Surface exclusion of IncC conjugative plasmids and their relatives

Nicolas Rivard[1#a], Malika Humbert[1¶], Kévin T. Huguet[1#b¶], Aurélien Fauconnier[1¶], César Pérez Bucio[2], Eve Quirion[1], Vincent Burrus[1*]

[1]Département de biologie, Université de Sherbrooke, Sherbrooke, Quebec, Canada

[2]Instituto Tecnológico y de Estudios Superiores de Monterrey, Monterrey, Nuevo León, Mexico

[#a]Current address: Laboratoire d'Ingénierie des Systèmes Macromoléculaires, UMR7255, Institut de Microbiologie de la Méditerranée, Aix-Marseille Univ, CNRS, Marseille, France.

[#b]Current address: CNRS, LCPME, UMR 7564, Université de Lorraine, Villers-lès-Nancy, France

Short title: Surface exclusion of IncC conjugative plasmids

*Corresponding author

E-mail: Vincent.Burrus@USherbrooke.ca

[¶]These authors contributed equally to this work.

Abstract

The phenomenon of exclusion allows conjugative plasmids to selectively impede the entry of identical or related elements into their host cell to prevent the resulting instability. Entry exclusion blocks DNA translocation into the recipient cell, whereas surface exclusion destabilizes the mating pair. IncC conjugative plasmids largely contribute to the dissemination of antibiotic-resistance genes in Gammaproteobacteria. IncC plasmids are known to exert exclusion against their relatives, including IncC and IncA plasmids, yet the entry exclusion factor *eexC* alone does not account for the totality of the exclusion phenotype. In this study, a transposon-directed insertion sequencing approach identified *sfx* as necessary and sufficient for the remaining exclusion phenotype. Sfx is an exclusion factor unrelated to the ones described to date. A cell fractionation assay localized Sfx in the outer membrane. Reverse transcription PCR and beta-galactosidase experiments showed that *sfx* is expressed constitutively at a higher level than *eexC*. A search in Gammaproteobacteria genomes identified Sfx homologs encoded by IncC, IncA and related, untyped conjugative plasmids and an uncharacterized family of integrative and mobilizable elements that likely rely on IncC plasmids for their mobility. Mating assays demonstrated that *sfx* is not required in the donor for exclusion, ruling out Sfx as the exclusion target. Instead, complementation assays revealed that the putative adhesin TraN in the donor mediates the specificity of surface exclusion. Mating assays with TraN homologs from related untyped plasmids from *Aeromonas* spp. and *Photobacterium damselae* identified two surface exclusion groups, with each Sfx being specific of TraN homologs from the same group. Together, these results allow us to better understand the apparent incompatibility between IncA and IncC plasmids and to propose a mechanistic model for surface exclusion mediated by Sfx in IncC plasmids and related elements, with implications for the rampant dissemination of antibiotic resistance.

Author summary

Bacterial conjugation plays a pivotal role in the evolution of bacterial populations. The circulation of drug resistance genes bolsters the emergence of multidrug-resistant pathogens, with which contemporary medicine struggles to cope. Exclusion is a natural process preventing the redundant acquisition of a plasmid via conjugation by a host harbouring an identical or similar plasmid. Although exclusion has been known for the past half-century, the mechanisms involved remain poorly understood. This study describes an exclusion factor, Sfx, encoded by IncC, IncA and related conjugative plasmids and by unrelated integrative and mobilizable elements. We report that Sfx is a lipoprotein of the recipient that selectively inhibits conjugation based on the adhesin TraN expressed at the surface of the donor. We propose a mechanistic model for Sfx-mediated exclusion. Ultimately, a better understanding of exclusion could facilitate the design of conjugation inhibitors targeting mating pair formation to curb the circulation of drug-resistance genes in healthcare settings, agriculture, animal husbandry and food and drug production.

