Cassette recombination dynamics within chromosomal integrons are regulated by toxin-antitoxin systems

Egill Richard, Baptiste Darracq, Eloi Littner, Claire Vit, Clémence Whiteway, Julia Bos, Florian Fournes, Geneviève Garriss, Valentin Conte, Delphine Lapaillerie, et al.

MICROBIOLOGY

Cassette recombination dynamics within chromosomal integrons are regulated by toxin-antitoxin systems

Egill Richard[1,2], Baptiste Darracq[1,2], Eloi Littner[2,3,4], Claire Vit[1,2], Clémence Whiteway[1], Julia Bos[1], Florian Fournes[1], Geneviève Garriss[1], Valentin Conte[1], Delphine Lapaillerie[5,6], Vincent Parissi[5,6], François Rousset[7], Ole Skovgaard[8], David Bikard[7], Eduardo P. C. Rocha[3], Didier Mazel[1]*, Céline Loot[1]*

Integrons are adaptive bacterial devices that rearrange promoter-less gene cassettes into variable ordered arrays under stress conditions, thereby sampling combinatorial phenotypic diversity. Chromosomal integrons often carry hundreds of silent gene cassettes, with integrase-mediated recombination leading to rampant DNA excision and integration, posing a potential threat to genome integrity. How this activity is regulated and controlled, particularly through selective pressures, to maintain such large cassette arrays is unknown. Here, we show a key role of promoter-containing toxin-antitoxin (TA) cassettes as systems that kill the cell when the overall cassette excision rate is too high. These results highlight the importance of TA cassettes regulating the cassette recombination dynamics and provide insight into the evolution and success of integrons in bacterial genomes.

INTRODUCTION

Integrons are bacterial recombination systems that are capable of capturing, stockpiling, and reordering mobile elements called cassettes, thus regulating their expression. They were originally found to be the genetic systems responsible for the acquisition of antimicrobial resistance determinants in some mobile elements (*1*). Integrons contain a stable platform and a variable cassette array (Fig. 1A). The stable platform is defined by (i) the integrase gene (*intI*) under the control of its promoter P_{int}, (ii) the *attI* integration site, and (iii) the cassette P_C promoter. The cassettes in the array are generally composed of a promoter-less gene associated to a recombination site called *attC*. Only the first few cassettes (those closest to the P_C promoter) can be expressed, while the rest represent silent and valuable functions for the cell (*2*). These cassettes can be excised and then reintegrated at the *attI* integration site and therefore become expressed (*3*). This cassette shuffling ensures a combinatorial phenotypic diversity which allows bacteria to screen for the set of functions that would optimize its survival in a given environment. We previously found that the integrase, which catalyzes the cassette shuffling, is only expressed in stress conditions (*4*). For this reason, integrons are described as "on demand" adaptation systems (*5, 6*).

The integron integrase (IntI) has a very singular place among the broad family of the tyrosine recombinases (*7*). While it recognizes the *attI* site under its double-stranded form through its primary sequence, the *attC* sites are recombined as a single-stranded form (*8, 9*). Although both the bottom and the top strands of an *attC* site can form a secondary structure, the recombination of the bottom strand (*bs*) is about 10^3 more efficient than that of the top (*9*). This ensures a correct orientation of the cassette upon integration at the *attI* site, allowing its expression by the P_C promoter (*10*).

We distinguish two types of integrons: mobile integrons (MIs), carried by conjugative plasmids and mobilizable among bacteria, and sedentary chromosomal integrons (SCIs). In addition to their sedentary nature, the main feature that distinguishes SCIs from MIs is the size of their cassette array. While MIs do not typically store more than 10 cassettes, SCIs often contain dozens of cassettes—with the largest SCI from *Vibrio vulnificus* encoding 301—thus encompassing a substantial fraction of their host genomes (*11, 12*). SCI cassettes constitute an almost infinite variety of functions extending far beyond antibiotic resistance and linked to bacterial key adaptive functions (*13*). The high genetic capacitance of SCIs allows them to serve as a repository of cassettes for MIs to compile the most relevant repertoire of cassettes in a specific environment (*6, 14*). The remarkable adaptive success of the integron system is partly due to this intricate connection between SCIs and MIs. Despite all that, 25 years after their discovery, SCIs remain enigmatic genetic structures. We still ignore how such structures can accumulate so many silent cassettes and be maintained in bacteria to provide such an adaptive device for future stresses.

Among the cassettes of SCIs, a distinct type is notable, encoding toxin-antitoxin (TA) systems and containing their own promoter (*15–18*). TA systems encode a stable toxin that inhibits an essential cellular process and an unstable antitoxin that counteracts its cognate toxin. When the expression of the gene pair is stopped, the antitoxin is degraded, resulting in cell death or growth arrest (*19*). The presence of TA systems within SCI arrays likely prevents cassette deletion through large chromosomal rearrangements thus maintaining the array (*15, 16*). Nevertheless, this sole factor is not sufficient to understand how SCIs are kept so massive. In stress conditions, when the integrase is expressed, the cassette excision rate must be tightly regulated. To ensure integron cassette accumulation, the cassette excision rate must be kept below the integration rate, and to maintain the large integrons, the two rates must be at least similar. An imbalance in favor of the excision rate would lead in fine to the loss of most silent cassettes. Here, directly linked to their cassette nature, we demonstrated that the TAs play an

[1]Institut Pasteur, Université Paris Cité, CNRS UMR3525, Unité Plasticité du Génome Bactérien, 75015 Paris, France. [2]Sorbonne Université, ED515, F-75005 Paris, France. [3]Institut Pasteur, Université Paris Cité, CNRS UMR3525, Microbial Evolutionary Genomics, 75015 Paris, France. [4]DGA CBRN Defence, 91710 Vert-le-Petit, France. [5]University of Bordeaux, Fundamental Microbiology and Pathogenicity Laboratory, CNRS, UMR 5234, SFR TransBioMed, Bordeaux, France. [6]Viral DNA Integration and Chromatin Dynamics Network (DyNAVir), France. [7]Institut Pasteur, Université Paris Cité, CNRS UMR6047, Synthetic Biology, 75015 Paris, France. [8]Department of Science, Systems and Models, Roskilde University, Roskilde DK-4000, Denmark.
*Corresponding author. Email: celine.loot@pasteur.fr (C.L.); didier.mazel@pasteur.fr (D.M.)

additional role regulating the cassette recombination dynamics. We used as model the *Vibrio cholerae* N16961 strain responsible for the ongoing seventh cholera pandemic and which contain a particularly massive SCI [~130 kb and 179 cassettes (*20*)]. By developing a genetic assay to measure the SCI cassette dynamics at the population level, we demonstrate that the cassette excision rate is markedly increased all along the SCI when we "artificially" inverted it on-site. We show that the inversion of the SCI is also associated to a strong growth defect correlated with a high cell mortality rate in the presence of a functional integrase only. By successfully inactivating all the 19 TA cassettes within the inverted SCI (*17, 18*) (Fig. 1B), using highly multiplexed CRISPR base editing, we demonstrate that these cassettes are in a large part responsible for the observed mortality phenotype.

Our results highlight a role for TA cassettes in SCI maintenance via an integrase-dependent mechanism that we termed "cassette loss killing." Under conditions where the cassette excision rate is high and could compromise the maintenance of large SCIs, TA cassettes are also excised at a high rate as individual cassettes, thus killing cells. Through this, TAs exert a selective pressure that drives integrons

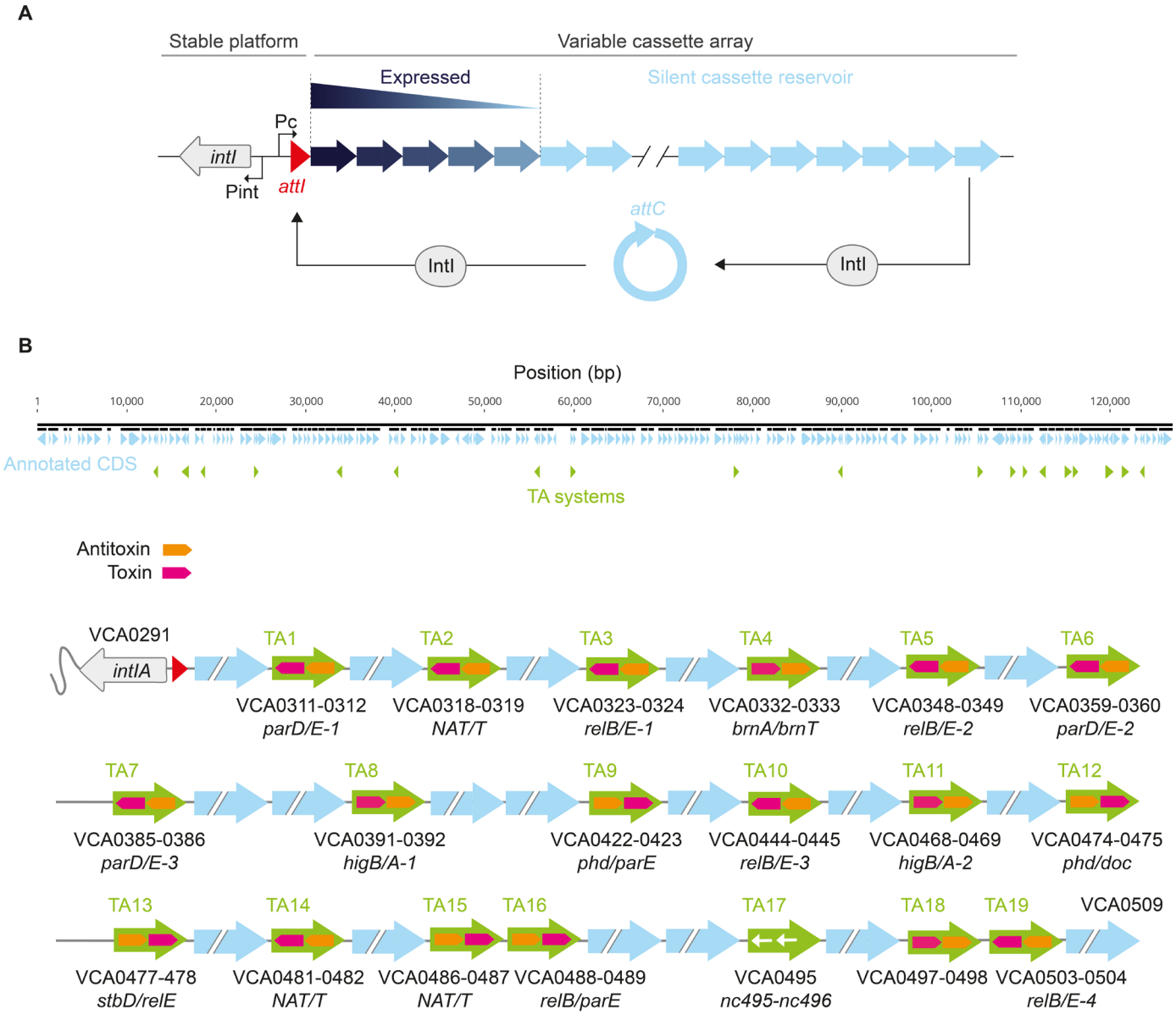

Fig. 1. The integron system and the TA cassettes. (**A**) Schematic representation of the integron. The stable platform consists of a gene coding for the integrase (*intI*, in gray) and its promoter P_{int}, the cassette integration site *attI* (in red) and the cassette promoter P_C driving the expression of the downstream cassettes along an expression gradient. Cassettes are displayed as arrows where the tip represent the *attC* site. Their color intensity represents their level of expression from dark blue (highly expressed) to light blue (not expressed). Nonexpressed cassettes can be excised through an intramolecular *attC* × *attC* reaction and be reintegrated in first position near the P_C promoter through an intermolecular *attC* × *attI* reaction and become expressed. (**B**) Representation of the Sedentary Chromosomal Integron of *V. cholerae* and the repartition of the various toxin-antitoxin (TA) systems numbered from 1 to 19. The detail of the TA systems (name, orientation in the array, and order of toxin/antitoxin) is represented. NAT/T, *N*-acetyltransferase/transcription factor.

to adopt a configuration in which the excision rate of their cassettes is kept low. This has probably led to the preferential orientation of SCIs with respect to the replication fork that we previously observed (*14*) and which limits the rate of cassette excision. The discovery of this additional type of regulation in integrons adds to the already known regulatory network. It enables silent cassettes to be accumulated and maintained, making them available to produce combinatorial phenotypic diversity and respond to future stresses.

