Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Metagenomic water quality monitoring with a portable laboratory

Kishor Acharya a, Adrian Blackburn a, Jemila Mohammed c, d, b, Alemseged Tamiru Haile c, Andualem Mekonnen Hiruy b, David Werner a, *

a School of Engineering, Newcastle University, Newcastle upon Tyne, NE1 7RU, United Kingdom
b Centre for Environmental Science, Addis Ababa University, PO.Box 1176, Addis Ababa, Ethiopia
c International Water Management Institute (IWMI), Addis Ababa, Ethiopia
d Addis Ababa Water and Sewerage Authority (AAWSA), Addis Ababa, Ethiopia

ARTICLE INFO

Article history:
Received 13 April 2020
Received in revised form 23 June 2020
Accepted 25 June 2020
Available online 10 July 2020

Keywords:
Water quality
Field kits
Fecal source tracking
Waterborne hazards
Metagenomics
MinION nanopore sequencing

ABSTRACT

We describe the technical feasibility of metagenomic water quality analysis using only portable equipment, for example mini-vacuum pumps and filtration units, mini-centrifuges, mini-PCR machines and the memory-stick sized MinION of Oxford Nanopore Technologies, for the library preparation and sequencing of 16S rRNA gene amplicons. Using this portable toolbox on site, we successfully characterized the microbiome of water samples collected from Birtley Sewage Treatment Plant, UK, and its environs. We also demonstrated the applicability of the portable metagenomics toolbox in a low-income country by surveying water samples from the Akaki River around Addis Ababa, Ethiopia. The 16S rRNA gene sequencing workflow, including DNA extraction, PCR amplification, sequencing library preparation, and sequencing was accomplished within one working day. The metagenomic data became available within 24–72 h, depending on internet speed. Metagenomic analysis clearly distinguished the microbiome of pristine samples from sewage influenced water samples. Metagenomic analysis identified the potential role of two bacterial genera not conventionally monitored, Arcobacter and Aeromonas, as predominant faecal pollution indicators/waterborne hazards. Subsequent quantitative PCR analysis validated the high Arcobacter butzleri abundances observed in the urban influenced Akaki River water samples by portable next generation sequencing with the MinION device. Overall, our field deployable metagenomics toolbox advances the capability of scientists to comprehensively monitor microbiomes anywhere in the world, including in the water, food and drinks industries, the health services, agriculture and beyond.

© 2020 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Water quality surveying is essential for achieving several United Nation Sustainable Development Goals (Alcamo, 2019), including clean water and sanitation (SDG 6), zero hunger (SDG 2) and good health (SDG 3). Robust and frequent water quality monitoring underpins safe water and food provision (Wright et al., 2004; Mazari-Hiriart et al., 2008), and good health (Schwarzenbach et al., 2010). Comprehensive water quality monitoring enables evidence-based management and protection of water resources, the design and monitoring of water treatment and sanitation systems, safe water use in agriculture, and the enforcement of regulations protecting consumers and the environment (WHO, 2017; UN, 2019). Most countries nowadays have basic water quality testing capacity at a national level, but water suppliers and surveillance agencies nonetheless often fail to meet requirements for the coverage, quality assurance and frequency of water testing (WHO, 2017). Peletz et al. identified transportation problems as a key factor for inconsistent water testing outcomes in Sub-Saharan Africa, especially where testing samples from rural locations (Peletz et al., 2018). Transportation of water samples from field sites to centralized laboratories adds costs and delays the availability of data for decision making. In addition, storage and transport may result in sample changes, especially in microbiological and DNA samples (Takahara et al., 2015), causing biased results. Portable methods can overcome these limitations and enable near-real time screening of water quality for rapid decision making (Peletz et al., 2018; Aarestrup and Woolhouse, 2020).

Some field kits are already available for basic microbial water
quality assessments such as faecal coliform and *Escherichia coli* counts by Membrane Filtration (MF) or Most Probable Number (MPN) (HACH, 2020b; HACH, 2020a). Furthermore, multiple sensor, spectrophotometric, fluorescence and microscopic based portable technologies can detect microbes and chemicals, including bacterial concentration (Grossi et al., 2013), pathogenic parasites (Mudanyali et al., 2010), fluoride (Hussain et al., 2017), faecal coliforms (Hakalehto and Heitto, 2012), biological oxygen demand (Baker et al., 2015), pesticides (Sicard et al., 2015) and heavy metals (Weindorf et al., 2012). However, these methods provide limited insight into the actual composition of water microbial communities. High-throughput molecular methods such as next generation sequencing (NGS) allow more comprehensive water microbiome characterisations. NGS enables screening different types of bacteria, including faecal indicators and putative pathogens, for example *Bacteroides*, *Streptococci*, and *Vibrio cholerae*, etc. (Acharya et al., 2019; Mrozik et al., 2019). Unfortunately, conventional sequencing platforms such as Illumina and Ion Torrent are large-sized instruments, that are difficult to transport, install and operate in settings, where availability of continuous power and cold chains, laboratory space, and trained personnel is limited (Quick et al., 2015). An exciting development is therefore the recent release of a low cost, field-deployable, memory-stick sized sequencer; the MinION from Oxford Nanopore Technologies (ONT) Ltd. Its portability offers opportunity to comprehensively survey microbial water quality in remote locations.