Introduction

IncC plasmids are large conjugative plasmids frequently associated with multidrug resistance phenotypes in a broad range of Gammaproteobacteria species [1]. IncC plasmids' ability to mobilize non-autonomous, unrelated integrative and mobilizable elements (IMEs) of the *Salmonella* Genomic Island 1 (SGI1) and MGI*Vch*Hai6 families exacerbates their importance in antibiotic resistance dissemination [2–5]. SGI1 and its multiple variants are important vehicles of antibiotic resistance genes and are frequently found in *Salmonella enterica* and a broad range of Gammaproteobacteria [6]. MGI*Vch*Hai6 conferred multidrug resistance to *Vibrio cholerae* non-O1/non O139 infecting cholera patients during the 2010 cholera outbreak in Haiti [4]. IncC, IncA and untyped IncA/C-like (ACL) plasmids (e.g., pAsa4c, pAhD4-1, pAQU1) share a syntenic set of genes involved in conjugation, DNA repair, and regulation. All seem regulated by closely related homologs of the master activator of transfer AcaCD [7–9]. IncC plasmid-encoded AcaCD promotes the excision and mobilization of MGI*Vch*Hai6-like IMEs [4,5]. AcaCD also stimulates the excision and replication of SGI1 and the expression of its *traN*, *traG* and *traH* genes, which modify the mating apparatus of IncC plasmids to enhance its dissemination [7,10,11].

Large conjugative plasmids are autonomous replicons that usually remain at a low-copy number. The entry into the same cell of a plasmid sharing similar replication or partitioning determinants promotes plasmid destabilization and loss, a phenomenon known as incompatibility [12]. Exclusion refers to the mechanisms employed by conjugative elements to preclude instability resulting from incompatibility by hindering redundant transfer into the cells they reside [13]. While the definite nature of exclusion mechanisms remains elusive, they can be categorized as entry and surface exclusion by whether they interfere with DNA transfer or mating pair stabilization [13,14]. Exclusion seems to be a staple of the conjugative plasmid lifestyle. Exclusion factors have been identified in the IncF, IncHI1, IncI, IncN, IncP, IncW and ColE1

families, plasmids of the pKPC_UVA01 replicon type and the enterococcal conjugative plasmid pCF10 [14–22]. Integrative and conjugative elements (ICEs) of the SXT/R391 family and ICE*Bs*1 of *Bacillus subtilis* also encode exclusion determinants [23,24]. Most exclusion factors identified to date are associated with the cytoplasmic membrane, where they would exert entry exclusion, with only the IncF TraT firmly established as a surface exclusion factor [13,14]. Like TraT, the IncHI1 EexB localizes in the outer membrane and is thought to engage in surface exclusion [15]. Likewise, the surface-exposed PrgA protease of pCF10 reduces aggregation and transfer [22].

Exclusion is not specific to the transferred DNA [25], and proteins in the donor have been identified as targets of exclusion factors in several systems. EexA and EexB of IncHI1 plasmids and TrbK of pKPK_UVA01 appear to act concurrently as exclusion factors and exclusion targets, as they are necessary in both the donor and recipient for exclusion to take place [15,21]. Other exclusion factors target different plasmid-encoded proteins. The IncI ExcA targets TraY, which is found in the cytoplasmic membrane of the donor [16]. Similarly, the IncF TraS and SXT/R391 Eex of the recipient seem to interact directly with the VirB6 homolog TraG of the donor, and the regions involved in both proteins are cytoplasmic [23,26]. In the cases above, the genes encoding the exclusion factor and its target are adjacent. In IncP, IncW and IncN plasmids, for which the exclusion target remains unknown, the exclusion gene abuts a gene coding for a VirB6 homolog, suggesting a widespread usage of this target beyond the IncF and SXT/R391 families [13].

