Atypical integrative element with strand-biased circularization activity assists interspecies antimicrobial resistance gene transfer from *Vibrio alfacsensis*

Lisa Nonaka,[1,2*] Michiaki Masuda,[1] Hirokazu Yano[3,4*]

[1]Department of Microbiology, Dokkyo Medical University School of Medicine, Mibu, Tochigi, Japan

[2]Faculty of Human Life Sciences, Shokei University, Kumamoto, Japan

[3]Faculty of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Japan

[4]Graduate School of Life Sciences, Tohoku University, Sendai, Japan

Short title: Integrative elements with strand-biased circularization activity

*Corresponding authors

Email: nonaka20@shokei-gakuen.ac.jp (LN); yano.hirokazu@ige.tohoku.ac.jp (HY)

Abstract

The exchange of antimicrobial resistance (AMR) genes between aquaculture and terrestrial microbial populations has emerged as a serious public health concern. However, the nature of the mobile genetic elements in marine bacteria is poorly documented. To gain insight into the genetic mechanisms underlying AMR gene transfer from marine bacteria, we mated a multidrug-resistant *Vibrio alfacsensis* strain with an *Escherichia coli* strain, and then determined the complete genome sequences of the donor and transconjugant strains. Sequence analysis revealed a multidrug resistance conjugative plasmid in the donor strain, which was integrated into the chromosome of the recipient. The plasmid backbone in the transconjugant chromosome was flanked by two copies of a 7.1 kb unclassifiable integrative element harboring a β-lactamase gene. The 7.1 kb element (named SE-6945) and previously reported Tn*6283* share four coding sequences, two of which encode the catalytic R-H-R-Y motif of tyrosine recombinases. Polymerase chain reaction and sequencing experiments revealed that SE-6945 generates a circular copy of its one specific strand without generating an empty donor site, in contrast to the movement of integron gene cassette or ICE/IMEs discovered to date. Copy numbers and location of SE-6945 in the chromosome affected antibiotic resistance levels of the transconjugants. SE-6945/Tn*6283*-related integrative elements are collectively termed SEs (strand-biased circularizing integrative elements). SEs were identified in the genomes of other *Vibrio* species. Overall, these results suggest that in addition to conjugative plasmids, SEs are involved in the spread of AMR genes among marine bacteria.

Introduction

Antimicrobials have been used globally in aquaculture to control fish diseases. Although this approach helps to maintain a stable supply of aquacultural products, misuse of antimicrobials has led to the emergence of antimicrobial-resistant microbes and the accumulation of resistance genes at aquaculture sites [1,2]. As the spread of multidrug-resistant (MDR) bacteria and pan-drug-resistant bacteria threatens human life [3,4], it is important to obtain clues about whether and how aquatic and terrestrial microbial populations exchange genetic materials.

Plasmids [5] and integrative and conjugative elements (ICE) [6] are DNA units that can move from one cell to another through the conjugation machinery. They can directly move between species [7,8], and indirectly mobilize genes on specific mobile elements [9–11]. Plasmids can be classified into families according to replicon type or mobilization machinery type for epidemiological purposes [12,13].

Transposons have been defined as specific DNA segments that can repeatedly insert into one or more sites in one or more genomes [14]. Several distinct strand exchange mechanisms and the routes used for their location changes, i.e., *transposition*, have been discovered to date [15]. In the comprehensive review by Curcio and Derbyshire [15], transposons were classified into five families based on the DNA strand exchange enzymes used. In prokaryotes, the mode of transposition was largely classified into 'cut-out paste-in' (often called 'excision integration' for the ICEs [6]), 'copy-out paste-in,' or 'copy-in' for which replication is necessary to complete transposition; e.g., Tn*3*, IS*91*. DNA insertion involving DNA replication can also be seen in the insertion of a folded single-strand DNA of Integron gene cassette into *attI* [16] and of phage CTX genome into *dif* [17]; both use tyrosine

63 recombinase(s) as the strand exchange enzyme. In transposon classification, ICEs are
64 classified as transposons that use Y-transposase (tyrosine recombinase) or S-transposase
65 (serine recombinase) [15]. Other than ICEs, DNA segments that can move to a location in
66 other cells via the excision – transfer (rolling-circle replication) – integration route have
67 been termed IME [18], MGI [10], CIME [19], or MTn [20] (hereafter collectively referred to
68 as IME). IMEs encode site-specific recombinase or DDE transposase as strand exchange
69 enzyme [21]. They have been recognized as degraded forms of ICEs [19]. ICEs/IMEs so far
70 studied in the laboratory seem to change their location via the 'excision integration' route
71 only [22–25]. Genes embedded in these mobile genetic elements are readily shared among
72 bacteria and provide genetic resources for microbial adaptation in changing environments.
73 To better understand the genetic mechanisms underlying the spread of antibiotic
74 resistance genes at aquaculture sites, we previously collected resistant bacteria from
75 sediments at an aquaculture site in Kagawa, Japan [26–28]. We identified a self-
76 transmissible MDR plasmid of the MOB_H-family, named pAQU1, that can replicate in both
77 the original host *Photobacterium damselae* subsp. *damselae* and in *Escherichia coli* [27].
78 Subsequently, another MOB_H-family MDR plasmid, named pSEA1, was identified in a *Vibrio*
79 *alfacsensis* isolate (previously identified as *V. ponticus*) [29]. pSEA1 carries a 12 kb
80 unclassifiable integrative element, Tn*6283* (renamed as SE-6283 in this study), in addition to
81 the antimicrobial resistance (AMR) genes. Although pSEA1 could not replicate in *E. coli* at
82 42 °C, following conjugation, it could integrate into the *E. coli* chromosome by homologous
83 recombination between two SE-6283 copies: one on pSEA1 and another that moved from
84 pSEA1 into the chromosome [29]. Further, SE-6283 was suggested to move via a circular
85 copy intermediate of one specific strand without generating an empty donor site,

presumably using tyrosine recombinases. However, it is not known whether interspecies gene transfer assisted by SE-6283-like elements is common in nature.

To obtain further insights into the mechanisms behind genetic exchange among aquaculture-associated bacteria, in this study we mated another MDR *Vibrio* isolate, *V. alfacsensis* 04Ya249, with *E. coli* strains in the laboratory and then determined the genome sequences of both the donor and transconjugant strains. We identified a SE-6283-like element involved in the integration of the multidrug resistance plasmid into *E. coli* chromosome. Based on the findings in this study, we have termed the currently unclassifiable integrative element, SE-6283, and its related elements as a new class of mobile elements: strand-biased circularizing integrative element (SE).

In this study, several types of intracellular DNA movements need to be discussed. For clarity, terms *transposition*, *integration*, and *insertion* are used here based on the following rules. We call the movement of a DNA segment from one location to the other location in one or two genomes, mediated by any DNA strand exchange enzymes, as transposition. The physical incorporation of a plasmid into a chromosome is called integration regardless of the molecular mechanisms and genomic locations involved. The physical incorporation of a DNA molecule into a specific site in another molecule mediated by a site-specific recombinase (so-called integrase) has traditionally been called integration, but we refer to it as an insertion to distinguish this event from the plasmid integration defined above. The term translocation has frequently been used when referring to rarely occurring location change of a large DNA segment between chromosomes in organisms carrying multiple chromosomes. We did not use translocation in this study as it is not relevant to this study.

Results

Identification of SE carrying a β-lactamase gene

We previously isolated *V. alfacsensis* strain 04Ya249 from sea sediment at an aquaculture site [26]. 04Ya249 showed resistance to erythromycin, tetracycline, and ampicillin [30]. To identify active mobile elements in this strain that are relevant to its AMR, the strain was mated with macrolide-sensitive *E. coli* JW0452, and the transconjugant was first selected in the presence of erythromycin at 42 °C. One transconjugant was named strain TJ249.

The complete genome sequences of strains 04Ya249 and TJ249 were determined using the PacBio RS II platform. The genome of strain 04Ya249 consists of four replicons: two chromosomes, a 306 kb putative conjugative plasmid pSEA2, and an 82 kb putative conjugative plasmid pVA249. Strain 04Ya249 has a very similar genome architecture to the database strain *V. alfacsensis* CAIM 1831 based on gene synteny (S1 Fig A) and average nucleotide identity (S1 Fig B). Seven AMR genes were identified on plasmid pSEA2 (S2 Fig A). One β-lactamase gene (*CARB-19* allele in the CARD database [31]) was located within the 7.1-kb repeat region found in both chromosome 1 and plasmid pSEA2 (S2 Fig B), which contained four protein-coding sequences in addition to a β-lactamase gene. These four coding sequences showed homology to the four coding sequences clustered at one end of SE-6283 from strain 04Ya108 (Fig 1A), although gene product identities between SE-6945 and SE-6283 counterparts were very low: 22.5 % for IntA, 20.5 % for CDS2, 23.2 % for IntB, and 26.6 % for CDS4 in blastp.

This 7.1-kb region was confirmed to be active in generating its circular copy of top strand in *E. coli* and *V. alfacsensis* and inserting it into new target locations in *E. coli* in this study. The 7.1kb region was designated as a new member of SE, SE-6945. The four

conserved coding sequences (SE core genes) were tentatively called *intA*, CDS2, *intB*, and CDS4. The biochemical functions of their products were estimated using Pfam and InterProScan in EMBL. The *intA* and *intB* products generated a hit on phage_integrase (CL0382) in Pfam and were further aligned with known tyrosine recombinases using structure-aware alignment program PROMAL3D [32] to confirm the presence of conserved motifs. The *intA* and *intB* products contained the catalytic R-H-R-Y motif (G-R motif in box I and H-R-H-Y motif in box II (Fig 1B)) conserved in tyrosine recombinases [33], though the N-terminal region of IntB is extraordinarily long among tyrosine recombinases. CDS2 and CDS4 products did not generate hits in the Pfam or InterProScan databases. The SEs carry imperfect 19 bp inverted repeat motifs termed C and C′ at two ends (Fig. 1 A), which might be equivalent to core-type sites [34] used in strand change mediated by tyrosine recombinase(s).

Fig 1. Genetic features of SEs. (A) Upper panel: Genetic organization of SE-6283 and SE-6945. Lower panel: Nucleotide sequences of imperfect inverted repeat motifs C and C′. The nicking sites used for circularization were six bases upstream of the 5′-end of motif C and the 3′-end of motif C′ (black arrowheads in the upper panel). No strand exchange was detected on the bottom strand (white arrowheads). (B) Presence of R-H-R-Y motif in IntA and IntB. IntA and IntB sequences are aligned with sequences of known tyrosine recombinases based on secondary structure prediction using PROMAL 3D [32]. The aligned regions corresponding to box I and box II in [33] are shown. (C) Schematic representation of SE-6945 and pSEA2 insertion locations in *E. coli* TJ249.

155 As pSEA2 showed > 99.9 % nucleotide identity on the aligned region to pSEA1
156 (accession no. LC081338.1) from strain 04Ya108 [29], we also determined the complete
157 genome of strain 04Ya108 for comparison with strain 04Ya249. As predicted from Southern
158 hybridization [29], strain 04Ya108 possessed two copies of SE-6283, one in chromosome 1
159 and the other in plasmid pSEA1. Strain 04Ya108 also carries SE-6945 in the same two
160 locations as strain 04Ya249 (S1 Fig A, S2 Fig B).

161 The chromosome of *E. coli* TJ249 contains two notable insertions (Fig 1C), a smaller
162 insertion of SE-6945 alone immediately downstream of the putative transcription
163 termination site for *yjjN* (*yjjNt*) located between *yjjN* and *yjjO* (in GenBank accession no.
164 U00096.3), and a larger insertion into *insJ*, containing a nearly complete copy of pSEA2 and
165 an additional copy of SE-6945. The 14.2 kb *mphA-cat* gene region flanked by repeats in
166 pSEA2 (contig 1 in S2 Fig A) was detected as a separate circular contig in the TJ249 genome
167 assembly. These indicate the movement between different genomic locations (pSEA2, *yjjNt*,
168 *InsJ*), thereby transposition, of SE-6945 in *E. coli*.

169 The border regions between SE-6945 and its target locations were termed *attL* (*intA*-
170 proximal side) and *attR* (*intA*-distal side) (Fig 2 A). The nucleotide sequence alignments of
171 *attL* and *attR* from two *E. coli* chromosomal locations and two locations in 04Ya249
172 genomes (pSEA2 and chr 1) are shown in S3 Fig and S4 Fig, respectively. Sequence
173 comparison of the *attL*, *attR*, and unoccupied target site in the host (*attB*) in *E. coli* (Fig 2B)
174 revealed that the 6 bp next to the 3′ end of C' in *attR* in *E. coli* (5'-TTTTCT-3') was identical
175 to the 6 bp next to the 5′ end of C at *attL* on pSEA2, while the 6 bp next to the 5′ end of C
176 in *attL* (5′-TTCTTT-3′ in *insJ*, 5′-TTTTTT-3′ in *yjjNt*) was from *attB* (S3 Fig, S4 Fig, Fig 2B).
177 Therefore, SE-6945 inserts itself, ending with 5′-GTA-3′ (Fig 1 A) along with an additional 6

178 bp next to the 5′ end of C at *attL* on the donor molecule into the target site. The 6 bp from
179 the donor molecule is placed next to the 3′ end of C′ in the newly formed *attR*.
180

181 **Target site preference of SE-6945**

182 After the isolation and genome sequence determination of TJ249, additional mating
183 experiments between *E. coli* JW0452 and *V. alfacsensis* 04Ya249 were performed to test
184 whether pSEA2 integration into a specific location of *E. coli* chromosome accompanying SE-
185 6945 transposition is reproducible. Independent mating was performed more than 17 times,
186 and transconjugant was screened in the presence of Tc alone, Ap alone, or both Ap and Tc.
187 SE-6945 insertion pattern was investigated by Southern hybridization for 16 randomly
188 chosen Tc resistant transconjugants, and one Tc sensitive but Ap-resistant transconjugant
189 obtained from Ap alone selection (S5 Fig). Among 16 Tc resistant transconjugant strains, 12
190 strains generated *intA*-containing fragments originating from the *yjjNt*-SE-6945 junction or
191 the *insJ*-SE-6945 junction. These 16 strains also generated pSEA2-derived fragments (S5 Fig).
192 The remaining four strains generated pSEA2-derived fragments and contained only one
193 copy of *intA*, suggesting SE-6945-independent pSEA2 integration into chromosome;
194 autonomous replication of pSEA2 seems unlikely given the high frequency of pSEA2
195 integration into the chromosome in other transconjugant strains. Strain LN52, obtained
196 from Ap alone selection, carried only one SE-6945 copy, and it was inserted into *yjjNt* (S5
197 Fig). Taken together, these observations suggest that pSEA2 integration into *E. coli*
198 chromosome generally accompanies the SE-6945 transposition, and that SE-6945 prefers
199 *yjjNt* or *insJ*, but not random sites in the *E. coli* genome.

