

1 Targeted Sequencing of CTX-M Alleles in Seattle Area Wastewater

2 Angelo Q. Ong<sup>1</sup>, Sarah E. Philo<sup>1</sup>, Anysiah Taylor<sup>2</sup>, Ruohan Hu<sup>2</sup>, John Scott Meschke<sup>1</sup>, Erica R.

3 Fuhrmeister<sup>1,3\*</sup>

4

5 <sup>1</sup>Department of Environmental and Occupational Health Sciences, University of Washington,

6 Seattle, Seattle, WA, USA

7 <sup>2</sup>Department of Global Health, University of Washington, Seattle, Seattle, WA, USA

8 <sup>3</sup>Civil and Environmental Engineering, University of Washington, Seattle, Seattle, WA, USA

9

10 \*Correspondence to:

11 Dr. Erica Fuhrmeister, [efuhrm@uw.edu](mailto:efuhrm@uw.edu)

12 4225 Roosevelt Way

13 Seattle, WA 98105

14

15

16

17

18

19

20

21

22

23 **Abstract**

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

40

41

42

## 43 Introduction

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

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

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

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

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

## 87 **Methods**

### 88 ***Wastewater collection, concentration, and extraction***

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

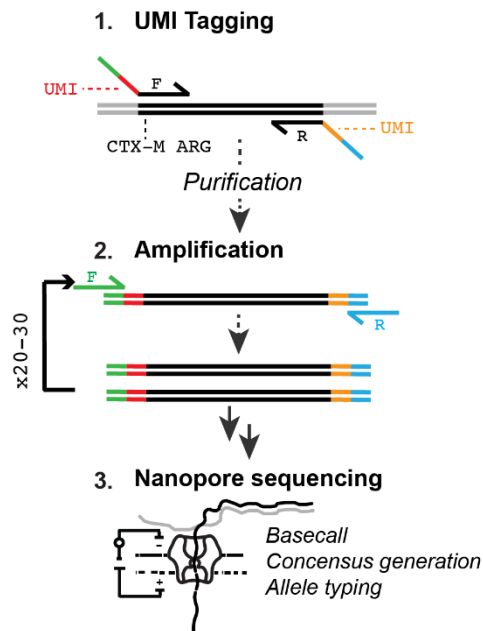
93 mixed well prior to concentration. For the other WWTPs, one grab sample was collected for a  
94 single day. All samples were processed and concentrated within one week of the collection date.  
95 0.5 L of primary influent wastewater was concentrated using 5% skimmed milk, shaken,  
96 centrifuged, and pellets resuspended in 6 mL of 1X PBS. DNA and RNA were coextracted from  
97 280  $\mu$ L of concentrated sample using a QIAmp Viral RNA Mini extraction kit (Qiagen) with no  
98 DNase treatment. One sample from the first week of each month for 12 months was analyzed for  
99 CTX-M variants across the three WWTPs (36 samples total).

### 100 ***Gene Quantification***

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

### 111 ***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  $\mu$ L (qPCR  
119 results in Figure S1). The mastermix consisted of 0.2  $\mu$ 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  $\mu$ M forward and reverse primers, 0.5 mM  
125 dNTPs, 1X SuperFI II buffer, 1U SuperFI II polymerase, and molecular grade water. 20  $\mu$ L of the  
126 previous PCR product was used as input DNA for a total reaction volume of 50  $\mu$ L. PCR products  
127 were loaded with 6X TriTrack DNA Loading Dye (ThermoScientific) and amplification was

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

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

### 139 ***DNA Sequencing***

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

### 145 ***Data Analysis***

146 Fast5 files were basecalled with guppy (v6.5.7) using the R10.4.1 super accuracy model  
147 (dna\_r10.4.1\_e8.2\_400bps\_sup.cfg) with a minimum quality score of 9. Basecalled reads were  
148 demultiplexed with cutadapt (v2.6)<sup>30</sup> to search for forward and reverse barcodes at each read end  
149 with the following parameters: an error rate of 0.15, a minimum length of 800 bps, and maximum  
150 length of 1100 bps. Demultiplexed reads were run through an existing ssUMI pipeline (v0.3.2)<sup>28,31</sup>  
151 to produce consensus sequences based on dual UMIs. The longread\_umi ssumi\_std command