RESULTS

SCIs are specifically enriched in TA systems

To better understand the role of the TAs as cassettes in SCI arrays and to assess whether SCI size and TA number are correlated, we systematically explored all TA carried by SCIs. We used IntegronFinder 2.0 (*12*) to detect all of the sedentary integrons in complete genomes of the RefSeq National Center for Biotechnology Information (NCBI) database. We identified 398 genomes with at least one SCI [defined as an integron comprising strictly more than 10 cassettes, as suggested in (*12*)]. In each of these genomes, we screened TA systems with TASmania (*21*). All the data of the analysis are presented in table S1. Then, we compared the number of toxins and antitoxins detected in the SCIs with the total number of cassettes in each of the 398 genomes. We found a positive correlation between the two, indicating that larger SCIs have more TA cassettes (Fig. 2A). To assess the specific enrichment of TAs in SCIs when compared to the rest of the genome, we performed in each isolate a contingency table analysis. Of the 398 isolates, 246 verified the hypothesis of a TA enrichment within SCI ($P < 0.05$, Fisher test with Benjamini-Hochberg correction). Of the remaining isolates, 87 genomes lack genes encoding toxins or antitoxins and 65 encode at least one of them but do not show significant enrichment. As most of the genomes with a significant overrepresentation of TAs in SCIs belonged to the *Vibrio* genus (Fig. 2A), we decided to analyze them thoroughly. Among the 280 *Vibrio* genomes, 225 verified the hypothesis of a TA enrichment within SCI (Fig. 2B). This may reflect a particularly strong effect in this genus, but we cannot exclude that it simply results from other SCIs having too few cassettes (median: 21.5 *attC*) to provide a statistically significant signal of overrepresentation compare to the *Vibrio* genus (median: 123.5 *attC*). We studied the patterns of this overrepresentation in function of the phylogeny of *Vibrio*. The cladogram used to group species was adapted from the phylogeny reconstructed by Sawabe and colleagues (*22*). We observe that the enrichment signal is scattered across the tree, highlighting the important role of TAs in a broad range of *Vibrio* species (Fig. 2C). This association holds true with the very stringent Bonferroni correction, although with a less sparse distribution on the phylogenetic tree, centered around a clade comprising *V. cholerae* (fig. S1). In this species, some TA systems (in the ParDE family) are conserved in most isolates (82 of 86). This conservation may suggest a long period of coevolution of integrons and TA systems, at least in *V. cholerae*. Although it does not exclude the regular turnover of some other TA systems, TA cassettes may have contributed to the evolution of SCIs to a large extent.

The inversion of the SCI in *V. cholerae* markedly increases the cassette excision rate in the array

Evolutionary factors have enabled the integron cassette accumulation and maintenance to lead to the emergence of massive and stable chromosomal integrons. To specify these factors and determine whether TA cassettes have a role in this evolution process, we decided to render the cassette array of large integrons unstable and determine how these "genetically modified" integron structures can be counterselected. To do this, we inverted the *V. cholerae* SCI on place (SCI Inv strain) and reinverted it back to its original orientation (SCI Reinv strain) (Fig. 3A, fig. S2A, and Materials and Methods). We previously establish that when the recombinogenic *bs* of *attC* sites are located on the lagging strand template, where discontinuous replication leaves single-strand gaps, this favors *attC* site folding and cassette excision (*14, 23*). We therefore hypothesized that the natural SCI orientation, where the *bs* of *attC* sites are located on the leading strand template, disfavors cassette excision events and that the SCI inversion would be a way to make the cassette array unstable.

We first monitored the effect of SCI inversion on cassette shuffling by cultivating overnight *V. cholerae* with the SCI in both orientations, in the presence or absence of IntIA (Fig. 3B). A polymerase chain reaction (PCR) was then performed on independent clones with primers located in the *attIA* integration site and in the first cassette of the array (VCA0292) with an expected size of 275 bp in the absence of shuffling (Fig. 3B). Very few shuffling events could be detected in the SCI Reinv clones (3 of 48), whereas we observed large amplicons in 46 of 48 of the SCI Inv clones, testifying to the integration of multiple cassettes in the *attIA* site (Fig. 3C). To study the effect of SCI inversion on cassette integration rate in the *attIA* site alone (uncoupled from cassette excision), we performed our classical conjugation assay which delivers cassettes into *V. cholerae* receptor strains (Materials and Methods) (*3*). We obtained an identical integration rate of these cassettes in the *attIA* site whatever the orientation of the SCI (fig. S3). Our results therefore demonstrated that inversion of the SCI does not alter the rate of cassette integration but does lead to a substantial increase in cassette excision.

To monitor the cassette excision rate at the population level in strains with SCI in both orientations, we designed an SCI-wide cassette excision assay. Like the previous shuffling assay, we grew *V. cholerae* strains overnight in the absence or presence of integrase but, this time, in the absence of the *attIA* site to focus solely on excision events. To assess cassette rearrangements along the array, we used long-read sequencing (Materials and Methods and Fig. 3D). Each read captured the state of a segment of the SCI in a single cell and allowed an easy detection of cassette deletion on this segment. For each population, a total of ~80,000 reads with an average length of ~6.4 kb was produced. For the *V. cholerae* population corresponding to the SCI inverted strain expressing integrase, we identified 478 cassette excision events, encompassing 120 cassettes where excision was observed at least once leading to a cassette excision rate of 3×10^{-2} (Fig. 3E). For the populations of strains with the SCI in the native orientation with and without integrase and in the strain with an inverted SCI in the absence of the integrase, the excision rates are below 7×10^{-5} (i.e., the limit of detection; Fig. 3E and Materials and Methods). These results reveal a direct link between the inversion of the SCI and a highly increased cassette excision frequency.

The inversion of the SCI in *V. cholerae* leads to a growth defect in the presence of a functional integrase

If a single inversion is sufficient to increase the excision cassette rate and empty a SCI of all its cassettes in a few generations, then the massive and silent structures that are SCIs would probably not exist.

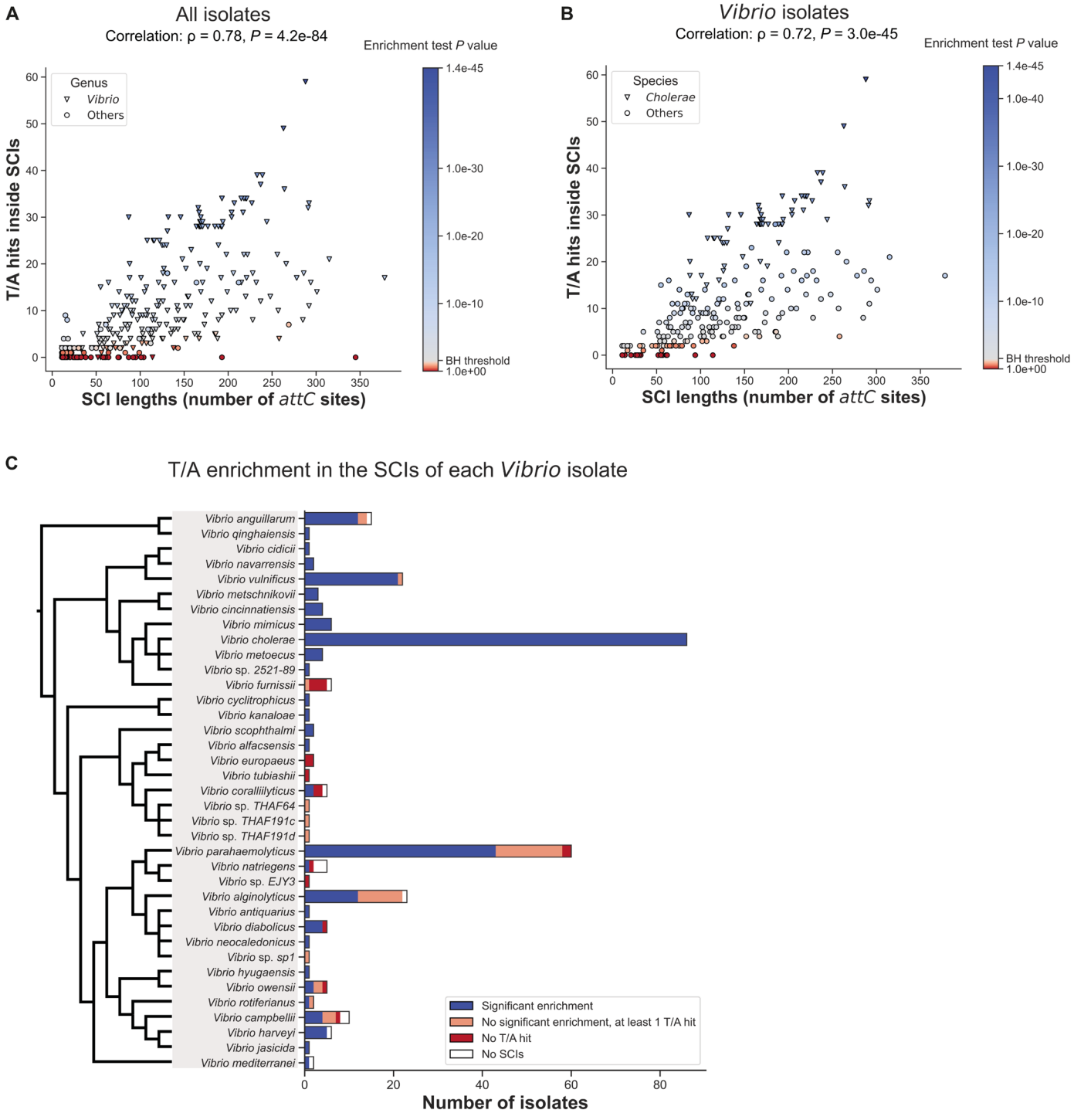

Fig. 2. TA landscape in SCIs. (**A** and **B**) Each point represents a single NCBI RefSeq isolate harboring at least one SCI (>10 *attC* sites). The *x* axis displays the total number of *attC* sites in the (potentially multiple) SCI of the isolate. The *y* axis exhibits the total number of TASmania toxin and antitoxin HMM hits in the SCI of the isolate (*e*-value < 10^{-3}). The Spearman nonparametric rank correlation coefficient is shown at the top of the graph, together with its significance *P* value. Point hue indicates the *P* value of Fisher's test for TASmania hits enrichment in SCI compared to the rest of the genome (scale on the right of the graph, significant enrichment in blue). The Benjamini-Hochberg correction for multiple testing was applied (see below for the corresponding computed thresholds). (A) All the isolates harboring at least one SCI are displayed. The shape of each point on the graph represents whether the isolate belongs to the *Vibrio* genus (triangles) according to the NCBI taxonomy or to another genus (circles). Benjamini-Hochberg correction threshold: 0.031. (B) Same graph as (A), but only with isolates belonging to the *Vibrio* genus. The shape of each point on the graph represents whether the isolate belongs to the *V. cholerae* species (triangle) or to another species (round). Benjamini-Hochberg correction threshold: 0.036. (**C**) Distribution of *Vibrio* isolates by species according to the toxin/antitoxin enrichment inside the SCI they harbor. T, toxin; A, antitoxin; BH, Benjamini-Hochberg correction.

