



**HAL**  
open science

## Chromosomal Integrans are Genetically and Functionally Isolated Units of Genomes

Paula Blanco, Filipa Trig da Roza, Laura Toribio-Celestino, Lucía García-Pastor, Niccolò Caselli, Francisco Ojeda, Baptiste Darracq, Ester Vergara, Álvaro San Millán, Ole Skovgaard, et al.

► **To cite this version:**

Paula Blanco, Filipa Trig da Roza, Laura Toribio-Celestino, Lucía García-Pastor, Niccolò Caselli, et al.. Chromosomal Integrans are Genetically and Functionally Isolated Units of Genomes. 2023. pasteur-04305666

**HAL Id: pasteur-04305666**

**<https://pasteur.hal.science/pasteur-04305666>**

Preprint submitted on 24 Nov 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

# 1 Chromosomal Integrons are Genetically and Functionally 2 Isolated Units of Genomes

3

4 Paula Blanco<sup>1,2</sup>, Filipa Trigo da Roza<sup>1,2</sup>, Laura Toribio-Celestino<sup>3</sup>, Lucía García-Pastor<sup>1,2</sup>,  
5 Niccolò Caselli<sup>4</sup>, Francisco Ojeda<sup>1,2</sup>, Baptiste Darracq<sup>6,7</sup>, Ester Vergara<sup>1,2</sup>, Álvaro San  
6 Millán<sup>3</sup>, Ole Skovgaard<sup>5</sup>, Didier Mazel<sup>6</sup>, Céline Loot<sup>6\*</sup> and José Antonio Escudero<sup>1,2\*</sup>.

7 \*Correspondence: [jaescudero@ucm.es](mailto:jaescudero@ucm.es) and [celine.loot@pasteur.fr](mailto:celine.loot@pasteur.fr)

8 <sup>1</sup> Molecular Basis of Adaptation. Departamento de Sanidad Animal. Universidad  
9 Complutense de Madrid, Madrid, Spain.

10 <sup>2</sup> VISAVET Health Surveillance Centre, Universidad Complutense de Madrid, Madrid,  
11 Spain.

12 <sup>3</sup> Department of Microbial Biotechnology, Centro Nacional de Biotecnología–CSIC,  
13 Madrid, Spain.

14 <sup>4</sup> Departamento de Química Física. Universidad Complutense de Madrid, Madrid, Spain.

15 <sup>5</sup> Department of Science and Environment, Roskilde University, Roskilde, Denmark.

16 <sup>6</sup> Institut Pasteur, Université Paris Cité, CNRS UMR3525, Unité Plasticité du Génome  
17 Bactérien, 75015 Paris, France.

18 <sup>7</sup>Sorbonne Université, ED515, F-75005, Paris, France.

19

20

21 ORCID:

22 Paula Blanco: <https://orcid.org/0000-0003-0769-2233>

23 Filipa Trigo da Roza: <https://orcid.org/0000-0002-8345-0415>

24 Lucía García-Pastor: <https://orcid.org/0000-0001-7462-1860>

25 Laura Toribio-Celestino: <https://orcid.org/0000-0002-8848-1691>

26 Niccolò Caselli: <https://orcid.org/0000-0003-2609-3834>

27 Baptiste Darracq: <https://orcid.org/0000-0002-6257-4048>

28 Ester Vergara: <https://orcid.org/0000-0003-1323-4749>

29 Álvaro San Millán: <https://orcid.org/0000-0001-8544-0387>

30 Ole Skovgaard: <https://orcid.org/0000-0002-4860-1847>

31 Didier Mazel: <https://orcid.org/0000-0001-6482-6002>

32 Céline Loot: <https://orcid.org/0000-0003-3774-249X>

33 José Antonio Escudero: <https://orcid.org/0000-0001-8552-2956>

34

35

36

37

38

39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58

## ABSTRACT

Integrans are genetic elements that increase the evolvability of bacteria by capturing new genes and stockpiling them in arrays. Sedentary chromosomal integrans (SCIs), can be massive and highly stabilized structures encoding hundreds of genes, whose function remains generally unknown. SCIs have co-evolved with the host for aeons and are highly intertwined with their physiology from a mechanistic point of view. But, paradoxically, other aspects, like their variable content and location within the genome, suggest a high genetic and functional independence. In this work, we have explored the connection of SCIs to their host genome using as a model the Superintegron (SI), a 179-cassette long SCI in the genome of *Vibrio cholerae* N16961. We have relocated and deleted the SI using SeqDelTA, a novel method that allows to counteract the strong stabilization conferred by toxin-antitoxin systems within the array. We have characterized in depth the impact in *V. cholerae*'s physiology, measuring fitness, chromosome replication dynamics, persistence, transcriptomics, phenomics and virulence. The deletion of the SI did not produce detectable effects in any condition, proving that -despite millions of years of co-evolution-, SCIs are genetically and functionally isolated units of genomes.

## 59 INTRODUCTION

60 Integrons are genetic elements that allow bacteria to adapt swiftly to changing  
61 environments<sup>1-3</sup>. They recruit new genes encoded in small mobile genetic elements called  
62 integron cassettes (IC) (or gene cassettes), stocking them in an array, to form a genetic  
63 memory of adaptive functions. According to the needs of their hosts, integrons can also  
64 modulate cassette expression by reshuffling their order within the array<sup>4,5</sup>. The integron  
65 platform comprises the gene encoding the integrase (*intI*) -the recombinase that governs  
66 all reactions-, the integration site where cassettes are incorporated (*attI* site), and two  
67 promoters in opposite orientation that drive the expression of the integrase ( $P_{int}$ ) and that  
68 of cassettes in the array ( $P_c$ ), which are generally promoterless. Integrons are well known  
69 for their role in the rise of multidrug resistance and the spread of resistance genes among  
70 clinically relevant Gram-negative species<sup>6-8</sup>. Encoded on plasmids, these Mobile  
71 Integrons (MIs) act as vehicles for more than 170 resistance genes against the most  
72 relevant families of antibiotics<sup>9,10</sup>. Yet, MIs are only a small subset of all integrons that  
73 have been mobilized from the chromosomes of environmental bacteria by transposons.  
74 The vast majority of integrons in nature are Sedentary Chromosomal Integrons (SCIs),  
75 which are found in a variety of phyla and have an ancient origin<sup>11-13</sup>. These SCIs are a  
76 highly variable region of the genome, representing a hotspot for genetic diversity<sup>14</sup>. On  
77 the antipodes of what is observed for MIs, the functions encoded in sedentary  
78 chromosomal integrons are generally not related to antimicrobial resistance and are  
79 mostly unknown<sup>15</sup>. The colossal repertoire of cassettes in the environment holds a high  
80 potential for biotechnology that is yet to be explored<sup>16</sup>.

81 The paradigm of SCI is the Superintegron (SI)<sup>17</sup>, a massive structure encoded in  
82 the secondary chromosome of *Vibrio cholerae*, the causative agent of cholera disease. In  
83 strain N16961, where it was first described, the SI is 126 kb long, embodying 3% of the  
84 genome of this bacterium, and containing 179 integron cassettes. Although the model of  
85 integron predicts that the SI should be mostly silent, a recent RNAseq analysis shows that  
86 many cassettes are expressed at biologically relevant levels (Blanco et al. under review).  
87 The functions of cassettes are mostly cryptic, except for the notable exception of 19  
88 cassettes that encode toxin-antitoxin (TA) systems<sup>18,19</sup> and a chloramphenicol resistance  
89 gene<sup>20</sup>. TA cassettes are scattered along the SI and serve to stabilize the array through a  
90 post-segregational killing mechanism<sup>21,22</sup>. Despite being the best studied SCI, the  
91 knowledge on the Superintegron is limited for two reasons: first, the long cassette array  
92 interferes with the genetic tools commonly used in the field to deliver recombination

93 experiments that allow to characterize the system; and second, its integrase cannot be  
94 studied in heterologous genetic backgrounds because it seems to require additional host-  
95 factors<sup>23–25</sup>.

96 Sedentary Chromosomal Integrations have co-evolved with the genome of the host  
97 for millions of years<sup>26,27</sup>. Indeed, large SCIs like the superintegron are ubiquitous in  
98 *Vibrio* species, and the phylogenetic signal of integrases mimics that of the species<sup>1</sup>. This  
99 suggests that their acquisition predates the radiation of the genus, an event that occurred  
100 more than 300 million years ago<sup>28</sup>. The timescale of this co-evolution process has allowed  
101 for the intertwining between host physiology and integron activity at many different  
102 levels. A blatant example is that integrase expression is under the control of the host's  
103 SOS response. Yet many other aspects are rather more subtle and represent better the  
104 depth of such connections<sup>29</sup>. Examples of this are i) the balanced interplay between the  
105 integrase and single strand binding (SSB) proteins of the host on *attC sites*, that regulate  
106 *attC* site folding to avoid its structuring unless needed for recombination<sup>30,31</sup>, or ii) the  
107 fact that SCIs have a specific orientation related to the origin of replication, to maximize  
108 their carrying capacity<sup>22,32</sup>. Yet, a striking observation is that some aspects of SCIs remain  
109 extremely plastic, such as their content and their location -despite their size, SCIs can be  
110 found in either chromosome depending on the species of *Vibrio*. Hence, it remains unclear  
111 whether the intertwining between host and integron is restricted to the functioning of the  
112 platform and the mechanistic of recombination ( $P_{int}$  and  $P_c$  promoters and *attC* sites), or  
113 if it is extensive to the array of functions encoded in the variable part. In other words, we  
114 ignore if cassettes are streamlined to act as plug and play add-ons to the pool of functions  
115 encoded in the genome without interacting with it, or whether they can modify or interfere  
116 with the host physiology in any other way.

117 In this work we have addressed the question of whether chromosomal integrations  
118 are genetically and functionally isolated units of the genome by relocating and deleting  
119 the Superintegron from *V. cholerae*. To overcome the high stabilization of the structure  
120 by the TAs, we have designed an approach called SeqDelTA (Sequential Deletion of  
121 Toxin Antitoxin Systems) that exploits natural competence and homologous  
122 recombination to deliver a set of 15 consecutive allelic replacements. At each step, a toxin  
123 is inactivated without altering the cognate antitoxin, deleting a piece of the SI. After the  
124 last allelic replacement, we produce the scarless deletion of the SI through a counter-  
125 selectable suicide vector. The resulting strain has been sequenced, and unintended  
126 mutations located elsewhere in the genome have been corrected. Incidentally, this strain

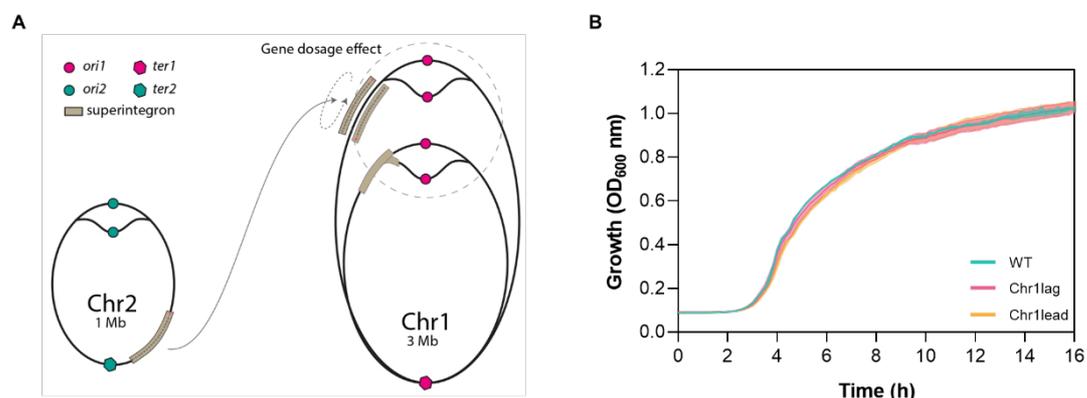
127 represents a milestone in the study of chromosomal integrons, since it provides, for the  
128 first time, the possibility of delivering experimental studies using common tools -without  
129 interference from the Superintegron. To understand how the deletion of the SI affects *V.*  
130 *cholerae*, we have characterized in depth this strain from a variety of perspectives,  
131 including chromosome replication dynamics, fitness, persistence, transcriptomics,  
132 phenomics and virulence assays. We observe no significant variations in the physiology  
133 of the  $\Delta$ SI strain compared to the wild type. We conclude that sedentary chromosomal  
134 integrons are functionally isolated units of the genome, despite eons of co-evolution with  
135 their hosts. This has strong implications in the type of functions one can expect to find in  
136 integron cassettes.

