Single-cell resolution genetic association analysis of heterogeneous bacterial communities by utilizing droplet digital PCR

Lara Ambrosio Leal Dutra[1,4], Matti Jalasvuori[1], Ole Franz[1], Kimi Nurminen[1], Pauliina Salmi[1,3], Marja Tiirola[1], Reetta Penttinen[1,2]

[1] Department of Biological and Environmental Science and Nanoscience Center, FI-40014 University of Jyväskylä, Jyväskylä, Finland

[2] Department of Biology, FI-20014 University of Turku, Turku, Finland

[3] Faculty of Information and Technology, FI-40014 University of Jyväskylä, Jyväskylä, Finland

[4] Department of Virology, Faculty of Medicine, FI-00014 University of Helsinki, Helsinki, Finland

Abstract

Microbial communities often respond to various challenges, such as the presence of antibiotics, as a whole. Dissecting these community-level effects into separate acting entities requires the identification of organisms that carry functional genes for the observed feature. However, unculturable microbes abound in various environments, hence making the identification challenging. Here we present a cultivation-free technique that can be utilized to link functional genes with carrying bacterial species at single-cell resolution. The developed protocol is relatively simple to use, utilizes commercially available droplet microfluidics devices, does not require toxic reagents (as compared to some previous methods), eliminates invalid signals emerging from extracellular DNA, and allows simultaneous analysis of community diversity via 16S rRNA gene sequencing. The method can be customized for any given genetic trait to accurately identify its hosting subpopulation from a heterogeneous and potentially uncultivable bacterial community.

Keywords: 16S rRNA gene; droplet digital PCR; droplet microfluidics; functional genes; microbial diversity analysis; single-cell.

Introduction

Microbes are ubiquitous organisms that play important roles in systems shaping life on our planet. Yet, microbes rarely exist in isolation. A single grain of sand contains tens to hundreds of thousands of bacteria of various species (Probandt *et al.* 2017), one gram of dry faecal matter has a billion microbial cells and even very slowly growing microbial mats in lava caves have hundreds of distinct taxonomical units in their thick slimy layers (Lavoie *et al.* 2017). Within these communities, bacteria form interaction networks where individual cells are often unable to reproduce in the absence of their contemporaries (Browne *et al.* 2016). Ultimately, genetic, metabolic and ecological links between microbes shape the communities and their behaviour (Stubbendieck *et al.* 2016). Bacterial systems may, for example, degrade hazardous waste materials, confer resistance to antibiotics and regulate the turnover of limiting nutrients for subsequent levels in trophic cascades. Further, horizontal gene transfer is common in microbial systems and opportunistic genetic characteristics may rapidly switch into more suitable hosts in response to environmental conditions (Mattila *et al.* 2017). Disentangling these communities into individual cells that can still be studied with precision and without the influence of potential confounding factors, such as remnant DNA from dead cells, is challenging.

Various methodological advances have been achieved in recent years to help resolve the roles of individual and potentially unculturable microbes within communities. 16S ribosomal RNA sequencing is the current standard for providing an overall picture of the diversity of microbes – living or dead – in the studied sample. However, it fails to associate genetic characteristics to the species. In contrast, total DNA sequencing can cover a lot of the overall genetic variation, but usually without a definitive answer to the question of who acted as the cellular vehicle for any particular gene – or if that gene was indeed actually a part of a living organism at all or merely a remnant from a deceased microbe (Lennon *et al.* 2018). Further, (mobile) plasmids, in particular, replicate separately from chromosomes in bacterial cells, and hence, metagenomic sequencing fails to identify the host of the potentially crucially important genetic element (Jalasvuori 2012). To link genetic traits with particular species, a number of microfluidic techniques have been utilized to identify target traits (such as the ability to cause a disease or carry bacteriophages) from heterogeneous bacterial samples (Zeng *et al.*

60 2010; Tadmor *et al.* 2011). Ottesen and colleagues utilized a chip containing 1176
61 microchambers to divide termite midgut samples to single-cell resolution and used two
62 fluorescent probes to determine the chambers containing 16S rRNA and/or targeted functional
63 gene (Ottesen *et al.* 2006). By using a needle to extract the contents of chambers that expressed
64 both fluorescent signals, they were able to associate 16S rRNA sequence with the functional
65 gene. Alternatively, fixed bacterial samples have been used for fluorescent probe hybridization
66 (Baptista *et al.* 2014). Hybridization methods, however, are relatively laborious and unable to
67 study a large number of cells. In order to study hundreds of thousands of cells from
68 uncharacterized bacterial samples, Spencer and colleagues developed a method to trap
69 individual cells into tiny acrylamide compartments prior to PCR amplification (Spencer *et al.*
70 2016). In this so-called epicPCR, a very localized PCR reaction is utilized to amplify targeted
71 functional gene and combine the gene in concatenation PCR with the 16S ribosomal RNA
72 gene. By sequencing the paired products, the cellular hosts harbouring the studied genetic trait
73 can be identified. While epicPCR is a promising tool for large-scale single-cell resolution
74 analyses, it still requires specific knowhow of emulsion preparation and subsequent analytical
75 steps.

76 Droplet microfluidics utilizes microchannels for the formation of nanolitre-scale
77 droplets. In droplet digital PCR (ddPCR), PCR reagents are enclosed into droplets surrounded
78 and stabilized by oil, enabling thousands of PCR reactions to occur simultaneously. The
79 amplified DNA is detected by using double-stranded DNA-binding fluorescent dye and
80 measured from each droplet one by one. As the sample is randomly distributed across the
81 droplets, both empty (negative) and DNA-containing (positive) droplets are formed. The
82 system is extremely sensitive and capable of detecting even a single copy within a droplet.
83 Finally, the fraction of positive droplets is compared to the total number of droplets which is
84 then used with Poisson statistics for absolute quantification of target DNA in the original
85 sample (Pinheiro *et al.* 2012; Hindson *et al.* 2011). DNA can be delivered into droplets also
86 within intact cells instead of pure DNA, allowing ddPCR to be applied for bacterial
87 communities. This principle has been harnessed to trace enterohemorrhagic *Escherichia coli*
88 (EHEC) in food production samples (Capobianco *et al.* 2020; McMahon *et al.* 2017). In these

89 studies, single cells were delivered into droplets, and after the cell lysis initiated by the
90 denaturation step of ddPCR, EHEC was detectable via distinct amplification of two virulence
91 genes and their colocalization in the same droplet with fluorescent probes (McMahon *et al.*
92 2017).

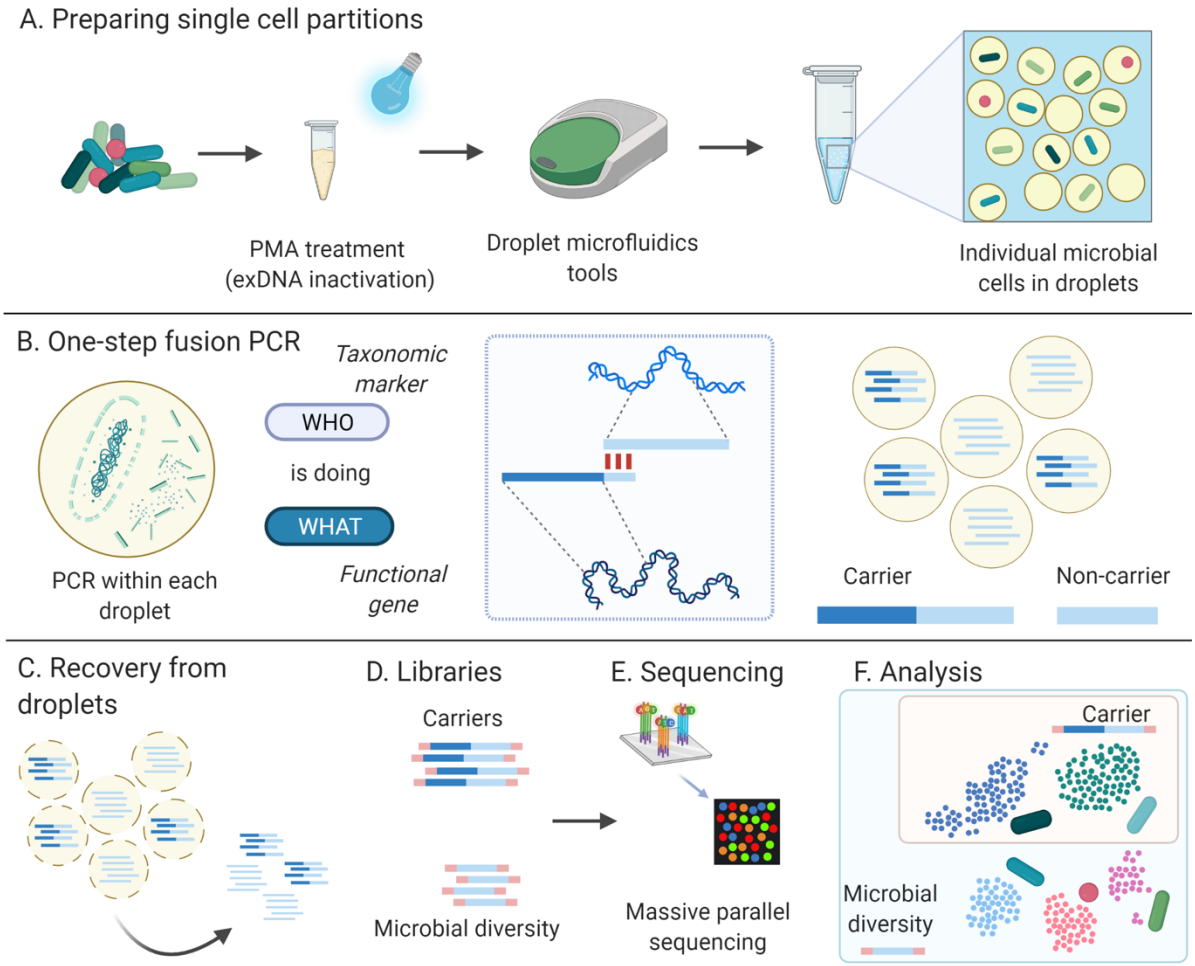
93 In this study, we present a new adapted technique for investigating bacterial
94 communities at single-cell resolution. The described method utilizes commercially available
95 ddPCR devices (based on droplet microfluidics), is fast to perform, requires only non-toxic
96 reagents, contains steps to efficiently inactivate extracellular DNA from samples, employs
97 paired fusion PCR to link taxonomic marker to the target gene and allows accurate
98 fluorescence-based verification of single-cell resolution prior to fusion amplicon sequencing.
99 We demonstrate that the method is applicable for precise detection and identification of carriers
100 of a beta-lactamase resistance gene (*bla*CTX-M-14) in the heterogenous bacterial sample, and
101 thus, can be utilized for tracing the movements of mobile genetic elements such as plasmids
102 within the bacterial community. As such, the introduced single-cell ddPCR method provides a
103 convenient approach for different laboratories to identify specific genetic traits from multi-
104 species communities.
105

