Uncovering bacterial hosts of class 1 integrons in an urban coastal aquatic environment with a single-cell fusion-polymerase chain reaction technology

Qin Qi*‡, Timothy M Ghaly‡, Anahit Penesyan, Vaheesan Rajabal, Jeremy AC Stacey, Sasha G Tetu, Michael R Gillings

* Correspondence: qin.qi@mq.edu.au

‡ These authors contributed equally

AUTHOR ADDRESS

School of Natural Sciences, Macquarie University, 14 Eastern Road, Sydney, NSW 2109, Australia

Abstract

Horizontal gene transfer (HGT) is a key driver of bacterial evolution via transmission of genetic materials across taxa. Class 1 integrons are genetic elements that correlate strongly with anthropogenic pollution and contribute to the spread of antimicrobial resistance (AMR) genes via HGT. Despite their significance to human health, there is a shortage of robust, culture-free surveillance technologies for identifying uncultivated environmental taxa that harbour class 1 integrons. We developed a modified version of epicPCR (emulsion, paired isolation and concatenation polymerase chain reaction) that links class 1 integrons amplified from single

bacterial cells to taxonomic markers from the same cells in emulsified aqueous droplets. Using this single-cell genomic approach and Nanopore sequencing, we successfully assigned class 1 integron gene cassette arrays containing mostly AMR genes to their hosts in coastal water samples that were affected by pollution. Our work presents the first application of epicPCR for targeting variable, multi-gene loci of interest. We also identified the *Rhizobacter* genus as novel hosts of class 1 integrons. These findings establish epicPCR as a powerful tool for linking taxa to class 1 integrons in environmental bacterial communities and offer the potential to direct mitigation efforts towards hotspots of class 1 integron-mediated dissemination of AMR.

Synopsis

We present a novel single-cell genomic surveillance technology for identifying environmental bacterial hosts of a class of mobile genetic elements that are linked to anthropogenic pollution and contribute to the dissemination of antimicrobial resistance.

Keywords

Integrons, antimicrobial resistance (AMR), anthropogenic pollution, class 1 integron gene cassettes, horizontal gene transfer (HGT), mobile genetic elements, environmental bacterial hosts, single-cell genomics

Introduction

Horizontal gene transfer (HGT) is an important source of genomic variability upon which selection can act [1-3]. Microbial adaptation via HGT most commonly occurs among closely

44 related taxa or between more distantly related taxa found in similar environments [4, 5]. A key
45 example is the dissemination of antimicrobial resistance (AMR) genes via conjugation,
46 bacteriophage transduction, and natural transformation by extracellular DNA [6-9]. To understand
47 the dynamics of HGT, we need to identify the sources of horizontally transferred genes and the
48 range of taxa they subsequently inhabit [10-12]. This is challenging given the sheer diversity of
49 microbial taxa and the difficulty of culturing many environmental microbes [13-16]. Metagenomic
50 sequencing of bacterial communities does not require prior culturing of microbes [17, 18].
51 However, mobile genetic elements are often associated with repetitive DNA sequences and
52 exhibit complex interactions with genome rearrangements, which limit our ability to attribute
53 laterally transferred genes to host genomes due to assembly and/or taxonomic binning
54 difficulties [19, 20].

55 epicPCR (emulsion, paired isolation and concatenation PCR) is a fusion PCR-based technique
56 that overcomes these problems by co-amplifying taxonomic markers with functional genes of
57 interest from single uncultured bacterial cells. Single cells are trapped in droplets generated
58 by emulsifying an aqueous phase containing PCR reagents and environmental bacteria in an
59 emulsion oil [21-26]. This compartmentalises the fusion PCR into parallel, single cell-level
60 reactions. Within each droplet, chimeric amplicons that combine genes of interest with
61 taxonomic markers corresponding to diverse phylogenetic groups are generated for high-
62 throughput sequencing. To date, epicPCR has been successfully applied to target functional
63 genes including dissimilatory sulfite reductase *dsrB* genes [21, 24], antibiotic resistance genes [22,
64 23, 26, 27], class 1 integron associated genes (*intI1*, *sul1*, Δ*qacE*) [23, 27, 28], and bacteriophage-borne
65 ribonucleotide reductase *RNR* genes [25]. epicPCR amplicons typically comprise short fragments
66 of genes of interest linked to the V4 hypervariable region of 16S rRNA markers. Illumina
67 sequencing of recovered amplicons demonstrates the presence of the genes of interest in a

68 variety of bacterial taxa. However, no studies have so far applied epicPCR to target variable,
69 multi-gene loci of interest.

70 Class 1 integrons contribute to the dissemination of AMR genes via HGT in natural ecosystems
71 and clinical settings [29, 30]. Integrons are genetic elements that comprise an integron-integrase
72 gene (*intI*), an *attI* recombination site, and a promoter (P_C) that drives the expression of variable
73 gene cassettes [31-33]. IntI catalyses the integration and excision of gene cassettes via *attC* sites
74 that can be recombined with *attI* or other *attC* sites [33-36]. All class 1 integrons contain highly
75 conserved integron-integrase (*intI1*) and *attI1* sequences (Figure 1A). Clinical class 1 integrons
76 are strongly associated with AMR gene cassettes that confer resistance to antibiotics,
77 disinfectants, or heavy metals [29, 31, 37]. Mobile genetic elements such as broad host-range
78 plasmids and transposons have acquired class 1 integrons, resulting in the proliferation of class
79 1 integrons across diverse bacterial hosts [31, 38]. Recently, it has been reported that phage-
80 plasmids can also disseminate class 1 integrons [39].

81 Bacterial hosts of class 1 integrons have previously been isolated and characterised through
82 colony PCR screening and fosmid library preparation [29, 40-43]. The necessity to pre-isolate
83 environmental bacteria severely biases host identification towards the cultivated fraction and
84 hinders our ability to attribute relative abundances of class 1 integrons to individual taxa. To
85 overcome these limitations, we developed a modified experimental and bioinformatic pipeline
86 for epicPCR that associates class 1 integrons with their bacterial hosts. Class 1 integron gene
87 cassette arrays are co-amplified with the V4 hypervariable region of the 16S rRNA marker
88 genes from individual bacterial cells (Figure 1B-C). The fused amplicons are sequenced using
89 Nanopore long-read technology. We applied stringent bioinformatic filtering to ensure that the
90 associations between the class 1 integrons and the taxonomic markers were *bona fide.* We
91 successfully tested our epicPCR pipeline on water samples from a coastal environment and
92 linked six unique class 1 integrons with variable gene cassettes to seven distinct environmental

93 bacterial taxa. While most of the gene cassettes encode AMR genes, we also linked a gene
94 cassette encoding a cupin-domain containing protein of unknown functions to two
95 *Alphaproteobacteria* hosts. Furthermore, our results demonstrated that *Rhizobacter* is a novel
96 host of class 1 integrons. Overall, our study shows that epicPCR is a powerful tool for
97 determining the environmental hosts of class 1 integrons with the potential to access previously
98 untapped carriers of these genetic elements. This paves the way for future studies that seek to
99 pinpoint the "sources" and "sinks" of class 1 integrons.

100

101 ## Materials and Methods

102 **Isolation of bacteria from an urban coastal location**

103 Coastal water samples were collected in three 50 mL biological replicates after rain in a rocky
104 intertidal zone downstream from a stormwater outlet at Shark Point (location coordinates: -
105 33.91, 151.27) near Sydney, NSW, Australia (Figure S1). The samples were filtered using
106 EASYstrainer cell strainers with a mesh size of 40 µm (Greiner AG, Germany) and centrifuged
107 at 3220 *g* for 15 minutes. Cell pellets were resuspended in phosphate buffered saline (PBS)
108 solution at pH 7.4. Resuspended cells were stored as 25% glycerol stocks.

109

110 **Bacterial cell count estimates by fluorescence microscopy**

111 Approximately 10^5-10^6 cells were fixed and stained in PBS containing 4% formaldehyde and
112 2 µg/mL 4′,6-diamidino-2-phenylindole (DAPI) for 30 min in the dark with rotation at room
113 temperature. Re-suspended cells were diluted in PBS, transferred to a haemocytometer and
114 imaged under the x40 lens of an Olympus BX63 fluorescence microscope. Differential
115 interference contrast (DIC) and DAPI fluorescence channel images were overlaid and analysed

116 with ImageJ (National Institutes of Health, United States). Bacterial cell densities were
117 calculated from the average number of cells in five 0.2 mm x 0.2 mm square grids of the
118 counting chamber for each sample (Figure S2).

119

120 **Emulsion, paired-isolation and concatenation PCR (epicPCR)**

121 Our hybrid epicPCR procedures were adapted from the protocols of Sakowski *et al* [25] and
122 Diebold *et al* [26]. Briefly, epicPCR was performed in three technical replicates for three
123 biological replicates. Approximately 150,000 cells from the bacterial glycerol stocks were
124 pelleted and vortexed in a 75 µL PCR reagent mix containing GC buffer (1x), dNTP mix (0.4
125 mM), Phusion Hot-Start II polymerase (0.05 U/µL) (Thermo Scientific, United States), bovine
126 serum albumin solution (1 µg/µL) (Promega, United States), Lucigen Ready-Lyse lysozyme
127 (500 U/µL) (LGC Biosearch Technologies, United States) and oligonucleotide primers R926
128 (2 µM), *intI1*_outer (1 µM), and R519-*qacE* bridging primer (0.04 µM) (Table S2). To
129 demonstrate the absence of false associations between 16S rRNA markers and class 1 integrons
130 originating from different cells, we spiked one set of technical replicates with cells of a class 1
131 integron-free *Escherichia coli* MG1655 strain at a total population frequency of approximately
132 10%.

