

1 **Harnessing non-standard nucleic acids for highly sensitive icosaplex (20-plex) detection of**
2 **microbial threats**

3
4 Hinako Kawabe¹, Luran Manfio², Sebastian Magana Pena², Nicolette A. Zhou³, Kevin M.
5 Bradley^{2,4}, Cen Chen^{2,4}, Chris McLendon⁴, Steven A. Benner^{2,4}, Karen Levy³, Zunyi Yang^{2,4*},
6 Jorge A. Marchand^{1,5*}, Erica R. Fuhrmeister^{3,6*}

7 ¹Chemical Engineering, University of Washington, Seattle, WA, 98195, USA

8 ²Foundation for Applied Molecular Evolution (FfAME), 13709 Progress Blvd, Alachua, FL
9 32615, USA

10 ³Department of Environmental and Occupational Health Sciences, University of Washington,
11 Seattle, Seattle, WA, 98195, USA

12 ⁴Firebird Biomolecular Sciences LLC, 13709 Progress Blvd, Box 17, Alachua, FL 32615, USA

13 ⁵Molecular Engineering and Science Institute, University of Washington, Seattle, Seattle, WA,
14 98195, USA

15 ⁶Civil and Environmental Engineering, University of Washington, Seattle, Seattle, WA, 98195,
16 USA

17
18 *co-corresponding authors

19 Zunyi Yang, zyang@ffame.org

20 Jorge A. Marchand, jmarcha@uw.edu

21 Erica R. Fuhrmeister, efuhrm@uw.edu

22 **Abstract:** Environmental surveillance and clinical diagnostics heavily rely on the polymerase
23 chain reaction (PCR) for target detection. A growing list of microbial threats warrants new PCR-
24 based detection methods that are highly sensitive, specific, and multiplexable. Here, we
25 introduce a PCR-based icosaplex (20-plex) assay for detecting 18 enteropathogen and two
26 antimicrobial resistance genes. This multiplexed PCR assay leverages the self-avoiding
27 molecular recognition system (SAMRS) to avoid primer dimer formation, the artificially
28 expanded genetic information system (AEGIS) for amplification specificity, and next-generation
29 sequencing for amplicon identification. We benchmarked this assay using a low-cost, portable
30 sequencing platform (Oxford Nanopore) on wastewater, soil, and human stool samples. Using
31 parallelized multi-target TaqMan Array Cards (TAC) to benchmark performance of the 20-plex
32 assay, there was 74% agreement on positive calls and 97% agreement on negative calls.
33 Additionally, we show how sequencing information from the 20-plex can be used to further
34 classify allelic variants of genes and distinguish sub-species. The strategy presented offers
35 sensitive, affordable, and robust multiplex detection that can be used to support efforts in
36 wastewater-based epidemiology, environmental monitoring, and human/animal diagnostics.

37

38 **Keywords:** Enteric pathogens, wastewater-based epidemiology, multiplex PCR, SAMRS,
39 AEGIS, next-generation sequencing

40

41

42

43

44

45 Introduction

46 Emerging infectious diseases, coupled with rising antibiotic resistance, are a threat to global
47 public health.¹ The diversity of pathogens that can cause illness necessitates pathogen detection
48 methods that can identify multiple genetic targets. There is a need for low-cost, multi-target,
49 detection methods, especially in regions where the burden of infectious diseases is high,
50 resources are constrained, and there is a high diversity in the pathogens that are present.

51 A high burden of diarrheal illness exists in low- and middle-income countries (LMICs).² The
52 types of enteric pathogens (bacteria, virus, protozoa, helminths) present in LMICs contributing to
53 disease are geographically diverse and location specific.³⁻⁵ As an example of geographic
54 diversity, previous studies used highly parallelized quantitative polymerase chain reaction
55 (qPCR) to survey a wide range of pathogens around the world.⁶ In Mozambique, *Shigella* spp.
56 and *Giardia* spp. were the most prevalent pathogens whereas *Campylobacter* spp. and *Giardia*
57 spp. dominated infections in children across eight other settings in South America, sub-Saharan
58 Africa, and Asia.^{3,5}

59 The need for multi-target detection assays is not limited to LMICs. In high-income countries
60 (HICs), respiratory illnesses are common and diarrheal illnesses predominate through foodborne
61 outbreaks.^{7,8} Characteristically, urban areas in HICs have sewerage sanitation systems, which
62 allow for active monitoring of the infectious disease burden through wastewater-based
63 epidemiology (WBE).⁹ As an early example, Poliovirus was isolated from wastewater samples as
64 part of the World Health Organization's Global Polio Eradication Initiative (GPEI) to monitor
65 for emergence/reemergence of the virus.¹⁰⁻¹² Recently, SARS-CoV-2 was monitored in
66 wastewater at $\approx 14,000$ sites in 59 countries in March 2021.¹³ These latest efforts rely on methods
67 such as qPCR and digital PCR (dPCR) to determine the level of infections in a population.¹⁴⁻¹⁶

68 Looking beyond the COVID-19 pandemic, monitoring programs are eager to expand
69 surveillance to include enteric pathogens, other respiratory viruses, sexually transmitted
70 infections, arboviruses, and antimicrobial resistance.¹⁷ However, expanding the range of targets
71 for surveillance is both labor-intensive and costly, underscoring the urgent need for new multi-
72 target capabilities that can efficiently and affordably address this challenge.

73 Complementing WBE efforts, other environmental surveillance strategies are premised on
74 the environment (e.g., soil, water, air, fomites) serving as an intermediary between infected
75 hosts.¹⁸ Environmental detection can facilitate surveying disease burden. For example, a study in
76 Kenya found that positive detection of helminths (*Ascaris lumbricoides*, *Trichuris trichiura*, and
77 *Necator americanus*) in a household's soil was significantly associated with cases of helminth
78 infections of household members.¹⁹ As a result, qPCR-based surveillance of helminths in soil is a
79 promising alternative to more invasive stool-based surveillance.

80 Environmental detection is also used for determining dominant transmission pathways of
81 pathogens. During the COVID-19 pandemic, studies sampled fomites using qPCR for SARS-
82 CoV-2 RNA and concluded that fomites were unlikely to be a dominant transmission
83 pathway.^{20,21} Similar methods have been proposed to survey the burden of antimicrobial
84 resistance across the globe by sampling soil and water.²²⁻²⁴ To expand the scale and scope of
85 environmental surveillance, highly multiplexed detection assays that can capture the diversity of
86 possible microbial threats of interest are needed. Despite this pressing need, methods for multi-
87 target detection generally fall short in achieving meaningful reductions in assay cost,
88 necessitating a more selective approach in deciding which targets to prioritize in monitoring,
89 surveillance, and detection efforts.

90 PCR-based multi-target detection strategies are typically limited by the number of targets that
91 can be simultaneously amplified and identified. In conventional multiplexed PCR amplification,

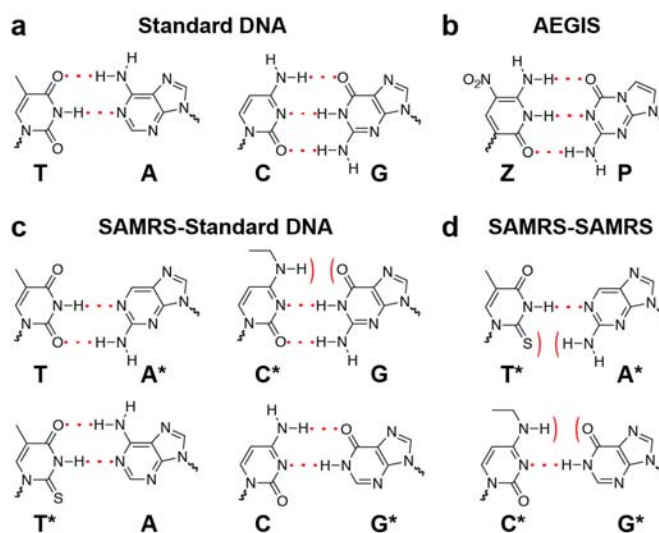
92 increasing the number of targets increases the likelihood of off-target reactions (e.g., primer
93 dimer formation, non-specific amplification product). These off-target reactions result in reduced
94 sensitivity or assay failure.²⁵ Further, highly multiplexed PCR assays tend to have a narrow
95 tolerance for changing reaction conditions and sample composition. For example, adding primers
96 for new targets to an established multiplexed assay can result in primer cascading failure.
97 Performing an assay in highly heterogeneous sample matrices, such as environmental DNA
98 extracts, can also reduce PCR efficiency and result in false negatives. Additionally, multiplexed
99 fluorescence-based detection assays, such as qPCR, are typically constrained (up to 5 targets) by
100 the limited orthogonality of fluorescent reporter dye spectra.²⁶

101 Other detection technologies have approached the ‘many-target’ problem by scaling down
102 reaction volumes and parallelizing reactions in microfluidic devices.²⁷ One widely used
103 commercial platform, the TaqManTM Array Card (TAC), parallelizes qPCR reactions into micro-
104 scale ($\approx 1.5 \mu\text{L}$) reactions.²⁸ Even with this compartmentalization, qPCR and dPCR-based
105 platforms still encounter significant challenges in scaling and accessibility due to the high capital
106 costs of equipment, high variable costs for consumables, and the substantial cost and personnel
107 time required for adding additional targets. To achieve highly multiplexed detection, beyond
108 what is accessible with qPCR and dPCR, alternative solutions are needed.

109 Research in synthetic biology has produced various non-standard nucleotides that can be
110 used to circumvent major obstacles associated with scaling multiplexed PCR amplification to
111 larger ‘*n*’-plex reactions (**Figure 1**). The non-standard nucleic acids from the Self-Avoiding
112 Molecular Recognition Systems (SAMRS: A*, T*, G*, and C*) are structurally modified
113 versions of the standard DNA nucleobases (A, T, G, and C).²⁹ Though structurally distinct,
114 SAMRS nucleobases maintain an ability to base pair with standard DNA nucleobases (**Figure**
115 **1c**) but not with their SAMRS complement. For nucleic acid amplification, primers modified

116 with SAMRS components anneal and amplify natural DNA/RNA. Conversely, formation of
117 SAMRS:SAMRS pairs (T*:A* and C*:G*, **Figure 1d**) are thermodynamically disfavored. In
118 amplification assays with SAMRS-containing primers, this results in a decrease in off-target
119 primer-primer interactions. Selectively inserting SAMRS bases into primer sequences has been
120 shown to reduce primer dimer formation and increase multiplexed assay sensitivity.^{30,31}

121 Additionally, non-standard nucleic acids from the Artificially Expanded Genetic Information
122 System (AEGIS) can be used to improve primer binding specificity.³² AEGIS nucleotides (Z:P
123 pair, **Figure 1b**), such as Z (6-amino-5-nitro-3-(1'- β -d-2'-deoxyribofuranosyl)-2(1H)pyridine)
124 and P (2-amino-8-(1'- β -d-2'-deoxyribofuranosyl)), can form highly-specific base pairs *orthogonal*
125 to the standard, natural set (T:A and C:G).^{33,34} Since AEGIS bases are not found in nature, primers
126 containing AEGIS nucleotides can be used to amplify template targets containing
127 complementary AEGIS sequences while avoiding off-target amplification.³⁵



128

129 **Figure 1. Structures and hydrogen-bonding interactions for standard DNA, AEGIS, and**
130 **SAMRS nucleobases.** (a) Structures of standard DNA hydrogen bonding base pairs. (b)
131 Structures of Z and P AEGIS bases that form base pairs orthogonal to the standard DNA bases.
132 (c) SAMRS bases (*) form base pairs with their natural standard DNA complements. (d)
133 The strategic removal of hydrogen bonding groups hinders SAMRS bases from base pairing with
134 their SAMRS complement. Dotted lines indicate hydrogen bonding between base pairs, and
135 curved lines indicate a lack of hydrogen bond formation between base pairs.

136 Various detection assays have been developed that leverage properties of SAMRS and
137 AEGIS bases for viral pathogen detection, including arboviruses (e.g., Zika, Dengue,
138 Chikungunya),^{36,37} coronaviruses (e.g., RSV, MERS-CoV, Influenza A/B, SARS-CoV-
139 2),^{38,39} human papillomavirus (HPV),⁴⁰ and norovirus.⁴¹ Despite proving utility of SAMRS and
140 AEGIS nucleobases in assay design, these assays were either non-multiplexed or multiplexed but
141 required an expensive strategy for target readout (XMAP Luminex array detection).

142 In this work, we develop an ‘*icosaplex*’ (20-plex) PCR-based sequencing assay able to detect
143 20 enteric pathogen and antimicrobial resistance gene targets. This 20-plex assay greatly expands
144 on prior work that targeted viruses to new pathogens (bacteria, protozoa, and helminths) and to
145 environmental sample types. The 20-plex assay amalgamates individual primer sets for 20
146 targets from a previous collection of work and achieves effective multiplexing by incorporating
147 SAMRS-AEGIS nucleotides into primers chosen for biological reasons. To circumvent the target
148 identification limitations of PCR, we leveraged nanopore sequencing (Oxford Nanopore
149 Technologies), a third-generation sequencing method that is inexpensive, portable, and can
150 provide sequencing results in real time. This study is the first to use SAMRS-AEGIS primers for
151 highly multiplexed PCR in combination with nanopore sequencing for microbial surveillance
152 applications. Sequencing information provides additional insight into gene alleles and subspecies
153 that would otherwise be missed through presence/absence methods. The target panel and
154 detection method were chosen for application areas in environmental detection and surveillance
155 efforts in resource constrained settings, such as LMICs. We benchmarked performance of the 20-
156 plex assay in three sample matrices: wastewater, soil, and human feces.

