Uncovering bacterial hosts of class 1 integrons in an urban coastal aquatic environment with a single-cell fusion-PCR 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 PCR) 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 related taxa or between more distantly related taxa found in similar environments [4, 5]. A key

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

54 epicPCR (emulsion, paired isolation and concatenation PCR) is a fusion PCR-based technique
55 that overcomes these problems by co-amplifying taxonomic markers with functional genes of
56 interest from single uncultured bacterial cells. Single cells are trapped in droplets generated
57 by emulsifying an aqueous phase containing PCR reagents and environmental bacteria in an
58 emulsion oil [21-26]. This compartmentalises the fusion PCR into parallel, single cell-level
59 reactions. Within each droplet, chimeric amplicons that combine genes of interest with
60 taxonomic markers corresponding to diverse phylogenetic groups are generated for high-
61 throughput sequencing. To date, epicPCR has been successfully applied to target functional
62 genes including dissimilatory sulfite reductase *dsrB* genes [21, 24], antibiotic resistance genes [22,
63 23, 26, 27], class 1 integron associated genes (*intI1*, *sul1*, Δ*qacE*) [23, 27, 28], and bacteriophage-borne
64 ribonucleotide reductase *RNR* genes [25]. epicPCR amplicons typically comprise short fragments
65 of genes of interest linked to the V4 hypervariable region of 16S rRNA markers. Illumina
66 sequencing of recovered amplicons demonstrates the presence of the genes of interest in a
67 variety of bacterial taxa. However, no studies have so far applied epicPCR to target variable,
68 multi-gene loci of interest.

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

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

94 *Alphaproteobacteria* hosts. Furthermore, our results demonstrated that *Rhizobacter* is a novel
95 host of class 1 integrons. Overall, our study shows that epicPCR is a powerful tool for
96 determining the environmental hosts of class 1 integrons with the potential to access previously
97 untapped carriers of these genetic elements. This paves the way for future studies that seek to
98 pinpoint the "sources" and "sinks" of class 1 integrons.

99

100 ## Materials and Methods

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

102 Coastal water samples were collected in three 50 mL biological replicates after rain in a rocky
103 intertidal zone downstream from a stormwater outlet at Shark Point (location coordinates: -
104 33.91, 151.27) near Sydney, NSW, Australia (Supplementary Figure 1). The samples were
105 filtered using EASYstrainer cell strainers with a mesh size of 40 µm (Greiner AG, Germany)
106 and centrifuged at 3220 *g* for 15 minutes. Cell pellets were resuspended in phosphate buffered
107 saline (PBS) solution at pH 7.4. Resuspended cells were stored as 25% glycerol stocks.

108

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

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

116 calculated from the average number of cells in five 0.2 mm x 0.2 mm square grids of the
117 counting chamber for each sample (Supplementary Figure 2).

118

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

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

132 ABIL emulsion oil was prepared by supplementing mineral oil with 4% ABIL EM 90 (Redox,
133 Australia) and 0.05% Triton X-100 (Promega, United States). 425 µL ABIL oil was added to
134 the PCR mix, which was immediately emulsified at 4 ms^{-1} for 45 s using the FastPrep-24 bead
135 beating system (MP Biomedicals, United States). The water-in-oil emulsion was aliquoted into
136 8 portions, incubated at 37˚C for 10 min, and was subjected to the following PCR amplification
137 conditions: 98˚C for 5 minutes; 38 cycles of 98°C for 10 s, 59°C for 20 s, and 72°C for 2 min;
138 and finally 72°C for 5 min. The *intI1*_outer primer and the R519-*qacE* bridging primer bind to
139 *intI1* in the 5'-conserved segment (CS) and *qacE* in the 3'-CS of class 1 integrons respectively

140 (Figure 1B). The 18 bp overhang introduced by the bridging primer to the *qacE* end allows the
141 intermediate PCR products to act as a long primer and amplify the V4-V5 hypervariable region
142 with the R926 reverse primer.