200 To obtain further insights into the target site of SE-6945, bacterial genomes carrying
201 SE-6945 were searched based on blastn in NCBI nucleotide collection (nr/nt) database using
202 the SE-6945 core gene region (*intA*-CDS2-*intB*-CDS4) as the query. The genome of *Vibrio*
203 *harveyi* strain WXL538 contained a segment showing 88.9% nucleotide sequence identity to
204 the SE-6945 core genes. Genome structure comparisons between WXL538 and 04Ya249
205 revealed that WXL538 carries a 16.4 kb SE-6945-like SE (tentatively termed SE-VhaWXL538)
206 containing a different set of cargo genes (Fig 2 A). SE-VhaWXL538 is inserted into the
207 promoter region of a *yaaA* gene encoding an H2O2_YaaD_Superfamily protein (NCBI
208 conserved domain accession no. cl01143) (locus_tag F9277_14400 in CP045070.1). This
209 location corresponds to the SE-6945 insertion location in *V. alfacsensis* chromosome 1 (Fig 2
210 A). Sequences of unoccupied target sites (*attB*) in *E. coli*, *V. alfacsensis*, and *V. harveyi* are
211 shown in Fig 2B. The conservation level of nucleotide in *attB* is presented in Fig 2C as
212 Sequence logos information contents [35]. The central 6 bp putative crossover region in *attB*
213 was T rich. However, other motifs, such as imperfect inverted repeat structure observed in
214 *attC/attI* of integron (IntI target) [36,37], or *dif* of CTX phage integration site (XerC/XerD
215 target) [17], could not be detected in *attB, attL* (S3 Fig), or *attR* (S4 Fig) for SE-6945 in the
216 current data set.
217
218 **Fig 2. Locations of SE-6945 and its related elements.** (A) Comparisons of genetic
219 organization around SE insertion locations. GenBank accession numbers (*Vibrio* strains only)
220 and the coordinate of genomic segments compared are shown below the strain name. (B)
221 Alignment of *attB* sequences. The sequences are retrieved from SE-free genomes. Since the
222 SE-6945-free pSEA2 ancestor could not be found in the database, the ancestral *attB*
223 sequence on pSEA2 in the putative crossover region shown as x is unpredictable. Nicking is

224 expected to take place next to the central box: one site on the left side on the top strand.
225 (C) Sequence logos representation of conservation levels of *attB* sequence.
226

227 Transposition of SE-6945 and plasmid integration in mating

228 The direct repeat structure of SE-6945 in the TJ249 chromosome suggested that
229 pSEA2 was integrated into the recipient chromosome through the transposition of SE-6945,
230 with subsequent integration of pSEA2 into the chromosome by homologous recombination
231 between SE-6945 copies: one on the chromosome and the other on pSEA2 (referred to as a
232 two-step gene transfer mechanism in the previous study [29]). If this scenario holds,
233 insertion of SE-6945 alone should occur more frequently than pSEA2 integration upon
234 mating, and pSEA2 integration would be reduced in frequency or abolished altogether in the
235 chromosome of a *recA*-null mutant recipient. To test these possibilities, we conducted
236 additional mating assays.
237 Quantitative mating assays were performed using a rifampicin-resistant recipient
238 strain, JW0452rif, and its isogenic strain carrying *recA* deletion JW0452Δ*recA*rif. The transfer
239 of Tc Ap resistance from *V. alfacsensis* 04Ya249 to *E. coli* JW0452rif was detected only in 2
240 out of 4 replicated mating experiments, and the transfer frequency was determined to be
241 approximately 10^{-9} per donor, indicating a very rare event, generating 0–3 transconjugant
242 colonies per mating (Fig 3). The transfer of Ap resistance alone was observed at >300-fold
243 higher frequency than the transfer of Tc Ap-resistant transconjugants (compare the first and
244 the second groups from left in Fig 3), indicating that insertion of SE-6945 alone occurs more
245 frequently than pSEA2 integration during mating.

The transfer of Tc Ap resistance to JW0452Δ*recA*rif could not be detected in all four independent mating assays (below detection limit of 1.02×10^{-9} per donor). Unexpectedly transfer frequency of Ap resistance was also reduced to the 1/17 level of JW0452rif in JW0452Δ*recA*rif (compare the second and the fourth groups from the left in Fig 3). Thus, the direct effect of *recA* knockout on pSEA2 integration could not be evaluated due to the reduced detection sensitivity of gene transfer. However, when LN52rif (which already carries one copy of SE-6945 and wild type *recA*) was used as the recipient, Tc Ap-resistant transconjugants were obtained at a 56-fold higher frequency (in mean) than that observed when using JW0452rif as the recipient (compare the first and fifth groups in Fig 3). This indicates that the presence of the 7.1 kb long homology stretches in both target DNA and pSEA2 promotes pSEA2 integration in the recipient.

The Tn*3*-like cointegrate formation [15] might happen if SE-6945 could simultaneously introduce nick at one site on the top strand at *attL* (or *attR*) and one site on the bottom strand at *attR* (or *attL*), and DNA-protein crosslink (DPC) repair [38] could clean the linked tyrosine recombinase to produce a 3′-OH priming site. If SE-mediated cointegrate formation occurs, then Tc Ap-resistant transconjugant should emerge at a comparable frequency to Ap-resistant transconjugant, which was not the case in our observation.

Collectively, these results and the nature of tyrosine recombinases support the model that pSEA2 integration mainly occurs through a two-step gene transfer mechanism; transposition followed by homologous recombination than through a one-step Tn*3*-like cointegrate formation process.

269 **Fig 3. Resistance gene transfer from *V. alfacsensis*.** Transfer frequency (Y-axis) is the $\log_{10}$-
270 transformed value of the transconjugant colony-forming units (CFU) divided by the donor
271 CFU. Four replicate mating experiments were performed. No transconjugant was detected
272 in any of the four replicate experiments using JW0452Δ*recA*rif with Tc Ap selection or in the
273 two replicate experiments using JW0452rif with Tc Ap selection (indicated by open circles
274 placed at the bottom). In one experiment on LN52rif, the donor CFU was not evaluated, but
275 the transconjugant CFU was obtained at a comparable frequency to that observed in the
276 other three replicate experiments. The detection limit of this experiment was around $1.02 \times$
277 10^{-9} per donor.

278

279 **Strand-biased circularization of SE-6945**

280 We previously [29] revealed the following unique features of SE-6283: (i) it does not
281 generate an unoccupied donor site upon its circularization *in vivo*, and (ii) the circular form
282 of SE-6283 is generated using only one strand as a template at least in *E. col*. We note that
283 this assertion is based on the notion that tyrosine recombinases have a nature to not
284 produce blunt ends at *att* sites on double-strand DNA during strand exchange. Although SE-
285 6283 encodes tyrosine recombinases and not a single DDE transposase, the outcomes
286 described above, which result only from a figure-eight structure containing a single-strand
287 DNA bridge (Path 1 in Fig 4A), is reminiscent of the first step of the 'copy-out paste-in'
288 transposition of ISs [39] (IS*3* route), rather than the excision of ICEs (ICE/IME route) or the
289 single-strand circle generation of Integron gene cassette by *attC* × *attC* recombination (gene
290 cassette route), both of which generate empty donor sites [24,40,41]. If the figure-eight
291 structure containing a single-strand DNA bridge is subject to replication without DPC repair,

292 the products would be a single-strand SE circle, the original copy of donor molecule, and the
293 unjoined chromosome, which is eventually lost (Fig 4A bottom left). To test whether circular
294 copy production of SE-6945 also follows Path 1, we created a pSEA2-free 04Ya249 derivative
295 strain LN95 carrying only a single copy of SE-6945 in the chromosome and analyzed the
296 production of its circular molecule in this strain and the sequence of strand exchange
297 products.

298 Pairs of primers were designed such that one primer anneals to the inside of the
299 integrative element and the other anneals to the outside of the element (Fig 4A), amplifying
300 *attL* (*intA* side, Fig 4B, product size 304 bp), *attR* (*bla* side, 248 bp), empty donor site (*attB,*
301 345 bp), occupied *attB* (7.4 kb), and *attS* (joint of motif C and C' formed on the circular SE-
302 6945, 206 bp). Primer set 1-4 in Fig 4 A were used to detect both empty *attB* and occupied
303 *attB*.

304 The circular form of SE-6945 was detected in LN95 (Fig 4B) as well as the parent
305 strain 04Ya249 (S6 Fig A). This suggests that the *cis*-elements (*attL*/*attR*) on chromosome 1
306 and strand exchange enzymes encoded by the chromosomal copy of SE-6945 are functional
307 in *V. alfacsensis*. The empty *attB* (345 bp) could not be detected under PCR conditions with
308 short extension times (Fig 4 B(i)), while a larger product of occupied *attB* (7.4 kb) was
309 detected when the Long-PCR condition was used (Fig 4 B(ii)).

310

311 **Fig 4. PCR and sequencing analyses of SE-6945 circle.** (A) Design of PCR experiment to
312 detect SE-6945 circle. Upper panel: primer annealing locations (triangles in gray) on SE-6945
313 and the *attB* region. The number corresponds to the primer set numbers in B. Primer
314 number indicates the following primers in Table 2: 1, LN_181_Valc_chr1_attB_L; 2,
315 LN_142_junction2; 3, LN_143_junction2; 4, LN183_Valc_chr1_attB_R2. Lower panel:

316 possible routes of SE movement. Two strands of DNA are shown in separate lines. Lines in
317 orange are SEs. In Path 1 (solid arrows), one strand is nicked, and the SE side is circularized
318 while the host DNA side of the strand remains unjoined, leading to a figure-eight structure
319 containing a single-strand DNA bridge. The figure-eight form can be subject to replication,
320 giving rise to a double-strand SE circle (if the recombinase bound 3′ end is cleaned, for
321 example, by DPC repair [38]) (copy-out route), or single-strand DNA circle (if the 3′ end is
322 not cleaned). In Path 2, strand exchange generates a figure-eight containing a double-strand
323 DNA bridge. After going through Path 2, there are two possible outcomes: (i) another round
324 of strand exchange at bottom strand leading to circular a double-strand SE as well as a
325 circular donor molecule carrying empty donor site (ICE/IME route), (ii) replication on the
326 figure-eight leading to single-strand SE circle and donor molecule with empty *attB*, and the
327 original donor molecule (gene cassette route). (B) The result of PCR. Numbers in the primer
328 set correspond to primer numbers in panel A. PCR products were electrophoresed in 2%
329 agarose gel (i) or 0.8% agarose gel (ii). (C) Putative nicking sites in *attL* and *attR*. The SE DNA
330 is shown in orange. Nicking and ligation at positions shown in arrowheads in the same color
331 can generate *attS* with a 6 bp spacer. (D) Observed *attS* sequences. *attS* type 1 can be
332 generated from strand exchange at the top strand on chromosome 1 or bottom strand on
333 pSEA2, while *attS* type 1 can be generated from strand exchange at the top strand on pSEA2
334 or bottom strand on chromosome 1.
335
336 To estimate the copy number of the circular form of SE-6945 in the cell population,
337 we first searched for Illumina reads of strain LN95 that spanned *attS* or the "hypothetical"
338 *attB*. Although next-generation sequencing reads were obtained at 229× chromosome
339 coverage, no reads spanning *attS* or *attB* were detected [42]. Quantitative PCR further

340 revealed that the mean *attS* to *gyrB* ratio was 0.0012 (S6 Fig B). This relative copy number
341 was consistent with the result from a previous study detecting the circular form of SE-6283
342 in *E. coli*, which showed an *attS* to chromosome ratio of 0.001 [29].

343 To determine the nicking site on *attL* and *attR* (Fig 4D), we cloned the PCR products
344 of *attS* and then sequenced ten cloned molecules. Based on the *attL*, *attR*, and the original
345 *attB* sequences in the TJ249 chromosome (S3 Fig, S4 Fig), we hypothesized four nicking sites
346 (arrowheads in Fig 4D) in *attL* and *attR* on pSEA2 and chromosome 1 in strain 04Ya249. The
347 cloned *attS* PCR products of 04Ya249 contained a 6 bp spacer sequence between C and C'
348 (Fig 4E). This observation supports the hypothesis that nicking occurs 6 bp upstream of the
349 5′ end of C side and 3′ end of C' side and/or 6 bp upstream of the 5′ end of C' side and 3′
350 end of C side. Two sequences, 5′-TTTTTT-3′ (*attS* type 1) and 5′-TTTTCT-3′ (*attS* type 2)
351 were detected at a 1:9 ratio in the spacer region between C and C' in strain 04Ya249
352 (examples of the Sanger sequencing trace files have been posted in figshare [42]). These
353 two types of spacer sequences can arise from both pSEA2 and chromosome 1 (Fig 4D).
354 However, only one sequence (*attS* type 1) was detected in the pSEA2-free strain LN95 (Fig
355 4E), indicating that strand exchange occurs only on the top strand in *attL* × *attR*
356 recombination on chromosome 1 (Fig 4C). The occurrence of *attS* type 1 and *attS* type 2 in
357 04Ya249 (Fig 4E) agrees with the top strand exchange in *attL* × *attR* recombinations on
358 chromosome 1 and pSEA2, respectively (Fig 4D). Collectively, these observations are
359 consistent with the model that SE-6945 follows Path 1 (Fig 4 A) without undergoing strand
360 exchange of the bottom strand in *V. alfacsensis*.