133 ABIL emulsion oil was prepared by supplementing mineral oil with 4% ABIL EM 90 (Redox,
134 Australia) and 0.05% Triton X-100 (Promega, United States). 425 µL of ABIL oil was added
135 to the PCR mix, which was immediately emulsified at 4 ms^{-1} for 45 s using the FastPrep-24
136 bead beating system (MP Biomedicals, United States). The water-in-oil emulsion was
137 aliquoted into 8 portions, incubated at 37˚C for 10 min, and was subjected to the following
138 PCR amplification conditions: 98˚C for 5 minutes; 38 cycles of 98°C for 10 s, 59°C for 20 s,
139 and 72°C for 2 min; and finally 72°C for 5 min. The *intI1*_outer primer and the R519-*qacE*

140 bridging primer bind to *intI1* in the 5'-conserved segment (CS) and *qacE* in the 3'-CS of class
141 1 integrons respectively (Figure 1B). The 18 bp overhang introduced by the bridging primer to
142 the *qacE* end allows the intermediate PCR products to act as a long primer and amplify the V4-
143 V5 hypervariable region with the R926 reverse primer.

144 The emulsified PCR mix for each technical replicate was pooled, vortexed in 900 µL of
145 isobutanol and 200 µL of 5M NaCl and centrifuged for 1 min at ~20,000*g* with a soft brake.
146 The aqueous phase at the bottom of the tubes was extracted, purified using the Monarch PCR
147 & DNA Cleanup Kit (New England Biolabs, United States), and eluted with nuclease-free
148 water. Sera-Mag Select magnetic beads (Cytiva, United States) were added to the eluted PCR
149 products to deplete <300 bp DNA fragments. The size-selected DNA was used as template in
150 100 µL Phusion PCR mix containing GC buffer (1x), dNTP mix (0.4 mM), Phusion Hot-Start
151 II polymerase (0.04 U/µL), AP27_short primer (0.8 µM), *intI1*_nested primer (0.4 µM),
152 forward and reverse blocking primers (0.32 µM each). The PCR mix was subjected to the
153 following thermocycling conditions in 8 aliquots: 98˚C for 30 s; 35 cycles of 98°C for 10 s,
154 59°C for 20 s and 72°C for 1 min 50 s; followed by a final step of 72°C for 5 min. epicPCR
155 products were visualised on 1% agarose gel by electrophoresis and treated with Sera-Mag
156 Select magnetic beads to deplete epicPCR products containing cassette-less integrons.

157 To confirm that fusion between class 1 integron and 16S DNA fragments has occurred (Figure
158 S3A), 2 ng of each purified epicPCR product was amplified using *qacE*_F (0.4 µM) and
159 AP28_short (0.4 µM) primers in GoTaq polymerase master mix (Promega, United States)
160 under the following thermocycling conditions: 95˚C for 30 s; 30 cycles of 95°C for 30 s, 55°C
161 for 30 s and 72°C for 30 s; with a final step of 72°C for 5 min. Successful amplification of the
162 *qacE*-16S rRNA gene chimeric region should produce ~390 bp gel bands.

163

164 **Sanger sequencing of epicPCR products from the *Rhizobacter* genus**

165 To confirm that class 1 integrons can be found in the *Rhizobacter* genus, we pooled
166 approximately 150,000 cells from the three biological replicates and performed the epicPCR
167 procedures in three technical replicates with the minor modifications shown in Figure S3B: the
168 R926 and AP27_short primers were replaced with *Rhizobacter*-specific primers
169 *Rhizobacter*_16S_ R926 and *Rhizobacter*_16S_ R806 respectively, which were designed based
170 on the 16S rRNA sequence of *R. gummiphilus* strain NS21 (NCBI accession no. CP015118)
171 [44]. *qacE*_F was used as the nested forward primer. The resulting 389 bp epicPCR products were
172 verified by Sanger sequencing (Marcogen Inc, South Korea).

173

174 **Oxford Nanopore long-read sequencing**

175 Purified epicPCR amplicons were treated with the NEBNext Ultra II End Repair/dA-Tailing
176 Module (New England Biolabs, United States), purified with JetSeq Clean magnetic beads
177 (Meridian Bioscience, United States), and eluted using nuclease-free water. End-repaired DNA
178 was barcoded using the Native Barcoding Expansion kit (ONT, United Kingdom) and
179 Blunt/TA Ligase Master Mix (New England Biolabs, United States). Barcoded samples were
180 multiplexed and ligated to sequencing adaptor molecules in Adapter Mix II using the NEBNext
181 Quick T4 DNA ligase (New England Biolabs, United States). The ligation products were
182 washed twice with Short Fragmentation Buffer and eluted with Elution Buffer. The sequencing
183 library was loaded into a FLO-MIN106D flow cell R9.4.1 and a MinION Mk1B sequencer
184 (ONT, United Kingdom). The minimum threshold for read length was set to 1 kB on the
185 MinKNOW operating software. Basecalling of FAST5 raw data was performed with the high
186 accuracy option using Guppy v6.1.2 (ONT, United Kingdom). The native barcoded samples
187 were demultiplexed using default parameters.

188

189 **Analysis of Nanopore sequencing data of epicPCR amplicons**

190 All filtering and processing steps were carried out using an in-house pipeline
191 (https://github.com/timghaly/Int1-epicPCR). First, the pipeline quality-filters reads using
192 NanoFilt v2.8.0 [45], removing those with an average read quality of <7 and read length <670 bp,
193 which represents the minimum length of an epicPCR product with a cassette-less integron.
194 Quality-filtered reads are then oriented and trimmed with the final nested forward and reverse
195 primer sequences using Pychopper v2.7.0 (https://github.com/epi2me-labs/pychopper).
196 Pychopper identifies both primers in each read using edlib v1.2.3 [46], orients the reads based on
197 the forward and reverse primer sequences, and discards reads that do not contain both primers
198 in the correct orientations. The pipeline then clusters the primer-oriented reads into amplicon-
199 specific clusters using isONclust v0.0.6.1 [47]. Error correction was performed on each cluster
200 using isONcorrect v0.0.8 [48], which can jointly use all cassette arrangements of the same
201 integron that occur in different clusters and allow efficient error correction even for amplicons
202 with low sequencing depths. A consensus sequence is then generated for each cluster using
203 spoa v4.0.7 [49]. All consensus sequences are then pooled, while removing any redundancies,
204 including reverse complement redundancies, using dedupe from BBTools v35
205 (https://github.com/kbaseapps/BBTools). Next, the consensus sequences are screened for the
206 R519-*qacE* bridging primer and the 157 bp region (covering the 5'-end of *intI1* and all of *attI1*)
207 using blastn from BLAST v2.2.31. Retained sequences must contain both regions. Any
208 sequences containing more than one hit to either of these sequences are considered unintended
209 chimeras and discarded. Finally, sequences are screened for the correctly fused 16S rRNA gene
210 fragment using Metaxa2 v2.2.3 [50]. The final output of the pipeline for each sample is a set of
211 full-length, primer-oriented amplicon consensus sequences that contain a complete *attI1*
212 sequence, R519-*qacE* bridging primer, and the V4 hypervariable region of the 16S rRNA gene.

213 Cd-hit v4.8.1 [51] was used to cluster chimeric epicPCR products comprising class 1 integron
214 gene cassette arrays and V4 hypervariable regions of the 16S rRNA gene with ≥99% pairwise
215 identity in nucleotide sequences in at least 3 epicPCR replicates. For epicPCR products that
216 were found in fewer than 3 replicates, the NCBI Genome Workbench was used to perform
217 blastn searches of these sequences against a local database created using all the Nanopore reads
218 obtained in this study. Reads with ≥98% pairwise identity in nucleotide sequences and ≥98%
219 coverage were aligned using the Geneious bioinformatic software (Biomatters, New Zealand),
220 and the consensus sequences were generated from these alignments within each set of epicPCR
221 replicates. Consensus sequences that could be found in at least 3 epicPCR replicates and
222 showed ≥99% pairwise identity in nucleotide sequences and ≥99% identical sites were added
223 to the list of epicPCR products that had already been shortlisted by the Cd-hit algorithm.
224 IntegronFinder 2.0 [parameters: --local-max --gbk --promoter-attI --calin-threshold 1] was
225 used to predict the ORFs and *attC* sites [52].

226

227 **Taxonomic classification for bacterial hosts of class 1 integrons**

228 The sequences of V4 hypervariable regions of 16S rRNA gene sequences obtained in this study
229 were searched against the SILVA 16S rRNA gene database website [53] using the Alignment,
230 Classification and Tree (ACT) tool with a cut-off threshold of 0.80. The V4 region sequences
231 were aligned using the Geneious multiple alignment tool using default parameters and cost
232 matrix = 93% similarity. A phylogenetic tree was constructed with PhyML (v3.0) using the
233 maximum likelihood method based on the Hasegawa-Kishino-Yano (HKY85) model with
234 bootstrapping (number of replicates = 100) [54]. To verify genus-level taxonomic classifications,
235 the V4 hypervariable regions of 16S rRNA for the *Rhizobacter* and *Aquabacterium*-like hosts
236 found in this study were searched against the NCBI Nucleotide Collection by blastn with the

237 exclusion of "uncultured/environmental sample sequences". The 16S rRNA sequences of all
238 identical matches were searched against the SILVA 16S rRNA gene database. All the reference
239 genomes of *Rhizobacter* and *Aquabacterium* species were downloaded from the NCBI RefSeq
240 database (last accessed on 11 November 2022). Nucleotide alignment was performed against
241 the full 16S rRNA gene sequences from the RefSeq genomes using the Geneious bioinformatic
242 software.