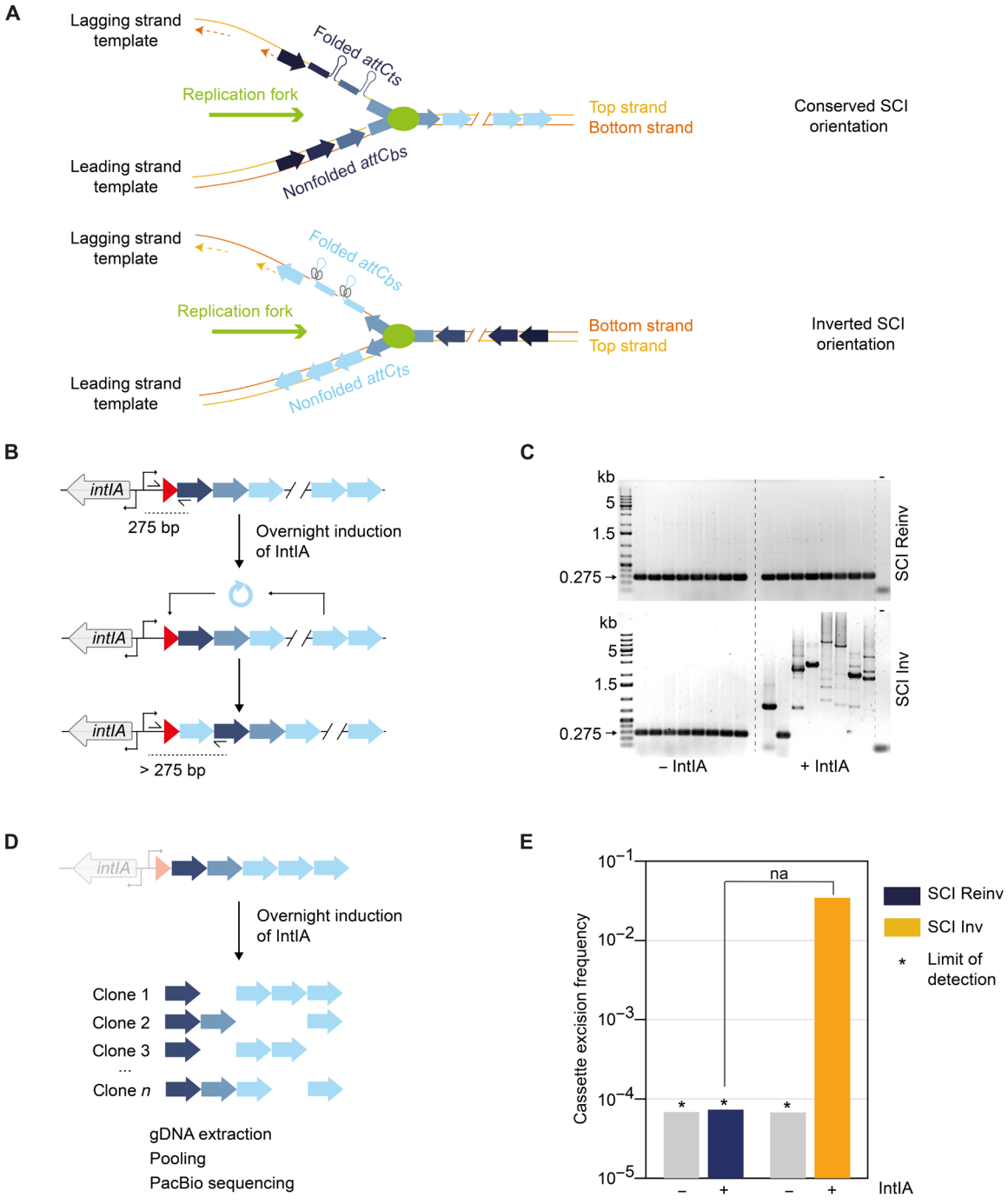

Fig. 3. Cassette shuffling and excision assays in the SCI inverted and reinverted *V. cholerae* strains. (**A**) Mechanistic insight on the issue of integron orientation. An array of cassettes is represented while it is replicated. In the lagging strand, Okazaki fragments are represented as dotted lines, leaving stretches of single-stranded DNA on the correspondent template. In their conserved orientation, SCI recombinogenic $attC_{bs}$ are carried by the continuously replicated leading strand template, while in the inverted orientation, they are carried by the discontinuously replicated lagging strand template. (**B**) Schematic representation of the assay designed to detect cassette shuffling. (**C**) PCR analysis of the cassettes shuffled in first position in individual clones in the presence (+ IntIA) or in the absence of the integrase (– IntIA) for one overnight. The expected PCR product in the absence of shuffling is 275 base pairs (bp). Longer bands are evidence of shuffling events. Multiple bands may come from the random slippage of the DNA polymerase in the presence of highly repeated sequences (multiple *attC* sites in one amplicon). (**D**) Schematic representation of the assay designed to detect cassette excision. A strain of *V. cholerae* Δ*attIA* is used to initiate a culture composed of clones experiencing various cassette excision events. Clones from this heterogeneous population are represented (from 1 to *n*), each with different cassettes missing. (**E**) Cassette excision frequency in the array of the SCI. The cassette excision frequency is calculated for the SCI inverted (SCI Inv) strain expressing the integrase (Materials and Methods). No excision event could be detected in the other conditions; hence, the limit of detection (*) is represented instead. The gray bars represent the frequencies obtained in the absence of integrase in the SCI Inv and SCI Reinverted (Reinv) strains. Statistical comparisons (Student's *t* test) are as follows: na (not applicable).

We reasoned that there may be counterselective forces in situations in which the cassette array is unstable, preventing, for example, a spontaneous inversion of the SCI. We set out to measure growth of the SCI Inv strain (Fig. 4A). While we did not observe any growth defect in the SCI Inv strain in the absence of integrase, we found an important growth defect upon expression of the integrase in this strain compared to the SCI Reinv strain (Fig. 4A). Growth rate is ~35% lower in the SCI Inv strain expressing the integrase compared to the controls (Fig. 4B). We observed the same growth defect while inducing the endogenous integrase through the SOS response using sublethal concentrations of ciprofloxacin (Fig. 4C), confirming our previous conclusions in this more natural setting. To further evaluate the cost of the inversion of the SCI, we set out a competition experiment where SCI Inv and Reinv strains were cocultivated for 24 hours starting with an initial 1:1 ratio (Fig. 4D). While no substantial deviation to the original ratio between SCI Inv and Reinv strains could be observed in the absence of integrase, the expression of the integrase in those strains quickly led to a strong disadvantage of the SCI Inv strain compared to the SCI Reinv. This confirms the high fitness cost of the inversion of the SCI of *V. cholerae* in the presence of the integrase. Expressing a catalytically inactive integrase that binds to the $attC_{bs}$ (fig. S4) without cleaving, it did not produce any growth defect (Fig. 4E), ruling out a scenario where the mere binding of more accessible $attC_{bs}$ in the SCI Inv strain could explain the growth defect. Hence, the cleavage activity of the integrase was necessary to induce a growth defect in the SCI Inv. Yet, the integrase can cleave both *attIA* and *attC* sites so to determine whether it was the increased shuffling or the cassette excision that imposed such a strong cost of SCI inversion, we performed growth curve in cells lacking the *attIA* site (Fig. 4F). We still observed a growth defect, hence confirming that it is the increased cassette excision rate that explained the growth defect of the SCI Inv strain.

The growth defect in the SCI Inverted strain in the presence of integrase is associated with increased cell death

The growth defect previously observed in the integrase-expressing SCI Inv strain could result from a slower division rate of each individual cell or a higher mortality rate at each generation. To discriminate between these two possibilities, we needed information at the single-cell level. We therefore set out to observe in wide-field microscopy live cells of our different strains growing on an agarose pad mimicking the liquid medium used for generating the growth curves. An exponentially growing culture of the SCI Inv and Reinv strains was diluted to display individual cells onto the agarose pads. We recorded the development of microcolonies for 170 min (Fig. 5A and movies S1 to S4), and the number of cells per microcolony in

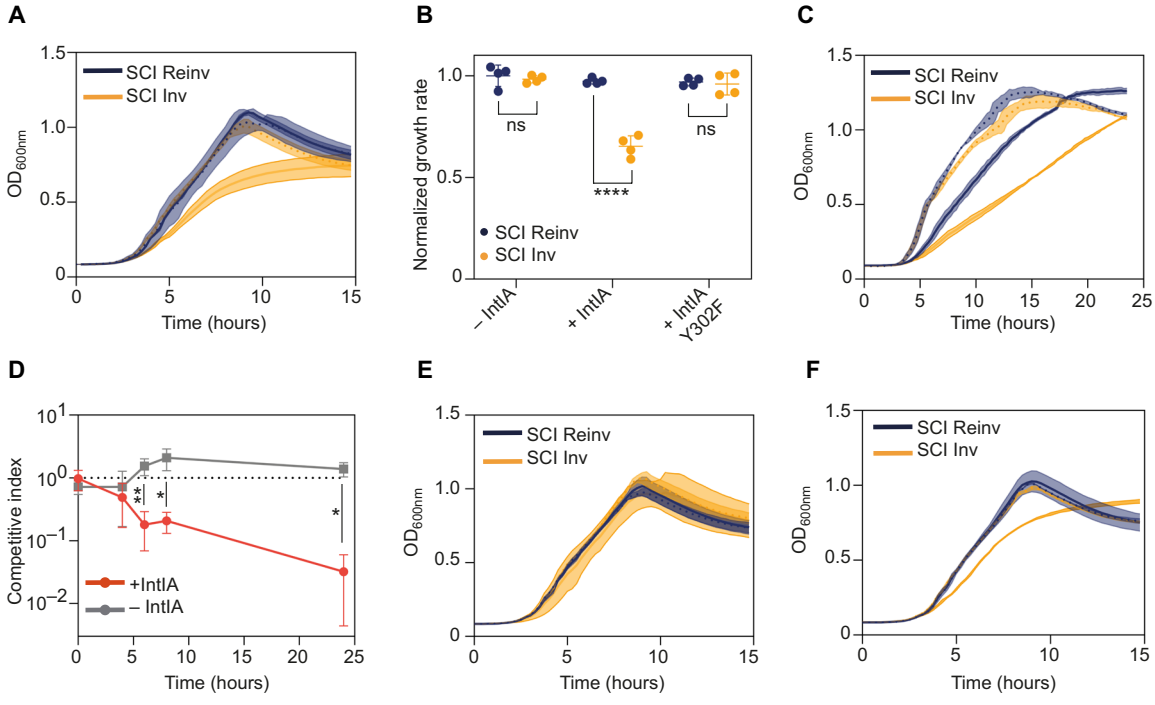

Fig. 4. Growth of the SCI inverted and reinverted *V. cholerae* strains. (**A**) Growth curve of SCI Inv (orange) or SCI Reinv (blue) strains in the presence (full lines) or in the absence of IntIA (dotted lines). Curve corresponds to the mean of four independent experiments ($n = 4$), and the shade corresponds to the SEs at each time point. (**B**) Normalized growth rate of SCI Inv and SCI Reinv strains in the absence of integrase (– IntIA) and in the presence of the WT (+ IntIA) or the catalytically inactive integrase (+ IntIA Y302F). Growth rates are normalized by the mean growth rate of the SCI Reinv strain in the absence of IntIA. Horizontal bars show the mean of four independent experiments ($n = 4$, individual plots), and error bars show the SD. Statistical comparisons (Student's *t* test) are as follows: ns (not significant) and ****$P < 0.0001$; all two-sided. (**C**) As in (A) but in the presence (full lines) or in the absence (dotted lines) of a subinhibitory concentration of ciprofloxacin inducing the endogenous IntIA expression. (**D**) Competitive index of the SCI Inv strain compared to the Reinv strain in the absence (gray) or in the presence of IntIA (red) in function of time (24 h of cocultures). Each point on the curves represents the mean of three independent experiments ($n = 3$). Error bars show the SD. Statistical comparisons (Student's *t* test) showed not significant differences between + IntIA and – IntIA for 0 and 4 hours and significant differences for 6, 8, and 24 hours as follows: *$P < 0.05$ and **$P < 0.01$. (**E**) As in (A) but in the presence of the catalytically inactive IntIA (full lines) or in the absence of IntIA (dotted lines). (**F**) As in (A) but in SCI Inv and Reinv strains lacking the *attIA* site. For (C), (E), and (F), curve corresponds to the mean of three independent experiments ($n = 3$).

the SCI Inv and Reinv strains expressing the integrase was determined as a measure of growth. We observed that the SCI Inv strain displayed increased heterogeneity in growth across microcolonies resulting in a broader distribution of microcolony sizes at the end of the run (from $t = 120$ min; Fig. 5B and movies S3 and S4). This reflects the fact that, whereas the SCI Reinv colonies generally follow an expected developmental trajectory, in the SCI Inv strain, many microcolonies stop growing after a few divisions or do not grow at all. Some events of cell lysis, characterized by sudden release by some cells of diffracting matter (*24*), could also be observed specifically in the SCI Inv strain (Fig. 5C and movie S5). These observations favor a model where the expression of the integrase in the SCI Inv strain induces cell death in a part of the population.

We tested this model in a live and dead assay where a differential coloration of the cells allows discrimination between living (green) and dead (red) cells in a population (Fig. 5D). We observed a clear increase of the proportion of dead cells in culture of SCI Inv strain expressing the integrase compared to the control conditions. We quantified the cell mortality and observed a high percentage of dead cells (~16%) in the SCI Inv strain expressing the integrase and a basal mortality rate (~1 to 5%) in the other strains (Fig. 5E and table S2). Expression of a catalytically inactive integrase in the SCI Inv strain resulted to a mortality rate similar to the basal level (Fig. 5E and table S2), consistent with the observation that was made in the previous growth measurement (Fig. 4, B and E). These data indicated that inversion of the SCI led to a high mortality rate that was dependent on the activity of the integrase.

