**Description of Urban Stormwater Microbial Source Tracking and Antibiotic Resistance Gene Dataset**

**Overview:**

This dataset includes water quality measurements from an urban stream (Site A) and three municipal separate storm sewer system outfalls (Sites B-D) situated in a mixed-use urban catchment located in the District of Columbia, USA (**Figure 1**). The total catchment area is 61 ha with outfall drainage areas ranging from 4.69 ha (Outfall Site D) to 37.4 ha (Outfall Site C). Catchment and outfall boundaries were defined with the Spatial Analysis Hydrology tool using outfall and elevation data from the National Hydrology Dataset (<http://datagateway.nrcs.usda.gov>) combined with local stormwater transit drain and pipe data (Department of Energy and Environment, Washington, DC). Outfall drainage area residential land use ranged from 2.35 ha (Site D) to 17.1 ha (Site C). Residential metrics were defined with ArcGIS ArcMap (version 10.3; ESRI, Redlands, CA).



**Figure 1:** Geographic information system (GIS) map location of urban stream, outfall sites, and residential land use assignments. The underlying street map is from OpenStreetMap (data available under the Open Database License: <https://www.openstreetmap.org/copyright>).

A total of 124 samples were collected across 31 sampling events over 12-months. Routine sampling occurred twice per month. Seven additional sampling events were conducted immediately after cumulative precipitation from a single storm exceeded 2.54 mm following a 72 h dry period based on the United States Environmental Protection Agency recommended storm event sampling guidance [1]. Urban stream and outfall samples were always collected on the same day within a short time (< 1 h). Samples were collected in sterile 1 L containers, immediately stored on ice, and transported to the laboratory (< 8 h). A field blank consisting of molecular grade water was included for each sampling day.

For each sample, 100 mL was filtered through a 0.45 µm polycarbonate filter (Fisher Scientific, Pittsburg, PA), placed in a 2 mL screw cap tube containing a silica bead mill matrix (GeneRite, North Brunswick, NJ), and stored at -80℃ (<18-months). DNA extraction was performed with the DNA-EZ RW02 kit (GeneRite) as previously described [2]. DNA was eluted with 100 µL elution buffer into low-adhesion microcentrifuge tubes and stored at 4℃ prior to qPCR amplification (< 24 h). Three method extraction blanks with molecular grade water substituted for DNA extract were performed to monitor for potential DNA contamination for each sample batch.

Genetic testing employed eleven clinically relevant ARG qPCR assays targeting *bla*CMY-2, *qnrA*, *bla*KPC, *bla*OXA-48, *bla*NDM, *bla*VIM, *mcr-1*, *mecA*, *sul1*, *tetW*, and *vanA* (**Table 1**). The HF183/BacR287 qPCR assay was used to characterize human fecal pollution levels [3-5]. The Sketa22 qPCR assay was used as a sample processing control (SPC) [6].

**Table 1:** Clinically relevant antibiotic resistance gene (ARG) qPCR assay oligonucleotide sequences.

|  |  |  |
| --- | --- | --- |
| **qPCR Target** | **Primer and Probe Sequences (5’ 🡪 3’)** | **Reference** |
| *bla*OXA-48 | F: GATTATGGTAATGAGGACATTTCGGGCR: CATATCCATATTCATCGCAAAAAACCACACProbe: FAM-CCATTGGCTTCGGTCAGCATGGCTTGTTT-ZEN/IBFQ | [7] |
| *bla*KPC | F: GCAGCGGCAGCAGTTTGTTGATTR: GTAGACGGCCAACACAATAGGTGCProbe: FAM-CAGTCGGAGACAAAACCGGAACCTGC-ZEN/IBFQ |
| *bla*NDM | F: CCAGCAAATGGAAACTGGCGACR: ATCCAGTTGAGGATCTGGGCGProbe: FAM-ACCGAATGTCTGGCAGCACACTTC-ZEN/IBFQ |
| *bla*VIM | F: TTGCTTTTGATTGATACAGCGTGGGGR: GTACGTTGCCACCCCAGCCProbe: FAM-TCTCGCGGAGATTGAAAAGCAAATTGGACTTCC-ZEN/IBFQ |
| *bla*CMY-2 | F: AGACGTTTAACGGCGTGTTGR: TAAGTGCAGCAGGCGGATACProbe: FAM-TATCGCCCGCGGCGAAAT-ZEN/IBFQ | [8] |
| *mcr-1* | F: CATCGCGGACAATCTCGGR: AAATCAACACAGGCTTTAGCACProbe: FAM-AACAGCGTGGTGATCAGTAGCAT-ZEN/IBFQ | [9] |
| *sul1* | F: CCGTTGGCCTTCCTGTAAAGR: TTGCCGATCGCGTGAAGTProbe: FAM-CGAGCCTTGCGGCGG-ZEN/IBFQ | [10] |
| *mecA* | F: AACCACCCAATTTGTCTGCCR: TGATGGTATGCAACAAGTCGTAAAProbe: FAM-CCTTGTTTCATTTTGAGTTCTGCAGTACCGG-ZEN/IBFQ | [11] |
| *tetW* | F: GCAGAGCGTGGTTCAGTCTR: GACACCGTCTGCTTGATGATAATProbe: FAM-TTCGGGATAAGCTCTCCGCCGA-ZEN/IBFQ | [12] |
| *vanA* | F: CTGTGAGGTCGGTTGTGCGR: TTTGGTCCACCTCGCCAProbe: FAM-CAACTAACGCGGCACTGTTTCCCAAT-ZEN/IBFQ | [13] |
| *qnrA* | F: GGATGCCAGTTTCGAGGAR: CCTGAACTCTATGCCAAAGCProbe: FAM-CACTTCAGCTATGCCGATCTGCGCGAT-ZEN/IBFQ | [14] |

In addition to water sampling, 3,743 quality control samples, including field blanks, method blanks, and no template controls were collected and processed in an identical manner as water samples and are reported in a separate data file [Shanks\_MST ARG\_FINAL\_04152025.xlsx].