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

156 To confirm that fusion between class 1 integron and 16S DNA fragments has occurred
157 (Supplementary Figure 3A), 2 ng of each purified epicPCR product was amplified using
158 *qacE*_F (0.4 µM) and AP28_short (0.4 µM) primers in GoTaq polymerase master mix
159 (Promega, United States) under the following thermocycling conditions: 95˚C for 30 s; 30
160 cycles of 95°C for 30 s, 55°C for 30 s and 72°C for 30 s; with a final step of 72°C for 5 min.
161 Successful amplification of the *qacE*-16S rRNA gene chimeric region should produce ~390 bp
162 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 Supplementary
168 Figure 3B: the 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 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 have 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 (Supplementary Figure 1). We anticipated that class 1 integron-associated environmental
248 bacteria would be detected at this location, because surface runoff was expected to carry
249 pollutants from the surrounding urban areas. A preponderance of cells displayed unicellular
250 and planktonic morphologies (Supplementary Figure 2). We applied modified epicPCR
251 procedures to amplify class 1 integron gene cassette arrays from single bacterial cells
252 encapsulated by PCR reagent droplets homogenously dispersed in an oil emulsion and
253 successfully generated chimeric epicPCR products containing class 1 integrons and 16S rRNA
254 gene sequences (Figure 1). These epicPCR products were Nanopore sequenced. The median
255 number of reads per sample was 232,402. After stringent bioinformatic filtering, we identified
256 six class 1 integrons in seven distinct bacterial hosts (Figure 2). All seven combinations of
257 epicPCR products were observed in at least three independent replicates ($n \geq 3$, ≥99% pairwise
258 nucleotide identity and ≥99% identical sites relative to the consensus sequence for each
259 epicPCR product). In addition, there were five combinations of class 1 integrons and hosts that
260 were detected in fewer than 3 replicates and were consequently excluded. Taxonomic

261 classifications were obtained using the SILVA ribosomal RNA gene database (Supplementary
262 Table S1) [53]. *Alphaproteobacteria* (*n*=2) and *Gammaproteobacteria* (*n*=5) were the two major
263 classes of hosts represented in our dataset, while *Burkholderiales* (*n*=4) was the most common
264 bacterial order. These observations are broadly consistent with previous findings that class 1
265 integrons are associated with *Gammaproteobacteria* [29, 42, 43]. The finding of integrons in two
266 *Alphaproteobacteria* is notable, because *Alphaproteobacteria* are rarely known to be
267 associated with integrons [55]. Currently, there are very few examples of *Alphaproteobacteria*
268 species that harbour class 1 integrons [56-58]. This underscores the potential of epicPCR
269 approaches to discover novel host-function 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 Supplementary Table S3. We compared their nucleotide sequences against
273 reference class 1 integron gene cassettes from the NCBI database and found that both the
274 coding and non-coding regions of these gene cassettes were highly conserved relative to the
275 reference sequences [34, 59]. In our dataset, there are six different *attC* recombination sites with
276 100% nucleotide sequence identity to those in the respective reference gene cassettes in Table
277 1. We also aligned all the nucleotide sequences in the 5'-CS and 3'-CS regions of all the class
278 1 integrons from this study and found that both the 157 bp *intI1/attI1* and 190 bp *attC*/*qacE*
279 regions in the chimeric amplicons were identical in all but one exception (Supplementary
280 Figure 4). Overall, these class 1 integron structures are typical of those recovered from clinical
281 bacterial isolates [29, 60, 61].

282

283

284

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

286 Cupins belong to a superfamily of proteins with a beta-barrel fold and were originally found as
287 highly conserved domains in plant proteins [62]. We found a gene cassette encoding a cupin barrel
288 domain of unknown functions in two *Alphaproteobacteria* hosts. One host was unclassified at
289 the order level (Figure 2A), while the other belonged to the *Rhodobacteraceae* family (Figure
290 2B). This cupin-domain-encoding gene was previously observed in a class 1 integron gene
291 cassette from an unknown host [59]. Based on identical nucleotide sequence matches from the
292 NCBI nucleotide database, this gene cassette was detected in class 1 integrons, non-class 1
293 integrons, and CALINs (clusters of *attC*s lacking an associated integron-integrase) in the
294 chromosomes or plasmids of several *Gammaproteobacteria* hosts (Supplementary Table S4).
295 This observation suggests selective pressure for maintaining this gene cassette in both class 1
296 and non-class 1 integron genetic contexts. To the best of our knowledge, its association with
297 an *Alphaproteobacteria* host has not been reported previously.

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

310 spectinomycin [34, 67-70]. In the *Rhodocyclaceae*-family host, we found the same *aadA2* gene
311 cassette with a SNP in the non-coding region and a non-synonymous mutation (V257A) in the
312 coding region (Figure 2G and Supplementary Table S3). The phylogenetic relationship
313 between the seven hosts and the distribution of class 1 integron gene cassettes amongst them
314 are summarised in Figure 3.

315

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

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

327 We found two *Rhizobacter* isolates with identical nucleotide sequence in the 16S rRNA V4
328 hypervariable region to that of the *Rhizobacter* host in our dataset (NCBI accession no.
329 AB835065.1 and LC734193.1) [78, 79]. To experimentally verify that the association between the
330 *Rhizobacter* and class 1 integron was *bona fide*, we designed *Rhizobacter*-specific 16S rRNA
331 gene primers that target the same primer binding sites in the 16S rRNA gene as R926 and R806.
332 By replacing the *intI1*_nested primer with *qacE*_F (the primer positions are indicated in
333 Supplementary Figure 3A) in our epicPCR experimental pipeline, we successfully obtained the

334 expected chimeric epicPCR product featuring a *qacE* fragment and the *Rhizobacter*
335 phylogenetic marker. This provides the first experimental evidence of an association between
336 class 1 integrons and the *Rhizobacter* genus.