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

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

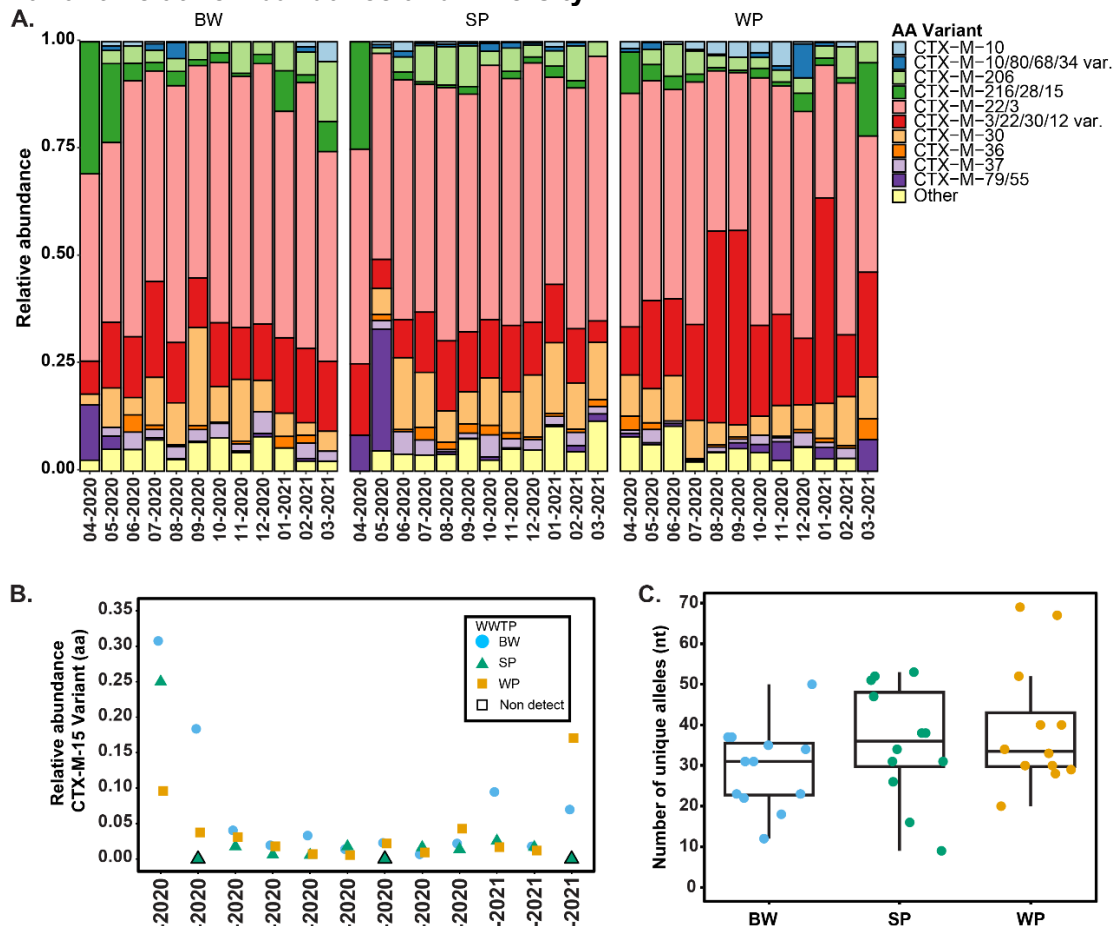
169 To test for differences in alpha diversity in WWTPs, we used a Kruskal-Wallis test to  
170 compare the median number of unique alleles and Shannon diversity indices. Flow values,  
171 measured in millions of gallons per day (MGD) at each sample date were used to categorize  
172 samples as either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