243

244 Results and Discussion

245 Coastal water samples were collected in three biological replicates from the rockpools
246 downstream of a stormwater outlet at Shark Point, Sydney, New South Wales, Australia
247 (Figure S1). We anticipated that class 1 integron-associated environmental bacteria would be
248 detected at this location, because surface runoff was expected to carry pollutants from the
249 surrounding urban areas. A preponderance of cells displayed unicellular and planktonic
250 morphologies (Figure S2). We applied modified epicPCR procedures to amplify class 1
251 integron gene cassette arrays from single bacterial cells encapsulated by PCR reagent droplets
252 homogenously dispersed in an oil emulsion and successfully generated chimeric epicPCR
253 products containing class 1 integrons and 16S rRNA gene sequences (Figure 1). These
254 epicPCR products were Nanopore sequenced. The median number of reads per sample was
255 232,402. After stringent bioinformatic filtering, we identified six class 1 integrons in seven
256 distinct bacterial hosts (Figure 2). All seven combinations of epicPCR products were observed
257 in at least three independent replicates ($n \geq 3$, ≥99% pairwise nucleotide identity and ≥99%
258 identical sites relative to the consensus sequence for each epicPCR product). In addition, there
259 were five combinations of class 1 integrons and hosts that were detected in fewer than 3
260 replicates and were consequently excluded. Taxonomic classifications were obtained using the

261 SILVA ribosomal RNA gene database (Table S1) [53]. *Alphaproteobacteria* (*n*=2) and
262 *Gammaproteobacteria* (*n*=5) were the two major classes of hosts represented in our dataset,
263 while *Burkholderiales* (*n*=4) was the most common bacterial order. These observations are
264 broadly consistent with previous findings that class 1 integrons are associated with
265 *Gammaproteobacteria* [29, 42, 43]. The finding of integrons in two *Alphaproteobacteria* is notable,
266 because *Alphaproteobacteria* are rarely known to be associated with integrons [55]. Currently,
267 there are very few examples of *Alphaproteobacteria* species that harbour class 1 integrons [56-
268 58]. This underscores the potential of epicPCR approaches to discover novel host-function
269 associations.

270 Six different gene cassettes were detected, five of which encoded well-characterised AMR
271 genes (Table 1). The nucleotide sequences of all the gene cassettes reported in this study are
272 summarised in Table S3. We compared their nucleotide sequences against reference class 1
273 integron gene cassettes from the NCBI database and found that both the coding and non-coding
274 regions of these gene cassettes were highly conserved relative to the reference sequences [34, 59].
275 In our dataset, there are six different *attC* recombination sites with 100% nucleotide sequence
276 identity to those in the respective reference gene cassettes in Table 1. We also aligned all the
277 nucleotide sequences in the 5'-CS and 3'-CS regions of all the class 1 integrons from this study
278 and found that both the 157 bp *intI1/attI1* and 190 bp *attC*/*qacE* regions in the chimeric
279 amplicons were identical in all but one exception (Figure S4). Overall, these class 1 integron
280 structures are typical of those recovered from clinical bacterial isolates [29, 60, 61].

281

282 **Class 1 gene cassettes detected by epicPCR**

283 Cupins belong to a superfamily of proteins with a beta-barrel fold and were originally found as
284 highly conserved domains in plant proteins [62]. We found a gene cassette encoding a cupin barrel

285 domain of unknown functions in two *Alphaproteobacteria* hosts. One host was unclassified at
286 the order level (Figure 2A), while the other belonged to the *Rhodobacteraceae* family (Figure
287 2B). This cupin-domain-encoding gene was previously observed in a class 1 integron gene
288 cassette from an unknown host [59]. Based on identical nucleotide sequence matches from the
289 NCBI nucleotide database, this gene cassette was detected in class 1 integrons, non-class 1
290 integrons, and CALINs (clusters of *attC*s lacking an associated integron-integrase) in the
291 chromosomes or plasmids of several *Gammaproteobacteria* hosts (Table S4). This observation
292 suggests selective pressure for maintaining this gene cassette in both class 1 and non-class 1
293 integron genetic contexts. To the best of our knowledge, its association with an
294 *Alphaproteobacteria* host has not been reported previously.

295 As expected, the majority of the class 1 integron gene cassettes we detected are AMR related.
296 The dihydrofolate reductase encoded by *dfrA5* confers high levels of trimethoprim resistance
297 in Gram-negative bacteria [63]. We found an association between the *dfrA5* gene cassette and a
298 *Comamonadaceae*-family host (Figure 2C). The $blaOXA_{10}$ gene encodes a class D beta-
299 lactamase that hydrolyses the semi-synthetic antibiotic oxacillin [64, 65]. In our dataset, the
300 $blaOXA_{10}$ gene cassette was detected in a *Moraxellaceae*-family host (Figure 2D). Genes *qacF*
301 and *qacG2* encode small multidrug resistance (SMR) efflux pumps that confer resistance to
302 quaternary ammonium compounds (QACs) present in a wide range of disinfectants and
303 detergents [31, 66]. We found a *qacG2* gene cassette in an *Aquabacterium*-like host (Figure 2E)
304 and a *qacF* gene cassette at the proximal position in the gene cassette array in a *Rhizobacter*
305 host (Figure 2F). Also found in the class 1 integron of the *Rhizobacter* host is an *aadA2* gene
306 cassette. *aadA* genes confer resistance to the aminoglycoside antibiotics streptomycin and
307 spectinomycin [34, 67-70]. In the *Rhodocyclaceae*-family host, we found the same *aadA2* gene
308 cassette with a SNP in the non-coding region and a non-synonymous mutation (V257A) in the
309 coding region (Figure 2G and Table S3). The phylogenetic relationship between the seven hosts

310 and the distribution of class 1 integron gene cassettes amongst them are summarised in Figure
311 3.

312

313 **Discovery of *Rhizobacter* as novel carriers of class 1 integrons**

314 *Rhizobacter* is an environmental bacterial genus. Isolates of *Rhizobacter* species have been
315 successfully isolated from the rhizospheres of the *Panax* [71, 72] and *Bergenia* genera [73], as well
316 as from freshwater sediment [74]. There are still many questions regarding their potential
317 ecological roles. For example, *R. gummiphilus* naturally degrades rubber [44, 75, 76], which implies
318 beneficial roles in promoting nutrient recycling and bioremediation in soil. On the other hand,
319 *R. daucus* is thought to be an agricultural plant pathogen that causes galls in carrot root tissues
320 [77]. Associations between the *Rhizobacter* genus and class 1 integrons have not been reported
321 previously. In all currently available NCBI RefSeq genomes for the *Rhizobacter* genus (*n* =
322 11), we detected various non-class 1 integrons but did not find any bioinformatic evidence of
323 class 1 integrons.

324 We found two *Rhizobacter* isolates with an identical nucleotide sequence in the 16S rRNA V4
325 hypervariable region to that of the *Rhizobacter* host in our dataset (NCBI accession no.
326 AB835065.1 and LC734193.1) [78, 79]. To experimentally verify that the association between the
327 *Rhizobacter* and class 1 integron was *bona fide*, we designed *Rhizobacter*-specific 16S rRNA
328 gene primers that target the same primer binding sites in the 16S rRNA gene as R926 and R806.
329 By replacing the *intI1*_nested primer with *qacE*_F (the primer positions are indicated in Figure
330 S3A) in our epicPCR experimental pipeline, we successfully obtained the expected chimeric
331 epicPCR product featuring a *qacE* fragment and the *Rhizobacter* phylogenetic marker. This
332 provides the first experimental evidence of an association between class 1 integrons and the
333 *Rhizobacter* genus.

334 A recent study applied a modelling approach to understand the emergence and spread of novel
335 antimicrobial resistance genes [80]. The dispersal of antibiotic resistance genes from
336 environmental bacteria to humans has been highlighted as a major knowledge gap that
337 contributes to uncertainties in these models. Acquiring more experimental data for these gene
338 mobilisation and HGT events can improve future quantitative risk assessment models for
339 antibiotic resistance and our ability to curb transfer of antibiotic resistance genes between
340 bacteria. epicPCR has the potential to contribute to this endeavour given its strengths in
341 identifying environmental microbes that cannot be cultivated under laboratory conditions.

342 The seven bacterial hosts described in this study are expected to be a subset of all the class 1
343 integron hosts in our samples. Our bridging primer R519-*qacE* cannot bind to class 1 integrons
344 that lack *qacE*/Δ*qacE*. These include many chromosomal class 1 integrons and atypical class
345 1 integrons with IS*26*-mediated deletions in their 3'-CS [29, 41]. To detect these class 1 integrons
346 and their hosts requires new bridging primers that bind to the 3'-end of their gene cassette
347 arrays. In addition, the presence of epicPCR products that were detected in fewer than 3
348 replicates in this study could be confirmed by replacing the two *intI1*-targeting forward primers
349 (*intI1*_outer and *intI1*_nested) with specific primers that target the 5'-end of specific gene
350 cassettes of interest. The degenerate bases in the 16S rRNA gene-targeting primers also mean
351 that certain hosts are less likely to be detected than others due to primer bias. Another inherent
352 limitation is the variable resolution of taxonomic classifications when using 16S rRNA as
353 phylogenetic marker.

