

Targeted Sequencing of CTX-M Alleles in Seattle Area Wastewater

Angelo Q. Ong¹, Sarah E. Philo¹, Anysiah Taylor², Ruohan Hu², John Scott Meschke¹, Erica R. Fuhrmeister^{1,3*}

¹Department of Environmental and Occupational Health Sciences, University of Washington, Seattle, Seattle, WA, USA

²Department of Global Health, University of Washington, Seattle, Seattle, WA, USA

³Civil and Environmental Engineering, University of Washington, Seattle, Seattle, WA, USA

*Correspondence to:

Dr. Erica Fuhrmeister, efuhrm@uw.edu

4225 Roosevelt Way

Seattle, WA 98105

Abstract

Extended-spectrum-beta-lactamases (ESBLs) are a growing group of antimicrobial resistance (AMR) enzymes that can result in severe clinical outcomes. The CTX-M gene, which encodes for ESBLs in bacteria, confers resistance to third generation cephalosporins and is of high clinical concern. We developed a targeted, long-read sequencing method utilizing unique molecular identifiers to generate accurate, full length CTX-M gene sequences from wastewater. We characterized CTX-M in 36 samples from three Seattle area wastewater treatment plants from April 2020 to March 2021. We identified a core community of alleles that persisted across time and treatment plant. The CTX-M-15 containing protein variant (CTX-M-15/216/28) was detected in all but three samples and made up, at most, 30% of detected CTX-M alleles. We observed significant diversity across the CTX-M gene at the nucleic acid level, although most nucleotide mutations were synonymous - resulting in two to three amino acid variants across 19 loci. By average relative abundance, 23% of protein variants were novel, defined as those not represented in the CARD database. This method provides information (full length gene sequences) that cannot be obtained through other culture-independent methods. This flexible approach can be expanded to additional targets and implemented in settings where AMR surveillance is a priority, such as hospital wastewater.

Introduction

The emergence of antimicrobial resistance (AMR) is an existential threat to human and animal health. In 2019, 4.95 million human deaths were associated with bacterial antimicrobial resistance.¹ The expansion of extended-spectrum beta-lactamases (ESBLs) in multiple genera of bacteria has further worsened the landscape of antibiotic resistance.² ESBLs are enzymes that confer resistance to a wide array of beta-lactam antibiotics targeting the bacterial cell wall. CTX-M (cefotaximase-München-lactamase) is an antibiotic resistance gene (ARG) that encodes for ESBLs and has high efficacy against third generation beta-lactam cephalosporins (3GC). 3GCs are primarily reserved for difficult-to-treat infections such as sexually transmitted infections,³ neonatal sepsis, and meningitis.⁴ Resistance to 3GCs such as ceftriaxone has been observed in many taxa including *E. coli* and *Salmonella* bacteria harboring CTX-M genes.^{5,6} CTX-M is a highly diversified gene with different alleles conferring varying levels of resistance to beta-lactams. A notable example is CTX-M-15, a protein variant that has increased activity against the 3GC ceftazidime compared to other CTX-M alleles.^{7,8}

CTX-M is globally distributed with predominant alleles differing across geographic regions. CTX-M-15 is the most dominant allele in most regions around the world with exceptions to CTX-M-14 in East Asia and CTX-M-2 in South America.⁹ Although there is limited clinical literature for CTX-M in Africa compared to other continents⁹, CTX-M-15 is the most prevalent allele described.¹⁰ Clonal relationships have been demonstrated in virulent ST131 *E. coli* isolates containing CTX-M-15 in eight countries across three continents.¹¹ Despite its clinical significance, CTX-M is not limited to clinical settings as has been detected in a variety of veterinary,¹² agricultural,^{13,14} and environmental settings.^{15,16}

As new CTX-M alleles, subgroups, and clusters are continuously being identified in bacterial isolates,^{17–20} there is a need to characterize alleles at the population-level. Surveillance of ARGs in populations can be done through wastewater-based epidemiology (WBE). WBE has

been used to characterize circulation of pathogens and ARGs at different scales (e.g. city, neighborhood, building).^{21,22} ESBLs have been surveilled in wastewater through culture-based, PCR, and DNA sequencing methods.^{23–25} Metagenomic sequencing is commonly used to provide broad characterization of all ARGs found in wastewater. However, current metagenomic methods are prone to oversampling highly abundant ARGs and can lack sensitivity.^{26,27} In addition, short-read sequencing methods (approx. 300 bps) are unable to provide adequate breadth of coverage for ARGs, such as CTX-M, that are approximately 900 base pairs in length. Several clinically significant, low abundant ARGs, such as CTX-M and their alleles, are often overlooked in wastewater. Targeted sequencing increases sequencing depth of low abundance targets and long-read sequencing allows for high breadth of coverage of ARGs.

A notable limitation of long-read sequencing methods (i.e. Nanopore) is the high error rate. Unique molecular identifiers (UMIs) have previously been developed for full length 16S long-read sequencing. High error rates are corrected by tagging template molecules with UMIs during PCR. Then, UMIs in the sequenced amplicons are clustered and binned to generate consensus sequences.²⁸ The resulting consensus sequences are highly accurate, span the length of the gene, and enable characterization at the single nucleotide level. In this study, we first developed a targeted, long-read sequencing method that utilizes UMIs to sequence the clinically relevant ARG, CTX-M in wastewater and then demonstrate the utility of this method by characterizing the diversity of CTX-M alleles in Seattle-area wastewater.