157 **Results and Discussion**

158 **Pathogen and Antimicrobial Resistance Gene Target Selection**

159 We developed a multiplexable assay that can detect a broad range of microbial threats
160 relevant to global health. We chose 20 genes of interest that encompass a wide range of
161 enteropathogens (bacteria, protozoa, and one helminth) and two clinically important
162 antimicrobial resistance genes (ARGs) (**Table 1**). QPCR assays for these 20 targets have
163 previously been reported (**Table S1**).

164 Twelve of our targets are genes specific to pathogenic *E. coli*. These represent five of the
165 major pathogenic *E. coli* subtypes and *Shigella* spp. In LMICs, pathogenic *E. coli* is a leading
166 cause of diarrheal illness.^{5,42,43} The enterotoxigenic *E. coli* (ETEC) subtype is associated with
167 moderate-to-severe diarrhea which can lead to additional severe clinical outcomes.⁴³ LMICs also
168 experience a high incidence of soil-transmitted helminth infections.⁴⁴ It is estimated that 738
169 million people globally are infected with helminths of the genus *Ascaris*.⁴⁵ In HICs like the
170 United States, *Campylobacter* spp., non-typhoidal *Salmonella*, *Shigella* spp., and *Giardia*
171 *intestinalis*, are leading causes of reported foodborne illnesses.⁴⁶ In both LMICs and HICs,
172 pathogenic bacteria pose an even larger threat to human health if they acquire antimicrobial
173 resistance activity. *Bla*_{NDM} and *mcr-1* are globally distributed ARGs that confer resistance to the
174 last line of defense antibiotics reserved for difficult-to-treat infections.^{47,48} Many of the 20 gene
175 targets chosen in our panel were used by other studies to detect pathogens and ARGs in human
176 feces,^{3,6} wastewater,⁴⁹ and environmental samples.⁵⁰⁻⁵²

177

178 **Table 1. Enteric pathogen and antimicrobial resistance gene panel in the 20-plex assay.** The
179 20-plex assay is designed to detect 18 enteropathogen genes and two antimicrobial resistance
180 genes. Amplicon lengths for each PCR product are reported.

| Gene | Target | Len. (bp) |
|------|--------------------------------------|-----------|
| stx1 | Shiga toxin-producing <i>E. coli</i> | 132 |

| | | |
|-------|----------------------------------|-----|
| stx2 | (STEC) | 93 |
| aaiC | | 215 |
| aatA | Enter aggregative <i>E. coli</i> | 237 |
| aggR | (EAEC) | 95 |
| eae | Enteropathogenic <i>E. coli</i> | 102 |
| bfpA | (EPEC) | 110 |
| LT | | 62 |
| STh | Enterotoxigenic <i>E. coli</i> | 147 |
| STp | (ETEC) | 136 |
| ipaH | | 64 |
| virF | Enteroinvasive <i>E. coli</i> | 101 |
| | (EIEC) / <i>Shigella</i> spp. | |
| ttr | <i>Salmonella enterica</i> | 95 |
| hipO | <i>Campylobacter jejuni</i> | 122 |
| GlyA | <i>Campylobacter coli</i> | 125 |
| CR18S | <i>Cryptosporidium</i> spp. | 126 |
| G18S | <i>Giardia</i> spp. | 63 |
| ITS1 | <i>Ascaris</i> spp. | 88 |
| mcr-1 | Colistin resistance gene | 108 |
| NDM | Beta-lactam resistance gene | 109 |

181

182 Design and validation of 20-plex primers

183 Primer sequences from previously reported PCR and qPCR assays for the 20 gene targets
184 were used as a starting point for 20-plex PCR primer design (**Table S1**). Initial 40 primer
185 sequences (1 forward, 1 reverse for each gene target) were chosen to accommodate a single
186 annealing temperature (60 °C) during PCR cycling. At these temperatures and at a high relative
187 abundance of primer to target, various cross-primer interactions can occur to form primer dimers
188 and off-target amplicons, each reducing assay sensitivity (**Figure 2a**). To combat primer dimer
189 formation (**Figure 2b**), we modified all 40 standard DNA primers with SAMRS nucleobases
190 using the PrimerCompare software developed at the Foundation for Applied Molecular
191 Evolution (FfAME). PrimerCompare took standard DNA primer sequences that have proven
192 targets, primer concentrations, salt concentrations, and thermodynamic parameters (maximum
193 ΔG for hairpins and dimers) as inputs to simulate potential primer-primer interactions. These
194 interactions include self-dimerization, cross-primer dimerization, and hairpin structures.

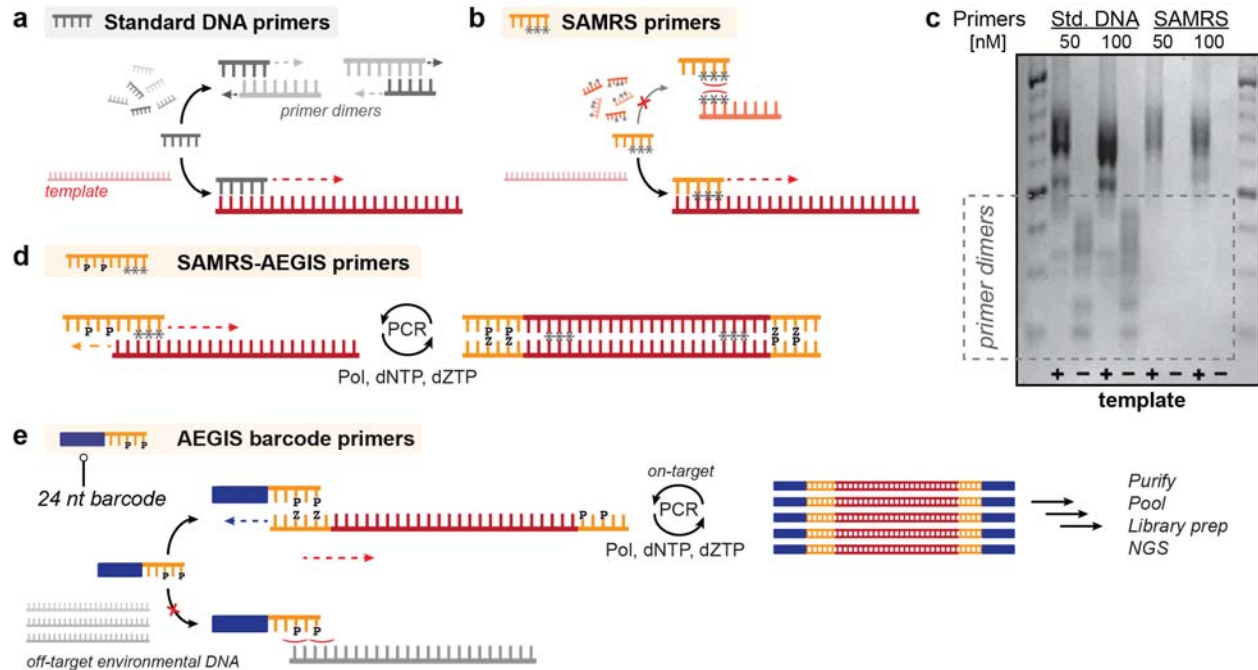
195 SAMRS containing primers for the 20 targets (**Table 1**) were then synthesized and
196 validated in single target PCR amplification reactions (**Figure S1**). A qPCR melt curve for each
197 primer set, with and without added synthetic DNA template, showed no primer dimer formations
198 in the no template controls (NTC, **Figure S2**). We then tested the effectiveness of the SAMRS
199 primers compared to standard DNA primers at reducing primer dimers in multiplexed reaction
200 conditions (**Table S2-S5**). 20-plex PCR was performed with and without synthetic template
201 added (**Table S6**). By agarose gel electrophoresis, primer dimer products were observed with
202 standard DNA primers, but not SAMRS-containing primers (**Figure 2c, Figure S3**). While both
203 standard DNA and SAMRS-containing primers showed target amplification in the 20-plex, the
204 identity of the individual amplicons could not be resolved through gel electrophoresis alone.

205 While incorporation of SAMRS bases helped decrease primer dimers in multiplex PCR
206 reactions, addition of a 5'-overhang tag to primers could be used for downstream barcoding and
207 attachment of sequencing adapters (**Table S2**). To avoid off-target amplification, 5'-overhang tag
208 sequences should be distinct from sequences that could be present in samples of interest. The
209 generalized design of a 5' -overhang tag, however, is challenging due to the unknown
210 metagenetic composition of many sample matrices (e.g., wastewater, soil, surface water, fomites,
211 feces). We overcame this obstacle by introducing non-standard AEGIS **P** nucleobases into the 5'
212 -overhang tag. Since AEGIS bases form a highly specific orthogonal base pair to the standard
213 DNA bases (**Z:P, Figure 1b, Figure 2d**), AEGIS containing primers should solely bind to
214 complementary AEGIS-tagged regions.

215 Various design choices were made to minimize design complexity and reagents that
216 would be required for performing multiplexed assays that use AEGIS components. First, the
217 AEGIS tag sequences used a 5-letter alphabet composed of the standard DNA bases (A, T, G, C)
218 and one of the AEGIS bases (P). In amplification reactions, end-users would therefore only need

219 access to complementary nucleotide triphosphate, dZTP, rather than both dZTP and dPTP.
220 Second, we chose to use a single AEGIS tag sequence for both forward and reverse primers. We
221 previously observed that using a single tag sequence in multiplexed PCR reactions reduced
222 overall primer dimer formation and increased detection sensitivity (data not shown). The final
223 AEGIS tag sequence (AGCPCTCGPTTC) was selected due to low propensity for hairpin
224 formation, as determined computationally. This AEGIS tag sequence is appended to the 5'-end of
225 the 40 SAMRS-containing primers used in this work (**Table S2**).

226 To multiplex samples, we then created 10 unique barcoding primers that contained a 24-
227 nt barcode region using sequences from an Oxford Nanopore Technologies barcoding kit. These
228 barcoding primers contained the barcode sequence and a downstream region homologous to the
229 common 5'-tag of the 20-plex SAMRS-AEGIS primers (**Table S3**). The universal 5' AEGIS tag
230 thus serves as the priming region for the barcoding primers either in the same PCR reaction (one-
231 pot amplification) or a subsequent PCR reaction (sequential amplification, **Figure 2e**). Though
232 barcoding primers discussed in this work were designed to be compatible with Oxford Nanopore
233 demultiplexing workflows, a similar design strategy can be used for barcoding applications on
234 other sequencing platforms (**Table S3**).



235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250

Figure 2. Role of SAMRS and AEGIS nucleotides in multiplex PCR design. (a) Standard DNA primers in a PCR reaction can dimerize and cross prime, consuming available primers and dNTPs. (b) SAMRS bases can be inserted in primer sequences to avoid primer dimer formation. (c) PCR amplification of all 20 targets using synthetic templates in nuclease-free water using standard DNA or SAMRS primers shows standard DNA primer dimerization, particularly prominent when no template is present (Figure S3, Table S2-S6). SAMRS primers show no visible dimerization. (d) The first round of amplification in our SAMRS-AEGIS 20-plex reaction uses primers containing SAMRS bases in the target-binding region, and the AEGIS P base in an overhang tag region. The corresponding Z triphosphate (dZTP) is included during amplification. (e) The second amplification reaction uses primers containing the AEGIS P base to bind to the tag region added during the first round. These primers also contain 24-nt barcode overhangs. Here, the AEGIS bases prevent non-specific amplification due to their lack of pairing with standard DNA bases. After this second PCR amplification, samples are purified, pooled, and prepared for next-generation sequencing.

251 Optimization of a sequential ‘two-step’ 20-plex PCR reaction

252 A unique challenge of multiplexing in complex samples of unknown metagenomic
253 composition is that gene targets are not present in equimolar amounts. For certain sample types,
254 targets in the same sample could be present at gene copy numbers that vary by orders of
255 magnitude. If barcoding and target amplification occur in one reaction, rather than sequentially,
256 higher abundance targets will bias amplification and consume barcoding primers, reducing assay
257 sensitivity for lower abundance targets. We tested this hypothesis by performing both one-pot

258 and sequential amplification of two targets using synthetic templates, *stx2* (present at 10 or 10²
259 copies/ μ L) and *aaiC* (10⁴ or 10⁵ copies/ μ L). When compared to one-pot PCR, performing
260 barcoding in a separate PCR reaction (sequential PCR) decreased differences in abundances of
261 *stx2* and *aaiC* amplicons (**Figure S4**).

262 Subsequently, PCR optimization was used to identify optimal reaction and cycling
263 conditions. For the optimal number of cycles in each step, we found 40 cycles (as is used for
264 qPCR) during the first round of amplification followed by 15 cycles in the second round for
265 barcoding minimized amplification bias and maximized barcoded targets over other
266 combinations tested (**Figure S4, S5**). In the first round of PCR, we found a uniform
267 concentration of each primer (0.2 μ M of each primer, 8 μ M total primer) minimized observed
268 amplification bias (**Figure S6**). Under these optimized 20-plex PCR reaction and cycling
269 conditions, primer dimers were still observed with standard DNA primers, but not with SAMRS-
270 AEGIS primers (**Figure S7**).