106 **RESULTS**

107 **A microbial single-cell screening from a community**

108 In this study, we developed an easy-to-use microbial single-cell screening method to identify
109 or track hosts of any gene of interest in heterogenous communities. Our experimental setup
110 focused on tracking an antibiotic resistance gene (ARG) *bla*CTX-M-14 (hereafter *bla*). A
111 fragment of *bla* is paired with the 16S rRNA gene in fusion PCR to identify taxonomic groups
112 that harbour the ARG. For enabling analysis of individual cells, the bacterial samples are
113 partitioned into compartments by utilizing QX200™ Droplet Digital PCR System from Bio-
114 Rad. This automated system employs droplet microfluidics for generating equal-sized 1
115 nanolitre droplets that encapsulate both individual bacterial cells and PCR reagents (Figure 1).

116 During thermal cycling, the bacterial DNA is released from cells via lysis and used as a
117 template for paired amplification of target ARG and 16S rRNA gene. As single-cell PCR-based
118 methods are highly sensitive to any contaminating DNA present in the droplet but outside of
119 the cell, the protocol includes steps for inactivating extracellular DNA. These steps were
120 observed to be crucial for the accurate detection of genetic content that is indeed bound within
121 cellular membranes. By sequencing the paired amplicons, the studied bacterial community can
122 be allocated into gene-carriers and gene-free individuals. Four bacterial species were employed
123 in the development and optimization of the described method: *Escherichia coli* (carrying the
124 target *bla* gene in a conjugative plasmid pEC13), *Pseudomonas fluorescens, Salmonella*
125 Typhimurium and *Klebsiella pneumoniae*.



126

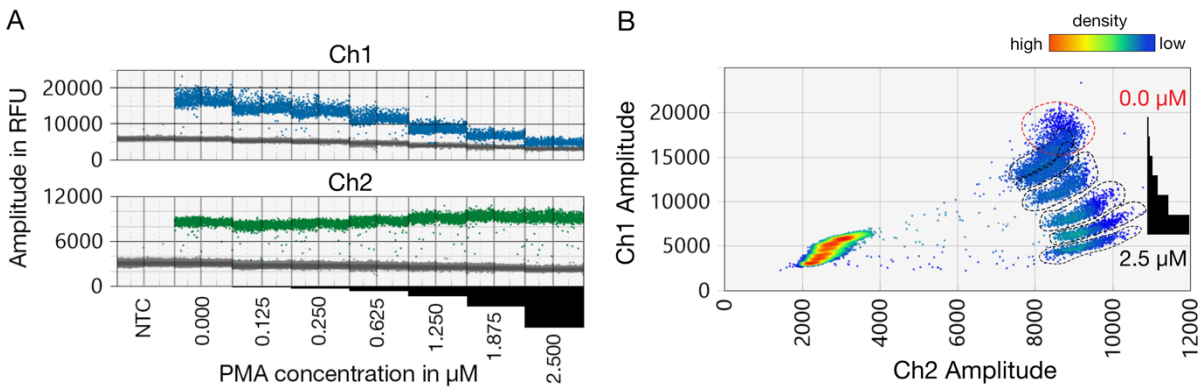
127 **Figure 1. Overview of the single-cell ddPCR method workflow.** A) Bacterial samples are
128 prepared and partitioned using droplet microfluidics technology to obtain a single cell per
129 droplet. This step includes treatment with PMA to inactivate exDNA (extracellular DNA), the
130 addition of PCR reagents and, finally, the generation of single-cell -containing droplets with a
131 droplet microfluidics device. B) One-step fusion PCR assay is carried out within each droplet.
132 In droplets containing a carrier cell, *i.e.* a cell that potentially has the functional gene, the
133 taxonomic marker gene and the functional gene are amplified and fused into a concatemer PCR
134 product. The taxonomic marker is the sole product in droplets containing non-carrier bacterial
135 cells. C) The produced amplicons are recovered from droplets by breaking the emulsion using
136 liquid nitrogen method (Dutra *et al.* 2020). D) The recovered amplicons are used to construct
137 sequencing libraries and E) sequenced by massive parallel sequencing. F) The sequence
138 analysis determines the bacterial population carrying the functional gene as well as the overall
139 bacterial diversity in the given sample. Created with BioRender.com.
140

141 **Extracellular DNA (exDNA) inactivation from cell samples**

142 Microfluidic encapsulation traps any content in the sample, including free DNA. Preliminary
143 analysis (not shown) indicated that free DNA (*i.e.* DNA outside of bacterial cells) is abundant
144 in bacterial suspensions even after DNAse treatment and/or repeated washing of the cells. As
145 such, this DNA forms a notable confounding factor for single-cell analyses that aim to reliably
146 identify a specific DNA target that is indeed within the cellular membranes (*i.e.* a single viable
147 cell). To this end, we ensured that the process only amplifies DNA originating from a unique
148 cell by treating the samples with propidium monoazide (PMA). PMA is a DNA binding dye
149 that covalently attaches to double stranded DNA (dsDNA) via photoactivation, thus preventing
150 primer annealing and polymerization to occur. Cell membranes are impermeable to PMA,
151 hence only DNA external to intact cells is rendered non-amplifiable.

152 ***PMA treatment – fluorescence emission spectral shift***

153 After initial screening, we hypothesized that PMA may influence PCR. As such, the maximum
154 non-inhibitory concentration of PMA in PCR amplification was assessed by adding increasing
155 concentrations of photoactivated PMA to ddPCR EvaGreen -based assay (Supplementary
156 Figure 1). The reactions targeting the 16S rRNA region (amplified with 27F and 338R primers)
157 of digested *E. coli* DNA were prepared with 0.125 µM up to 2.5 µM of PMA. The Relative
158 Fluorescence Unit (RFU) was obtained from each droplet by the Droplet Reader through
159 fluorescence detection channels 1 (Ch1) and 2 (Ch2). Analysis of RFU showed that two droplet
160 clusters could be seen: a cluster of negative droplets where amplification had not occurred (low
161 RFU) and a cluster of positive droplets with successful amplification (high RFU). Nonetheless,
162 in Ch1, a pronounced drop in RFU from positive droplets cluster could be seen when increasing
163 concentrations of PMA were used (Figure 2A). The same pattern could not be observed in Ch2
164 data. Here, instead, the positive droplets cluster showed an increase in RFU in higher PMA
165 concentrations. The analysis of Ch1 *versus* Ch2 (Figure 2B) showed a clear distinction between
166 positive and negative droplet clusters. These results suggest that PMA does not inhibit PCR
167 reactions when used in concentrations up to 2.5 µM but affects the fluorescence emission
168 spectrum that disturbs the detection of the DNA binding dye (EvaGreen) from Ch1.

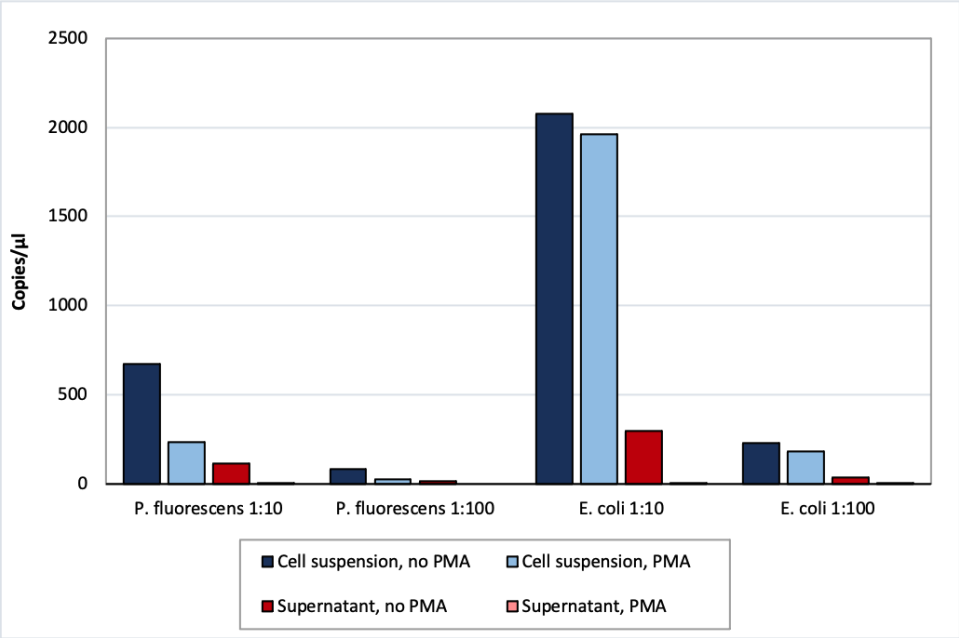


169

170 **Figure 2. Effect of PMA concentration on RFU reading.** Light-inactivated PMA in final
171 concentration ranging from 0.125 to 2.5 µM was added to 16S rRNA ddPCR. After

172 amplification, the fluorescence of each droplet was measured through both detection channels.
173 The increasing PMA concentration resulted in a drop in RFU when measured via channel 1
174 (Ch1) (A, upper panel), but not with channel 2 (Ch2) (A, lower panel). The positive droplets
175 were identified by two-dimensional plotting of RFU values detected from Ch1 and Ch2
176 channels (B).