137

## 138 RESULTS

### 139 Genomic position of the SI does not alter *V. cholerae* growth.

140 The genomic location of SCIs in *Vibrio* species is remarkably plastic given their  
141 size. We ignore whether SCI relocations in different species have been innocuous for the  
142 cell or have entailed complex changes that required subsequent adaptation. To test this,  
143 here we used the  $\Phi$  HK recombination machinery to relocate the SI. We moved it from  
144 the right arm of chromosome 2 -close to the replication terminus-, to chromosome 1, near  
145 the origin of replication, where we placed it in both orientations. This relocation could  
146 have a potential impact in cell growth because of the higher gene dosage effect near the  
147 Ori of chromosome 1. Additionally, it could also generate conflicts between transcription  
148 and translation machineries depending on the orientation of the array. We measured  
149 growth of both mutants and could not observe differences with the WT, suggesting that  
150 SCIs are genetically independent units of the genome.

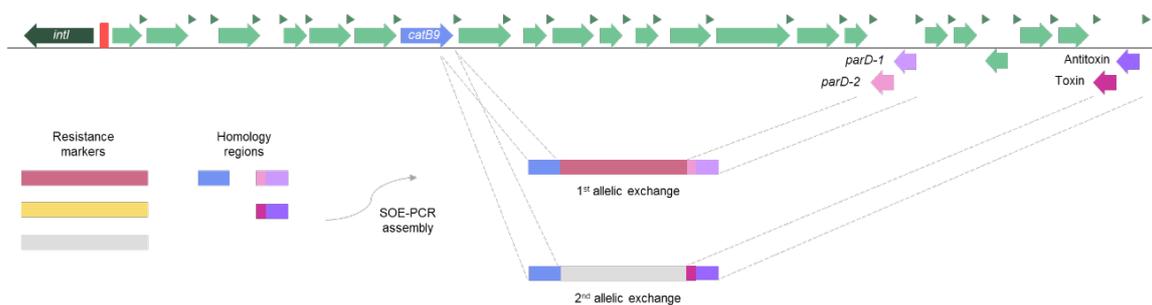


151 **Figure 1.** (A) Scheme of the relocation of the SI from Chr2 to Chr1, where it is inserted in both  
152 orientations (modified from<sup>33</sup>). (B) Growth curves of WT and relocated-SI strains (Chr1lag and  
153 Chr1lead). No significant growth differences are observed.

154

## 155 **SeqDelTA, a tool to erase chromosomal integrons.**

156 To further assess the level of genetic and functional independence, we sought to  
157 delete the superintegron. The deletion of the SI in *V. cholerae* is a longstanding milestone  
158 in the field, that should allow to deliver the kind of experiments that have advanced the  
159 field in the last years using the Class 1 integron as a model. Previous efforts to delete the  
160 SI were hampered by unknown TA systems and by a one-step deletion approach. The  
161 recent full characterization of TA systems in the SI<sup>18,19</sup> has enabled the development of  
162 SeqDelTA, a multi-step approach that allows to sequentially inactivate TA systems while  
163 deleting, at each step, the cassette cargo between two TAs (Figure 2).



164 **Figure 2.** Schematic representation of two successive allelic replacements using SeqDelTA. The  
165 left homology region is conserved during the deletion process, allowing to recycle the resistance  
166 markers at each step. The right homology region is redesigned at each step to sequentially inactivate  
167 the TA modules. This allows to inactivate a TA system by knocking the toxin but leaving the  
168 antitoxin intact. A total of 18 sequential replacements were performed. The last deletion step was  
169 carried out using a suicide plasmid.

170

171 Briefly, we take advantage of *V. cholerae*'s natural competence to provide linear  
172 DNA fragments in which a resistance marker is flanked by left and right homology  
173 regions (LHR and RHR). These HRs contain the sequence of specific parts of the  
174 superintegron to produce allelic replacements. The superintegron spans from gene  
175 VCA0291 -the integrase- to VCA0506 (nucleotides 309,750 to 435,034). In the design of  
176 the allelic exchanges, we used a conserved LHR targeting the *catB9* chloramphenicol  
177 resistance cassette in 9<sup>th</sup> position of the array (VCA0300). The rationale for this choice is  
178 that, since the strain is susceptible to chloramphenicol, this region of the SI is expected  
179 to be silent, being therefore a good starting point in the 5' end of the array, free of  
180 interference with the P<sub>c</sub> promoter. In contrast, the RHR changes at each step of the  
181 deletion, targeting the next TA downstream the array. RHRs were designed to produce  
182 crossovers that maintained the antitoxin of each system intact while deleting the toxin  
183 gene. Keeping a common LHR while advancing the deletion through changes in RHRs

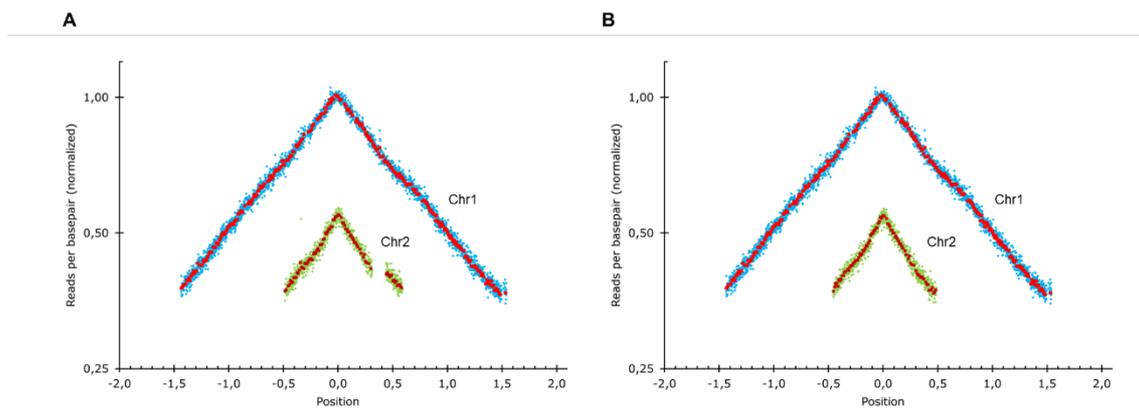
184 allowed the removal of the resistance marker introduced in the previous step, easily  
185 alternating resistance markers. This scheme for SeqDeITA was generally kept constant.  
186 In the cases in which the orientation and/or the order of the genes in the TA system were  
187 not favorable, we used alternative LHRs and maintained two resistance markers in  
188 consecutive allelic replacements. Then, a third replacement allowed to remove both  
189 markers. After each allelic exchange, several colonies growing on the appropriate  
190 antibiotic were verified through PCR and checked phenotypically for the loss of the  
191 previous marker where applicable. The growth curve of PCR-verified clones was then  
192 analyzed to avoid the hitchhiking of mutations with important deleterious effects. A clone  
193 without growth defects was chosen to continue the deletion process. After the last allelic  
194 replacement, we had erased all cassettes from VCA0300 (*catB9*) to VCA0503 (the last  
195 toxin). We then used a counter-selectable integrative vector (pMP7<sup>34</sup>) to deliver the  
196 scarless deletion of the remainder of the SI -including the integrase, the first 8 and the last  
197 2 cassettes from the array, as well as a transposase inserted immediately downstream the  
198 last cassette (VCA0291 to VCA0508). A PCR from the borders of the integron confirmed  
199 the deletion of the SI.

200

## 201 **Genome Sequencing and Marker Frequency Analysis**

202 We performed genome sequencing on the obtained *V. cholerae*  $\Delta$ SI using  
203 Illumina short reads and confirmed the scarless deletion of the SI. We also used this  
204 approach to perform Marker Frequency Analysis (MFA), a technique that allows to assess  
205 the dynamics of chromosome replication. MFA compares the abundance of reads across  
206 the genome between exponential and stationary growth conditions. In fast growing  
207 bacteria as *V. cholerae*, this comparison reveals the gene dosage gradient from the origin  
208 to the terminus of replication as an inverted “V”. MFA allows to observe the relative  
209 timing of replication of both chromosomes, and the speed of replication across the  
210 genome<sup>35</sup>. As shown in Figure 3, no major alterations in chromosome replication  
211 dynamics could be observed, compared to the WT strain.

212



213 **Figure 3.** (A) Marker frequency analysis (MFA) of the two chromosomes of *V. cholerae*  $\Delta$ SI  
214 mapped against *V. cholerae* WT genome and (B) against *V. cholerae*  $\Delta$ SI genome. The gap on  
215 chromosome 2 in panel A reflects the superintegrons deletion. The genome position is represented  
216 relative to the *oriC* (set to 0). Blue and green rhombus represent the average of 1 kbp windows,  
217 and red and garnet squares the average of 10 kbp windows in Chr1 and Chr 2, respectively.  
218

219 MFA can reveal the presence of large chromosomal rearrangements (such as  
220 inversions), through inconsistencies in the slope of the inverted V shape, but lacks  
221 resolution for smaller rearrangements like IS movements. We therefore sequenced *V.*  
222 *cholerae*  $\Delta$ SI using MinION long reads to detect the presence of any rearrangement that  
223 could go undetected by the assembly of short reads. Combination of small and long reads  
224 allowed to obtain a high-quality genome sequence in which we could rule out genomic  
225 rearrangements while detecting the presence of unintended single nucleotide  
226 polymorphisms (SNPs). The *V. cholerae*  $\Delta$ SI strain bore three unintended mutations in  
227 *rocS* (VC0653), *cry2* (VC01392) and *rpoS* (VC0534) genes. RocS is a diguanylate  
228 cyclase involved in the smooth-to-rugose switch in our strain. A101 had a single base  
229 deletion that altered the reading frame of *rocS*. We could observe the loss of motility and  
230 increase in biofilm formation described in *rocS*<sup>-</sup> mutants<sup>36,37</sup>. We were able to map the  
231 emergence of this mutation to the third deletion step of SeqDelTA. Interestingly, 12  
232 clones from two independent experiments had been kept from this deletion step, of which  
233 11 had mutations in *rocS*. We found 8 different types of mutations among them, pointing  
234 to an unusually high mutation rate. Indeed, the sequence of *rocS* is rich in homopolymeric  
235 tracts, and is likely to act as a contingency locus, allowing for a frequent smooth-to-rugose  
236 switch<sup>38</sup>. *cry2* presented a non-synonymous (Pro129Lys) mutation. This gene encodes a  
237 protein annotated as a photolyase, for which no experimental evidence of its function  
238 could be found. RpoS is an alternative sigma factor related to stress response and entry  
239 in stationary phase. It is a major regulator of cell physiology and influences many relevant

240 phenotypes, such as oxidative stress, natural competence, motility, or colonization<sup>39-41</sup>.  
241 *V. cholerae*  $\Delta$ SI harbored a single-base insertion in *rpoS* giving rise to a frameshift.