361 Nicking at one strand on the double-strand DNA and its circularization is often seen in
362 the reactions mediated by HUH endonucleases (relaxase/rolling-circle replication initiator)
363 [43]. To confirm that observed nicking at *attL* and *attR* is independent of HUH

364 endonucleases produced from plasmids or other mobile elements in *V. alfaccensis*, the
365 production of *attS* was additionally investigated using *E. coli* strain LN52 that carries a single
366 copy of SE-6945 but lacks pSEA2. *attS* was produced even in LN52 (S6 Fig C). Therefore,
367 strand-biased (top strand only) *attL* × *attR* recombination is likely mediated by DNA strand
368 exchange enzymes encoded by SE-6945.
369

370 The insertion copy number and location of SE-6945 affects β-lactam
371 resistance levels of the host cell

372 The naturally occurring strains 04Ya249 and 04Ya108 carry two copies of SE-6945 in
373 genomes. Furthermore, 04Ya108 also carries two copies of SE-6283. However, it remains
374 unclear whether increased SE copies in a genome confer an advantage to the host cell.
375 Three copies of SE-6945 in the transconjugant TJ249 were initially detected by genome
376 sequencing. Southern hybridization analysis of other JW0452 transconjugants revealed one
377 to two copies of SE-6945 in their genomes (S5 Fig). Location of SE-6945 insertion became
378 evident based on Southern hybridization for 15 among 19 transconjugants. Therefore, we
379 used these transconjugant strains to test whether multiple insertions of SE-6945 or its
380 insertion location can increase the antibiotic resistance phenotype of the transconjugants.
381 Sixteen *E. coli* strains with or without SE-6945 were grown in Luria-Bertani broth
382 with ampicillin, and then the minimum inhibitory concentrations (MICs) of ampicillin were
383 determined using the broth dilution method following the Clinical Laboratory Standards
384 Institute (CLSI) guidelines. The MIC was highest in the clone carrying three copies of SE-6945,
385 followed by the group of transconjugants carrying two copies and the group carrying one
386 copy, and the MIC was the lowest in JW0452 (Fig 5). Although within-group variation was

387 also observed, the increased copy number of SE-6945 in the genome generally resulted in
388 the increased β-lactam resistance of the host cell.
389 MIC of individual strain is shown in Table S1. When focused on a class of MIC 500,
390 there was a trend that strains with SE-6945 insertion into *insJ* show a higher MIC than
391 strains with insertion into *yjjNt*. Thus, the location of SEs can also affect phenotype variation,
392 possibly due to the differences in the expression levels of their cargo genes.
393
394 **Fig 5. The copy number of SE-6945 affects the resistance level.** Sixteen transconjugants and
395 a control strain (JW0452) were divided into four groups shown in distinct colors based on
396 the SE-6945 copy number in the genome. MIC of ampicillin is shown according to copy
397 number.
398

399 Target site preference of SE-6283

400 We previously reported SE-6283 insertion into the coding region of the *bcp* gene encoding
401 peroxiredoxin (locus_tag b2480) in *E. coli.* SE-6283 also carries a *bcp* homolog (locus_tag
402 VYA_19800 in AP024165.1) as if it complements the disrupted *bcp* gene in the host genome
403 upon insertion [29]. Genome sequencing of strain 04Ya108 in this study revealed that one
404 SE-6283 copy was inserted into a *bcp* gene homolog in chromosome 1 (Fig 6 A). To obtain
405 further insight into the target site preference of SE-6283, SE-6283-like elements in public
406 databases were screened using blastn. SE-6283-like SEs were identified in three other *Vibrio*
407 species (Fig 6 A). The sizes of SE-6283 and its related SEs ranged from 13.8 kb (SE-6283) to
408 41.3 kb (SE-ValK09K1 in Fig 6 A). Although the motifs C and C' of the three SE-6283-like
409 elements were slightly different from those of SE-6283 (see *attL* in S7 Fig and *attR* in S8 Fig),

410 all are inserted into *bcp* orthologs in chromosome 1: chromosome 1 (or I) of *V. alginolyticus*
411 strain K09K1 is incorrectly named as chromosome II in CP017919. Comparison of *attL*, *attR*,
412 and unoccupied *attB* found in the SEs-free genome of each *Vibrio* species suggests the
413 insertion of SE-6283-like elements and the 6 bp sequence into the central 6 bp putative
414 crossover region in *attB* (Fig 6 B). Nucleotide sequences around putative nicking sites in *attB*
415 for SE-6283 were G and A-rich and did not show inverted repeat structure like the case of
416 *attB* for SE-6945.
417
418 **Fig 6. Location of SE-6283 and its related SEs.** (A) Comparisons of genetic organization
419 around SE elements in six genomic locations. GenBank accession numbers and the
420 coordinates of genomic segments compared are shown below the strain name. Except SE-
421 6283 on pSEA1, all elements are inserted into a *bcp* gene, while they all carry a *bcp* gene. (B)
422 Alignment of *attB* sequences from different bacterial species. (C) Sequence logos showing
423 conservation levels of *attB*. The arrowhead indicates the putative nicking site on the top
424 strand.
425

426 # Discussion

427 *Vibrio* is one of the major bacterial genera found in marine sediments [44], and is among the
428 most common microbiota of wild and farmed shrimp [45] and fish [46,47]. Several members
429 of the genus *Vibrio* are pathogens of fishes reared in aquaculture [48], while other subsets
430 of *Vibrio* species, including *V. cholerae*, *V. parahaemolyticus,* and *V. vulnificus*, which are
431 ubiquitous in relatively low-salinity sea water, are seafood-borne human pathogens [48,49].
432 Thus, *Vibrio* can be considered as a key genus linking AMR genes between aquatic

433 environments and human-associated environments. Indeed, the accumulation of AMR
434 genes in this genus has attracted increased research attention, particularly in *V. cholerae*
435 [50,51]. However, direct experimental evidence for AMR gene transfer from *Vibrio* species
436 to other human-relevant bacteria is limited.

437 Known genetic mechanisms of AMR gene transfer from *Vibrio* include A/C plasmids
438 [52,53], unclassified conjugative plasmids [54], pAQU1-type MOB_H-family conjugative
439 plasmids [55], pSEA1-type MOB_H-family conjugative plasmids [29], ICEs [56], IMEs [57], and
440 a combination of a sedentary chromosomal integron and a conjugative plasmid carrying an
441 integron [58]. The host range of pSEA1-type MOB_H plasmids, discovered in *V. alfacsensis* (for
442 which only one complete genome was available until the present study), is unknown. In this
443 study, we investigated how a pSEA1-like plasmid can contribute to AMR gene transfer in a
444 laboratory setting.

445 In contrast to A/C plasmids, autonomous replication of pSEA2 and pSEA1 was
446 difficult to achieve in *E. coli,* since transconjugant selection generally yielded *E. coli* clones
447 carrying plasmid DNA integrated into the chromosome. The current most likely mechanism
448 behind this phenomenon is transposition of SE followed by homologous recombination,
449 which enabled AMR gene transfer beyond the plasmid's replication host range. This mode of
450 horizontal gene transfer may be important among marine bacteria, since the plasmid
451 conjugation host range is expected to be wider than the replication host range [59,60].

452 pSEA1 was found to carry two SEs (SE-6283 and SE-6945). However, in the previous
453 study [29], only transposition of SE-6283 was observed, likely because we did not recognize
454 SE-6945 on pSEA1 and did not intend to detect its transposition. The newly identified SE-
455 6945 is the smallest SE identified to date (Fig 2 and Fig 6), and the only known SE harboring
456 an AMR gene. As SEs can have more than one target site in a single genome, they may

457 contribute to microbial adaptation to the antimicrobials used at aquaculture sites, with
458 diverse mechanisms beyond mediating the horizontal transmission of AMR genes, such as
459 increasing the resistance level or gene redundancy preceding evolutionary innovation of an
460 AMR gene [61].

461 The discovery of SE-6945 highlights four potential core genes (*intA*, CDS2, *intB*, and
462 CDS4) present in SEs. However, the specific roles of these gene products in their
463 transposition steps remain to be determined. Among the transposon moves classified by
464 Curcio and Derbyshire [15], 'copy-out' is currently the most consistent with our
465 observations of SEs regarding the first step of its location change. One model for strand
466 exchange on single-strand DNA mediated by tyrosine recombinase, leading to a single-
467 strand circle, has been proposed for Integron gene cassette excision (replicative resolution
468 model in [41]). In the replicative resolution model, strand exchange between two folded
469 *attC* bottom strands and the following replication generates a gene cassette circle and an
470 empty *attC* [40]. SE circle formation barely accompanies the generation of empty *attB* (Fig 4
471 and [29]). Thus the rejoining of the host DNA side at nicking sites is somehow prevented,
472 possibly by the action of multiple proteins involved: i.e., CDS2, CDS4 products, unlike the
473 movement of the gene cassette. Currently, we do not have information on the fate of the
474 unjointed donor molecule. Unlike DDE transposase, tyrosine recombinases usually do not
475 generate a free 3′ OH end and are covalently linked to the phosphate of substrate DNA (Fig
476 4A) [62]. Therefore, DPC repair-equivalent process [38], which might require both host and
477 SE-encoding factors, is necessary to achieve the IS*3*-like copy-out process. Nonetheless, we
478 speculate that the alternate single-strand DNA circle generation route, accompanying
479 chromosome loss (bottom left in Fig 4B), is a less likely route as it does not give any
480 advantage to target site-selective and putatively non-transferrable SEs left in dying cells.

481 While strand-biased *attL* × *attR* recombination is repeatedly observed for the two
482 SEs, not much information has been obtained regarding their integration step. While 6 bp
483 from the 5′ end of C at *attL* on the donor molecule is incorporated into a circular SE copy,
484 the incorporated 6 bp tends to be placed next to the 3′ end of C′ at *attR* after insertion into
485 the target site (see *attL* on pSEA2 and *attRs* on TJ249, and previously detected SE-6283
486 transposition from pSEA1 to *bcp* in *E. coli*). Thus, there seems to be a strict rule in the order
487 of strand exchange at the integration step, but determining the precise order of nicking and
488 nicking sites on the two strands during the integration step requires further investigations.
489 A notable difference between IS and SE is the strong strand bias unique to SE upon
490 strand exchange: IS can generate a single-strand bridge using either of the two strands (but
491 not both simultaneously) [63,64], while SEs use only top strand. The dissimilarity between
492 SE and ICE/IMEs is that ICE/IMEs encode the recombination directionality factor Xis [24],
493 while SEs do not encode Xis homologs. CDS2 or CDS4 products may function exclusively in
494 either the initial circularization step or the integration as the recombination directionality
495 factor. Another dissimilarity between SE and ICE/IME is that the former does not encode
496 HUH endonuclease. There is still a possibility that SEs carry *oriT* recognized by a family of
497 relaxase and are mobilized like CIME [20], though we could not identify *oriT* in SE-6945 and
498 SE-6283 by local similarity search alone.
499 We propose that interplay between SEs and MDR conjugative plasmids contribute to
500 AMR gene transmission in marine bacteria. Further biochemical and bioinformatic studies
501 on SEs are needed to reveal the mechanisms of gene transfer among the genus *Vibrio* and
502 other aquaculture-associated bacteria.

503 To refer to individual SE members, we used digits that are compatible with the
504 existing transposon database [65]. However, further debate is needed regarding handling
505 SEs in mobile DNA databases.
506

507 Materials and methods

508 Strains, plasmids, and culture media

509 Strains and plasmids used in the study are listed in Table 1. Strain 04Ya108 was previously
510 identified as *Vibrio ponticus* based on 16S rRNA gene sequence similarity. However,
511 determination of the complete sequence in this study revealed that strain 04Ya108 shows
512 >96% average nucleotide identity to *V. alfacsensis* strain CAIM 1831 (DSM 24595) (S1 Fig)
513 [66]. Therefore, strain 04Ya108 was newly classified as *V. alfacsensis*. Strain LN95 is a pSEA2-
514 free tetracycline-susceptible derivative of strain 04Ya249. This strain was generated through
515 a repeated batch culture of 04Ya249, and subsequent single-colony isolation. The absence
516 of pSEA2 in LN95 was confirmed by next-generation sequencing.
517

518 **Table 1. Strains and plasmids used.**

Strains or plasmids	Relevant genotype, phenotype, descriptions[a c]	Reference
E. coli		
DH5α	F^-, Φ80d*lacZ*ΔM15, Δ(*lacZYA-argF*)U169, *deoR*, *recA*1, *endA*1, *hsdR*17(r_K^-, m_K^+), *phoA*, *supE*44, λ^-, *thi*-1, *gyrA*96, *relA*1	Takara Bio Inc., Shiga, Japan
JW0452	λ- *rph-1* Δ(*rhaD-rhaB*)*568 hsdR514* Δ*lacZ4787*(::*rrnB-3*) Δ*(araD-araB)567* Δ*acrA748*::*kan*; Km^r	[67]
JW0452rif	Spontaneous rifampicin-resistant mutant of JW0452	This study
W0452Δ *recA*	JW0452Δ*recA::cat*; Km^r, Cm^r	This study
JW0452Δ *recA*rif	Spontaneous rif resistant mutant of JW0452Δ*recA*; Km^r, Cm^r, Rif^r	This study

TJ249	JW0452 *insJ*::SE-6945::pSEA2 *yjjNt*::SE-6945; Ap^r Tc^r Ery^r Cm^r	This study
LN28	JW0452 chr::pSEA2[b]; Ap^r Tc^r Ery^r Cm^r	This study
LN29	JW0452 *yjjNt*::SE-6945::pSEA2; Ap^r Tc^r Ery^r Cm^r	This study
LN30	JW0452 *yjjNt*::SE-6945::pSEA2; Ap^r Tc^r Ery^r Cm^r	This study
LN32	JW0452 chr::pSEA2; Ap^r Tc^r Ery^r Cm^r	This study
LN33	JW0452 *insJ*::SE-6945::pSEA2; Ap^r Tc^r Ery^r Cm^r	This study
LN34	JW0452 *insJ*::SE-6945::pSEA2; Ap^r Tc^r Ery^r Cm^r	This study
LN35	JW0452 *insJ*::SE-6945::pSEA2; Ap^r Tc^r Ery^r Cm^r	This study
LN36	JW0452 chr::pSEA2[b]; Ap^r Tc^r Ery^r Cm^r	This study
LN37	JW0452 *insJ*::SE-6945::pSEA2; Ap^r Tc^r Ery^r Cm^r	This study
LN52	JW0452 *yjjNt*::SE-6945; Ap^r	This study
LN52rif	Spontaneous rifampicin-resistant mutant of LN52; Rif^r, Ap^r	This study
LN58	JW0452 *insJ*::SE-6945::pSEA2; Ap^r Tc^r Ery^r Cm^r	This study
LN82	JW0452 *yjjNt*::SE-6945::pSEA2; Ap^r Tc^r Ery^r Cm^r	This study
LN84	JW0452 *insJ*::SE-6945::pSEA2; Ap^r Tc^r Ery^r Cm^r	This study
LN86	JW0452 chr::pSEA2; Ap^r Tc^r Ery^r Cm^r	This study
LN88	JW0452 *yjjNt*::SE-6945::pSEA2; Ap^r Tc^r Ery^r Cm^r	This study
LN89	JW0452 *insJ*::SE-6945::pSEA2; Ap^r Tc^r Ery^r Cm^r	This study
LN90	JW0452 *insJ*::SE-6945::pSEA2; Ap^r Tc^r Ery^r Cm^r	This study
LN91	JW0452 *yjjNt*::SE-6945::pSEA2; Ap^r Tc^r Ery^r Cm^r	This study
Vibrio alfacsensis		
04Ya249	Aquaculture isolate harboring conjugative plasmid pSEA2; Tc^r, Ap^r, Ery^r Cm^r	[26]
04Ya108	Aquaculture isolate harboring conjugative plasmid pSEA1; Tc^r, Ap^r, Ery^r Cm^r	[26]
LN95	pSEA2 free derivative of strain 04Ya249; Ap^r	This study
Plasmids		
pGEM-T	Cloning vector; pUC replicon, Ap^r	Promega, Madison, WI, USA
pGEM-attS	pGEM-T carrying the 206 bp *attS* fragment; pUC replicon, Ap^r	This study
pGEM-int	pGEM-T carrying the 103 bp *intA* fragment; pUC replicon, Ap^r	This study
pGEM-gyrB	pGEM-T carrying the 110 bp *gyrB* fragment from strain 04Ya249; pUC replicon, Ap^r	This study
pKD46	Lambda-Red recombinase expression plasmid; pSC101ts replicon, Ap^r	[68]
pKD3	Donor of the *cat* gene for the Lambda-Red recombination experiment; R6K replicon, Cm^r	[68]