Methods

Wastewater collection, concentration, and extraction

Primary influent wastewater was grab sampled weekly from three Seattle-area wastewater treatment plants (WWTPs) between April 2020 to March 2021 as part of a previous study.²⁹ All grab samples were transported on ice and stored at 4°C prior to processing. For Brightwater WWTP, multiple grab samples from different time points for a single day were composited and

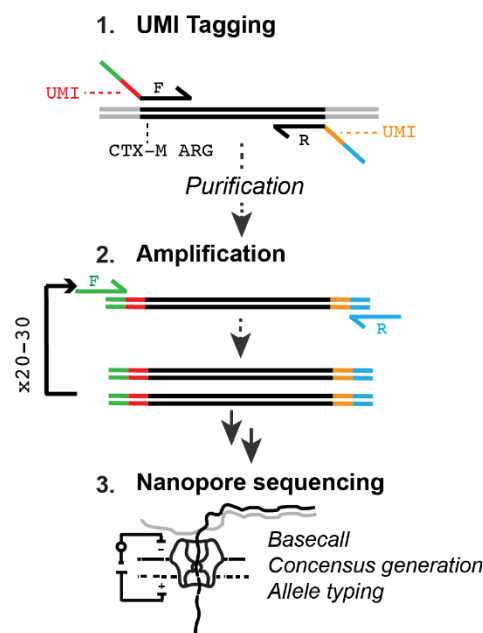
mixed well prior to concentration. For the other WWTPs, one grab sample was collected for a single day. All samples were processed and concentrated within one week of the collection date. 0.5 L of primary influent wastewater was concentrated using 5% skimmed milk, shaken, centrifuged, and pellets resuspended in 6 mL of 1X PBS. DNA and RNA were coextracted from 280 μ L of concentrated sample using a QIAmp Viral RNA Mini extraction kit (Qiagen) with no DNase treatment. One sample from the first week of each month for 12 months was analyzed for CTX-M variants across the three WWTPs (36 samples total).

Gene Quantification

qPCR was performed on wastewater nucleic acid extracts to quantify the gene copies (gc) of CTX-M in each sample (alleles captured in Table S1). CTX-M quantities were used to dilute samples to 5000 gc input in Unique Molecular Identifier PCR (described below). A final concentration of 0.9 μ M for forward and reverse primers, 0.25 μ M of probe, and 1X of TaqMan Fast Advanced MM (Invitrogen) was used in each reaction (Table S2). 2 μ L of sample was used with 18 μ L of qPCR mastermix and samples were run on a StepOne Plus (Applied Biosystems). Undiluted and ten-fold diluted samples were run in duplicate. A gBlock (Integrated DNA Technologies) was used to generate a standard curve, from 10^1 to 10^6 gc per reaction, in triplicate. Triplicate negative controls with 18 μ L of qPCR mastermix and 2 μ L of molecular grade water were included. Cycling conditions for all thermocycling are available in Table S3.

Unique Molecular Identifier PCR

112



113 **Figure 1:** Schematic of UMI PCR Protocol. UMIs are added to each template molecular. Tagged molecules
114 are subsequently amplified and sequenced. Consensus sequences are generated from binning by UMIs.

115 To add unique molecular identifiers (UMIs) to CTX-M genes and amplify the gene and
116 UMIs, two sequential PCR reactions were performed (**Figure 1**). The first reaction used two cycles
117 and primers targeting the CTX-M gene contained an 18 nt random sequence (Table S2). 5000
118 gcs of CTX-M were used as input to the PCR reaction for a total reaction volume of 25 μ L (qPCR
119 results in Figure S1). The mastermix consisted of 0.2 μ M forward and reverse CTX-M primer, 0.5
120 mM dNTPs, 1X SuperFI II buffer (Invitrogen), 1U SuperFI II DNA polymerase, and molecular
121 grade water. 930 bp PCR products were cleaned using 0.9X magnetic beads (Sergi Lab Supplies)
122 with 80% ethanol.

123 The second reaction of UMI PCR was performed with primers to amplify the UMIs along
124 with the template molecules. Mastermix consisted of 0.2 μ M forward and reverse primers, 0.5 mM
125 dNTPs, 1X SuperFI II buffer, 1U SuperFI II polymerase, and molecular grade water. 20 μ L of the
126 previous PCR product was used as input DNA for a total reaction volume of 50 μ L. PCR products
127 were loaded with 6X TriTrack DNA Loading Dye (ThermoScientific) and amplification was

visualized on 1% agarose gels stained with GelGreen (Biotium). PCR products were cleaned as described above.

PCR products were subsequently barcoded to sequence multiple samples on a MinION (Table S4). The final concentration of mastermix for barcoding PCR consisted of 0.2 μ M forward and reverse primer, 0.5 mM dNTPs, 1X SuperFI II buffer, 1U SuperFI DNA polymerase, and molecular grade water. 20 μ L of the previous PCR product was used as input DNA for a total reaction volume of 100 μ L. All PCR thermocycling was performed on a T100 thermal cycler (BioRad Laboratories). PCR products were cleaned as described above. For two of six batches, a ratio of 0.5X beads was used to exclude low molecular weight DNA bands. DNA concentrations for PCR products were measured using the Qubit 1X dsDNA High Sensitivity assay kit (Invitrogen) on a DeNovix Fluorometer and then 70-100 fmols were pooled for sequencing.