177 Further, we studied whether PMA inactivates exDNA in samples containing bacterial
178 cells. Cell suspensions of *E. coli* and *P. fluorescens* and their filtrates (suspension filtered
179 through a 0.2 µm filter) were treated with 2 µM PMA and used as a template in ddPCR
180 amplifying a region within the gene for 16S rRNA. For controls, the corresponding samples
181 were treated with sterile water instead of PMA. PMA treatment reduced the obtained 16S rRNA
182 concentration by as much as 71 % in *P. fluorescens* and 22 % in *E. coli* cell samples (Figure
183 3). This suggests that exDNA is abundant even in pure bacterial cultures and potentially causes
184 erroneous PCR amplification if the DNA is present inside a droplet together with the trapped
185 bacterial cell. To verify the complete inactivation of exDNA, we filtered out the bacterial cells
186 and exposed the filtrate to PMA. Less than 0.12 % of the original contaminating DNA was
187 amplifiable after PMA treatment (Figure 3).



188

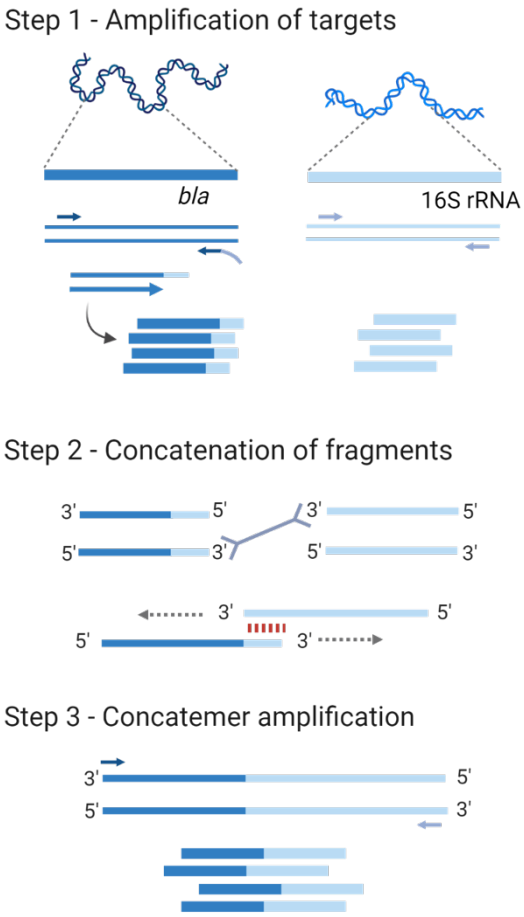
189 **Figure 3. Effect of PMA treatment on inactivation of exDNA.** *Pseudomonas fluorescens*
190 and *Escherichia coli* cell dilution (1:10 or 1:100 of an overnight-grown bacterial culture) and
191 their filtrated (0.2 µm) supernatant treated with 10 µM PMA or sterile water as control. 16S
192 rRNA was amplified with single-cell ddPCR method and is presented as copies/µl.
193

194 **One-step fusion PCR**

195 The developed method relies on linking two targeted DNA fragments together by using fusion
196 PCR. To produce the concatemer, we employed one-step fusion PCR with four primers. In this
197 approach, the amplification of both targets occurs within the same reaction. Introduction of a
198 fusion primer into the reaction prompts the attachment of the distinct target fragments into a
199 unified DNA molecule (Figure 4). The fusion primer consists of a priming sequence for a given
200 target sequence added by additional bases at the 5'-end that encodes the complementary reverse
201 sequence of the forward primer of the other target. The dynamics of the reaction is driven by a
202 difference in the ratio of forward to reverse primer (within a pair) and the concentration of each
203 primer pair. The outer primers that flank the concatemer sequence are used in higher
204 concentrations than the inner primers. Amplification of the target sequence will eventually
205 diminish due to the depletion of inner primers, thus the matching overhang will prompt the
206 fusion of both target sequences into a paired concatemer. The excess of concatemer flanking
207 primers will provide amplification of the linked sequence.

208 In our experiments, we used the 16S rRNA universal primers, 27F and 338R, to amplify
209 a variable region that can be used to identify the used bacterial strains. As for ARG target, we
210 designed primers 293F and 320R for amplifying a 101 bp region of *bla* gene. Furthermore, we
211 modified *bla* reverse primer 320R with a 5'-extension of the reverse complementary sequence
212 of 27F (therefore producing a fragment of 121 bp). The complementary sequence to be
213 produced during PCR was to act as primer and allow annealing of *bla*-27F overhang onto 16S
214 rRNA fragment, thus fusing both target sequences as a concatemer.

215



216

217 **Figure 4. One-step fusion PCR with four primers.** The amplification and fusion of 16S
218 rRNA and beta-lactamase gene *bla* are carried out in the same reaction. The fusion process is
219 driven by a fusion primer that is one of the *bla* primers with a 5' -end extension that encodes
220 the reverse complement sequence of the 16S rRNA primer 27F. In step 1, both target regions
221 are simultaneously and independently amplified, generating the intermediate product of the
222 fusion: the *bla* fragment with the tail complementary to 27F. In step 2, the tail will prompt the
223 fusion of 16S and *bla* products. Lower concentration of the inner primers (27F and the modified
224 *bla* primers that bind to the fusion region) facilitates the reaction dynamic to produce the
225 concatemer. After the depletion of the inner primers, in step 3, the amplification of the fusion
226 product is carried out by using the excess of external primers. Created with BioRender.com.

227

228 ***Optimization of primer sets***

229 The optimization of one-step fusion PCR was carried out by screening the primer ratio and
230 concentration to assess the efficiency of concatemer production. First, the pairs were tested
231 separately to verify the production of each target in different reaction conditions. After this, the
232 primers were combined in a single ddPCR reaction to ensure the amplification of the
233 concatemer. The separation of positive and negative clusters was evaluated to define the
234 concentration and ratio to be used in the following assays. Clustering was more efficient at
235 lower ratios and with higher primer concentrations (Supplementary Figure 3).

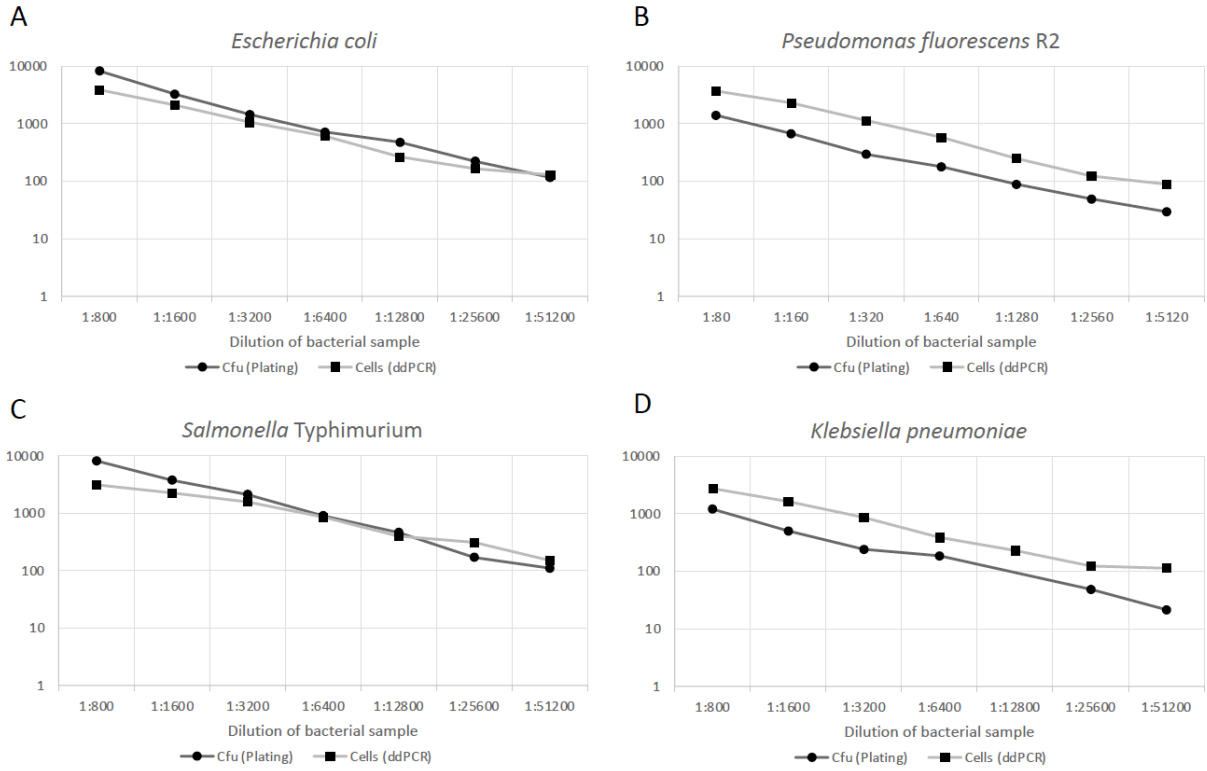
236 In ddPCR, the amplification of DNA fragments of distinct sizes is indicated by RFU
237 level. Thus, the occurrence and separation of four droplet clusters were sought: a cluster with
238 very low RFU was expected for negative droplets; two clusters with intermediate RFU were
239 expected to result from droplets where a single target was amplified, either a low RFU cluster
240 for *bla* (121 bp product) or higher RFU cluster for 16S rRNA (350 bp product). The cluster
241 with the highest RFU indicated the amplification of the fused concatemer (of 450 bp) (for
242 detailed information, see Supplementary Figure 3). Finally, the conditions for one step fusion
243 reaction in ddPCR were defined by using 225 nM of 16S rRNA primers in 1:5
244 (forward:reverse) ratio, and 200 nM of *bla* primers in 10:1 (forward:reverse) ratio.