519 *a*. Antimicrobial resistance phenotypes (^{r}) listed were confirmed by experiments.
520 Abbreviations are as follows: Ap, ampicillin; Tc, tetracycline: Ery, erythromycin; Cm,
521 chloramphenicol; Rif, rifampicin; *kan*, kanamycin resistance gene. *yjjNt* denotes the putative
522 transcription termination region for the *yjjN* gene.
523 *b.* Integration of pSEA2 without generating a new junction of SE and chromosome.
524 Integration location of pSEA2 is unknown.
525 *c*. 19 transconjugants listed (TJ249, LN28, LN29, LN30, LN32, LN33, LN34, LN35, LN36, LN37,
526 LN52, LN58, LN82, LN84, LN86, LN88, LN89, LN90, LN91) were obtained from 19
527 independent mating experiments. Selection conditions used in transconjugant screening are
528 as follows: Tc 42 °C for LN28, LN29 LN30, LN32, LN33, LN34, LN35, LN36, and LN37; Ap 42 °C
529 for LN52; Ap Tc 42 °C for LN58 LN82, LN84, LN86, LN88, LN89, LN90, and LN91; Ery 42 °C for
530 TJ249.
531
532 *E. coli* strain LN52, which carries a single copy of SE-6945, is a JW0452
533 transconjugant obtained by ampicillin selection at 42 °C. TJ249 is also a JW0452
534 transconjugant obtained by erythromycin selection at 42 °C. The presence of a single SE-
535 6945 copy in the LN52 genome was confirmed by Southern hybridization (S5 Fig). The *recA*-
536 null mutant of JW0452, JW0452Δ*recA*, was constructed using the lambda-Red method [68].
537 A DNA fragment containing the 5′ and 3′ sequences of the *recA* gene and a
538 chloramphenicol resistance gene was amplified by PCR using primers YO-175 and
539 RecA_stop_primingsite_2 (Table 2) and plasmid pKD3 [68] as a template. The PCR products
540 (700 ng) were introduced into electrocompetent cells of JW0452 carrying pKD46 by
541 electroporation using a Gene Pulser Xcell™ (BioRad, Hercules, CA, USA). The occurrence of
542 recombination at the expected site was confirmed by PCR using primers CAT-584 and

543 BW25113_2815723f (Table 2). The absence of the *recA* gene was also confirmed by PCR
544 using primers LN192_recA1 and LN193_recA2.
545 *Vibrio alfacsensis* strains were cultured in BD Bacto™ brain heart infusion medium
546 (BD237500; Becton, Dickinson, and Company, Franklin Lakes, NJ, USA) supplemented with
547 up to 2% NaCl at 25 °C. *E. coli* strains were cultured in BD Difco™ LB Broth, Miller
548 (BD244520; Becton, Dickinson, and Company) at 37 °C or 42 °C. BD Difco™ Mueller Hinton
549 Broth (BD 275730; Becton, Dickinson, and Company) was used for antibiotic susceptibility
550 testing of *E. coli*. BD Difco™ Marine Broth 2216 (BD279110; Becton, Dickinson, and
551 Company) was used for filter mating. Solid media were prepared by adding 1.5% agar to the
552 broth. Antibiotics were added to the medium at the following concentrations when
553 required: erythromycin (Nacalai Tesque, Kyoto, Japan), 100 μg/ml; tetracycline (Nacalai
554 Tesque), 10 μg/ml; rifampicin (Sigma-Aldrich, St. Louis, MO, USA), 50 μg/ml; ampicillin
555 (Nacalai Tesque), 100 μg/ml.
556
557 **Table 2. Oligonucleotides used.**

Name	Sequence (5′ to 3′)	Purpose
LN112	GGGTTACCTTCCCAATGCGT	Southern hybridization probe for SE-6945 *intA*
LN113	CGACTGTTGGTAGCGACTGT	Southern hybridization probe for SE-6945 *intA*
LN_142_junction2	AAGATGGTAAAAGTGTTCCA	Detection of *attS* (206 bp) by qPCR and *attL* (304 bp) by standard PCR
LN_143_junction2	TTTGTGTGTAGCCCTTGTG	Detection of *attS* (206 bp) by

		qPCR and *attR* (248 bp) by standard PCR
LN_150_intA2	GGTTATGTGGAGAAGTTGCC	Detection of *intA* (103 bp) by qPCR
LN_151_intA2	TGAGTTCGGTTTCTTGCTTC	Detection of *intA* (103 bp) by qPCR
LN_181_Valc_chr1_attB_L	CGAGGGTAAAGTGCCAACAT	Detection of chromosomal *attB* (345 bp if empty, 7.4 kb if occupied) and *attL* (304 bp) by standard PCR
LN183_Valc_chr1_attB_R2	ACATCAGCAGGAGTTAGTTG	Detection of chromosomal *attB* (345 bp, 7.4 kb) and *attR* (248 bp) by standard PCR
LN184_04Ya249_gyrBf1	AACAGAATTGCACCCAGAAG	Detection of *gyrB* (110 bp) by qPCR
LN185_04Ya249_gyrBr1	GAAGACCGCCTGATACTTTG	Detection of *gyrB* (110 bp) by qPCR
YO-175	CAGAACATATTGACTATCCGGTATTACCCGGCATGACAGGAGTAAAAATGTGTAGGCTGGAGCTGCTTCG	Amplification of *cat* gene for Lambda-Red recombination
RecA_stop_primingsite_2	ATGCGACCCTTGTGTATCAAACAAGACGATTAAAAATCTTCGTTAGTTTCCATATGAATATCCTCCTTA	Amplification of *cat* gene for Lambda-Red recombination
CAT-584	AAGCCATCACAAACGGCATG	Confirmation of *cat* gene insertion

BW25113_2815723f	AATACGCGCAGGTCCATAAC	Confirmation of *cat* gene insertion
LN192_recA1	GTTCCATGGATGTGGAAACC	Confirmation of *recA* deletion
LN193_recA2	ATATCGACGCCCAGTTTACG	Confirmation of *recA* deletion

558

559 ## Conjugation

560 The donor *Vibrio* strain and *E. coli* recipient strains were grown overnight at 25 °C and 37 °C,
561 respectively. A 500 μl aliquot of each culture was mixed, centrifuged, and resuspended in 50
562 μl of Luria-Bertani broth. The cell mixture was spotted on a 0.45 μm pore-size nitrocellulose
563 filter (Merck, Millipore Ltd., Tullagreen, Ireland) placed on marine broth agar, and allowed
564 to mate for 24 h at 25 °C. After incubation, the cell mixture on the filter was serially diluted
565 in 1× phosphate-buffered saline, and then 100 μl of the mixture was plated on an
566 appropriate agar medium to measure the CFU. The transconjugant strain TJ249 was
567 obtained by the first mating experiment of strain 04Ya249 with *E. coli* strain JW0452 as the
568 donor and recipient, respectively, followed by Ery selection at 42 °C. A high temperature
569 (42 °C) was used to suppress the growth of the *Vibrio alfacsensis* strain on LB agar. After
570 genome sequence determination of TJ249, more than 17 independent mating experiments
571 were performed to obtain insights into the variation of pSEA2 integration patterns using
572 JW0452 as the recipient, and transconjugants were screened at 42 °C in the presence of Tc
573 alone, Ap alone, or both Ap and Tc. The AMR phenotype of each transconjugant was further
574 determined after single-colony isolation using Ery, Cm, Ap, and Tc (Table 1).

575 Quantitative mating assays (Fig 3) were performed to address the involvement of
576 homologous recombination in pSEA2 integration using strains JW0452rif, JW0452Δ*recA*rif,
577 and LN95rif as recipients. Selection conditions for transconjugants were Rif Ap Tc 42 °C or
578 Rif Ap 42 °C. The *Vibrio* donor strain was selected on brain heart infusion agar with 2% NaCl
579 supplemented with Tc at 25 °C. CFUs of donor, recipient, and transconjugant were
580 determined after 24 hours of incubation.
581

582 **PCR experiments**

583 The PCR detection of *attL*, *attR,* and *attS* was performed using TaKaRa Ex Taq (TaKaRa Bio
584 Inc., Kusatsu, Japan), and genomic DNA of 04Ya249, LN95, and LN52 prepared using QIAGEN
585 DNeasy Blood & Tissue kit (QIAGEN GmbH, Hilden, Germany). PCR premix was prepared in a
586 50 μl scale, and 1 μl of the template (10 ng/μl) was added to the reaction mix throughout
587 the study. The PCR cycle used was 94 °C for 1 min, followed by 35 cycles of 94 °C for 30 s
588 and 60 °C for 30 s. As the empty *attB* was not amplified using TaKaRa Ex Taq, we conducted
589 additional Long-PCR using LA Taq polymerase (TaKaRa Bio Inc.) to detect occupied *attB* and
590 confirm the functionality of primers used for detection of empty *attB*. The PCR cycles used
591 for Long-PCR were 1 cycle of 94 °C for 1 min, followed by 35 cycles of 95 °C for 1 min, 65 °C
592 for 1 min, and 72 °C for 1 min. PCR products of *attS* were further cloned using the pGEM-T
593 easy vector system (Promega, Madison, WI, USA) and DH5α competent cells to analyze the
594 sequence variation of PCR products.
595 Quantitative PCR was performed using THUNDERBIRD® SYBR qPCR Mix (Toyobo,
596 Osaka, Japan) and a CFX connect Real-Time system (BioRad, Hercules, CA, USA) and a two-
597 step PCR protocol of 40 cycles of 95 °C for 5 s and 60 °C for 30 s. Target quantity was

598 estimated based on a standard curve of the control plasmid DNA (pGEM-*gyrB*, pGEM-*intA*,
599 pGEM-*attS*) constructed by TA cloning. Primers used and their purposes are listed in Table 2
600 and Fig 4 legend.
601

602 Southern hybridization

603 The insertion copy number and insertion location of SE-6945 in transconjugants (JW0452
604 derivatives) were analyzed by Southern hybridization using the 5′ end of *intA*, amplified
605 using primers LN112 and LN113 (Table 2) and the PCR DIG Synthesis Kit (Roche, Basel,
606 Switzerland) as a probe. Genomic DNA (2.5 μg) was double digested with either *Nde*I and
607 *Sph*I or *Nde*I and *Hin*dIII (New England Biolabs, Ipswich, MA, USA), then electrophoresed in
608 0.8% agarose gel. Separated DNA fragments were transferred to the nitrocellulose
609 membranes (Amersham Hybond-N+; GE Healthcare UK Ltd., Buckinghamshire, UK) by
610 capillary transfer. The probe was detected using CDT-star® (Roche). Restriction map for the
611 DNA segments around SE-6945 insertion locations and the expected size of DNA segments
612 with probe signal are shown in S5 Fig A.
613

614 Antimicrobial susceptibility testing

615 To examine the antimicrobial susceptibility of transconjugants harboring one to three SE-
616 6945 copies, the MIC of ampicillin was determined using the broth dilution method in 96-
617 well microtiter plate format according to standard M07 of the CLSI [69]. The antibiotic
618 concentrations tested were 1000, 500, 250, 125, 62.5, 31.25, 16, 8, 4, and 2 μg/ml. The test
619 plates were incubated at 35 °C for 24 h.
620

621 ## Genome sequencing

622 Genomic DNA was extracted from 250–500 μl of *V. alfacsensis* strains 04Ya249, 04Ya108, or
623 *E. coli* strain TJ249 culture using the QIAGEN DNeasy Blood & Tissue kit (QIAGEN GmbH).
624 The extracted genomic DNA was sequenced on the PacBio RS II platform using P6C4
625 chemistry at Macrogen (Tokyo, Japan) in 2016. Genome assembly was conducted using
626 HGAP v.3 [70] for *Vibrio* strains and Flye v 2.8.3-b1695 for *E. coli* TJ249 [71]. Reads were
627 obtained at >120× coverage for the chromosomes of each strain. Reads and the genome
628 sequence of TJ249 have been posted to figshare [72]. Illumina reads of the pSEA2-free strain
629 LN95 were obtained using the TruSeq PCR-free library and NovaSeq 6000 platform at
630 NovogeneAIT Genomics Singapore Pte., Ltd. (Singapore) to confirm the loss of pSEA2 and to
631 identify the circular form of SE-6945. The mapping results have been posted to figshare [42].
632 Genomes were compared using MUMmer3.23 [73] and GenomeMatcher [74]. AMR genes in
633 the assembly were searched using AMRFinderPlus [75]. The average nucleotide identity was
634 determined using fastANI [76].
635

636 **Searches for SE-6945 and SE-6283-related SEs**

637 Nucleotide sequences of the SE core gene region (*intA*, CDS2, *intB*, CDS4) (6432 bp for SE-
638 6283, 5541 bp for SE-6945) were used as queries in blastn searches against the NCBI
639 nucleotide collection (nr/nt) database (August 3, 2021). Hit subjects showing > 85 %
640 nucleotide identity with > 99% query coverage were regarded as close relatives of SE-6945
641 or SE-6283, and were further investigated as to their insertion locations.
642

643 ## Data availability

Sequence data generated in this study and their availability are summarized in Table S2. The complete genome sequences of strains 04Ya108 and 04Ya249 were submitted to GenBank/DDBJ/EMBL under accession numbers AP024165 (chr1 of 04Ya108), AP024166 (chr2 of 04Ya108), AP024167 (pSEA1 of 04Ya108), AP024168 (pYa108 of 04Ya108), AP019849 (chr1 of 04Ya249), AP019850 (chr2 of 04Ya249), AP019851 (pSEA2 of 04Ya249), and AP019852 (pVA249 of 04Ya249). The raw reads for strains 04Ya108, 04Ya249, and LN95 are available from Sequence Read Archive under accession numbers DRA011098, DRA008632, and DRA011762, respectively.