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

## 182 Results and Discussion

### 183 Variant Relative Abundance and Diversity



184 **Figure 2:** A.) Relative abundance of the top 10 most common protein variants across all samples. B.)  
 185 Relative abundance of the CTX-M-15 containing protein variant (CTX-M-216/28/15), which has high clinical  
 186 relevance, across all samples. C.) Number of unique nucleotide alleles by WWTP.  
 187  
 188

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

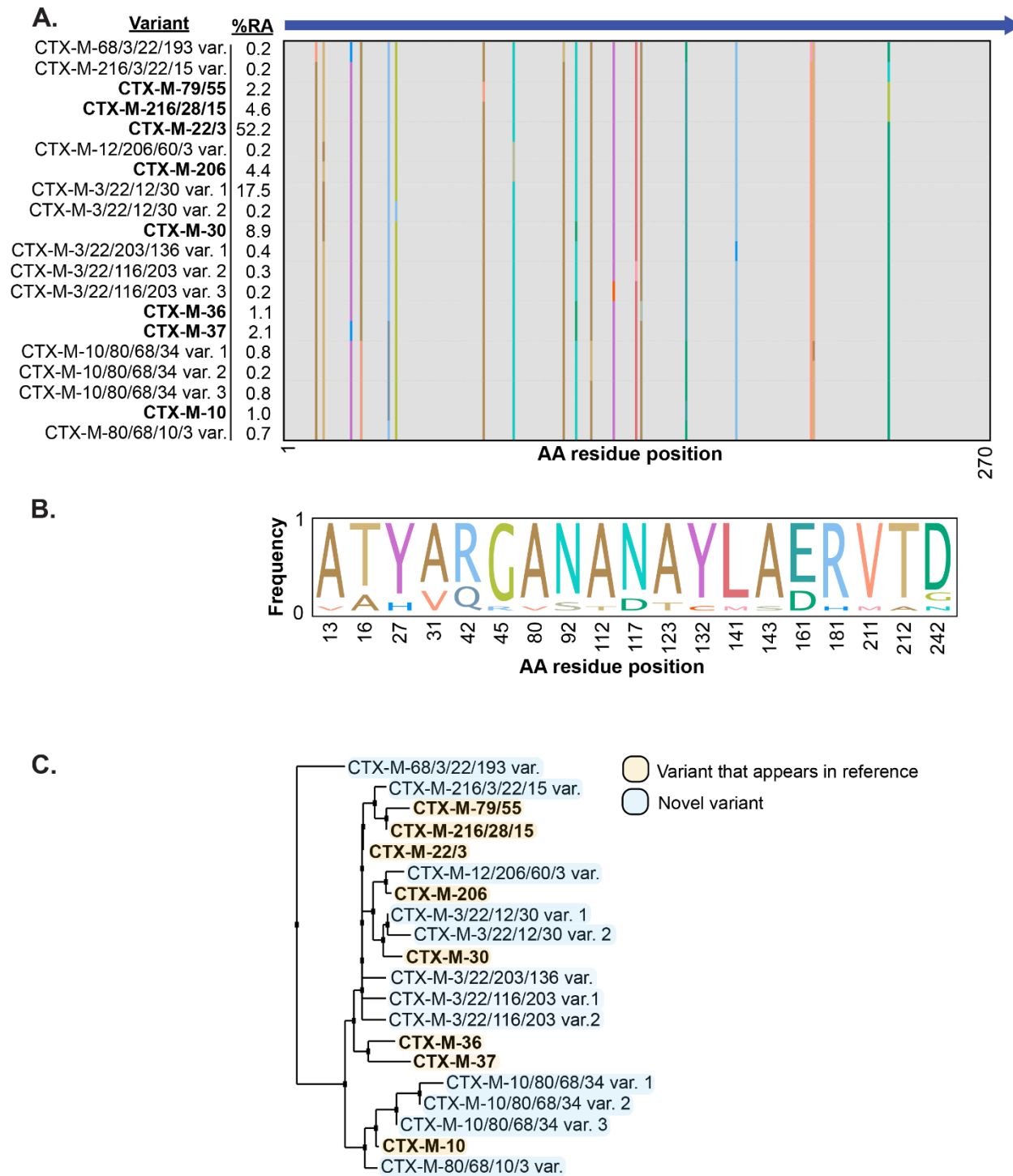
192 M-22 was also detected previously in a study of clinically significant (isolated from a sterile body  
193 site or a significant quantity of growth from nonsterile sites) *E. coli* isolates from pediatric patients  
194 in Seattle.<sup>40</sup> The second most abundant protein variant (CTX-M-3/22/30/12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