DNA Sequencing

Pooled PCR products were prepared for sequencing using Oxford Nanopore's ligation sequencing kit (SQK-LSK-114) according to the manufacturer's protocol. After adaptor ligation, short fragment buffer and a bead ratio of 0.8X was used in the final cleanup step. R10.4.1 flow cells were run on a MinION MK1B for 36 hours. Sequencing reads are available in NCBI under BioProject PRJNA1107162.

Data Analysis

Fast5 files were basecalled with guppy (v6.5.7) using the R10.4.1 super accuracy model (dna_r10.4.1_e8.2_400bps_sup.cfg) with a minimum quality score of 9. Basecalled reads were demultiplexed with cutadapt (v2.6)³⁰ to search for forward and reverse barcodes at each read end with the following parameters: an error rate of 0.15, a minimum length of 800 bps, and maximum length of 1100 bps. Demultiplexed reads were run through an existing ssUMI pipeline (v0.3.2)^{28,31} to produce consensus sequences based on dual UMIs. The longread_umi ssumi_std command

was used with the following parameters: -s 200, -e 200, -E 0.1, -m 800, -M 1100, -f AAGGTTGGCCAGGCTACCCAAAAC, -F CGACGCTAATACATCGCG, -r CAAGCAGAAGACGGCATAACGAGAT, -R ATGGTTAAAAAATCACTGCGCCAGT, -c 3, -p 2. In brief, USEARCH (v11.0.667)³² was used to identify UMI clusters at the beginning and end of each read. Reads were then grouped by UMI pairs and chimeras were removed. Grouped reads were polished using cycles of Racon (v1.4.10)³³ and Medaka (v0.11.5)³⁴ resulting in a single consensus sequence for each UMI pair bin.

ARGs were identified in resulting consensus sequences using Minimap2 (v2.22)³⁵ to map against the Comprehensive Antibiotic Resistance Database (CARD) (v3.2.6).³⁶ Nucleotide alleles and amino acid variants were named based on the closest database match in CARD. If sequences mapped at 100% identity, the name of the allele in CARD was retained. If sequences did not map at 100% identity to any CARD sequence, the name of the closest match (highest percent identity) was retained with a “variant” suffix. More than one allele was included in the name if sequences aligned equally to multiple database matches. Alleles that occurred only once in the entire dataset were removed from the analysis to avoid spurious associations. Data were visualized and analyzed in R (4.2.3)³⁷ with phyloseq (1.42.0).³⁸ Alignment of alleles, frequency plot, and phylogenetic tree were generated using Jalview (2.11.3.2).³⁹

To test for differences in alpha diversity in WWTPs, we used a Kruskal-Wallis test to compare the median number of unique alleles and Shannon diversity indices. Flow values, measured in millions of gallons per day (MGD) at each sample date were used to categorize samples as either high or low flow. Samples with a flow value lower than the median were categorized as low flow while samples with a flow value greater than the median were categorized as high flow (West Point median: 67 MGD, Brightwater: 16 MGD, South Plant: 62 MGD). We used a Wilcoxon-Ranked Sum test to determine if there was a statistical difference in alpha diversity between samples collected during high and low flow.

To validate performance of the PCR, sequencing, and data analysis methods, we performed an experiment where two alleles with 1 bp difference were spiked into wastewater and underwent the processing pipeline described above. Our analysis pipeline correctly identified the spiked-in alleles at the single nucleotide level (Figure S2 and additional method details in the supporting information).