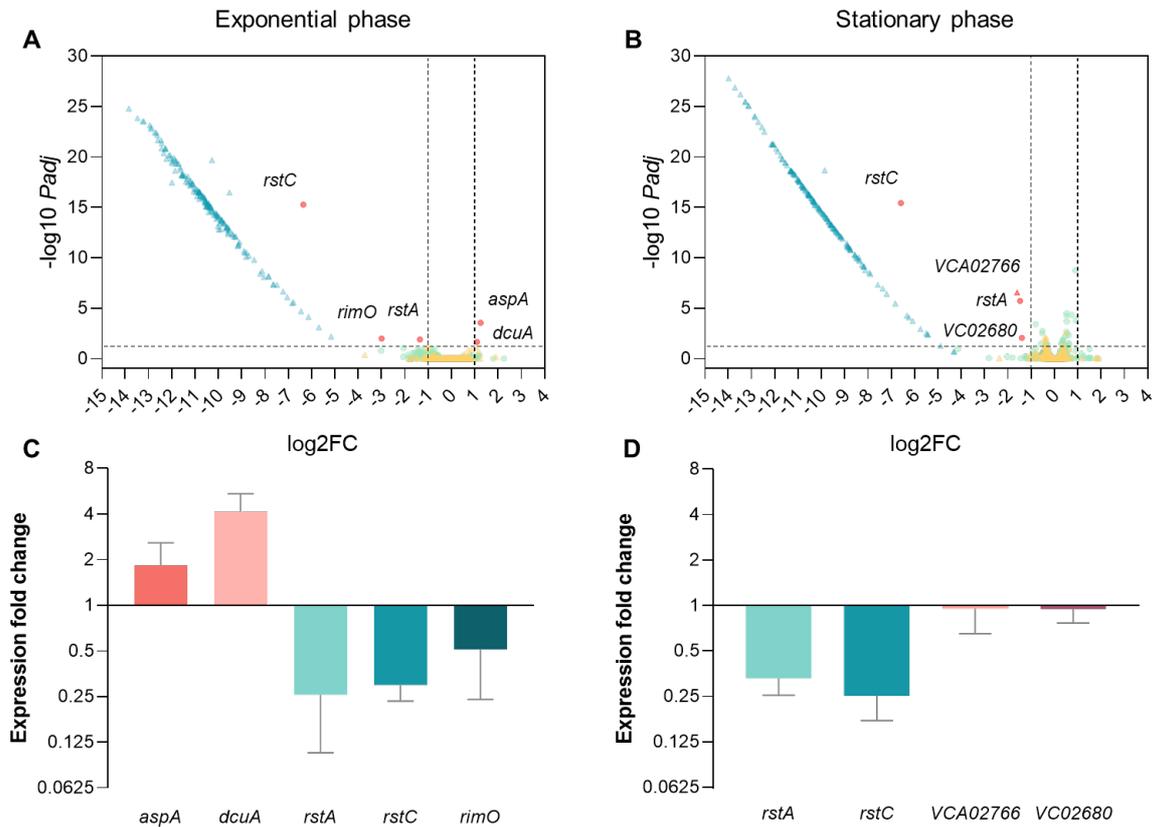
242 To understand the impact of the deletion of the SI in the bacterium's physiology  
243 we need to avoid the interference of the mutations in these three genes. To do so, we  
244 restored the WT sequence of the three alleles using the counterselectable suicide vector  
245 pMP7. We then verified the restoration of the three genes, as well as the absence of other  
246 mutations arising during this process again through whole genome sequencing using short  
247 and long reads. As a result, we could confirm that the  $\Delta$ SI strain is genetically identical  
248 to the WT except for the absence of the SI and the loss of one of the two copies of satellite  
249 phage RS1.

250

### 251 **Transcriptomics**

252 Regulatory networks are generally finely tuned and often affected by minor  
253 genetic changes. To address the level of functional isolation of the SI, we sought to assess  
254 the changes in expression patterns of the genome in the absence of the SI. Indeed, the  
255 deletion of the SI could have pleiotropic effects through direct and indirect links to  
256 regulatory networks within the cell. To unveil such effects, we have performed  
257 transcriptomic analysis of the  $\Delta$ SI and the parental strain in exponential and stationary  
258 phase. As expected, all genes contained within the SI gave a clear signal of decreased  
259 expression. Unexpectedly, the transcriptome of the rest of the genome remained  
260 fundamentally unaltered, with few exceptions that were confirmed through RT-qPCR  
261 using independent RNA extractions (Figure 4A and 4B).

262



263 **Figure 4.** (A) Volcano plots showing gene expression measured by RNA-seq in *V. cholerae*  $\Delta SI$   
 264 in comparison to the WT strain in exponential and (B) stationary growth phase. Green dots  
 265 represent chromosome 1 genes; yellow triangles represent chromosome 2 genes, while blue  
 266 triangles represent those genes from the superintegron; red dots or triangles represent  
 267 downregulated or upregulated genes from chromosome 1 or 2, respectively. The vertical dashed  
 268 lines indicate the  $\log_2$  fold change (FC) cutoffs, and the horizontal dashed line indicates the  
 269 threshold of the  $P_{adj}$  value ( $< 0.05$ ). (C) Validation of the differentially expressed genes by  
 270 quantitative PCR (qPCR) during exponential and (D) stationary growth phase. Differential  
 271 expression values represent the fold change in gene expression compared to the WT strain.  
 272 Error bars indicate standard deviation of two biological replicates with three technical replicates each.  
 273

274 The loss of one of the tandem copies of RS1 in the  $\Delta SI$  strain led to the  
 275 downregulation in both exponential growth and stationary phase of *rstA* (VC01411) and  
 276 *rstC* (VC01409), that was confirmed through RT-qPCR on independent samples. In the  
 277 WT strain, *rstA* is found three times as part of the  $\Phi$  CTX (VC01397) and two tandem  
 278 copies of the satellite phage  $\Phi$  RS1 (VC01407, VC01411), while *rstC* is exclusively  
 279 found in the latter in two copies (VC01409, VC01413).<sup>394041</sup> This downregulation through  
 280 the loss of a  $\Phi$  RS1 copy serves as an additional validation of transcriptomic results.

281 Besides prophage-related genes, in exponential growth phase we also found a  
 282 two-fold upregulation of *aspA* (VC00205; aspartate ammonia-lyase) and *dcuA*  
 283 (VC00204; C4-dicarboxilase transporter) and a 6-fold downregulation of *rimO*  
 284 (VC00523; ribosomal protein S12 methylthiotransferase). *aspA* and *dcuA* colocalize in

285 *V. cholerae* chromosome 1. While *aspA* catalyzes the reversible conversion of L-aspartate  
286 to fumarate<sup>42</sup>, *dcuA* is involved in the L-aspartate/fumarate antiport<sup>43</sup>. Additionally,  
287 RNAseq data showed genes VC02680 and VCA02766 to be down-regulated 2-fold  
288 approximately in stationary phase, but RT-qPCR failed to confirm such downregulation  
289 (Figure 4C and 4D).

290 A gene set enrichment analysis (GSEA) can reveal the collective up- or  
291 downregulation of related genes, even if individually they show no significant changes in  
292 expression patterns. GSEA was performed here on the group of differentially expressed  
293 genes, revealing that the  $\Delta$ SI strain showed a mild enrichment of the iron ion  
294 transmembrane transport process (GO:0034755; NES = 2,126432; *P*<sub>adj</sub> = 0,00714409)  
295 and the arginine biosynthetic process (GO:0006526; NES = 1,959273; *P*<sub>adj</sub> =  
296 0,02539212) in exponential phase and of isoleucine biosynthetic process in stationary  
297 phase (GO: 0009097; NES = 2,020381; *P*<sub>adj</sub> = 0,0129091). Collectively, these data  
298 indicate that deleting the superintegron does not have important consequences at the  
299 transcriptome level in *V. cholerae*.

300

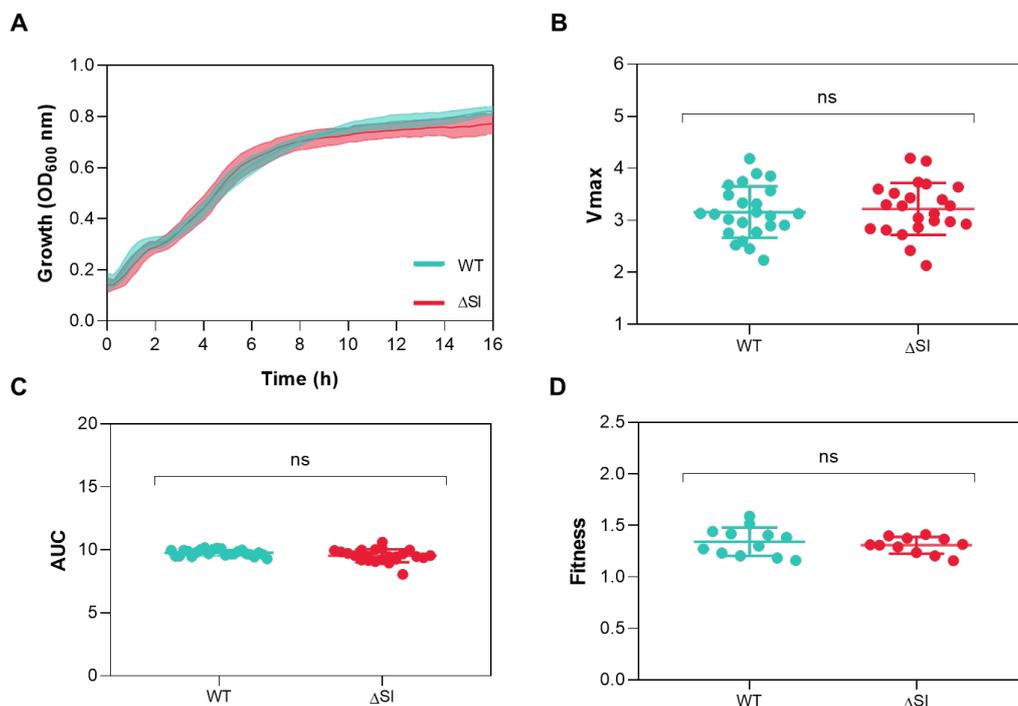
### 301 **Phenotype search**

302 The working model of integrons suggested that most of the array of the  
303 superintegron is silent, and only the cassettes closest to the P<sub>c</sub> are being expressed. Hence,  
304 it would be unlikely to find many phenotypic changes after the deletion of the SI. Yet this  
305 seems not to be true, with independent reports showing both the presence of transcription  
306 start sites scattered along the superintegron, and a large proportion of cassettes expressed  
307 at biologically relevant levels (Blanco *et al.* under review.), in accordance with our  
308 RNAseq results<sup>44</sup>. To address the potential emergence of phenotypic changes, and  
309 therefore the discovery of cassette functions, we implemented a broad approach to  
310 characterize a maximum of phenotypes.

311

312 Growth and Fitness

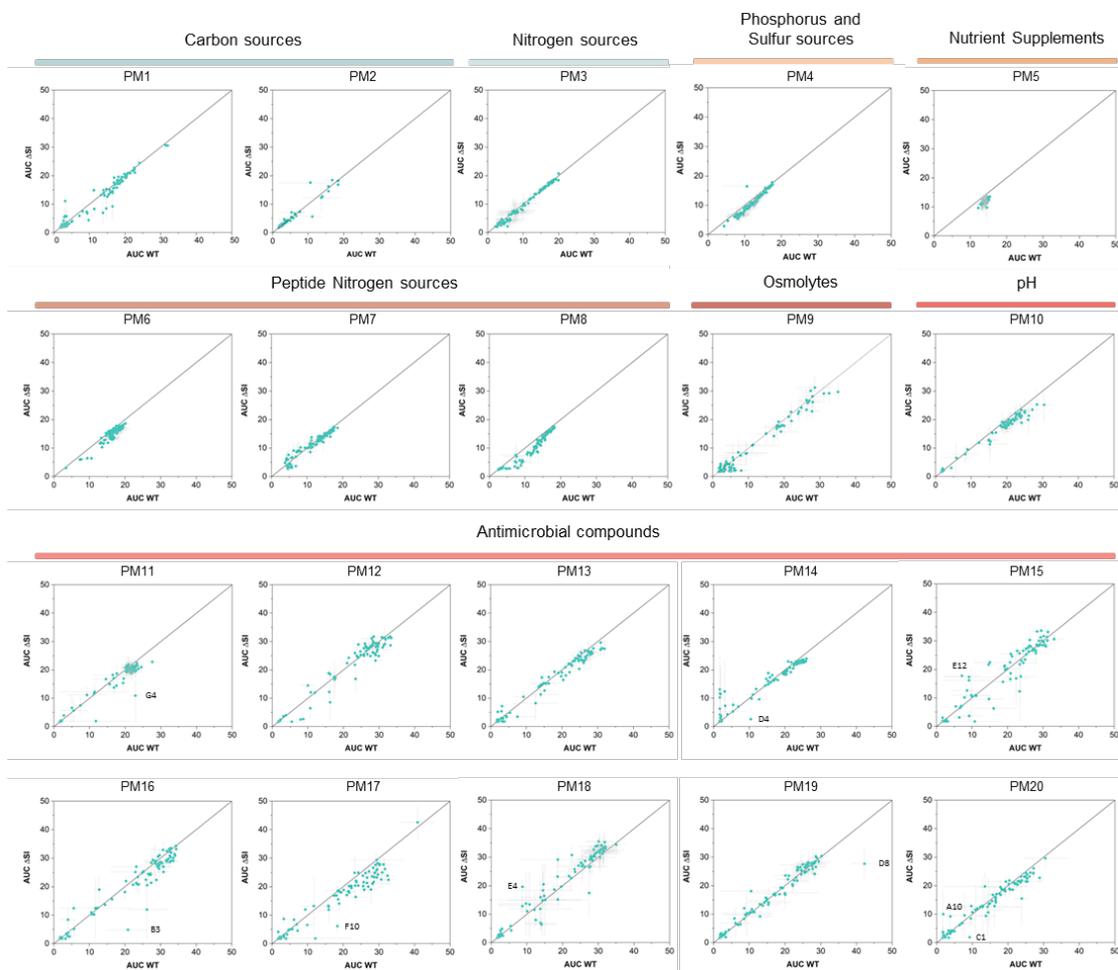
313 Deletion of *V. cholerae* superintegron implies the loss of the 3% of the bacterial  
314 genome. Thus, it is expected that elimination of this genomic load would result in some  
315 growth and/or fitness consequences. To explore this possibility, we performed growth  
316 curves of 24 independent *V. cholerae*  $\Delta$ SI colonies and compared with the WT in rich LB  
317 medium. As shown in Figure 5A, growth curves of both strains present similar shapes  
318 after 16 h of growth. We did not observe significant growth differences between strains  
319 measured as Vmax (Figure 5B). We then measured the area under de curve (AUC), which  
320 combines more information about the growth curve, such as lag phase, Vmax and the  
321 carrying capacity. The results showed similar AUC values for both strains (Figure 5C).  
322 To confirm that the deletion of the superintegron does not influence bacterial fitness, we  
323 increased the sensitivity by performing competition assays using flow cytometry. Using  
324 a fluorescent *E. coli* strain as competitor, we found no significant differences between the  
325 fitness values given by *V. cholerae* WT and  $\Delta$ SI. Altogether, these results indicate that  
326 deletion of the *V. cholerae* superintegron does not affect significantly bacterial growth.