The fitness defect in the SCI Inverted strain is largely explained by the excision of TA systems

The fact that SCI inversion results in increased cassette excision and that this is associated with increased growth arrest and cell death

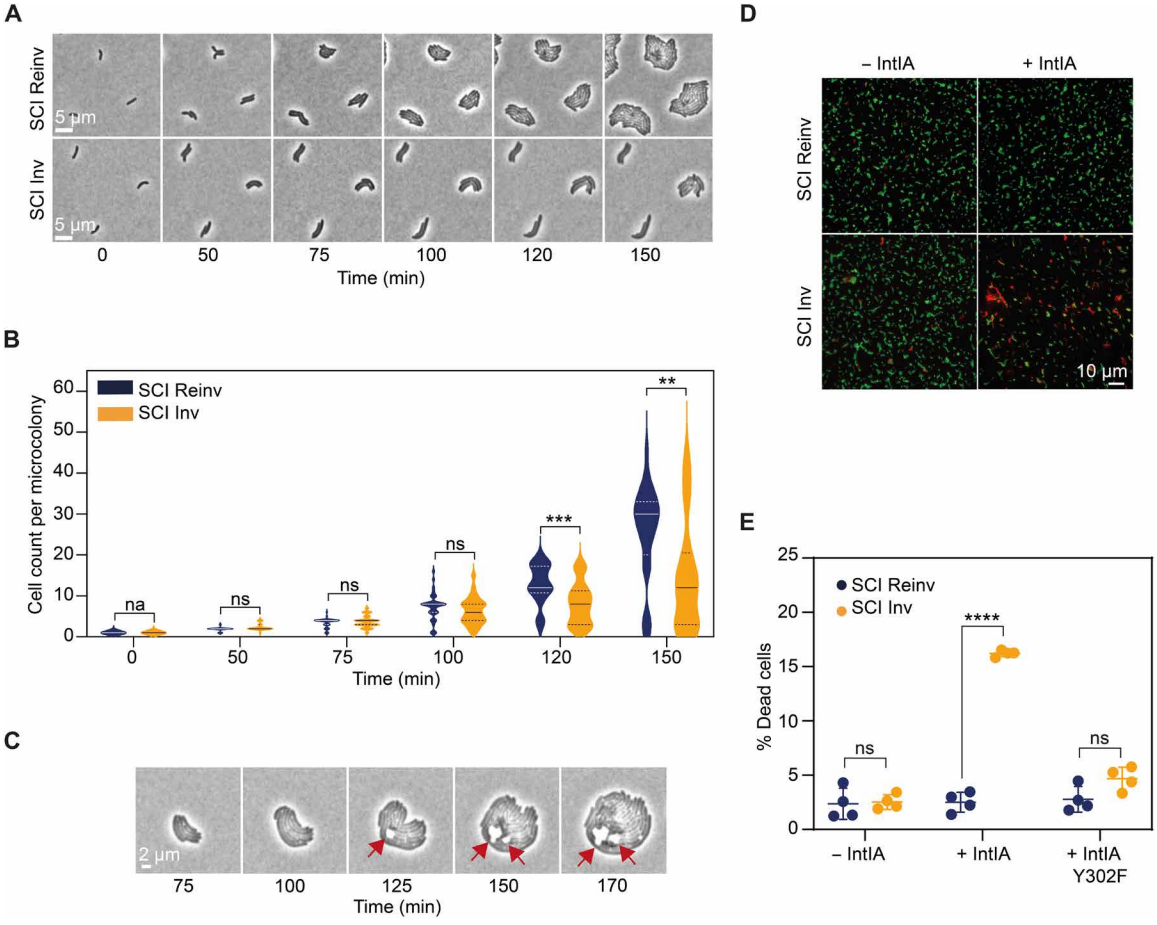

Fig. 5. Cell viability of the SCI inverted and reinverted *V. cholerae* strains. (**A**) Bright-field microscopy images taken from 170-min time-lapse series of live cells growing onto a Mops-rich agarose pad. A representative sample of three microcolonies of SCI Inv or SCI Reinv strains is shown. Scale bars, 5 µm. (**B**) Microcolony growth of the SCI Inv and Reinv strains onto Mops-rich agarose pads. Violin plots show the distribution of the number of cells per microcolony (27 for SI Inv and 34 for SCI Reinv strains) tracked during 170 min. Median (full line) and quartiles (dotted line) are shown. Statistical comparisons (Mann-Whitney test) are as follows: na (not applicable), ns (not significant), ***$P < 0.001$, and **$P < 0.01$. (**C**) Representative example of cell lysis events (red arrows) captured during the time-lapse experiments of live SCI Inv cells growing on Mops-rich agar pads and expressing the integrase. Lysis events were not observed in the SCI Reinv strain expressing the integrase. Scale bar, 2 µm. (**D**) Fluorescence microscopy images resulting from the live and dead assay on the SCI Inv and Reinv strains expressing or not the integrase. Green fluorescence indicates live cells (SYTO-9), whereas red fluorescence [propidium iodide (PI)] indicates dead cells. Scale bar, 10 µm. (**E**) Quantification of cell death as a measure of cells positively stained with PI over the total number of cells counted (~2500 cells per replicate). Horizontal bars show the mean of four independent experiments ($n = 4$, individual plots), and error bars show the SD. Statistical comparisons (Student's *t* test) are as follows: ns (not significant) and ****$P < 0.0001$; all two-sided.

dependent on integrase catalytic activity led us to hypothesize that the increased excision of some cassettes could be deleterious for the cell. The increased excision of the cassettes encoding TA systems, alongside all the nonexpressed cassettes, could be at the origin of the increased growth arrest and cell death observed in the SCI Inv strain. Analysis of *attC* sequences and pfold determination indicate that TA systems are as likely to be excised as other cassettes (Materials and Methods and fig. S5A). In addition, we observe excision of the TA-encoding cassettes when they are duplicated (fig. S5B), suggesting that the nonduplicated TA cassettes can also be excised but result in such a large fitness defect that we cannot observe their excision in our dataset. By combining prior approaches (*25, 26*), we designed a CRISPR-based cytosine base editors that allowed us to inactivate almost all the TA cassettes of the array in one go. We designed 12 guides to target 15 toxins (3 of which are duplicated) such that a STOP codon was introduced, resulting in at least a one-third truncation of each toxin (Materials and Methods and fig. S6). The remaining four toxins, for which we were unable to design a suitable guide, were inactivated by introducing a stop codon through the more classical allelic exchange method (*27*). It resulted in *V. cholerae* strains where all the TA within the SCI were inactivated (TAi).

We next assessed the effect of TA inactivation on growth of both SCI Inv and Reinv strains. In the absence of integrase (Fig. 6A and fig. S7A), we do not observe substantial differences between all the strains whether the SCI is inverted or not. In the presence of integrase, the growth defect characteristic of the SCI Inv strain appears to be largely rescued upon inactivation of the TA systems (Fig. 6B and fig. S7A). Performing a competition assay, we found again, in the presence of integrase, the SCI Inv to be severely out competed by the SCI Reinv strain, but this was alleviated in the absence of functional TA systems (Fig. 6C). Moreover, the SCI Inv TAi outcompetes the SCI Inv strain by more than a factor of three in the presence of integrase (Fig. 6D). A cell viability assay and a quantification of dead cells showed that the improved fitness of the SCI Inv TAi strain in the presence of integrase compared with the SCI Inv strain was associated with decreased cell death (Fig. 6, E to H, and table S2). For more clarity, we reported in these figures the previous results presented in Fig. 5 [i.e., TA wild type (WT)].

Last, although the restoration of growth is important in the absence of TA systems in the SCI Inv strain, the phenotypic rescue is not complete. We reasoned that the intense recombination activity throughout the 130 kb inverted SCI could also affect the replication process, explaining this gap. To address this, we performed marker frequency analysis (MFA), a technique that gives access to the replication pattern by deep sequencing of genomic DNA (gDNA) extracted from an exponentially growing bacterial culture (*28–30*). Because of the absence of an active replication termination mechanism in *V. cholerae*, a replication slow-down in one replichore should lead to a shift of the termination of replication locus (*ter*) (*31*). We did not observe a *ter* shift in the *chr2* containing the inverted SCI in the presence of the WT integrase (fig. S8A), meaning that replication is not altered in this strain. As negative controls, we tested the SCI Reinv strains in the presence and absence of the WT integrase. We also tested the Y302F integrase, which, although catalytically inactive, is still able to bind the *attC* sites. We did not observe a *ter* shift in any of those strains. Last, to be able to probe even a very minor effect on replication, we decided to relocate the SCI near the *ori* of chr 1 (*ori1*; fig. S2B and Materials and Methods). The *V. cholerae* SCI is endogenously located near the native *ter* of *chr2*, which only allows to detect a slight effect of its inversion on the replication pattern. On the contrary, any replication slow-down when the SCI is located near *ori1* such that its $attC_{bs}$ are located on the lagging strand template (SCI lag, equivalent to the SCI Inv) compared to the other orientation (SCI lead, equivalent to the SCI Reinv) would lead to a substantial shift of the downstream termination of chromosome 1 (*ter1*). We did not observe neither a *ter1* shift nor a replication slow-down in the *chr1* region after the SCI in lag orientation compared to the lead one and in the presence of integrase (fig. S8A), while the relocated SCI lag strain presents a large fitness defect in these conditions (fig. S8B). All these observations rule out any effect of a disruption of the replication process by the intense cassette dynamics in the SCI inverted strains expressing integrase.

In an attempt to explain the incomplete rescue of growth in the SCI Inv TAi, we tested the SOS response level of induction by performing, in our strains, a digital reverse transcription PCR (RT-dPCR) using the *recA* gene expression as proxy of the SOS induction (*32*). We demonstrated that the SOS response is largely activated in both SCI Inv and SCI Inv TAi strains in the presence of integrase, while we do not see any SOS induction in the native orientation of the SCI (fig. S7B). Thus, these results show that inversion of the SCI leads to a stress for the cell when the integrase is expressed, independently of TA systems. We sequenced single clones stemming from the culture of four SCI Inv TAi-independent strains to assess their cassette content after an overnight culture in the presence of the integrase (approximately 10 generations). While two clones had lost 9 and 11% of their cassette content (i.e., respectively 16 and 20 cassettes), the other two had lost 83 and 52% of their cassette content (i.e., respectively 104 and 94 cassettes). This confirms the intense cassette recombination activity when the SCI is inverted (see also Fig. 3, C and E). Although the precise origin of the specific SOS induction observed in SCI Inv strains is unknown, it is dependent on the integrase activity on cassette recombination, as the recombination rate is visibly much higher in this context than in the native one. The SOS induction could explain why we did not completely rescue the growth in the SCI Inv TAi strain.

We conclude that the growth defect in the SCI inverted strains is mostly due to the increased excision of TA cassettes acting as "addictive" systems killing the cell. Nonetheless, we cannot exclude a slight effect of SOS induction linked to the intense cassette dynamics.

DISCUSSION

The recent and notable evolutionary success of the integron system to annihilate antibiotic treatment of Gram-negative pathogen relies on the complementarity and interplay between MIs and SCIs. There is ample evidence that the large SCIs are the source of the antibiotic resistance cassettes disseminated by the small and mobile ones (*33, 34*). The SCI arrays may carry up to more than 300, mostly silent, cassettes and represent several percent of bacterial genomes (3% in *V. cholerae*) (*5, 6*). Although the accumulation of many different functions represents an obvious long-term evolutionary advantage, since each function could be expressed in response to critical stresses, the mechanism underlying the second-order selection of vast silent cassette arrays remains unknown. Here, we demonstrate that TA cassettes in SCIs regulate the cassette recombination dynamics and stabilize massive SCIs.

TA systems were found as two-gene modules carried by plasmids, ensuring their maintenance through a process named