354 We established epicPCR as a valuable surveillance platform for uncovering novel associations
355 between class 1 integrons with diverse cassette arrays and their bacterial hosts in environmental
356 water samples. Using this single-cell genomic approach, we successfully identified
357 *Alphaproteobacteria* as novel hosts of a gene cassette encoding a cupin-domain protein of
358 unknown functions, and the *Rhizobacter* genus as novel carriers of class 1 integrons. It is

359 envisioned that this technology can be deployed to analyse water samples from sewage
360 treatment plants, as well as effluents from industrial, clinical, and agricultural settings where
361 antimicrobial usage is high. These surveillance efforts will hopefully lead to a timelier
362 detection of hotspots of class 1 integron mediated AMR transmission and more targeted
363 environmental mitigation measures. epicPCR also paves the way for future studies that seek to
364 uncover the "sources" and "sinks" of integron gene cassettes in bacterial communities, as well
365 as possible barriers to integron-mediated HGT that may also exist [81]. With greater diversity in
366 combinations of gene cassettes and bacterial hosts that are likely to be captured through future
367 epicPCR studies, we will be able to address outstanding questions and gain more insights into
368 possible origins of novel gene cassettes, which have so far remained the greatest mystery in the
369 field of integron research [32, 35, 82].

370

371 **Figure Caption**

372 **Figure 1**. Class 1 integrons and their co-amplification with 16S rRNA marker genes via
373 epicPCR.

374 **(A)** The class 1 integron carried by the IncW plasmid R388 has the canonical structure of
375 clinical class 1 integrons. *intI1* and *attI1* sequences are highly conserved in the 5'-conserved
376 segments (CS). The 3'-CS comprise a truncated quaternary ammonium compound (QAC)
377 resistance gene (Δ*qacE*) and a sulfonamide resistance gene (*sul1*). Between the 5'-CS and 3'-
378 CS are variable integron gene cassette arrays that contain *attC* recombination sites and gene
379 cassettes. Each gene cassette is typically associated with a single open-reading frame (ORF).

380 **(B)** During the first stage of epicPCR, the *intI1*_outer primer and the R519_*qacE* bridging
381 primer target *intI1* and *qacE* sequences in the 5'-CS and 3'-CS respectively and amplify the
382 class 1 integron. A short overhang introduced by the bridging primer to the *qacE* fragment end

383 allows the intermediate PCR product to act as a long primer and co-amplify the V4-V5
384 hypervariable regions of the 16S rRNA marker gene in the presence of the R926 reverse primer.

385 **(C)** During the second stage of epicPCR, the *intI1*_nested and AP27_short nested primers
386 further amplify the chimeric amplicons pooled from the previous step. AP27_short introduces
387 a non-degenerate tail sequence (in blue) to the 3'-end of the amplicons. During standard PCR
388 verification of chimeric amplicon formation (Figure S3A), this tail sequence is targeted by the
389 AP28_short non-degenerate primer, which is used in conjunction with the *qacE*_F forward
390 primer. The final chimeric epicPCR amplicons contain highly conserved features of class 1
391 integrons, including a 157 bp *intI1-attI1* region in the 5'-CS and the *qacE*/Δ*qacE* fragment in
392 the 3'-CS.

393

394 **Figure 2**: Class 1 integrons identified through epicPCR alongside the taxonomic classification
395 of their respective bacterial hosts. The *intI1* gene fragment is displayed in blue, the *attI1* site in
396 purple, and all the gene cassette ORFs in green. The six different *attC* sites are shown in red.

397

398 **Figure 3**: Phylogenetic tree constructed using the maximum likelihood method for the V4
399 hypervariable regions of 16S rRNA genes for the seven hosts of class 1 integrons identified in
400 this study. The integron gene cassettes associated with *Alphaproteobacteria* (in blue) and
401 *Gammaproteobacteria* (in red) are grouped accordingly. Gene cassettes that were found in
402 *Gammaproteobacteria* hosts of the *Burkholderiales* order are shown in yellow.

403

404

405

406 # Tables

407 **Table 1**. Comparison between the class 1 integron gene cassettes sequenced in this work and
408 annotated reference class 1 integron gene cassettes.

ORF in gene cassette	Predicted function of ORF gene product	NCBI reference gene cassette	% Nucleotide identity to NCBI reference gene cassette sequence
Cupin-domain-encoding gene	Unknown	FJ172388	99.8% nucleotide identity; 1 synonymous mutation in the coding region
dfrA5	Trimethoprim-resistant dihydrofolate reductase	X12868	100% nucleotide identity
blaOXA$_{10}$	Oxacillin-hydrolysing beta-lactamase	XXU37105	
qacG2	Quaternary ammonium compound efflux pump	AF327731	99.6% nucleotide identity; 2 non-synonymous mutations (S13F and A57G) in the coding region
qacF	Quaternary ammonium compound efflux pump	AF034958	99.6% nucleotide identity; 1 non-synonymous mutation (S10A) in the coding region
aadA2	Aminoglycoside (3'') adenylyltransferase conferring resistance to streptomycin and spectinomycin	X68227	In the *Rhizobacter* host: 100% nucleotide identity
			In the *Rhodocyclaceae* host: 99.6% nucleotide identity; 1 SNP in the non-coding region, 1 non-synonymous mutation (V257A) in the coding region

409

410 # Associated Content

411 **Supporting Information**: Additional experimental results and methods on the assignment of
412 taxonomy for bacterial hosts (Table S1); oligonucleotide primers (Table S2); nucleotide
413 sequences of gene cassettes (Table S3); hosts of the cupin-domain-encoding gene cassette
414 (Table S4); sampling location and photograph (Figure S1); examples of fluorescence
415 microscopy images for cell density estimation (Figure S2); additional epicPCR amplicons
416 (Figure S3); nucleotide sequence alignments of conserved regions in sequenced class 1
417 integrons (Figure S4) (PDF)

Author Contributions

Conceptualisation of study: TMG, QQ, SGT, MRG

Experimental investigation and validation: QQ, AP, JACS

Bioinformatic analysis: TMG, QQ, VR

Project management and supervision: MRG, SGT

Writing the original draft of the manuscript: QQ

Reviewing and editing the manuscript: All authors

Funding Sources

This work was funded by the Australian Research Council Discovery Project DP200101874 grant to MRG and SGT.

Acknowledgement

The authors thank Professor Marko Virta and his research group at the Department of Microbiology, University of Helsinki for sharing their expertise on epicPCR. We would like to acknowledge the Macquarie University Faculty of Science and Engineering Microscope Facility for access to its instrumentation and staff. We also thank Mr Richard Miller (IT Systems and HPC, Macquarie University) for his technical advice and support.

438 # References

439 (1) Jain, R.; Rivera, M. C.; Moore, J. E.; Lake, J. A. Horizontal gene transfer accelerates genome
440 innovation and evolution. Mol Biol Evol 2003, 20 (10), 1598-1602. DOI: 10.1093/molbev/msg154.

441 (2) Wiedenbeck, J.; Cohan, F. M. Origins of bacterial diversity through horizontal genetic transfer and
442 adaptation to new ecological niches. FEMS Microbiol Rev 2011, 35 (5), 957-976. DOI: 10.1111/j.1574-
443 6976.2011.00292.x.

444 (3) Boto, L. Horizontal gene transfer in evolution: Facts and challenges. Proc Biol Sci 2010, 277 (1683),
445 819-827. DOI: 10.1098/rspb.2009.1679.

446 (4) Beiko, R. G.; Harlow, T. J.; Ragan, M. A. Highways of gene sharing in prokaryotes. Proc Natl Acad
447 Sci U S A 2005, 102 (40), 14332-14337. DOI: 10.1073/pnas.0504068102.

448 (5) Gillings, M. R.; Holley, M. P.; Stokes, H. W. Evidence for dynamic exchange of *qac* gene cassettes
449 between class 1 integrons and other integrons in freshwater biofilms. FEMS Microbiol Lett 2009, 296
450 (2), 282-288. DOI: 10.1111/j.1574-6968.2009.01646.x.

451 (6) Sørensen, S. J.; Bailey, M.; Hansen, L. H.; Kroer, N.; Wuertz, S. Studying plasmid horizontal
452 transfer *in situ*: A critical review. Nat Rev Microbiol 2005, 3 (9), 700-710. DOI: 10.1038/nrmicro1232.

453 (7) Gillings, M. R. Lateral gene transfer, bacterial genome evolution, and the Anthropocene. Ann N Y
454 Acad Sci 2017, 1389 (1), 20-36. DOI: 10.1111/nyas.13213.

455 (8) Brown-Jaque, M.; Calero-Cáceres, W.; Muniesa, M. Transfer of antibiotic-resistance genes via
456 phage-related mobile elements. Plasmid 2015, 79, 1-7. DOI: 10.1016/j.plasmid.2015.01.001.

457 (9) Lerminiaux, N. A.; Cameron, A. D. S. Horizontal transfer of antibiotic resistance genes in clinical
458 environments. Can J Microbiol 2019, 65 (1), 34-44. DOI: 10.1139/cjm-2018-0275.

459 (10) Hall, J. P.; Wood, A. J.; Harrison, E.; Brockhurst, M. A. Source-sink plasmid transfer dynamics
460 maintain gene mobility in soil bacterial communities. Proc Natl Acad Sci U S A 2016, 113 (29), 8260-
461 8265. DOI: 10.1073/pnas.1600974113.

462 (11) Zhou, H.; Beltrán, J. F.; Brito, I. L. Functions predict horizontal gene transfer and the emergence
463 of antibiotic resistance. Sci Adv 2021, 7 (43), eabj5056. DOI: 10.1126/sciadv.abj5056.