327 **Figure 5.** (A) Growth curves of *V. cholerae* WT (blue) and  $\Delta$ SI (red) strains in LB. Growth  
328 parameters including (B) Vmax and (C) the area under the curve (AUC) were extracted from growth  
329 curves in LB by measuring the optical density at 600 nm (OD<sub>600</sub>). Values correspond to the  
330 measurement of 24 independent colonies. (D) Relative fitness of *V. cholerae* WT and  $\Delta$ SI strains  
331 compared with *E. coli* DH5 $\alpha$  PcS::gfp. Competition assays were performed in LB by inoculating a  
332 cells ratio of 1:1. Fitness values were determined from 12 independent experiments by flow  
333 cytometry. The p-values were calculated by comparing each measure with that of the WT strain  
334 using unpaired t-test. Ns: not significant.

335 Biolog Phenotype Microarrays

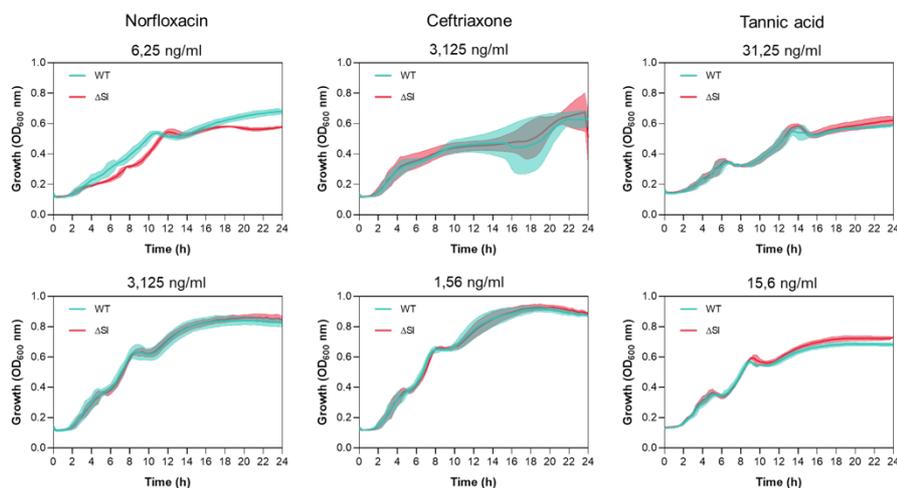
336 To address if we can observe phenotypic differences in the absence of the SI, we  
337 have used Biolog Phenotype Microarrays® (PMs), a phenomics technology that allows  
338 the testing of bacterial respiration and growth in nearly 2.000 conditions<sup>45</sup>. PMs test  
339 carbon utilization, nitrogen sources, phosphorus and sulfur compounds, biosynthetic  
340 pathways, osmotic, ionic and pH effects, as well as a broad set of chemicals and antibiotic  
341 compounds at four different concentrations. By using tetrazolium violet as a redox dye in  
342 the mixture, we monitor bacterial respiration for 24 hours and compare the metabolism  
343 of two independent replicates of the WT and  $\Delta$ SI strain in all the available conditions  
344 (Supplementary Figure 1).



345 **Figure 6.** Comparison of the area under the curves (AUC) values of *V. cholerae* WT and  $\Delta$ SI in  
346 different substrates given by Biolog Phenotype Microarrays (PM1-PM20). The AUC values are the  
347 mean of two biological replicates for each strain. Error bars indicate the standard error. The points  
348 that are above the bisector indicate those conditions where the absence of the superintegron is  
349 beneficial for the bacteria, while those points that lay under the bisector indicate the conditions  
350 where the presence of the superintegron is beneficial for the bacteria.  
351

352

353 To search for differences between the two strains, the mean AUCs given by both  
354 replicates of WT and  $\Delta$ SI in a specific compound is graphically compared using Matlab  
355 software (R2022a). In Figure 6, we show the correlation of AUC values for WT and  $\Delta$ SI  
356 strains in each condition. Distance to the origin of the axis represents metabolic activity,  
357 and deviations from the diagonal represents differences between strains. Generally, dots  
358 fall within the diagonal, showing that both strains behave similarly in the presence of the  
359 different compounds, which also supports the similarity in growth rates across a broad  
360 variety of conditions. It is of note, that the  $\Delta$ SI shows a potential decrease in respiration  
361 when using dipeptides and tripeptides as nitrogen sources (Plates 6 and 8) (see  
362 discussion). Within antimicrobial-containing plates (PM11 to 20), we generally did not  
363 found differences between the WT and  $\Delta$ SI. However, we have highlighted some isolated  
364 cases in Figure 6 where the AUC of one of the strains is at least the double of the other  
365 and the standard deviation does not cross the bisector. These compounds are: ceftriaxone  
366 (G4, PM11), cadmium chloride (D4, PM14), methyl viologen (E12, PM15), norfloxacin  
367 (B3, PM16), tannic acid (F10, PM17), sodium metasilicate (E4, PM18), iodonitro  
368 tetrazolium violet (D8, PM19), thioridazine (C1, PM20), and benserazide (A10, PM20).  
369 Despite the lack of consistent results with other molecules of the same family, we  
370 believed these results deserved further verification. We selected three compounds  
371 (ceftriaxone, norfloxacin and tannic acid) and measured the MIC, but found identical  
372 values for both strains: norfloxacin 50 ng/ml; ceftriaxone 25 ng/ml; tannic acid 250 ng/ml.  
373 We also performed growth curves using subinhibitory concentrations of the three  
374 compounds in MH, but could not reproduce the results obtained using phenotype  
375 microarrays (Figure 7), suggesting that phenotypes.

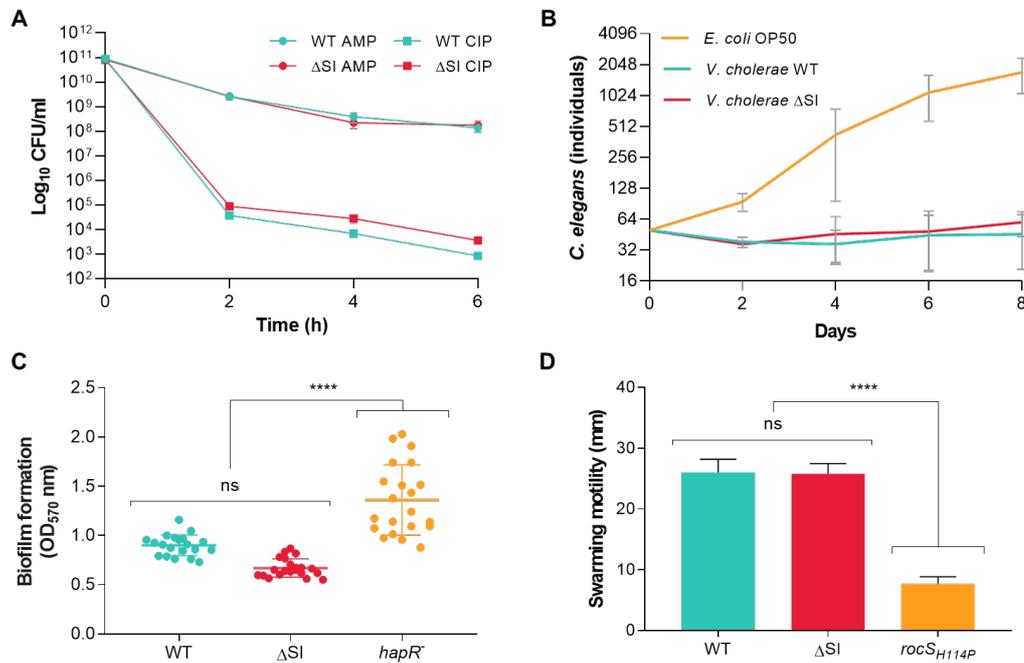


376 **Figure 7.** Growth curves of *V. cholerae* WT (blue) and  $\Delta$ SI (red) strains in MH medium containing  
377 subinhibitory concentrations of norfloxacin, ceftriaxone or tannic acid. Colored shading of the  
378 curves represents the standard deviation of at least three biological replicates.

379 Persistence, virulence, biofilm formation and swarming motility

380 Persistence is the ability of a bacterial population (or a part of it) to survive  
381 exposure to bactericidal drugs<sup>46</sup>, without becoming resistant to the antibiotic. Toxin-  
382 antitoxin systems have been described to play a role in triggering persistence in bacteria  
383 by reversibly intoxicating the bacterium and producing a transient arrest of metabolic  
384 activity that renders antibiotics ineffective. This role is controversial, at least for some  
385 TAs, since several studies were later shown to be biased by the activity of prophages<sup>47</sup>.  
386 Given the deletion of a variety of TS systems in the  $\Delta$ SI mutant we decided to test  
387 persistence upon exposure to high concentrations of antibiotics. We challenged growing  
388 cultures of *V. cholerae* WT and  $\Delta$ SI to 10-fold the minimum inhibitory concentration  
389 (MIC) of ampicillin and ciprofloxacin, and viability was measured along a 6 h-period  
390 (Figure 8A). Similar decreases in the number of viable cells were observed for both  
391 strains, of varying intensities depending on the antibiotic molecule, suggesting that toxin-  
392 antitoxin systems encoded in the superintegron do not play a role in antibiotic persistence  
393 in *V. cholerae*.

394 *V. cholerae* expresses well-known virulence factors to colonize and cause  
395 infection in the mammalian host, as the cholera toxin or the toxin-coregulated pili<sup>48,49</sup>.  
396 The nematode *Caenorhabditis elegans* has been widely used as an invertebrate infection  
397 model to screen and identify virulence factors of several human pathogens, including *V.*  
398 *cholerae*<sup>50,51</sup>. Toxin-antitoxin systems in the superintegron are reported to be associated  
399 with virulence. For instance, the RelBE TA systems of the superintegron play a role in  
400 intestine colonization in a mouse model<sup>52</sup>. However, evidence on the implication of other  
401 elements in the superintegron in the virulence of *V. cholerae* is lacking. To address this  
402 question, we performed a *C. elegans* killing assay. Hermaphrodites L4 worms raised on  
403 *Escherichia coli* OP50 were transferred onto bacterial lawns of WT and  $\Delta$ SI strains. The  
404 number of *C. elegans* individuals were counted every 48 h for 8 days. *E. coli* OP50 was  
405 used a control. As shown in Figure 8B, we did not observe any differences in lethality  
406 between *V. cholerae* WT and the  $\Delta$ SI mutant.