245

246 **Partition of single cells into droplets**

247 In order to verify whether single cells could be delivered to droplets with the ddPCR
248 microfluidic system, we designed an assay where PMA-treated dilutions of *E. coli,*
249 *Pseudomonas fluorescens, Salmonella* Typhimurium and *Klebsiella pneumoniae* cells were
250 used as template in the ddPCR targeting the 16S rRNA gene. Serial plating assay was prepared
251 in parallel, and the obtained cell density compared to that estimated with the ddPCR. With all
252 tested bacterial strains, the bacterial cell density was comparable when measured with serial

253 plating and 16S rRNA single-cell ddPCR; the decreasing trend was detectable between
254 adjacent dilutions. However, in *P. fluorescens* and *K. pneumoniae*, a higher cell concentration
255 was consistently observed with ddPCR in comparison to plating (Figure 5).



256

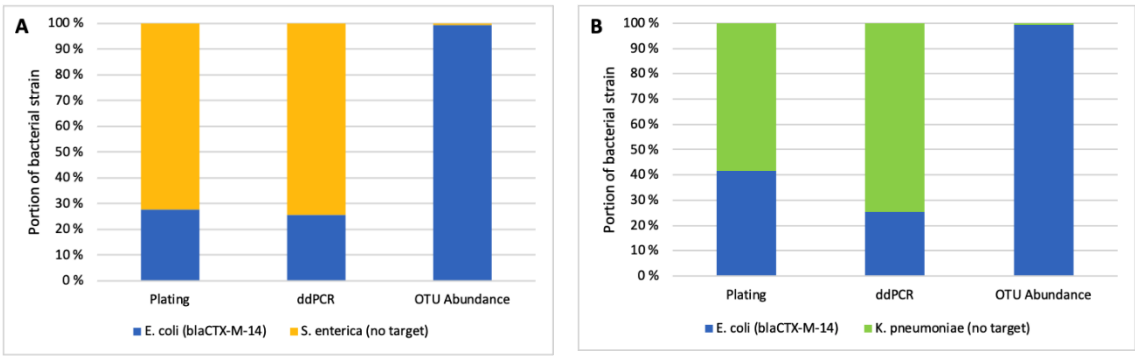
257 **Figure 5. Single cells of *Escherichia coli, Pseudomonas fluorescens, Salmonella***
258 **Typhimurium and *Klebsiella pneumoniae* within droplets.** A dilution series of overnight-
259 grown bacterial cultures were enclosed into droplets and amplified with 16S rRNA single-cell
260 ddPCR. The cell number per ddPCR sample was determined by 16S rRNA ddPCR and is
261 presented as cells. Parallel serial plating of bacterial cultures on LB agar was performed to
262 measure cfu (colony forming units) per sample.

263

264 **Identifying the ARG carriers by concatemer sequencing**

265 To investigate how effectively the single-cell ddPCR method identifies target gene hosts from
266 a heterogeneous sample, *E. coli* cells carrying *bla*CTX-M-14 were mixed with *S.* Typhimurium
267 and *K. pneumoniae* cells not carrying the gene. Bacterial cell suspensions were first pre-treated

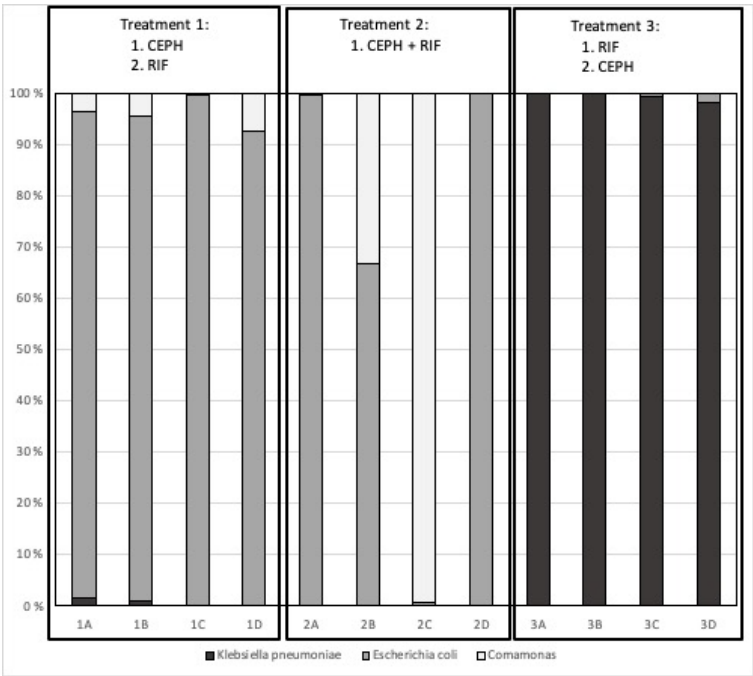
268 with PMA and immediately added to ddPCR reaction to include either a single strain or two
269 strains: *E. coli* with either *S.* Typhimurium or *K. pneumoniae.* The portion of each strain was
270 determined by i) serial plating of each bacterial pure culture, ii) single-cell ddPCR amplifying
271 16S rRNA and iii) OTU abundance (determined via sequencing of bla_16S concatemer
272 produced in single-cell ddPCR) (Figure 6). The portion of *E. coli* cells in the samples ranged
273 from 28-61% when estimated by separate serial plating of the pure cultures. A comparable ratio
274 was obtained with ddPCR from a mix of *E. coli* and S. Typhimurium cells, however from *E.*
275 *coli* and *K. pneumoniae* cell mixes, ddPCR resulted in higher portion of *K. pneumoniae* cells.
276 The mixed cell samples were delivered into single-cell ddPCR where bla_16S concatemers
277 were produced and proceeded to library preparation and sequencing. The 16S rRNA part of the
278 concatemer was used for identifying the host strain associated with *bla*CTX-M-14 and the
279 abundance of *E. coli*, *S.* Typhimurium or *K. pneumoniae* hosts was determined in each sample.
280 In the mixed samples, the portion of noncarrier (*S.* Typhimurium or *K. pneumoniae*) hosts was
281 0.47-1.15%, suggesting that single-cell ddPCR process effectively diminish the occurrence of
282 false positive host sequences.



284 **Figure 6. Portions of *Escherichia coli, Salmonella* Typhimurium (A) and *Klebsiella***
285 ***pneumoniae* (B) cells in mixed heterogeneous samples determined by serial plating, 16S**
286 **rRNA single-cell ddPCR and bla_16S concatemer single-cell ddPCR and sequencing.** Cell
287 density (colony forming units/ml) from each pure culture was measured separately by serial
288 dilution and plate counting. 16S rRNA single-cell ddPCR was used to measure the cell density

289 of each pure culture separately to determine the portion of each bacterial strain in mixed
290 sample. Single-cell ddPCR with primers producing bla_16S concatemer followed by fusion
291 amplicon sequencing was utilized to identify and quantify the *bla*CTX-M-14 carriers in mixed
292 samples. The fraction of *bla*CTX-M-14 hosts is presented as OTU abundance.
293
294 **Horizontal transfer of betalactamase *bla*CTX-M-14 in synthetic bacterial community**
295 To demonstrate the usage of the here-developed method for tracking horizontal transfer of
296 conjugative plasmid within microbial community, we set up a synthetic bacterial community,
297 consisting of *Klebsiella pneumoniae, Comamonas* sp. and *Salmonella* Typhimurium strains,
298 all of which were resistant to rifampicin. *E. coli* harboring conjugative plasmid pEC13 that
299 confers cephalothin resistance to its host via *bla*CTX-M-14 gene, was harnessed as plasmid
300 donor as it is previously shown to possess evolutionary rescue potential (Ruotsalainen *et al.*
301 2020; Mattila *et al.* 2017). Namely, pEC13 is able to transfer to new bacterial hosts even during
302 antibiotic treatment, reviving susceptible bacteria from its proximity. The synthetic
303 communities were either sequentially or simultaneously exposed to cephalothin and rifampicin
304 antibiotics as follows: *K. pneumoniae, Comamonas* sp. and *Salmonella* Typhimurium were co-
305 cultured (N=4) in 1) cephalothin, 2) cephalothin and rifampicin (simultaneous exposure) or 3)
306 rifampicin. *E. coli* carrying pEC13 was introduced into the communities after 30 minutes,
307 allowing evolutionary rescue to occur. After 1 h, treatments 1 and 3 with sequential antibiotic
308 treatment, were supplied with rifampicin and cephalothin, respectively. Synthetic communities
309 were cultivated 3.5 h after which pEC13 transconjugant population was identified via the
310 developed single cell ddPCR method (Figure 7).
311 The bacterial population carrying pEC13 was strongly affected by the order by which
312 the antibiotics were introduced into the population (Figure 7). In treatment 1, where the
313 community was first exposed to cephalothin, the plasmid carrier population was dominated by

314 the original host, *E. coli*, whereas in treatment 3 (first exposure to rifampicin) the most
315 abundant plasmid carrier species was *K. pneumoniae*. Simultaneous exposure to both
316 antibiotics led to relatively stochastic results, the most abundant hosts being *E. coli* and
317 *Comamonas* sp. However, further pairwise conjugation tests showed that pEC13 conjugation
318 between *E. coli* and *Comamonas* sp. was unlikely (data not shown).