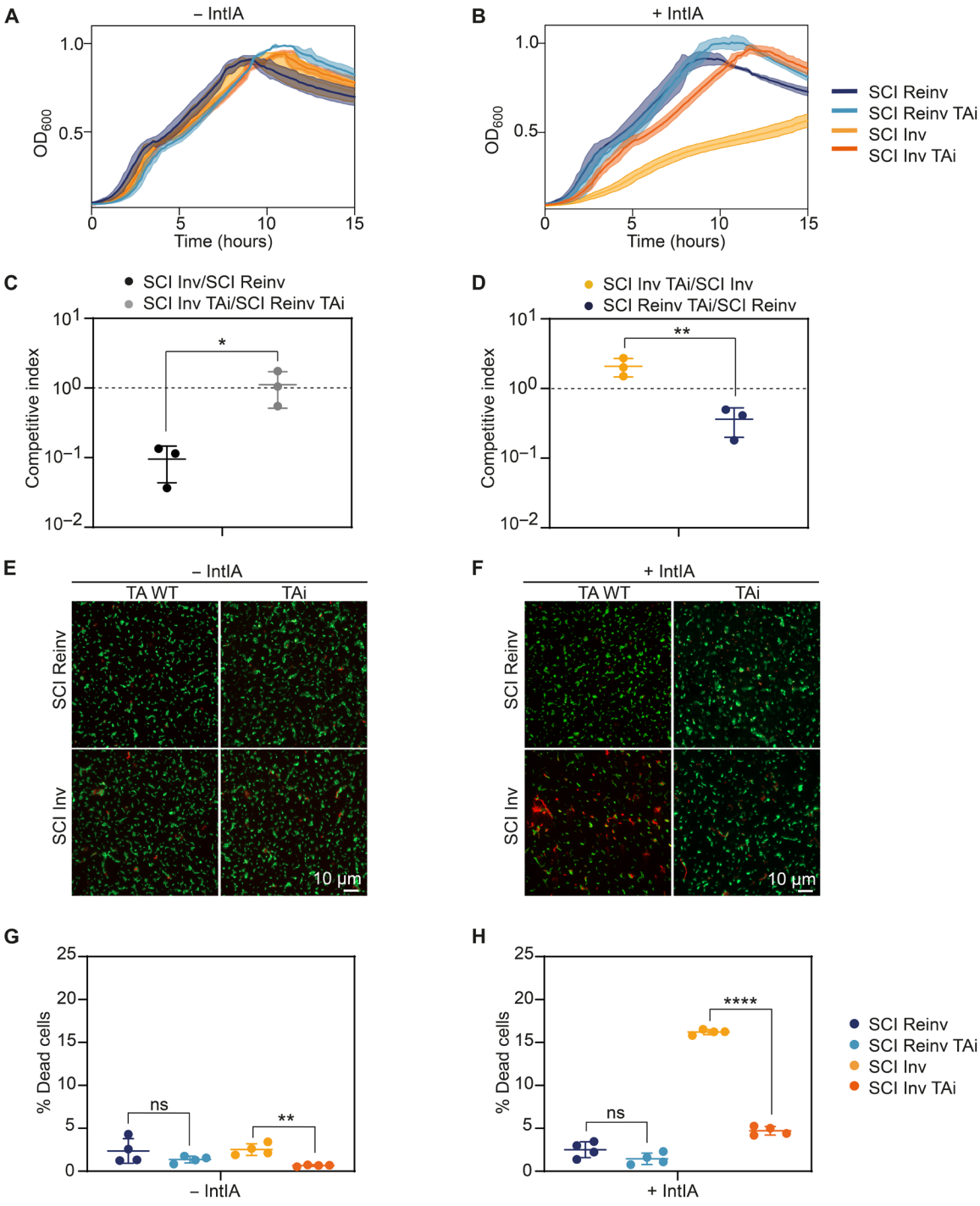

Fig. 6. Impact of the TA cassette inactivation on *V. cholerae* cell growth and viability. (**A** and **B**) Growth curve of SCI Inv or Reinv strains with TA wild type (WT) or inactivated (TAi), in the absence (A) or in the presence of the IntIA integrase (B). (**C**) Competitive index of the SCI Inv strain compared to the SCI Reinv strain (black) and of the SCI Inv TAi strain compared to the SCI Reinv TAi strain (gray) after 24 hours of cocultures. (**D**) Competitive index of the SCI Inv TAi strain compared to the SCI Inv strain (orange) and of the SCI Reinv TAi strain compared to the SCI Reinv strain (blue) after 24 hours of cocultures. For both (C) and (D), horizontal bars show the mean of three independent experiments ($n = 3$, individual plots), and error bars show the SD. Statistical comparisons (Student's t test) are as follows: ns (not significant), $^{*}P < 0.05$, and $^{**}P < 0.01$; all two-sided. (**E** and **F**) Fluorescence microscopy images resulting from the live and dead assay on the SCI Inv and Reinv strains with TA WT or inactivated (TAi), in the absence (E) or in the presence of the IntIA integrase (F). Green fluorescence indicates live cells (SYTO-9), whereas red fluorescence (PI) indicates dead cells. Scale bars, 10 µm. (**G** and **H**) Quantification of cell death as a measure of cells positively stained with PI over the total number of cells counted (~2500 cells per replicate) in the absence (G) or in the presence of the IntIA integrase (H). Horizontal bars show the mean of four independent experiments ($n = 4$, individual plots), and error bars show the SD. Statistical comparisons (Student's t test) are as follows: ns (not significant), $^{**}P < 0.01$, and $^{****}P < 0.0001$; all two-sided.

"post-segregation killing" (*35, 36*). They have also been discovered within chromosomes, primarily in genomic islands such as prophages, integrative and conjugative elements, or transposons (*37*). For example, they have been found identified as components of some members of the Tn*3* family transposons, potentially contributing to stable invasion of these transposons during transposition (*38*). In the case of *V. cholerae*, their presence along the SCI cassette array was proposed to avoid large scale rearrangements of the cassette array leading to the simultaneous loss of dozens of cassettes (*5, 15*). The high prevalence of repeated sequences within SCIs might lead to massive loss of cassettes by homologous recombination in the absence of regularly interspaced TA modules within the array. Strangely, all TA systems described so far in SCIs were found systematically associated with their own *attC* sites, turning them into cassettes (*5, 17, 18, 39*). This is puzzling since they all contain a promoter and they do not need to be integrated into *attI* sites to be expressed. Here, we propose an additional stabilizing activity of TA modules directly linked to their association with *attC* sites and which notably interconnects cassette excision frequency with cell viability (Fig. 7). While the excision of most cassettes of the array is completely harmless for the cell because they do not contain a promoter and are therefore not expressed, the excision of TA cassettes immediately imposes a prohibiting fitness cost. Since TA cassettes have supposedly the same chance to be excised than any other, this means that any excision event can potentially be lethal for the cell or at least lead to growth arrest. Although it seems counterintuitive for the host to accumulate these "time bombs" that could kill it every time the integrase becomes expressed, we argue that this could be beneficial at the population level. We call this process cassette loss killing, where cassette excision and loss lead to the death of a fraction of the population that is proportional to the overall cassette excision rate (Fig. 7). The strength of this process is that the sensor (the *attC* site) and the effector (TA system) are precisely combined into one piece (a TA cassette). This process represents an emergent property of TA systems once associated with an *attC* site, which proves to be a very effective way to ensure the stability of integron by driving the evolution of cassette excision rate toward low values. The very strong fitness cost of cassette excision imposed by the random chance of excising a TA must translate into counterselection of any condition in which the excision rate is highly increased. Here, we showed that the presence of TA cassettes heavily penalizes the inversion of the SCI of *V. cholerae* which highly increases the cassette excision rate. This makes a lot of sense with respect to integron evolution. This explains our recent in silico analyses which reveal that most of SCIs have a specific orientation relative to the chromosome replication (*14*). It fits the observations made previously that there are many TAs in large SCIs (*15*) and the current identification of their significant and perhaps ancient overrepresentation in the SCI

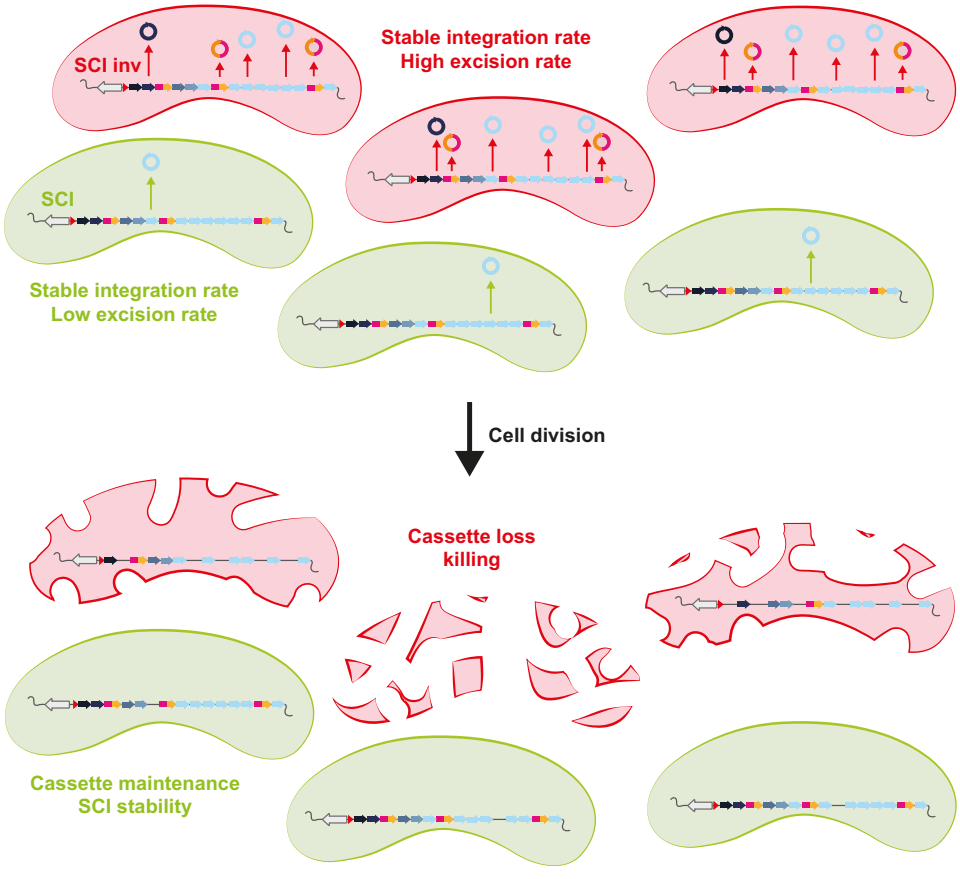

Fig. 7. Model of the cassette loss killing process in an integron array containing TA cassettes. In conditions for which the excision cassette rate is low, the TA cassettes (in pink and orange) are seldom lost. If the excision rate is high, then the TA cassettes are highly excised and kill cells. This cassette loss killing process ensures the SCI cassette maintenance.

of many *Vibrio* species compared to the rest of the genome. This long coevolution of *Vibrio* species, particularly *V. cholerae*, with their integron-encoded TA systems may have enabled SCI to have access to a vast repertoire of cassettes and thus expand their genetic capacitance. Notably, among the 1423 complete MIs carried by plasmids, only one TA cassette was found [specifically in the MI of the 248 kb plasmid from the bacterium *Comamonas testosteroni* (GCF_014076475.1)] (*40*). This very low percentage of TA cassettes in MIs perhaps means that if their recruitment can take place, then the too high cassette excision frequencies in MIs compared with SCIs (*14*) lead to the death of the cells containing and therefore excising TA cassettes. This suggests that the addictive nature of TA cassettes does not allow them to be accumulated in MIs, further confirming their role in maintaining stability of chromosomal integrons rather than being purely selfish elements.

Last, in addition to their known role in plasmid stability or viral defense (*19*), TAs associated with integron recombination sites represent special cassettes largely widespread in the large SCIs and acting as selective pressure to ensure the interplay between host genome organization and genome plasticity through that we called cassette loss killing.

MATERIALS AND METHODS

Bacterial strains, plasmids, and primers

The different bacterial strains, plasmids, and primers that were used in this study are described, respectively, in tables S3 to S5.

Media

V. cholerae and *Escherichia coli* strains were grown in LB at 37°C. *V. cholerae* strains containing a plasmid with a thermosensitive origin of replication were grown at 30°C. Thymidine (dT) and diaminopimelic acid (DAP) were supplemented, when necessary, to a final concentration of 300 μM. Glucose (Glc), L-arabinose (Ara), and fructose were added respectively at final concentrations of 10, 2, and 10 g/liter. 5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside (X-gal) was used at a final concentration of 100 μM. 2,4-Diacetylphloroglucinol (DAPG) was used at a final concentration of 50 μM. Antibiotics were respectively used at the following concentrations (for *V. cholerae* and *E. coli* respectively): carbenicillin (Carb, 100 μg/ml; 100 μg/ml), chloramphenicol (Cm, 5 μg/ml; 25 μg/ml), kanamycin (Km, 25 μg/ml; 25 μg/ml), rifampicin (Rif,1 μg/ml; 150 μg/ml), spectinomycin (Sp, 100 μg/ml or 200 μg/ml in the presence of glucose; 50 μg/ml), zeocin (Zeo, 25 μg/ml; 50 μg/ml). To avoid catabolic repression during various tests using arabinose as inducer for the expression of the integrase, cells were grown in a synthetic rich medium: Mops rich supplemented with fructose (1%) as a carbon source and arabinose (0.2%) as an inducer.

SCI inverted, reinverted, and relocated strain constructions

The inversion and relocation of the SCI were performed using a genetic tool developed in the lab (*41*) and designed to target the relocation of chromosomal DNA with bacteriophage attachment sites (HK_{022} and λ).

SCI inversion

Two DNA fragments were inserted respectively upstream and downstream of the SCI in the N16961 $hapR^+$ strain (8637). The upstream fragment contained a kanamycin resistance gene (*aph*) and the 5′ part of a carbenicillin resistance gene (*bla*) associated with the $attR_{HK}$ site. The downstream fragment contained a zeocin resistance gene (*ble*) and the partner 3′ part of the carbenicillin resistance gene associated with the $attL_{HK}$ site (fig. S2A). The strain containing the upstream and downstream fragments before inversion is referred to as the parental strain. The expression of the HK_{022} integrase and excisionase [as a directional factor, pA401; (*41*)] in the parental strain led to the $attR_{HK} \times attL_{HK}$ reaction resulting in the inversion of the SCI and to the reconstitution of the full copy of the *bla* gene, making it possible to select on carbenicillin for clones with an inverted SCI (fig. S2B). As a control for the following experiments, we reinverted the SCI to its original orientation using solely the HK_{022} integrase [p8507; (*42*)] to perform the $attP_{HK} \times attB_{HK}$ reaction in the SCI Inv strain.

Parental strain Successive natural transformations of 8637 strain with PCR fragments produced from pF384 (fragment Km^R), pF912 (fragment Zeo^R), and pF849 (fragment Sp^R) were performed. We used o4286 and o4302 pairwise primers to produce the Km^R fragment and, o4657 and o4662 to produce the Zeo^R fragment.

SCI inverted strain Transformation of I857 to I859 strains by the pA401 plasmid were performed at 30°C and in presence of Glc 1% to repress the HK integrase excisionase expression. Transformants were selected on Sp (the marker resistance carried by the plasmid) and Glccontaining plates at 30°C. The protocol of transformation is described above. SCI inversion was performed by cultivating the obtained transformant clones during 12 hours (overnight) at 30°C in the presence of Sp and Ara 0.2% (to induce the HK integrase excisionase expression). SCI inverted clones are selected by plating the resulting culture on plates containing Carb and at 37°C (to favor the loss of the pA401 thermosensitive replication plasmid).

SCI reinverted strain Transformation of the SCI inverted strain by the p8507 plasmid was performed at 30°C (to repress the thermoinducible HK integrase promoter). Transformants were selected on Sp (the marker resistance carried by the plasmid) containing plates at 30°C. The protocol of transformation is described above. SCI reinversion was performed by cultivating the obtained transformant clones up to optical density at 600 nm (OD_{600}) ~ 0.3 at 30°C and by shifting the temperature to 37°C during 90 min. SCI reinverted clones are selected by plating the resulting culture on plate without Carb at 42°C (to favor the loss of the p8507 thermosensitive replication plasmid). The reinversion of the SCI was checked by confirming the carbenicillin sensitivity of several obtained clones (by plating them on Carb containing plates).