407 **Figure 8.** (A) Persistence assay of *V. cholerae* WT and  $\Delta$ SI mutant challenged with 10-fold the  
408 MIC of ampicillin and ciprofloxacin, individually. CFUs/ml are plotted against time. Error bars indicate  
409 standard deviations from three independent biological replicates. (B) Virulence assay of *V. cholerae* WT  
410 and  $\Delta$ SI mutant in a *C. elegans* model. *E. coli* OP50 was used as a negative control. Error bars indicate  
411 standard deviations from three independent biological replicates. (C) Biofilm formation of *V. cholerae* WT  
412 and  $\Delta$ SI mutant. A *V. cholerae hapR* (frameshift) mutant was included as a positive control of biofilm  
413 formation. Data from at least 20 independent colonies is represented. (D) Swarming motility assay of  
414 *V. cholerae* WT and  $\Delta$ SI mutant. A *V. cholerae rocS* (H114P) mutant was included as a non-motile control  
415 strain. Error bars indicate standard deviation of at least three biological replicates. P values (\*\*\*\*) were  
416 calculated by pairwise comparison using the one-way ANOVA test ( $p < 0,0001$ ). Ns: not significant.  
417

418 Biofilm formation and motility are key behaviors in the lifestyle of *V. cholerae*,  
419 with strong implications in its survival in the environment, pathogenesis, dispersal, and  
420 natural competence. Motility is involved in regulating biofilm formation by participating  
421 in important processes, as surface attachment or biofilm dispersal<sup>53</sup>. We investigated the  
422 role of the SI by measuring biofilm formation and swarming motility. As shown in Figure  
423 8C and 8D, neither process was significantly affected in the  $\Delta$ SI strain. As controls, we  
424 included a *V. cholerae* strain with frameshift mutation in *hapR* (a regulator of quorum  
425 sensing) leading to an increase in biofilm production, and a strain with an amino acid  
426 change mutation in *rocS* as a non-motile control.  
427

428

## 429 DISCUSSION

430 The working model of integrons is based on semiconservative recombination and  
431 the strong expression of cassettes upon integration, highlighting that cassettes must be  
432 adaptive to be maintained in the array. This, together with the intimate intertwining  
433 between integrons and their hosts (through the SOS response controlling the expression  
434 of the integrase<sup>54</sup>, or through host factors<sup>29,55</sup>) put integrons at the service of the host to  
435 provide adaptation on demand<sup>2,55,56</sup>. A clear example of this is their role in the rise and  
436 spread of multidrug resistance during the last decades. In Vibrionaceae, integrons provide  
437 genetic diversity upon which natural selection can act<sup>57</sup>. It is therefore assumed that  
438 sedentary chromosomal integrons provide great adaptability too. Yet a better  
439 understanding of how they do so, remains out of reach because the functions encoded in  
440 the cassettes of chromosomal integrons are mostly unknown<sup>15</sup>. Apart from antibiotic  
441 resistance genes and toxin-antitoxin systems, only a handful of cassettes have been  
442 characterized (for a review see<sup>33</sup> and<sup>58</sup>).

443 Chromosomal integrons have co-evolved with their hosts during aeons. It is  
444 therefore plausible that hosts have co-evolved with integrons and have somehow adapted  
445 to the presence of such massive structures. Also, the functions encoded in cassettes could  
446 be intimately linked to the host's physiology, somehow interfering or modulating  
447 housekeeping functions. Indeed, *in silico* analysis of cassettes in Vibrionaceae revealed  
448 potential functions like information storage, cellular processes, and metabolism<sup>15</sup>. This is  
449 certainly at odds with the variable nature of the array of cassettes, that would imply that  
450 cassettes provide exclusively “self-contained” or independent functions, that do not  
451 interfere with those encoded in the rest of the genome. Here we have tested the hypothesis  
452 that SCIs are genetically and functionally independent of the host by relocating and  
453 deleting the superintegron of *V. cholerae* and testing a broad variety of phenotypes of  
454 importance in the lifestyle of this bacterium. The result is surprisingly clear: the  $\Delta$ SI  
455 mutant behaves almost identically to the WT strain in most conditions tested, and  
456 differences (like those observed in the use of di-peptides) are, at best, very mild. This tilts  
457 the balance towards the “plug-and-play” model of cassette function.

458 Focusing on the subtle differences that we have observed, we note that the deletion  
459 of the SI has a limited impact on gene expression, with only three genes (*aspA*, *dcuA* and  
460 *rimO*) being mildly affected. *aspA* and *dcuA* are involved in nitrogen metabolism. *AspA*  
461 carries out the reversible conversion of L-aspartate to fumarate, releasing ammonia for

462 nitrogen assimilation<sup>59</sup>. This enzyme has a specific interaction with DcuA, which  
463 functions as a L-aspartate/fumarate antiporter under both aerobic and anaerobic  
464 conditions<sup>60</sup>. This interaction has been proposed to form a metabolon, where DcuA  
465 facilitates the uptake of aspartate, while AspA converts it to fumarate<sup>61</sup>. The GSEA  
466 analysis revealed an enrichment in arginine and isoleucine biosynthetic processes in the  
467  $\Delta$ SI strain. Arginine synthesis is derived from glutamate, which, in turn, originates from  
468 L-aspartate. Additionally, the precursor molecule for isoleucine biosynthesis is pyruvate,  
469 a product of glycolysis, and an indirect byproduct of the aspartate metabolism<sup>62</sup>. These  
470 findings, coupled with the observation that the  $\Delta$ SI mutant exhibits an impaired growth  
471 when utilizing certain dipeptides and tripeptides as nitrogen sources, may suggest a  
472 potential connection between these processes. However, it is important to note that these  
473 pathways are not directly linked, and further experimental data would be needed to  
474 ascertain whether a change in metabolic balance has occurred within the cell. Given the  
475 subtle signal from our data, we remain cautious about the biological significance of these  
476 observations.

477         The obtention of a  $\Delta$ SI mutant of *V. cholerae* is a long-awaited milestone in the  
478 field, unblocking the study of sedentary chromosomal integrons. Getting rid of the  
479 interference of the cassettes within the SI allows to use this strain as a chassis to easily  
480 perform many of the experiments that have been delivered in *E. coli* with the class 1  
481 integron, allowing to enrich the field with data from other integron models. It is also the  
482 perfect chassis to unveil cassette functions in a classical *in trans* approach, indeed one of  
483 the key questions that remains unanswered in the field and one with an important  
484 biotechnological potential.

485

486

487

488 **DATA AVAILABILITY**

489 The RNA-seq data included in this publication have been deposited in NCBI's Gene Expression  
490 Omnibus<sup>63</sup> and are accessible through GEO Series accession number GSE247496.

491

492 **SUPPLEMENTARY DATA**

493 Supplementary Data are available online at the publisher's web.

494

495 **AUTHOR CONTRIBUTIONS**

496 Paula Blanco: conceptualization, formal analysis, methodology, validation,  
497 writing-original draft. Filipa Trigo: methodology, analysis and validation. Laura Toribio-  
498 Celestino: methodology, analysis, validation. Lucía García-Pastor: methodology,  
499 analysis and validation. Niccolò Caselli: methodology, analysis and validation. Francisco  
500 Ojeda: methodology, analysis and validation. Baptiste Darracq: methodology, analysis  
501 and validation. Ole Skovgaard: methodology, analysis and validation. Álvaro San  
502 Millán: formal analysis, writing- review and editing. Didier Mazel: conceptualization,  
503 writing- review and editing. Céline Loot: conceptualization, analysis, validation, writing-  
504 review and editing. José Antonio Escudero: conceptualization, methodology, formal  
505 analysis, validation, writing- original draft, review and editing.

506

507 **ACKNOWLEDGEMENTS**

508 Authors would like to thank Laurence Van Melderren, Nathalie Balaban and  
509 David Bikard for helpful discussion, and Claire Vit for technical assistance.

510

511 **FUNDING ENTITIES**

512 This work is supported by the European Research Council (ERC) through a Starting Grant  
513 [ERC grant no. 803375-KRYPTONINT;]; Ministerio de Ciencia, Innovación y  
514 Universidades [BIO2017-85056-P, PID2020-117499RB-100]; JAE is supported by the  
515 Atracción de Talento Program of the Comunidad de Madrid [2016-T1/BIO-1105 and  
516 2020-5A/BIO-19726]; PB is supported by the Juan de la Cierva program [FJC 2020-  
517 043017-I]; FTR is supported by the Portuguese Fundação para Ciência e a Tecnologia  
518 [SFRH/BD/144108/2019]; LTC and ASM are supported by the European Research  
519 Council (ERC) under the European Union's Horizon 2020 research and innovation  
520 program (ERC grant no. 757440-PLASREVOLUTION). BD, DM and CL are supported

521 by Institut Pasteur, the Centre Nationale de la Recherche Scientifique (CNRS-UMR3525)  
522 and the Fondation pour la Recherche Médicale (FRM Grant No. EQU202103012569)

523

524 **CONFLICTS OF INTEREST**

525 Universidad Complutense de Madrid and Institut Pasteur have filed a patent covering  
526 SeqDelTA and the *V. cholerae*  $\Delta$ SI strain B522. Patent number: P202330947. Inventors:  
527 PB, FTR, DM, JAE.

## 528 MATERIALS AND METHODS

### 529 SI relocation from Chr2 to Chr1

530 All strains, plasmids and primers used for the SI relocation are listed in  
531 Supplementary Table S1, S4, and S5. To relocate the whole superintegron, we used a  
532 genetic tool developed in the lab based on the recombination of two bacteriophage  
533 attachment sites<sup>64</sup>. Prophage excision from the host chromosome relies on site-specific  
534 recombination between two sequences flanking the phage, termed attachment sites *attL*  
535 and *attR*. This recombination is carried out by phage-specific recombinases “Int” and  
536 their cognate RDF (recombination directionality factor), or excisionase, Xis. The tool  
537 developed in the lab uses the attachment sites from bacteriophages HK and  $\lambda$   
538 (Supplementary Figure 2). The *attL/attR* pairs of each site were associated with parts of  
539 a genetic marker that is reconstituted when the sites recombine: HK sites were associated  
540 with parts of a *bla* gene (ampicillin resistance) and the  $\lambda$  sites with parts of *lacZ*. Co-  
541 expression of the HK and  $\lambda$  integrases/excisionases triggers the recombination of the  
542 partner *attL/attR* sites, which allows for the reconstitution of fully functional *bla* and *lacZ*  
543 markers and the selection of clones in which the SCI relocation has occurred. Blue  
544 colonies growing in LB supplemented with carbenicillin and X-gal (5-bromo-4-chloro-3-  
545 indolyl- $\beta$ -D-galactopyranoside) were PCR verified and used for growth curves.

546

### 547 SeqDelTA

548 All strains, plasmids, and primers for the deletion of the SI are listed in  
549 Supplementary Table S2, S3, S4, and S5. To generate the  $\Delta$ SI mutant of *V. cholerae*  
550 N16961, we developed SeqDelTA, as depicted in Figure 2. After identifying the 19 toxin-  
551 antitoxin (TA) systems contained in the *V. cholerae* N16961 superintegron, we prepared  
552 successive constructs designed with homology regions and a selective marker. A left  
553 homology region (LHR) was designed and maintained throughout the procedure. This  
554 region corresponded to the *catB9* gene (VCA0300), and was amplified with primers VCA  
555 299 F and LRHI VCA 300 R. For the successive deletions, different right homology  
556 regions (RHD) were designed to maintain antitoxin gene but to eliminate part of the  
557 corresponding toxin gene, ensuring that the event was not lethal to the bacteria (Figure 2  
558 and Table S2). As selective markers, three different antibiotic resistance genes (Zeo<sup>R</sup>,  
559 Cm<sup>R</sup>, and Carb<sup>R</sup>) were used and amplified with the same primers, facilitating the  
560 consequent design and assembly of the HRs. These three fragments (LHR, resistance  
561 marker, and RHR) were then assembled by SOE-PCR, placing the resistance marker

562 between the LHR and the RHR. Each successive construct was introduced by natural  
563 transformation (described below) into the *Vibrio cholerae* strain obtained from the  
564 previous TA system deletion. This strategy allowed the removal of the resistance marker  
565 introduced in the previous step while advancing the deletion of TA systems by changes  
566 in the RHR. When necessary, due to gene orientation and/or gene order in the TA system,  
567 alternative LHRs were designed, and two resistance markers were maintained in  
568 consecutive allelic replacements before the new step allowed deletion of both (Table S2).