Using the ONT 16S rRNA portable sequencing kit, one can analyse up to 48 samples and identify more than 100,000 bacteria from each sample on a MinION with two flow cells costing £1000 (Acharya et al., 2019; Urban et al., 2020). In comparison, conventional bench-top sequencing machine investment costs are about £50,000. The portable MinION NGS platform has already been used in laboratory settings for surveying public and animal health, and water quality (Quick et al., 2017; Hu et al., 2018; Rames and Macdonald, 2018; Tyler et al., 2018; Theuns et al., 2018). *Theuns* et al. used MinION NGS as a diagnostic tool and revealed porcine kobuvirus as the main enteric virus causing swine disease (*Theuns* et al., 2018). *Rames and Macdonald* detected Enteroviruses (EV) in wastewater (WW) samples with MinION sequencing, after spiking EV into the WW (*Rames and Macdonald*, 2018). Hu et al. traced faecal contamination in urban stormwater by MinION shotgun sequencing (*Hu* et al., 2018). *Urban* et al. monitored freshwater quality by nanopore sequencing (*Urban* et al., 2020). *Acharya* et al. combined MinION 16S rRNA amplicon sequencing with other methods to survey drinking water quality in informal settlements in Nepal (*Acharya* et al., 2019). *Korber* et al. used MinION NGS for real-time mutation tracking of SARS-CoV-2 (*Korber* et al., 2020). *Quick* et al. used the MinION in Guinea and successfully demonstrated real-time genomic surveillance of *Ebola*, an emerging infectious disease (*Quick* et al., 2016). However, to the best of our knowledge, no previous study has used the MinION platform to survey water quality on site without relying on supplementary stationary equipment in the researchers’ well-equipped laboratories for the sequencing library preparation.

Our study aim was to assemble all the equipment needed for metagenomic analysis of water samples, from sampling to sequencing library preparation and data generation and analysis, in a portable toolbox fitting into one or two suitcases. We aimed to demonstrate the technical feasibility of using this toolbox for metagenomic analysis in two case study applications: 1) on site water quality monitoring in a small wastewater treatment plant, as recently envisioned in a Science Policy Forum article on Global Health (*Aarestrup and Woolhouse*, 2020), and 2) water quality surveying in a low income country with only limited sanitation coverage and laboratory resources.

### 2. Method and site descriptions

#### 2.1. UK case study site and sampling description

The first case study was conducted in April 2019 at Northumbrian Water’s sewage treatment plant (STP) at Birtley, a small town in North East England. This STP received wastewater from a population equivalent of around 30,000, which was treated by both activated sludge and trickling filter processes. Trickling filter effluent was then co-treated with polluted mine water in a constructed wetland, named Lamesley Mine Water Treatment (LMWT). This case study was to demonstrate on-site metagenomic water testing facilities. Water samples were collected in duplicates from five sampling locations (Fig. 1): 1) Birtley STP trickling filter effluent (STP_Eff), 2) Mine water effluent (Mine_Eff) from Kibblesworth mining, 3) effluent from LMWT (Reed_Eff), 4) River Team upstream from the LMWT discharge (RivUpS) and 5) River Team downstream from the LMWT discharge into the river (RivDownS). Water from each site was sampled aseptically. Two litres of water from each site were collected in two sterile 1L bottles. The water sample processing and analysis from each sampling site was done in duplicate (i.e. n = 10). All the activities required for sequencing were conducted on site using solely equipment from the portable toolbox.

#### 2.2. Ethiopia case study site description

The Akaki River, which runs through Ethiopia’s capital, Addis Ababa, was our case study site to demonstrate metagenomic water analysis in a low-income country. The Akaki River catchment is highly polluted from urban activities (e.g. untreated domestic, commercial, and industrial discharge), which affects livelihoods of the downstream communities and ecosystem services (e.g. agricultural production in irrigated fields). The water samples were collected from four locations (Fig. 1): S1) upstream of Addis Ababa city on the Legedadi stream which drains into the Big Akaki River, S3) at the outlet of the Kebena River which drains the old and densely populated part of Addis Ababa into the Big Akaki River, S4) downstream of the confluence of the Kebena and Big Akaki Rivers, and S5) in the downstream part of the Big Akaki River before it enters an area with increased agricultural land use. A room with a working bench was provided by Addis Ababa Water and Sewerage Authority (AAWSA). Microbial analyses for the four grab samples were conducted in duplicate (n = 8) according to the methods described below (note: SQK-16S024 was used instead of SQK-RAB204). The sequencing data processing and analysis were conducted at the International Water Management Institute (IWMI), Addis Ababa, Ethiopia.

#### 2.3. Portable laboratory composition

The portable tools required for metagenomic analysis of water samples included a vacuum pump, filtration unit, mini-centrifuge, MinION, PCR machine, and associated consumables and small equipment items such as pipettes and tips and biohazard waste bags (Table 1). These items could be readily packed into the check-in sized luggage of two travellers (leaving enough space for their personal items) to be transported to the case study sites by road or air and road. Upon arrival at Birtley STP, or at AAWSA, a working bench in rudimentary on-site water testing laboratories was cleaned and sterilized with 70% ethanol before setting up the portable equipment. Bench space and access to an electrical power supply were the only site assets utilized for the metagenomics work. In addition, an internet connection was required later for data processing using ONT’s cloud-based bioinformatics platform.
The portable laboratory included equipment and consumables for physicochemical water quality analysis to provide metadata (Table 1). A portable HACH spectrophotometer with cuvette tests was used at Birtley in the UK to quantify different ions such as nitrate, nitrite, ammonium and fluoride. Due to airline restrictions on hazardous chemicals within check-in luggage, the cuvette-test based HACH spectrophotometric methods could not be used in Ethiopia, and chemical parameters such as nitrate, phosphate and ammonium were semi-quantitatively analysed using LaMotte test strips (LaMotte Europe Ltd, Warwick, UK).

2.4. Metagenomics water quality analysis

The overall workflow for water quality surveying in this study is shown in Fig. 2. For total DNA extraction, typically 250 mL of water were filtered through 0.22 μm membranes (Sartorius UK Limited, Surrey, UK). For S1, STP_Eff and Mine_Eff, 500, 100, and 1000 mL water were filtered to compensate for expected differences in the biomass concentrations of these samples. The total DNA was immediately extracted from the biomass in the membrane using a PowerWater DNA Isolation Kit as per the manufacturer’s instructions (QIAGEN, Crawley, UK). DNA concentration was measured using a Qubit dsDNA HS Assay Kit (Life Technologies, UK). The sequencing library for 16S rRNA gene sequencing was generated using Guppy (Version; v2.3.5) software (ONT, Oxford, UK) producing.fastq files. This step converted the electrical signals generated by a DNA strand passing through the nanopore into the corresponding base sequence. The sequencing and controlled using ONT’s MinKNOW software.