Six water quality parameters were measured directly from collected samples. Water quality parameters included turbidity (NTU), pH, conductivity (µS/cm), water temperature (℃), dissolved oxygen (mg/L), and total suspended solids (g/L). Parameters were measured with a YSI 6920 Multiparameter Water Quality Meter (Xylem Analytics, Yellow Springs, OH).

*E. coli* enumeration was performed within eight hours of sample collection with Colilert IDEXX defined substrate technologies (IDEXX Laboratories, Inc. Westbrook, ME, USA).

Precipitation was measured with a reference gauge (WH24B Wireless Weather Station, Fine Off-Set Electronics, Ltd) based on proximity to catchment (< 5 km). Cumulative precipitation of 12 h prior to a sampling event (mm) was used for all statistical analyses due to the nature and impact of precipitation events in urban catchments.

The results for antibiotic resistance gene and microbial source tracking measurements along with other water quality and precipitation datasets and a data dictionary are contained in a separate file [Shanks\_MST ARG\_FINAL\_04152025.xlsx].

**Sampling Period Included in this Data Release:** November 25, 2019 through November 30, 2020

**Point of Contact:** Orin Shanks US EPA / ORD / CEMM / WECD / BMB

**Disclaimer:** This research dataset has been reviewed in accordance with U.S. Environmental Protection Agency (U.S. EPA), Office of Research and Development, and approved for release. Mention of brand names or vendors does not constitute an endorsement of products or services by the U.S. EPA.

**References:**

1. USEPA, *NPDES Storm Water Sampling Guidance Document*. 1992, United States Environmental Protection Agency: Washington, DC.

2. Li, X., et al., *Large-scale implementation of standardized quantitative real-time PCR fecal source identification procedures in the Tillamook Bay Watershed.* PLoS ONE, 2019. **14**(6): p. e0216827.

3. Bernhard, A.E. and K.G. Field, *A PCR assay to discriminate human and ruminant feces on the basis of host differences in Bacteroides-Prevotella genes encoding for 16S rRNA.* Applied and Environmental Microbiology, 2000. **66**(10): p. 4571-4574.

4. Green, H.C., et al., *Improved HF183 quantitative real-time PCR assay for characterization of human fecal pollution in ambient surface water samples.* Applied and Environmental Microbiology, 2014. **80**(10): p. 3086-3094.

5. USEPA, *Method 1697.1: Characterization of human fecal pollution in water by HumM2 TaqMan quantitative polymerase chain reaction (qPCR) assay*, E.O.o. Water, Editor. 2021, United States Environmental Protection Agency: Washington DC.

6. Haugland, R.A., et al., *Evaluation of genetic markers from the 16S rRNA gene V2 region for use in quantitative detection of selected Bacteroidales species and human fecal waste by real-time PCR.* Systematic and Applied Microbiology, 2010. **33**: p. 348-357.

7. Ellington, M.J., et al., *Multicentre evaluation of a real-time PCR assay to detect genes encoding clinically relevant carbapenemases in cultured bacteria.* International Journal of Antimicrobial Agents, 2016. **47**: p. 151-154.

8. Schmidt, G.V., et al., *Sampling and pooling methods for capturing herd level antibiotic resistance in swine feces using qPCR and CFU approaches.* PLoS ONE, 2015. **10**(6): p. e0131672.

9. Yang, D.L., et al., *The occurrence of colistin resistance gene mcr-1 in the Haihe River (China).* International Journal of Environmental Research and Public Health, 2017. **14**: p. 576.

10. Calero-Caceres, W., et al., *Sludge as a potential important source of antibiotic resistance genes in both the bacterial and bacteriophage fractions.* Environmental Science & Technology, 2014. **48**: p. 7602-7611.

11. Kelley, K., et al., *Detection of methicillin-resistant Staphylococcus aureus by duplex droplet digital PCR assay.* Journal of Clinical Microbiology, 2013. **51**(7): p. 2033-2039.

12. Smith, M.S., et al., *Quantification of tetracycline resistance genes in feedlot lagoons by real-time PCR.* Applied and Environmental Microbiology, 2004. **70**(12): p. 7372-7377.

13. Volkmann, H., et al., *Detection of clinically relevant antibiotic-resistance genes in municipal wastewater using real-time PCR (TaqMan).* Journal of Microbiological Methods, 2004. **56**: p. 277-286.

14. Oliveira, M., et al., *The environmental contribution to the dissemination of carbapenem and (fluoro)quinolone resistance genes by discharged and reused wastewater effluents: the role of cellular and extracellular DNA.* Water Research, 2020. **182**: p. 116011.