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

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

357 We established epicPCR as a valuable surveillance platform for uncovering novel associations
358 between class 1 integrons with diverse cassette arrays and their bacterial hosts in environmental
359 water samples. Using this single-cell genomic approach, we successfully identified
360 *Alphaproteobacteria* as novel hosts of a gene cassette encoding a cupin-domain protein of
361 unknown functions, and the *Rhizobacter* genus as novel carriers of class 1 integrons. It is
362 envisaged that this technology can be deployed to analyse water samples from sewage
363 treatment plants, as well as effluents from industrial, clinical, and agricultural settings where
364 antimicrobial usage is high. These surveillance efforts will hopefully lead to a timelier
365 detection of hotspots of class 1 integron mediated AMR transmission and more targeted
366 environmental mitigation measures. epicPCR also paves the way for future studies that seek to
367 uncover the "sources" and "sinks" of integron gene cassettes in bacterial communities, as well
368 as possible barriers to integron-mediated HGT that may also exist [81]. With greater diversity in
369 combinations of gene cassette and bacterial hosts that are likely to be captured through future
370 epicPCR studies, we will be able to address outstanding questions and gain more insights into
371 possible origins of novel gene cassettes, which has so far remained the greatest mystery in the
372 field of integron research [32, 35, 82].

373

374 **Figure Caption**

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

377 **(A)** The class 1 integron carried by the IncW plasmid R388 has the canonical structure of
378 clinical class 1 integrons. *intI1* and *attI1* sequences are highly conserved in the 5'-conserved
379 segments (CS). The 3'-CS comprise a truncated quaternary ammonium compound (QAC)
380 resistance gene (Δ*qacE*) and a sulfonamide resistance gene (*sul1*). Between the 5'-CS and 3'-

381 CS are variable integron gene cassette arrays that contain *attC* recombination sites and gene
382 cassettes. Each gene cassette is typically associated with a single open-reading frame (ORF).

383 **(B)** During the first stage of epicPCR, the *intI1*_outer and R519_*qacE* bridging primers target
384 *intI1* and *qacE* sequences in the 5'-CS and 3'-CS respectively and amplify the class 1 integron.
385 A short overhang introduced by the bridging primer to the *qacE* fragment end allows the
386 intermediate PCR product to act as a long primer and co-amplify the V4-V5 hypervariable
387 regions of the 16S rRNA marker gene in the presence of the R926 reverse primer.

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

396

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

400

401 **Figure 3**: Phylogenetic tree constructed using the maximum likelihood method for the V4
402 hypervariable regions of 16S rRNA genes for the seven hosts of class 1 integrons identified in
403 this study. The integron gene cassettes associated with *Alphaproteobacteria* (in blue) and

404 *Gammaproteobacteria* (in red) are grouped accordingly. Gene cassettes that were found in
405 *Gammaproteobacteria* hosts of the *Burkholderiales* order are shown in yellow.

406

407 # Tables

408 **Table 1**. Comparison between the class 1 integron gene cassettes sequenced in this work and
409 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

410

411

412

413

414

415 ## Associated Content

416 **Supporting Information**: Additional experimental results and methods on the assignment of
417 taxonomy for bacterial hosts (Table S1); oligonucleotide primers (Table S2); nucleotide
418 sequences of gene cassettes (Table S3); hosts of the cupin-domain-encoding gene cassette
419 (Table S4); sampling location and photograph (Supplementary Figure 1); examples of
420 fluorescence microscopy images for cell density estimation (Supplementary Figure 2);
421 additional epicPCR amplicons (Supplementary Figure 3); nucleotide sequence alignments of
422 conserved regions in sequenced class 1 integrons (Supplementary Figure 4) (PDF)

423

424

425 ## Author Contributions

426 Conceptualisation of study: TMG, QQ, SGT, MRG

427 Experimental investigation and validation: QQ, AP, JACS

428 Bioinformatic analysis: TMG, QQ, VR

429 Project management and supervision: MRG, SGT

430 Writing the original draft of the manuscript: QQ

431 Reviewing and editing the manuscript: All authors

432

433

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.

References

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

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

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

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

455 (5) Gillings, M. R.; Holley, M. P.; Stokes, H. W. Evidence for dynamic exchange of *qac* gene cassettes
456 between class 1 integrons and other integrons in freshwater biofilms. FEMS Microbiol Lett 2009, 296
457 (2), 282-288. DOI: 10.1111/j.1574-6968.2009.01646.x.