Acknowledgments

We thank the National Bioresource Project of the National Institute of Genetics, Japan, for providing the *E. coli* strains BW25113, JW2669, and JW0452. We thank Atsushi Ota and Fumito Maruyama at Hiroshima University for their support for the annotation on the TJ249 assembly. We thank Yuichi Otsuka at Saitama University for helpful discussions and Yuta Sugimoto at Ehime University for experimental support. We thank Kenji K Kojima at Genetic Information Research Institute for the discussion on transposon classification, and Yasukazu Daigaku at Tohoku University for the discussion on DNA protein cross link repair. We thank Editage for language editing (job number LINON_4).

References

1. Cabello FC, Godfrey HP, Tomova A, Ivanova L, Dölz H, Millanao A et al. (2013) Antimicrobial use in aquaculture re-examined: its relevance to antimicrobial resistance and to animal and human health. Environ Microbiol 15: 1917-1942.

667 2. Cabello FC, Godfrey HP, Buschmann AH, Dölz HJ (2016) Aquaculture as yet another
668 environmental gateway to the development and globalisation of antimicrobial
669 resistance. Lancet Infect Dis 16: e127-e133.
670 3. Magiorakos AP, Srinivasan A, Carey RB, Carmeli Y, Falagas ME, Giske CG et al. (2012)
671 Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an
672 international expert proposal for interim standard definitions for acquired resistance.
673 Clin Microbiol Infect 18: 268-281.
674 4. Karakonstantis S, Kritsotakis EI, Gikas A (2020) Pandrug-resistant Gram-negative
675 bacteria: a systematic review of current epidemiology, prognosis and treatment
676 options. J Antimicrob Chemother 75: 271-282.
677 5. Rodríguez-Beltrán J, DelaFuente J, León-Sampedro R, MacLean RC, San Millán Á (2021)
678 Beyond horizontal gene transfer: the role of plasmids in bacterial evolution. Nat Rev
679 Microbiol 19: 347-359.
680 6. Johnson CM, Grossman AD (2015) Integrative and Conjugative Elements (ICEs): What
681 They Do and How They Work. Annu Rev Genet 49: 577-601.
682 7. Yano H, Rogers LM, Knox MG, Heuer H, Smalla K, Brown CJ et al. (2013) Host range
683 diversification within the IncP-1 plasmid group. Microbiology (Reading) 159: 2303-
684 2315.
685 8. Dahmane N, Libante V, Charron-Bourgoin F, Guédon E, Guédon G, Leblond-Bourget N
686 et al. (2017) Diversity of integrative and conjugative elements of *Streptococcus*
687 *salivarius* and their intra- and interspecies Transfer. Appl Environ Microbiol 83:
688 e00337-17.

689 9. Cabezón E, Lanka E, de la Cruz F (1994) Requirements for mobilization of plasmids
690 RSF1010 and ColE1 by the IncW plasmid R388: *trwB* and RP4 *traG* are interchangeable.
691 J Bacteriol 176: 4455-4458.
692 10. Carraro N, Matteau D, Luo P, Rodrigue S, Burrus V (2014) The master activator of
693 IncA/C conjugative plasmids stimulates genomic islands and multidrug resistance
694 dissemination. PLoS Genet 10: e1004714.
695 11. Daccord A, Ceccarelli D, Burrus V (2010) Integrating conjugative elements of the
696 SXT/R391 family trigger the excision and drive the mobilization of a new class of *Vibrio*
697 genomic islands. Mol Microbiol 78: 576-588.
698 12. Garcillán-Barcia MP, Francia MV, de la Cruz F (2009) The diversity of conjugative
699 relaxases and its application in plasmid classification. FEMS Microbiol Rev 33: 657-687.
700 13. Villa L, Carattoli A (2020) Plasmid Typing and Classification. Methods Mol Biol 2075:
701 309-321.
702 14. Roberts AP, Chandler M, Courvalin P, Guédon G, Mullany P, Pembroke T et al. (2008)
703 Revised nomenclature for transposable genetic elements. Plasmid 60: 167-173.
704 15. Curcio MJ, Derbyshire KM (2003) The outs and ins of transposition: from mu to
705 kangaroo. Nat Rev Mol Cell Biol 4: 865-877.
706 16. Bouvier M, Demarre G, Mazel D (2005) Integron cassette insertion: a recombination
707 process involving a folded single strand substrate. EMBO J 24: 4356-4367.
708 17. Val ME, Bouvier M, Campos J, Sherratt D, Cornet F, Mazel D et al. (2005) The single-
709 stranded genome of phage CTX is the form used for integration into the genome of
710 Vibrio cholerae. Mol Cell 19: 559-566.
711 18. Doublet B, Boyd D, Mulvey MR, Cloeckaert A (2005) The Salmonella genomic island 1 is
712 an integrative mobilizable element. Mol Microbiol 55: 1911-1924.

713 19. Pavlovic G, Burrus V, Gintz B, Decaris B, Guédon G (2004) Evolution of genomic islands
714 by deletion and tandem accretion by site-specific recombination: ICE*St1*-related
715 elements from *Streptococcus thermophilus*. Microbiology (Reading) 150: 759-774.
716 20. Achard A, Leclercq R (2007) Characterization of a small mobilizable transposon,
717 MTnSag1, in Streptococcus agalactiae. J Bacteriol 189: 4328-4331.
718 21. Guédon G, Libante V, Coluzzi C, Payot S, Leblond-Bourget N (2017) The obscure world
719 of integrative and mobilizable elements, highly widespread elements that pirate
720 bacterial conjugative systems. Genes (Basel) 8: E337.
721 22. Caparon MG, Scott JR (1989) Excision and insertion of the conjugative transposon
722 Tn*916* involves a novel recombination mechanism. Cell 59: 1027-1034.
723 23. Auchtung JM, Lee CA, Monson RE, Lehman AP, Grossman AD (2005) Regulation of a
724 *Bacillus subtilis* mobile genetic element by intercellular signaling and the global DNA
725 damage response. Proc Natl Acad Sci U S A 102: 12554-12559.
726 24. Burrus V, Waldor MK (2003) Control of SXT integration and excision. J Bacteriol 185:
727 5045-5054.
728 25. Naito M, Ogura Y, Itoh T, Shoji M, Okamoto M, Hayashi T et al. (2016) The complete
729 genome sequencing of *Prevotella intermedia* strain OMA14 and a subsequent fine-
730 scale, intra-species genomic comparison reveal an unusual amplification of conjugative
731 and mobile transposons and identify a novel *Prevotella*-lineage-specific repeat. DNA
732 Res 23: 11-19.
733 26. Nonaka L, Ikeno K, Suzuki S (2007) Distribution of tetracycline resistance gene, *tet(M)*,
734 in Gram-positive and Gram-negative bacteria isolated from sediment and seawater at
735 a coastal aquaculture site in Japan. Microbes Environ 22: 355-364.

736 27. Nonaka L, Maruyama F, Miyamoto M, Miyakoshi M, Kurokawa K, Masuda M (2012)
737 Novel conjugative transferable multiple drug resistance plasmid pAQU1 from
738 *Photobacterium damselae* subsp. *damselae* isolated from marine aquaculture
739 environment. Microbes Environ 27: 263-272.
740 28. Nonaka L, Maruyama F, Suzuki S, Masuda M (2015) Novel macrolide-resistance genes,
741 *mef*(C) and *mph*(G), carried by plasmids from *Vibrio* and *Photobacterium* isolated from
742 sediment and seawater of a coastal aquaculture site. Lett Appl Microbiol 61: 1-6.
743 29. Nonaka L, Yamamoto T, Maruyama F, Hirose Y, Onishi Y, Kobayashi T et al. (2018)
744 Interplay of a non-conjugative integrative element and a conjugative plasmid in the
745 spread of antibiotic resistance via suicidal plasmid transfer from an aquaculture *Vibrio*
746 isolate. PLoS One 13: e0198613.
747 30. Neela FA, Nonaka L, Suzuki S (2007) The diversity of multi-drug resistance profiles in
748 tetracycline-resistant Vibrio species isolated from coastal sediments and seawater. J
749 Microbiol 45: 64-68.
750 31. Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A et al. (2020)
751 CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic
752 resistance database. Nucleic Acids Res 48: D517-D525.
753 32. Pei J, Kim BH, Grishin NV (2008) PROMALS3D: a tool for multiple protein sequence and
754 structure alignments. Nucleic Acids Res 36: 2295-2300.
755 33. Nunes-Düby SE, Kwon HJ, Tirumalai RS, Ellenberger T, Landy A (1998) Similarities and
756 differences among 105 members of the Int family of site-specific recombinases.
757 Nucleic Acids Res 26: 391-406.
758 34. Landy A (1989) Dynamic, structural, and regulatory aspects of lambda site-specific
759 recombination. Annu Rev Biochem 58: 913-949.

760 35. Schneider TD, Stephens RM (1990) Sequence logos: a new way to display consensus
761 sequences. Nucleic Acids Res 18: 6097-6100.
762 36. Stokes HW, O'Gorman DB, Recchia GD, Parsekhian M, Hall RM (1997) Structure and
763 function of 59-base element recombination sites associated with mobile gene
764 cassettes. Mol Microbiol 26: 731-745.
765 37. Stokes HW, Hall RM (1989) A novel family of potentially mobile DNA elements
766 encoding site-specific gene-integration functions: integrons. Mol Microbiol 3: 1669-
767 1683.
768 38. Zhang H, Xiong Y, Chen J (2020) DNA-protein cross-link repair: what do we know now.
769 Cell Biosci 10: 3.
770 39. Chandler M, Fayet O, Rousseau P, Ton Hoang B, Duval-Valentin G (2015) Copy-out-
771 paste-in transposition of IS*911*: a major transposition pathway. Microbiol Spectr 3:
772 40. Collis CM, Hall RM (1992) Site-specific deletion and rearrangement of integron insert
773 genes catalyzed by the integron DNA integrase. J Bacteriol 174: 1574-1585.
774 41. Escudero JA, Loot C, Nivina A, Mazel D (2015) The integron: adaptation on demand.
775 Microbiol Spectr 3: MDNA3-0019.
776 42. Yano H, Nonaka L (2021) Illumina reads analysis of *Vibrio alfacsensis* pSEA2-free strain
777 LN95. URL: https://doiorg/106084/m9figshare14338883
778 43. Chandler M, de la Cruz F, Dyda F, Hickman AB, Moncalian G, Ton-Hoang B (2013)
779 Breaking and joining single-stranded DNA: the HUH endonuclease superfamily. Nat
780 Rev Microbiol 11: 525-538.
781 44. Chen J, McIlroy SE, Archana A, Baker DM, Panagiotou G (2019) A pollution gradient
782 contributes to the taxonomic, functional, and resistome diversity of microbial
783 communities in marine sediments. Microbiome 7: 104.

784 45. Cornejo-Granados F, Lopez-Zavala AA, Gallardo-Becerra L, Mendoza-Vargas A, Sánchez
785 F, Vichido R et al. (2017) Microbiome of Pacific Whiteleg shrimp reveals differential
786 bacterial community composition between Wild, Aquacultured and AHPND/EMS
787 outbreak conditions. Sci Rep 7: 11783.
788 46. Llewellyn MS, McGinnity P, Dionne M, Letourneau J, Thonier F, Carvalho GR et al.
789 (2016) The biogeography of the atlantic salmon (*Salmo salar*) gut microbiome. ISME J
790 10: 1280-1284.
791 47. Egerton S, Culloty S, Whooley J, Stanton C, Ross RP (2018) The Gut Microbiota of
792 Marine Fish. Front Microbiol 9: 873.
793 48. Thompson FL, Iida T, Swings J (2004) Biodiversity of Vibrios. Microbiol Mol Biol Rev 68:
794 403-431.
795 49. Takemura AF, Chien DM, Polz MF (2014) Associations and dynamics of *Vibrionaceae* in
796 the environment, from the genus to the population level. Front Microbiol 5: 38.
797 50. Verma J, Bag S, Saha B, Kumar P, Ghosh TS, Dayal M et al. (2019) Genomic plasticity
798 associated with antimicrobial resistance in *Vibrio cholerae*. Proc Natl Acad Sci U S A
799 116: 6226-6231.
800 51. Osei Sekyere J, Reta MA (2020) Genomic and resistance epidemiology of Gram-
801 negative bacteria in Africa: a systematic review and phylogenomic analyses from a one
802 health perspective. mSystems 5: e00897-20.
803 52. Carraro N, Sauvé M, Matteau D, Lauzon G, Rodrigue S, Burrus V (2014) Development
804 of pVCR94ΔX from *Vibrio cholerae*, a prototype for studying multidrug resistant IncA/C
805 conjugative plasmids. Front Microbiol 5: 44.