569 The last deletion was performed using the R6K origin of replication-plasmid  
570 pMP7 as described in<sup>34,65</sup>. pMP7 replication is dependent of the cell  $\pi$  protein and encodes  
571 the *ccdB* toxin under the control of the inducible promoter P<sub>BAD</sub>. Five hundred bp of each  
572 side of the superintegron were amplified with primers RHI link pMP7 F and LRHI+D R  
573 for the LHR, and RHD F and RHD link pMP7 R for the RHR. The two fragments were  
574 then joined by SOE-PCR and digested with enzymes NaeI y HindIII. Finally, the insert  
575 was ligated with the empty pMP7 backbone, giving rise to pMP7\_Δ*intIA attIA zeo*<sup>R</sup>. We  
576 started the cloning steps by transforming the pMP7\_Δ*intIA attIA zeo*<sup>R</sup> plasmid into *ccdB*-  
577 resistant *E. coli* π3813 (*thy::pyr* +) competent cells. Transformants were selected in LB  
578 agar plates containing Cm (25 μg/ml), thymidine (dT; 0,3 mM) and glucose (1%). For  
579 conjugal transfer of pMP7\_Δ*intIA attIA zeo*<sup>R</sup> into *V. cholerae*, the plasmid was first  
580 transformed into the donor strain *E. coli* β3914 (*dap::pir* +). Transformants were selected  
581 in LB agar plates supplemented with Cm (25 μg/ml), 2,6-diaminopimelic acid (DAP; 0,3  
582 mM) and glucose (1%). Conjugation assay was performed by growing both donor (*E. coli*  
583 β3914) and recipient (*V. cholerae* A066) strains to a OD<sub>600</sub> of 0,2. Then, cells were mixed  
584 in a ratio of 1/10 (donor/recipient) and transferred to a mating filter placed on a LB agar  
585 plate supplemented with DAP (0,3 mM) and glucose (1%) and conjugation was  
586 performed overnight at 37°C. Integration of the pMP7\_Δ*intIA attIA zeo*<sup>R</sup> into the *Vibrio*  
587 genome was selected by washing the filter in 5 ml of LB and plating serial dilutions on  
588 Cm (2,5 μg/ml) and glucose (1%) plates, but lacking DAP. Afterwards, Cm-resistant  
589 colonies were grown in liquid LB medium and plated in LB agar plates supplemented  
590 with L-arabinose (0,2%) to express *ccdB* and select for the second crossover that implies  
591 the excision of the pMP7 backbone. After this process, we Sanger-sequenced the  
592 superintegron region and selected a *V. cholerae* ΔSI strain (A101).

593

594 **Natural transformation assays**

595 Transformation assays were performed as previously described in<sup>66</sup> with some  
596 modifications. Briefly, strains were inoculated from frozen stocks into LB broth and  
597 grown rolling overnight at 30°C. Overnight cultures were diluted 1:100 and grown to an  
598 OD<sub>600</sub> of 1.0. Cells were then washed and resuspended in Instant Ocean<sup>®</sup> 0,5x (IO), and  
599 inoculated in a final volume of 1mL IO with 150 µl of chitin (Apollo Scientific<sup>®</sup>) slurry.  
600 Cells were incubated in static conditions at 30°C for ~18 hours. Then, 550 µl of  
601 supernatant was removed from the reactions, and the purified transforming DNA (tDNA)  
602 was added. DNA and cells were incubated for ~18 hours in static conditions at 30°C.  
603 When needed, reactions were outgrown by adding 1 m LB broth and shaking at 30°C for  
604 2 hrs. Cells were plated in LB agar with the corresponding antibiotic.

605

### 606 **DNA isolation, WGS, and data analysis**

607 Genomic DNA of the obtained *V. cholerae* ΔSI strain (A101) was extracted using  
608 the DNeasy<sup>®</sup> Blood and Tissue Kit (QIAGEN) following the manufacturer's protocol.  
609 The DNA quantity was determined using a BioSpectrometer (Eppendorf). DNA  
610 sequencing was performed at Institut Pasteur using Illumina paired-end reads, at MiGS  
611 (Pittsburgh, US), now SeqCoast Genomics (Portsmouth, US) and inhouse at MBA  
612 (Madrid) using MinIon. Data analysis was accomplished with Geneious Prime software  
613 and the genetic variants were identified by mapping the generated reads to the *V. cholerae*  
614 N16961 reference genome (Accession number: chromosome 1 CP028827.1;  
615 chromosome 2 CP028828.1).

616

### 617 **Marker Frequency Analysis (MFA)**

618 Marker Frequency Analysis was performed as described in<sup>35</sup>. Cells were grown  
619 in LB at 37°C with shaking and genomic DNA was isolated using the DNeasy<sup>®</sup> Tissue  
620 Kit (Qiagen) from exponential and stationary phase cultures. Genomic DNA was  
621 sequenced with Illumina technology. Short reads were aligned to the reference genome  
622 using Bowtie2<sup>67</sup> and R2R. The number of reads starting per bp (N) were calculated for 1  
623 kbp and 10 kbp windows. The exponential phase data were normalized with a factor  
624 calculated as the deviation of observed N from calculated N for each window of the  
625 stationary phase culture data. Any window including repeated sequences was excluded  
626 and N was plotted as a function of the absolute position, centered on the replication  
627 origins.

628

## 629 **Correction of *V. cholerae* $\Delta$ SI mutations**

630 All the strains, plasmids and primers are listed in Supplementary Table S3, S4 and  
631 S5. Three mutations were found in *V. cholerae*  $\Delta$ SI strain (A101): *rocS* (VC0653), *cry2*  
632 (VC01392) and *rpoS* (VC0534). While *rocS* and *rpoS* mutations led to frameshifts, *cry2*  
633 contained a non-synonymous mutation. To revert the mutations to the WT variant, allelic  
634 exchanges were performed using the pMP7, as previously described. To provide  
635 homology for the allelic replacement, 1.000 bp-fragments of the N16961 strain genes  
636 *rocS*, *rpoS* and *cry2* were amplified with primers *rocS*\_pMP7 F/R, *rpoS*\_pMP7 F/R, and  
637 *cry2*\_pMP7 F/R, respectively, which contain 20 bp-homology fragments with the pMP7  
638 cloning site; pMP7 backbone was amplified with oligos pMP7\_bb\_Gibson F/R. Cloning  
639 of the amplified fragments was performed using Gibson assembly<sup>68</sup>, giving rise to  
640 pMP7\_*rocS*, pMP7\_*rpoS*, and pMP7\_*cry2*. We performed the mating and allelic  
641 exchanges as previously mentioned, starting with pMP7\_*rocS*, then pMP7\_*rpoS*, and  
642 finally, pMP7\_*cry2*. After each step, we Sanger-sequenced the targeted gene and selected  
643 a *V. cholerae* strain corrected for *rocS*<sup>+</sup> (A677), *rocS*<sup>+</sup>, *rpoS*<sup>+</sup> (A684), and finally *rocS*<sup>+</sup>,  
644 *rpoS*<sup>+</sup>, *cry2*<sup>+</sup> (B522). We verified by Illumina whole-genome sequencing that *V. cholerae*  
645  $\Delta$ SI (B522) did not contain any other unintended mutations.

646

## 647 **RNA extraction and preparation for RNA-seq**

648 Total RNA was extracted from *V. cholerae* WT and  $\Delta$ SI cultures grown in LB at  
649 37°C in both exponential (OD<sub>600</sub> 0,8) and stationary (OD<sub>600</sub> 2,8) growth phases using the  
650 RNeasy® Mini Kit (QIAGEN), following the manufacturer's protocol. To eliminate any  
651 residual DNA, RNA was treated using the TURBO DNA-free™ Kit (Invitrogen). RNA  
652 concentration was measured using a BioSpectrometer (Eppendorf), and RNA integrity  
653 was determined using a Qubit™ 4 fluorometer (Invitrogen) with the RNA IQ Assay Kit  
654 (Invitrogen). Ribodepletion RNA library and sequencing was performed at the Oxford  
655 Genomics Centre using a NovaSeq6000 sequencing system (Illumina). Three biological  
656 samples per condition were sequenced.

657

## 658 **RNA-seq data analysis**

659 Raw Illumina reads were trimmed and adapter-removed using Trim Galore v0.6.6  
660 (<https://github.com/FelixKrueger/TrimGalore>) with a quality threshold of 20 and  
661 removing reads shorter than 50 bp. Trimmed paired reads were mapped to the *V. cholerae*  
662 N16961 reference genome (Accession number: chromosome 1 CP028827.1;

663 chromosome 2 CP028828.1) using BWA-MEM v0.7.17<sup>69</sup>. featureCounts from the  
664 Rsubread v2.10.2 package<sup>70</sup> was used to obtain read counts per annotated feature,  
665 including CDS, ncRNA, tmRNA, RNase P, tRNA, antisense RNA and SRP RNA.  
666 Differential expression analysis was performed from raw counts using DESeq2 v1.36.0<sup>71</sup>,  
667 by comparing the expression data of the deletion mutant against the WT strain.

668 Differentially expressed genes (DEGs) of CDSs were annotated with Gene  
669 Ontology (GO) terms retrieved from UniProt on July 7, 2023<sup>72</sup>. Gene Set Enrichment  
670 Analysis was performed to identify sets of biological processes, molecular functions, or  
671 cellular components with different expression profiles between the superintegron deletion  
672 mutant and WT strains, in both exponential and stationary phases. For this, a pre-ranked  
673 list of DEGs, ordered by log<sub>2</sub> fold changes, was provided to the GSEA function of the  
674 clusterProfiler v4.8.1<sup>73</sup>. Genes belonging to the superintegron (positions 312,057-  
675 438,942 bp of chromosome 2) were not included to prevent biased enrichment results  
676 caused by the acute downregulation of the superintegron genes. GSEA was run with  
677 default parameters, correcting for multiple tests with the Benjamini-Hochberg procedure.

678

### 679 **Quantitative real-time PCR**

680 Independent samples of total RNA from *V. cholerae* WT and  $\Delta$ SI cultures grown  
681 in LB at 37°C in both exponential (OD<sub>600</sub> 0,8) and stationary (OD<sub>600</sub> 2,8) were isolated as  
682 described previously. To eliminate any residual DNA, RNA was treated using the  
683 TURBO DNA-free™ Kit (Invitrogen) and RNA concentration was measured using a  
684 BioSpectrometer (Eppendorf). Before cDNA synthesis, a previous DNA wipe-out step  
685 was performed using the QuantiTect® Reverse Transcription Kit (QIAGEN) following  
686 the manufacturer's instructions; cDNA synthesis was carried out afterwards using the  
687 same kit using the temperature steps: 42°C for 15 min; 95 °C for 3 min.

688 RT-qPCR was performed in an Applied Biosystems QuantStudio 3 using the Fast  
689 SYBR™ Green Master Mix (Thermo Fischer Scientific). Primers used are listed in  
690 Supplementary Table S5, and *gyrA* was used as housekeeping gene. Relative changes in  
691 gene expression for the superintegron deletion mutant with respect to the WT were  
692 determined according to the threshold cycle method ( $2^{-\Delta\Delta CT}$ ). Mean values were obtained  
693 from two independent biological replicates with three technical replicates each.

694

695 **Note on the controls used for the search for phenotypes.**

696 The original *V. cholerae* N16961 isolate contains mutations in the *hapR* gene and  
697 is not naturally competent. The deletion of the superintegron has been performed in a  
698 well-known derivative strain in which natural competence is restored by the insertion of  
699 a transposon carrying a functional *hapR* gene (strain A001 in the Supplementary Table  
700 S3). It is known that *V. cholerae* cultures have a tendency towards the *hapR*<sup>-</sup> genotype.  
701 Through genome sequencing we observed this phenomenon in our stock of the parental  
702 strain, which seemed to rise through homologous recombination between both *hapR*  
703 alleles. Since *hapR*<sup>-</sup> strains have certain phenotypes, such as distinct growth curves or a  
704 decreased virulence, we performed most phenotypic experiments using as a control a  
705 strain of *V. cholerae* N16961, in which mutations in *hapR* have been corrected without  
706 the addition of a second allele (A096 in our strain list). In RNAseq and MFA experiments  
707 were we could use as a control the parental strain (A001), by confirming through  
708 sequence analysis that all replicates contained a functional *hapR* allele.

709

710 **Growth curves**

711 Overnight cultures were prepared by inoculating single colonies from LB agar  
712 plates in LB liquid medium and incubating for 16-20 h at 37°C. Cultures were then diluted  
713 1:1.000 in fresh medium and 200 µl were transferred into 96-well plates (Nunc, Thermo  
714 Scientific). OD<sub>600</sub> was measured every 15 min for 16 h using a Biotek Synergy HTX plate  
715 reader (Agilent). V<sub>max</sub> and area under the curve (AUC) were determined using Gen5  
716 software and MATLAB, version R2022a (Mathworks), respectively. At least 24  
717 independent replicates were included in the experiment and significant differences were  
718 assessed by performing unpaired t-test.