319

320 **Figure 7. Carrier population of *bla*CTX-M-14 resistance gene after sequential or**
321 **simultaneous antibiotic exposure.** Rifampicin resistant mock communities consisting of
322 *Klebsiella pneumoniae, Comamonas* sp. and *Salmonella* Typhimurium were exposed to
323 antibiotics in three different treatments (N=4): 1) cephalothin (CEPH), 2) cephalothin and
324 rifampicin (RIF) or 3) rifampicin. After 30 minutes, *Escherichia coli* (pEC13 with *bla*CTX-M-
325 14) was introduced into the communities. Second antibiotic exposure was given after 1 h with
326 1) rifampicin and 3) cephalothin. Horizontal transfer of pEC13 within the community was
327 determined with single cell ddPCR (*bla*CTX-M-14 and 16S rRNA fusion) after 3.5 h co-culture
328 and is presented as the frequency of plasmid host species.

329

330

331 **DISCUSSION**

332 Characterizing microbial communities at the resolution of their true operating units, individual
333 cells, is essential for understanding the functions and biological interactions that take place in
334 complex systems. In this study, we provide a stepwise pipeline that utilizes droplet
335 microfluidics for partitioning bacterial samples into single-cell -containing compartments. The
336 method employs fusion PCR that links 16S rRNA taxonomic marker with the target functional
337 gene (here an antibiotic resistance gene *bla*CTX-M-14), thus enabling the identification of
338 bacteria carrying the target gene after sequencing the paired amplicons.

339 Microbial communities, albeit natural or experimental, are supplemented with exDNA
340 that can prevent accurate analysis of genetic material that solely resides within the boundaries
341 of individual viable cells. exDNA does not only originate from lysed cells, but it is also actively
342 released from bacteria because of different functionalities in the microbial system, such as to
343 aid biofilm formation or mediate horizontal gene transfer (see Nagler *et al.* 2018 and Ibáñez de
344 Aldecoa *et al.* 2017 for review). To exemplify, Carini and colleagues demonstrated that
345 extracellularly located relic DNA may significantly bias the observed microbial diversity in
346 soil samples (Carini *et al.* 2016). This bias, however, was shown to be corrected by treating the
347 soil samples with PMA. PMA is a DNA-intercalating dye that only enters dead cells or cells
348 with compromised membranes and hence it has been used in microbiology to identify viable
349 bacteria. The covalent binding of PMA to DNA occurs via photoactivation, preventing the
350 DNA from serving as a template for PCR (Nocker *et al.* 2007). While exDNA skews the
351 community analyses when conducted on total microbial DNA, it may also disrupt single-cell
352 methodologies that rely on DNA.

353 To investigate the prevalence of exDNA in overnight cultures of *E. coli* and *P.*
354 *fluorescens*, we filtered out the cells and performed 16S rRNA ddPCR for the remaining
355 supernatant. Cell-free supernatant contributed 14-17 % of the total 16S rRNA gene

356 concentration when compared to the corresponding sample where cells were still present. In a
357 preliminary analysis, DNAse treatment was not sufficient in reducing the amplification of
358 exDNA. Hence, the possibility to employ PMA in ddPCR was evaluated. This was done by
359 determining a concentration that did not inhibit the PCR, but effectively blocked the
360 amplification of exDNA in cell-containing and filtrated cell-free samples. The presence of
361 PMA in ddPCR caused a shift in the emitted fluorescence spectrum, however, an accurate
362 categorization of the droplets into positive and negative clusters was still achieved with the 2-
363 dimensional plot of the droplet-emitted fluorescence using two detection channels. As such,
364 potential errors stemming from exDNA can be significantly reduced or eliminated with PMA
365 also in single-cell methods where individual cells and exDNA are trapped within the same
366 compartment.

367 Production of paired concatemer requires a special primer design that allows
368 amplification and fusion of two physically distant genetic regions. The here-employed protocol
369 includes four primers for amplifying both 16S rRNA and target *bla* regions that are then
370 combined with a fusion primer. Four-primer setup changes the reaction dynamics compared to
371 the three-primer protocol utilized in comparable methods such as epicPCR (Spencer *et al.*
372 2016). In the three-primer protocol, the production of fused concatemer is dependent on the
373 presence of the target gene; in case the trapped bacterial cell does not carry it, neither the 16S
374 rRNA gene nor fusion amplicons are produced. The addition of a fourth primer (in this case,
375 27F) not only enables the production of 16S rRNA gene fragments of noncarrier cells, but also
376 increases the yield of the fusion product (data not shown). Higher outer primer concentration
377 (compared to inner primers) may enhance the production of fusion concatemer (Liu *et al.*
378 2018). Thus, we tested several primer ratios and total concentrations in order to achieve the
379 most efficient production of bla_16S concatemer. The results showed that the best fusion
380 concatemer production and droplet separation was obtained when outer primer concentrations

381 were several times higher compared to that of inner primers (forward to reverse primer ratios
382 10:1 for *bla* and 1:5 for 16S rRNA primers, where *bla* forward and 16S rRNA reverse primers
383 acted as outer primers). The application of the here-described system for alternative genetic
384 targets may require optimization of primer concentrations and ratios in order to enhance
385 amplification for better DNA yield and for allowing sufficient separation of negative droplets
386 from positive ones in the subsequent fluorescence analysis.

387 In this method, the droplets are generated with a microfluidic device that produces
388 equal-sized droplets with a volume of one nanolitre. Bacterial cells are encapsulated into oil-
389 surrounded droplets during droplet generation. The accuracy of ddPCR to quantify DNA
390 concentrations assumes that targets are randomly distributed into droplets of equal size. In the
391 presented approach, individual cells (instead of gene copies) serve as the quantifiable unit. As
392 such, we first assured the applicability of single-cell ddPCR individually for *E. coli, P.*
393 *fluorescens, K. pneumoniae* and *Salmonella* Typhimurium. To investigate the capability of the
394 method to differentiate *bla*-carriers from non-carriers, *E. coli bla*-carriers were mixed with non-
395 carriers in different ratios and combinations. These mixtures were immediately divided into
396 droplets and the droplets were transferred to a thermal cycler. We employ a prolonged initial
397 denaturation at 95 °C in the ddPCR thermal cycling protocol to ensure cell lysis and
398 accessibility of PCR reagents to the released DNA. As noted above, during PCR, the four
399 primers enable the production of both fusion concatemer from carrier cells and the sole 16S
400 amplicon from noncarriers. As fluorescence is measured after PCR from each droplet one by
401 one, each cell-containing droplet can be detected and hence the ratio of positive to negative
402 droplets is quantified unlike in methods that rely solely on statistical probabilities. The
403 advantage of the QX200 system is that, in addition to droplet fluorescence measurement, it also
404 inspects droplet shape and size and excludes abnormal droplets from the analysis. In
405 comparable methods, such as in epicPCR and OIL-PCR (Spencer *et al.* 2016; Diebold *et al.*

406 2021), partitions of various sizes are formed, hence producing compartments with various
407 ratios of PCR reagents to the target DNA. This, in turn, may result in notably differing amounts
408 of amplified DNA that are produced from a single target cell. As such, the here-described
409 method can provide also a somewhat more reliable proxy of the number of positive targets in
410 the initial sample, not only their diversity. Nevertheless, Poisson distribution is used to select
411 a suitable number of cells to be loaded into a single ddPCR reaction. In other words, a cell
412 concentration that statistically should enclose only a single cell in a droplet while most droplets
413 remain cell-free is used to seed the droplet generation. We confirmed single-cell resolution by
414 parallel single-cell 16S ddPCR to reveal the potential maximum number of positive droplets
415 produced from the sample and proceeded with only those samples that had a target
416 concentration within the confidence interval.

417 After single-cell ddPCR, the DNA is recovered from all droplets in a single step, thus
418 mixing their contents. At this point, the overall DNA concentration is typically extremely low
419 and needs downstream amplification and barcoding steps to enable amplicon sequencing. Four-
420 primer setup produces single 16S rRNA gene products that can serve as a primer in the
421 subsequent processing and hence may result in false positive host sequences. To prevent cross-
422 binding of fusion amplicons and 16S rRNA products originating from distinct droplets, we
423 used blocking primers in the library preparation steps as described by Spencer *et al.* (2016).
424 Excessive concentration of blocking primers binding to both strands of the 27F region was used
425 to prevent the amplification to start from the 16S rRNA region (*i.e.* in the middle of fusion
426 product, single 16S rRNA gene amplicon or genomic DNA), as in this phase, only fused
427 concatemers are to be enriched. Sequencing of full-length fusion amplicons from mixed cell
428 samples revealed that the vast majority of *bla* host sequences belonged to *E. coli*, which was
429 the only true host carrying the resistance gene. To conclude, our results suggest that the steps
430 for eliminating signals that would potentially cause biased host diversity were sufficient to

431 diminish the false host sequences. However, it should be noted that especially in the case of
432 unknown complex communities, the extremely rare hosts of the target functional gene may be
433 shadowed by the background 16S rRNA gene sequences.