SCI relocation

Two DNA fragments were inserted respectively upstream and downstream of the SCI in the N16961 $hapR^+$ strain (8637). The upstream fragment contained the $attR_{HK}$ site associated with the 5′ part of a carbenicillin resistance gene (*bla*) and a kanamycin resistance gene (*aph*). The downstream fragment contained a zeocin resistance gene (*ble*) and the 5′ part of a *lacZ* gene associated with the $attL_{\lambda}$ site (fig. S2B). In this strain, another DNA fragment was inserted in the *chr1* near the *ori1* between VC0018 and VC0019 in both orientations. This fragment contained the 3′ part of the carbenicillin resistance gene associated to the $attL_{HK}$ site, the spectinomycin resistance gene (*aadA7*), and the $attR_{\lambda}$ site associated with the 3′ part of the *lacZ* gene. The strain containing all these fragments before relocation is referred to as the parental strain. The expression of the HK_{022} and λ integrases and excisionases (pF930) in the parental strain led to the $attR_{HK} \times attL_{HK}$ and to the $attR_{\lambda} \times attL_{\lambda}$

reactions resulting in the relocation of the SCI and to the reconstitution of the full copy of the *bla* and *lacZ* genes, making it possible to select on Carb for clones with relocated SCI (fig. S2B). Depending of the orientation of the *chr1* fragment, the SCI will be oriented in lagging or leading orientation.

Parental strain Successive natural transformations of 8637 strain with PCR fragments produced from pF851 (fragment Km^R) and pF850 (fragment Zeo^R) were performed. We used o4286 and o4302 pairwise primers to produce the Km^R fragment, o4659 and o4662 pairwise primers to produce the Zeo^R fragment, and o4665 and o4667 to produce the Sp^R fragment.

SCI relocated strains Transformation of strains by the pF930 plasmid were performed at 30°C and in the presence of Glc 1% to repress the HK and λ integrase excisionase expression. Transformants were selected on Cm (the marker resistance carried by the plasmid) and Glc containing plates at 30°C. The protocol of transformation is described above. SCI relocation was performed by cultivating the obtained transformant clones during 12 hours (overnight) at 30°C in the presence of Cm and Ara 0.2% (to induce the HK and λ integrase excisionase expression). SCI relocated clones are selected by plating the resulting culture on plates containing Carb and at 37°C (to favor the loss of the pF930 thermosensitive replication plasmid).

Automated growth curve measurements

Overnight cultures of the indicated strain were diluted 1:1000 and then grown for 20 hours in the indicated medium. Bacterial preparations were distributed by triplicate in 96-well microplates. Growth-curve experiments were performed using a Tecan Infinite microplate reader, with absorbance measurements (600 nm) taken at 10-min intervals. Maximum growth rates during exponential phase were directly obtained using the "GrowthRates" R package. The package can be downloaded from Comprehensive R Archive Network (CRAN; https://cran.r-project.org/package=growthrates). Curve corresponds to the mean of at least three independent experiment ($n = 3$), each with an average of at least two technical replicates and, the shade corresponds to the SEs at each time point.

TAs inactivation by allelic exchange

We performed allelic exchange to construct N16961 lacking the TA systems that could not be targeted using the base-editing tool. To this purpose, we constructed and used different variants of the pMP7 vector, respectively, pP897, pP898, pP900, and pP902. We followed the same protocols as previously described (*27*). Briefly, the suicide vector pMP7 contains a R6K origin of replication, and its replication is then dependent on the presence of the Π protein in the host cell. The Π3813 cell, a *pir* + CcdB–resistant *E. coli* strain, was used for cloning the different pMP7 plasmids. Once constructed, these vectors were transformed into the β3914 donor strain to deliver by conjugation the pMP7 vector into the desired recipient strain. Homology regions corresponding to the genomic DNA from the recipient strain have been cloned in the different pMP7 vector to allow the integration of the plasmid by homologous recombination. The only way for pMP7 vector to replicate into recipient strains is then to integrate into the host genome after a first crossover. After conjugation, integration of the entire pMP7 vector were then selected by plating cells on Cm plates lacking DAP. Next, cells were grown in the presence of Ara 0.2% to express the CcdB toxin.

The expression of this toxin allows to kill cells in which the second crossover that leads to the excision of pMP7 backbone did not take place. This method allows us to obtain mutants of *V. cholerae* that are devoid of any antibiotic resistance marker. Since we did not perform deletions, the verification of the gene replacement by PCR only was prohibited. For this reason, the STOP containing version of the toxin was also associated to a Bam HI restriction site to allow an easier screening of the correct gene replacement by digesting the appropriate PCR and looking for a restriction profile. After that, the PCR products containing the Bam HI sites were checked by sequencing. The primers used for PCR screening and sequencing are described table S5.

Golden Gate Assembly

By default, the vector expressing dead Cas9 (dCas9) carries a "random" guide that does not target any locus in *E. coli* nor in *V. cholerae* but that do contain two Bsa I restriction sites (*43*). To change the guide, we perform a "Golden Gate Assembly" as such: two oligonucleotides have to be designed in the form 5′-TAGTNNNNNNNNNNNNNNNNNNN-3′ and 5′-TTTGNNNNNNNNNNNNNNNNNNN-3′ where the 20 "N" is the targeted sequence. For the annealing, 6 µl of each oligo is mixed with 2 µl of T4 DNA ligase buffer and 0.4 µl of T4 Polynucleotide Kinase (PNK) for a total volume of 20 µl and then incubated for 30 min at 37°C. We then add 1 µl of NaCl (1 M), incubate 5 min at 95°C, and lastly let slowly cool down to room temperature for at least 4 hours.

To insert the desired guide into the dCas9 expressing vector, we prepare the following mix: 2 µl of the dCas9 plasmid, 2 µl of annealed oligos, 1 µl of Cutsmart buffer, 1 µl of Bsa I enzyme, 1 µl of adenosine triphosphate (1 mM), 1 µl of ligase, and 2 µl of H_2O. We then incubate the mix in a thermocycler with the following steps: 3 min at 37°C (for digestion), 4 min at 16°C (for ligation), and alternate between those steps 25 times. We finish with 1 cycle of 5 min at 50°C and 5 min at 80°C for enzymes inactivation. At this step, the mix is ready to be transformed in a cloning *E. coli* strain, and the successful single guide RNA insertions are screened by PCR using the o2406 and o3812 primers.

Base editing

The tool used for base editing is derived from the catalytically dCas9 previously described (*25*). The construction consists in a gene fusion between dCas9 and CDA1, a cytidine deaminase that catalyzes the deamination of cytidine, resulting in a uridine base which will later be replicated as a thymine (C ➔ U ➔ T). A linker of 100 amino acids separated the dCas9 and the CDA1. An uracil glycosylase inhibitor was also fused to CDA1 to increase the base editing rate as described by Banno and colleagues (*25*) as well as a Leucine Alanine Valine (LVA) degradation tag to decrease the toxicity the construct. Because of a low-efficiency transformation rate of this construct in *V. cholerae*, the construct was delivered by conjugation. Hence, the appropriate construct was first transformed in a β3914 donor strain. Then, both the donor and receptor strains were cultivated to exponential phase, mixed 1:1 in a 2-ml tube and centrifugated (6000 rpm for 6 min).

The pellet was spread on a membrane on a plate containing DAP to sustain growth of the donor strain and DAPG for induction of base editing, and the plate was incubated for 3 hours at 37°C. After incubation, the membrane was resuspended in 5 ml of LB and vortexed for 30 s to resuspend the conjugated cells which were then plated on appropriate media: Cm + Glc 1% + X-gal to obtain both

the colony-forming units per milliliter (CFU/ml) after conjugation and the base editing efficiency, or Rif + Cm + Glc 1% to obtain the rate of apparition rifampicinresistant clones as a proxy for global mutation rate. Blue colonies had a functional *lacZ* gene, while the white colonies were synonym for a mutated *lacZ* gene. All targeted sites were checked by PCR and sequencing (see table S5). Then, the whole genomes of the TAi strains were sequenced (Illumina).

SCI cassette shuffling assay

A single clone of each strain was isolated on plate and then inoculated for an overnight culture in LB + Sp + Glc 1%. The next day, the cells were inoculated at 1/50th for 2 hours in Mops rich + Sp + fructose 1% + Ara 0.2% as a preinduction step until they were at exponential phase. They were then inoculated at 1/1000th and grew for 20 hours until stationary phase. Cultures were streaked on 1% agarose plates containing Mueller Hinton (MH) + Sp + Glc 1%, and single clones were isolated. A total of 48 PCR per condition was performed using an oligo (o5778; see table S5) hybridizing in the promoter of *intIA* (stable part) and one (o1401) hybridizing in the first cassette of the array (variable part). A representative sample of the PCR profile for each condition was migrated in 1% agarose for visualization.

Suicide conjugation assay

This assay has been previously described by Vit and colleagues (*3*). We used the suicide conjugative vector pSW23T (*44*) that allows the delivery of the bottom strand of the $attC_{aadA7}$ recombination site (pD060). For each condition assay, at least 16 recombinant clones were isolated on appropriate plates and analyzed by PCR. To determine precisely whether the pSW23T vector has been inserted into the *attIA* site of the SCI, we used 5778 and SWend primers. These primers hybridize respectively in a sequence upstream of *attIA* in *V. cholerae* chromosome 2 or downstream of $attC_{aadA7}$ in the pSW23T vector. For each condition assay, at least three PCR reactions were purified and sequenced to confirm the integration point.

SCI cassette excision assay

A single clone of each strain was isolated on plate and then inoculated for an overnight culture in LB + Sp + Glc 1%. The next day, the cells were inoculated at 1/50th for 2 hours in Mops rich + Sp + fructose 1% + Ara 0.2% as a preinduction step until they were at exponential phase. They were then inoculated at 1/1000th and grew for 20 hours until stat phase. Last, 1 ml of this culture was used to inoculate a 100-ml culture of LB + Sp + Glc 1% to filter the dead cells of the latter culture. The resulting culture constitutes a mixed population where each individual might have experienced a diverse set of cassette excision events. Bulk DNA was extracted from these mixed population (Qiagen genomic DNA extraction kit) and sequenced using the Pacific Bioscience (PacBio) long-read sequencing technology. Mapping was performed using minimap2 (*45*), and the cassette deletions were detected using a homemade R program. In order not to confuse the excision rate with the random chance of an excision event to be propagated, the cultures of the different strains were replicated eight times independently.

The cassette excision frequency corresponds to the total number of cassette excision events divided by the sequencing depth across the SCI. For the SCI inverted strain expressing integrase, we detected 478 excision events encompassing the entire SCI. The calculated frequency corresponds to: $N_{exc}/(SCI_{cov} + N_{exc})$; where N_{exc} is the number of excision events and SCI_{cov} is the mean coverage for each cassette multiplied by the number of cassettes in the SCI (i.e., 179). We therefore obtained a cassette excision frequency of $478/(13269.16 + 478) = 3.48 \times 10^{-2}$.

For the SCI inverted strain without integrase and the SCI reinverted strains with and without integrase, for which we did not detect any excision event, we calculated the limit of detection. It corresponds to: $1/SCI_{cov}$. Note that the SCI read coverage obtained for each condition is, respectively, 13518.08, 14737.07, and 14550.91, corresponding to detection limits of, respectively, 7.40×10^{-5}, 6.79×10^{-5}, and 6.87×10^{-5}.

attC sites pfold predictions

The calculation of the probability of an *attC* site to fold into a recombinogenic structure (pfold) was performed as described by Vit *et al.* (*46*) [i.e., using the RNAfold program from ViennaRNA Package (*47*) (http://rna.tbi.univie.ac.at, category "RNAfold Server")].

Protein purification

IntIA proteins were expressed in BL21 bacteria (*48*). Cells were lysed in 25 mM tris (pH 8), 1 M NaCl, 1 mM phenylmethylsulfonyl fluoride, lysozyme, and protease inhibitors containing buffer. Proteins were purified by nickel affinity chromatography with a 200 mM imidazole elution. His-tag was removed by incubating proteins at 4°C overnight with Tobacco Etch Virus (TEV) protease (1 mg per 15 mg of total proteins containing a 5 mM final concentration of β mercapto-ethanol). After dilution to 150 mM NaCl, proteins were purified on an SP column equilibrated with 25 mM tris (pH 8) and 150 mM NaCl. A gradient of 150 mM to 1 M NaCl was performed. Proteins were then brought to a final concentration of 1 mM dithiothreitol (DTT) and 10% glycerol.