719 Growth curves in the presence of subinhibitory concentration of norfloxacin,  
720 ceftriaxone and tannic acid were performed by diluting the compounds at the desired  
721 concentrations in MH medium. Overnight cultures of *V. cholerae* WT and ΔSI were  
722 diluted 1:1000 in the antimicrobial-containing MH media in 96-well plates (Nunc,  
723 Thermo Scientific). OD<sub>600</sub> was measured every 20 min for 24 h using a Biotek Synergy  
724 HTX plate reader (Agilent). At least three biological replicates were included in the  
725 experiment.

726

727 **Competition assays**

728 Competition assays were performed by flow cytometry to measure the fitness  
729 values of *V. cholerae* WT and  $\Delta$ SI mutant relative to an *E. coli* DH5 $\alpha$  strain as a common  
730 competitor carrying a pSU38 plasmid with the *gfp* under the control of the constitutive  
731 promoter P<sub>C</sub>S. Competition procedure was performed as described in<sup>74</sup> with some  
732 modifications. Briefly, pre-cultures were prepared by inoculating single colonies from a  
733 LB agar plate in 200  $\mu$ l of liquid LB in a 96-well plate (Nunc, Thermo Scientific). After  
734 22 h of growth at 37°C with 250 rpm shaking, cultures were mixed at 1:1 proportion and  
735 diluted 1:400 in fresh medium. In order to confirm the initial proportions of *E. coli*  
736 carrying *gfp* and the non-fluorescent *V. cholerae* strains, cells were diluted 1:400 in NaCl  
737 (0,9%) and 30.000 events per sample were recorded using a CytoFLEX S flow cytometer  
738 (Beckman Coulter). Bacterial mixtures were competed for 22 h at 37°C with 250 rpm  
739 shaking, and final proportions were determined as stated above. The fitness values of both  
740 *V. cholerae* WT and  $\Delta$ SI, relative to *E. coli* P<sub>C</sub>S::*gfp* were calculated using the formula:  
741  $w = \ln(N_{\text{final,gfp-}}/N_{\text{initial,gfp-}}) / \ln(N_{\text{final,gfp+}}/N_{\text{initial,gfp+}})$ , where *w* is the relative fitness of the non  
742 GFP-tagged *V. cholerae* strains,  $N_{\text{initial,gfp-}}$  and  $N_{\text{final,gfp-}}$  are the numbers of non GFP-  
743 tagged *V. cholerae* at the beginning and end of the competition, and  $N_{\text{initial,gfp+}}$  and  $N_{\text{final,gfp+}}$   
744 are the numbers of *E. coli* P<sub>C</sub>S::*gfp* cells at the beginning and end of the competition,  
745 respectively. Twelve biological replicates were performed for each competition  
746 experiment and significant differences were assessed by performing unpaired t-test.

747

#### 748 **Phenotypic microarray assay**

749 Phenotype Microarrays (Biolog) PM1-PM20 were tested as previously described  
750 in<sup>45</sup> with some modifications. Briefly, single colonies of *V. cholerae* WT and  $\Delta$ SI were  
751 inoculated in liquid LB medium and grown overnight at 37 °C. For PM1 and PM2 plates  
752 inoculation, 120  $\mu$ l of Redox Dye mix D (100x), 1 ml of sterile water and 880  $\mu$ l of a 85%  
753 T cell suspension in NaCl (0.9%) were added to 10 ml of IF-0 (1.2x), resulting in a final  
754 volume of 12 ml per plate. For PM3-PM8 plates inoculation, which requires an  
755 appropriate carbon source, 120  $\mu$ l of sodium pyruvate (2M) was added as an additive and  
756 mixed with 120  $\mu$ l of Redox Dye mix D (100x), 880  $\mu$ l of sterile water, and 880  $\mu$ l of a  
757 85% T cell suspension in NaCl (0.9%). Then, this solution was added to 10 ml of IF-0  
758 (1.2x), resulting in a final volume of 12 ml per plate. For PM9-PM20 plates inoculation,  
759 120  $\mu$ l of Redox Dye mix D (100x), 1 ml of sterile water and 880  $\mu$ l of 1:200 dilution of  
760 the 85% T cell suspension in NaCl (0.9%) were added to 10 ml of IF-10 (1.2x), resulting  
761 in a final volume of 12 ml per plate. Each plate was inoculated with 100  $\mu$ l per well of

762 the respective prepared solution and then sealed using Breathe-Easy® sealing membranes  
763 (Sigma-Aldrich). Plates were then transferred to a Biostack Microplate Stacker (Agilent)  
764 to process several plates per day and incubated at 37°C using a Memmert IF750  
765 (Mettler). OD<sub>590</sub>, which indicates the reduction of the tetrazolium dye, was measured  
766 every 20 min during 24 h using a Biotek Synergy HTX plate reader (Agilent).  
767 Experiments were conducted in duplicate and growth characteristics were determined by  
768 calculating the AUC using MATLAB, version R2022a (Mathworks).

769

#### 770 **Persistence assay**

771 Three colonies of *V. cholerae* WT and  $\Delta$ SI were inoculated in LB liquid medium  
772 and grown overnight at 37°C. Overnight cultures were diluted 1:100 in 1 ml of fresh LB  
773 medium in 24-well plates and grown at 37°C on 150 rpm shaker for 3h. After 3 h  
774 incubation, cultures were challenged individually with 10-fold MIC of ciprofloxacin  
775 (MIC = 0,008 µg/ml) or ampicillin (MIC = 16 µg/ml) and incubated for 6 h at 37°C on  
776 150 rpm shaker. The CFUs/ml were enumerated at times 0 h (before adding the  
777 antibiotics), 2 h, 4 h, and 6 h by serially diluting the cells and plating on LB agar.

778

#### 779 ***Caenorhabditis elegans* virulence assay**

780 *C. elegans* N2 Bristol individuals were routinely maintained at 20°C on potato  
781 dextrose agar (PDA; Sigma-Aldrich) plates seeded with *E. coli* OP50. Overnight LB  
782 cultures of the *V. cholerae* strains and *E. coli* OP50 were diluted 1:100 and grown at 37°C  
783 until an OD<sub>600</sub> of 0,8 was reached. Then, 50 µl from each strain were spread on 6-cm-  
784 diameter plates containing PDA medium. Plates were incubated for 16-20 h at 37°C to  
785 form a bacterial lawn and 50 L4-stage hermaphrodite individuals were then placed on the  
786 PDA plates and incubated at 20°C. Live *C. elegans* were scored every 48 h for 8 days.  
787 An individual was considered dead when it no longer responded to touch. *E. coli* OP50  
788 was used as a negative control of virulence. At least, three independent experiments were  
789 performed.

790

#### 791 **Biofilm formation assay**

792 Overnight cultures were diluted 1:100 in fresh LB medium. Hundred microliters  
793 were transferred to into 96-well Serocluster™ plates (Costar) and incubated at 37°C for  
794 20 h without shaking. After incubation, all the wells were washed with sterile distilled  
795 water for three times and dried. Biofilm mass was stained with 125 µl cristal violet (1%)

796 per well and incubated for 15 min at room temperature. Cristal violet was removed by  
797 performing three washes with sterile distilled water. The stained biofilm mass was then  
798 detached from the wells using acetic acid (30 %) and transferred to a new 96-well plate  
799 (Nunc, Thermo Scientific). OD<sub>570</sub> was measured using a Biotek Synergy HTX plate  
800 reader (Agilent). At least 20 independent experiments were performed per strain. One  
801 way-ANOVA was used to determine significant differences among the three groups.

802

### 803 **Swarming motility assay**

804 Swarming motility assays were performed as described in<sup>75</sup>. Briefly, 3 µl of  
805 overnight cultures were soaked on motility agar plates containing 10 g/l tryptone, 5 g/l  
806 NaCl, and 0,25% bacto-agar. The swarming motility halos were measured and compared  
807 after growth at 37°C for 6 h. At least three independent experiments were performed per  
808 strain. One way-ANOVA was used to determine significant differences among the three  
809 groups.

810

### 811 **REFERENCES**

812

- 813 1. Mazel, D. Integrons: Agents of bacterial evolution. *Nat Rev Microbiol* **4**, 608–  
814 620 (2006).
- 815 2. Escudero\*, J. A., Loot\*, C., Nivina, A. & Mazel, D. The Integron: Adaptation On  
816 Demand. *Microbiol Spectr* **3**, (2015).
- 817 3. Stokes, H. W. & Hall, R. M. A novel family of potentially mobile DNA elements  
818 encoding site-specific gene-integration functions: integrons. *Mol Microbiol* **3**,  
819 1669–1683 (1989).
- 820 4. Souque, C., Escudero, J. A. & Maclean, R. C. Integron activity accelerates the  
821 evolution of antibiotic resistance. *Elife* **10**, 1–47 (2021).
- 822 5. Barraud, O. & Ploy, M. C. Diversity of class 1 integron gene cassette  
823 rearrangements selected under antibiotic pressure. *J Bacteriol* **197**, 2171–2178  
824 (2015).
- 825 6. Mitsuhashi, S., Harada, K., Hashimoto, H. & Egawa, R. On the drug-resistance  
826 of enteric bacteria. 4. Drug-resistance of Shigella prevalent in Japan. *Jpn J Exp*  
827 *Med* **31**, 47–52 (1961).
- 828 7. Gillings, M. *et al.* The Evolution of Class 1 Integrons and the Rise of Antibiotic  
829 Resistance. *J Bacteriol* **190**, 5095–5100 (2008).
- 830 8. Zhu, Y.-G. *et al.* Microbial mass movements. *Science (1979)* **357**, 1099–1100  
831 (2017).
- 832 9. Hipólito, A. *et al.* The expression of aminoglycoside resistance genes in integron  
833 cassettes is not controlled by riboswitches. *Nucleic Acids Res* (2022)  
834 doi:10.1093/nar/gkac662.
- 835 10. Hipólito, A., García-Pastor, L., Vergara, E., Jové, T. & Escudero, J. A. Profile  
836 and resistance levels of 136 integron resistance genes. *npj Antimicrobials and*  
837 *Resistance* **1**, 13 (2023).

- 838 11. Rowe-Magnus, D. A. *et al.* The evolutionary history of chromosomal super-  
839 integrons provides an ancestry for multiresistant integrons. *Proc Natl Acad Sci U*  
840 *S A* **98**, 652–657 (2001).
- 841 12. Néron, B. *et al.* IntegronFinder 2.0: Identification and Analysis of Integrons  
842 across Bacteria, with a Focus on Antibiotic Resistance in Klebsiella.  
843 *Microorganisms* **10**, (2022).
- 844 13. Cury, J., Jové, T., Touchon, M., Néron, B. & Rocha, E. P. Identification and  
845 analysis of integrons and cassette arrays in bacterial genomes. *Nucleic Acids Res*  
846 **44**, 4539–4550 (2016).
- 847 14. Clark, C. A., Purins, L., Kaewrakon, P., Focareta, T. & Manning, P. A. The  
848 *Vibrio cholerae* 01 chromosomal integron. *Microbiology (N Y)* **146**, 2605–2612  
849 (2000).
- 850 15. Boucher, Y., Labbate, M., Koenig, J. E. & Stokes, H. W. Integrons: mobilizable  
851 platforms that promote genetic diversity in bacteria. *Trends Microbiol* **15**, 301–  
852 309 (2007).
- 853 16. Ghaly, T. M. *et al.* Methods for the targeted sequencing and analysis of integrons  
854 and their gene cassettes from complex microbial communities. *Microb Genom* **8**,  
855 (2022).
- 856 17. Mazel, D., Dychinco, B., Webb, V. A. & Davies, J. A distinctive class of  
857 integron in the *Vibrio cholerae* genome. *Science (1979)* **280**, 605–608 (1998).
- 858 18. Iqbal, N., Guérout, A. M., Krin, E., Le Roux, F. & Mazel, D. Comprehensive  
859 functional analysis of the 18 *Vibrio cholerae* N16961 toxin-antitoxin systems  
860 substantiates their role in stabilizing the superintegron. *J Bacteriol* **197**, 2150–  
861 2159 (2015).
- 862 19. Krin, E. *et al.* Systematic transcriptome analysis allows the identification of new  
863 type I and type II Toxin/Antitoxin systems located in the superintegron of *Vibrio*  
864 *cholerae*. *Res Microbiol* 103997 (2022) doi:10.1016/j.resmic.2022.103997.
- 865 20. Rowe-Magnus, D. A., Guerout, A. M. & Mazel, D. Bacterial resistance evolution  
866 by recruitment of super-integron gene cassettes. *Mol Microbiol* **43**, 1657–1669  
867 (2002).
- 868 21. Szekeres, S., Dauti, M., Wilde, C., Mazel, D. & Rowe-Magnus, D. A.  
869 Chromosomal toxin-antitoxin loci can diminish large-scale genome reductions in  
870 the absence of selection. *Mol Microbiol* **63**, 1588–1605 (2007).
- 871 22. Richard, E. *et al.* Cassette recombination dynamics within chromosomal  
872 integrons are regulated by toxin–antitoxin systems. *bioRxiv* 2022.08.03.502626  
873 (2023) doi:10.1101/2022.08.03.502626.
- 874 23. Biskri, L., Bouvier, M., Guérout, A. M., Boissard, S. & Mazel, D. Comparative  
875 study of class 1 integron and *Vibrio cholerae* superintegron integrase activities. *J*  
876 *Bacteriol* **187**, 1740–1750 (2005).
- 877 24. Vit, C. *et al.* Cassette recruitment in the chromosomal Integron of *Vibrio*  
878 *cholerae*. *Nucleic Acids Res* **49**, 5654–5670 (2021).
- 879 25. Richard, E., Darracq, B., Loot, C. & Mazel, D. Unbridled Integrons: A Matter of  
880 Host Factors. *Cells* **11**, 925 (2022).
- 881 26. Rowe-Magnus, D. A. & Mazel, D. Integrons: Natural tools for bacterial genome  
882 evolution. *Curr Opin Microbiol* **4**, 565–569 (2001).
- 883 27. Escudero, J. A. *et al.* Unmasking the ancestral activity of integron integrases  
884 reveals a smooth evolutionary transition during functional innovation. *Nat*  
885 *Commun* **7**, (2016).