271 Finally, we incorporated nanopore sequencing, a low capital cost, portable sequencing
272 platform, as a read-out for detection of the SAMRS-AEGIS 20-plex reaction. Amplification was
273 performed on the 20 synthetic template mixtures at two initial concentrations for each target: 10
274 and 10⁴ copies/ μ L. Samples were sequenced on a MinION flow cell, basecalled, and
275 demultiplexed. All 20 targets were detectable by nanopore sequencing at initial template
276 concentrations of 10 and 10⁴ copies/ μ L (**Figure S8, S9, Table S6**). In both reaction conditions,
277 less reads were observed for three assay targets: *stx1*, *STh*, *aatA*. Though additional optimization
278 (e.g., adjusting primer concentrations) could be performed to improve relative amplification of
279 these three targets, many factors in environmental samples that cannot be controlled likely play a
280 larger role in determining differential amplification. For example, the absolute and relative

281 abundance of each target in real samples cannot be optimized. For design simplicity, we opted to
282 continue with equimolar SAMRS-AEGIS primer concentrations.

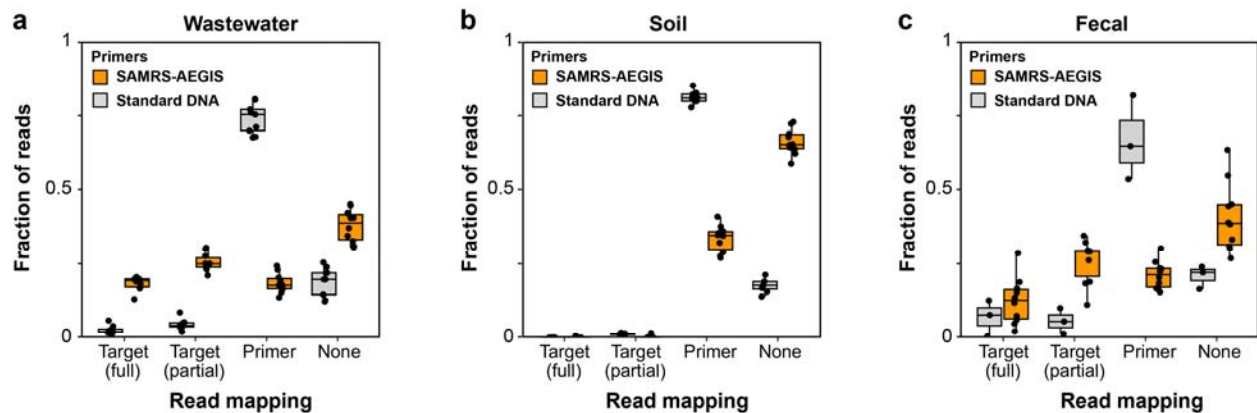
283 As designed, this SAMRS-AEGIS 20-plex PCR reaction overcomes challenges that must
284 be addressed for sensitive detection of multiple targets in environmental samples. Pathogen and
285 antimicrobial resistance genes can be in low abundance,⁵²⁻⁵⁴ necessitating modifications that
286 avoid primer dimerization. Inclusion of 1-3 SAMRS nucleotides in the seed region of the 20-plex
287 PCR was effective at eliminating detectable primer dimer formation as seen by both gel
288 electrophoresis and qPCR-based melting curve analysis. Target species in environmental
289 samples are often differentially abundant and many times orders of magnitude different.⁵⁴ We
290 performed two PCR reactions sequentially - Reaction 1 uses 40 cycles to detect low abundance
291 species or amplicons with low amplification efficiency, while Reaction 2 uses AEGIS
292 nucleotides to introduce 24-nt sample barcodes for nanopore sequencing in order to minimize
293 background, non-specific amplification. With equimolar amounts of all SAMRS-AEGIS primers,
294 this workflow was sensitive enough to detect all 20 targets using synthetic templates at 10
295 copies/ μ L for each target by nanopore sequencing.

296

297 **20-plex assay performance in environmental samples**

298 Previously, we hypothesized that AEGIS nucleotides in the primers could help avoid
299 non-specific 'background' amplification in environmental samples. To test this hypothesis, we
300 compared the sequencing outputs of the 20-plex assay using SAMRS-AEGIS primers to standard
301 DNA primers, in three sample types: wastewater, soil, and human feces (**Figure S10-S14, Table**
302 **S7-9**). Nanopore sequencing reads were demultiplexed and binned into one of four categories:
303 (1) fully map to target; (2) partially map to target; (3) map to primer regions, but not target; (4)
304 unmapped. For all sample types, the SAMRS-AEGIS 20-plex assay had significantly more reads

305 align to targets and fewer reads mapping to only primer regions compared to standard DNA 20-
306 plex assay (**Figure 3**). Though variations between sample matrices were readily observable, the
307 SAMRS-AEGIS 20-plex assay had between 1.8 – 7.5 times more read alignments to the full-
308 length targets compared to reads derived from the standard DNA 20-plex assay. Conversely, the
309 standard DNA 20-plex assay resulted in an average of 2.4 – 4 times more reads aligning only to
310 primers, but not target, compared to reads from the SAMRS-AEGIS 20-plex assay.
311



312 **Figure 3. Outcome of nanopore reads in SAMRS-AEGIS and standard DNA 20-plex assays**
313 **in wastewater, soil, and fecal matrices.** Read mapping fractions for each sample, separated by
314 sample matrix type: (a) wastewater ($n = 10$ SAMRS-AEGIS; $n = 9$ standard DNA), (b) soil ($n =$
315 10 SAMRS-AEGIS; $n = 8$ standard DNA), and (c) fecal ($n = 10$ SAMRS-AEGIS, $n = 3$ standard
316 DNA). For each sample, reads are binned into one of four categories: “Target (full)” = aligns to
317 full target sequence with at least 95% coverage; “Target (partial)” partially aligns to target
318 sequence with < 95% coverage; “Primer” maps to primer regions (priming site, barcode) but not
319 full target; “None” none of the prior bins. Fractions of reads within each sample that fall into
320 each bin are plotted (points) with boxplot overlaid to show the distribution of fractions
321 observed across each sample type-assay combination. Reads mapping to G18S, LT, and ipaH
322 were excluded from this analysis as they were detected in the NTC of the standard DNA 20-plex
323 assay. Box shows interquartile range (25th to 75th percentiles) with the median and whiskers
324 extending to 1.5 times the interquartile range.
325

326
327 The observed increase in on-target alignment of the SAMRS-AEGIS 20-plex assay
328 highlights the importance of non-standard nucleotides as an indispensable component of this 20-
329 plex assay. Reads aligning to only primers constituted the majority of reads from the standard
330 DNA 20-plex assay. Reads in this category constitute a mixture of off-target products, including

331 non-specific amplicons of environmental DNA and primer dimers. Two purification steps
332 involved in preparing the nanopore sequencing library involve steps that partially remove primer
333 dimers that would have been present. As such, a lower fraction of read in the ‘map to primer
334 only’ category could be traced to primer dimers. For sequencing-based detection, the sensitivity
335 of this assay is dependent on sequencing depth. Minimizing wasted sequencing effort on off-
336 target amplicons is critical for minimizing assay costs since it allows users to multiplex more
337 samples per sequencing flow cell.

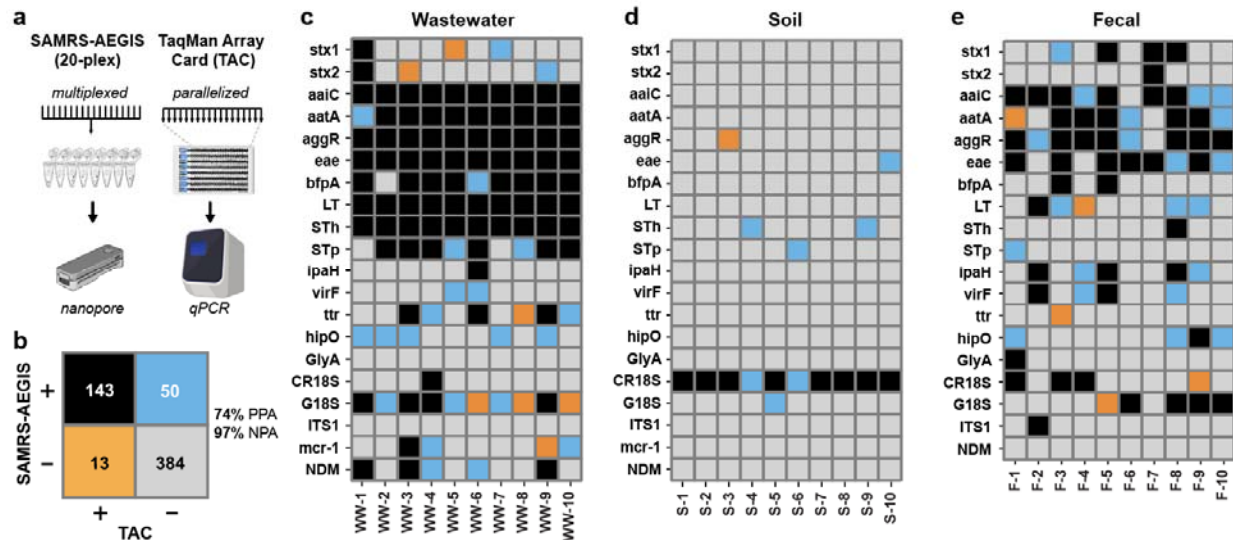
338

339 **Comparison between 20-plex assay and parallelized detection with TaqMan™ Array Cards**

340 To evaluate the performance of the SAMRS-AEGIS 20-plex assay against an established
341 method (**Figure 4a**), we compared assay results obtained from the 20-plex assay to those from
342 TaqMan™ Array Cards (TAC). TAC assays are qPCR-based and use a highly parallelized
343 architecture to detect multiple targets. Due to their convenience, sensitivity, and potential for
344 semi-quantitative detection, TAC assays are widely used in diagnostic and environmental
345 surveillance settings.^{6,51} Unlike the SAMRS-AEGIS 20-plex, TAC assays also require a
346 fluorescent probe for target identification. Both assays allow for sample multiplexing, with TAC
347 assays limited to eight samples per card. For these comparisons, 10 samples from the SAMRS-
348 AEGIS 20-plex reactions were multiplexed in a single MinION nanopore sequencing flow cell.

349 Comparing target detection between TAC and the SAMRS-AEGIS 20-plex assay, we
350 observed a 74% PPA (positive percent agreement) and a 97% NPA (negative percent agreement)
351 between these two methods (**Figure 4b**). Among the discrepancies, 13 out of 63 cases involved
352 targets detected by TAC but not by the SAMRS-AEGIS 20-plex assay, while 50 out of 63 cases
353 involved targets detected by the SAMRS-AEGIS 20-plex assay but not by TAC (**Figure 4b-e**).
354 To confirm the read-to-target assignments in the SAMRS-AEGIS 20-plex assay was not due to

355 mapping error or improper demultiplexing, reads were aligned against reference sequences and
 356 manually inspected for processing errors. All 50 targets identified in the SAMRS-AEGIS 20-
 357 plex assay, but not in the TAC assay, could be fully mapped to properly barcoded reads.
 358 Additionally, no reads in the no-template controls (NTCs) for the SAMRS-AEGIS 20-plex assay
 359 could be mapped to assay targets.
 360



361 **Figure 4. Comparison between SAMRS-AEGIS 20-plex assay and TaqMan™ array cards**
 362 **(TAC) in wastewater, soil, and fecal matrices.** (a) The SAMRS-AEGIS 20-plex assay uses a
 363 highly multiplexed architecture for multi-target detection, while TAC uses a highly parallelized
 364 architecture. TAC and SAMRS-AEGIS 20-plex assays were run on 30 environmental samples
 365 (10 wastewater, 10 soil, 10 fecal). (b) Total assays positive (+) and negative (-) for TAC and
 366 SAMRS-AEGIS 20-plex. Percent positive agreement (PPA) and percent negative agreement
 367 (NPA) shown. Individual assays positive and negative for TAC and SAMRS-AEGIS 20-plex in
 368 (c) wastewater, (d) soil, and (e) fecal samples. Square color indicates assay result, following the
 369 color coding used in the total assay result matrix. The TAC *mcr-1* assay was not available for
 370 fecal samples.
 371
 372

373 These results provide evidence that the SAMRS-AEGIS 20-plex has sensitivity in the
 374 tested environmental sample matrices similar to TAC; for certain targets, the SAMRS-AEGIS
 375 20-plex may be more sensitive. One possible reason for the observed differences by method is
 376 the assayed template input: TAC uses a maximum of 21 ng template per singleplexed well, while
 377 the 20-plex assay uses a maximum of 100 ng template in the first round of PCR. Another

378 possible explanation stems from differences in cycling conditions: TAC uses 40 qPCR cycles,
379 while the SAMRS-AEGIS 20-plex assay uses 40 cycles for first round amplification and 15 for
380 second round of tagged amplification (net: 45 cycles accounting for dilution).

381

382 **SAMRS-AEGIS 20-plex assays reveal additional information about microbial threats**

383 We next asked whether additional information is provided by the sequencing data
384 obtained from the SAMRS-AEGIS 20-plex assay. For each sample and each positive gene target
385 with at least 10 mapped reads, consensus sequences were generated and dereplicated to explore
386 diversity across samples.

387 STh, also known as ST1b, encodes a heat stable enterotoxin produced by enterotoxigenic
388 *E. coli* (ETEC).⁵⁵ Of the 12 samples that were positive by STh, three unique STh variants could
389 be identified (**Figure 5a**). These variants closely match unnamed STh variants present in
390 databases (**Supplementary File 1**).

391 For antimicrobial resistance genes, *mcr-1* and *bla_{NDM}*, we sought to identify different
392 alleles that could be amplified using the SAMRS-AEGIS 20-plex primer set by mapping reads
393 from positive samples to all alleles in the Comprehensive Antibiotic Resistance Database
394 (CARD).⁵⁶ While reads mapped to more than one allele sequence, the putative alleles are highly
395 similar (e.g., *mcr-1.20* and *mcr-1.14* have 1 bp different in the amplicon region) and could not be
396 distinguished from nanopore sequencing error. One strategy to distinguish highly similar alleles
397 within the same sample with nanopore sequencing is incorporating unique molecular identifiers
398 (UMIs) in the primers.⁵⁷ Alternatively, higher accuracy sequencing platforms such as Illumina or
399 PacBio could be used.