The surface exclusion factor TraT of IncF plasmids reduces the percentage of mating aggregates, and five contiguous residues determine specificity between exclusion groups [14,27]. Harrison *et al*. proposed that TraT destabilizes the mating pair by interacting with the pilus tip; however, it was eventually shown that the pilin does not define exclusion specificity

[27,28]. The discovery of interactions between TraT and the outer membrane protein OmpA in the recipient and between OmpA in the recipient and the putative adhesin TraN in the donor suggested surface exclusion could instead result from TraT preventing docking of TraN to OmpA, thus hindering mating pair stabilization [29–31]. The identification of TraN receptor specificity groups across IncF plasmids makes this hypothesis extremely appealing, yet *traN* substitutions failed to flip exclusion specificity between F and R100-1 [30,32,33].

IncC plasmids share with IncA plasmids the same entry exclusion system, *eexC*/*traG*, and inhibit each other's entry into the cell in which they reside [9,34]. During the characterization of *eexC*/*traG*, we reported that knocking-out *eexC* of the IncC plasmid pVCR94 in the recipient only reduced the exclusion phenotype in mating assays using an IncC$^+$ donor. Likewise, the deletion of a 45.6-kb fragment encompassing 33 open reading frames located between *traN* and *traF* led to a statistically significant reduction of exclusion. Hence, another mechanism likely prevents redundant transfer between cells containing identical plasmids.

In this report, we investigate the cause of the incomplete abolition of exclusion by a recipient strain bearing an IncC plasmid lacking the entry exclusion gene *eexC*. Using a transposon-directed insertion sequencing (TraDIS) approach, we identified IncC-borne candidate genes whose disruption in recipient cells facilitated the acquisition of an incoming mobilizable plasmid from an IncC$^+$ donor. Further analyses revealed that a single open reading frame, hereafter named *sfx*, flanked by *traN* and *acaB*, accounts for the residual exclusion phenotype. Deletion of *sfx* with *eexC* completely abolished exclusion. Signal peptide prediction and a subcellular localization assay suggest the Sfx protein is an outer-membrane lipoprotein and support a role as a surface exclusion factor. Sfx belongs to an uncharacterized family of broadly distributed surface exclusion proteins encoded by IncC plasmids, their relatives and several putative IMEs. Surface exclusion assays designed to test *sfx*/*traN* gene pairs from ACL plasmids revealed that

sfx and *traN* are part of a surface exclusion system, enabling us to propose a model of surface exclusion of ACL plasmids.

Results

vcrx085 impedes conjugative transfer between IncC$^+$ cells

To identify additional factors responsible for the inhibition of transfer between IncC$^+$ cells, we used transposon-directed insertion sequencing (TraDIS) aimed at mapping Tn*5* insertions that enhance the entry of the mobilizable plasmid pClo transferred from an IncC$^+$ donor strain into a Tn*5*$^+$ IncC$^+$ recipient library (Fig 1A). pClo was used as a proxy for conjugative transfer to bypass IncC plasmid incompatibility. Briefly, we constructed a high-density mini-Tn*5* (Sp^R) insertion library in *E. coli* GG56 carrying pVCR94^{Kn} Δ*eexC*::*cat*, a kanamycin-resistant variant of pVCR94 unable to exert entry exclusion in recipient cells [9]. This set of mutants formed the input library containing Tn*5* insertions in the chromosome or pVCR94 that allowed plasmid replication and maintenance. The input library was then used as recipient in a mating assay with a donor strain carrying pVCR94^{Kn} and the mobilizable plasmid pClo. GG56 transconjugant colonies that acquired pClo made up the output library. We expected that genes or sequences of pVCR94^{Kn} Δ*eexC*::*cat* more frequently disrupted in the output than in the input (high insertion index ratio) were those impeding the mobilization of pClo between IncC$^+$ cells. Since our assays monitored pClo entry only, we did not consider the role of chromosomal genes in the recipient.