806 53. Folster JP, Katz L, McCullough A, Parsons MB, Knipe K, Sammons SA et al. (2014)
807 Multidrug-resistant IncA/C plasmid in *Vibrio cholerae* from Haiti. Emerg Infect Dis 20:
808 1951-1953.
809 54. Liu M, Wong MH, Chen S (2013) Molecular characterisation of a multidrug resistance
810 conjugative plasmid from *Vibrio parahaemolyticus*. Int J Antimicrob Agents 42: 575-
811 579.
812 55. Nonaka L, Maruyama F, Onishi Y, Kobayashi T, Ogura Y, Hayashi T et al. (2014) Various
813 pAQU plasmids possibly contribute to disseminate tetracycline resistance gene *tet*(M)
814 among marine bacterial community. Front Microbiol 5: 152.
815 56. Beaber JW, Burrus V, Hochhut B, Waldor MK (2002) Comparison of SXT and R391, two
816 conjugative integrating elements: definition of a genetic backbone for the mobilization
817 of resistance determinants. Cell Mol Life Sci 59: 2065-2070.
818 57. Carraro N, Rivard N, Ceccarelli D, Colwell RR, Burrus V (2016) IncA/C conjugative
819 plasmids mobilize a new family of multidrug resistance islands in clinical *Vibrio*
820 *cholerae* non-O1/non-O139 isolates from Haiti. mBio 7: e00509-16.
821 58. Rowe-Magnus DA, Guerout AM, Mazel D (2002) Bacterial resistance evolution by
822 recruitment of super-integron gene cassettes. Mol Microbiol 43: 1657-1669.
823 59. Encinas D, Garcillán-Barcia MP, Santos-Merino M, Delaye L, Moya A, de la Cruz F
824 (2014) Plasmid conjugation from proteobacteria as evidence for the origin of
825 xenologous genes in cyanobacteria. J Bacteriol 196: 1551-1559.
826 60. Kishida K, Inoue K, Ohtsubo Y, Nagata Y, Tsuda M (2017) Host range of the conjugative
827 transfer system of IncP-9 naphthalene-catabolic plasmid NAH7 and characterization of
828 its *oriT* region and relaxase. Appl Environ Microbiol 83: e02359-16.

829 61. Elliott KT, Cuff LE, Neidle EL (2013) Copy number change: evolving views on gene
830 amplification. Future Microbiol 8: 887-899.
831 62. Jayaram M, Ma CH, Kachroo AH, Rowley PA, Guga P, Fan HF et al. (2015) An overview
832 of tyrosine site-specific recombination: from an Flp perspective. Microbiol Spectr 3:
833 63. Polard P, Chandler M (1995) An *in vivo* transposase-catalyzed single-stranded DNA
834 circularization reaction. Genes Dev 9: 2846-2858.
835 64. Rousseau P, Loot C, Turlan C, Nolivos S, Chandler M (2008) Bias between the left and
836 right inverted repeats during IS*911* targeted insertion. J Bacteriol 190: 6111-6118.
837 65. Tansirichaiya S, Rahman MA, Roberts AP (2019) The Transposon Registry. Mob DNA
838 10: 40.
839 66. Gomez-Gil B, Roque A, Chimetto L, Moreira APB, Lang E, Thompson FL (2012) *Vibrio*
840 *alfacsensis* sp. nov., isolated from marine organisms. Int J Syst Evol Microbiol 62: 2955-
841 2961.
842 67. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M et al. (2006) Construction of
843 *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol
844 Syst Biol 2: 2006.0008.
845 68. Datsenko KA, Wanner BL (2000) One-step inactivation of chromosomal genes in
846 *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci U S A 97: 6640-6645.
847 69. Wayne PA (2018) Methods for Dilution Antimicrobial Susceptibility Tests for Bacteria
848 That Grow Aerobically. 11th ed. CLSI standard M07. Clinical and Laboratory Standards
849 Institute.
850 70. Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C et al. (2013)
851 Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing
852 data. Nat Methods 10: 563-569.

71. Kolmogorov M, Yuan J, Lin Y, Pevzner PA (2019) Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol 37: 540-546.

72. Yano H, Nonaka L (2021) Complete genome of transconjugant TJ249. URL: https://doiorg/106084/m9figshare13332467

73. Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C et al. (2004) Versatile and open software for comparing large genomes. Genome Biol 5: R12.

74. Ohtsubo Y, Ikeda-Ohtsubo W, Nagata Y, Tsuda M (2008) GenomeMatcher: a graphical user interface for DNA sequence comparison. BMC Bioinformatics 9: 376.

75. Feldgarden M, Brover V, Haft DH, Prasad AB, Slotta DJ, Tolstoy I et al. (2019) Validating the AMRFinder tool and resistance gene database by using antimicrobial resistance genotype-phenotype correlations in a collection of isolates. Antimicrob Agents Chemother 63: e00483-19.

76. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S (2018) High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun 9: 5114.

Supporting information

S1 Fig. Comparison of the genome structure between strain 04Ya249 with 04Ya108, CAIM 1831 or TJ249. Structure comparison was performed using nucmer in MUMmer3 [73]. Purple dots indicate a match on the Watson strand (5′ to 3′ on the top strand in GenBank file), and light blue indicates a match on the Crick strand (5’ to 3’ on the bottom strand). (B) Average nucleotide identity (ANI) between two strains as determined by fastANI [76]. The

875 commands used were as follows: (A) $nucmer -minmatch

876 60../../data/04Ya249_submission.fas../../data/reference.fas

877 $mummerplot -x "[0,6000000]" -y "[0,22000000]" -postscript -p test out.delta; (B) $fastANI -

878 q../../data/04Ya249_submission.fas -r../../data/CAIM1831_Refseq.fas -o

879 04Ya249vsCAIM1831.txt

880 $fastANI -q../../data/04Ya249_submission.fas -r../../data/04Ya108_submission.fas -o

881 04Ya249vs04Ya108.txt

882 $fastANI -q../../data/04Ya249_submission.fas -r../../data/04Ya108_submission.fas -o

883 04Ya249vs04Ya108.txt

884

885 **S2 Fig. Genetic map of pSEA2**. (A) Locations of antimicrobial resistance (AMR) genes and SE-

886 6945 in pSEA2. AMR genes were inferred using AMRFinderPlus [75]. Genes are visualized

887 using CLC Sequence Viewer (Qiagen, Hilden Germany). Green arc denotes the region

888 detected as a circular contig (contig1) in Pacbio RSII reads assembly of transconjugant strain

889 TJ249. (B) Location of SEs in pSEA1 and pSEA2. Four SE core genes are indicated by four

890 distinct colors. Red pentagons are the β-lactamase gene.

891

892 **S3 Fig. *attL* sequences formed in the SE-6945 insertion regions.** Sequences are derived

893 from 04Ya249 genomes (AP019851.1, AP019849.1), TJ239 genome (*insJ*, *yjjNt*), and *Vibrio*

894 *harveri* strain WXL538 (CP045070.1) carrying a SE-6945-like element carrying a different set

895 of cargo genes. Sequences in red are a part of the mobile DNA unit of the SE element. 6-bp

896 underlined sequence is incorporated into *attS* during SE circle formation. 19 bp motif C

897 highlighted in yellow form an imperfect inverted repeat with the motif C' located the other

898 terminus of SE. Sequence highlighted in cyan is a putative transcription termination motif.
899 Asterisks indicate conserved nucleotides.
900
901 **S4 Fig. *attR* sequences formed in the SE-6945 insertion regions.** Sequences are derived
902 from 04Ya249 genomes (AP019851.1, AP019849.1), TJ239 genome (*insJ*, *yjjNt*), and *Vibrio*
903 *harveri* strain WXL538 (CP045070.1). Sequences in red are a part of the mobile DNA unit of
904 the SE element. The 6 bp underlined sequences originate from the original *attB* or SE circle,
905 and they are not incorporated into the *attS* of *the* SE circle.
906
907 **S5 Fig. Inference of SE-6945 insertion number and its insertion location by Southern**
908 **hybridization.** (A) *in silico* restriction map of DNA segments around SE-6945 insertion
909 locations in strain 04Ya249, TJ249, and a hypothetical *E. coli* strain carrying a pSEA2
910 insertion into SE-6945 integrated into *yjjNt*. Locations of SE-6945 and *intA* probe
911 hybridization region are indicated by the yellow square and black filled square, respectively,
912 above the restriction map. Horizontal black lines indicate the fragments detected by the
913 probe and their sizes. (B) pSEA2-free *Vibrio* strain LN95 and parental strain 04Ya249. (C)
914 Nineteen *E. coli* transconjugants obtained from 19 independent mating assays. Upper panel:
915 digestion with *Nde*I and *Sph*I. Lower panel: digestion with *Nde*I and *Hin*dIII. The color of the
916 strain name indicates the pattern of SE/pSEA2 insertion: blue, three SE copies (*yjjNt*::SE-
917 6945, *insJ*::SE-6945::pSEA2); red, two SE copies (*insJ*::SE-6945::pSEA2); purple, (*yjjNt:*::SE-
918 6945::pSEA2); lime, one SE copy (*yjjNt*::SE-6945 or chr::pSEA2); Strains with asterisk (*) are
919 expected to carry pSEA2 in an unknown chromosomal location without generating a SE-
920 6945-chromosome junction.
921

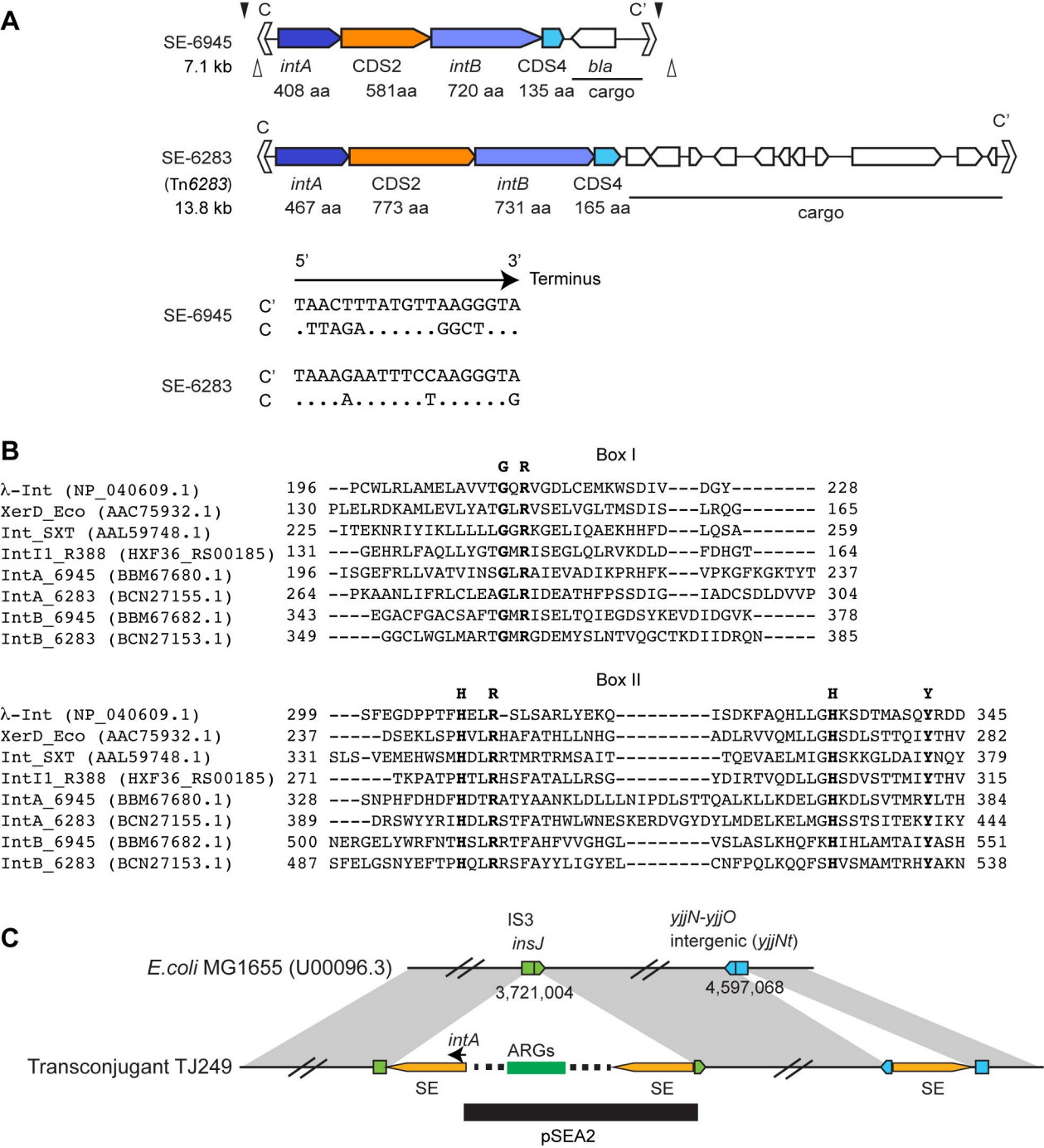
S6 Fig. PCR detection of SE-6945 circle in *V. alfacsensis* strain 04Ya249 and *E. coli* strain LN52. (A) PCR detection of chromosomal *attL*, *attR*, *attS* in 04Ya249. (B) Copy number of *attS* in pSEA2 free *V. alfacsensis* strain LN95 by qPCR. Copy number is represented as relative copy number. Primers used are shown in Table 2. (C) PCR detection of *attS* in pSEA2 free *E. coli* strain LN52.