2.5. Sequencing data processing and statistical analysis

The raw reads (i.e. HDF5 raw signals) were base-called with GUPPY (Version; v2.3.5) software (ONT, Oxford, UK) producing .fastq files. This step converted the electrical signals generated by a DNA strand passing through the nanopore into the corresponding base sequence. Base-called data were then uploaded to the EPI2ME interface (v. 2.59.1896509), a platform for cloud-based analysis of base-called MinION data. Data interpretation was performed with the FASTQ 16S workflow, using a quality score ≥7 for filtering. The FASTQ 16S workflow revealed the taxonomic classification of base-called reads along with their frequency. The online data analysis platform EPI2ME of ONT enabled easy exploration of the metagenomic data, but did not provide functionalities for statistical data evaluation. For further processing, the taxonomic classification and quality of barcoded reads was downloaded from the EPI2ME dashboard as a CSV file which contained information on run and read IDs and read accuracy, barcodes, and NCBI taxa IDs for classified reads. The CSV file was processed with Matlab © scripts for 1) generating root level OTU tables, by matching NCBI taxa IDs to lineages and counting the number of reads per NCBI taxa ID, with and without rarefaction; 2) combining root level OTU tables from different runs into a single table; 3) creating OTU tables with grouping of reads at genera level; 4) extracting species or genera of interest from OTU tables; 5) multivariate data analysis, including principal component and cluster analysis and ANOSIM. These Matlab © scripts are available under a Creative Commons ShareA-like licence upon request from the corresponding author. For the overall microbial community analysis, an equal number of reads above the quality threshold (90,000 and 100,000, respectively, for Akaki River and Birtley STP samples) was drawn without replacement from each barcode (i.e. sample). Multivariate data analysis was performed using square root transformed relative abundance data (Hellinger transformation), of reads classified to at least genera level and grouped at this level. Cluster analysis was performed using average Euclidean distance for the linkage tree, and principal component analysis (PCA) was performed using default Matlab © PCA settings. ANOSIM was performed with the Fathom Toolbox for Matlab © developed by the Marine Resource Assessment Program at the University of South Florida’s College of Marine Science (Jones, 2015).

2.6. Metadata collection with portable tools

To provide context water samples were also analysed for pH, electrical conductivity, salinity, total dissolved solid and temperature on-site using a portable probe from EXTECH INSTRUMENTS (Boston, USA). For chemical analysis, HACH cuvette test kits were used in the UK to measure nitrite (NO₂⁻), nitrate (NO₃⁻), ammonia (NH₃–N), and fluoride (F⁻), following the manufacturer’s instructions (HACH LANGE LTD, Manchester, UK). Alkalinity was measured using a HACH digital titrator, following the manufacturer’s instructions (HACH LANGE LTD, Manchester, UK). In Ethiopia, physicochemical parameters pH, alkalinity, nitrite, nitrate, phosphate and ammonium were semi-quantitatively analysed using LaMotte test strips (LaMotte Europe Ltd, Warwick, UK).

2.7. Additional microbiology work

MinION microbiology work

MinION sequencing is ideal for screening, but might result in some false positive outcomes, especially at species level, because of limited read accuracy, or because of non-viable microorganisms in...
Table 1
List of equipment, probes, consumables and reagents used in portable toolbox for comprehensive water quality surveying.