202 The CTX-M-15 containing protein variant (CTX-M-15/216/28) was detected in all but three  
203 samples (SP in May 2020, October 2020, and March 2021). The relative abundance of the CTX-  
204 M-15/216/28 protein variant was as high as 30% for BW and 25% for SP in April 2020, measured  
205 by consensus sequences. For most samples, the relative abundance was less than 5% of  
206 consensus sequences (**Figure 2B**). While CTX-M-15 is globally disseminated, it is also locally  
207 significant due to its appearance in both clinical settings and in agriculture. The previously  
208 referenced study of clinically significant isolates from Seattle pediatric patients detected CTX-M-  
209 15 in 10 out of 49 *E. coli* isolates. The same study also found that most of the CTX-M-15 positive  
210 isolates were clones of a ST131 pandemic strain.<sup>40</sup> In another previous study in Washington state,  
211 CTX-M-15 was the most frequently detected allele (50 out of 99 isolates) in CTX-M positive *E.*  
212 *coli* isolates from dairy cattle.<sup>42</sup>

213 There were no statistically significant differences between WWTPs in the number of  
214 unique alleles observed (Kruskal-Wallis p-value=0.40) (**Figure 2C**) or Shannon diversity index (p-  
215 value=0.44) (Figure S4) at either the nucleotide or amino acid level (Figure S5). Within each  
216 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***



223

224 **Figure 3:** A.) Location of observed mutations across the top 20 protein variants (amino acid level) of the  
 225 CTX-M gene. B.) Frequency of amino acid substitutions at loci with multiple residues based on unique  
 226 variants and not abundance weighted C.) Phylogenetic tree of the top 20 protein variants using nearest  
 227 neighbor joining.  
 228

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

239 The consensus variant is the most abundant protein variant CTX-M-3/22 (average RA of  
240 52.5%). Despite a larger diversity observed across the CTX-M gene at the nucleic acid level, most  
241 nucleotide mutations were synonymous – resulting in just two amino acid variants at all but one  
242 position, 242, where we observed three variants. Of these, the majority of amino acid substitutions  
243 are neutral (e.g., A13V, T16A) or conservative (e.g., Y27H, R42Q) as classified by BLOSUM62  
244 scoring.<sup>45</sup> Three were unfavorable or non-conservative (G45R, Y132C, D242G). 12 of the top 20  
245 occurring variants were found to be novel – yet closely match to CTX-M variants in CARD (**Figure**  
246 **3C**). While many of our protein variants are not observed in CARD, they are observed in other  
247 studies of non-clinical isolates. For example, the second most abundant protein variant (CTX-M-  
248 3/22/30/12 var. 1) was also found in wastewater.<sup>41</sup> Similarly, CTX-M-10/80/68/34 var. 1 was found  
249 in a *Kluyvera intermedia* isolated from an aquaculture study.<sup>46</sup> Our work highlights the challenges  
250 with using common ARG databases, populated with clinical isolates, when there is an abundance  
251 of diverse alleles found in nature that are not represented. Our method offers a rapid approach to  
252 surveil the diverse alleles in environmental samples.

### 253 **Method Versatility**

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

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

271 Our study has some limitations. The UMI PCR primers used in this study primarily capture  
272 the CTX-M-1 cluster and were not able to identify all CTX-M variants including clinically relevant  
273 alleles that have been previously identified locally (e.g. CTX-M-27).<sup>40</sup> In addition, some alleles  
274 were grouped (e.g., CTX-M-15) because they cannot be distinguished from others (e.g., CTX-M-  
275 28 or CTX-M-216) due to single base pair mutations outside the reverse primer binding region.  
276 Primer design could be optimized with degenerate bases to detect a broader suite of alleles.  
277 Finally, we cannot distinguish between anthropogenic or environmental sources of variants. For  
278 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.