400 Beyond toxins and ARGs, the SAMRS-AEGIS 20-plex also targeted the 18S rRNA gene
401 to identify protozoan pathogens. The CR18S assay targets the 18S rRNA gene in

402 *Cryptosporidium* spp.⁵⁸ 12 of the 14 CR18S SAMRS-AEGIS 20-plex assays that were positive
403 for CR18S were also positive by TAC. Consensus sequences revealed six unique 18S alleles
404 from these samples. Three of these alleles mapped at 100% identity to previously observed
405 variants found in sequence databases, including an unnamed *Cryptosporidium sp.* isolate, a
406 *Cryptosporidium meleagridis* isolate, and a *Cryptosporidium hominis* isolate (**Figure 5b**). The
407 remaining variants were found to map with lower homology (approximately 90% ID) to
408 uncultured alveolates.

409 Finally, we were able to observe gene variants belonging to two subspecies of
410 *Campylobacter jejuni*. The *hipO* gene encodes for hippurate hydrolase in *C. jejuni*.⁵⁹ From
411 sequencing we were able to observe eight unique variants (**Figure 5c**). Four of these variants,
412 with nucleotides T33/T37 in the amplicon, closely map to *C. jejuni* subspecies *jejuni* while the
413 other four, with nucleotides C33/C37, closely map to *C. jejuni* subsp. *doylei*. Two of the *C. jejuni*
414 subsp. *jejuni hipO* variants and two of the *C. jejuni* subsp. *doylei hipO* variant matched with
415 100% ID to previously observed *hipO* genes. Two of the conserved polymorphisms in the *hipO*
416 *doylei* variant overlap with the probe binding region for the TAC assay (probe: T48/G47; *doylei*
417 C38/A47). Only one of eight samples positive for *hipO* with the SAMRS-AEGIS 20-plex was
418 positive using TAC.

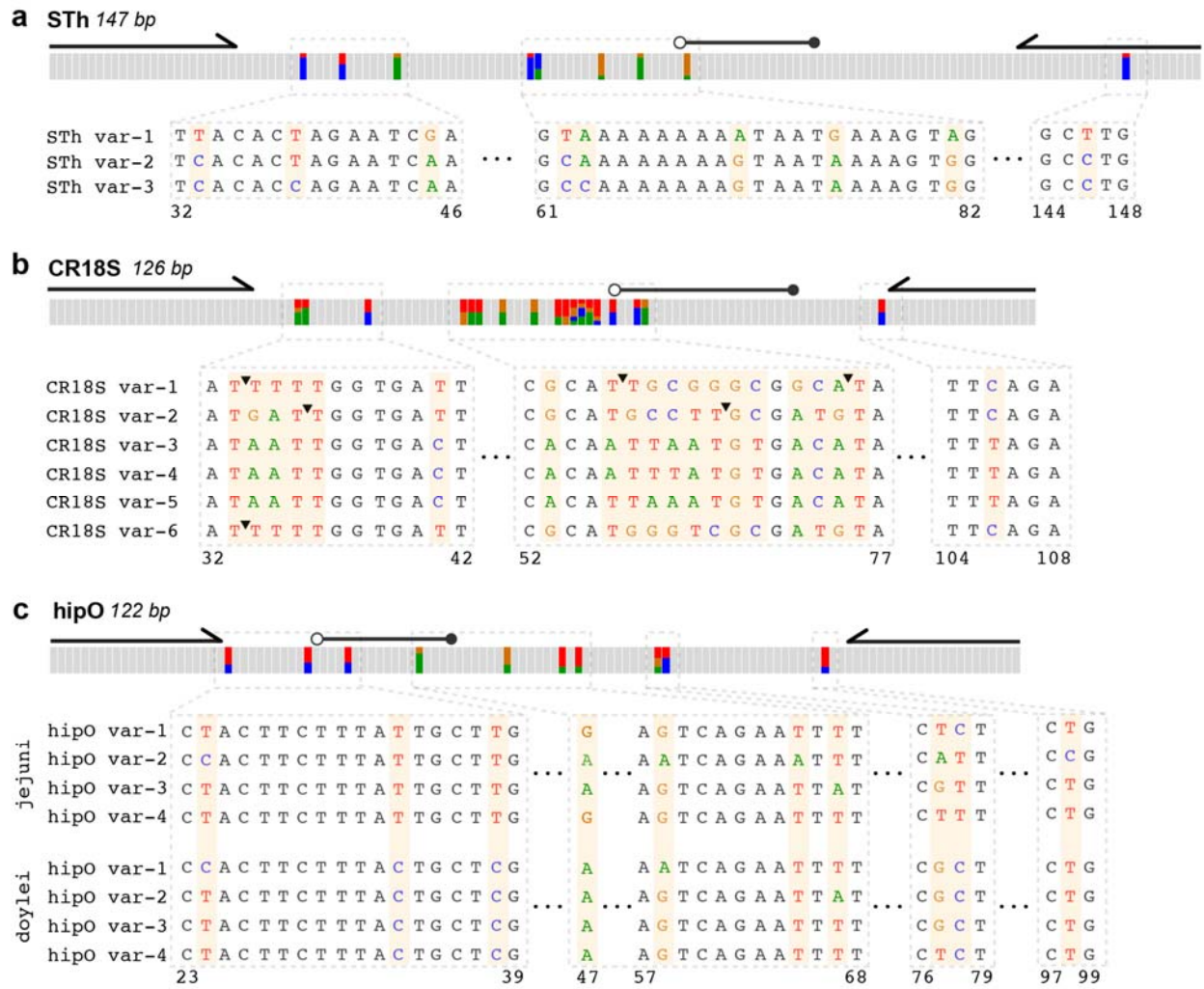


Figure 5. Unique consensus sequences of assay targets identified with the SAMRS-AEGIS 20-plex assay. Unique consensus sequences for targets are aligned against their respective reference sequence. Alignments are shown for (a) STh, an enterotoxin gene in *E. coli*; (b) CR18S, 18S rRNA gene of *Cryptosporidium*; and (c) *hipO*, a pathogen gene in *C. jejuni*. Primers and probe locations of the corresponding qPCR assay are marked above each target. Regions of interest are expanded to show single-base resolution. The top-alignment bar shows positions with high variation from consensus alignments in color, color coded by base, and gray otherwise. ▼ - indicates the presence of an insertion.

Sequencing results from the SAMRS-AEGIS 20-plex provided additional insight regarding microbial threats that would have otherwise been missed through a presence/absence-based approach. *C. jejuni* is an important pathogen in LMICs.⁶⁰ A 2016 study of eight birth cohorts across South America, sub-Saharan Africa, and Asia found that 85% of children are

434 carriers of *Campylobacter* spp. before the age of one.⁶⁰ *C jejuni* is also an important food-borne
435 pathogen in HICs primarily transmitted via poultry products.⁶¹ More samples were positive for
436 *C. jejuni* (*hipO*) than *C. coli* (*GlyA*) in both wastewater samples from WA and child fecal
437 samples from Ecuador. Within *C. jejuni*, two subspecies exist that display differing phenotypic
438 and clinical case presentations. The lesser-known *C. jejuni* subsp. *doylei* is more associated with
439 bacteremia and is known to cause gastritis, in addition to enteritis.⁶² Consensus sequences from
440 our 20-plex assay showed half of the *hipO* amplicons were more similar to *C. jejuni* subsp.
441 *doylei* than *C. jejuni* subsp. *jejuni* (**Supplementary File 1**), though some samples contained a
442 mixture of both species. The ability to distinguish these two subspecies is a notable feature of the
443 20-plex and provides important information that is useful for both LMIC and HIC settings.

444 Two assays in this work target eukaryotic 18S rRNA genes: CR18S (*Cryptosporidium*
445 *spp.*) and G18S (*Giardia spp.*). Of positive G18S samples from the 20-plex assay, all consensus
446 sequences mapped with 100% identity to *Giardia intestinalis*, the causative species of disease in
447 humans.⁶³ We observed six variants in *Cryptosporidium* 18S rRNA amplicons, (**Supplementary**
448 **File 1**), three of which mapped to uncultured alveolates at a lower identity (approx. 90% ID).
449 While the alveolate genus and species is unknown, the positive detection in both SAMRS-
450 AEGIS 20-plex and TAC is possibly the result of off-target amplification of related, but non-
451 pathogenic organisms in the alveolate genera. Given the high variation in the observed amplicons
452 for the CR18S genes, more specific assays that target *C. hominis* and *C. parvum*,⁶⁴ the causative
453 species of illness in humans, may be warranted.⁶⁵ Nonetheless, the sequencing used in the
454 SAMRS-AEGIS 20-plex assay provides a general means to interrogate variants within a sample
455 and distinguish between false positives and true positives.

456 Though the SAMRS-AEGIS 20-plex assay proved to be sensitive and more information-
457 rich than probe-based amplification strategies, there are some notable limitations. Probe-based

458 strategies that are qPCR-based, such as the TAC assay and dPCR are quantitative. As developed,
459 the 20-plex assay is not capable of relative quantification due to the use of sequential rounds of
460 PCR, nor can the assay be used for absolute quantification since it relies on sequencing.
461 Additionally, sensitivity of assay targets when multiplexing is highly dependent on the
462 differential abundance of target species. Optimization of concentrations for specific targets,
463 combined with *a priori* knowledge of expected environmental abundances, may be required for
464 improving sensitivity in certain sample types. Though the use of nanopore sequencing is well-
465 suited for resource limited settings, the low nominal basecalling accuracy (95%) limits our
466 ability to resolve multiple alleles within the same sample. To distinguish between multiple alleles
467 in a single sample, much higher sequencing coverage with nanopore or other higher accuracy
468 NGS (e.g., Illumina) would be required. Finally, as the 20-plex assay is built around detection of
469 extracted genetic material from samples, it is not suited for detection of viable organisms.

470 Despite these limitations, the SAMRS-AEGIS 20-plex assay strategy presented is a
471 promising alternative to conventional multi-target PCR detection methods. We estimate that
472 multiplexing 10 samples and 20 targets (20-plex) on a single nanopore MinION flow cell would
473 cost approximately \$4.00 per target and \$80.00 per sample (**Table S10, Supplementary File 2**),
474 which is similar to the per-target and per-sample costs of 20 parallelized assays on the TAC
475 platform. Where the SAMRS-AEGIS 20-plex assay design excels is at *scale*. Assuming a fixed
476 read coverage, sequencing the 20-plex reactions on an Illumina NovaSeq S4 would drop assay
477 costs to approximately \$0.55 per target or \$11.00 per sample. Setting aside variable costs, using
478 nanopore sequencing as a readout still offers a lower entry barrier in capital costs compared to
479 qPCR, dPCR, and other NGS platforms, making it well-suited for work in resource constrained
480 settings, such as LMICs.

481 Lastly, we highlight that the strategy presented in this work is not limited to the 20 targets
482 described here. SAMRS-AEGIS primers are premised on orthogonality, offering an element of
483 modularity for target choices that can be adapted to specific geographic contexts or modified to
484 include emerging threats. Coupled with the additional insight gained from sequencing, this
485 approach has the potential to significantly enhance our understanding of pathogens and antibiotic
486 resistance globally, paving the way for more effective public health interventions.

487

488 **Methods**

489 **Sample collection and nucleic acid extraction.** 10 wastewater samples were obtained from
490 treatment plants from Washington (WA) State. 25-50 mL of wastewater was centrifuged at 5000
491 x g for 20 minutes at 4 °C. The resulting pellet was resuspended in 200 µL of supernatant.
492 Nucleic acids from 200 µL of resuspended wastewater solids were extracted using AllPrep
493 PowerViral DNA/RNA Kit (Qiagen, Hilden, Germany), omitting the use of β-mercaptoethanol.
494 Purified wastewater DNA was eluted in RNase-free water to a final volume of 100 µL.

495 10 soil samples were collected from three dog parks located in Seattle, WA. Nucleic
496 acids from 0.25 g of each sample were extracted using DNeasy PowerSoil Pro Kit (Qiagen)
497 following standard protocol. Purified soil DNA was eluted in Solution C6 (10 mM Tris-HCl
498 buffer) to a final volume of 50 µL.

499 10 fecal samples from children were obtained from the ECoMiD cohort study in
500 northwest Ecuador.⁶⁶ The child stool samples were collected at 18 months of age. Nucleic acids
501 were extracted from 0.22 grams of stool samples using a modified QIAamp Fast DNA Stool
502 Mini Kit (Qiagen). Purified fecal DNA was eluted in Buffer ATE (10 mM Tris-HCl, 0.1 mM
503 EDTA, 0.04% NaN₃) to a final volume of 200 µL. The ECoMiD study protocol was approved by

504 the institutional review boards of the University of Washington (UW; IRB STUDY00014270),
505 Emory University (IRB00101202), and the Universidad San Francisco de Quito (2018–022M).
506 The study protocol was also reviewed and approved by the Ministry of Health of Ecuador
507 (MSPCURI000253-4).