Fig 1. Identification of an additional IncC exclusion factor. (A) Overview of the TraDIS workflow to identify exclusion factors in pVCR94^{Kn} Δ*eexC*::*cat*. (B) Mini-Tn5 reads near *vcrx085*. The tracks plot the number of reads from two independent replicates as a function of position in pVCR94^{Kn} Δ*eexC*::*cat* for both the input and output libraries. ORFs with similar functions are colour-coded as indicated in the panel. This panel was created using the UCSC Genome Browser (http://genome.ucsc.edu). (C) Nine loci of unknown function show a mini-Tn*5* insertion index ratio greater than 5 (dashed line), calculated as the ratio of insertion counts between the output and input libraries. ORFs with similar functions are colour-coded as indicated in Fig S1.

(D) *sfx* (*vcrx085*) is the only functional exclusion factor among candidates identified in (B) and suffices alongside *eexC* (*eex*) to account for the totality of IncC exclusion. *E. coli* VB112 (Rf^r) containing pVCR94^{Kn} and pClo (Ap^r) served as the donor strain, and *E. coli* GG56 (Nx^r) containing the specified derivatives of pVCR94^{Sp} served as the recipient strains. When indicated, *eexC* or *sfx* were expressed in the recipients from single-copy, chromosomally integrated pAH56 (Kn^r) in the presence of IPTG or together from pBeloBAC11 (Cm^r) under the control of their respective native promoters. Transconjugants containing pClo were selected as the Nx^r Ap^r colonies. Crossbars show the mean and standard error of the mean of three independent experiments. One-way ANOVA (p=1.6e-25) with a Tukey-Kramer post-test was used on the log_{10}-transformed values to compare the means. Statistical significance (S2 File) is shown as a compact letter display for pairwise comparisons where means grouped under identical letters are not statistically different. Exclusion indices, shown at the bottom of each crossbar, are calculated as the frequency of transfer into the empty recipient divided by the frequency of transfer into the indicated mutant.

We identified nine candidate genes in pVCR94^{Kn} Δ*eexC*::*cat* with insertion index ratios above the arbitrary threshold of 5 (Table 1, Fig 1B and 1C, S1 Fig). Among these genes, only *acr2* (*vcrx150*) is a known repressor of conjugative transfer [7]. To validate the ability of the candidates to inhibit incoming conjugative transfer, we tested mobilization of pClo into the corresponding deletion mutants of pVCR94^{Sp} used as recipients. All but one had no impact on transfer of pClo (Fig 1D). The deletion of *vcrx085* (hereafter referred to as *sfx* for surface exclusion) improved conjugation to the same level similar as the deletion of the entry exclusion factor *eexC*. Like *eexC*, we could complement the *sfx* deletion in *trans* (Fig 1D, pAH56). Both mutations combined completely abolished exclusion, allowing conjugation to a level comparable to an empty recipient. Conversely, the standalone expression of *sfx* and *eexC* from their native

promoter in the absence of an IncC plasmid in the recipient matched the full exclusion phenotype of the wildtype IncC plasmid (Fig 1D, pBeloBAC11).

Table 1. Candidate genes identified by TraDIS.

ORF	Fold change	Predicted size (aa)	Predicted product[a]
vcrx018	6.39	119	Hypothetical protein
vcrx019	5.17	148	Fe^{3+}-siderophore ABC transporter permease
vcrx062	8.67	194	Hypothetical protein
vcrx085 (*sfx*)	6.22	286	Lipoprotein (Sec/SPII)
vcrx107	6.07	114	Hypothetical protein
vcrx108	6.08	100	Hypothetical protein
vcrx111	6.83	184	Hypothetical protein
vcrx137	6.63	163	Hypothetical protein
acr2	6.50	139	Transcriptional repressor

[a]As predicted by NCBI blastp or SignalP 6.0.

Deletion of *eexC*, *sfx*, or both did not significantly impact the fitness of the cells (similar CFU counts/ml) when compared to wild-type pVCR94^{Sp} or empty cells (S2 Fig). Conversely, the overexpression of *eexC* or *sfx* from pBAD30 slightly decreased the fitness.