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

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

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

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

466 (10) Hall, J. P.; Wood, A. J.; Harrison, E.; Brockhurst, M. A. Source-sink plasmid transfer dynamics
467 maintain gene mobility in soil bacterial communities. Proc Natl Acad Sci U S A 2016, 113 (29), 8260-
468 8265. DOI: 10.1073/pnas.1600974113.

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

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

474 (13) Cárdenas, J. P.; Quatrini, R.; Holmes, D. S. Genomic and metagenomic challenges and
475 opportunities for bioleaching: A mini-review. Res Microbiol 2016, 167 (7), 529-538. DOI:
476 10.1016/j.resmic.2016.06.007.

477 (14) Thompson, L. R.; Sanders, J. G.; McDonald, D.; Amir, A.; Ladau, J.; Locey, K. J.; Prill, R. J.;
478 Tripathi, A.; Gibbons, S. M.; Ackermann, G.; Navas-Molina, J. A.; Janssen, S.; Kopylova, E.; Vázquez-
479 Baeza, Y.; González, A.; Morton, J. T.; Mirarab, S.; Xu, Z. Z.; Jiang, L.; Haroon, M. F.; Kanbar, J.;

480 Zhu, Q.; Song, S. J.; Kosciolek, T.; Bokulich, N. A.; Lefler, J.; Brislawn, C. J.; Humphrey, G.; Owens,
481 S. M.; Hampton-Marcell, J.; Berg-Lyons, D.; McKenzie, Fierer, V.; N.; Fuhrman, J. A.; Clauset, A.;
482 Stevens, R. L.; Shade, A.; Pollard, K. S.; Goodwin, K. D.; Jansson, J. K.; Gilbert, J. A.; Knight, R.; The
483 Earth Microbiome Project Consortium. A communal catalogue reveals Earth's multiscale microbial
484 diversity. Nature 2017, 551 (7681), 457-463. DOI: 10.1038/nature24621.

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

489 (16) Stewart, E. J. Growing unculturable bacteria. J Bacteriol 2012, 194 (16), 4151-4160. DOI:
490 10.1128/JB.00345-12.

491 (17) Douglas, G. M.; Langille, M. G. I. Current and promising approaches to identify horizontal gene
492 transfer events in metagenomes. Genome Biol Evol 2019, 11 (10), 2750-2766. DOI:
493 10.1093/gbe/evz184.

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

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

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

501 (21) Spencer, S. J.; Tamminen, M. V.; Preheim, S. P.; Guo, M. T.; Briggs, A. W.; Brito, I. L.; D, A. W.;
502 Pitkanen, L. K.; Vigneault, F.; Virta, M. P.; Alm, E. J. Massively parallel sequencing of single cells by
503 epicPCR links functional genes with phylogenetic markers. ISME J 2016, 10 (2), 427-436. DOI:
504 10.1038/ismej.2015.124.

505 (22) Roman, V. L.; Merlin, C.; Virta, M. P. J.; Bellanger, X. EpicPCR 2.0: Technical and
506 methodological improvement of a cutting-edge single-cell genomic approach. Microorganisms 2021, 9
507 (8), 1649. DOI: 10.3390/microorganisms9081649.

508 (23) Hultman, J.; Tamminen, M.; Parnanen, K.; Cairns, J.; Karkman, A.; Virta, M. Host range of
509 antibiotic resistance genes in wastewater treatment plant influent and effluent. FEMS Microbiol Ecol
510 2018, 94 (4), fiy038. DOI: 10.1093/femsec/fiy038.

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

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

517 (26) Diebold, P. J.; New, F. N.; Hovan, M.; Satlin, M. J.; Brito, I. L. Linking plasmid-based beta-
518 lactamases to their bacterial hosts using single-cell fusion PCR. eLife 2021, 10: e66834. DOI:
519 10.7554/eLife.66834.

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

522 (28) Wei, Z.; Feng, K.; Wang, Z.; Zhang, Y.; Yang, M.; Zhu, Y. G.; Virta, M. P. J.; Deng, Y. High-
523 throughput single-cell technology reveals the contribution of horizontal gene transfer to typical
524 antibiotic resistance gene dissemination in wastewater treatment plants. Environ Sci Technol 2021, 55
525 (17), 11824-11834. DOI: 10.1021/acs.est.1c01250.

526 (29) Gillings, M.; Boucher, Y.; Labbate, M.; Holmes, A.; Krishnan, S.; Holley, M.; Stokes, H. W. The
527 evolution of class 1 integrons and the rise of antibiotic resistance. J Bacteriol 2008, 190 (14), 5095-
528 5100. DOI: 10.1128/JB.00152-08.