508

509 **SAMRS-AEGIS Primer Design.** We selected 40 standard DNA primers from 20 qPCR assays
510 reported in previous literature (**Table S1**). These primers were modified with SAMRS
511 nucleobases to prevent primer dimer in a 20-plex PCR assay. SAMRS modifications were
512 designed using an iterative approach. Software developed at FfAME (PrimerCompare) took all
513 40 standard DNA primers along with primer, salt, and Mg⁺⁺ concentrations (200 nM, 60 mM,
514 and 2 mM, respectively) and output potential primer-primer interactions including self-
515 dimerization and hairpin structures. Using filters in the software, we concentrate on only the
516 most detrimental structures with sufficiently low ΔG values for hairpins and dimers, as well as
517 dimers with 3' to 3' overlaps within a short footprint (4-8 nt). These become our primary SAMRS
518 substitution regions. We then identified between 1-3 bases for SAMRS substitutions in the 3'-
519 overlap region that can destabilize the largest proportion of predicted structures. PrimerCompare
520 incorporates SAMRS nearest neighbor thermodynamic data and allows us to run the SAMRS
521 modified set as input to evaluate if further substitutions are required, along with checking the
522 Tms and ΔG s of modified primers. This process continues until an optimal set of primers is
523 designed.

524 Once all 40 standard DNA primers were modified with SAMRS, we added a common
525 AEGIS tag to the 5'-end. The 5' overhang sequence facilitates the attachment of barcode and
526 sequencing adapters in PCR. The AEGIS tag (AGCPCTCGPTTC) was designed to allow 2

527 AEGIS bases separated by 3 or more standard DNA bases and to have a T_m of at least 60 °C.
528 The designed 40 SAMRS-AEGIS primers are listed in **Table S2**. Further, AEGIS barcode
529 sequences used for sample multiplexing were designed by concatenating a barcode from Oxford
530 Nanopore Technologies Native Barcoding Kit (SQK-NBD112.24) with the AEGIS tag sequence,
531 which are listed in **Table S3**.

532

533 **SAMRS-AEGIS Primer Synthesis.** SAMRS or AEGIS containing oligonucleotides were
534 synthesized on Mermade 12 instruments, using standard phosphoramidite methods with minor
535 changes to the coupling time of AEGIS phosphoramidites (2 min for AEGIS, 1 min for standard
536 DNA bases and SAMRS). Solid support was a Mermade style column packed with controlled
537 pore glass (CPG) at 1000 Å pore size. Oligonucleotides were synthesized as either DMT-on or
538 DMT-off, followed by diethylamine wash (10% in ACN) at the end of the synthesis. DMT-off
539 oligonucleotides were deprotected in aqueous ammonium hydroxide (28-33% NH_3 in water) at
540 either 65 °C for 3 hours or 55 °C overnight, purified by ion-exchange HPLC (Dionex DNAPac
541 PA-100, 22x250 column), and desalted over SepPak C18 cartridges (Waters Corp., Milford,
542 MA). Oligonucleotides synthesized as DMT-on were deprotected using the same method,
543 followed by purification on Glen-Pak cartridges (GlenResearch, Sterling, VA). The purity of
544 each oligonucleotide was analyzed by analytical ion-exchange HPLC (Dionex DNAPac PA-100,
545 2x250 column). The oligonucleotides were sent out for ESI mass spectrometry (Novatia LLC,
546 Newtown, PA) to confirm their molecular weights.

547 **Sequential Multiplex PCR reaction and cycling conditions.** Unless otherwise specified, first
548 round PCR was performed at 20 μL scale and contained 1X Quantitect Multiplex PCR NoROX
549 master mix (Qiagen), 0.2 μM final concentration of each primer (40 primers listed in **Table S2**,

550 **S4** for SAMRS-AEGIS and standard DNA assays, respectively), and nucleic acid template. For
551 reactions that used synthetic templates, $10 - 10^5$ gene copies/ μL of synthetic templates (IDT,
552 **Table S6**) were added as specified. Synthetic templates were ordered as IDT gBlock Gene
553 Fragments, except LT, ipaH, G18S, and ITS1, which were ordered as two single-stranded oligos.
554 Oligos for LT, ipaH, G18S, and ITS1 were annealed by adding $20 \mu\text{M}$ of each oligo in 100 mM
555 of NaCl and 10 mM Tris-EDTA (pH 8.0) buffer and incubating at $90 \text{ }^\circ\text{C}$ for 3 minutes, then
556 cooling at $0.1 \text{ }^\circ\text{C/s}$ until reaching $20 \text{ }^\circ\text{C}$. For reactions using environmental or fecal DNA
557 extracts, up to 100 ng of DNA extract or $5 \mu\text{L}$ of volume were added (**Table S7**). First round
558 PCR reactions with SAMRS-AEGIS primers also contained a 0.05 mM final concentration of
559 dZTP (Firebird Biomolecular Sciences, Alachua, Fl). No template control (NTC) reactions were
560 run in parallel with samples, with the template volume replaced by nuclease-free water.

561 With the exception of experiments where cycling conditions are explicitly varied, first
562 round amplification in sequential PCR was amplified using the following cycling conditions:
563 initial denaturation at $95 \text{ }^\circ\text{C}$ for 15 min; followed by 40 cycles of (1) $95 \text{ }^\circ\text{C}$ for 30 s and (2) $60 \text{ }^\circ\text{C}$
564 for 60 s; a final extension 72°C for 5 min; then holding step at $12 \text{ }^\circ\text{C}$.

565 $1 \mu\text{L}$ of the PCR product was then used as the template for a second PCR reaction. The
566 second PCR reaction contained template, 1X Quantitect Multiplex PCR NoROX master mix
567 (Qiagen), $2 \mu\text{M}$ of 24-mer barcoding primer (**Table S3, S5**) in $30 \mu\text{L}$ of volume, or $20 \mu\text{L}$ when
568 specified. For reactions that contained SAMRS-AEGIS primers, 0.05 mM of dZTP was added.
569 No template control reactions for each barcoding primer were run in parallel with the samples,
570 with the template volume replaced by nuclease-free water.

571 The second round PCR reactions were amplified using the following cycling conditions:
572 $95 \text{ }^\circ\text{C}$ for 15 min, followed by 15 cycles of (1) $95 \text{ }^\circ\text{C}$ for 30 s and (2) $60 \text{ }^\circ\text{C}$ for 60 s, followed by
573 72°C for 5 min and a final holding step at $12 \text{ }^\circ\text{C}$. After each round of PCR, amplicons were

574 analyzed by gel electrophoresis on a 3% (w/v) agarose gel stained with GelGreen, and visualized
575 using a blue light transilluminator.

576

577 **Nanopore library preparation and data acquisition.** Prior to library preparation, all barcoded
578 samples were purified using magnetic DNA-binding beads (Sergi Lab Supplies, Seattle, WA)
579 with a 2:1 bead-to-sample ratio (v/v). Samples were washed twice with 70% ethanol, and eluted
580 in nuclease-free water to a final volume of 12 μ L. Purified DNA was quantified on a DeNovix
581 Fluorometer, and barcoded samples were pooled equally by weight. A subset of SAMRS-AEGIS
582 and standard DNA NTCs were also sequenced. Nanopore sample preparation followed standard
583 MinION Genomic DNA by Ligation protocol using the SQK-LSK114 kit with the two following
584 modifications: 1) During the DNA repair and end prep step, the NEBNext FFPE Repair Mix was
585 omitted to avoid potential SAMRS-AEGIS removal by repair enzymes. The volume of the repair
586 mix was replaced by nuclease-free water. 2) To preserve short fragments, the magnetic DNA-
587 binding bead-to-sample ratio was increased to 2:1. Up to 1.3 pmol of pooled samples were
588 loaded into the flowcell. MinION flow cells used in this work were from the R10.4.1 series.
589 Nanopore flow cells were used once per sample without washing, and data collection proceeded
590 for 72 h. A summary of nanopore sequencing runs is shown in **Table S8**.

591

592 **Nanopore data collection, basecalling, and processing.** Nanopore data acquisition was
593 performed using MinKNOW version 23.07.12. Data was collected in FAST5 format for
594 experiments with synthetic templates, and POD5 format for environmental/fecal samples.
595 FAST5 files were converted to POD5 format using the pod5 package (ONT, version 0.3.2).⁶⁷
596 Raw POD5 data files were basecalled using Dorado (ONT, version 0.6.2+14a7067) using the
597 super accurate model (dna_r10.4.1_e8.2_400bps_sup@v4.2.0) and a minimum q-score threshold

598 of 7.⁶⁸ Sample barcodes were demultiplexed using the Dorado demux command with the “--no-
599 trim” flag and a custom barcode configuration file that contained the 24-nt barcodes used in this
600 work.

601 Demultiplexed reads were aligned to a database containing barcoded reference sequences
602 using BLAST Command Line Tool (blastn, NCBI, version 2.9.0+) with the following flags: --
603 outfmt 10, --max-target-seqs 1.⁶⁹ After alignment, top hits for each read with at least 95%
604 coverage were stored as an initial match. The resulting reads were then passed through a more
605 stringent alignment using bowtie2 (version 2.3.5.1) with the following flags: --very-sensitive, --
606 local.⁷⁰ Bowtie2 alignment reference sequences contained target sequences without the barcode
607 region for the G18S, eae, CR18S, LT, and ipaH assays, and without the barcode or priming
608 region for the remaining assays. For sub-analysis of *hipO* alleles, both *hipO* variants from *C.*
609 *jejuni* subsp. *jejuni* and *C. jejuni* subsp. *doylei* were included in the reference sequences.

610 Reads that passed bowtie2 alignment were further aligned to the fully barcoded target
611 sequences. Consensus sequences for each target within each sample were generated from these
612 aligned reads using medaka (ONT, version 1.12.1) commands: ‘consensus’ and ‘stitch’.⁷¹ For the
613 ‘consensus’ command, no lower limit on the number of sequences required to generate a
614 consensus was placed at consensus generation stage. However, only alignments generated from
615 at least 10 sequences were used for downstream analysis. The ‘stitch’ command used the
616 following flag: --no-fillgaps. Consensus alignment % ID was calculated using the BLAST
617 Command Line (blastn) with the following flags: --outfmt 10, --max-target-seqs 1.

618
619 **TaqMan™ Array Card assays.** 1X TaqMan™ Fast Advanced PCR master mix (Thermo Fisher
620 Scientific) was used for all assays. A maximum amount of either 1400 ng or 20 µL of DNA was
621 loaded into a TaqMan™ Custom Plated Assay Microarray Card (**Table S7**). Six samples were

622 run on each card with a positive and negative control. For the positive control, 1×10^3 copies/ μ L
623 of synthetic templates from IDT containing all 20 targets were used (sequences provided in
624 **Table S6**). For the negative control, volume of template DNA was replaced by nuclease-free
625 water. Before running, the loaded card was spun down twice at 300 x g for 1 min. The
626 TaqMan™ Array Card Sealer was then used to seal the card. The QuantStudio 7 Flex System
627 (Thermo Fisher Scientific) was used, with qPCR cycling conditions set at 92 °C for 10 min,
628 followed by 40 cycles of 95 °C for 1 s and 60 °C for 20 s. Data analysis was performed using
629 Design & Analysis Software (version 2.8.0). The fecal samples were run as part of the ECoMiD
630 study using TAC cards with AgPath-ID One-Step RT-PCR master mix and did not have the *mcr-*
631 *I* assay.

632

633 **Comparison of read distributions between the SAMRS-AEGIS 20-plex and standard DNA**
634 **20-plex assays.** Of the 30 environmental samples collected in this work, 10 wastewater, 10 soil,
635 and 10 fecal samples were processed by the SAMRS-AEGIS 20-plex assay while 9 wastewater,
636 8 soil, and 3 fecal samples were processed by the standard DNA 20-plex assay. For both assay
637 results, nanopore reads were demultiplexed then binned into one of four categories. Reads were
638 classified as “Target (full)” if they successfully mapped to the intended target following the
639 pipeline outlined in the “Nanopore data collection, basecalling, and processing” section. This
640 pipeline used an initial 95% query mapping filter to remove partial alignments. Reads that
641 mapped to barcodes, primer region, and target amplicon region, but with <95% coverage, were
642 binned as “Target (partial)”. Reads that did not align to the target amplicon region, but did align
643 to barcode and primer regions were binned as “primer”. Reads in this category could include
644 primer dimers and other non-specific amplification products. The remaining reads, which did not
645 fall under the previous categories, were binned as “None”. Reads mapping to G18S, LT, and

646 ipaH were excluded from this analysis since they were detected in the NTC of the standard DNA
647 20-plex assay. Visuals for read bins were generated using R (version 4.3.2).

648

649 **Comparison of target detection between the SAMRS-AEGIS 20-plex assay and the**
650 **TaqMan™ array cards in environmental samples.** 10 wastewater, 10 soil, and 10 fecal
651 samples were analyzed for the presence of gene targets by both the SAMRS-AEGIS 20-plex
652 assay and TaqMan™ array cards (TAC). For the SAMRS-AEGIS 20-plex assay, an assay was
653 considered positive if at least one read successfully mapped to its target according to the pipeline
654 described in the “Nanopore data collection, basecalling, and processing” section; otherwise, it
655 was considered negative. For TAC, an assay was considered positive if at least one of the two
656 replicates in a card reported a Ct value <40; otherwise, it was considered negative. For TAC
657 assays in fecal samples, the *mcr-1* assay was not available and was excluded from analysis.
658 Agreement and disagreement between SAMRS-AEGIS and TAC for each assay and across all
659 samples were visualized on a plotted matrix. Plots were generated using Python (version 3.8.0).