464 (12) Abe, K.; Nomura, N.; Suzuki, S. Biofilms: Hot spots of horizontal gene transfer (HGT) in aquatic
465 environments, with a focus on a new HGT mechanism. FEMS Microbiol Ecol 2020, 96 (5), fiaa031.
466 DOI: 10.1093/femsec/fiaa031.

467 (13) Cárdenas, J. P.; Quatrini, R.; Holmes, D. S. Genomic and metagenomic challenges and
468 opportunities for bioleaching: A mini-review. Res Microbiol 2016, 167 (7), 529-538. DOI:
469 10.1016/j.resmic.2016.06.007.

470 (14) Thompson, L. R.; Sanders, J. G.; McDonald, D.; Amir, A.; Ladau, J.; Locey, K. J.; Prill, R. J.;
471 Tripathi, A.; Gibbons, S. M.; Ackermann, G.; Navas-Molina, J. A.; Janssen, S.; Kopylova, E.; Vázquez-
472 Baeza, Y.; González, A.; Morton, J. T.; Mirarab, S.; Xu, Z. Z.; Jiang, L.; Haroon, M. F.; Kanbar, J.;
473 Zhu, Q.; Song, S. J.; Kosciolek, T.; Bokulich, N. A.; Lefler, J.; Brislawn, C. J.; Humphrey, G.; Owens,
474 S. M.; Hampton-Marcell, J.; Berg-Lyons, D.; McKenzie, Fierer, V.; N.; Fuhrman, J. A.; Clauset, A.;
475 Stevens, R. L.; Shade, A.; Pollard, K. S.; Goodwin, K. D.; Jansson, J. K.; Gilbert, J. A.; Knight, R.; The
476 Earth Microbiome Project Consortium. A communal catalogue reveals Earth's multiscale microbial
477 diversity. Nature 2017, 551 (7681), 457-463. DOI: 10.1038/nature24621.

478 (15) Bahram, M.; Hildebrand, F.; Forslund, S. K.; Anderson, J. L.; Soudzilovskaia, N. A.; Bodegom, P.
479 M.; Bengtsson-Palme, J.; Anslan, S.; Coelho, L. P.; Harend, H.; Huerta-Cepas, J.; Medema, M. H.;
480 Maltz, M. R.; Mundra, S.; Olsson, P. A.; Pent, M.; Põlme, S.; Sunagawa, S.; Ryberg, M.; Tedersoo, L.;
481 Bork, P. Nature 2018, 560 (7717), 233-237. DOI: 10.1038/s41586-018-0386-6.

482 (16) Stewart, E. J. Growing unculturable bacteria. J Bacteriol 2012, 194 (16), 4151-4160. DOI:
483 10.1128/JB.00345-12.

484 (17) Douglas, G. M.; Langille, M. G. I. Current and promising approaches to identify horizontal gene
485 transfer events in metagenomes. Genome Biol Evol 2019, 11 (10), 2750-2766. DOI:
486 10.1093/gbe/evz184.

487 (18) Seiler, E.; Trappe, K.; Renard, B. Y. Where did you come from, where did you go: Refining
488 metagenomic analysis tools for horizontal gene transfer characterisation. PLoS Comput Biol 2019, 15
489 (7), e1007208. DOI: 10.1371/journal.pcbi.1007208.

490 (19) Ghurye, J. S.; Cepeda-Espinoza, V.; Pop, M. Metagenomic assembly: Overview, challenges and
491 applications. Yale J Biol Med 2016, 89 (3), 353-362.

492 (20) Darmon, E.; Leach, D. R. Bacterial genome instability. Microbiology and molecular biology
493 reviews: MMBR 2014, 78 (1), 1-39. DOI: 10.1128/MMBR.00035-13.

494 (21) Spencer, S. J.; Tamminen, M. V.; Preheim, S. P.; Guo, M. T.; Briggs, A. W.; Brito, I. L.; D, A. W.;
495 Pitkanen, L. K.; Vigneault, F.; Virta, M. P.; Alm, E. J. Massively parallel sequencing of single cells by
496 epicPCR links functional genes with phylogenetic markers. ISME J 2016, 10 (2), 427-436. DOI:
497 10.1038/ismej.2015.124.

498 (22) Roman, V. L.; Merlin, C.; Virta, M. P. J.; Bellanger, X. EpicPCR 2.0: Technical and
499 methodological improvement of a cutting-edge single-cell genomic approach. Microorganisms 2021, 9
500 (8), 1649. DOI: 10.3390/microorganisms9081649.

501 (23) Hultman, J.; Tamminen, M.; Parnanen, K.; Cairns, J.; Karkman, A.; Virta, M. Host range of
502 antibiotic resistance genes in wastewater treatment plant influent and effluent. FEMS Microbiol Ecol
503 2018, 94 (4), fiy038. DOI: 10.1093/femsec/fiy038.

504 (24) Qin, H.; Wang, S.; Feng, K.; He, Z.; Virta, M. P. J.; Hou, W.; Dong, H.; Deng, Y. Unraveling the
505 diversity of sedimentary sulfate-reducing prokaryotes (SRP) across Tibetan saline lakes using epicPCR.
506 Microbiome 2019, 7 (1), 71. DOI: 10.1186/s40168-019-0688-4.

507 (25) Sakowski, E. G.; Arora-Williams, K.; Tian, F.; Zayed, A. A.; Zablocki, O.; Sullivan, M. B.;
508 Preheim, S. P. Interaction dynamics and virus-host range for estuarine actinophages captured by
509 epicPCR. Nat Microbiol 2021, 6 (5), 630-642. DOI: 10.1038/s41564-021-00873-4.

510 (26) Diebold, P. J.; New, F. N.; Hovan, M.; Satlin, M. J.; Brito, I. L. Linking plasmid-based beta-
511 lactamases to their bacterial hosts using single-cell fusion PCR. eLife 2021, 10: e66834. DOI:
512 10.7554/eLife.66834.

513 (27) Sarekoski, A. K. The use of epicPCR to study antimicrobial resistance genes and their bacterial
514 hosts in human impacted environments in West Africa. University of Helsinki: 2020.

515 (28) Wei, Z.; Feng, K.; Wang, Z.; Zhang, Y.; Yang, M.; Zhu, Y. G.; Virta, M. P. J.; Deng, Y. High-
516 throughput single-cell technology reveals the contribution of horizontal gene transfer to typical
517 antibiotic resistance gene dissemination in wastewater treatment plants. Environ Sci Technol 2021, 55
518 (17), 11824-11834. DOI: 10.1021/acs.est.1c01250.

519 (29) Gillings, M.; Boucher, Y.; Labbate, M.; Holmes, A.; Krishnan, S.; Holley, M.; Stokes, H. W. The
520 evolution of class 1 integrons and the rise of antibiotic resistance. J Bacteriol 2008, 190 (14), 5095-
521 5100. DOI: 10.1128/JB.00152-08.

522 (30) Domingues, S.; Da Silva, G. J.; Nielsen, K. M. Global dissemination patterns of common gene
523 cassette arrays in class 1 integrons. Microbiology 2015, 161 (7), 1313-1337. DOI:
524 10.1099/mic.0.000099.

525 (31) Gillings, M. R.; Gaze, W. H.; Pruden, A.; Smalla, K.; Tiedje, J. M.; Zhu, Y. G. Using the class 1
526 integron-integrase gene as a proxy for anthropogenic pollution. ISME J 2015, 9 (6), 1269-1279. DOI:
527 10.1038/ismej.2014.226.

528 (32) Ghaly, T. M.; Geoghegan, J. L.; Tetu, S. G.; Gillings, M. R. The peril and promise of integrons:
529 beyond antibiotic resistance. Trends Microbiol 2020, 28 (6), 455-464. DOI: 10.1016/j.tim.2019.12.002.

530 (33) Escudero, J. A.; Loot, C.; Mazel, D. Integrons as adaptive devices. In Molecular Mechanisms of
531 Microbial Evolution, Rampelotto, P. H. Ed.; Springer International Publishing, 2018; 199-239.

532 (34) Partridge, S. R.; Tsafnat, G.; Coiera, E.; Iredell, J. R. Gene cassettes and cassette arrays in mobile
533 resistance integrons. FEMS Microbiol Rev 2009, 33 (4), 757-784. DOI: 10.1111/j.1574-
534 6976.2009.00175.x.

535 (35) Escudero, J. A.; Loot, C.; Nivina, A.; Mazel, D. The Integron: Adaptation on demand. Microbiol
536 Spectr 2015, 3 (2), MDNA3-0019-2014. DOI: 10.1128/microbiolspec.MDNA3-0019-2014.

537 (36) Recchia, G. D.; Stokes, H. W.; Hall, R. M. Characterisation of specific and secondary
538 recombination sites recognised by the integron DNA integrase. Nucleic Acids Res 1994, 22 (11), 2071-
539 2078. DOI: 10.1093/nar/22.11.2071.

540 (37) Zheng, W.; Huyan, J.; Tian, Z.; Zhang, Y.; Wen, X. Clinical class 1 integron-integrase gene - A
541 promising indicator to monitor the abundance and elimination of antibiotic resistance genes in an urban
542 wastewater treatment plant. Environ Int 2020, 135, 105372. DOI: 10.1016/j.envint.2019.105372.

543 (38) Jechalke, S.; Schreiter, S.; Wolters, B.; Dealtry, S.; Heuer, H.; Smalla, K. Widespread
544 dissemination of class 1 integron components in soils and related ecosystems as revealed by cultivation-
545 independent analysis. Front Microbiol 2013, 4, 420. DOI: 10.3389/fmicb.2013.00420.