## 285 References

- 286 1. Murray CJL, Ikuta KS, Sharara F, et al. Global burden of bacterial antimicrobial resistance in 2019:  
287 a systematic analysis. *The Lancet*. 2022;399(10325):629-655. doi:10.1016/S0140-6736(21)02724-0
- 288 2. Rawat D, Nair D. Extended-spectrum  $\beta$ -lactamases in Gram Negative Bacteria. *J Glob Infect Dis*.  
289 2010;2(3):263-274. doi:10.4103/0974-777X.68531
- 290 3. Sexually Transmitted Infections | CDC Yellow Book 2024. Accessed March 14, 2024.  
291 <https://wwwnc.cdc.gov/travel/yellowbook/2024/posttravel-evaluation/sexually-transmitted-infections>
- 292 4. Arumugham VB, Gujarathi R, Cascella M. Third-Generation Cephalosporins. In: *StatPearls*.  
293 StatPearls Publishing; 2024. Accessed May 24, 2024. <http://www.ncbi.nlm.nih.gov/books/NBK549881/>
- 294 5. Tan TY, Ng LSY, He J, Hsu LY. CTX-M and *ampC*  $\beta$ -lactamases contributing to increased prevalence  
295 of ceftriaxone-resistant *Escherichia coli* in Changi General Hospital, Singapore. *Diagn Microbiol Infect Dis*.  
296 2010;66(2):210-213. doi:10.1016/j.diagmicrobio.2009.08.019
- 297 6. Li WC, Huang FY, Liu CP, et al. Ceftriaxone Resistance of Nontyphoidal *Salmonella enterica*  
298 Isolates in Northern Taiwan Attributable to Production of CTX-M-14 and CMY-2  $\beta$ -Lactamases. *J Clin*  
299 *Microbiol*. 2005;43(7):3237-3243. doi:10.1128/jcm.43.7.3237-3243.2005
- 300 7. Poirel L, Gniadkowski M, Nordmann P. Biochemical analysis of the ceftazidime-hydrolysing  
301 extended-spectrum  $\beta$ -lactamase CTX-M-15 and of its structurally related  $\beta$ -lactamase CTX-M-3. *J*  
302 *Antimicrob Chemother*. 2002;50(6):1031-1034. doi:10.1093/jac/dkf240
- 303 8. Karim A, Poirel L, Nagarajan S, Nordmann P. Plasmid-mediated extended-spectrum  $\beta$ -lactamase  
304 (CTX-M-3 like) from India and gene association with insertion sequence ISEcp1. *FEMS Microbiol Lett*.  
305 2001;201(2):237-241. doi:10.1111/j.1574-6968.2001.tb10762.x
- 306 9. Bevan ER, Jones AM, Hawkey PM. Global epidemiology of CTX-M  $\beta$ -lactamases: temporal and  
307 geographical shifts in genotype. *J Antimicrob Chemother*. 2017;72(8):2145-2155. doi:10.1093/jac/dkx146
- 308 10. Saravanan M, Ramachandran B, Barabadi H. The prevalence and drug resistance pattern of  
309 extended spectrum  $\beta$ -lactamases (ESBLs) producing *Enterobacteriaceae* in Africa. *Microb Pathog*.  
310 2018;114:180-192. doi:10.1016/j.micpath.2017.11.061
- 311 11. Nicolas-Chanoine MH, Blanco J, Leflon-Guibout V, et al. Intercontinental emergence of  
312 *Escherichia coli* clone O25:H4-ST131 producing CTX-M-15. *J Antimicrob Chemother*. 2008;61(2):273-281.  
313 doi:10.1093/jac/dkm464
- 314 12. Isgren CM, Edwards T, Pinchbeck GL, et al. Emergence of carriage of CTX-M-15 in faecal  
315 *Escherichia coli* in horses at an equine hospital in the UK; increasing prevalence over a decade (2008–  
316 2017). *BMC Vet Res*. 2019;15(1):268. doi:10.1186/s12917-019-2011-9
- 317 13. Abreu R, Castro B, Espigares E, et al. Prevalence of CTX-M-Type Extended-Spectrum  $\beta$ -  
318 Lactamases in *Escherichia coli* Strains Isolated in Poultry Farms. *Foodborne Pathog Dis*. 2014;11(11):868-  
319 873. doi:10.1089/fpd.2014.1796