660 Percent positive agreement (PPA) and percent negative agreement (NPA) was calculated
661 using the following formula:

662

$$\text{PPA} = \frac{P_{\text{SA_TAC}}}{P_{\text{SA_TAC}} + P_{\text{SA}}}$$

663

$$\text{NPA} = \frac{N_{\text{SA_TAC}}}{N_{\text{SA_TAC}} + N_{\text{SA}}}$$

664

665

666

PPA = positive percent agreement

667 P_{SA_TAC} = number of assays SAMRS-AEGIS and TAC positive

668 P_{SA} = number of assays SAMRS-AEGIS positive

669

670 NPA = negative percent agreement

671 N_{SA_TAC} = number of assays SAMRS-AEGIS and TAC negative

672 N_{SA} = number of assays SAMRS-AEGIS negative

673

674 **Identification and visualization of pathogen and antimicrobial resistance gene alleles.** Reads

675 were processed as described previously with the inclusion of *hipO* variant sub-analysis

676 specifications. Consensus sequences generated in each sample with a coverage ≥ 10 reads were

677 aligned to reference sequences of target gene, then dereplicated. Alignments were visualized

678 using Integrative Genomics Viewer (version 2.16.2).⁷² Regions of interest were manually

679 extracted and expanded for visualization. To identify if putative allele sequences had previously

680 been observed, BLASTN webserver was used to map consensus sequences against NCBI core

681 non-redundant nucleic acid database (core_nt).

682

683 **Data availability.** The demultiplexed nanopore sequencing basecalls (FASTQ) for each sample

684 analyzed in this work have been deposited in the sequence reads archive (SRA) under Bioproject

685 PRJNA1150247 (Table S9).

686 **Acknowledgments:** We thank the wastewater treatment plants for collecting samples for this

687 work.

688 **Author Contribution:** Project conceptualization was performed by ERF, JAM, and ZY.

689 Methodology for this work was developed by HK, SMP, LM, and ZY. SAMRS-AEGIS

690 oligonucleotides were synthesized by CC and CM. Fecal samples were contributed by KL and
691 NAZ. Laboratory experiments were performed by HK and NAZ. Data analysis was conducted by
692 HK, SMP, KB, ZY, JAM, and ERF. Visualization of data and results was performed by HK,
693 JAM, and ERF. This project was supervised by ERF. Writing of original draft was carried out by
694 HK, JAM, ERF, and ZY. Reviewing and editing of the manuscript was performed by all.

695 **Conflict of Interest:** S.A.B and Z.Y. own the intellectual property of AEGIS and SAMRS.
696 Many AEGIS and SAMRS components are commercially available from Firebird Biomolecular
697 Sciences, LLC (www.firebirdbio.com, Email: support@firebirdbio.com). The remaining authors
698 declare no competing interests.

699 **Funding:** LM, SMP, KMB, CC, SAB, and ZY were supported by the LRE Diagnostics grant
700 1R01AI135146-01A1. NAZ and KL were supported by R01AI137679. Laboratory infrastructure
701 and hardware used for this study was supported by the University of Washington
702 Interdisciplinary Center for Exposures, Diseases, Genomics, and Environment funded by the
703 NIEHS (P30ES007033). JM and HK were supported by University of Washington Royalty
704 Research Fund (RRF).

References.

705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760

- (1) Baker, R. E.; Mahmud, A. S.; Miller, I. F.; Rajeev, M.; Rasambainarivo, F.; Rice, B. L.; Takahashi, S.; Tatem, A. J.; Wagner, C. E.; Wang, L. F.; Wesolowski, A.; Metcalf, C. J. E. Infectious Disease in an Era of Global Change. *Nat. Rev. Microbiol.* Nature Research April 1, 2022, pp 193–205. <https://doi.org/10.1038/s41579-021-00639-z>.
- (2) Murray, C. J. L. Global Burden of 288 Causes of Death and Life Expectancy Decomposition in 204 Countries and Territories and 811 Subnational Locations, 1990–2021: A Systematic Analysis for the Global Burden of Disease Study 2021. *The Lancet* **2024**, *403* (10440), 2100–2132. [https://doi.org/https://doi.org/10.1016/S0140-6736\(24\)00367-2](https://doi.org/https://doi.org/10.1016/S0140-6736(24)00367-2).
- (3) Knee, J.; Sumner, T.; Adriano, Z.; Anderson, C.; Bush, F.; Capone, D.; Casmo, V.; Holcomb, D.; Kolsky, P.; Macdougall, A.; Molotkova, E.; Braga, J. M.; Russo, C.; Schmidt, W. P.; Stewart, J.; Zambrana, W.; Zuin, V.; Nalá, R.; Cumming, O.; Brown, J. Effects of an Urban Sanitation Intervention on Childhood Enteric Infection and Diarrhea in Maputo, Mozambique: A Controlled before-and-after Trial. *Elife* **2021**, *10*, e62278. <https://doi.org/10.7554/ELIFE.62278>.
- (4) Grembi, J. A.; Lin, A.; Karim, M. A.; Islam, M. O.; Miah, R.; Arnold, B. F.; Rogawski McQuade, E. T.; Ali, S.; Rahman, M. Z.; Hussain, Z.; Shoab, A. K.; Famida, S. L.; Hossen, M. S.; Mutsuddi, P.; Rahman, M.; Unicomb, L.; Haque, R.; Taniuchi, M.; Liu, J.; Platts-Mills, J. A.; Holmes, S. P.; Stewart, C. P.; Benjamin-Chung, J.; Colford, J. M.; Houpt, E. R.; Luby, S. P. Effect of Water, Sanitation, Handwashing, and Nutrition Interventions on Enteropathogens in Children 14 Months Old: A Cluster-Randomized Controlled Trial in Rural Bangladesh. *J. Infect. Dis.* **2023**, *227* (3), 434–447. <https://doi.org/10.1093/infdis/jiaa549>.
- (5) Platts-Mills, J. A.; Babji, S.; Bodhidatta, L.; Gratz, J.; Haque, R.; Havt, A.; McCormick, B. J. J.; McGrath, M.; Olortegui, M. P.; Samie, A.; Shakoor, S.; Mondal, D.; Lima, I. F. N.; Hariraju, D.; Rayamajhi, B. B.; Qureshi, S.; Kabir, F.; Yori, P. P.; Mufamadi, B.; Amour, C.; Carreon, J. D.; Richard, S. A.; Lang, D.; Bessong, P.; Mduma, E.; Ahmed, T.; Lima, A. A. A. M.; Mason, C. J.; Zaidi, A. K. M.; Bhutta, Z. A.; Kosek, M.; Guerrant, R. L.; Gottlieb, M.; Miller, M.; Kang, G.; Houpt, E. R.; Chavez, C. B.; Trigo, D. R.; Flores, J. T.; Vasquez, A. O.; Pinedo, S. R.; Acosta, A. M.; Ahmed, I.; Alam, D.; Ali, A.; Rasheed, M.; Soofi, S.; Turab, A.; Yousafzai, A. K.; Bose, A.; Jennifer, M. S.; John, S.; Kaki, S.; Koshy, B.; Muliyl, J.; Raghava, M. V.; Ramachandran, A.; Rose, A.; Sharma, S. L.; Thomas, R. J.; Pan, W.; Ambikapathi, R.; Charu, V.; Dabo, L.; Doan, V.; Graham, J.; Hoest, C.; Knobler, S.; Mohale, A.; Nayyar, G.; Psaki, S.; Rasmussen, Z.; Seidman, J. C.; Wang, V.; Blank, R.; Tountas, K. H.; Swema, B. M.; Yarrot, L.; Nshama, R.; Ahmed, A. M. S.; Tofail, F.; Hossain, I.; Islam, M.; Mahfuz, M.; Chandyo, R. K.; Shrestha, P. S.; Shrestha, R.; Ulak, M.; Black, R.; Caulfield, L.; Checkley, W.; Chen, P.; Lee, G.; Murray-Kolb, L. E.; Schaefer, B.; Pendergast, L.; Abreu, C.; Costa, H.; Moura, A. Di; Filho, J. Q.; Leite, Á.; Lima, N.; Maciel, B.; Moraes, M.; Mota, F.; Oriá, R.; Quetz, J.; Soares, A.; Patil, C. L.; Mahopo, C.; Mapula, A.; Nesamvuni, C.; Nyathi, E.; Barrett, L.; Petri, W. A.; Scharf, R.; Shrestha, B.; Shrestha, S. K.; Strand, T.; Svensen, E. Pathogen-Specific Burdens of Community Diarrhoea in Developing Countries: A Multisite Birth Cohort Study (MAL-ED). *Lancet Glob. Health* **2015**, *3* (9), e564–e575. [https://doi.org/10.1016/S2214-109X\(15\)00151-5](https://doi.org/10.1016/S2214-109X(15)00151-5).
- (6) Liu, J.; Platts-Mills, J. A.; Juma, J.; Kabir, F.; Nkeze, J.; Okoi, C.; Operario, D. J.; Uddin, J.; Ahmed, S.; Alonso, P. L.; Antonio, M.; Becker, S. M.; Blackwelder, W. C.; Breiman, R. F.; Faruque, A. S. G.; Fields, B.; Gratz, J.; Haque, R.; Hossain, A.; Hossain, M. J.; Jarju, S.; Qamar, F.; Iqbal, N. T.; Kwambana, B.; Mandomando, I.; McMurry, T. L.; Ochieng, C.; Ochieng, J. B.; Ochieng, M.; Onyango, C.; Panchalingam, S.; Kalam, A.; Aziz, F.; Qureshi, S.; Ramamurthy, T.; Roberts, J. H.; Saha, D.; Sow, S. O.; Stroup, S. E.; Sur, D.; Tamboura, B.; Taniuchi, M.; Tennant, S. M.; Toema, D.; Wu, Y.; Zaidi, A.; Nataro, J. P.; Kotloff, K. L.; Levine, M. M.; Houpt, E. R. Use of Quantitative Molecular Diagnostic Methods to Identify Causes of Diarrhoea in Children: A Reanalysis of the GEMS Case-Control Study. *The Lancet* **2016**, *388* (10051), 1291–1301. [https://doi.org/10.1016/S0140-6736\(16\)31529-X](https://doi.org/10.1016/S0140-6736(16)31529-X).
- (7) Tokars, J. I.; Olsen, S. J.; Reed, C. Seasonal Incidence of Symptomatic Influenza in the United States. *Clin. Infect. Dis.* **2018**, *66* (10), 1511–1518. <https://doi.org/10.1093/cid/cix1060>.
- (8) Khabbaz, R. F.; Moseley, R. R.; Steiner, R. J.; Levitt, A. M.; Bell, B. P. Challenges of Infectious Diseases in the USA. *The Lancet*. Elsevier B.V. 2014, pp 53–63. [https://doi.org/10.1016/S0140-6736\(14\)60890-4](https://doi.org/10.1016/S0140-6736(14)60890-4).