546 (39) Pfeifer, E.; Bonnin, R. A.; Rocha, E. P. C. Phage-plasmids spread antibiotic resistance genes
547 through infection and lysogenic conversion. mBio 2022, 13 (5), e0185122. DOI: 10.1128/mbio.01851-
548 22.

549 (40) Waldron, L. S.; Gillings, M. R. Screening foodstuffs for class 1 integrons and gene cassettes. J Vis
550 Exp 2015, (100), e52889. DOI: 10.3791/52889.

551 (41) Dawes, F. E.; Kuzevski, A.; Bettelheim, K. A.; Hornitzky, M. A.; Djordjevic, S. P.; Walker, M. J.
552 Distribution of class 1 integrons with IS*26*-mediated deletions in their 3'-conserved segments in
553 *Escherichia coli* of human and animal origin. PLoS One 2010, 5 (9), e12754. DOI:
554 10.1371/journal.pone.0012754.

555 (42) Stokes, H. W.; Nesbø, C. L.; Holley, M.; Bahl, M. I.; Gillings, M. R.; Boucher, Y. Class 1 integrons
556 potentially predating the association with Tn*402*-like transposition genes are present in a sediment
557 microbial community. J Bacteriol 2006, 188 (16), 5722-5730. DOI: 10.1128/JB.01950-05.

558 (43) Rosewarne, C. P.; Pettigrove, V.; Stokes, H. W.; Parsons, Y. M. Class 1 integrons in benthic
559 bacterial communities: Abundance, association with Tn*402*-like transposition modules and evidence

560 for coselection with heavy-metal resistance. FEMS Microbiol Ecol 2010, 72 (1), 35-46. DOI:
561 10.1111/j.1574-6941.2009.00823.x.

562 (44) Kasai, D.; Imai, S.; Asano, S.; Tabata, M.; Iijima, S.; Kamimura, N.; Masai, E.; Fukuda, M.
563 Identification of natural rubber degradation gene in *Rhizobacter gummiphilus* NS21. Biosci Biotechnol
564 Biochem 2017, 81 (3), 614-620. DOI: 10.1080/09168451.2016.1263147.

565 (45) De Coster, W.; D'Hert, S.; Schultz, D. T.; Cruts, M.; Van Broeckhoven, C. NanoPack: Visualizing
566 and processing long-read sequencing data. Bioinformatics 2018, 34 (15), 2666-2669. DOI:
567 10.1093/bioinformatics/bty149.

568 (46) Šošic, M.; Šikic, M. Edlib: A C/C ++ library for fast, exact sequence alignment using edit distance.
569 Bioinformatics 2017, 33 (9), 1394-1395. DOI: 10.1093/bioinformatics/btw753.

570 (47) Sahlin, K.; Medvedev, P. *De novo* clustering of long-read transcriptome data using a greedy,
571 quality value-based algorithm. J Comput Biol 2020, 27 (4), 472-484. DOI: 10.1089/cmb.2019.0299.

572 (48) Sahlin, K.; Medvedev, P. Error correction enables use of Oxford Nanopore technology for
573 reference-free transcriptome analysis. Nat Commun 2021, 12 (1), 2. DOI: 10.1038/s41467-020-20340-
574 8.

575 (49) Vaser, R.; Sović, I.; Nagarajan, N.; Šikić, M. Fast and accurate *de novo* genome assembly from
576 long uncorrected reads. Genome Res 2017, 27 (5), 737-746. DOI: 10.1101/gr.214270.116.

577 (50) Bengtsson-Palme, J.; Hartmann, M.; Eriksson, K. M.; Pal, C.; Thorell, K.; Larsson, D. G.; Nilsson,
578 R. H. METAXA2: Improved identification and taxonomic classification of small and large subunit
579 rRNA in metagenomic data. Mol Ecol Resour 2015, 15 (6), 1403-1414. DOI: 10.1111/1755-
580 0998.12399.

581 (51) Li, W.; Godzik, A. Cd-hit: A fast program for clustering and comparing large sets of protein or
582 nucleotide sequences. Bioinformatics 2006, 22 (13), 1658-1659. DOI: 10.1093/bioinformatics/btl158.

583 (52) Neron, B.; Littner, E.; Haudiquet, M.; Perrin, A.; Cury, J.; Rocha, E. P. C. IntegronFinder 2.0:
584 Identification and analysis of integrons across bacteria, with a focus on antibiotic resistance in
585 *Klebsiella*. Microorganisms 2022, 10 (4), 700. DOI: 10.3390/microorganisms10040700.

586 (53) Quast, C.; Pruesse, E.; Yilmaz, P.; Gerken, J.; Schweer, T.; Yarza, P.; Peplies, J.; Glöckner, F. O.
587 The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools.
588 Nucleic Acids Res 2013, 41 (Database issue), D590-596. DOI: 10.1093/nar/gks1219.

589 (54) Guindon, S.; Dufayard, J. F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New algorithms
590 and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0.
591 Syst Biol 2010, 59 (3), 307-321. DOI: 10.1093/sysbio/syq010.

592 (55) Cury, J.; Jove, T.; Touchon, M.; Neron, B.; Rocha, E. P. Identification and analysis of integrons
593 and cassette arrays in bacterial genomes. Nucleic Acids Res 2016, 44 (10), 4539-4550. DOI:
594 10.1093/nar/gkw319.

595 (56) Zhang, Y.; Lu, J.; Wu, J.; Wang, J.; Luo, Y. Potential risks of microplastics combined with
596 superbugs: Enrichment of antibiotic resistant bacteria on the surface of microplastics in mariculture
597 system. Ecotoxicol Environ Saf 2020, 187, 109852. DOI: 10.1016/j.ecoenv.2019.109852.

598 (57) Zhang, Y.; Yang, X.; Huang, L.; Zhang, R. Whole-genome analysis and description of an IMP-8-
599 producing *Ochrobactrum anthropi*. J Glob Antimicrob Resist 2022, 29, 275-277. DOI:
600 10.1016/j.jgar.2022.03.016.

601 (58) Martini, M. C.; Quiroga, M. P.; Pistorio, M.; Lagares, A.; Centrón, D.; Del Papa, M. F. Novel
602 environmental class 1 integrons and cassette arrays recovered from an on-farm bio-purification plant.
603 FEMS Microbiol Ecol 2018, 94 (3), fix190. DOI: 10.1093/femsec/fix190.

604 (59) Gillings, M. R.; Xuejun, D.; Hardwick, S. A.; Holley, M. P.; Stokes, H. W. Gene cassettes encoding
605 resistance to quaternary ammonium compounds: a role in the origin of clinical class 1 integrons? ISME
606 J 2009, 3 (2), 209-215. DOI: 10.1038/ismej.2008.98.

607 (60) Martinez-Freijo, P.; Fluit, A. C.; Schmitz, F. J.; Grek, V. S.; Verhoef, J.; Jones, M. E. Class I
608 integrons in Gram-negative isolates from different European hospitals and association with decreased
609 susceptibility to multiple antibiotic compounds. J Antimicrob Chemother 1998, 42 (6), 689-696. DOI:
610 10.1093/jac/42.6.689.

611 (61) van Essen-Zandbergen, A.; Smith, H.; Veldman, K.; Mevius, D. Occurrence and characteristics of
612 class 1, 2 and 3 integrons in *Escherichia coli*, *Salmonella* and *Campylobacter* spp. in the Netherlands.
613 J Antimicrob Chemother 2007, 59 (4), 746-750. DOI: 10.1093/jac/dkl549.

614 (62) Dunwell, J. M.; Culham, A.; Carter, C. E.; Sosa-Aguirre, C. R.; Goodenough, P. W. Evolution of
615 functional diversity in the cupin superfamily. Trends Biochem Sci 2001, 26 (12), 740-746. DOI:
616 10.1016/s0968-0004(01)01981-8.

617 (63) Krucinska, J.; Lombardo, M. N.; Erlandsen, H.; Estrada, A.; Si, D.; Viswanathan, K.; Wright, D.
618 L. Structure-guided functional studies of plasmid-encoded dihydrofolate reductases reveal a common
619 mechanism of trimethoprim resistance in Gram-negative pathogens. Commun Biol 2022, 5 (1), 459.
620 DOI: 10.1038/s42003-022-03384-y.

621 (64) Hall, L. M.; Livermore, D. M.; Gur, D.; Akova, M.; Akalin, H. E. OXA-11, an extended-spectrum
622 variant of OXA-10 (PSE-2) beta-lactamase from *Pseudomonas aeruginosa*. Antimicrob Agents
623 Chemother 1993, 37 (8), 1637-1644. DOI: 10.1128/AAC.37.8.1637.

624 (65) Evans, B. A.; Amyes, S. G. OXA β-lactamases. Clin Microbiol Rev 2014, 27 (2), 241-263. DOI:
625 10.1128/CMR.00117-13.

626 (66) Gaze, W. H.; Zhang, L.; Abdouslam, N. A.; Hawkey, P. M.; Calvo-Bado, L.; Royle, J.; Brown, H.;
627 Davis, S.; Kay, P.; Boxall, A. B.; Wellington, E. M. H. Impacts of anthropogenic activity on the ecology
628 of class 1 integrons and integron-associated genes in the environment. ISME J 2011, 5 (8), 1253-1261.
629 DOI: 10.1038/ismej.2011.15.

630 (67) Antunes, P.; Machado, J.; Peixe, L. Dissemination of *sul3*-containing elements linked to class 1
631 integrons with an unusual 3' conserved sequence region among *Salmonella* isolates. Antimicrob Agents
632 Chemother 2007, 51 (4), 1545-1548. DOI: 10.1128/AAC.01275-06.