| Items                                         | Number | Product code | Company                      |
|-----------------------------------------------|--------|--------------|------------------------------|
| **Equipment and Probes**                      |        |              |                              |
| Conductivity/PH/Temperature/Salinity probe    | 1      | EC510        | PASS LTD                     |
| DR1900 Portable Spectrophotometer to quantify anions (e.g., nitrate, nitrites, ammonium, fluoride) | 1      | DR1900-02L   | HACH                         |
| Test kit, alkalinity, model AL-DT             | 1      | 2663700      | HACH                         |
| Thermocycler to perform PCR reaction for sequencing library preparation | 1      | MF-QP-1016-01 | miniPCR supplied by Cambio |
| Fluorometer to quantify DNA and sequencing library | 1      | Qubit 4 Q3238 | Life Technologies            |
| Laptop to run sequencing and perform data processing | 1      | Corel7 vPro 7th Gen (ThinkPad) | Lenovo                     |
| Scientific Industries SN™ Vortex-Genie™ 2 (used in DNA extraction) | 1      | 15557335     | Fisher Scientific            |
| Vortex Adapter for 6 (5 & 15 mL) tubes (used in DNA extraction) | 1      | 13000-V1-5   | Qagen                        |
| MinION to perform DNA sequencing              | 1      | n/a          | Oxford Nanopore Technologies |
| Pipettes                                      | 2      | SPS-KIT (P2, 10, 100, 200, 1000) | SciQuip Ltd                 |
| Fisherbrand™ Mini-Centrifuge 100-240V, 50/60Hz Universal Plug, Grey (used in DNA extraction and sequencing library preparation) | 1      | 15358266     | Fisher Scientific            |
| DynaMag™-PCR Magnet (used in library clean-up) | 1      | 492025       | Thermo Fisher Scientific     |
| Mini Diaphragm Vacuum Pump LABOPORT™ (to filter water samples) | 2      | 181–3134     | VWR                          |
| Nalgene Reusable Filter Unit 250 × 250ml (to filter water samples) | 5      | FIL8250      | Scientific Laboratory Supplies |
| **Reagents**                                  |        |              |                              |
| NuclEase Free Water (to make up volumes in DNA extraction and sequencing library preparation) | 1      | AM9906       | Life Sciences                |
| Absolute Ethanol (to clean up sequencing library) 10 mL n/a | 1 | Q32851 | Life Sciences |
| Qubit dsDNA HS Asssay Kit (to quantify DNA and sequencing library) | 1 | M0544L | New England Biolabs |
| NEBNext Ultra™ II Q5™ Master Mix (to perform PCR for sequencing library preparation) | 1 | SQK-RAB204 | Oxford Nanopore Technologies |
| 16S Barcoding kit* (to prepare sequencing library) | 1 | SQK-160024  | Oxford Nanopore Technologies |
| 16S Barcoding Kit 1–24* (to prepare sequencing library) | 1 | SQK-160024  | Oxford Nanopore Technologies |
| Agencourt AMPure XP (to clean up sequencing library) | 1 | A63880 | Beckman Coulter |
| Ampules, m-endo broth, plastic, pk/50 (media to enumerate total coliform) | 1 | 2373550 | HACH |
| Ampules, plastic, m-FC 2 mL, pk/50 (media to enumerate faecal coliform) | 1 | 2373250 | HACH |
| DNeasy PowerWater Kit (50) (Reagents to extract DNA) | 1 | 14900-50-NF | Qagen |
| **Consumables**                               |        |              |                              |
| CN Membrane; white - black; sterile; 0.45 µm; 47 mm; 100pcs (filter membrane used for TC and FC enumeration) | 1 | 11406-47—ACN | Sartorius |
| MiniION Flowcells (to load the sequencing library for DNA sequencing) | 2 | FLO-MIN106D | Oxford Nanopore Technologies |
| CN Membrane; white - black; sterile; 0.2 µm; 47 mm; 100pcs (to capture the biomass during filtration for DNA extraction) | 1 | 11407-47—ACN | Sartorius |
| Polystyrene Petri dish, w/pad (PALL), 100pk (to collect the nutrient media for performing TC and FC enumeration) | 1 | 1417199 | HACH |
| Bottles, wide neck, with screw cap, Nalgene® (to collect the water samples) | 4 | 215–2550 | VWR |
| Pipette tips (1–10 µL) are used in chemical analysis, DNA extraction and sequencing library preparation for dispensing samples and reagents. | 2 | 732–1487 | VWR |
| Pipette tips (20–200 µL) | 2 | 732–1489 | VWR |
| Pipette tips (200–1000 µL) | 2 | 732–1491 | VWR |
| Microtube storage boxes to store the DNA samples | 2 | 525–0925 | VWR |
| PCR-Cooler, 0.2 mL, Starter-Set (1xpink, 1xblue) | 2 | 479–1160 | VWR |
| Molecular grade water to make up volume and dilution | 2 | AM9906 | Thermofisher Scientific |
| Nitrate cuvette test 0.23–13.5 mg/L NO−3, 25 tests to quantify nitrate | 1 | LCK339 | HACH |
| Nitrate cuvette test 5–35 mg/L NO−3, 25 tests to quantify nitrate | 1 | LCK340 | HACH |
| Nitrite cuvette test 0.015–0.6 mg/L NO−2, 25 tests to quantify nitrite | 1 | LCK341 | HACH |
| Nitrite cuvette test 0.6–6.0 mg/L NO−2, 25 tests to quantify nitrite | 1 | LCK342 | HACH |
| Ammonium cuvette test 0.015–2.0 mg/L NH−4, 25 tests to quantify ammonium | 1 | LCK304 | HACH |
| Ammonium cuvette test 2.0–47.0 mg/L NH−4, 25 tests to quantify ammonium | 1 | LCK303 | HACH |
| Fluoride cuvette test 0.1–2.5 mg/L F, 25 tests to quantify fluoride | 1 | LCK323 | HACH |
| INSTA-TEST NATURAL WATER 5-WAY TEST STRIPS/25/BTL to analyse pH, nitrate, nitrite, hardness, alkalinity | 1 | 3038-G | LaMotte Europe |
| INSTA-TEST LR PHOSPHATE ENV STR 0–2500PPB W/TBUEACA/CAP 25/BTL to analyse phosphate | 1 | 3021-G | LaMotte Europe |
| INSTA-TEST AMMONIA TEST STRIPS/25/BTL to analyse ammonia | 1 | 3023-G | LaMotte Europe |

*Used only in Ethiopia, ‡used only in UK, and *used in both the UK and Ethiopia. In addition to the items listed, PPE (laboratory coats, safety glasses, nitrile gloves), tweezers and hazardous waste bags were included in the items needed for the fieldwork.

discharged water samples. Further validation of results is therefore recommended while the technology is still under development and evaluation (Acharya et al., 2019). For context and validation, total coliform (TC) and faecal coliform counts (FC) in water samples were determined by conventional membrane filtration, incubation (37 °C; TC and 44 °C; FC, for 18–24 h), and plate counting following Standard Methods for the Examination of Water and Wastewater (APHA 2015). Enumeration of coliform bacteria after culturing is
the WHO recommended method for monitoring microbial water quality. The filtration was performed using the portable tools in Table 1. For microbial enumeration the plates were incubated in stationary incubators, since they were available in Newcastle University and AAWSA laboratories. In principle, these conventional microbial culturing assays could also have been completed using commercial field kits (HACH, 2019). To validate the very high Arcobacter butzleri signature observed in Ethiopia with the portable metagenomic toolbox, we additionally performed at Newcastle University qPCR on extracted DNA from Akaki River water samples. For this validation, we targeted the ciaB gene (see Table S1 for the methodology).

3. Results and discussion

3.1. Costs and timelines for metagenomics water quality analysis with the portable laboratory

The initial investment for the portable equipment, including the sequencing devise and computer for the toolbox used in this study was approximately £10,000. The reagents and consumables for metagenomics analysis of 10 samples costed approximately £1200, or £120 per sample, while those for basic water chemistry and microbiology analysis costed approximately £26 per sample (Table S2). With the recent release of a 24-sample barcoding kit by ONT, per sample costs of the metagenomics analysis can be reduced to about £63 per sample (Table S2). In developing countries, the VAT and freight costs can vary and will significantly impact the per sample analysis cost.