Results and Discussion

Variant Relative Abundance and Diversity

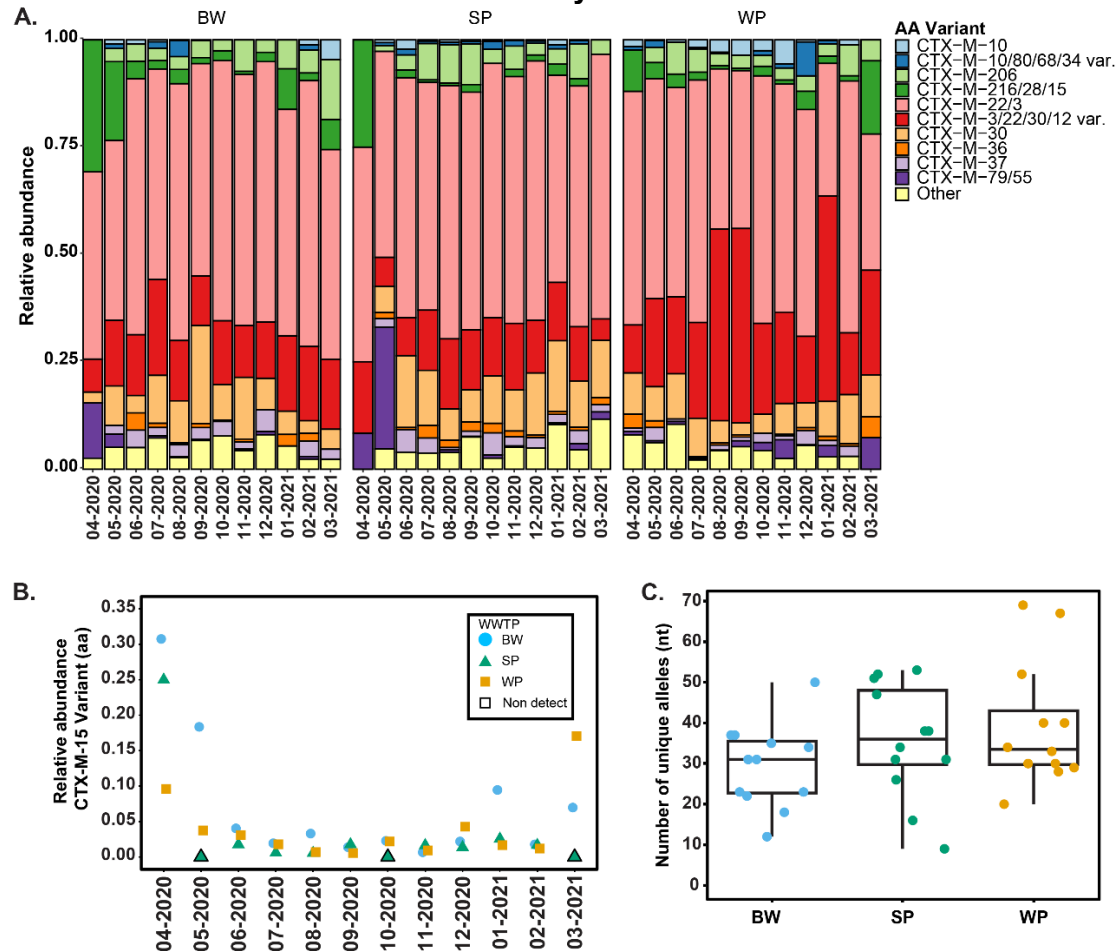


Figure 2: A.) Relative abundance of the top 10 most common protein variants across all samples. B.) Relative abundance of the CTX-M-15 containing protein variant (CTX-M-216/28/15), which has high clinical relevance, across all samples. C.) Number of unique nucleotide alleles by WWTP.

We observed a core composition of highly abundant CTX-M protein variants across all three WWTPs (**Figure 2A**). The CTX-M-22/3/211 nucleotide allele and CTX-M-22/3 protein variant were the most abundant regardless of time and WWTP (Figure 2A and Figure S3). CTX-

M-22 was also detected previously in a study of clinically significant (isolated from a sterile body site or a significant quantity of growth from nonsterile sites) *E. coli* isolates from pediatric patients in Seattle.⁴⁰ The second most abundant protein variant (CTX-M-3/22/30/12 var.) was novel, defined as those that did not map 100% to the CARD database, and one amino acid/5 bps away from CTX-M-22/3. Protein variants CTX-M-30 and CTX-M-206 were also frequently detected across all WWTPs and sampling dates (average relative abundance (RA) of 8.9 and 4.4%, respectively). In a previous study of municipal wastewater in Southern California using cloning and sanger sequencing, the authors also identified a high relative abundance of CTX-M-3 and CTX-M-30 gene alleles,⁴¹ suggesting that there may be established alleles of CTX-M in human populations or that are established in wastewater sewer systems.

The CTX-M-15 containing protein variant (CTX-M-15/216/28) was detected in all but three samples (SP in May 2020, October 2020, and March 2021). The relative abundance of the CTX-M-15/216/28 protein variant was as high as 30% for BW and 25% for SP in April 2020, measured by consensus sequences. For most samples, the relative abundance was less than 5% of consensus sequences (**Figure 2B**). While CTX-M-15 is globally disseminated, it is also locally significant due to its appearance in both clinical settings and in agriculture. The previously referenced study of clinically significant isolates from Seattle pediatric patients detected CTX-M-15 in 10 out of 49 *E. coli* isolates. The same study also found that most of the CTX-M-15 positive isolates were clones of a ST131 pandemic strain.⁴⁰ In another previous study in Washington state, CTX-M-15 was the most frequently detected allele (50 out of 99 isolates) in CTX-M positive *E. coli* isolates from dairy cattle.⁴²

There were no statistically significant differences between WWTPs in the number of unique alleles observed (Kruskal-Wallis p-value=0.40) (**Figure 2C**) or Shannon diversity index (p-value=0.44) (Figure S4) at either the nucleotide or amino acid level (Figure S5). Within each WWTP, there were no significant differences in number of unique alleles/variants or Shannon

217 diversity index between high and low flow (Figures S4 and S5). Each WWTP differs in the
 218 population served and land use based on the number of medical facilities, pastureland area, and
 219 industrial waste permits (Figure S6 and Table S5). Despite the different contributions to influent
 220 wastewater there were no differences in alpha diversity, which may indicate geographically
 221 consistent trends.