633 (68) Chang, C. Y.; Chang, L. L.; Chang, Y. H.; Lee, T. M.; Chang, S. F. Characterisation of drug
634 resistance gene cassettes associated with class 1 integrons in clinical isolates of *Escherichia coli* from
635 Taiwan, ROC. J Med Microbiol 2000, 49 (12), 1097-1102. DOI: 10.1099/0022-1317-49-12-1097.

636 (69) Clark, N. C.; Olsvik, O.; Swenson, J. M.; Spiegel, C. A.; Tenover, F. C. Detection of a
637 streptomycin/spectinomycin adenylyltransferase gene (*aadA*) in *Enterococcus faecalis*. Antimicrob
638 Agents Chemother 1999, 43 (1), 157-160. DOI: 10.1128/AAC.43.1.157.

639 (70) Binh, C. T.; Heuer, H.; Kaupenjohann, M.; Smalla, K. Diverse *aadA* gene cassettes on class 1
640 integrons introduced into soil *via* spread manure. Res Microbiol 2009, 160 (6), 427-433. DOI:
641 10.1016/j.resmic.2009.06.005.

642 (71) Yoon, M. H.; Ten, L. N.; Im, W. T.; Lee, S. T. *Methylibium fulvum* sp. nov., a member of the
643 *Betaproteobacteria* isolated from ginseng field soil, and emended description of the genus *Methylibium*.
644 Int J Syst Evol Microbiol 2007, 57 (Pt 9), 2062-2066. DOI: 10.1099/ijs.0.64909-0.

645 (72) Stackebrandt, E.; Verbarg, S.; Frühling, A.; Busse, H. J.; Tindall, B. J. Dissection of the genus
646 *Methylibium*: Reclassification of *Methylibium fulvum* as *Rhizobacter fulvus* comb. nov., *Methylibium*
647 *aquaticum* as *Piscinibacter aquaticus* gen. nov., comb. nov. and *Methylibium subsaxonicum* as
648 *Rivibacter subsaxonicus* gen. nov., comb. nov. and emended descriptions of the genera *Rhizobacter* and
649 *Methylibium*. Int J Syst Evol Microbiol 2009, 59 (Pt 10), 2552-2560. DOI: 10.1099/ijs.0.008383-0.

650 (73) Wei, L.; Si, M.; Long, M.; Zhu, L.; Li, C.; Shen, X.; Wang, Y.; Zhao, L.; Zhang, L. *Rhizobacter*
651 *bergeniae* sp. nov., isolated from the root of *Bergenia scopulosa*. Int J Syst Evol Microbiol 2015, 65
652 (Pt 2), 479-484. DOI: 10.1099/ijs.0.070318-0.

653 (74) Jin, L.; Ko, S. R.; Ahn, C. Y.; Lee, H. G.; Oh, H. M. *Rhizobacter profundi* sp. nov., isolated from
654 freshwater sediment. Int J Syst Evol Microbiol 2016, 66 (5), 1926-1931. DOI: 10.1099/ijsem.0.000962.

(75) Imai, S.; Yoshida, R.; Endo, Y.; Fukunaga, Y.; Yamazoe, A.; Kasai, D.; Masai, E.; Fukuda, M. *Rhizobacter gummiphilus* sp. nov., a rubber-degrading bacterium isolated from the soil of a botanical garden in Japan. J Gen Appl Microbiol 2013, 59 (3), 199-205. DOI: 10.2323/jgam.59.199.

(76) Linh, D. V.; Gibu, N.; Tabata, M.; Imai, S.; Hosoyama, A.; Yamazoe, A.; Kasai, D.; Fukuda, M. Complete genome sequence of natural rubber-degrading, Gram-negative bacterium, *Rhizobacter gummiphilus* strain NS21^{T}. Biotechnol Rep (Amst) 2019, 22, e00332. DOI: 10.1016/j.btre.2019.e00332.

(77) Goto, M.; Kuwata, H. *Rhizobacter daucus* gen. nov., sp. nov., the causal agent of carrot bacterial gall. Int J Syst Evol Microbiol 1988, 38 (3), 233-239. DOI: 10.1099/00207713-38-3-233.

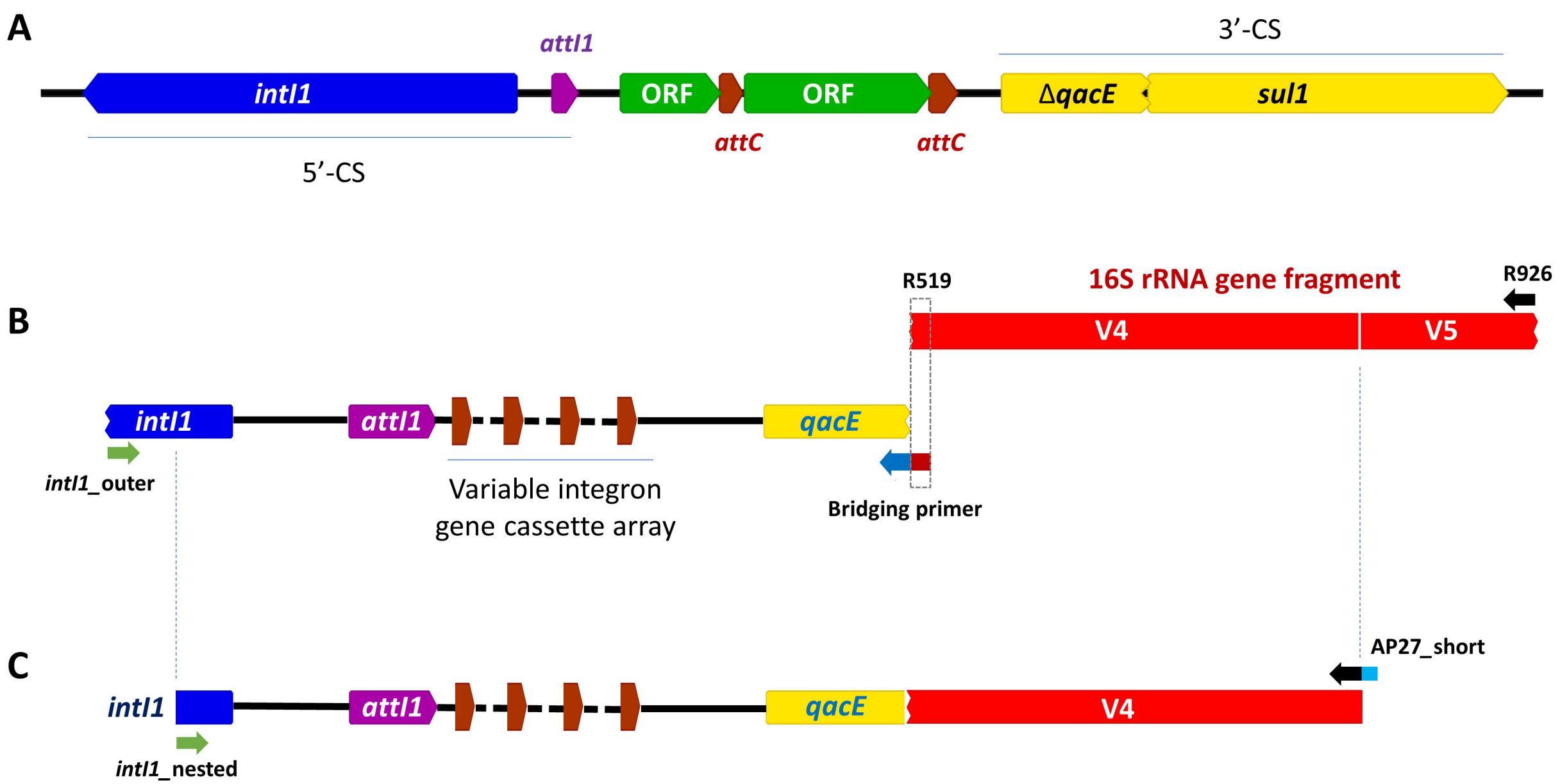
(78) Watanabe, M.; Igarashi, K.; Kato, S.; Kamagata, Y.; Kitagawa, W. Critical effect of H_2O_2 in the agar plate on the growth of laboratory and environmental strains. Microbiol Spectr 2022, e0333622. DOI: 10.1128/spectrum.03336-22.