The flow-cells, PCR reagents, and 16S rRNA sequencing kit required cold storage. These items were placed into a polystyrene box with the cool packs used by ONT to ship their flow cells to customers. That packaging maintained a temperature of about 4 °C for approximately 2 days. For fieldwork in Ethiopia, this box was sealed and placed in a holdall with the other consumables in a check-in luggage of the travellers from the UK. For custom clearance in Ethiopia, we attached an accompanying letter from IWMI.

A timeline for each activity in the UK and Ethiopia starting with the portable laboratory set-up, sampling, sample processing, data generation to data analysis is provided in Fig. 2 and in Table S3 in supporting information. The practical work at Birtley STP took 8 h. Within this time we collected, processed and analysed samples from five locations in duplicate (n = 10), completed the physico-chemical tests (Table S2), including chemical analysis by portable spectrophotometry, and then obtained comprehensive water quality results including a description of the water microbiomes within 24 h (Fig. 2). However, factors such as number of personnel involved, location and accessibility of sampling sites, number of equipment, computer specification and internet speed will determine the time required for water quality surveying with the portable laboratory. In this case, a team of three experienced people completed predefined tasks. The sampling sites were within close vicinity of Birtley STP and easily accessible. In contrast, sampling in Ethiopia took approximately 5 h to sample four different sites in the Akaki River, as the sampling locations were more dispersed, and traffic congested in parts of Addis Ababa. Filtration can be another time-consuming step. The overall filtration time was determined by turbidity, volume and the number of water samples to be filtered, and the number and power of the vacuum pumps used. At Birtley STP, we filtered water samples from five sites in duplicate (n = 10) in 1 h by using two pumps simultaneously, and a separate, pre-sterilized filtration unit for each sample. In Ethiopia, filtration of water samples from four sites in duplicate (n = 8) with a single vacuum pump took approximately 2 h.

Library preparation, including PCR, is a key process in 16S rRNA sequencing (Quick et al., 2016). Our portable PCR thermocycler (miniPCR bio™, USA) had a maximum capacity of 16 samples per run which would become a bottleneck, if the samples exceeded 16.
ONT recently released a 16S rRNA sequencing kit with 24 barcodes (SQK-16S024) to multiplex and sequence up to 24 samples in a single workflow. This would still generate $10^5$–$10^6$ reads per sample, but would ideally require 24 sample capacity for the thermocycler.

As a high throughput method, the MiniION generated gigabytes of raw sequencing data. In the UK case study, within 4 h, approximately 2 million 16S rRNA genes were sequenced and generated raw data in fast5 format, which were decoded by the base-calling software Guppy within 8 h. The time required for base calling depends on the specification of the computer or laptop used, and this can be done without an internet connection. The laptop used in this study had an Intel(R) Core(TM) i7 processor with 16 GB RAM memory and 8 cores. The MinKNOW software from ONT to control the sequencing device has an option to base call raw reads as they are being generated. However, we used an offline version of MinKNOW, a special version more suitable for work in remote locations, that does not require internet connection to start the sequencing run, but lacks an option for live base-calling. The base-called data (i.e., fastq) were further processed using the 16S-workflow available in the cloud-based data analysis platform EPI2ME. Within 1 h, we identified and quantified bacteria of interest using the intuitive graphical interface of the EPI2ME website for data exploration. Internet speed is the major factor in determining the EPI2ME 16S workflow processing time. The download and upload speeds of the internet at the time we processed our data in the UK were 96.1 and 94.1 Mbps, respectively. The download and upload speed of the internet in Addis Ababa, Ethiopia, were 10.1 Mbps and 0.87 Mbps, respectively, and it took approximately 6 h to process approximately 1.5 million reads.

The CSV file was downloaded after completion of the EPI2ME 16S workflow within about 1 min in the UK and 10 min in Ethiopia. The interpretation and further statistical analysis of this file was performed within 20 min, both in UK and Ethiopia, using our Matlab scripts, with the speed solely dependent on the laptop specifications. These scripts rapidly generated useful OTU tables at root and genera level, and performed multivariate data analysis, such as PCA and cluster analysis (Fig. 3), and ANOSIM.

On-site water quality analysis using portable probes was completed after sampling with minimal delay. Chemical water quality analysis using a portable spectrophotometer (in the UK) and test strips (in Ethiopia), and a portable alkalinity kit was performed in parallel to the microbiology work, by one of three team members. The water filtration and microbial plating to enumerate total and faecal coliform using conventional methods was completed within 2 h of downtime, whilst waiting for the PCR thermocycler reaction to complete, which took approximately 2.5 h. The plates were incubated at the appropriate temperatures (37 °C and 44 °C, respectively for total and faecal coliform) overnight and enumerated within 24 h.