Electromobility shift assay

Each reaction contained 500 ng of poly[d(I-C)], 12 mM Hepes-NaOH (pH 7.7), 12% glycerol, 4 mM tris-HCl (pH 8.0), 60 mM KCl, 1 mM EDTA, bovine serum albumin (0.06 mg/ml), 1 mM DTT, 10% Tween 20, 0.01 nM specified 32P-labeled DNA oligonucleotide (table S5), and the specified quantities of purified His-tagged IntIA and IntIAY302F, in a final volume of 20 µl. The samples were incubated at 30°C for 10 min without the *attC* probe, followed by 20 min with the *attC* probe, and then loaded to a 5% native polyacrylamide gels with 0.5× tris-borate EDTA as running buffer. The gels were visualized using a Molecular Dynamics intensification screen and a Typhoon FLA 9500 laser scanner.

Competition assay

The SCI Inv strains are resistant to Carb, which is not the case for the SCI Reinv strains (see above). We used this difference to perform a competition assay. One clone of each strain was isolated on plate and used to inoculate an overnight culture. Before starting the competition assay, the cells were precultured at 1/50th in LB + Sp + Glc 1% until exponential phase to be not biased by the dead cells from the overnight culture. Cells were then mixed at 1:1 ratio at 1/200th in Mops rich + fructose 1% + Sp + Ara 0.2% for induction of the integrase. The coculture was plated on petri dishes containing MH + Sp + Glc 1% to get the total CFU/ml and in parallel on MH + Carb + Sp + Glc 1% plates to get the number of SCI Inv CFU/ml in that same culture. We plated at t_0, t_0 + 2 hours,

t_0 + 4 hours, t_0 + 6 hours, t_0 + 8 hours, and t_0 + 24 hours. The competitive index (*I*) was calculated as such: $I = CFU_{SCI\ Inv}/CFU_{SCI\ Reinv} - CFU_{SCI\ Inv}$ so that an index of 1 represents a situation where the ratio of SCI Inv and SCI Reinv is 1:1 (neutrality). A high index indicates that SCI Inv has a competitive advantage compare to its Reinv counterpart and vice versa.

Live and dead assay

Overnight cultures were performed in LB + Sp + Glc 1%. Then, day cultures were inoculated at 1:1000 in Mops rich + Sp + fructose 1% + Ara 0.2% medium to OD_{600} ~ 0.8. Then, 50 µl of the latter culture was dyed using the LIVE/DEAD *Bac*Light Bacterial Viability Kit, which is a mix of propidium iodide (PI) and SYTO-9 dyes. PI specifically stains the dead cells in red. SYTO-9 specifically stains viable cells in green. Cell viability was then assessed by microscopy. Stained cells were placed of 1.4% agarose pad containing Mops rich and observed using a Zeiss ApoTome inverted wide-field microscope. Snapshot images were taken with a 40- and 50-ms exposure time (fluorescein isothiocyanate and DsRED channel, respectively), using a Plan Apo 63× objective (numerical aperture = 1.4, +optovar 1.6×) using a Hamamatsu scientific complementary metal-oxide semiconductor (sCMOS) ORCA-Flash 4.0 v3 (Institut Pasteur imaging facility imagopole). The result is presented as the proportion of dead cells (red) in the population.

Microscopy setup for live imaging

Bacterial cells were grown to mid-exponential phase in liquid Mops rich + Sp + fructose 1% + Ara 0.2% medium and then transferred to 1.4% agarose-padded slides containing Mops rich + Sp + fructose 1% + Ara 0.2%. A coverslip was placed on top of the agarose pad and sealed with a vaseline:lanolin:paraffin mix (ratio 1:1:1) to prevent pad evaporation. Slides were incubated at 37°C under an inverted wide-field microscope (Zeiss ApoTome) for time-lapse video recording. Video frames were taken at 1-min interval time for a total duration of 170 min, using a Plan Apo 63× objective (numerical aperture = 1.4) using a Hamamatsu sCMOS ORCA-Flash 4.0 v3 and autofocus module (Definite Focus, Zeiss) (Pasteur Institute Imaging Facility Imagopole). A total of eight movies were recorded and analyzed across all strains.

Digital reverse transcription PCR

For RNA extraction, overnight cultures of three or four biological replicates (depending on strains) were diluted 1:1000 in Mops rich + Sp + fructose 1% + Ara 0.2% and grown with agitation at 37°C until an OD_{600} ~ 0.3 (exponential phase). Two milliliters of these cultures was centrifuged, and the supernatant was removed. Pellet were resuspended in 2:1.5 TRIzol (Invitrogen). Next, 300 µl of chloroform was added to the samples following mix by vortexing. Samples were then centrifuged at 4°C for 10 min. A volume of 1:1 of 70% ethanol was mixed with the upper phase before transfer on column (RNeasy Mini kit, Qiagen), and RNA was extracted according to the manufacturer's directions. Samples were then subjected to deoxyribonuclease treatment using the TURBO DNA-free Kit (Ambion) according to the manufacturer's instructions. RNA concentration of the samples was measured with a NanoDrop spectrophotometer and diluted to a final concentration of 0.2 ng/µl. Quantitative RT-PCR reactions were prepared with 1 µl of diluted RNA samples using the qScript XLT1-Step RT-qPCR ToughMix (Quanta Biosciences, Gaithersburg, MD, USA) and loaded on Ruby chips. RT-dPCR was conducted on a Naica Geode. Image acquisition was performed using the Naica Prism3 reader. Images were then analyzed using Crystal Reader software and the Crystal Miner software. Values were normalized against expression of the housekeeping gene *gyrA* as previously described (*49*). Primers and probes used for amplification are described in table S5.

Marker frequency analysis

MFA was performed as described in (*30*). Briefly, cells were grown in Mops rich + Sp + Ara 0.2% + fructose 1% at 37°C under agitation. Genomic DNA was prepared using the DNeasy Tissue Kit (Qiagen) from 30 ml of exponentially growing cells (OD_{450} ~ 0.15). The remaining culture was washed twice and was kept under agitation in Mops rich + Sp + Glc 1% at 37°C to repress expression of IntIA and to prepare gDNA from cells in stationary phase (24 hours of culture). Libraries were prepared for Illumina sequencing (150 bp, paired end). The resulting reads were mapped with bowtie2 (*50*). The genomic coverage of the sequencing data stemming from the exponentially growing cultures of each strain was normalized with the corresponding data in stationary phase. The normalized coverage is represented in function of the relative position to the terminus of replication of chromosome 1 (*ter1*) as colored (blue and green) dots for 1-kb bins and red dots for 10-kb bins. Dots are not shown when at least one position has a coverage of zero within the bin (typically due to repetitive sequences that prevents correct reads mapping). The gradient of coverage from *ori* to *ter* serves as a proxy for the replication dynamics, and the slope translates the local speed of the replication fork (*29*).

Comparative genomics analysis of TA systems within SCI

Genome dataset and SCI annotation

To study the enrichment of TA systems within SCI across bacterial genomes, we used the dataset and integron predictions of IntegronFinder 2.0 (*12*). Briefly, this dataset consists of 21,105 complete genomes retrieved from NCBI RefSeq on 30 March 2021. We defined SCIs as complete integrons with at least 11 predicted *attC* sites. Several isolates of *V. cholerae* appeared to carry large (>10 *attC* sites) CALINs (cluster of *attC* sites lacking an integrase) (*12, 51*). A closer investigation showed that these CALINs were preceded by a pseudogenized integrase, which relates them with SCIs. These elements could have been recently inactivated, or the putative pseudogenes could result from sequencing artifacts. Hence, we included in our analysis all the isolates carrying these CALINs. To simplify the results, we grouped together SCIs and CALINs, without distinction, under the global term SCIs.

Toxin and antitoxin prediction

In each genome harboring at least one SCI, we predicted toxins and antitoxins with TASmania, a predictor with very low false-negative rate (*21*). Our goal was to retrieve even unusual TA systems that would be rarely found outside SCIs and hence missed by more conservative predictors. More precisely, we downloaded toxin and antitoxin HMM models from TASmania's website (accessed on 25 November 2021) and screened for them with the hmmscan command of HMMER 3.3 (*52*) (November 2019; http://hmmer.org/). As the command was run with default parameters, no threshold was used to filter hits in the first screening. To reduce the number of false positives, we evaluated TASmania's performance on the genome of *V. cholerae* N16961's secondary chromosome, known to harbor 19 TA systems (*17, 18*), all located within the SCI. We observed that

applying an *e*-value threshold of 10^{-3} allowed to keep all the hits corresponding to characterized TA systems [as described by Iqbal and colleagues (*17*)] while eliminating most of the other hits. Hence, we selected this value to filter TASmania hits in all the genomes comprised in our analysis, and we decided to consider as a toxin (respectively antitoxin) hit any protein matching a TASmania toxin (respectively antitoxin) profile with an *e*-value lower than 10^{-3}.

Contingency analysis of TA enrichment within SCI

For each genome, we performed a contingency table analysis to test for an enrichment of TA systems within SCIs. The contingency table was built as follows: In each genome, each protein was classified as, on one side, belonging or not to an SCI and, on the other side, containing a TASmania hit or none. A Fisher one-sided test for statistical enrichment significance was performed on each contingency table. To correct for multiple testing, the resulting *P* values were adjusted either with the Benjamini-Hochberg (main analysis) or the Bonferroni (complementary analysis) corrections. An isolate was claimed to harbor a SCI significantly enriched in TA when the adjusted *P* value was lower than 0.05. The Benjamini-Hochberg and Bonferroni corrections were recomputed in the analysis focusing on *Vibrio* isolates.

Vibrio genus cladogram

To identify a potential evolutionary trend explaining the TA enrichment within SCI across the *Vibrio* genus, we reconstructed a cladogram grouping *Vibrio* species together. More precisely, we took as basis the phylogeny published by Sawabe and colleagues (*22*) and removed the species that did not harbor any SCI. The species comprised in our dataset that were absent of the phylogeny were then added as a sister branch of the closest relative species.

Supplementary Materials

This PDF file includes:

Figs. S1 to S8
Tables S2 to S5
Legend for table S1
Legends for movies S1 to S5

Other Supplementary Material for this manuscript includes the following:

Movies S1 to S5
Table S1

REFERENCES AND NOTES

1. H. W. Stokes, R. M. Hall, A novel family of potentially mobile DNA elements encoding site-specific gene-integration functions: Integrons. *Mol. Microbiol.* **3**, 1669–1683 (1989).
2. C. M. Collis, R. M. Hall, Expression of antibiotic resistance genes in the integrated cassettes of integrons. *Antimicrob. Agents Chemother.* **39**, 155–162 (1995).
3. C. Vit, E. Richard, F. Fournes, C. Whiteway, X. Eyer, D. Lapaillerie, V. Parissi, D. Mazel, C. Loot, Cassette recruitment in the chromosomal Integron of Vibrio cholerae. *Nucleic Acids Res.* **49**, 5654–5670 (2021).
4. E. Guerin, G. Cambray, N. Sanchez-Alberola, S. Campoy, I. Erill, S. da Re, B. Gonzalez-Zorn, J. Barbé, M. C. Ploy, D. Mazel, The SOS response controls integron recombination. *Science* **324**, 1034 (2009).
5. J. A. Escudero, C. Loot, A. Nivina, D. Mazel, The Integron: Adaptation on Demand. *Microbiol. spectr.* **3**, MDNA3-0019-2014 (2015).
6. E. Richard, B. Darracq, C. Loot, D. Mazel, Unbridled integrons: A matter of host factors. *Cell* **11**, (2022).
7. J. A. Escudero, C. Loot, V. Parissi, A. Nivina, C. Bouchier, D. Mazel, Unmasking the ancestral activity of integron integrases reveals a smooth evolutionary transition during functional innovation. *Nat. Commun.* **7**, 10937 (2016).
8. C. Johansson, M. Kamali-Moghaddam, L. Sundstrom, Integron integrase binds to bulged hairpin DNA. *Nucleic Acids Res.* **32**, 4033–4043 (2004).
9. M. Bouvier, G. Demarre, D. Mazel, Integron cassette insertion: A recombination process involving a folded single strand substrate. *EMBO J.* **24**, 4356–4367 (2005).
10. A. Nivina, J. A. Escudero, C. Vit, D. Mazel, C. Loot, Efficiency of integron cassette insertion in correct orientation is ensured by the interplay of the three unpaired features of attC recombination sites. *Nucleic Acids Res.* **44**, 7792–7803 (2016).
11. G. Cambray, A. M. Guerout, D. Mazel, Integrons. *Annu. Rev. Genet.* **44**, 141–166 (2010).
12. B. Neron, E. Littner, M. Haudiquet, A. Perrin, J. Cury, E. P. C. Rocha, IntegronFinder 2.0: Identification and analysis of integrons across bacteria, with a focus on antibiotic resistance in klebsiella. *Microorganisms* **10**, 700 (2022).
13. M. Buongermino Pereira, T. Österlund, K. M. Eriksson, T. Backhaus, M. Axelson-Fisk, E. Kristiansson, A comprehensive survey of integron-associated genes present in metagenomes. *BMC Genomics* **21**, 495 (2020).
14. C. Loot, A. Nivina, J. Cury, J. A. Escudero, M. Ducos-Galand, D. Bikard, E. P. C. Rocha, D. Mazel, Differences in integron cassette excision dynamics shape a trade-off between evolvability and genetic capacitance. *MBio* **8**, e02296–e02216 (2017).
15. S. Szekeres, M. Dauti, C. Wilde, D. Mazel, D. A. Rowe-Magnus, Chromosomal toxin-antitoxin loci can diminish large-scale genome reductions in the absence of selection. *Mol. Microbiol.* **63**, 1588–1605 (2007).
16. D. A. Rowe-Magnus, A. M. Guerout, L. Biskri, P. Bouige, D. Mazel, Comparative analysis of superintegrons: Engineering extensive genetic diversity in the vibrionaceae. *Genome Res.* **13**, 428–442 (2003).
17. N. Iqbal, A. M. Guerout, E. Krin, F. Le Roux, D. Mazel, Comprehensive functional analysis of the 18 vibrio cholerae N16961 toxin-antitoxin systems substantiates their role in stabilizing the superintegron. *J. Bacteriol.* **197**, 2150–2159 (2015).
18. E. Krin, Z. Baharoglu, O. Sismeiro, H. Varet, J. Y. Coppée, D. Mazel, Systematic transcriptome analysis allows the identification of new type I and type II Toxin/Antitoxin systems located in the superintegron of Vibrio cholerae. *Res. Microbiol.* **174**, 103997 (2023).
19. D. Jurenas, N. Fraikin, F. Goormaghtigh, L. Van Melderen, Biology and evolution of bacterial toxin-antitoxin systems. *Nat. Rev. Microbiol.* **20**, 335–350 (2022).
20. D. Mazel, B. Dychinco, V. A. Webb, J. Davies, A distinctive class of integron in the Vibrio cholerae genome. *Science* **280**, 605–608 (1998).
21. H. Akarsu, P. Bordes, M. Mansour, D. J. Bigot, P. Genevaux, L. Falquet, TASmania: A bacterial Toxin-Antitoxin Systems database. *PLoS Comput. Biol.* **15**, e1006946 (2019).
22. T. Sawabe, Y. Ogura, Y. Matsumura, G. Feng, A. R. Amin, S. Mino, S. Nakagawa, T. Sawabe, R. Kumar, Y. Fukui, M. Satomi, R. Matsushima, F. L. Thompson, B. Gomez-Gil, R. Christen, F. Maruyama, K. Kurokawa, T. Hayashi, Updating the Vibrio clades defined by multilocus sequence phylogeny: proposal of eight new clades, and the description of Vibrio tritonius sp. nov. *Front. Microbiol.* **4**, 414 (2013).
23. C. Loot, D. Bikard, A. Rachlin, D. Mazel, Cellular pathways controlling integron cassette site folding. *EMBO J* **29**, 2623–2634 (2010).
24. F. Wong, A. Amir, Mechanics and dynamics of bacterial cell lysis. *Biophys. J.* **116**, 2378–2389 (2019).
25. S. Banno, K. Nishida, T. Arazoe, H. Mitsunobu, A. Kondo, Deaminase-mediated multiplex genome editing in Escherichia coli. *Nat. Microbiol.* **3**, 423–429 (2018).
26. A. C. Reis, S. M. Halper, G. E. Vezeau, D. P. Cetnar, A. Hossain, P. R. Clauer, H. M. Salis, Simultaneous repression of multiple bacterial genes using nonrepetitive extra-long sgRNA arrays. *Nat. Biotechnol.* **37**, 1294–1301 (2019).
27. F. Le Roux, J. Binesse, D. Saulnier, D. Mazel, Construction of a Vibrio splendidus mutant lacking the metalloprotease gene vsm by use of a novel counterselectable suicide vector. *Appl. Environ. Microbiol.* **73**, 777–784 (2007).
28. K. S. Lang, A. N. Hall, C. N. Merrikh, M. Ragheb, H. Tabakh, A. J. Pollock, J. J. Woodward, J. E. Dreifus, H. Merrikh, Replication-transcription conflicts generate R-loops that orchestrate bacterial stress survival and pathogenesis. *Cell* **170**, 787–799.e18 (2017).
29. O. Skovgaard, M. Bak, A. Lobner-Olesen, N. Tommerup, Genome-wide detection of chromosomal rearrangements, indels, and mutations in circular chromosomes by short read sequencing. *Genome Res.* **21**, 1388–1393 (2011).
30. M. E. Val, M. Marbouty, F. de Lemos Martins, S. P. Kennedy, H. Kemble, M. J. Bland, C. Possoz, R. Koszul, O. Skovgaard, D. Mazel, A checkpoint control orchestrates the replication of the two chromosomes of Vibrio cholerae. *Sci. Adv.* **2**, e1501914 (2016).
31. E. Galli, J. L. Ferat, J. M. Desfontaines, M. E. Val, O. Skovgaard, F. X. Barre, C. Possoz, Replication termination without a replication fork trap. *Sci. Rep.* **9**, 8315 (2019).
32. E. Krin, S. A. Pierlé, O. Sismeiro, B. Jagla, M. A. Dillies, H. Varet, O. Irazoki, S. Campoy, Z. Rouy, S. Cruveiller, C. Médigue, J. Y. Coppée, D. Mazel, Expansion of the SOS regulon of Vibrio cholerae through extensive transcriptome analysis and experimental validation. *BMC Genomics* **19**, 373 (2018).
33. D. A. Rowe-Magnus, A. M. Guerout, D. Mazel, Bacterial resistance evolution by recruitment of super-integron gene cassettes. *Mol. Microbiol.* **43**, 1657–1669 (2002).
34. F. Le Roux, M. Zouine, N. Chakroun, J. Binesse, D. Saulnier, C. Bouchier, N. Zidane, L. Ma, C. Rusniok, A. Lajus, C. Buchrieser, C. Médigue, M. F. Polz, D. Mazel, Genome sequence of Vibrio splendidus: An abundant planctonic marine species with a large genotypic diversity. *Environ. Microbiol.* **11**, 1959–1970 (2009).
35. K. Gerdes, P. B. Rasmussen, S. Molin, Unique type of plasmid maintenance function: Postsegregational killing of plasmid-free cells. *Proc. Natl. Acad. Sci. U. S. A.* **83**, 3116–3120 (1986).

36. T. Ogura, S. Hiraga, Mini-F plasmid genes that couple host cell division to plasmid proliferation. *Proc. Natl. Acad. Sci. U. S. A.* **80**, 4784–4788 (1983).
37. F. Hayes, L. Van Melderen, Toxins-antitoxins: Diversity, evolution and function. *Crit. Rev. Biochem. Mol. Biol.* **46**, 386–408 (2011).
38. G. Lima-Mendez, D. Oliveira Alvarenga, K. Ross, B. Hallet, L. van Melderen, A. M. Varani, M. Chandler, Toxin-antitoxin gene pairs found in Tn3 Family transposons appear to be an integral part of the transposition module. *MBio* **11**, e00452–e00420 (2020).
39. A. M. Guerout, N. Iqbal, N. Mine, M. Ducos-Galand, L. Van Melderen, D. Mazel, Characterization of the phd-doc and ccd toxin-antitoxin cassettes from Vibrio superintegrons. *J. Bacteriol.* **195**, 2270–2283 (2013).
40. J. Huyan, Z. Tian, Y. Zhang, H. Zhang, Y. Shi, M. R. Gillings, M. Yang, Dynamics of class 1 integrons in aerobic biofilm reactors spiked with antibiotics. *Environ. Int.* **140**, 105816 (2020).
41. M. E. Val, O. Skovgaard, M. Ducos-Galand, M. J. Bland, D. Mazel, Genome engineering in vibrio cholerae: A feasible approach to address biological issues. *PLOS Genet.* **8**, e1002472 (2012).
42. M. Rossignol, L. Moulin, F. Boccard, Phage HK022-based integrative vectors for the insertion of genes in the chromosome of multiply marked Escherichia coli strains. *FEMS Microbiol. Lett.* **213**, 45–49 (2002).
43. F. Rousset, J. Cabezas-Caballero, F. Piastra-Facon, J. Fernández-Rodríguez, O. Clermont, E. Denamur, E. P. C. Rocha, D. Bikard, The impact of genetic diversity on gene essentiality within the Escherichia coli species. *Nat. Microbiol.* **6**, 301–312 (2021).
44. G. Demarre, A. M. Guérout, C. Matsumoto-Mashimo, D. A. Rowe-Magnus, P. Marlière, D. Mazel, A new family of mobilizable suicide plasmids based on broad host range R388 plasmid (IncW) and RP4 plasmid (IncPα) conjugative machineries and their cognate *Escherichia coli* host strains. *Res. Microbiol.* **156**, 245–255 (2005).
45. H. Li, Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).
46. C. Vit, C. Loot, J. A. Escudero, A. Nivina, D. Mazel, Integron identification in bacterial genomes and cassette recombination assays. *Methods Mol. Biol.* **2075**, 189–208 (2020).
47. R. Lorenz, S. H. Bernhart, C. Höner zu Siederdissen, H. Tafer, C. Flamm, P. F. Stadler, I. L. Hofacker, ViennaRNA package 2.0. *Algorithms Mol. Biol.* **6**, 26 (2011).
48. D. MacDonald, G. Demarre, M. Bouvier, D. Mazel, D. N. Gopaul, Structural basis for broad DNA-specificity in integron recombination. *Nature* **440**, 1157–1162 (2006).
49. M. Lo Scrudato, M. Blokesch, The regulatory network of natural competence and transformation of Vibrio cholerae. *PLOS Genet.* **8**, e1002778 (2012).
50. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
51. J. Cury, T. Jove, M. Touchon, B. Neron, E. P. Rocha, Identification and analysis of integrons and cassette arrays in bacterial genomes. *Nucleic Acids Res.* **44**, 4539–4550 (2016).
52. S. R. Eddy, Profile hidden Markov models. *Bioinformatics* **14**, 755–763 (1998).

Acknowledgments: We thank S. H. Sternberg for critical reading of the manuscript. We thank J. Teramoto and A. Kondo for sharing construct for base editing. We thank M. Monot, L. Ma, J. P. Da Fonseca, and T. Cokelaer from the Biomics platform, C2RT, Institut Pasteur, Paris, France, supported by France Génomique (ANR-10-INBS-09) and IBISA. We thank L. Fruchard for help with RT-dPCR and C. Garcia Rodriguez for help with microscopy. **Funding:** This work was supported by Institut Pasteur, Centre National de la Recherche Scientifique (CNRS-UMR 3525), Fondation pour la Recherche Médicale (FRM grant no. EQU202103012569 to D.M.), ANR Chromintevol (ANR-21-CE12-0002-01 to C.L.), French Government's Investissement d'Avenir program Laboratoire d'Excellence "Integrative Biology of Emerging Infectious Diseases" (ANR-10-LABX-62-IBEID to D.M.), Ministère de l'Enseignement Supérieur et de la Recherche, and the Direction Générale de l'Armement (DGA). **Author contributions:** Conceptualization: E.R., C.L., D.M., D.B., and E.P.C.R. Methodology: E.R., C.L., J.B., and V.P. Investigation: E.R., B.D., E.L., C.V., C.W., J.B., F.F., G.G., V.C., D.L., V.P., F.R., O.S., and C.L. Supervision: C.L., D.M., D.B., and E.P.C.R. Writing—original draft: E.R., C.L., and D.M. Writing—review and editing: E.R., C.L., and D.M. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The sequences data are publicly available in The European Nucleotide Archive (ENA). For the SCI Inv TAi and SCI Reinv TAi strains, the accession numbers are ERS16645193 (R555), ERS16645194 (S507), ERS16645195 (S513), and ERS16645196 (R556). For the MFA analysis, the accession numbers are ERR11475429, ERR11475428, ERR11475427, ERR11475426, ERR11475425, ERR11475424, ERR11475423, and ERR11475422 (SCI Inv and SCI Reinv with and without the WT integrase); ERS16808486, ERS16808487, ERS16808488, and ERS16808489 (SCI Inv and SCI Reinv with Y302F integrase); ERS16808490, ERS16808491, ERS16808492, and ERS16808493 (reloc SCI lag and reloc SCI lead with the WT integrase).For the PacBio sequencing, the accession numbers are ERR11475405, ERR11475404, ERR11475403, and ERR11475402. For the SCI Inv TAi strains 1-day evolution, the accession numbers are ERS16808481, ERS16808482, ERS16808484, and ERS16808485. The dCas9-sgRNA plasmid expression system (pFR48) was derived from plasmid pJF1, which can be provided by the Institut Pasteur pending scientific review and a completed material transfer agreement. Requests for the materials should be submitted to C.L.

Submitted 21 June 2023
Accepted 14 December 2023
Published 12 January 2024
10.1126/sciadv.adj3498

ScienceAdvances

Cassette recombination dynamics within chromosomal integrons are regulated by toxin-antitoxin systems

Egill Richard, Baptiste Darracq, Eloi Littner, Claire Vit, Clémence Whiteway, Julia Bos, Florian Fournes, Geneviève Garriss, Valentin Conte, Delphine Lapaillerie, Vincent Parissi, François Rousset, Ole Skovgaard, David Bikard, Eduardo P. C. Rocha, Didier Mazel, and Céline Loot

Sci. Adv. **10** (2), eadj3498. DOI: 10.1126/sciadv.adj3498

Science Advances (ISSN 2375-2548) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.