529 (30) Domingues, S.; Da Silva, G. J.; Nielsen, K. M. Global dissemination patterns of common gene
530 cassette arrays in class 1 integrons. Microbiology 2015, 161 (7), 1313-1337. DOI:
531 10.1099/mic.0.000099.

532 (31) Gillings, M. R.; Gaze, W. H.; Pruden, A.; Smalla, K.; Tiedje, J. M.; Zhu, Y. G. Using the class 1
533 integron-integrase gene as a proxy for anthropogenic pollution. ISME J 2015, 9 (6), 1269-1279. DOI:
534 10.1038/ismej.2014.226.

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

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

539 (34) Partridge, S. R.; Tsafnat, G.; Coiera, E.; Iredell, J. R. Gene cassettes and cassette arrays in mobile
540 resistance integrons. FEMS Microbiol Rev 2009, 33 (4), 757-784. DOI: 10.1111/j.1574-
541 6976.2009.00175.x.

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

544 (36) Recchia, G. D.; Stokes, H. W.; Hall, R. M. Characterisation of specific and secondary
545 recombination sites recognised by the integron DNA integrase. Nucleic Acids Res 1994, 22 (11), 2071-
546 2078. DOI: 10.1093/nar/22.11.2071.

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

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

553 (39) Pfeifer, E.; Bonnin, R. A.; Rocha, E. P. C. Phage-plasmids spread antibiotic resistance genes
554 through infection and lysogenic conversion. mBio 2022, 13 (5), e0185122. DOI: 10.1128/mbio.01851-
555 22.

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

558 (41) Dawes, F. E.; Kuzevski, A.; Bettelheim, K. A.; Hornitzky, M. A.; Djordjevic, S. P.; Walker, M. J.
559 Distribution of class 1 integrons with IS*26*-mediated deletions in their 3'-conserved segments in
560 *Escherichia coli* of human and animal origin. PLoS One 2010, 5 (9), e12754. DOI:
561 10.1371/journal.pone.0012754.

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

565 (43) Rosewarne, C. P.; Pettigrove, V.; Stokes, H. W.; Parsons, Y. M. Class 1 integrons in benthic
566 bacterial communities: Abundance, association with Tn*402*-like transposition modules and evidence
567 for coselection with heavy-metal resistance. FEMS Microbiol Ecol 2010, 72 (1), 35-46. DOI:
568 10.1111/j.1574-6941.2009.00823.x.

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

572 (45) De Coster, W.; D'Hert, S.; Schultz, D. T.; Cruts, M.; Van Broeckhoven, C. NanoPack: Visualizing
573 and processing long-read sequencing data. Bioinformatics 2018, 34 (15), 2666-2669. DOI:
574 10.1093/bioinformatics/bty149.

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

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

579 (48) Sahlin, K.; Medvedev, P. Error correction enables use of Oxford Nanopore technology for
580 reference-free transcriptome analysis. Nat Commun 2021, 12 (1), 2. DOI: 10.1038/s41467-020-20340-
581 8.

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

584 (50) Bengtsson-Palme, J.; Hartmann, M.; Eriksson, K. M.; Pal, C.; Thorell, K.; Larsson, D. G.; Nilsson,
585 R. H. METAXA2: Improved identification and taxonomic classification of small and large subunit
586 rRNA in metagenomic data. Mol Ecol Resour 2015, 15 (6), 1403-1414. DOI: 10.1111/1755-
587 0998.12399.

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

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

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

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

599 (55) Cury, J.; Jove, T.; Touchon, M.; Neron, B.; Rocha, E. P. Identification and analysis of integrons
600 and cassette arrays in bacterial genomes. Nucleic Acids Res 2016, 44 (10), 4539-4550. DOI:
601 10.1093/nar/gkw319.

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

605 (57) Zhang, Y.; Yang, X.; Huang, L.; Zhang, R. Whole-genome analysis and description of an IMP-8-
606 producing *Ochrobactrum anthropi*. J Glob Antimicrob Resist 2022, 29, 275-277. DOI:
607 10.1016/j.jgar.2022.03.016.

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

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

614 (60) Martinez-Freijo, P.; Fluit, A. C.; Schmitz, F. J.; Grek, V. S.; Verhoef, J.; Jones, M. E. Class I
615 integrons in Gram-negative isolates from different European hospitals and association with decreased
616 susceptibility to multiple antibiotic compounds. J Antimicrob Chemother 1998, 42 (6), 689-696. DOI:
617 10.1093/jac/42.6.689.

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

621 (62) Dunwell, J. M.; Culham, A.; Carter, C. E.; Sosa-Aguirre, C. R.; Goodenough, P. W. Evolution of
622 functional diversity in the cupin superfamily. Trends Biochem Sci 2001, 26 (12), 740-746. DOI:
623 10.1016/s0968-0004(01)01981-8.