- 761 (9) Kilaru, P.; Hill, D.; Anderson, K.; Collins, M. B.; Green, H.; Kmush, B. L.; Larsen, D. A.
762 Wastewater Surveillance for Infectious Disease: A Systematic Review. *Am. J. Epidemiol.* Oxford
763 University Press February 1, 2023, pp 305–322. <https://doi.org/10.1093/aje/kwac175>.
- 764 (10) Asghar, H.; Diop, O. M.; Weldegebriel, G.; Malik, F.; Shetty, S.; El Bassioni, L.; Akande, A. O.;
765 Al Maamoun, E.; Zaidi, S.; Adeniji, A. J.; Burns, C. C.; Deshpande, J.; Oberste, M. S.; Lowther, S.
766 A. Environmental Surveillance for Polioviruses in the Global Polio Eradication Initiative. *J. Infect.*
767 *Dis.* **2014**, *210* (suppl 1), S294–S303. <https://doi.org/10.1093/infdis/jiu384>.
- 768 (11) Kitakawa, K.; Kitamura, K.; Yoshida, H. Monitoring Enteroviruses and SARS-CoV-2 in
769 Wastewater Using the Polio Environmental Surveillance System in Japan. *Appl. Environ.*
770 *Microbiol.* **2023**, *89* (4), e01853-22. <https://doi.org/10.1128/aem.01853-22>.
- 771 (12) Zhou, N. A.; Fagnant-Sperati, C. S.; Komen, E.; Mwangi, B.; Mukubi, J.; Nyangao, J.; Hassan, J.;
772 Chepkurui, A.; Maina, C.; van Zyl, W. B.; Matsapola, P. N.; Wolfaardt, M.; Ngwana, F. B.;
773 Jeffries-Miles, S.; Coulliette-Salmond, A.; Peñaranda, S.; Shirai, J. H.; Kossik, A. L.; Beck, N. K.;
774 Wilmouth, R.; Boyle, D. S.; Burns, C. C.; Taylor, M. B.; Borus, P.; Meschke, J. S. Feasibility of
775 the Bag-Mediated Filtration System for Environmental Surveillance of Poliovirus in Kenya. *Food*
776 *Environ. Virol.* **2020**, *12*, 35–47. <https://doi.org/10.1007/s12560-019-09412-1>.
- 777 (13) Naughton, C. C.; Roman, F. A.; Alvarado, A. G. F.; Tariqi, A. Q.; Deeming, M. A.; Kadonsky, K.
778 F.; Bibby, K.; Bivins, A.; Medema, G.; Ahmed, W.; Katsivelis, P.; Allan, V.; Sinclair, R.; Rose, J.
779 B. Show Us the Data: Global COVID-19 Wastewater Monitoring Efforts, Equity, and Gaps. *FEMS*
780 *Microbes* **2023**, *4*, xtad003. <https://doi.org/10.1093/femsmc/xtad003>.
- 781 (14) Ahmed, W.; Smith, W. J. M.; Metcalfe, S.; Jackson, G.; Choi, P. M.; Morrison, M.; Field, D.;
782 Gyawali, P.; Bivins, A.; Bibby, K.; Simpson, S. L. Comparison of RT-qPCR and RT-dPCR
783 Platforms for the Trace Detection of SARS-CoV-2 RNA in Wastewater. *ACS ES&T Water* **2022**, *2*
784 (11), 1871–1880. <https://doi.org/10.1021/acsestwater.1c00387>.
- 785 (15) Peccia, J.; Zulli, A.; Brackney, D. E.; Grubaugh, N. D.; Kaplan, E. H.; Casanovas-Massana, A.; Ko,
786 A. I.; Malik, A. A.; Wang, D.; Wang, M.; Warren, J. L.; Weinberger, D. M.; Arnold, W.; Omer, S.
787 B. Measurement of SARS-CoV-2 RNA in Wastewater Tracks Community Infection Dynamics.
788 *Nat. Biotechnol.* **2020**, *38* (10), 1164–1167. <https://doi.org/10.1038/s41587-020-0684-z>.
- 789 (16) Kilpatrick, D. R.; Yang, C.-F.; Ching, K.; Vincent, A.; Iber, J.; Campagnoli, R.; Mandelbaum, M.;
790 De, L.; Yang, S.-J.; Nix, A.; Kew, O. M. Rapid Group-, Serotype-, and Vaccine Strain-Specific
791 Identification of Poliovirus Isolates by Real-Time Reverse Transcription-PCR Using Degenerate
792 Primers and Probes Containing Deoxyinosine Residues. *J. Clin. Microbiol.* **2009**, *47* (6), 1939–
793 1941. <https://doi.org/10.1128/JCM.00702-09>.
- 794 (17) Adams, C.; Bias, M.; Welsh, R. M.; Webb, J.; Reese, H.; Delgado, S.; Person, J.; West, R.; Shin,
795 S.; Kirby, A. The National Wastewater Surveillance System (NWSS): From Inception to
796 Widespread Coverage, 2020–2022, United States. *Sci. Total Environ.* **2024**, *924*, 171566.
797 <https://doi.org/10.1016/j.scitotenv.2024.171566>.
- 798 (18) Wagner, E. G.; Lanoix, J. N. *Excreta Disposal for Rural Areas and Small Communities*. World
799 Health Organization; 1958.
- 800 (19) Manuel, M.; Amato, H. K.; Pilotte, N.; Chieng, B.; Araka, S. B.; Edoux Eric Siko, J.; Harris, M.;
801 Nadimpalli, M.; Janagaraj, V.; Houngebegnon, P.; Rajendiran, R.; Thamburaj, J.; Puthupalayam
802 Kaliappan, S.; Sirois, A. R.; Walch, G.; Oswald, W. E.; Asbjornsdottir, K. H.; Galagan, S. R.;
803 Walson, J. L.; Williams, S. A.; F Luty, A. J.; Njenga, S. M.; Ibikounlé, M.; Ajjampur, S. S.;
804 Pickering, A. J. Soil Surveillance for Monitoring Soil-Transmitted Helminth Infections: Method
805 Development and Field Testing in Three Countries. *medRxiv* **2023**.
806 <https://doi.org/10.1101/2023.09.26.23296174>.
- 807 (20) Harvey, A. P.; Fuhrmeister, E. R.; Cantrell, M. E.; Pitol, A. K.; Swarthout, J. M.; Powers, J. E.;
808 Nadimpalli, M. L.; Julian, T. R.; Pickering, A. J. Longitudinal Monitoring of SARS-CoV-2 RNA
809 on High-Touch Surfaces in a Community Setting. *Environ. Sci. Technol. Lett.* **2020**, *8* (2), 168–
810 175. <https://doi.org/10.1021/acs.estlett.0c00875>.
- 811 (21) Pitol, A. K.; Julian, T. R. Community Transmission of SARS-CoV-2 by Surfaces: Risks and Risk
812 Reduction Strategies. *Environ. Sci. Technol. Lett.* **2021**, *8* (3), 263–269.
813 <https://doi.org/10.1021/acs.estlett.0c00966>.
- 814 (22) Genné-Bacon, E. A.; Bascom-Slack, C. A. The PARE Project: A Short Course-Based Research
815 Project for National Surveillance of Antibiotic-Resistant Microbes in Environmental Samples. *J.*
816 *Microbiol. Biol. Educ.* **2018**, *19* (3). <https://doi.org/10.1128/jmbe.v19i3.1603>.

- 817 (23) Liguori, K.; Keenum, I.; Davis, B. C.; Calarco, J.; Milligan, E.; Harwood, V. J.; Pruden, A.
818 Antimicrobial Resistance Monitoring of Water Environments: A Framework for Standardized
819 Methods and Quality Control. *Environ. Sci. Technol.* **2022**, *56* (13), 9149–9160.
820 <https://doi.org/10.1021/acs.est.1c08918>.
- 821 (24) Hart, A.; Warren, J.; Wilkinson, H.; Schmidt, W. Environmental Surveillance of Antimicrobial
822 Resistance (AMR), Perspectives from a National Environmental Regulator in 2023. *Euro. Surveill.*
823 **2023**, *28* (11). <https://doi.org/10.2807/1560-7917.ES.2023.28.11.2200367>.
- 824 (25) Elnifro, E. M.; Ashshi, A. M.; Cooper, R. J.; Klapper, P. E. Multiplex PCR: Optimization and
825 Application in Diagnostic Virology. *Clin. Microbiol. Rev.* **2000**, *13* (4), 559–570.
- 826 (26) Qiagen. QIAcuity Application Guide Version 2. In *QIAcuity User Manual Extension*; 2023.
- 827 (27) Kodani, M.; Yang, G.; Conklin, L. M.; Travis, T. C.; Whitney, C. G.; Anderson, L. J.; Schrag, S. J.;
828 Taylor, T. H.; Beall, B. W.; Breiman, R. F.; Feikin, D. R.; Njenga, M. K.; Mayer, L. W.; Oberste,
829 M. S.; Tondella, M. L. C.; Winchell, J. M.; Lindstrom, S. L.; Erdman, D. D.; Fields, B. S.
830 Application of TaqMan Low-Density Arrays for Simultaneous Detection of Multiple Respiratory
831 Pathogens. *J. Clin. Microbiol.* **2011**, *49* (6), 2175–2182. <https://doi.org/10.1128/JCM.02270-10>.
- 832 (28) Applied Biosystems. *TaqMan™ Gene Expression Assays—TaqMan™ Array Cards; User Guide*;
833 2021.
- 834 (29) Hoshika, S.; Chen, F.; Leal, N. A.; Benner, S. A. Self-Avoiding Molecular Recognition Systems
835 (SAMRS). *Nucleic acids symp. ser.* **2008**, *52* (1), 129–130. <https://doi.org/10.1093/nass/nrn066>.
- 836 (30) Hoshika, S.; Chen, F.; Leal, N. A.; Benner, S. A. Artificial Genetic Systems: Self-avoiding DNA
837 in PCR and Multiplexed PCR. *Angew. Chem. Int. Ed. Engl.* **2010**, *49* (32), 5554–5557.
838 <https://doi.org/10.1002/anie.201001977>.
- 839 (31) Yang, Z.; Le, J. T.; Hutter, D.; Bradley, K. M.; Overton, B. R.; McLendon, C.; Benner, S. A.
840 Eliminating Primer Dimers and Improving SNP Detection Using Self-Avoiding Molecular
841 Recognition Systems. *Biol. Methods Protoc.* **2020**, *5* (1), bpa004.
842 <https://doi.org/10.1093/biomethods/bpaa004>.
- 843 (32) Benner, S. A. Understanding Nucleic Acids Using Synthetic Chemistry. *Acc. Chem. Res.* **2004**, *37*
844 (10), 784–797. <https://doi.org/10.1021/ar040004z>.
- 845 (33) Yang, Z.; Hutter, D.; Sheng, P.; Sismour, A. M.; Benner, S. A. Artificially Expanded Genetic
846 Information System: A New Base Pair with an Alternative Hydrogen Bonding Pattern. *Nucleic*
847 *Acids Res.* **2006**, *34* (21), 6095–6101. <https://doi.org/10.1093/nar/gkl633>.
- 848 (34) Hoshika, S.; Leal, N. A.; Kim, M.-J.; Kim, M.-S.; Karalkar, N. B.; Kim, H.-J.; Bates, A. M.;
849 Watkins, N. E.; SantaLucia, H. A.; Meyer, A. J.; DasGupta, S.; Piccirilli, J. A.; Ellington, A. D.;
850 SantaLucia Jr., J.; Georgiadis, M. M.; Benner, S. A. Hachimoji DNA and RNA: A Genetic System
851 with Eight Building Blocks. *Science (1979)* **2019**, *363* (6429), 884–887.
852 <https://doi.org/10.1126/science.aat0971>.
- 853 (35) Yang, Z.; Chen, F.; Chamberlin, S. G.; Benner, S. A. Expanded Genetic Alphabets in the
854 Polymerase Chain Reaction. *Angew. Chem. Int. Ed.* **2010**, *49* (1), 177–180.
855 <https://doi.org/10.1002/anie.200905173>.
- 856 (36) Glushakova, L. G.; Bradley, A.; Bradley, K. M.; Alto, B. W.; Hoshika, S.; Hutter, D.; Sharma, N.;
857 Yang, Z.; Kim, M. J.; Benner, S. A. High-Throughput Multiplexed XMAP Luminex Array Panel
858 for Detection of Twenty Two Medically Important Mosquito-Borne Arboviruses Based on
859 Innovations in Synthetic Biology. *J. Virol. Methods* **2015**, *214*, 60–74.
860 <https://doi.org/10.1016/j.jviromet.2015.01.003>.
- 861 (37) Glushakova, L. G.; Alto, B. W.; Kim, M. S.; Bradley, A.; Yaren, O.; Benner, S. A. Detection of
862 Chikungunya Viral RNA in Mosquito Bodies on Cationic (Q) Paper Based on Innovations in
863 Synthetic Biology. *J. Virol. Methods* **2017**, *246*, 104–111.
864 <https://doi.org/10.1016/j.jviromet.2017.04.013>.
- 865 (38) Yaren, O.; Glushakova, L. G.; Bradley, K. M.; Hoshika, S.; Benner, S. A. Standard and AEGIS
866 Nicking Molecular Beacons Detect Amplicons from the Middle East Respiratory Syndrome
867 Coronavirus. *J. Virol. Methods* **2016**, *236*, 54–61. <https://doi.org/10.1016/j.jviromet.2016.07.008>.
- 868 (39) Glushakova, L. G.; Sharma, N.; Hoshika, S.; Bradley, A. C.; Bradley, K. M.; Yang, Z.; Benner, S.
869 A. Detecting Respiratory Viral RNA Using Expanded Genetic Alphabets and Self-Avoiding DNA.
870 *Anal. Biochem.* **2015**, *489*, 62–72. <https://doi.org/10.1016/j.ab.2015.08.015>.
- 871 (40) Wang, Y.; Jiao, W.-W.; Wang, Y.; Wang, Y.-C.; Shen, C.; Qi, H.; Shen A-Dong. Graphene Oxide
872 and Self-Avoiding Molecular Recognition Systems-Assisted Recombinase Polymerase