- 320 14. Davis MA, Sisco WM, Jones LP, et al. Recent Emergence of *Escherichia coli* with Cephalosporin  
321 Resistance Conferred by blaCTX-M on Washington State Dairy Farms. *Appl Environ Microbiol.*  
322 2015;81(13):4403-4410. doi:10.1128/AEM.00463-15
- 323 15. Hartmann A, Amoureux L, Locatelli A, et al. Occurrence of CTX-M Producing *Escherichia coli* in  
324 Soils, Cattle, and Farm Environment in France (Burgundy Region). *Front Microbiol.* 2012;3.  
325 doi:10.3389/fmicb.2012.00083
- 326 16. Amos GCA, Hawkey PM, Gaze WH, Wellington EM. Waste water effluent contributes to the  
327 dissemination of CTX-M-15 in the natural environment. *J Antimicrob Chemother.* 2014;69(7):1785-1791.  
328 doi:10.1093/jac/dku079
- 329 17. Yu K, Huang Z, Lan R, et al. Genomic Characterisation of Opportunistic Pathogen *Kluyvera* Reveals  
330 a Novel CTX-M Subgroup. *Microorganisms.* 2023;11(12):2836. doi:10.3390/microorganisms11122836
- 331 18. Mendonça J, Guedes C, Silva C, et al. New CTX-M Group Conferring  $\beta$ -Lactam Resistance: A  
332 Compendium of Phylogenetic Insights from Biochemical, Molecular, and Structural Biology. *Biology.*  
333 2022;11(2):256. doi:10.3390/biology11020256
- 334 19. Li C, Zhang Z, Xu X, et al. Molecular Characterization of Cephalosporin-Resistant *Salmonella*  
335 Enteritidis ST11 Isolates Carrying blaCTX-M from Children with Diarrhea. *Foodborne Pathog Dis.*  
336 2021;18(10):702-711. doi:10.1089/fpd.2020.2878
- 337 20. In Vivo Evolution of CTX-M-215, a Novel Narrow-Spectrum  $\beta$ -Lactamase in an *Escherichia coli*  
338 Clinical Isolate Conferring Resistance to Mecillinam | Antimicrobial Agents and Chemotherapy. Accessed  
339 April 29, 2024. <https://journals.asm.org/doi/full/10.1128/aac.00562-20>
- 340 21. Mao K, Zhang K, Du W, Ali W, Feng X, Zhang H. The potential of wastewater-based epidemiology  
341 as surveillance and early warning of infectious disease outbreaks. *Curr Opin Environ Sci Health.*  
342 2020;17:1-7. doi:10.1016/j.coesh.2020.04.006
- 343 22. Prieto Riquelme MV, Garner E, Gupta S, et al. Demonstrating a Comprehensive Wastewater-  
344 Based Surveillance Approach That Differentiates Globally Sourced Resistomes. *Environ Sci Technol.*  
345 2022;56(21):14982-14993. doi:10.1021/acs.est.1c08673
- 346 23. Davidova-Gerzova L, Lausova J, Sukkar I, et al. Hospital and community wastewater as a source of  
347 multidrug-resistant ESBL-producing *Escherichia coli*. *Front Cell Infect Microbiol.* 2023;13. Accessed  
348 February 13, 2024. <https://www.frontiersin.org/articles/10.3389/fcimb.2023.1184081>
- 349 24. Puljko A, Milaković M, Križanović S, et al. Prevalence of enteric opportunistic pathogens and  
350 extended-spectrum cephalosporin- and carbapenem-resistant coliforms and genes in wastewater from  
351 municipal wastewater treatment plants in Croatia. *J Hazard Mater.* 2022;427:128155.  
352 doi:10.1016/j.jhazmat.2021.128155
- 353 25. Kutilova I, Medvecký M, Leekitcharoenphon P, et al. Extended-spectrum beta-lactamase-  
354 producing *Escherichia coli* and antimicrobial resistance in municipal and hospital wastewaters in Czech  
355 Republic: Culture-based and metagenomic approaches. *Environ Res.* 2021;193:110487.  
356 doi:10.1016/j.envres.2020.110487