624 (63) Krucinska, J.; Lombardo, M. N.; Erlandsen, H.; Estrada, A.; Si, D.; Viswanathan, K.; Wright, D.
625 L. Structure-guided functional studies of plasmid-encoded dihydrofolate reductases reveal a common

626 mechanism of trimethoprim resistance in Gram-negative pathogens. Commun Biol 2022, 5 (1), 459.
627 DOI: 10.1038/s42003-022-03384-y.

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

631 (65) Evans, B. A.; Amyes, S. G. OXA β-lactamases. Clin Microbiol Rev 2014, 27 (2), 241-263. DOI:
632 10.1128/CMR.00117-13.

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

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

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

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

646 (70) Binh, C. T.; Heuer, H.; Kaupenjohann, M.; Smalla, K. Diverse *aadA* gene cassettes on class 1
647 integrons introduced into soil *via* spread manure. Res Microbiol 2009, 160 (6), 427-433. DOI:
648 10.1016/j.resmic.2009.06.005.

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

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

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

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

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

665 (76) Linh, D. V.; Gibu, N.; Tabata, M.; Imai, S.; Hosoyama, A.; Yamazoe, A.; Kasai, D.; Fukuda, M.
666 Complete genome sequence of natural rubber-degrading, Gram-negative bacterium, *Rhizobacter*
667 *gummiphilus* strain NS21^{T}. Biotechnol Rep (Amst) 2019, 22, e00332. DOI:
668 10.1016/j.btre.2019.e00332.

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

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

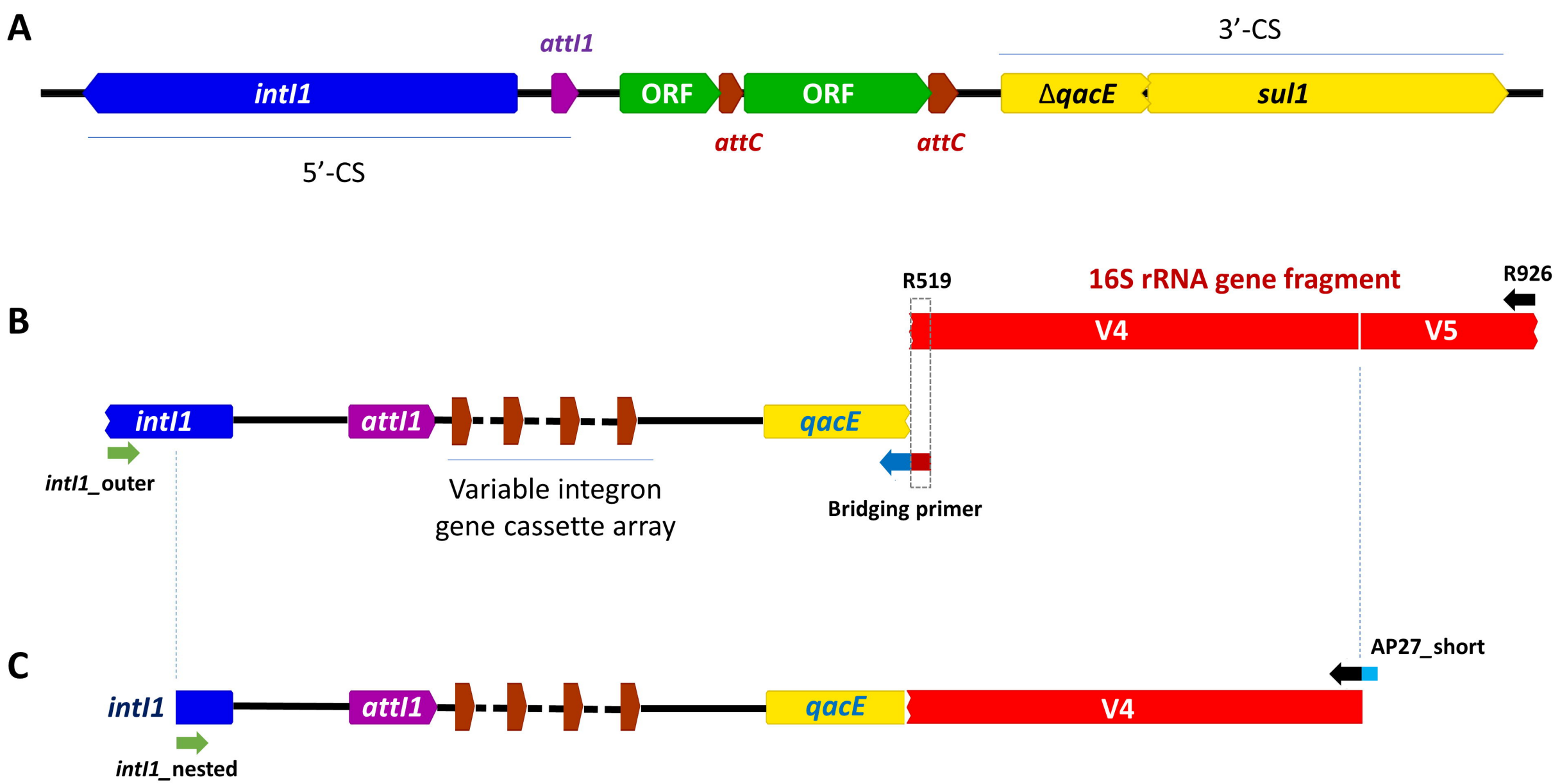
674 (79) Tanaka, T.; Kawasaki, K.; Daimon, S.; Kitagawa, W.; Yamamoto, K.; Tamaki, H.; Tanaka, M.;
675 Nakatsu, C. H.; Kamagata, Y. A hidden pitfall in the preparation of agar media undermines
676 microorganism cultivability. Appl Environ Microbiol 2014, 80 (24), 7659-7666. DOI:
677 10.1128/AEM.02741-14.