- 873 Amplification Coupled with Lateral Flow Bioassay for Nucleic Acid Detection. *Microchim Acta*
874 **2020**, 187 (667), 1–11. <https://doi.org/10.1007/s00604-020-04637-5>.
- 875 (41) Yaren, O.; Bradley, K. M.; Moussatche, P.; Hoshika, S.; Yang, Z.; Zhu, S.; Karst, S. M.; Benner, S.
876 A. A Norovirus Detection Architecture Based on Isothermal Amplification and Expanded Genetic
877 Systems. *J. Virol. Methods* **2016**, 237, 64–71. <https://doi.org/10.1016/j.jviromet.2016.08.012>.
- 878 (42) Montero, L.; Smith, S. M.; Jesser, K. J.; Paez, M.; Ortega, E.; Peña-Gonzalez, A.; Soto-Girón, M.
879 J.; Hatt, J. K.; Sánchez, X.; Puebla, E.; Endara, P.; Cevallos, W.; Konstantinidis, K. T.; Trueba, G.;
880 Levy, K. Distribution of *Escherichia coli* Pathotypes along an Urban–Rural Gradient in Ecuador.
881 *Am. J. Trop. Med. Hyg.* **2023**, 109 (3), 559–567. <https://doi.org/10.4269/ajtmh.23-0167>.
- 882 (43) Kotloff, K. L.; Nataro, J. P.; Blackwelder, W. C.; Nasrin, D.; Farag, T. H.; Panchalingam, S.; Wu,
883 Y.; Sow, S. O.; Sur, D.; Breiman, R. F.; Faruque, A. S. G.; Zaidi, A. K. M.; Saha, D.; Alonso, P. L.;
884 Tamboura, B.; Sanogo, D.; Onwuchekwa, U.; Manna, B.; Ramamurthy, T.; Kanungo, S.; Ochieng,
885 J. B.; Omore, R.; Oundo, J. O.; Hossain, A.; Das, S. K.; Ahmed, S.; Qureshi, S.; Quadri, F.;
886 Adegbola, R. A.; Antonio, M.; Hossain, M. J.; Akinsola, A.; Mandomando, I.; Nhampossa, T.;
887 Acácio, S.; Biswas, K.; O’Reilly, C. E.; Mintz, E. D.; Berkeley, L. Y.; Muhsen, K.; Sommerfelt,
888 H.; Robins-Browne, R. M.; Levine, M. M. Burden and Aetiology of Diarrhoeal Disease in Infants
889 and Young Children in Developing Countries (the Global Enteric Multicenter Study, GEMS): A
890 Prospective, Case-Control Study. *The Lancet* **2013**, 382 (9888), 209–222.
891 [https://doi.org/10.1016/S0140-6736\(13\)60844-2](https://doi.org/10.1016/S0140-6736(13)60844-2).
- 892 (44) Pullan, R. L.; Smith, J. L.; Jasrasaria, R.; Brooker, S. J. Global Numbers of Infection and Disease
893 Burden of Soil Transmitted Helminth Infections in 2010. *Parasit. Vectors* **2014**, 7 (37), 1–19.
894 <https://doi.org/10.1186/1756-3305-7-37>.
- 895 (45) Holland, C.; Sepidarkish, M.; Deslyper, G.; Abdollahi, A.; Valizadeh, S.; Mollalo, A.; Mahjour, S.;
896 Ghodsian, S.; Ardekani, A.; Behniafar, H.; Gasser, R. B.; Rostami, A. Global Prevalence of *Ascaris*
897 Infection in Humans (2010–2021): A Systematic Review and Meta-Analysis. *Infec. Dis. Poverty*
898 **2022**, 11 (113). <https://doi.org/10.1186/s40249-022-01038-z>.
- 899 (46) Centers for Disease Control and Prevention. *Estimated Annual Number of Episodes of Illnesses*
900 *Caused by 31 Pathogens Transmitted Commonly by Food, United States*; 2019.
901 www.cdc.gov/EID/content/17/1/7-Techapp4.pdf.
- 902 (47) Wang, R.; Van Dorp, L.; Shaw, L. P.; Bradley, P.; Wang, Q.; Wang, X.; Jin, L.; Zhang, Q.; Liu, Y.;
903 Rieux, A.; Dorai-Schneiders, T.; Weinert, L. A.; Iqbal, Z.; Didelot, X.; Wang, H.; Balloux, F. The
904 Global Distribution and Spread of the Mobilized Colistin Resistance Gene *mcr-1*. *Nat. Commun.*
905 **2018**, 9 (1), 1179. <https://doi.org/10.1038/s41467-018-03205-z>.
- 906 (48) Khan, A. U.; Maryam, L.; Zarrilli, R. Structure, Genetics and Worldwide Spread of New Delhi
907 Metallo- β -Lactamase (NDM): A Threat to Public Health. *BMC Microbiol.* BioMed Central Ltd.
908 April 27, 2017. <https://doi.org/10.1186/s12866-017-1012-8>.
- 909 (49) Rao, G.; Capone, D.; Zhu, K.; Knoble, A.; Linden, Y.; Clark, R.; Lai, A.; Kim, J.; Huang, C.-H.;
910 Bivins, A.; Brown, J. Simultaneous Detection and Quantification of Multiple Pathogen Targets in
911 Wastewater. *PLOS Water* **2024**, 3 (2), e0000224. <https://doi.org/10.1371/journal.pwat.0000224>.
- 912 (50) Capone, D.; Berendes, D.; Cumming, O.; Holcomb, D.; Knee, J.; Konstantinidis, K. T.; Levy, K.;
913 Nalá, R.; Risk, B. B.; Stewart, J.; Brown, J. Impact of an Urban Sanitation Intervention on Enteric
914 Pathogen Detection in Soils. *Environ. Sci. Technol.* **2021**, 55 (14), 9989–10000.
915 <https://doi.org/10.1021/acs.est.1c02168>.
- 916 (51) Lappan, R.; Henry, R.; Chown, S. L.; Luby, S. P.; Higginson, E. E.; Bata, L.; Jirapanjawat, T.;
917 Schang, C.; Openshaw, J. J.; O’Toole, J.; Lin, A.; Tela, A.; Turagabeci, A.; Wong, T. H. F.;
918 French, M. A.; Brown, R. R.; Leder, K.; Greening, C.; McCarthy, D. Monitoring of Diverse Enteric
919 Pathogens across Environmental and Host Reservoirs with TaqMan Array Cards and Standard
920 qPCR: A Methodological Comparison Study. *Lancet Planet. Health* **2021**, 5 (5), e297–e308.
921 [https://doi.org/10.1016/S2542-5196\(21\)00051-6](https://doi.org/10.1016/S2542-5196(21)00051-6).
- 922 (52) Baker, K. K.; Senesac, R.; Sewell, D.; Sen Gupta, A.; Cumming, O.; Mumma, J. Fecal Fingerprints
923 of Enteric Pathogen Contamination in Public Environments of Kisumu, Kenya, Associated with
924 Human Sanitation Conditions and Domestic Animals. *Environ. Sci. Technol.* **2018**, 52 (18), 10263–
925 10274. <https://doi.org/10.1021/acs.est.8b01528>.
- 926 (53) Fuhrmeister, E. R.; Ercumen, A.; Pickering, A. J.; Jeanis, K. M.; Ahmed, M.; Brown, S.; Arnold, B.
927 F.; Hubbard, A. E.; Alam, M.; Sen, D.; Islam, S.; Kabir, M. H.; Kwong, L. H.; Islam, M.; Unicomb,
928 L.; Rahman, M.; Boehm, A. B.; Luby, S. P.; Colford, J. M.; Nelson, K. L. Predictors of Enteric

- 929 Pathogens in the Domestic Environment from Human and Animal Sources in Rural Bangladesh.
930 *Environ. Sci. Technol.* **2019**, *53* (17), 10023–10033. <https://doi.org/10.1021/acs.est.8b07192>.
- 931 (54) Gerba, C. P. Environmentally Transmitted Pathogens. In *Environmental Microbiology*; 2015; pp
932 509–550. <https://doi.org/10.1016/B978-0-12-394626-3.00022-3>.
- 933 (55) Bo□lin, I.; Wiklund, G.; Qadri, F.; Torres, O.; Bourgeois, A. L.; Savarino, S.; Svennerholm, A.-M.
934 Enterotoxigenic *Escherichia coli* with STh and STp Genotypes Is Associated with Diarrhea Both in
935 Children in Areas of Endemicity and in Travelers. *J. Clin. Microbiol.* **2006**, *44* (11), 3872–3877.
936 <https://doi.org/10.1128/JCM.00790-06>.
- 937 (56) Alcock, B. P.; Raphenya, A. R.; Lau, T. T. Y.; Tsang, K. K.; Bouchard, M.; Edalatmand, A.;
938 Huynh, W.; Nguyen, A.-L. V.; Cheng, A. A.; Liu, S.; Min, S. Y.; Miroshnichenko, A.; Tran, H.-K.;
939 Werfalli, R. E.; Nasir, J. A.; Oloni, M.; Speicher, D. J.; Florescu, A.; Singh, B.; Faltyn, M.;
940 Hernandez-Koutoucheva, A.; Sharma, A. N.; Bordeleau, E.; Pawlowski, A. C.; Zubyk, H. L.;
941 Dooley, D.; Griffiths, E.; Maguire, F.; Winsor, G. L.; Beiko, R. G.; Brinkman, F. S. L.; Hsiao, W.
942 W. L.; Domselaar, G. V.; McArthur, A. G. CARD 2020: Antibiotic Resistome Surveillance with the
943 Comprehensive Antibiotic Resistance Database. *Nucleic Acids Res.* **2020**, *48* (D1), D517–D525.
944 <https://doi.org/10.1093/nar/gkz935>.
- 945 (57) Karst, S. M.; Ziels, R. M.; Kirkegaard, R. H.; Sørensen, E. A.; McDonald, D.; Zhu, Q.; Knight, R.;
946 Albertsen, M. High-Accuracy Long-Read Amplicon Sequences Using Unique Molecular Identifiers
947 with Nanopore or PacBio Sequencing. *Nat. Methods* **2021**, *18* (2), 165–169.
948 <https://doi.org/10.1038/s41592-020-01041-y>.
- 949 (58) Liu, J.; Gratz, J.; Amour, C.; Kibiki, G.; Becker, S.; Janaki, L.; Verweij, J. J.; Taniuchi, M.; Sobuz,
950 S. U.; Haque, R.; Haverstick, D. M.; Houpt, E. R. A Laboratory-Developed TaqMan Array Card
951 for Simultaneous Detection of 19 Enteropathogens. *J. Clin. Microbiol.* **2013**, *51* (2), 472–480.
952 <https://doi.org/10.1128/JCM.02658-12>.
- 953 (59) Liu, J.; Gratz, J.; Amour, C.; Nshama, R.; Walongo, T.; Maro, A.; Mduma, E.; Platts-Mills, J.;
954 Boisen, N.; Nataro, J.; Haverstick, D. M.; Kabir, F.; Lertsethtakarn, P.; Silapong, S.;
955 Jeamwattanalert, P.; Bodhidatta, L.; Mason, C.; Begum, S.; Haque, R.; Praharaaj, I.; Kang, G.;
956 Houpt, E. R. Optimization of Quantitative PCR Methods for Enteropathogen Detection. *PLoS One*
957 **2016**, *11* (6), e0158199. <https://doi.org/10.1371/journal.pone.0158199>.
- 958 (60) Amour, C.; Gratz, J.; Mduma, E.; Svensen, E.; Rogawski, E. T.; McGrath, M.; Seidman, J. C.;
959 McCormick, B. J. J.; Shrestha, S.; Samie, A.; Mahfuz, M.; Qureshi, S.; Hotwani, A.; Babji, S.;
960 Trigoso, D. R.; Lima, A. A. M.; Bodhidatta, L.; Bessong, P.; Ahmed, T.; Shakoor, S.; Kang, G.;
961 Kosek, M.; Guerrant, R. L.; Lang, D.; Gottlieb, M.; Houpt, E. R.; Platts-Mills, J. A. Epidemiology
962 and Impact of *Campylobacter* Infection in Children in 8 Low-Resource Settings: Results from the
963 MAL-ED Study. *Clin. Infect. Dis.* **2016**, *63* (9), 1171–1179. <https://doi.org/10.1093/cid/ciw542>.
- 964 (61) Ford, L.; Healy, J. M.; Cui, Z.; Ahart, L.; Medalla, F.; Ray, L. C.; Reynolds, J.; Laughlin, M. E.;
965 Vugia, D. J.; Hanna, S.; Bennett, C.; Chen, J.; Rose, E. B.; Bruce, B. B.; Payne, D. C.; Francois
966 Watkins, L. K. Epidemiology and Antimicrobial Resistance of *Campylobacter* Infections in the
967 United States, 2005–2018. *Open Forum Infect. Dis.* **2023**, *10* (8), ofad378.
968 <https://doi.org/10.1093/ofid/ofad378>.
- 969 (62) Parker, C. T.; Miller, W. G.; Horn, S. T.; Lastovica, A. J. Common Genomic Features of
970 *Campylobacter jejuni* Subsp. *doylei* Strains Distinguish Them from *C. jejuni* Subsp. *jejuni*. *BMC*
971 *Microbiol.* **2007**, *7*, 50. <https://doi.org/10.1186/1471-2180-7-50>.
- 972 (63) Rojas-López, L.; Marques, R. C.; Svård, S. G. *Giardia Duodenalis*. *Trends in Parasitol.* **2022**, *38*
973 (7), 605–606. <https://doi.org/10.1016/j.pt.2022.01.001>.
- 974 (64) Hadfield, S. J.; Robinson, G.; Elwin, K.; Chalmers, R. M. Detection and Differentiation of
975 *Cryptosporidium* Spp. in Human Clinical Samples by Use of Real-Time PCR. *J. Clin. Microbiol.*
976 **2011**, *49* (3), 918–924. <https://doi.org/10.1128/JCM.01733-10>.
- 977 (65) Carey, C. M.; Lee, H.; Trevors, J. T. Biology, Persistence and Detection of *Cryptosporidium*
978 *parvum* and *Cryptosporidium hominis* Oocyst. *Water Res.* **2004**, *38* (4), 818–862.
979 <https://doi.org/10.1016/j.watres.2003.10.012>.
- 980 (66) Lee, G. O.; Eisenberg, J. N. S.; Uruchima, J.; Vasco, G.; Smith, S. M.; Van Engen, A.; Victor, C.;
981 Reynolds, E.; MacKay, R.; Jesser, K. J.; Castro, N.; Calvopiña, M.; Konstantinidis, K. T.; Cevallos,
982 W.; Trueba, G.; Levy, K. Gut Microbiome, Enteric Infections and Child Growth across a Rural–
983 Urban Gradient: Protocol for the ECoMiD Prospective Cohort Study. *BMJ Open* **2021**, *11* (10),
984 e046241. <https://doi.org/10.1136/bmjopen-2020-046241>.

- 985 (67) Oxford Nanopore Technologies. Pod5-File-Format. 2024. [https://github.com/nanoporetech/pod5-](https://github.com/nanoporetech/pod5-file-format)
986 file-format (accessed 2024-08-16).
- 987 (68) Oxford Nanopore Technologies. Dorado. 2024. <https://github.com/nanoporetech/dorado> (accessed
988 2024-08-16).
- 989 (69) Madden, T. The BLAST Sequence Analysis Tool. In *The NCBI handbook 2.5*; 2013; pp 425–436.
- 990 (70) Langmead, B.; Salzberg, S. L. Fast Gapped-Read Alignment with Bowtie 2. *Nat. Methods* **2012**, *9*
991 (4), 357–359. <https://doi.org/10.1038/nmeth.1923>.
- 992 (71) Oxford Nanopore Technologies. Medaka. 2022. <https://github.com/nanoporetech/medaka> (accessed
993 2024-08-16).
- 994 (72) Robinson, J. T.; Thorvaldsdóttir, H.; Winckler, W.; Guttman, M.; Lander, E. S.; Getz, G.; Mesirov,
995 J. P. Integrative Genomics Viewer. *Nat Biotechnol* **2011**, *29* (1), 24–26.
996 <https://doi.org/10.1038/nbt.1754>.
997
998

999