Supervision, writing—review and editing. All authors reviewed the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41545-021-00099-1>.

Correspondence and requests for materials should be addressed to K.A.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the

article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021