434 To demonstrate its applicability, we utilized the single cell ddPCR method for tracing
435 the horizontal transfer of conjugative resistance plasmid pEC13 within a simple synthetic
436 microbial community consisting of cephalothin resistant (rifampicin sensitive) *E. coli*(pEC13)
437 and rifampicin resistant (cephalothin sensitive) strains of *P. fluorescens, K. pneumoniae* and
438 *Salmonella* Typhimurium. The communities were exposed to three different treatments: 1)
439 cephalothin and after 1 h 30 min to rifampicin, 2) simultaneously to cephalothin and rifampicin
440 or 3) rifampicin and after 1 h 30 min to cephalothin. *E. coli*(pEC13 with *bla*CTX-M-14) was
441 introduced after 30 min of the first antibiotic exposure. Carriers of *bla*CTX-M-14, and thereby
442 the potential evolutionary rescue via pEC13, were identified with the single cell ddPCR. When
443 exposed to cephalothin, the susceptible synthetic population is expectedly diminished. Further,
444 the introduction of *E. coli* with a potential to rescue its accompanying bacteria with the
445 resistance-conferring plasmid, new transconjugant hosts do not seem to appear in the
446 community. In contrast, when rifampicin is administrated as the first antibiotic, the initial
447 resistant population is restored, allowing the spread of rescuing pEC13 plasmid into new hosts
448 before its original host (susceptible to rifampicin) is eliminated from the population. The
449 second antibiotic exposure to cephalothin selects for those bacteria that took in the resistance
450 plasmid and were mainly identified as *K. pneumoniae.* This is in accordance with our previous
451 work where pEC13 was shown to rescue *K. pneumoniae* cells even under lethal cephalothin
452 concentration (Ruotsalainen *et al.* 2020). The simultaneous exposure to both antibiotics is
453 lethal to all bacterial strains in the community, and therefore assumed not to give enough time
454 for conjugation events. Plasmid host population from this treatment was drastically inconsistent
455 between biological replicates, suggesting that the observed results were likely biased by dead

456 cells dominating the bacterial sample. Moreover, *Comamonas* was not observed to host pEC13
457 plasmid, however it was over-represented in some samples, suggesting that certain bacterial
458 behavior, such as secretion of extracellular vesicles, may have significant effects on sequencing
459 results when viable host cells are rare in the community. These results show the method is
460 useful for identifying the plasmid-carrying population which is shaped by the used antibiotics
461 as well as the resistance phenotypes of the initial community compartments. Hence, the method
462 has potential for future applications *e.g.* for studying plasmid-host interactions in more
463 complex microbial communities, such as the gut or soil microbiota – yet it is acknowledged
464 diverse sample types may require further pre-processing steps.

465 In summary, we have outlined a ddPCR-based protocol that enables both counting the
466 number of bacterial cells as well as providing a phylogenetic assignment of functional genes
467 or extrachromosomal elements (plasmids) to their hosts. Available in many molecular biology
468 laboratories, the ddPCR system offers a valuable droplet microfluidic device and necessary
469 plasticware for high-throughput genomic research of microbial communities.

470

471 **MATERIAL AND METHODS**

472 ***Bacterial strains and cultivation***

473 Bacterial strains with and without the target beta-lactam resistance gene were used in the
474 development of the gene tracking method presented in this study. *Escherichia coli* K-12
475 BL21Gold carrying the conjugative plasmid pEC13 with the target beta-lactamase gene
476 *bla*CTX-M-14 (Mattila *et al.* 2017) served as the antibiotic resistance gene carrier. *Klebsiella*
477 *pneumoniae* DSM681, *Salmonella enterica* serovar Typhimurium SL5676 and *Pseudomonas*
478 *fluorescens* (ATCC 13525) were used as a noncarriers. The bacteria were cultivated in Luria
479 Bertani Lennox-broth (LB) (Lennox *et al.* 1955), *E. coli, S.* Typhimurium and *K. pneumoniae*
480 at +37˚C with 120 or 220 rpm shaking, and *P. fluorescens* at room temperature (RT) with 150

481 rpm. LB-agar (1%) plates were used for quantifying bacterial densities from liquid cultures. *E.*
482 *coli, S.* Typhimurium and *K. pneumoniae* plates were incubated overnight at +37˚C and *P.*
483 *fluorescens* 2 days at RT. Media for *E. coli* was supplemented with ampicillin (150 µg/ml) in
484 order to select for pEC13 plasmid.

485 ***ddPCR reactions***

486 The QX200™ Droplet Digital™ PCR system (ddPCR, Bio-Rad Laboratories, Inc.) was used
487 to generate droplets and read the results following the manufacturer's instructions. The
488 reactions were run in C1000 Touch™ Thermal Cycler (Bio-Rad Laboratories, Inc) with 96
489 Deep-Well Reaction Module. The ddPCR reactions were kept overnight at 12 ˚C before reading
490 the droplets. The results were obtained from channels 1 (FAM) and 2 (HEX/VIC) and analysed
491 with QuantaSoft™ Analysis Pro 1.0 (Bio-Rad Laboratories, Inc). All ddPCR reactions were
492 prepared with QX200™ ddPCR™ EvaGreen® Supermix (2x).

493 ***16S rRNA gene amplification***

494 The amplification of the 16S rRNA gene was done using universal primers 27F and 338R (the
495 primers used in this study are listed in Table 1). After initial optimization for ddPCR (see
496 Supplementary Figure 2 for more information), the reactions were prepared using 1x ddPCR
497 supermix and 100 nM of each primer. The amplification conditions consisted of initial
498 incubation at 95 ˚C for 10 minutes, followed by 40 cycles of denaturation at 95 ˚C for 60
499 seconds, annealing at 55˚C for 60 seconds and elongation at 72 ˚C for 2 minutes, and a final
500 step for fluorescence stabilization with incubation at 4 ˚C for 5 minutes followed by incubation
501 at 90 ˚C for 5 minutes. The ramp rate between steps was 2 ˚C/sec until the last elongation
502 incubation, and 1 ˚C /sec for the fluorescence stabilization.

503 ***Propidium monoazide (PMA) for sample treatment***

504 Initially, the maximum non-inhibitory concentration (MNIC) of propidium monoazide (PMA)
505 in ddPCR reaction was tested. PMA (PMAxx™, Biotium) was diluted in PCR grade water to

506 10 µM of final concentration and photoactivated. The treatment consisted of initial incubation
507 in the dark for 10 minutes, followed by incubation for 10 minutes under blue LED light (460
508 nm). The solution was prepared in 500 µl thin-walled polypropylene Qubit™ assay tubes
509 (Invitrogen™), with final volume of 200 µl. Samples were kept on ice layer and mixed
510 occasionally during the treatment by inverting the tube. Lower concentrations were prepared
511 from the treated PMA solution and used in ddPCR, with final concentration ranging from 0.125
512 to 2.5 µM. *E. coli* DNA digested with FastDigest HindIII and NotI (Thermo Fisher Scientific,
513 MA, USA) was used as a template in ddPCR reactions optimized for 16S rRNA. The ddPCR
514 was prepared using 100 nM of each universal primer (27F and 338R) for 16S rRNA
515 amplification. The amplification program was carried out using 5 minutes of initial incubation
516 at 95 ˚C, and 40 cycles of denaturation at 95 ˚C for 30 seconds, annealing at 55 ˚C for 60
517 seconds and elongation at 72 ˚C for 2 minutes. The final fluorescence stabilization and ramp
518 rate were the same as mentioned above.
519 Next, cell samples prepared from *E. coli* and *P. fluorescens* cultures were treated with
520 PMA and used as template in ddPCR. Bacterial suspensions were freshly prepared for every
521 new assay by diluting an overnight culture up to 1000-fold in 1x phosphate buffer solution
522 (PBS). The filtrate was obtained by individually centrifuging 250 µl of each dilution at 5000 ×
523 *g* for 3 minutes through a 0.2 µm spin filter (Centrifugal Filter, Modified Nylon 0.2 µm, VWR).
524 Both sample types, cell suspension and filtrate, were treated with 10 µM of PMA as described
525 above. After PMA treatment, the cells were diluted in PCR grade water and used as template
526 in ddPCR reaction using 16S rRNA primers. The same reaction conditions used for the MNIC
527 assay were used here. The cell density of the samples was determined by serial dilution and
528 plate cultivation of cell suspensions, before and/or after PMA treatment.
529

530 ***One-step fusion PCR for concatemer production***

531 Specific primers were designed to detect the pEC13 plasmid and prompt the bla_16S
532 concatemer formation. The plasmid sequence was retrieved from the GeneBank database under
533 accession number KU932024.1. The class A beta lactamase gene (*bla*CTX-M-14) present in
534 the plasmid was the target sequence chosen and here referred to as *bla*. Primers were designed
535 using the online Primer Designer™ Tool (Thermo Fischer Scientific). To amplify a 101-bp *bla*
536 fragment, 293F was used as a forward primer together with a modified reverse primer rc27F-
537 bla_320R (Table 1). It contains *bla* reverse primer part (320R) and 5'-overhang encoding the
538 reverse complement sequence of 16S rRNA forward primer 27F. Fusion primer enables the
539 production of the 450-bp concatemer via the attachment of the amplified *bla* fragment on its
540 27F-overhang to 16S rRNA fragment (350 bp PCR product amplified by primers 27F and
541 338R) in the same reaction (Figure 2). Initially, both primer pairs (for amplifying *bla* and 16S
542 fragments) were screened separately for forward to reverse ratio and primer pair concentration
543 in ddPCR (Supplementary Figure 3). Reactions were prepared with up to a 100-fold difference
544 between forward and reverse primer molarity, and total concentration ranging from 100 to 400
545 nM. After the initial screening, both primer pairs were used in the same reaction to produce the
546 concatemer. The total concentration of 16S rRNA primers was fixed to 225 nM and the ratio
547 of forward to reverse in 1:5. *bla* primers were used in concentrations ranging from 100 to 300
548 nM and ratios from 10:1 to 30:1 (Supplementary Figure 3C-E). Finally, the primer ratios and
549 concentrations were defined for single-cell experiments. The further reactions for concatemer
550 production were prepared with 225 nM and a ratio of 1:5 (forward to reverse) of 16S rRNA
551 primers, and 200 nM and 10:1 (forward:reverse) of *bla* primers.