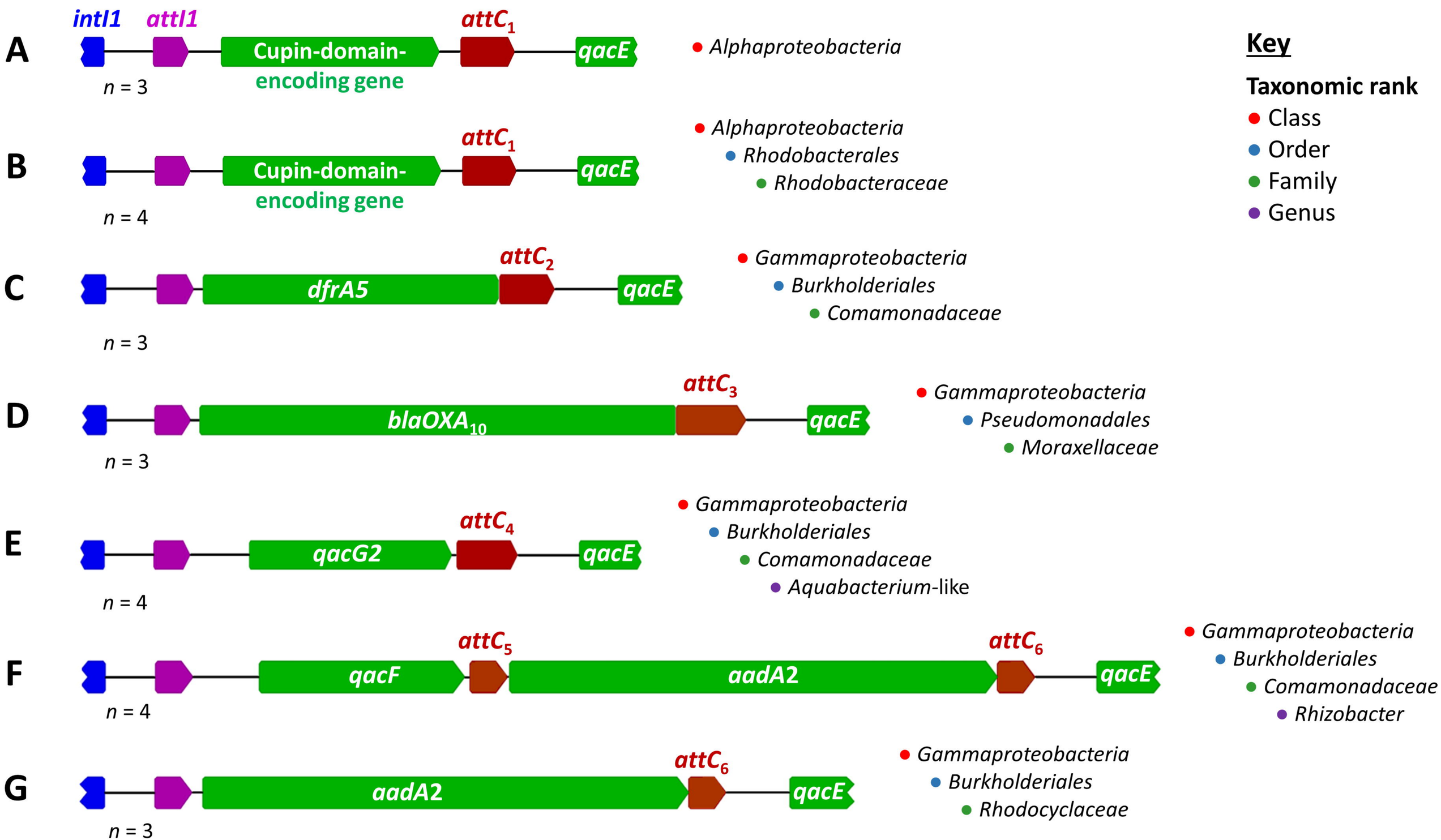
678 (80) Bengtsson-Palme, J.; Jonsson, V.; Hess, S. What is the role of the environment in the emergence
679 of novel antibiotic resistance genes? A modeling approach. Environ Sci Technol 2021, 55 (23), 15734-
680 15743. DOI: 10.1021/acs.est.1c02977.