Sfx is a constitutively expressed outer membrane protein

The SignalP 6.0 server predicts a lipoprotein signal peptide in the translation product of *sfx* (S3 Fig), suggesting Sfx is displayed on the cell surface [35]. A C-terminally 3xFLAG-tagged version of *sfx* was cloned under the control of the inducible promoter P_{tac}. The Sfx^{3xFLAG} polypeptide has a predicted size of 33.4 kDa before signal peptide cleavage and 31.7 kDa afterwards. A cell fractionation assay confirmed the presence of Sfx^{3xFLAG} in the inner and outer membrane fractions (Fig 2A). In contrast, the known outer membrane protein OmpC, used as a control, was only detected in the outer membrane fraction. This cellular localization remains consistent with the role of Sfx in surface exclusion. The accumulation of Sfx^{3xFLAG} in the inner membrane suggests that the high expression level of *sfx*^{3xFLAG} from P_{tac} on a multi-copy vector saturates the Lpt trafficking pathway. We could not undoubtedly assign the processed and unprocessed forms

in the inner membrane fraction, and two additional products appeared between 20 and 25 kDa on the western blot. These weaker bands are presumably degradation products of an unknown nature.

Fig 2. Expression of IncC surface exclusion is independent of the master activator of transfer AcaCD. (A) Cell fractionation assay. The soluble (S), inner membrane (IM), and outer membrane (OM) cell fractions of *E. coli* DH5α expressing 3xFLAG-tagged Sfx_{94} or 3xFLAG-tagged AcaC from pAH56 under the control of the IPTG-inducible P_{tac} promoter with (+) and without (-) IPTG were analyzed by SDS-PAGE and immunodetected using rabbit anti-FLAG or anti-OmpC antibodies. Molecular weight markers (M) are indicated on the left. A vertical white line separates non-adjacent lanes from a single blot. AcaC-3xFLAG (22 kDa), a known cytoplasmic protein [7] and OmpC (36 kDa), a known outer membrane protein [36], were used as controls in the fractionation. (B) Schematic representation of the *acaB-sfx* locus. The relative positions of reverse transcription primer 85RT and primers used to amplify *sfx* and *acaB* are shown. A dotted line shows the reverse transcription product. Stem loops represent predicted Rho-independent terminators. The translational *lacZ* fusion used for enzymatic assays was introduced after the third codon of *sfx*. (C) *sfx* is expressed constitutively and independently of the upstream AcaCD-dependent P_{acaB} promoter. A 2% agarose gel of the PCR products from an *sfx*-derived reverse transcribed RNA. Untreated RNA (RNA) and genomic DNA from *E. coli* GG56 (Nx^r) containing $pVCR94^{Kn}$ Δ*acaCD* (gDNA) were used as negative and positive controls, respectively. (D) β-galactosidase activity of GG56 (Nx^r) containing $pVCR94^{Kn}$ *eexC'-'lacZ*, $pVCR94^{Kn}$ *sfx'-'lacZ*, or $pVCR94^{Kn}$ *traN'-'lacZ* translational fusions with or without p*acaCD* and arabinose. Crossbars show the mean and standard error of the mean of three independent experiments. One-way ANOVA with a Tukey-Kramer post-test was used to compare the means.

Statistical significance is shown as a compact letter display for pairwise comparisons where means grouped under identical letters are not statistically different.