552

553 **Single bacterial cells in droplets**

554 To test whether single bacterial cells are suitable to be used as templates in ddPCR, overnight
555 cultures of *E. coli, P. fluorescens, K. pneumoniae,* and *S.* Typhimurium were diluted to reach

556 a cell density of approximately 1×10^7 cfu/ml (cfu; colony forming units). Bacterial dilution
557 was PMA-treated as described above with the final PMA concentration of 10 µM. A dilution
558 series of 1:2 was prepared (in PCR grade water) from the PMA-treated cell samples and used
559 as a template in ddPCR. The bacterial cell suspension always constituted ¼ of the total ddPCR
560 reaction volume. The maximum detectable cell number was determined by using 16S rRNA
561 primers 27F and 338R. The corresponding bacterial cell density of the samples was also
562 determined by dilution series and plating on LB plates. The protocol optimized for the
563 amplification of the 16S rRNA gene in ddPCR and specified in the above section was used.
564
565 **Detection of beta-lactamase host bacteria from mixed samples with single-cell ddPCR**
566 Overnight cultures of *E. coli, K. pneumoniae,* and *S.* Typhimurium were diluted for PMA-
567 treatment and a 1:2 dilution series was prepared as described above. Diluted bacterial samples
568 containing either one bacterial strain or a mix of two strains: beta-lactamase-carrying *E. coli*
569 and a non-carrier strain *K. pneumoniae* or *S.* Typhimurium was used as template in ddPCR.
570 Both 16S rRNA (amplified with 27F and 338R) and concatemer bla_16S (amplified with four
571 primers; 293F, rc27F-320R, 27F and 338R) were amplified from the same samples in separate
572 reactions. Two replicates were prepared for reading the droplets from the 16S rRNA reaction,
573 and three replicates were prepared for the one-step fusion reaction to be saved and used in the
574 following procedures. The ratio of positive:negative droplets was determined with Droplet
575 Reader and the samples containing a sufficient number of positive droplets according to 16S
576 rRNA amplification (and therefore verified to contain the cell density required for single cell
577 analysis) were proceeded to DNA recovery. The amplification conditions were the same as in
578 the protocol optimized for 16S rRNA gene amplification in ddPCR (described in the specific
579 section above).
580

581 **DNA recovery from droplets and library preparation for sequencing**

582 The amplified bla_16S concatemers were recovered from droplets and enriched before
583 preparing the sequencing libraries. The recovery was done by freeze-thawing the ddPCR
584 emulsion using liquid nitrogen (LN2), a method we described recently (Dutra *et al.* 2020).
585 After freeze-thawing, phase separation was achieved, and the amplicon-containing phase was
586 collected for further processing. The recovered concatemers were purified with the Agencourt
587 AMPure XP purification system (Beckman Coulter, Brea, CA) using a 0.9x ratio.

588 Sample enrichment was performed before library preparation by reamplifying the
589 recovered bla_16S amplicons with primers M13_bla.2_F and 338R (Table 1). The samples
590 were then barcoded with bc_M13 and bc_338R primers, both extended with unique 5'-barcode
591 sequences (bc). Both enrichment and barcoding qPCR reactions were prepared with 1x Maxima
592 SYBR Green/Fluorescein qPCR Master Mix (ThermoScientific^{TM}) in a final volume of 25 µl.
593 To prevent the undesired pairing of bla-27F overhang with other 16S rRNA fragments that
594 were present in the sample, the reactions included blocking primers block_27F and block_27R,
595 both in a final concentration of 1 µM. The product of barcoding qPCR was cleaned by gel
596 extraction using NucleoSpin Gel and PCR clean-up kit (MACHEREY-NAGEL, GmbH & Co,
597 KG, Germany), after 1% AGE electrophoresis run for 1 hour at 100 V. The fragment was
598 further analysed with High Sensitivity D1000 ScreenTape assay in 2200 TapeStation System
599 (Agilent Technologies, Inc, CA, USA). PCR conditions for enrichment and barcoding reactions
600 were as follows: the amplification run started with pre-incubation at 95 ºC for 5 minutes,
601 followed by 15 cycles of denaturation at 95 ºC for 30 seconds, annealing at 53 ºC for 30
602 seconds, and elongation at 72 ºC for 30 seconds (with plate read after each cycle). After the
603 run, the product was analysed in a 1.5 % AGE run for 1 hour at 110V. The barcoded and
604 purified bla_16S amplicons were proceeded to library preparation with NEBNext® Ultra™ II
605 DNA Library Prep Kit and sequenced with Illumina NovaSeq 6000 platform using SP flowcell

606 (PE250). Further sequence analysis of 250-bp paired-end read data was run with CLC
607 Genomics Workbench 12 (Qiagen Bioinformatics). The obtained reads were trimmed with *bla*
608 and 16S rRNA primer sequences bla_293F, bla_320R, 27F and 338R to confirm that they
609 contain both parts of the concatemer. The genetic region between 27F and 338R was extracted
610 and used for *de novo* OTU clustering with a 97 % similarity threshold. Rare OTUs (abundance
611 <100 reads) were excluded from the analysis. The reference sequence of each OTU was
612 exposed to BLAST search, using the 16S Ribosomal RNA sequences database. Each OTU was
613 identified according to the BLAST results. Finally, the abundance of identified bacterial hosts
614 carrying *bla*CTX-M-14 present in each sample was determined.

615 **Horizontal transfer of betalactamase under lethal antibiotic exposure**

616 Rifampicin-resistant strains of *K. pneumoniae, S.* Typhimurium and *Comamonas* sp. (bacterial
617 isolate originating from the human fecal sample as part of gut bacterial strain library collected
618 under research permit T300/2020 for Hospital District of Southwest Finland) were used to
619 construct a synthetic bacterial community. 5 µl of each bacterial culture, cultivated overnight
620 in LB medium without antibiotics, were added to 5 ml of fresh LB medium containing 1)
621 cephalothin 2) rifampicin and cephalothin or 3) rifampicin, with four replicates per treatment.
622 The bacterial communities were incubated at +37 ºC for 30 min with 120 rpm agitation, after
623 which *E. coli* harboring conjugative plasmid pEC13 with *bla*CTX-M-14 (conferring resistance
624 to cephalothin) was introduced by adding 5 µl of overnight grown culture into each bacterial
625 community. After 1 h incubation, a second antibiotic pulse was given to treatment 1 with
626 rifampicin and treatment 3 with cephalothin, resulting in the same final antibiotic concentration
627 in all the treatments (50 µg/ml of both cephalothin and rifampicin). The bacterial communities
628 were incubated for 3.5 h after which one cell sample from each community was taken for single
629 cell analysis for tracing the *bla*CTX-M-14 carrier cells.

630 Cell samples from each community were PMA-treated and diluted as described earlier.
631 bla_16S concatemers were produced by single-cell ddPCR. Each reaction was run in two
632 replicates. One reaction was read with Droplet Reader to ensure cell density and single-cell
633 resolution. The other replicate was used for DNA recovery with LN2. Recovered fusion
634 amplicons were first purified with QIAquick PCR purification kit (Qiagen) and proceeded to
635 library preparation first with re-amplification PCR (PCR conditions: 95 ºC 5 min, followed by
636 30 cycles of denaturation at 95 ºC for 30 seconds, annealing at 53 ºC for 30 seconds, and
637 elongation at 72 ºC for 30 seconds, with plate read after each cycle) followed by PCR
638 purification (sparQ PureMag Beads, Quantabio) and inspecting the correct amplicon size with
639 agarose gel electrophoresis (1.5%, 30 min, 120 V). Purified fusion amplicons were barcoded
640 with both 5' and 3' extensions in barcoding PCR (PCR conditions as described above but with
641 9 cycles). Barcoded amplicons were extracted from agarose gel (1.5%, 1 h, 100 V) by using
642 NucleoSpin® Gel and PCR Cleanup XS kit (Macherey-Nagel) and the amplicon size and DNA
643 concentration were checked with High Sensitivity D1000 ScreenTape assay in 2200
644 TapeStation System (Agilent Technologies, Inc, CA, USA). Sequencing libraries were
645 prepared with NEBNext® Ultra™ II DNA Library Prep Kit and the samples were sequenced
646 with Illumina NovaSeq 6000 platform using SP flowcell (PE250).
647 Further sequence analysis of 250-bp paired-end read data was run with CLC Genomics
648 Workbench 21 (Qiagen Bioinformatics) and Microbial Genomics module. Paired reads were
649 first merged and trimmed by quality (0.05 as threshold). Reads were then trimmed with forward
650 and reverse primer sequences of *bla*CTX-M-14 to ensure they contain the complete sequence
651 of the target gene, and then with 16S forward primer. Single reads and merged read sequences
652 not covering the complete fusion amplicon length were discarded from the analysis. Finally,
653 the remaining sequence between 16S rRNA forward (27F) and reverse (338R) primer binding
654 sites was used to determine the carriers of the betalactamase gene. Only reads containing all

655 the primer sequences and with the length of 300-315 bp (after trimming) were taken to the final
656 analysis. A local 16S database was constructed to represent the synthetic community
657 (consisting of *K. pneumoniae, S.* Typhimurium, *Comamonas* sp. and *E. coli* 16S rRNA gene
658 sequences) and used for reference-based OTU clustering with 99% similarity, allowing the
659 creation of new OTUs (99% taxonomy similarity). OTUs with a total abundance lower than
660 4000 or recognized as chimeras were left out from the final analysis. The subpopulation
661 carrying *bla*CTX-M-14 was assessed as OTU frequency (OTU abundance/total OTU
662 abundance).
663
664 **Table 1. Primers used in this study.**