- 886 28. Sawabe, T., Kita-Tsukamoto, K. & Thompson, F. L. Inferring the evolutionary  
887 history of vibrios by means of multilocus sequence analysis. *J Bacteriol* **189**,  
888 7932–7936 (2007).
- 889 29. Loot, C., Bikard, D., Rachlin, A. & Mazel, D. Cellular pathways controlling  
890 integron cassette site folding. *EMBO Journal* **29**, 2623–2634 (2010).
- 891 30. Loot, C. *et al.* The integron integrase efficiently prevents the melting effect of  
892 *Escherichia coli* single-stranded dna-binding protein on folded attC sites. *J*  
893 *Bacteriol* **196**, 762–771 (2014).
- 894 31. Grieb, M. S. *et al.* Dynamic stepwise opening of integron attC DNA hairpins by  
895 SSB prevents toxicity and ensures functionality. *Nucleic Acids Res* **45**, 10555–  
896 10563 (2017).
- 897 32. Loot, C. *et al.* Differences in integron cassette excision dynamics shape a trade-  
898 off between evolvability and genetic capacitance. *mBio* **8**, (2017).
- 899 33. Escudero, J. A. & Mazel, D. Genomic Plasticity of *Vibrio cholerae*. *Int Microbiol*  
900 **20**, 138–148 (2017).
- 901 34. Val, M.-E., Skovgaard, O., Ducos-Galand, M., Bland, M. J. & Mazel, D.  
902 Genome Engineering in *Vibrio cholerae*: A Feasible Approach to Address  
903 Biological Issues. *PLoS Genet* **8**, e1002472 (2012).
- 904 35. Skovgaard, O., Bak, M., Løbner-Olesen, A. & Tommerup, N. Genome-wide  
905 detection of chromosomal rearrangements, indels, and mutations in circular  
906 chromosomes by short read sequencing. *Genome Res* **21**, 1388–1393 (2011).
- 907 36. Rashid, M. H., Rajanna, C., Ali, A. & Karaolis, D. K. R. Identification of genes  
908 involved in the switch between the smooth and rugose phenotypes of *Vibrio*  
909 *cholerae*. *FEMS Microbiol Lett* **227**, 113–119 (2003).
- 910 37. Lim, B., Beyhan, S., Meir, J. & Yildiz, F. H. Cyclic-diGMP signal transduction  
911 systems in *Vibrio cholerae*: Modulation of rugosity and biofilm formation. *Mol*  
912 *Microbiol* **60**, 331–348 (2006).
- 913 38. Moxon, R., Bayliss, C. & Hood, D. Bacterial contingency loci: The role of  
914 simple sequence DNA repeats in bacterial adaptation. *Annual Review of Genetics*  
915 vol. 40 307–333 Preprint at  
916 <https://doi.org/10.1146/annurev.genet.40.110405.090442> (2006).
- 917 39. Wölflingseder, M., Tutz, S., Fengler, V. H., Schild, S. & Reidl, J. Regulatory  
918 interplay of RpoS and RssB controls motility and colonization in *Vibrio cholerae*.  
919 *International Journal of Medical Microbiology* **312**, (2022).
- 920 40. Dalia, A. B. RpoS is required for natural transformation of *Vibrio cholerae*  
921 through regulation of chitinases. *Environ Microbiol* **18**, 3758–3767 (2016).
- 922 41. Baharoglu, Z., Krin, E. & Mazel, D. RpoS Plays a Central Role in the SOS  
923 Induction by Sub-Lethal Aminoglycoside Concentrations in *Vibrio cholerae*.  
924 *PLoS Genet* **9**, (2013).
- 925 42. Falzone, C. J., Karsten, W. E., Conley, J. D. & Viola, R. E. L-Aspartase from  
926 *Escherichia coli*: substrate specificity and role of divalent metal ions.  
927 *Biochemistry* **27**, 9089–9093 (1988).
- 928 43. Strecker, A., Schubert, C., Zedler, S., Steinmetz, P. & Unden, G. DcuA of  
929 aerobically grown *Escherichia coli* serves as a nitrogen shuttle (L-  
930 aspartate/fumarate) for nitrogen uptake. *Mol Microbiol* **109**, 801–811 (2018).
- 931 44. Krin, E. *et al.* Expansion of the SOS regulon of *Vibrio cholerae* through  
932 extensive transcriptome analysis and experimental validation. *BMC Genomics*  
933 **19**, (2018).

- 934 45. Bochner, B. R., Gadzinski, P. & Panomitros, E. Phenotype microarrays for high-  
935 throughput phenotypic testing and assay of gene function. *Genome Res* **11**, 1246–  
936 1255 (2001).
- 937 46. Balaban, N. Q. *et al.* Definitions and guidelines for research on antibiotic  
938 persistence. *Nat Rev Microbiol* **17**, 441–448 (2019).
- 939 47. Harms, A., Fino, C., Sørensen, M. A., Semsey, S. & Gerdes, K. Prophages and  
940 Growth Dynamics Confound Experimental Results with Antibiotic-Tolerant  
941 Persister Cells. *mBio* **8**, (2017).
- 942 48. Zhu, J. *et al.* Quorum-sensing regulators control virulence gene expression in  
943 *Vibrio cholerae*. *Proc Natl Acad Sci U S A* **99**, 3129–3134 (2002).
- 944 49. Sang Ho, L., Hava, D. L., Waldor, M. K. & Camilli, A. Regulation and temporal  
945 expression patterns of *Vibrio cholerae* virulence genes during infection. *Cell* **99**,  
946 625–634 (1999).
- 947 50. Vaitkevicius, K. *et al.* A *Vibrio cholerae* protease needed for killing of  
948 *Caenorhabditis elegans* has a role in protection from natural predator grazing.  
949 *Proc Natl Acad Sci U S A* **103**, 9280–9285 (2006).
- 950 51. Sahu, S. N. *et al.* Genomic analysis of immune response against *Vibrio cholerae*  
951 hemolysin in *Caenorhabditis elegans*. *PLoS One* **7**, (2012).
- 952 52. Wang, Y. *et al.* Functional RelBE-Family Toxin-Antitoxin Pairs Affect Biofilm  
953 Maturation and Intestine Colonization in *Vibrio cholerae*. *PLoS One* **10**, (2015).
- 954 53. Teschler, J. K., Nadell, C. D., Drescher, K. & Yildiz, F. H. Mechanisms  
955 Underlying *Vibrio cholerae* Biofilm Formation and Dispersion. *Annu Rev*  
956 *Microbiol* **76**, 503–532 (2022).
- 957 54. Guerin, É. *et al.* The SOS response controls integron recombination. *Science*  
958 (1979) **324**, 1034 (2009).
- 959 55. Richard, E., Darracq, B., Loot, C. & Mazel, D. Unbridled Integrons: A Matter of  
960 Host Factors. *Cells* vol. 11 Preprint at <https://doi.org/10.3390/cells11060925>  
961 (2022).
- 962 56. Fonseca, É. L. & Vicente, A. C. Integron Functionality and Genome Innovation:  
963 An Update on the Subtle and Smart Strategy of Integrase and Gene Cassette  
964 Expression Regulation. *Microorganisms* **10**, 224 (2022).
- 965 57. Rowe-Magnus, D. A., Guerout, A. M., Biskri, L., Bouige, P. & Mazel, D.  
966 Comparative analysis of superintegrons: Engineering extensive genetic diversity  
967 in the vibrionaceae. *Genome Res* **13**, 428–442 (2003).
- 968 58. Rapa, R. A. & Labbate, M. The function of integron-associated gene cassettes in  
969 *Vibrio* species: The tip of the iceberg. *Front Microbiol* **4**, 1–7 (2013).
- 970 59. Rudolph, F. B. & Fromm, H. J. The purification and properties of aspartase from  
971 *Escherichia coli*. *Arch Biochem Biophys* **147**, 92–98 (1971).
- 972 60. Strecker, A., Schubert, C., Zedler, S., Steinmetz, P. & Uden, G. DcuA of  
973 aerobically grown *Escherichia coli* serves as a nitrogen shuttle (L-  
974 aspartate/fumarate) for nitrogen uptake. *Mol Microbiol* **109**, 801–811 (2018).
- 975 61. Schubert, C., Kim, N. Y., Uden, G. & Kim, O. Bin. C4-dicarboxylate  
976 metabolons: interaction of C4-dicarboxylate transporters of *Escherichia coli* with  
977 cytosolic enzymes. *FEMS Microbiol Lett* **369**, (2022).
- 978 62. Kanehisa, M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids*  
979 *Res* **28**, 27–30 (2000).
- 980 63. Edgar, R., Domrachev, M. & Lash, A. E. Gene Expression Omnibus: NCBI gene  
981 expression and hybridization array data repository. *Nucleic Acids Res* **30**, 207–10  
982 (2002).

- 983 64. Bland, M. J., Ducos-Galand, M., Val, M.-E. & Mazel, D. An att site-based  
984 recombination reporter system for genome engineering and synthetic DNA  
985 assembly. *BMC Biotechnol* **17**, 62 (2017).
- 986 65. Vit, C. *et al.* Cassette recruitment in the chromosomal Integron of *Vibrio*  
987 *cholerae*. *Nucleic Acids Res* **49**, 5654–5670 (2021).
- 988 66. Marvig, R. L. & Blokesch, M. Natural transformation of *Vibrio cholerae* as a tool  
989 - Optimizing the procedure. *BMC Microbiol* **10**, 155 (2010).
- 990 67. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat*  
991 *Methods* **9**, 357–359 (2012).
- 992 68. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several  
993 hundred kilobases. *Nat Methods* **6**, (2009).
- 994 69. Li, H. Aligning sequence reads, clone sequences and assembly contigs with  
995 BWA-MEM. (2013) doi:10.48550/arxiv.1303.3997.
- 996 70. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose  
997 program for assigning sequence reads to genomic features. *Bioinformatics* **30**,  
998 923–930 (2014).
- 999 71. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and  
1000 dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
- 1001 72. Bateman, A. *et al.* UniProt: the Universal Protein Knowledgebase in 2023.  
1002 *Nucleic Acids Res* **51**, D523–D531 (2023).
- 1003 73. Wu, T. *et al.* clusterProfiler 4.0: A universal enrichment tool for interpreting  
1004 omics data. *The Innovation* **2**, 100141 (2021).
- 1005 74. Herencias, C. *et al.* Collateral sensitivity associated with antibiotic resistance  
1006 plasmids. *Elife* **10**, 1–13 (2021).
- 1007 75. García-Pastor, L., Sánchez-Romero, M. A., Gutiérrez, G., Puerta-Fernández, E.  
1008 & Casadesús, J. Formation of phenotypic lineages in *Salmonella enterica* by a  
1009 pleiotropic fimbrial switch. *PLoS Genet* **14**, e1007677 (2018).
- 1010
- 1011
- 1012