3.2. Metagenomic data obtained with the portable laboratory

An overall microbial community analysis using rarefied (90,000 and 100,000, respectively for Akaki and Birtley STP settings) and square root transformed relative abundance data, grouped at genus level, is presented as PCA plots and dendograms in Fig. 3. We also performed multivariate data analysis without rarefaction and obtained virtually identical results (data not shown). The PCA plot for the UK case study site displayed a distinct separation of mine water and STP trickling filter effluent samples from the reed bed and river water samples along principal components 1 & 2 (Fig. 3 a). Furthermore, the PCA plot suggested that *Aeromonas* and *Acrobacter*, two genera known to be present in the human gut (Deodhar et al., 1991; Banting and Figueras, 2017), were prevalent in the STP trickling filter effluent water [STP_Eff] and shaped the dissimilarity between STP_Eff and the other samples along principal component 1 (PC1). Reed bed effluent was similar to the river water samples, with prevalence of genera such as *Polaromonas*, *Polynucleobacter* and *Flavobacterium*, which separated these samples from the STP_Eff and Mine_Eff samples in the PCA. The cluster analysis (Fig. 3 b) also separated the trickling filter effluent [STP_Eff], mine effluent [Mine_Eff] and the other three samples [Reed_Eff, RivUpS and RivDownS] into distinct clusters, with good agreement between replicates. Among the two factors analysed (i.e., sampling location and water sample types), only water sample types had a significant effect on the similarity of the samples (ANISOM, R = 1, p-value = 0.002). In the Akaki river catchment in Ethiopia, the cluster analysis and PCA in Fig. 3 c & d separated replicates from S1, S3, S4 and S5 into different clusters, with S1 having the least similarity to...
the other samples. A similar influence of the genus *Arcobacter* was observed along PC1, which separated the most upstream site S1 from the more urban influenced sites S3–S5. Again, there was good reproducibility between replicates. Finally, both factors (i.e., sampling sites and rural versus urban water sample types) had a significant effect on the similarity of the samples (ANOSIM: Sampling sites; R = 0.96, p-value = 0.01 & Water sample types; R = 1, p-value = 0.04).

Acharya et al. demonstrated the value of MinION sequencing data as a screening tool for putative pathogens, but warned that the limited read accuracy of nanopore sequencing may sometimes result in false positive outcomes, especially at species level (Acharya et al., 2019). Consequently, species level identities need to be interpreted with great caution, while more reliable information can be obtained at genus and family level (Acharya et al., 2019). We screened the water samples for the presence and relative abundance of faecal indicator genera (Fig. 4a & b), and gut (Table S4 in supporting information) and putative human pathogenic bacteria (Table S5 in supporting information). As expected, among the five sites in the UK study, the highest relative percentage abundance of human pathogenic bacteria were observed in sewage treatment plant effluent samples [STP_Eff] (0.591 ± 0.054%), but the highest relative abundance of bacteria also abundant in the human gut microbiome were observed in the River Team before mixing with reedbed polished STP effluent (i.e., RivUpS; 0.788 ± 0.015%) (Table S6 in supporting information). This may be due to storm water overflows and other diffuse discharges upstream from Birtley STP. Mine effluent [Mine_Eff] had the lowest relative percentage abundance of the aforementioned bacteria, as would be expected (Table S6 in supporting information). The relative percentage abundance of putative faecal indicator bacteria was significantly lower (*t*-test, two-tailed, p < 0.05) in the reed bed effluent [Reed_Eff] as compared to the wastewater treatment plant effluent [STP_Eff] (Fig. 4a), showing the effectiveness of such a polishing step to reduce faecal influence in water microbiomes. Molleda et al. previously demonstrated the removal of wastewater pathogens, including *Escherichia coli*, total coliforms, *Clostridium perfringens*, faecal *Streptococci*, *Giardia* cysts, *Cryptosporidium* oocysts and helminth eggs, in a constructed wetland (Molleda et al., 2008).

Sequence results from Ethiopia found faecal indicator bacteria, namely *Bacteroides*, *Prevotella* and bacteria from the family *Enterobacteriaceae*, in high relative abundance in the water from urban influenced locations 3, 4 and 5 (Fig. 4b). Among the four sites, the highest relative abundance of human pathogenic (Table S7 & Table S9 in supporting information) and gut bacteria (Table S8 & Table S9 in supporting information) were observed in water samples from site 3 and 4, and the lowest in water samples from the most upstream and rural site 1 (*t*-test, two-tailed, p < 0.05). Site 3 is located at the outlet of the Kebena River just before this river merges with the Bulbula River to form the Big Akaki River, while site 4 is in the Big Akaki River just after these rivers merge. The Kebena River flows through the densely populated urban areas of Addis Ababa, and receives both domestic and industrial discharges, and is the most polluted tributary of Big Akaki River (Eriksson Malin, 2019). Likewise, *Arcobacter butzleri* was the most predominant human pathogenic bacterium identified in these watersheds, and respectively contributed a very high 2.47, 2.33 and 0.86 percentage of the total read abundance in site 3, 4 and 5 (Fig. 4c).

3.3. Physicochemical metadata

Table S10 shows concentrations of different relevant anions and other physico-chemical parameters in the water from the 5 different sampling locations around Birtley STP. Water samples from STP_Eff contained high nitrate and ammonium concentrations (i.e., 168.08 mg/L and 2.49 mg/L), while this concentration...
decreased to 41.32 mg/L and 0.49 mg/L in the reed bed [Reed_Eff] polished effluent sample.

In the Ethiopian study, ammonia, phosphate and nitrite concentrations measured with the test strip method were higher at sites S3, S4 and S5 (Table S11) as compared to the upstream site S1, and exceeded local standards (EPA, 2003), providing strong indication of urban pollution at these downstream sites. The Kebeba River (S3) is much more polluted than the upstream river (S1). When the two rivers merge (S4), the chemical water quality of the Big Akaki River remained similar to site 3, identifying the Kebeba River as a major pollution source of the Big Akaki River. The dissolved oxygen, nitrate and nitrite concentration at site 5 were higher than that measured at site 4 (Table S11), indicating nitrification in the downstream Big Akaki River.

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

The total and faecal coliform bacteria plate counts (Tables S10 and S12) were well aligned with the metagenomics data produced with the MinION device. No cultivable coliforms were observed on the mine effluent membrane filter [Mine_Eff], while significant numbers of both, coliform bacteria and faecal coliform bacteria, were detected in the other Birtley STP water samples (Table S10). The concentration of both types of coliforms were significantly lower (t-test, two-tailed, p < 0.05) in the reed bed water samples as compared to STP_Eff water samples. In Ethiopia, faecal coliform bacteria plate counts were three orders of magnitude lower at the most upstream site S1 compared to the other three sites (Table S12). For validation of the Arcobacter butzleri hazard, which was predominant in Akaki River water samples according to the MinION data (2.47 ± 0.22% in S3, 2.33 ± 0.16% in S4 and 0.86 ± 0.16% in S5), we additionally performed qPCR on extracted DNA for the ciaB gene at Newcastle University after returning to the UK. This gene codes for the invasive antigen B, which is responsible in host cell invasion, and can be used as an indicator for the abundance of Arcobacter butzleri (Lehmann et al., 2015). The qPCR results for ciaB presented in Fig. 4d agreed with the MinION results for Arcobacter butzleri (Fig. 4c), by confirming the much lower presence of this pathogen in the upstream site S1, as compared to the more urban sites. Similarly, Acharya et al. previously found good agreement between patterns observed for Vibrio cholerae with MinION screening and qPCR results for the epsM gene, a toxin secreting gene in Vibrio cholerae (Acharya et al., 2019).