- 357 26. Bush SJ, Foster D, Eyre DW, et al. Genomic diversity affects the accuracy of bacterial single-  
358 nucleotide polymorphism–calling pipelines. *GigaScience*. 2020;9(2):giaa007.  
359 doi:10.1093/gigascience/giaa007
- 360 27. Vollmers J, Wiegand S, Kaster AK. Comparing and Evaluating Metagenome Assembly Tools from a  
361 Microbiologist’s Perspective - Not Only Size Matters! *PLOS ONE*. 2017;12(1):e0169662.  
362 doi:10.1371/journal.pone.0169662
- 363 28. Karst SM, Ziels RM, Kirkegaard RH, et al. High-accuracy long-read amplicon sequences using  
364 unique molecular identifiers with Nanopore or PacBio sequencing. *Nat Methods*. 2021;18(2):165-169.  
365 doi:10.1038/s41592-020-01041-y
- 366 29. Philo SE, Keim EK, Swanstrom R, et al. A comparison of SARS-CoV-2 wastewater concentration  
367 methods for environmental surveillance. *Sci Total Environ*. 2021;760:144215.  
368 doi:10.1016/j.scitotenv.2020.144215
- 369 30. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads.  
370 *EMBnet.journal*. 2011;17(1):10-12. doi:10.14806/ej.17.1.200
- 371 31. Lin X, Waring K, Tyson J, Ziels RM. High-accuracy meets high-throughput for microbiome profiling  
372 with near full-length 16S rRNA amplicon sequencing on the Nanopore platform. Published online June  
373 19, 2023:2023.06.19.544637. doi:10.1101/2023.06.19.544637
- 374 32. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*.  
375 2010;26(19):2460-2461. doi:10.1093/bioinformatics/btq461
- 376 33. Vaser R, Sović I, Nagarajan N, Šikić M. Fast and accurate de novo genome assembly from long  
377 uncorrected reads. *Genome Res*. 2017;27(5):737-746. doi:10.1101/gr.214270.116
- 378 34. nanoporetech/medaka. Published online April 24, 2024. Accessed April 29, 2024.  
379 <https://github.com/nanoporetech/medaka>
- 380 35. Li H. Minimap2: pairwise alignment for nucleotide sequences. Birol I, ed. *Bioinformatics*.  
381 2018;34(18):3094-3100. doi:10.1093/bioinformatics/bty191
- 382 36. McArthur AG, Waglechner N, Nizam F, et al. The Comprehensive Antibiotic Resistance Database.  
383 *Antimicrob Agents Chemother*. 2013;57(7):3348-3357. doi:10.1128/AAC.00419-13
- 384 37. R: The R Project for Statistical Computing. Accessed April 29, 2024. <https://www.r-project.org/>
- 385 38. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome  
386 Census Data | PLOS ONE. Accessed April 29, 2024.  
387 <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0061217>
- 388 39. Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. Jalview Version 2—a multiple  
389 sequence alignment editor and analysis workbench. *Bioinformatics*. 2009;25(9):1189-1191.  
390 doi:10.1093/bioinformatics/btp033
- 391 40. Weissman SJ, Adler A, Qin X, Zerr DM. Emergence of extended-spectrum  $\beta$ -lactam resistance  
392 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  $\beta$ -  
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, Sicho 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  $\beta$ -Lactamase  
402 Affect Cefotaxime Hydrolysis  $\nabla$ . *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  $\beta$ -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