222 ***Novel Variants in Wastewater***

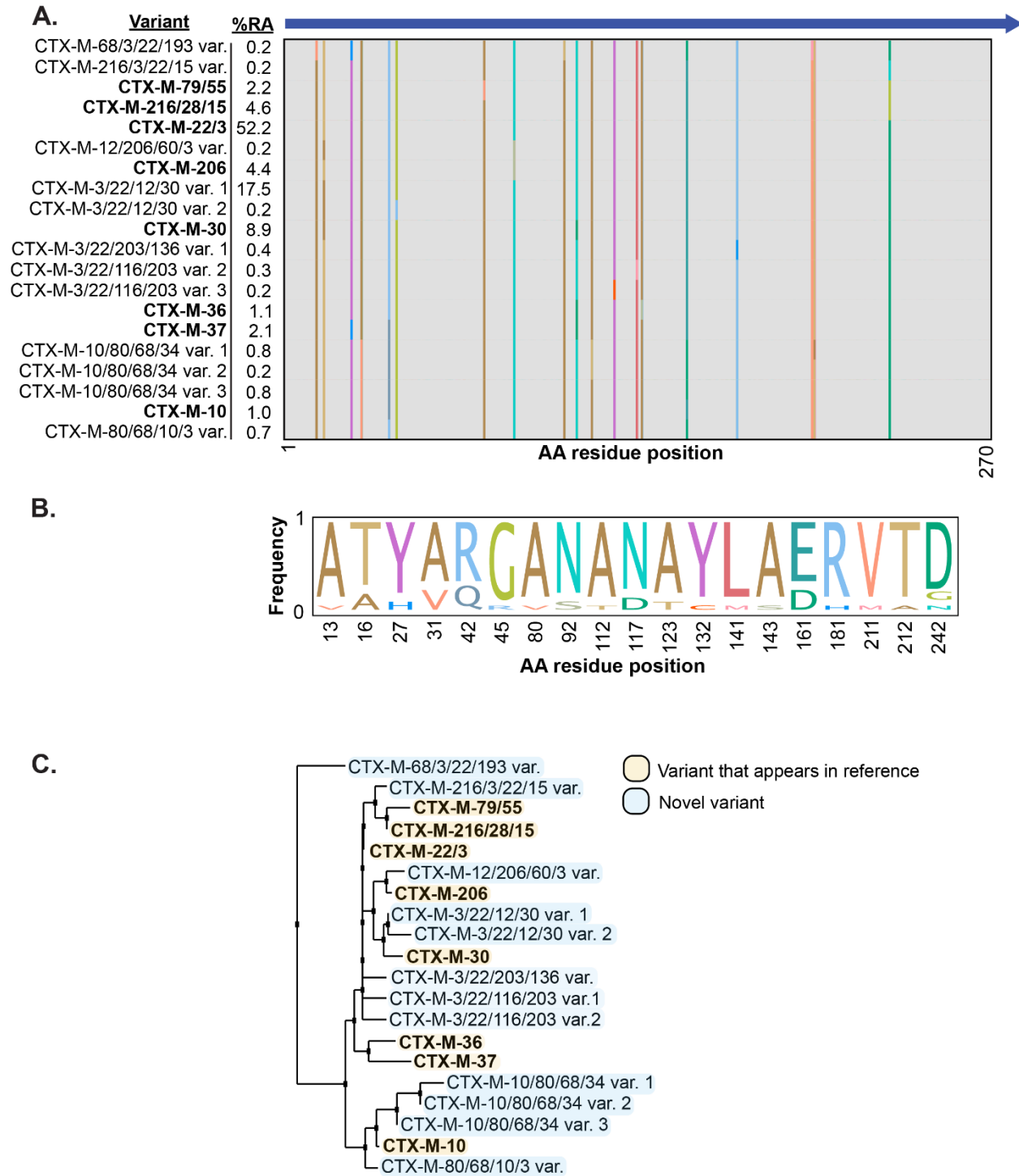


Figure 3: A.) Location of observed mutations across the top 20 protein variants (amino acid level) of the CTX-M gene. B.) Frequency of amino acid substitutions at loci with multiple residues based on unique variants and not abundance weighted C.) Phylogenetic tree of the top 20 protein variants using nearest neighbor joining.

The top 20 CTX-M protein variants, based on relative abundance, were aligned to visualize mutation hotspots (**Figure 3A**). Observed diversity of mutations were confined to 19 discrete loci on the CTX-M gene between residues 13-242 (**Figure 3B**). A subset of these observed mutations is associated with changes in CTX-M function.⁴³ For example, mutation D242G in CTX-M-15 has been implicated in conferring increased resistance to ceftazidime.⁴⁴ Residue D242 was previously characterized as a mutation hotspot in CTX-M-1 cluster variants.¹⁸ We identified two substitutions at this residue: 1.) D242G in CTX-M-79/55 and CTX-M-216/28/15 and 2.) D242N in CTX-M 216/3/22/15 var. Of the mutation hotspots in our work, many have been observed in CTX-M variants from clinical isolates (loci 13, 16, 31,80,112, 212, 242, 45,160),¹⁸ suggesting our method is capturing true allele diversity.

The consensus variant is the most abundant protein variant CTX-M-3/22 (average RA of 52.5%). Despite a larger diversity observed across the CTX-M gene at the nucleic acid level, most nucleotide mutations were synonymous – resulting in just two amino acid variants at all but one position, 242, where we observed three variants. Of these, the majority of amino acid substitutions are neutral (e.g., A13V, T16A) or conservative (e.g., Y27H, R42Q) as classified by BLOSUM62 scoring.⁴⁵ Three were unfavorable or non-conservative (G45R, Y132C, D242G). 12 of the top 20 occurring variants were found to be novel – yet closely match to CTX-M variants in CARD (**Figure 3C**). While many of our protein variants are not observed in CARD, they are observed in other studies of non-clinical isolates. For example, the second most abundant protein variant (CTX-M-3/22/30/12 var. 1) was also found in wastewater.⁴¹ Similarly, CTX-M-10/80/68/34 var. 1 was found in a *Kluyvera intermedia* isolated from an aquaculture study.⁴⁶ Our work highlights the challenges with using common ARG databases, populated with clinical isolates, when there is an abundance of diverse alleles found in nature that are not represented. Our method offers a rapid approach to surveil the diverse alleles in environmental samples.