3.5. Appraisal of the portable toolbox for water quality surveying

A recent Science Policy Forum article on Global Health highlighted that portable metagenomics enables near real time quantification of microbial hazards in sewage, even in remote and resource limited settings (Aarestrup and Woolhouse, 2020). The authors suggested that such portable metagenomics could become a surveillance tool in locations without microbiology laboratories (Aarestrup and Woolhouse, 2020). Using only portable tools for next generation sequencing, we first successfully demonstrated comprehensive water quality surveying on site at the Birtley STP by generating metagenomics data in less than 24 h. The same outcome was achieved within 3 days for the Akaki River in Addis Ababa, Ethiopia. The main difference between the two case studies was that sampling took longer and interacted speed was slower in Ethiopia compared to the UK. The 16S rRNA sequencing workflow, including DNA extraction, PCR amplification, sequencing library preparation and sequencing could be accomplished within a working day. Even in the resource limited environment of Ethiopia, our portable toolbox enabled us to comprehensively understand microbial water quality of the four studied sites of Akaki River within 3 days. To the best of our knowledge, this is the first report of water microbiomes generated through NGS of environmental DNA in Ethiopia. Several 16S rRNA gene sequences retrieved from water samples were associated with bacterial groups with public health relevance. However, beyond the technical feasibility demonstrated in this study, there are other obstacles to the uptake of portable metagenomics in low- and middle-income countries, which cannot be ignored. From our experience, the upskilling of local researchers is decidedly feasible and facilitated by modern communication technologies, and the level of funding required for the proposed methodologies is also often available. However, the in-country supply chains for the equipment and reagents needed are patchy, and purchasing procedures are complex and time consuming, which is an obstacle to the wider use of metagenomic methods in developing countries.

Overall, our proposed toolbox has potential to facilitate achieving the UN Sustainable Development Goals (SDG) 6; clean water and sanitation. For example, SDG 6.3 aims to protect both ecosystem health and human health by eliminating, minimizing and significantly reducing different streams of pollution into water bodies. Our toolbox will be useful for monitoring treatment systems and environmental water quality to document how safely managed sanitation reduces risks to public health. Our work at Birtley STP illustrated how polishing of trickling filter effluent in a constructed wetland achieved a microbiome in the final discharge similar to that of the receiving River Team. The method thus provided evidence for the benefits of such treatment. Contrariwise, the Akaki River microbiome was significantly altered by urban pollution inputs which caused greater prevalence of human gut bacteria and putative pathogens in the river water microbiomes. In the Akaki River, NGS led to the discovery of Arcobacter butzleri as a potentially very significant waterborne pathogen in this watershed, which would not have become apparent from the results of conventional methods such as faecal and total coliform plate counts. This pathogen is considered as a serious hazard to human health and known to cause gastroenteritis, diarrhoea and abdominal cramps in humans (Banting and Figueras, 2017). The findings of portable metagenomics can thus inform public health surveillance plans. Arcobacter butzleri is mostly present in a great variety of retail meats, including chicken, beef, pork, and lamb (Rivas et al., 2004; Banting and Figueras, 2017). The high prevalence of this bacterium in the studied watersheds could be linked to multiple factors such as the high number of grazing cattle, Ethiopian food habits and discharge of untreated sewage into the river. Interestingly, in Ethiopian cuisine there is a popular dish based on raw meat (Selleshe et al., 2014).

3.6. Future work

AAWSA is constructing several new wastewater treatment plants, and we intend to continue monitoring the Akaki River to provide evidence for the anticipated benefits of these significant wastewater treatment infrastructure investments. Our toolbox should be useful in many other future applications, including the rapid response to natural and human-made emergencies affecting water supplies. Sequencing allows the simultaneous screening, identification and monitoring of multiple indicator bacteria, which can then inform appropriate mitigation strategies. Adding a portable qPCR machine costing approximately £8000 to the current portable toolbox would expand its capacity to also estimate the absolute abundance of various bacteria present in the water samples via quantification of the 16S rRNA gene copy numbers. In addition, portable qPCR would enable quantification of marker
4. Conclusions

- Near real-time water quality surveying with a portable metagenomics toolbox is feasible in industrial (Birtley STP, UK) and resource limited settings (Addis Ababa, Ethiopia).
- However, lack of reliable supply chains for the equipment and reagents needed is a major obstacle to the wider use of metagenomic methods in low- and middle-income countries.
- The metagenomics data allows simultaneous screening of various putative pathogens and faecal indicators. This can subsequently guide the choice of complementary methods such as qPCR or culturing to more solidly establish the microbes associated with health hazards.
- Arcobacter is a genus of emerging interest for faecal pollution source tracking.
- Portable methods comprising metagenomic sequencing have potential to facilitate water safety monitoring plans in both low- and high-income countries.

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.

Acknowledgement

This project was funded by the United Kingdom’s Engineering and Physical Sciences Research Council (EPSRC), project EP/P028527/1. Additional support came from Water Security and Sustainable Development Hub funded by the UK Research and Innovation’s Global Challenges Research Fund (GCRF), grant number: ES/S008179/1, and the Royal Society, grant ICA/1/R191241. We thank Northumbrian Water, and the Addis Ababa Water and Sewerage Authority, for providing access to their facilities at Birtley STP, and in Addis Ababa, respectively. Data created during this research are openly available at DOI 10.25405/data.ncl.13588230. Please contact Newcastle Research Data Service at rdm@ncl.ac.uk for access instructions.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.watres.2020.116112.