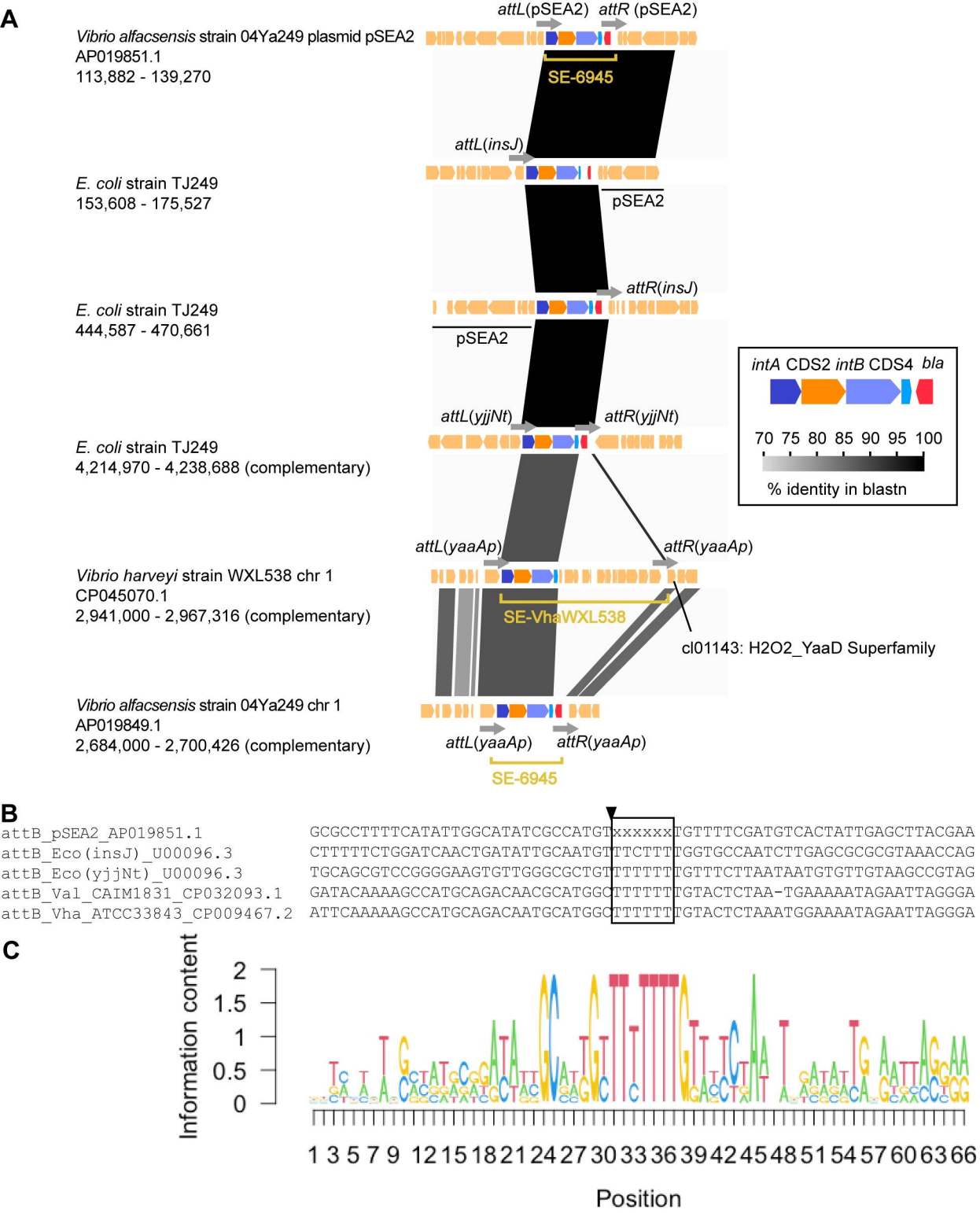
S7 Fig. *attL* sequences formed in the SE-6283 insertion regions. Sequences are derived from strain 04Ya108 (AP024165.1, AP024167.1), *E. coli* transconjugant TJ108W0 sequenced in our previous study [29], and *Vibrio* strains carrying SE-6283-like elements listed in Fig 5. Sequences in red are a part of the mobile DNA unit of SE. Incorporation of the 6 bp underlined sequence of TJ108W0 and pSEA1 into *attS* was demonstrated in the previous study.

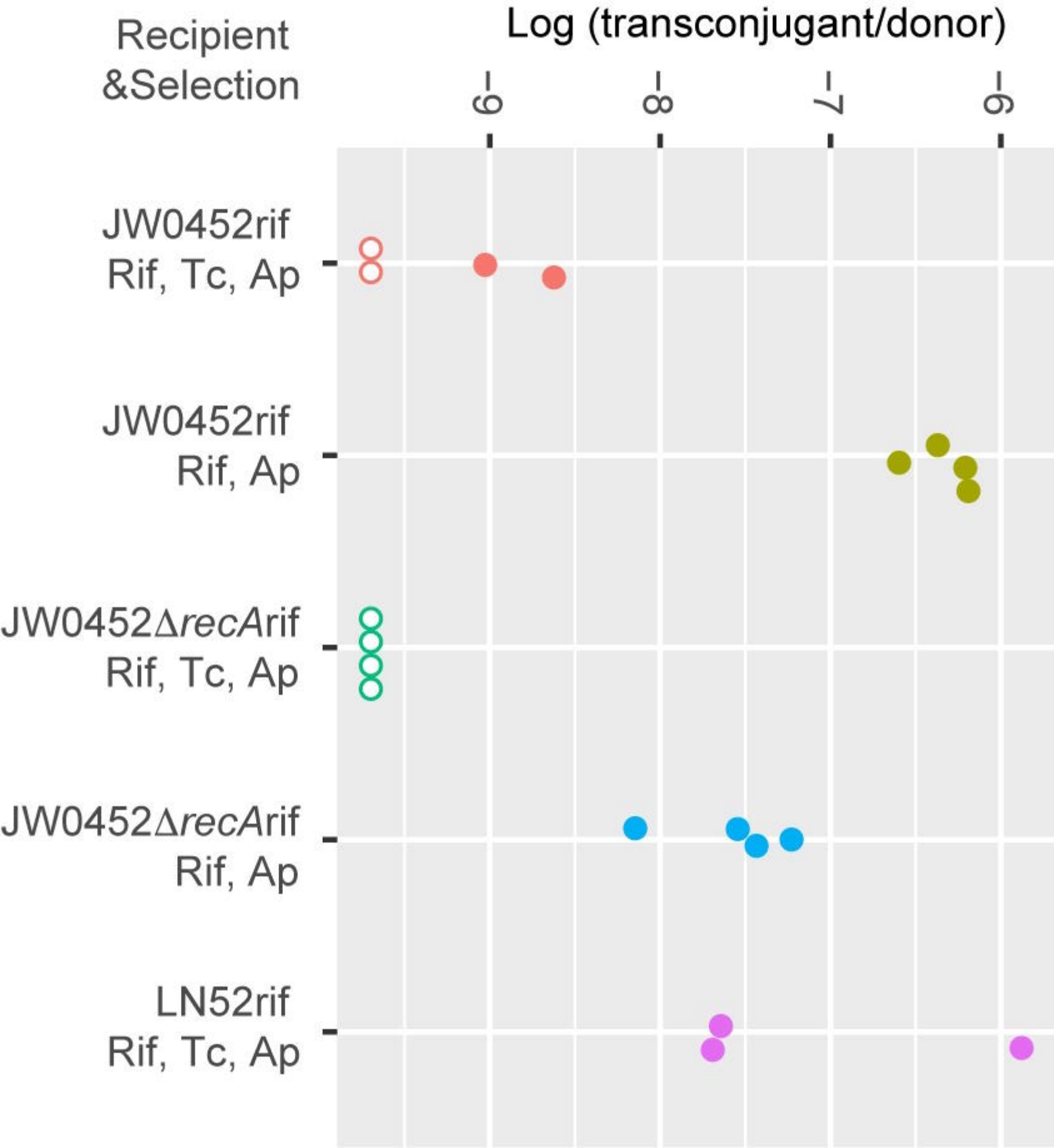
S8 Fig. *attR* sequences formed in the SE-6283 insertion regions. Sequences are derived from genomes listed in S7 Fig. Regarding SE-6283, the 6 bp underlined sequences in *attR* originate from inserting SE circle, and they are not incorporated into *attS* upon SE circularization.