Method Versatility

We demonstrate the utility of this method to characterize CTX-M in wastewater at a resolution that cannot be captured by short-read sequencing. In addition, UMI-based PCR does not necessitate bacterial cloning or culturing to characterize alleles. By coupling allele sequence information with longitudinal data, we can track how CTX-M alleles, and other ARGs-alike, change over time at the population level. Our method is versatile as it can be designed to capture allele clusters or a specific allele depending on the primer design. There is opportunity to multiplex a suite of ARGs for simultaneous allele characterization, including other globally important ESBL targets such as TEM and SHV.⁴⁷

Longitudinal tracking of alleles allows for the establishment of a baseline which can be further leveraged to identify emerging alleles at the population level. Portable allele and variant surveillance can be adapted for other settings such as hospital wastewater and low- and middle-income settings (LMICs). Surveilling hospital wastewater could aid in understanding the diversity of AMR and potentially inform targeted, clinical interventions.^{23,48} In areas with untreated wastewater, such as LMICs, understanding the discharged wastewater ARG alleles can help with public health interventions to reduce proliferation or exposure in the environment.^{49,50} In addition, by characterizing mutation hotspots within ARGs, we can potentially estimate mutations that could result in different phenotypic resistance.

Our study has some limitations. The UMI PCR primers used in this study primarily capture the CTX-M-1 cluster and were not able to identify all CTX-M variants including clinically relevant alleles that have been previously identified locally (e.g. CTX-M-27).⁴⁰ In addition, some alleles were grouped (e.g., CTX-M-15) because they cannot be distinguished from others (e.g., CTX-M-28 or CTX-M-216) due to single base pair mutations outside the reverse primer binding region. Primer design could be optimized with degenerate bases to detect a broader suite of alleles. Finally, we cannot distinguish between anthropogenic or environmental sources of variants. For example, biofilms and bacterial growth in sewage collection systems could contribute to observed

279 variants. There is a critical need to study the relative contributions from environmental versus
280 anthropogenic sources to AMR in wastewater influent.

281

282 **Acknowledgements**

283 We thank the local wastewater treatment plants for collecting samples and the UW Royalty
284 Research Fund for support.

References

1. Murray CJL, Ikuta KS, Sharara F, et al. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *The Lancet*. 2022;399(10325):629-655. doi:10.1016/S0140-6736(21)02724-0
2. Rawat D, Nair D. Extended-spectrum β -lactamases in Gram Negative Bacteria. *J Glob Infect Dis*. 2010;2(3):263-274. doi:10.4103/0974-777X.68531
3. Sexually Transmitted Infections | CDC Yellow Book 2024. Accessed March 14, 2024. <https://wwwnc.cdc.gov/travel/yellowbook/2024/posttravel-evaluation/sexually-transmitted-infections>
4. Arumugham VB, Gujarathi R, Cascella M. Third-Generation Cephalosporins. In: *StatPearls*. StatPearls Publishing; 2024. Accessed May 24, 2024. <http://www.ncbi.nlm.nih.gov/books/NBK549881/>
5. Tan TY, Ng LSY, He J, Hsu LY. CTX-M and *ampC* β -lactamases contributing to increased prevalence of ceftriaxone-resistant *Escherichia coli* in Changi General Hospital, Singapore. *Diagn Microbiol Infect Dis*. 2010;66(2):210-213. doi:10.1016/j.diagmicrobio.2009.08.019
6. Li WC, Huang FY, Liu CP, et al. Ceftriaxone Resistance of Nontyphoidal *Salmonella enterica* Isolates in Northern Taiwan Attributable to Production of CTX-M-14 and CMY-2 β -Lactamases. *J Clin Microbiol*. 2005;43(7):3237-3243. doi:10.1128/jcm.43.7.3237-3243.2005
7. Poirel L, Gniadkowski M, Nordmann P. Biochemical analysis of the ceftazidime-hydrolysing extended-spectrum β -lactamase CTX-M-15 and of its structurally related β -lactamase CTX-M-3. *J Antimicrob Chemother*. 2002;50(6):1031-1034. doi:10.1093/jac/dkf240
8. Karim A, Poirel L, Nagarajan S, Nordmann P. Plasmid-mediated extended-spectrum β -lactamase (CTX-M-3 like) from India and gene association with insertion sequence ISEcp1. *FEMS Microbiol Lett*. 2001;201(2):237-241. doi:10.1111/j.1574-6968.2001.tb10762.x
9. Bevan ER, Jones AM, Hawkey PM. Global epidemiology of CTX-M β -lactamases: temporal and geographical shifts in genotype. *J Antimicrob Chemother*. 2017;72(8):2145-2155. doi:10.1093/jac/dkx146
10. Saravanan M, Ramachandran B, Barabadi H. The prevalence and drug resistance pattern of extended spectrum β -lactamases (ESBLs) producing *Enterobacteriaceae* in Africa. *Microb Pathog*. 2018;114:180-192. doi:10.1016/j.micpath.2017.11.061
11. Nicolas-Chanoine MH, Blanco J, Leflon-Guibout V, et al. Intercontinental emergence of *Escherichia coli* clone O25:H4-ST131 producing CTX-M-15. *J Antimicrob Chemother*. 2008;61(2):273-281. doi:10.1093/jac/dkm464
12. Isgren CM, Edwards T, Pinchbeck GL, et al. Emergence of carriage of CTX-M-15 in faecal *Escherichia coli* in horses at an equine hospital in the UK; increasing prevalence over a decade (2008–2017). *BMC Vet Res*. 2019;15(1):268. doi:10.1186/s12917-019-2011-9
13. Abreu R, Castro B, Espigares E, et al. Prevalence of CTX-M-Type Extended-Spectrum β -Lactamases in *Escherichia coli* Strains Isolated in Poultry Farms. *Foodborne Pathog Dis*. 2014;11(11):868-873. doi:10.1089/fpd.2014.1796