References

Aarestrup, F.M., Woolhouse, M.E.J., 2020. Using sewage for surveillance of antimicrobial resistance. American Public Health Association/American Water Works Association/Water Environment Federation, Washington DC.

Baker, A., Cumberland, S.A., Bradley, C., Buckley, C., Bridgeman, J., 2015. To what extent can portable fluorescence spectroscopy be used in the real-time assessment of microbial water quality? Sci. Total Environ. 532, 14–19.

Banting, G., Figueras, M.J., 2017. Arcobacter. Global Water Pathogen Project.

Deodhar, L.P., Saraswathi, K., Varudkar, A., 1991. Aeromonas spp. and their association with human diarrhea disease. J. Clin. Microbiol. 29 (5), 853–856.

EPA, 2003. Guideline Ambient Environment Standards for Ethiopia. Environmental Protection Organization UNIDO, Addis Ababa.

Fernandez, R.I., Quinones-Falconi, F., 2008. Microbiological implications of periurban agriculture and water reuse in Mexico city. PloS One 3 (5), e80.

Hakalehto, E., Heitto, L., 2012. Minute microbial levels detection in water samples by portable microbe enrichment unit technology. Environ. Nat. Resour. 2, 80–88.

Hussain, I., Ahamad, K.U., Nath, P., 2017. Low-cost, robust, and field portable smartphone platform photometric sensor for fluoride level detection in drinking water. Anal. Chem. 89 (1), 767–775.

Jones, D.L., 2015. Fathom Toolbox for Matlab: Software for Multivariate Ecological and Oceanographic Data Analysis. College of Marine Science, University of South Florida. St. Petersburg, FL, USA.

Korber, B., Fischer, W., Gnarakar, S.G., Yoon, H., Theiler, J., Abfalterer, W., Foley, B., Giorgi, E.E., Bhattacharyya, T., Parker, M.D., 2020. Spike Mutation Pipeline Reveals the Emergence of a More Transmissible Form of SARS-CoV-2. bioRxiv.

Lehrmann, D., Alter, T., Lehrmann, L., Uherkova, S., Seidler, T., Golz, G., 2015. Prevalence, virulence gene distribution and genetic diversity of Arcobacter in food samples in Germany. Berl. Münchener Tierärztl. Wochenschr. 128 (3-4), 163–168.

Molleda, P., Blanco, I., Ansola, G., de Luis, E., 2008. Removal of wastewater pathogen indicators in a constructed wetland in Leon, Spain. Ecol. Eng. 33 (3-4), 252–257.

Mrozik, W., Vininthanatharat, S., Thongsamer, P., Pattanachan, P., Huch, M., Hach, 2020b. USEPA lauryl tryptose broth PourRite amplexes. Method 8074 Membrane Filtration 10. DOC316.53.01209. www.hach.com-asset-get-download-en.jsa.

Mountain Goals. Current Opinion in Environmental Sustainability 36, 126–136.

Rames, E., Macdonald, J., 2018. Evaluation of MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. Nat. Protoc. 13 (6), 1281.

Quick, J., Loman, N.J., Duraffour, S., Simpson, J.T., Severi, E., Cowley, L., Bore, J.A., Quick, J., Loman, N.J., Duraffour, S., Simpson, J.T., Severi, E., Cowley, L., Bore, J.A., Koundounou, R., Dudas, G., Mikhail, A., 2016. Real-time, portable genome sequencing for Ebola surveillance. Nature 530 (7589), 228–232.

Ramos, E., Macdonald, J., 2018. Evaluation of MinION nanopore sequencing for rapid enterovirus genotyping. Virus Res. 252, 8–12.

Rivas, L., Fegan, N., Vanderlinde, P., 2004. Identification and characterisation of Arcobacter buzzleri from meat. Int. J. Food Microbiol. 91 (1), 31–41.

Schwarzkofenbach, R.P., Egli, T., Hofstetter, T.B., van Gunten, U., Wehrli, B., 2010. Global water pollution and human health. In: Gadgil, A., Liverman, D.M. (Eds.), Annual Review of Environment and Resources, vol. 35. Annual Reviews, Palo Alto, pp. 109–136.

Siddiqui, M., Lee, C., Lee, M., 2014. Meat consumption culture in Ethiopia. Korean Biol. Conserv. 183, 64–69.

Singh, A.K., 2006. 136. Science of the Total Environment, 133923.
Matthijnssens, J., Maes, P., Nauwynck, H.J., 2018. Nanopore sequencing as a revolutionary diagnostic tool for porcine viral enteric disease complexes identifies porcine kobavirus as an important enteric virus. Sci. Rep. 8 (1), 9830.

UN, 2019. United Nations Sustainable Development Goals.

Urban, L., Holzer, A., Baronas, J.J., Hall, M., Braeuninger-Weimer, P., Scherm, M.J., Kunz, D.J., Perera, S.N., Martin-Herranz, D.E., Tipper, E.T., 2020. Freshwater Monitoring by Nanopore Sequencing. BioRxiv.

Weindorf, D.C., Zhu, Y., Chakraborty, S., Bakr, N., Huang, B., 2012. Use of portable X-ray fluorescence spectrometry for environmental quality assessment of peri-urban agriculture. Environ. Monit. Assess. 184 (1), 217–227.

WHO, 2017. Safely Managed Drinking Water - Thematic Report on Drinking Water 2017. World Health Organisation, Geneva, Switzerland.

Wright, J., Gundry, S., Conroy, R., 2004. Household drinking water in developing countries: a systematic review of microbiological contamination between source and point-of-use. Trop. Med. Int. Health 9 (1), 106–117.