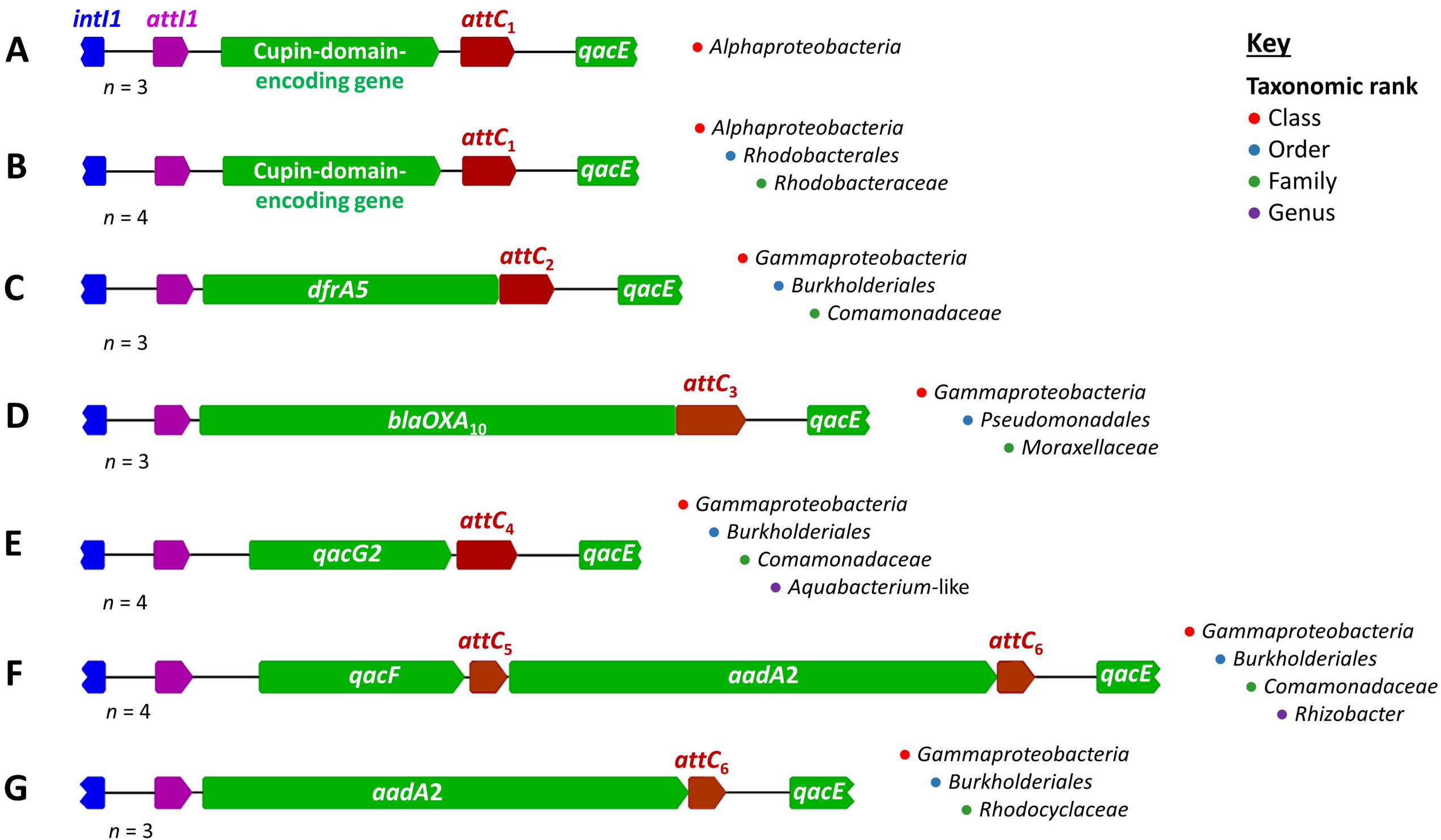
(79) Tanaka, T.; Kawasaki, K.; Daimon, S.; Kitagawa, W.; Yamamoto, K.; Tamaki, H.; Tanaka, M.; Nakatsu, C. H.; Kamagata, Y. A hidden pitfall in the preparation of agar media undermines microorganism cultivability. Appl Environ Microbiol 2014, 80 (24), 7659-7666. DOI: 10.1128/AEM.02741-14.

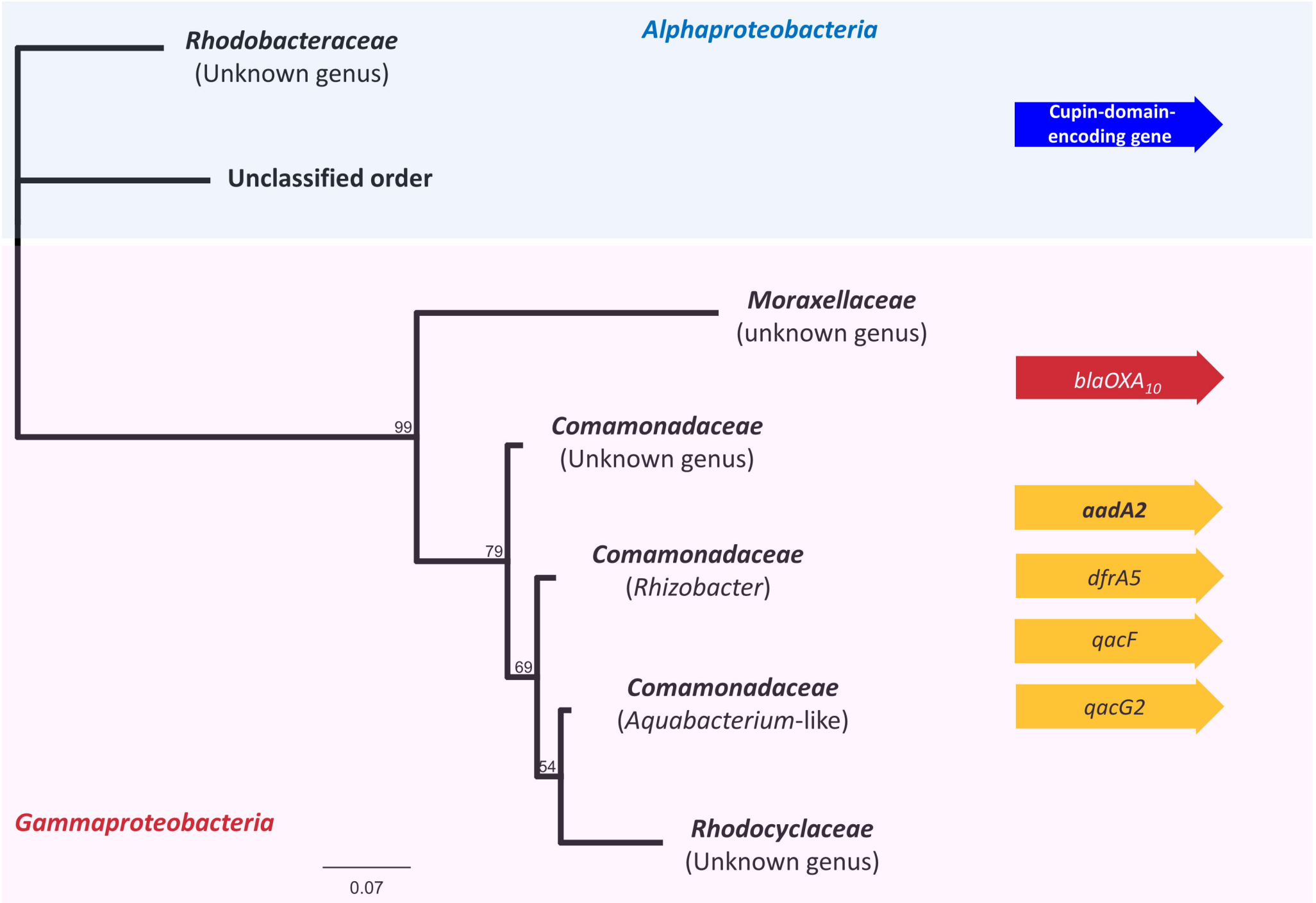
(80) Bengtsson-Palme, J.; Jonsson, V.; Hess, S. What is the role of the environment in the emergence of novel antibiotic resistance genes? A modeling approach. Environ Sci Technol 2021, 55 (23), 15734-15743. DOI: 10.1021/acs.est.1c02977.

(81) Kurland, C. G.; Canback, B.; Berg, O. G. Horizontal gene transfer: A critical view. Proc Natl Acad Sci U S A 2003, 100 (17), 9658-9662. DOI: 10.1073/pnas.1632870100.

(82) Ghaly, T. M.; Gillings, M. R.; Penesyan, A.; Qi, Q.; Rajabal, V.; Tetu, S. G. The natural history of integrons. Microorganisms 2021, 9 (11), 2212. DOI: 10.3390/microorganisms9112212.


A
attI1
3'-CS
intI1
ORF
ORF
ΔqacE
sul1
attC
attC
5'-CS
B
R519
16S rRNA gene fragment
R926
V4
V5
intI1
attI1
qacE
intI1_outer
Variable integron
gene cassette array
Bridging primer
C
AP27_short
intI1
attI1
qacE
V4
intI1_nested


A
intI1
attI1
Cupin-domain-encoding gene
attC1
qacE
n = 3
Alphaproteobacteria
B
Cupin-domain-encoding gene
attC1
qacE
n = 4
Alphaproteobacteria
Rhodobacterales
Rhodobacteraceae
C
dfrA5
attC2
qacE
n = 3
Gammaproteobacteria
Burkholderiales
Comamonadaceae
D
blaOXA10
attC3
qacE
n = 3
Gammaproteobacteria
Pseudomonadales
Moraxellaceae
E
qacG2
attC4
qacE
n = 4
Gammaproteobacteria
Burkholderiales
Comamonadaceae
Aquabacterium-like
F
qacF
attC5
aadA2
attC6
qacE
n = 4
Gammaproteobacteria
Burkholderiales
Comamonadaceae
Rhizobacter
G
aadA2
attC6
qacE
n = 3
Gammaproteobacteria
Burkholderiales
Rhodocyclaceae
Key
Taxonomic rank
Class
Order
Family
Genus


Alphaproteobacteria
Rhodobacteraceae
(Unknown genus)
Unclassified order
Cupin-domain-encoding gene
Moraxellaceae
(unknown genus)
blaOXA10
99
Comamonadaceae
(Unknown genus)
aadA2
79
Comamonadaceae
(Rhizobacter)
dfrA5
qacF
69
Comamonadaceae
(Aquabacterium-like)
qacG2
54
Gammaproteobacteria
Rhodocyclaceae
(Unknown genus)
0.07