14. Davis MA, Sisco WM, Jones LP, et al. Recent Emergence of *Escherichia coli* with Cephalosporin Resistance Conferred by blaCTX-M on Washington State Dairy Farms. *Appl Environ Microbiol*. 2015;81(13):4403-4410. doi:10.1128/AEM.00463-15
15. Hartmann A, Amoureux L, Locatelli A, et al. Occurrence of CTX-M Producing *Escherichia coli* in Soils, Cattle, and Farm Environment in France (Burgundy Region). *Front Microbiol*. 2012;3. doi:10.3389/fmicb.2012.00083
16. Amos GCA, Hawkey PM, Gaze WH, Wellington EM. Waste water effluent contributes to the dissemination of CTX-M-15 in the natural environment. *J Antimicrob Chemother*. 2014;69(7):1785-1791. doi:10.1093/jac/dku079
17. Yu K, Huang Z, Lan R, et al. Genomic Characterisation of Opportunistic Pathogen *Kluyvera* Reveals a Novel CTX-M Subgroup. *Microorganisms*. 2023;11(12):2836. doi:10.3390/microorganisms11122836
18. Mendonça J, Guedes C, Silva C, et al. New CTX-M Group Conferring β -Lactam Resistance: A Compendium of Phylogenetic Insights from Biochemical, Molecular, and Structural Biology. *Biology*. 2022;11(2):256. doi:10.3390/biology11020256
19. Li C, Zhang Z, Xu X, et al. Molecular Characterization of Cephalosporin-Resistant *Salmonella* Enteritidis ST11 Isolates Carrying blaCTX-M from Children with Diarrhea. *Foodborne Pathog Dis*. 2021;18(10):702-711. doi:10.1089/fpd.2020.2878
20. In Vivo Evolution of CTX-M-215, a Novel Narrow-Spectrum β -Lactamase in an *Escherichia coli* Clinical Isolate Conferring Resistance to Mecillinam | Antimicrobial Agents and Chemotherapy. Accessed April 29, 2024. <https://journals.asm.org/doi/full/10.1128/aac.00562-20>
21. Mao K, Zhang K, Du W, Ali W, Feng X, Zhang H. The potential of wastewater-based epidemiology as surveillance and early warning of infectious disease outbreaks. *Curr Opin Environ Sci Health*. 2020;17:1-7. doi:10.1016/j.coesh.2020.04.006
22. Prieto Riquelme MV, Garner E, Gupta S, et al. Demonstrating a Comprehensive Wastewater-Based Surveillance Approach That Differentiates Globally Sourced Resistomes. *Environ Sci Technol*. 2022;56(21):14982-14993. doi:10.1021/acs.est.1c08673
23. Davidova-Gerzova L, Lausova J, Sukkar I, et al. Hospital and community wastewater as a source of multidrug-resistant ESBL-producing *Escherichia coli*. *Front Cell Infect Microbiol*. 2023;13. Accessed February 13, 2024. <https://www.frontiersin.org/articles/10.3389/fcimb.2023.1184081>
24. Puljko A, Milaković M, Križanović S, et al. Prevalence of enteric opportunistic pathogens and extended-spectrum cephalosporin- and carbapenem-resistant coliforms and genes in wastewater from municipal wastewater treatment plants in Croatia. *J Hazard Mater*. 2022;427:128155. doi:10.1016/j.jhazmat.2021.128155
25. Kutilova I, Medvecký M, Leekitcharoenphon P, et al. Extended-spectrum beta-lactamase-producing *Escherichia coli* and antimicrobial resistance in municipal and hospital wastewaters in Czech Republic: Culture-based and metagenomic approaches. *Environ Res*. 2021;193:110487. doi:10.1016/j.envres.2020.110487