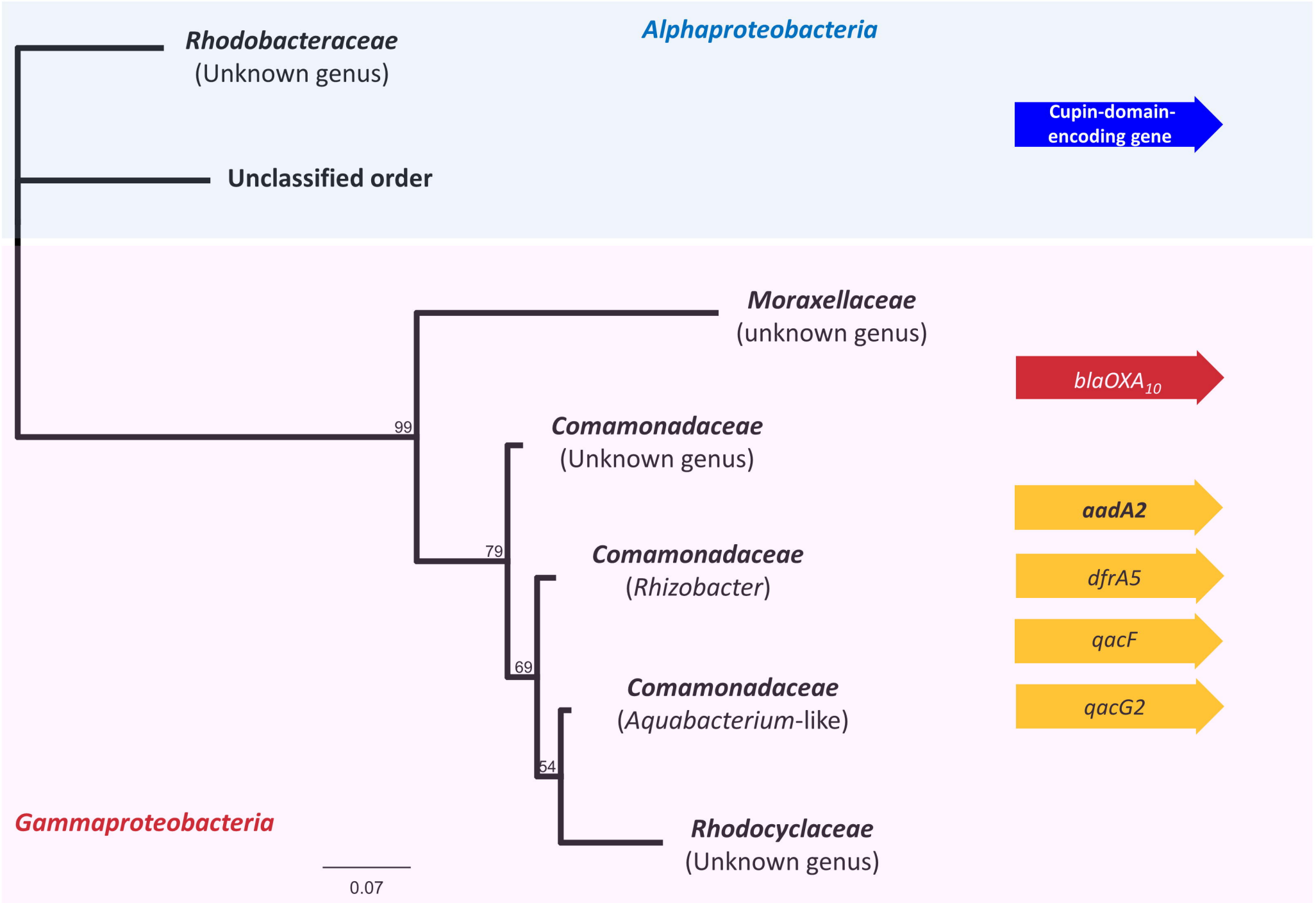
681 (81) Kurland, C. G.; Canback, B.; Berg, O. G. Horizontal gene transfer: A critical view. Proc Natl Acad
682 Sci U S A 2003, 100 (17), 9658-9662. DOI: 10.1073/pnas.1632870100.

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

685


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