sfx is located downstream of the transfer activator *acaB*, which is expressed from an AcaCD-inducible promoter (Fig 2B) [37].. Hence, we wondered whether *sfx* expression was controlled by AcaCD. To test this hypothesis, we constructed a translational fusion between the third codon of *sfx* and the eighth codon of *lacZ* in pVCR94^{Kn}. Expression of AcaCD did not impact the β-galactosidase activity of the *sfx'-'lacZ* fusion. A similarly constructed *eexC'-'lacZ* fusion yielded a ~150-fold lower β-galactosidase activity, regardless of the presence of AcaCD (Fig 2D). We used a similar *traN*'-'*lacZ* fusion as a control and observed a very low level of expression in IncC^{+} cells, which increased ~260 fold upon AcaCD overexpression. Furthermore, to test whether *sfx* and *acaB* are part of the same transcript, we attempted to amplify both loci from a cDNA synthesized from the 3' end of *sfx*. While *sfx* could be amplified regardless of AcaCD, *acaB* remained undetectable, suggesting that these two genes are expressed as independent mRNA transcripts (Fig 2C). A Rho-independent terminator predicted between *sfx* and *acaB* supports this idea (Fig 2B, S3B Fig). Altogether, our results indicate that *sfx* is constitutively expressed independently of AcaCD at a much higher level than *eexC*.

Genes encoding homologs of Sfx$_{94}$ are found in diverse genomic backgrounds

To explore the diversity of exclusion factors, we first used NCBI tblastn to identify Sfx$_{94}$ homologs in the ACL plasmids pAsa4c of *Aeromonas salmonicida*, pAhD4-1 of *Aeromonas hydrophila*, and pAQU1 of *Photobacterium damselae* that are known untyped relatives of IncA and IncC plasmids [8,9]. For pAsa4c and pAhD4-1, closely related homologs were found adjacent to the gene *traN*. For pAQU1, the search resulted in a single hit (*215*, hereafter referred to as *sfx*$_{AQU1}$) whose product shares only 39% identity over 273 amino acid residues with Sfx$_{94}$. *sfx*$_{AQU1}$ is surrounded by genes of unknown function. Sfx$_{AQU1}$ was manually added to

the dataset for downstream analyses. We next searched the GenBank database for homologs of Sfx_{94} and extracted the resulting sequences for further analyses. 130 unique proteins were identified (S3 File), most encoded by ACL plasmids (Fig 3A). 78 Sfx homologs share more than 90% identity with Sfx_{94}, including nine variants encoded by IncC plasmids. All of the Sfx homologs encoded by IncC plasmids are inside this cluster. Sfx_{94} belongs to an identical protein group of 193 representatives, all encoded by IncC or putative IncC plasmids. Sfx_{RA1} from the IncA plasmid pRA1 belongs to the Sfx_{94} cluster as they share 92% identity over 286 amino acid residues. All plasmids outside the Sfx_{94} cluster are ACL plasmids, including pAsa4c and pAhD4-1. In most ACL plasmids and with the rare exception of pAQU1, *sfx* lies downstream of the *dsbC-traC-trhF-traWU* operon between *traN* and the transfer activator gene *acaB* (Fig 3C).

Fig 3. Sfx homologs are found across a wide range of Gammaproteobacteria. Maximum likelihood phylogenetic analyses of Sfx (A) and TraN (B) homologs. Trees with the highest likelihood (9,621.78 and 28,204.23 for Sfx and TraN, respectively) are shown. Bootstrap supports are indicated as percentages at the branching points only when >80%. Branch lengths represent the number of substitutions per site over 269 and 813 amino acid positions for Sfx and TraN, respectively. Only one representative per cluster of similar proteins (>90%) is shown in each tree, except for $TraN_{Asa4c}$ and $TraN_{AhD4-1}$, which share 95% identity over the aligned sequences. S3 File provides additional details on host strains. (C) Schematic representation of the genetic context of *sfx* homologs in a selection of putative IMEs and ACL plasmids. Predicted *attL* and *attR* attachment sites flanking the IMEs are represented by blue bars. pAQU1 encodes homologs of *acaB*, *vcrx087-088-089-090*, *bet* and *exo* clustered in a distinct locus not shown here. ORFs with similar functions are colour-coded as indicated in the panel.