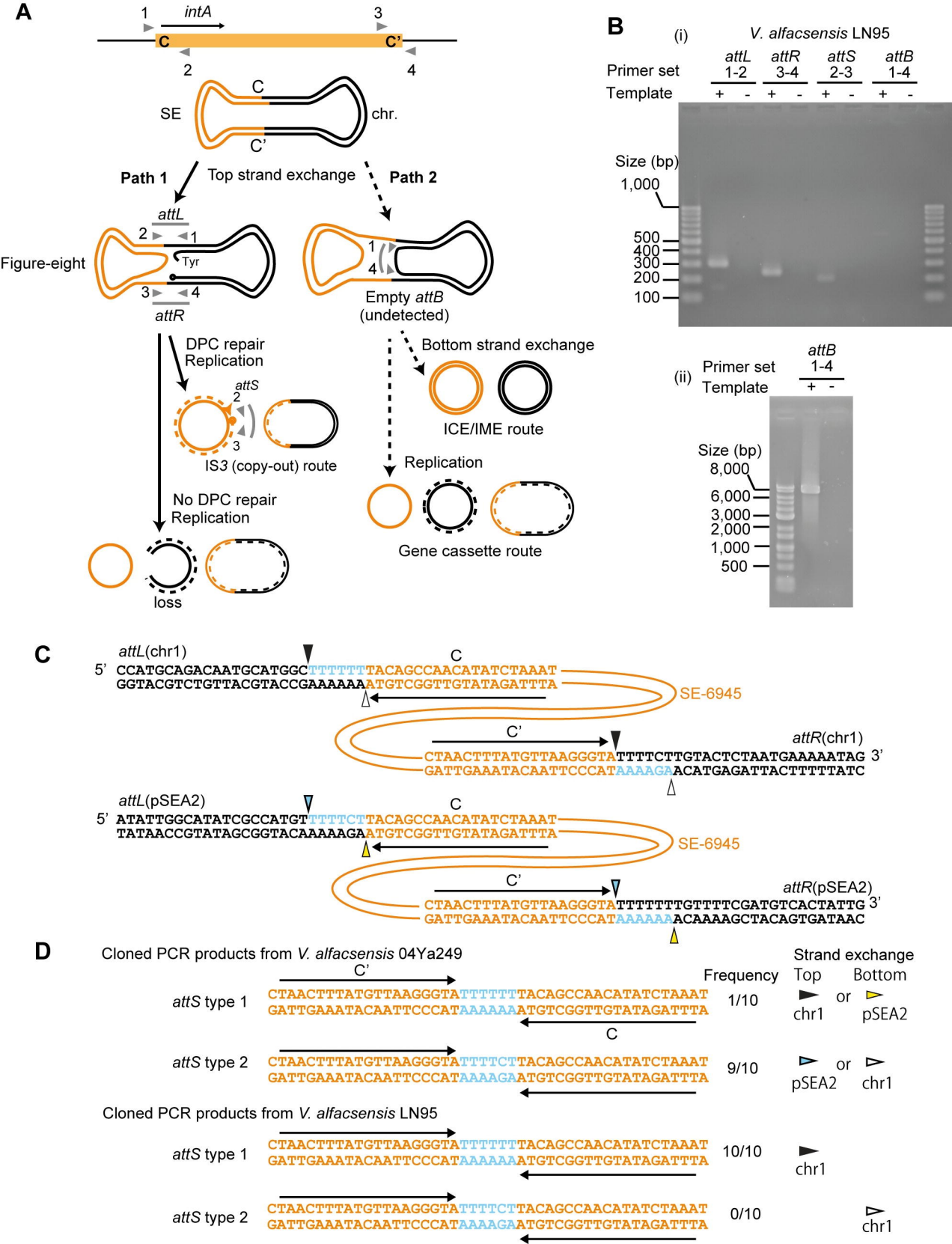
S1 Table. MIC of transconjugant strains

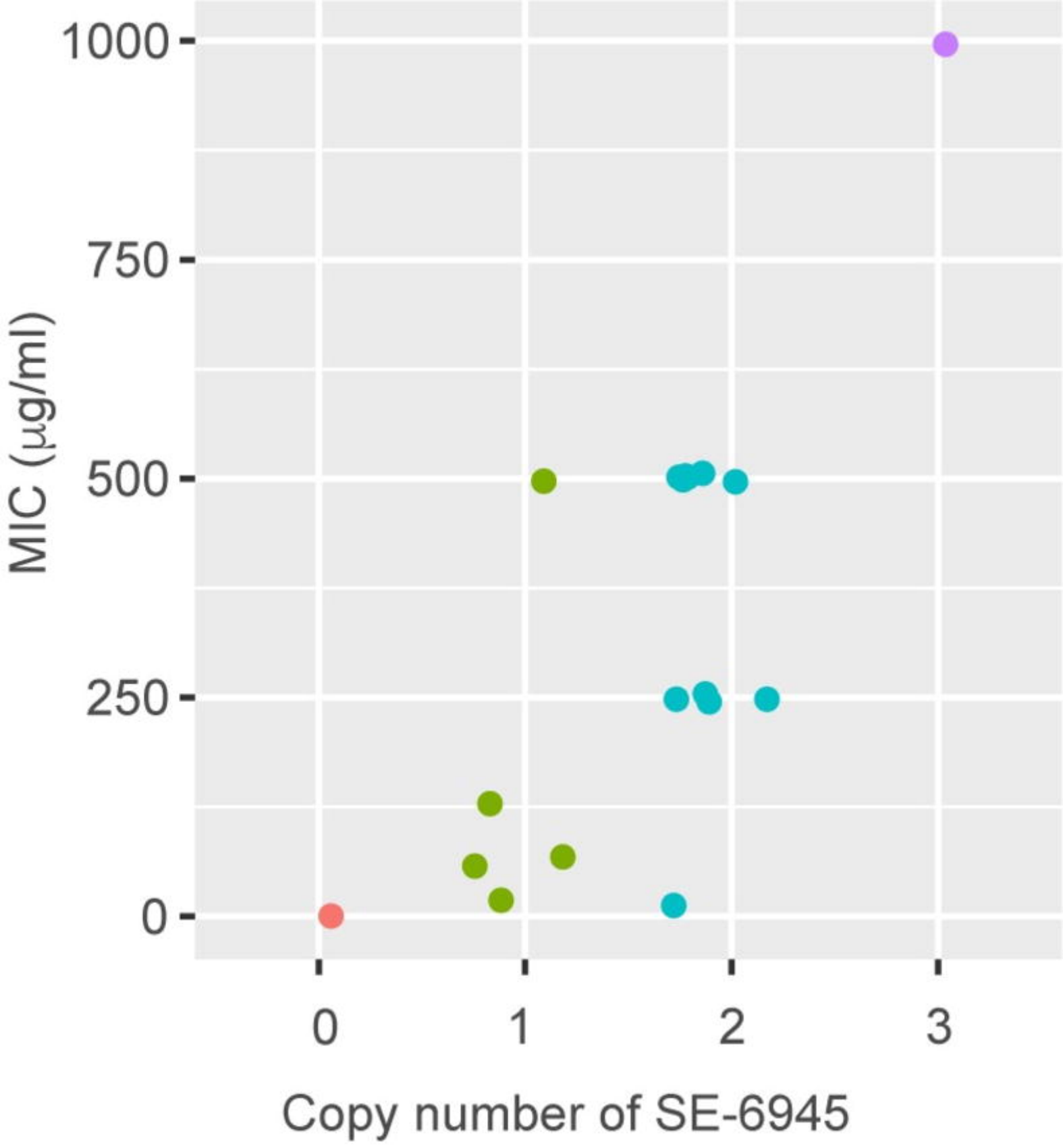
S2 Table. Data availability

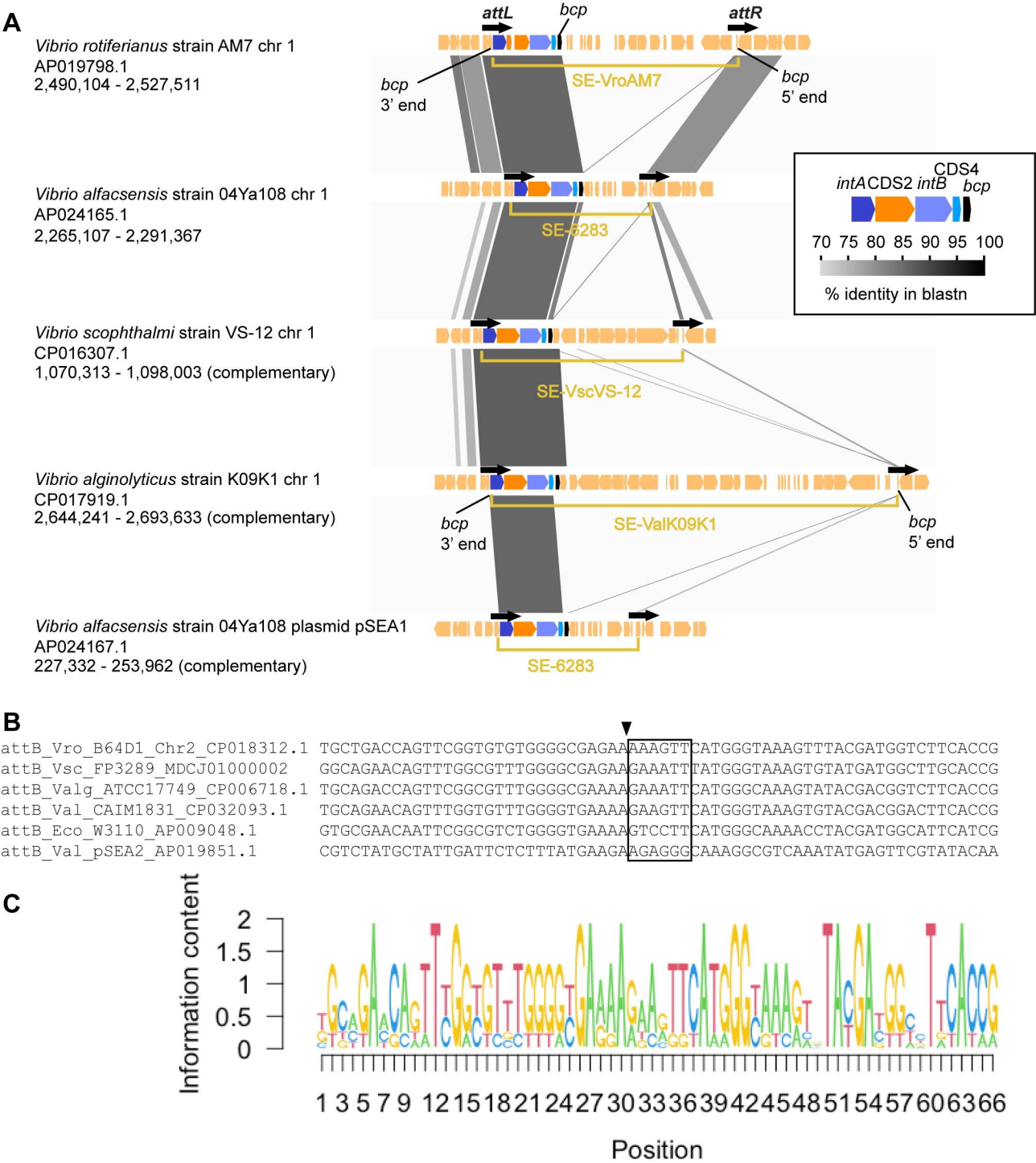

A
C
C'
SE-6945
7.1 kb
intA
408 aa
CDS2
581aa
intB
720 aa
CDS4
135 aa
bla
cargo
C
C'
SE-6283
(Tn6283)
13.8 kb
intA
467 aa
CDS2
773 aa
intB
731 aa
CDS4
165 aa
cargo
5'
3'
Terminus
SE-6945
C' TAACTTTATGTTAAGGGTA
C .TTAGA.....GGCT...
SE-6283
C' TAAAGAATTTCCAAGGGTA
C ....A.....T.....G
B
Box I
G R
λ-Int (NP_040609.1) 196 --PCWLRLAMELAVVTGQRVGDLCEMKWSDIV---DGY-------- 228
XerD_Eco (AAC75932.1) 130 PLELRDKAMLEVLYATGLRVSELVGLTMSDIS---LRQG------- 165
Int_SXT (AAL59748.1) 225 -ITEKNRIYIKLLLLLGGRKGELIQAEKHHFD---LQSA------- 259
IntI1_R388 (HXF36_RS00185) 131 ---GEHRLFAQLLYGTGMRISEGLQLRVKDLD---FDHGT------ 164
IntA_6945 (BBM67680.1) 196 -ISGEFRLLVATVINSGLRAIEVADIKPRHFK---VPKGFKGKTYT 237
IntA_6283 (BCN27155.1) 264 --PKAANLIFRLCLEAGLRIDEATHFPSSDIG---IADCSDLDVVP 304
IntB_6945 (BBM67682.1) 343 ----EGACFGACSAFTGMRISELTQIEGDSYKEVDIDGVK------ 378
IntB_6283 (BCN27153.1) 349 -----GGCLWGLMARTGMRGDEMYSLNTVQGCTKDIIDRQN----- 385
Box II
H R H Y
λ-Int (NP_040609.1) 299 ---SFEGDPPTFHELR-SLSARLYEKQ---------ISDKFAQHLLGHKSDTMASQYRDD 345
XerD_Eco (AAC75932.1) 237 -----DSEKLSPHVLRHAFATHLLNHG---------ADLRVVQMLLGHSDLSTTQIYTHV 282
Int_SXT (AAL59748.1) 331 SLS-VEMEHWSMHDLRRTMRTRMSAIT----------TQEVAELMIGHSKKGLDAIYNQY 379
IntI1_R388 (HXF36_RS00185) 271 ------TKPATPHTLRHSFATALLRSG---------YDIRTVQDLLGHSDVSTTMIYTHV 315
IntA_6945 (BBM67680.1) 328 ---SNPHFDHDFHDTRATYAANKLDLLLNIPDLSTTQALKLLKDELGHKDLSVTMRYLTH 384
IntA_6283 (BCN27155.1) 389 ----DRSWYYRIHDLRSTFATHWLWNESKERDVGYDYLMDELKELMGHSSTSITEKYIKY 444
IntB_6945 (BBM67682.1) 500 NERGELYWRFNTHSLRRTFAHFVVGHGL--------VSLASLKHQFKHIHLAMTAIYASH 551
IntB_6283 (BCN27153.1) 487 SFELGSNYEFTPHQLRRSFAYYLIGYEL--------CNFPQLKQQFSHVSMAMTRHYAKN 538
C
IS3
insJ
yjjN-yjjO
intergenic (yjjNt)
E.coli MG1655 (U00096.3)
3,721,004
4,597,068
intA
ARGs
Transconjugant TJ249
SE
SE
SE
pSEA2


A
attL(pSEA2) attR (pSEA2)
Vibrio alfacsensis strain 04Ya249 plasmid pSEA2
AP019851.1
113,882 - 139,270
SE-6945
attL(insJ)
E. coli strain TJ249
153,608 - 175,527
pSEA2
attR(insJ)
E. coli strain TJ249
444,587 - 470,661
pSEA2
attL(yjjNt)
attR(yjjNt)
E. coli strain TJ249
4,214,970 - 4,238,688 (complementary)
intA CDS2 intB CDS4 bla
70 75 80 85 90 95 100
% identity in blastn
attL(yaaAp)
attR(yaaAp)
Vibrio harveyi strain WXL538 chr 1
CP045070.1
2,941,000 - 2,967,316 (complementary)
SE-VhaWXL538
cl01143: H2O2_YaaD Superfamily
Vibrio alfacsensis strain 04Ya249 chr 1
AP019849.1
2,684,000 - 2,700,426 (complementary)
attL(yaaAp) attR(yaaAp)
SE-6945
B
attB_pSEA2_AP019851.1 GCGCCTTTTCATATTGGCATATCGCCATGTxxxxxxTGTTTTCGATGTCACTATTGAGCTTACGAA
attB_Eco(insJ)_U00096.3 CTTTTTCTGGATCAACTGATATTGCAATGTTTCTTTTGGTGCCAATCTTGAGCGCGCGTAAACCAG
attB_Eco(yjjNt)_U00096.3 TGCAGCGTCCGGGGAAGTGTTGGGCGCTGTTTTTTTTGTTTCTTAATAATGTGTTGTAAGCCGTAG
attB_Val_CAIM1831_CP032093.1 GATACAAAAGCCATGCAGACAACGCATGGCTTTTTTTGTACTCTAA-TGAAAAATAGAATTAGGGA
attB_Vha_ATCC33843_CP009467.2 ATTCAAAAAGCCATGCAGACAATGCATGGCTTTTTTTGTACTCTAAATGGAAAATAGAATTAGGGA
C
Information content
2
1.5
1
0.5
0
1 3 5 7 9 12 15 18 21 24 27 30 33 36 39 42 45 48 51 54 57 60 63 66
Position


Log (transconjugant/donor)
Recipient
&Selection
−9
−8
−7
−6
JW0452rif
Rif, Tc, Ap
JW0452rif
Rif, Ap
JW0452Δ*recA*rif
Rif, Tc, Ap
JW0452Δ*recA*rif
Rif, Ap
LN52rif
Rif, Tc, Ap


A
intA
1
3
C
C'
2
4
SE
chr.
Path 1
Top strand exchange
Path 2
attL
Figure-eight
Tyr
attR
Empty attB
(undetected)
DPC repair
Replication
attS
IS3 (copy-out) route
Bottom strand exchange
ICE/IME route
Replication
Gene cassette route
No DPC repair
Replication
loss
B
(i)
V. alfacsensis LN95
Primer set
attL 1-2
attR 3-4
attS 2-3
attB 1-4
Template
Size (bp)
1,000
500
400
300
200
100
(ii)
attB 1-4
8,000
6,000
3,000
2,000
1,000
500
C
attL(chr1)
5' CCATGCAGACAATGCATGGCTTTTTTTACAGCCAACATATCTAAAT
GGTACGTCTGTTACGTACCGAAAAAAATGTCGGTTGTATAGATTTA
SE-6945
attR(chr1)
CTAACTTTATGTTAAGGGTATTTTCTTGTACTCTAATGAAAAATAG 3'
GATTGAAATACAATTCCCATAAAAGAACATGAGATTACTTTTTATC
attL(pSEA2)
5' ATATTGGCATATCGCCATGTTTTTCTTACAGCCAACATATCTAAAT
TATAACCGTATAGCGGTACAAAAAGAATGTCGGTTGTATAGATTTA
SE-6945
attR(pSEA2)
CTAACTTTATGTTAAGGGTATTTTTTTGTTTTCGATGTCACTATTG 3'
GATTGAAATACAATTCCCATAAAAAAACAAAAGCTACAGTGATAAC
D
Cloned PCR products from V. alfacsensis 04Ya249
Frequency
Strand exchange
Top
Bottom
attS type 1
CTAACTTTATGTTAAGGGTATTTTTTTACAGCCAACATATCTAAAT
GATTGAAATACAATTCCCATAAAAAAATGTCGGTTGTATAGATTTA
1/10
chr1
or
pSEA2
attS type 2
CTAACTTTATGTTAAGGGTATTTTCTTACAGCCAACATATCTAAAT
GATTGAAATACAATTCCCATAAAAGAATGTCGGTTGTATAGATTTA
9/10
pSEA2
or
chr1
Cloned PCR products from V. alfacsensis LN95
attS type 1
CTAACTTTATGTTAAGGGTATTTTTTTACAGCCAACATATCTAAAT
GATTGAAATACAATTCCCATAAAAAAATGTCGGTTGTATAGATTTA
10/10
chr1
attS type 2
CTAACTTTATGTTAAGGGTATTTTCTTACAGCCAACATATCTAAAT
GATTGAAATACAATTCCCATAAAAGAATGTCGGTTGTATAGATTTA
0/10
chr1


1000
750
500
250
0
MIC (μg/ml)
0
1
2
3
Copy number of SE-6945


A
attL
bcp
attR
Vibrio rotiferianus strain AM7 chr 1
AP019798.1
2,490,104 - 2,527,511
bcp
3' end
SE-VroAM7
bcp
5' end
Vibrio alfacsensis strain 04Ya108 chr 1
AP024165.1
2,265,107 - 2,291,367
SE-6283
intA
CDS2
intB
CDS4
bcp
70 75 80 85 90 95 100
% identity in blastn
Vibrio scophthalmi strain VS-12 chr 1
CP016307.1
1,070,313 - 1,098,003 (complementary)
SE-VscVS-12
Vibrio alginolyticus strain K09K1 chr 1
CP017919.1
2,644,241 - 2,693,633 (complementary)
bcp
3' end
SE-ValK09K1
bcp
5' end
Vibrio alfacsensis strain 04Ya108 plasmid pSEA1
AP024167.1
227,332 - 253,962 (complementary)
SE-6283
B
attB_Vro_B64D1_Chr2_CP018312.1 TGCTGACCAGTTCGGTGTGTGGGGCGAGAAAAAGTTCATGGGTAAAGTTTACGATGGTCTTCACCG
attB_Vsc_FP3289_MDCJ01000002 GGCAGAACAGTTTGGCGTTTGGGGCGAGAAGAAATTTATGGGTAAAGTGTATGATGGCTTGCACCG
attB_Valg_ATCC17749_CP006718.1 TGCAGACCAGTTCGGCGTTTGGGGCGAAAAGAAATTCATGGGCAAAGTATACGACGGTCTTCACCG
attB_Val_CAIM1831_CP032093.1 TGCAGAACAGTTTGGTGTTTGGGGTGAAAAGAAGTTCATGGGTAAAGTGTACGACGGACTTCACCG
attB_Eco_W3110_AP009048.1 GTGCGAACAATTCGGCGTCTGGGGTGAAAAGTCCTTCATGGGCAAAACCTACGATGGCATTCATCG
attB_Val_pSEA2_AP019851.1 CGTCTATGCTATTGATTCTCTTTATGAAGAAGAGGGCAAAGGCGTCAAATATGAGTTCGTATACAA
C
Information content
2
1.5
1
0.5
0
1 3 5 7 9 12 15 18 21 24 27 30 33 36 39 42 45 48 51 54 57 60 63 66
Position