Primer name	Sequence 5'-3'	Target
27F (Lane, 1991)	AGAGTTTGATCMTGGCTCAG	16S rRNA forward primer
338R (Suzuki & Giovannoni 1996)	TGCTGCCTCCCGTAGGAGT	16S rRNA reverse primer
bla_293F	GAAACGTCTCATCGCCGATC	betalactamase forward primer
rc27F-bla_320R	CTGAGCCAKGATCAAACTCTGCAGTACAGCGACAATACCG	fusion primer; betalactamase reverse primer with 16S rc27F primer
M13_bla.2_F	TGTAAAACGACGGCCAGTCGAGCTGGGCAATCAATTTG	M13 adapter attachment and bla shortening
M13_27F	TGTAAAACGACGGCCAGTAGAGTTTGATCMTGGCTCAG	M13 adapter
27F_block	TTTTTTTTTTAGAGTTTGATCMTGGCTCAG/3SpC3/	blocking primer
rc27F_block	TTTTTTTTTTCTGAGCCAKGATCAAACTCT/3SpC3/	blocking primer
bc_M13F	[bc]TGTAAAACGACGGCCAGT	barcoding primer
bc_338R	[bc]TGCTGCCTCCCGTAGGAGT	barcoding primer

rc = reverse complementary; bc = barcode sequence

665
666
667 **Acknowledgements**

668 This study was funded by the Academy of Finland grants 347531 to M.J., 323063 to M.T.
669 and 322204 to R.P., ERC Consolidator Grant (615146) to M.T., Finnish Cultural Foundation
670 grant (00210277) to L.A.L.D. and Jane and Aatos Erkko Foundation (to M.J.).
671
672

673 **References**

674

675 Baptista, J. D., Lunn, M., Davenport, R. J., Swan, D. L., Read, L. F., Brown, M. R., Morais,
676 C., & Curtis, T. P. (2014). Agreement between amoA gene-specific quantitative PCR and
677 fluorescence in situ hybridization in the measurement of ammonia-oxidizing bacteria in
678 activated sludge. *Applied and environmental microbiology*, *80*(19), 5901–5910.
679 https://doi.org/10.1128/AEM.01383-14

680 Browne, H. P., Forster, S. C., Anonye, B. O., Kumar, N., Neville, B. A., Stares, M. D.,
681 Goulding, D., & Lawley, T. D. (2016). Culturing of 'unculturable' human microbiota reveals
682 novel taxa and extensive sporulation. *Nature*, *533*(7604), 543–546.
683 https://doi.org/10.1038/nature17645

684 Capobianco, J. A., Clark, M., Cariou, A., Leveau, A., Pierre, S., Fratamico, P., Strobaugh, T.
685 P., Jr, & Armstrong, C. M. (2020). Detection of Shiga toxin-producing Escherichia coli (STEC)
686 in beef products using droplet digital PCR. *International journal of food microbiology*, *319*,
687 108499. https://doi.org/10.1016/j.ijfoodmicro.2019.108499

688 Carini, P., Marsden, P. J., Leff, J. W., Morgan, E. E., Strickland, M. S., & Fierer, N. (2016).
689 Relic DNA is abundant in soil and obscures estimates of soil microbial diversity. *Nature*
690 *microbiology*, *2*, 16242. https://doi.org/10.1038/nmicrobiol.2016.242

691 Diebold, P. J., New, F. N., Hovan, M., Satlin, M. J. & Brito, I.L. (2021). Linking plasmid-
692 based beta-lactamases to their bacterial hosts using single-cell fusion PCR. *eLife* 10:e66834.
693 https://doi.org/10.7554/eLife.66834

694 Dutra, L., Franz, O., Puupponen, V. M., & Tiirola, M. (2020). DNA recovery from Droplet
695 Digital™ PCR emulsions using liquid nitrogen. *BioTechniques*, *69*(6), 450–454.
696 https://doi.org/10.2144/btn-2020-0076

697 Hindson, B. J., Ness, K. D., Masquelier, D. A., Belgrader, P., Heredia, N. J., Makarewicz, A.
698 J., Bright, I. J., Lucero, M. Y., Hiddessen, A. L., Legler, T. C., Kitano, T. K., Hodel, M. R.,
699 Petersen, J. F., Wyatt, P. W., Steenblock, E. R., Shah, P. H., Bousse, L. J., Troup, C. B., Mellen,
700 J. C., Wittmann, D. K., … Colston, B. W. (2011). High-throughput droplet digital PCR system
701 for absolute quantitation of DNA copy number. *Analytical chemistry*, *83*(22), 8604–8610.
702 https://doi.org/10.1021/ac202028g

703 Ibáñez de Aldecoa, A. L., Zafra, O., & González-Pastor, J. E. (2017). Mechanisms and
704 Regulation of Extracellular DNA Release and Its Biological Roles in Microbial
705 Communities. *Frontiers in microbiology*, *8*, 1390. https://doi.org/10.3389/fmicb.2017.01390

Lane D. J. (1991). 16S/23S rRNA sequencing, in Nucleic Acid Techniques In Bacterial Systematics. Eds Stackebrandt E., Goodfellow M. (Chichester: John Wiley & Sons;), 115–175.

Lavoie, K. H., Winter, A. S., Read, K. J., Hughes, E. M., Spilde, M. N., & Northup, D. E. (2017). Comparison of bacterial communities from lava cave microbial mats to overlying surface soils from Lava Beds National Monument, USA. *PloS one*, *12*(2), e0169339. https://doi.org/10.1371/journal.pone.0169339

Lennon, J. T., Muscarella, M. E., Placella, S. A., & Lehmkuhl, B. K. (2018). How, When, and Where Relic DNA Affects Microbial Diversity. *mBio*, *9*(3), e00637-18. https://doi.org/10.1128/mBio.00637-18

Lennox E. S. (1955). Transduction of linked genetic characters of the host by bacteriophage P1. *Virology*, *1*(2), 190–206. https://doi.org/10.1016/0042-6822(55)90016-7

Liu, Y., Chen, J., & Thygesen, A. (2018). Efficient One-Step Fusion PCR Based on Dual-Asymmetric Primers and Two-Step Annealing. *Molecular biotechnology*, *60*(2), 92–99. https://doi.org/10.1007/s12033-017-0050-7

Mattila, S., Ruotsalainen, P., Ojala, V., Tuononen, T., Hiltunen, T., & Jalasvuori, M. (2017). Conjugative ESBL plasmids differ in their potential to rescue susceptible bacteria via horizontal gene transfer in lethal antibiotic concentrations. *The Journal of antibiotics*, *70*(6), 805–808. https://doi.org/10.1038/ja.2017.41

McMahon, T. C., Blais, B. W., Wong, A., & Carrillo, C. D. (2017). Multiplexed Single Intact Cell Droplet Digital PCR (MuSIC ddPCR) Method for Specific Detection of Enterohemorrhagic *E. coli* (EHEC) in Food Enrichment Cultures. *Frontiers in microbiology*, *8*, 332. https://doi.org/10.3389/fmicb.2017.00332

Nagler, M., Insam, H., Pietramellara, G., & Ascher-Jenull, J. (2018). Extracellular DNA in natural environments: features, relevance and applications. *Applied microbiology and biotechnology*, *102*(15), 6343–6356. https://doi.org/10.1007/s00253-018-9120-4

Nocker, A., Sossa-Fernandez, P., Burr, M. D., & Camper, A. K. (2007). Use of propidium monoazide for live/dead distinction in microbial ecology. *Applied and environmental microbiology*, *73*(16), 5111–5117. https://doi.org/10.1128/AEM.02987-06

Ottesen, E. A., Hong, J. W., Quake, S. R., & Leadbetter, J. R. (2006). Microfluidic digital PCR enables multigene analysis of individual environmental bacteria. *Science (New York, N.Y.)*, *314*(5804), 1464–1467. https://doi.org/10.1126/science.1131370

Pinheiro, L. B., Coleman, V. A., Hindson, C. M., Herrmann, J., Hindson, B. J., Bhat, S., & Emslie, K. R. (2012). Evaluation of a droplet digital polymerase chain reaction format for DNA copy number quantification. *Analytical chemistry*, *84*(2), 1003–1011. https://doi.org/10.1021/ac202578x

Probandt, D., Eickhorst, T., Ellrott, A., Amann, R., & Knittel, K. (2018). Microbial life on a sand grain: from bulk sediment to single grains. *The ISME journal*, *12*(2), 623–633. https://doi.org/10.1038/ismej.2017.197

Ruotsalainen, P., Given, C., Penttinen, R. & Jalasvuori, M. (2020). Beta-Lactam Sensitive Bacteria Can Acquire ESBL-Resistance via Conjugation after Long-Term Exposure to Lethal Antibiotic Concentration. *Antibiotics.* 9:296. doi:10.3390/antibiotics9060296

Spencer, S. J., Tamminen, M. V., Preheim, S. P., Guo, M. T., Briggs, A. W., Brito, I. L., A Weitz, D., Pitkänen, L. K., Vigneault, F., Juhani Virta, M. P., & Alm, E. J. (2016). Massively parallel sequencing of single cells by epicPCR links functional genes with phylogenetic markers. *The ISME journal*, *10*(2), 427–436. https://doi.org/10.1038/ismej.2015.124

Stubbendieck, R. M., Vargas-Bautista, C., & Straight, P. D. (2016). Bacterial Communities: Interactions to Scale. *Frontiers in microbiology*, *7*, 1234. https://doi.org/10.3389/fmicb.2016.01234

Suzuki M. T., Giovannoni S. J. Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. *Appl Environ Microb.* 1996;**62**:625–630.

Tadmor, A. D., Ottesen, E. A., Leadbetter, J. R., & Phillips, R. (2011). Probing individual environmental bacteria for viruses by using microfluidic digital PCR. *Science (New York, N.Y.)*, *333*(6038), 58–62. https://doi.org/10.1126/science.1200758

Zeng, Y., Novak, R., Shuga, J., Smith, M. T., & Mathies, R. A. (2010). High-performance single cell genetic analysis using microfluidic emulsion generator arrays. *Analytical chemistry*, *82*(8), 3183–3190. https://doi.org/10.1021/ac902683t