26. Bush SJ, Foster D, Eyre DW, et al. Genomic diversity affects the accuracy of bacterial single-nucleotide polymorphism-calling pipelines. *GigaScience*. 2020;9(2):giaa007. doi:10.1093/gigascience/giaa007
27. Vollmers J, Wiegand S, Kaster AK. Comparing and Evaluating Metagenome Assembly Tools from a Microbiologist's Perspective - Not Only Size Matters! *PLOS ONE*. 2017;12(1):e0169662. doi:10.1371/journal.pone.0169662
28. Karst SM, Ziels RM, Kirkegaard RH, et al. High-accuracy long-read amplicon sequences using unique molecular identifiers with Nanopore or PacBio sequencing. *Nat Methods*. 2021;18(2):165-169. doi:10.1038/s41592-020-01041-y
29. Philo SE, Keim EK, Swanstrom R, et al. A comparison of SARS-CoV-2 wastewater concentration methods for environmental surveillance. *Sci Total Environ*. 2021;760:144215. doi:10.1016/j.scitotenv.2020.144215
30. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*. 2011;17(1):10-12. doi:10.14806/ej.17.1.200
31. Lin X, Waring K, Tyson J, Ziels RM. High-accuracy meets high-throughput for microbiome profiling with near full-length 16S rRNA amplicon sequencing on the Nanopore platform. Published online June 19, 2023:2023.06.19.544637. doi:10.1101/2023.06.19.544637
32. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*. 2010;26(19):2460-2461. doi:10.1093/bioinformatics/btq461
33. Vaser R, Sović I, Nagarajan N, Šikić M. Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Res*. 2017;27(5):737-746. doi:10.1101/gr.214270.116
34. nanoporetech/medaka. Published online April 24, 2024. Accessed April 29, 2024. <https://github.com/nanoporetech/medaka>
35. Li H. Minimap2: pairwise alignment for nucleotide sequences. Birol I, ed. *Bioinformatics*. 2018;34(18):3094-3100. doi:10.1093/bioinformatics/bty191
36. McArthur AG, Waglechner N, Nizam F, et al. The Comprehensive Antibiotic Resistance Database. *Antimicrob Agents Chemother*. 2013;57(7):3348-3357. doi:10.1128/AAC.00419-13
37. R: The R Project for Statistical Computing. Accessed April 29, 2024. <https://www.r-project.org/>
38. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data | PLOS ONE. Accessed April 29, 2024. <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0061217>
39. Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics*. 2009;25(9):1189-1191. doi:10.1093/bioinformatics/btp033
40. Weissman SJ, Adler A, Qin X, Zerr DM. Emergence of extended-spectrum β -lactam resistance among *Escherichia coli* at a US academic children's hospital is clonal at the sequence type level for CTX-

393 M-15, but not for CMY-2. *Int J Antimicrob Agents*. 2013;41(5):414-420.
394 doi:10.1016/j.ijantimicag.2013.01.006

395 41. Borgogna TR, Borgogna JL, Mielke JA, et al. High Diversity of CTX-M Extended-Spectrum β -
396 Lactamases in Municipal Wastewater and Urban Wetlands. *Microb Drug Resist*. 2016;22(4):312-320.
397 doi:10.1089/mdr.2015.0197

398 42. Afema JA, Ahmed S, Besser TE, Jones LP, Sisco WM, Davis MA. Molecular Epidemiology of Dairy
399 Cattle-Associated Escherichia coli Carrying blaCTX-M Genes in Washington State. *Appl Environ Microbiol*.
400 2018;84(6):e02430-17. doi:10.1128/AEM.02430-17

401 43. Pérez-Llarena FJ, Kerff F, Abián O, et al. Distant and New Mutations in CTX-M-1 β -Lactamase
402 Affect Cefotaxime Hydrolysis ∇ . *Antimicrob Agents Chemother*. 2011;55(9):4361-4368.
403 doi:10.1128/AAC.00298-11

404 44. Novais Â, Comas I, Baquero F, et al. Evolutionary Trajectories of Beta-Lactamase CTX-M-1 Cluster
405 Enzymes: Predicting Antibiotic Resistance. *PLOS Pathog*. 2010;6(1):e1000735.
406 doi:10.1371/journal.ppat.1000735

407 45. Henikoff S, Henikoff JG. Amino acid substitution matrices from protein blocks. *Proc Natl Acad Sci*
408 *U S A*. 1992;89(22):10915-10919.

409 46. Salgueiro V, Manageiro V, Rosado T, et al. Snapshot of resistome, virulome and mobilome in
410 aquaculture. *Sci Total Environ*. 2023;905:166351. doi:10.1016/j.scitotenv.2023.166351

411 47. Castanheira M, Simner PJ, Bradford PA. Extended-spectrum β -lactamases: an update on their
412 characteristics, epidemiology and detection. *JAC-Antimicrob Resist*. 2021;3(3):dlab092.
413 doi:10.1093/jacamr/dlab092

414 48. Parkins MD, Lee BE, Acosta N, et al. Wastewater-based surveillance as a tool for public health
415 action: SARS-CoV-2 and beyond. *Clin Microbiol Rev*. 2023;37(1):e00103-22. doi:10.1128/cmr.00103-22

416 49. Lamba M, Graham DW, Ahammad SZ. Hospital Wastewater Releases of Carbapenem-Resistance
417 Pathogens and Genes in Urban India. *Environ Sci Technol*. 2017;51(23):13906-13912.
418 doi:10.1021/acs.est.7b03380

419 50. Bougnom BP, Zongo C, McNally A, et al. Wastewater used for urban agriculture in West Africa as
420 a reservoir for antibacterial resistance dissemination. *Environ Res*. 2019;168:14-24.
421 doi:10.1016/j.envres.2018.09.022
422