The remaining Sfx proteins are encoded by putative IMEs integrated at four distinct chromosomal locations. First, ten of these IMEs belong to the MGI*Vch*Hai6 family and are

integrated at the 5' end of *dusA* (tRNA-dihydrouridine synthase A) in environmental and clinical *Aeromonas* strains [5]. Represented by IE*Ave*Tha1 (Fig 3C), all possess MGI*Vch*Hai6-like *mobI* and *oriT* and predicted AcaCD-responsive promoters (S2 Table). A second group represented by IE*Vch*Aut1 comprises six IMEs integrated at the 3' end of *yicC* (putative RNase adaptor protein YicC) in environmental and clinical *Vibrio* strains. In addition to a putative recombination directionality factor related to RdfM of MGI*Vfl*Ind1 [38], these IMEs carry a set of conserved genes, including a predicted type IV coupling protein (*traD*) and a putative relaxase of the MOB_F family (*traI*). Finally, two IMEs coding for an identical Sfx are integrated at the 3' end of *tRNA-TRP* and the 5' end of *rlmD* (23S rRNA uracil methyltransferase) in environmental isolates of *Shewanella chilikensis* and *Shewanella putrefaciens*, respectively. Despite sharing a conserved core set of genes with IE*Vch*Aut1, the Sfx protein of these two IMEs resembles and clusters with Sfx_{94} (Fig 3A). Although we can predict AcaCD binding sites across these IMEs (S2 Table), no other conserved features of the SGI1 and MGI*Vch*Hai6 families was found, suggesting they belong to an unrelated IME family mobilizable by IncC plasmids.

Surface exclusion specificity is determined by TraN

To test the role of Sfx in the donor, we monitored the transfer of pClo from donors carrying the Δ*eexC* and Δ*sfx* deletion mutants of $pVCR94^{Sp}$ toward an $IncC^+$ strain ($pVCR94^{Kn}$). No alleviation of exclusion, i.e., an increase in transfer frequency, was observed (Fig 4A). On the contrary, a modest yet significant decrease was observed for the Δ*sfx* and Δ*eexC* Δ*sfx* mutants, suggesting that augmented redundant transfer between donors was detrimental to transfer toward recipients. Hence, *sfx* is not self-targeting to promote surface exclusion.

Next, we focused on TraN as a potential target for surface exclusion, as the gene *traN* encodes a putative adhesin thought to stabilize the mating pair, and it is adjacent to *sfx* in most conjugative plasmids found but pAQU1 [30,31,39]. A search for homologs of $TraN_{94}$ yielded 412

homologous proteins (S3 File), most encoded by ACL plasmids. Others belonged to IncC-mobilizable IMEs of the SGI1 family integrated into *trmE* or *dusA*, with no clear relationship between TraN phylogeny and insertion site (Fig 3B). The cluster represented by pAQU1 differs significantly from the close ACL plasmids as the *dsbC-traC-trhF-traWUN* and *acaB*/*bet-exo* regions are not adjacent, and *sfx* lies elsewhere.

Fig 4. TraN directs surface exclusion specificity. (A) The absence of *eexC* and *sfx* in the donor decreases the frequency of transconjugant formation. pClo (Ap^r) was mobilized from *E. coli* VB112 (Rf^r) bearing the indicated deletion mutants of $pVCR94^{Sp}$ (donor, D) into *E. coli* GG56 (Nx^r) bearing $pVCR94^{Kn}$ (recipient). Transconjugants containing pClo were selected as the Nx^r Ap^r colonies. Crossbars show the mean and standard error of the mean of three independent experiments. One-way ANOVA (p=0.016) with a Tukey-Kramer post-test was used on the $\log_{10}$-transformed values to compare the means. Statistical significance (S2 File) is shown as a compact letter display for pairwise comparisons where means grouped under identical letters are not statistically different. (B) sfx_{94} is ineffective against the $traN_{94}$ distant homolog $traN_{AQU1}$. Pairwise t-tests were used on the $\log_{10}$-transformed values to compare the means. Exclusion indices are calculated by dividing the frequency of transfer of $pVCR94^{Sp}$ Δ*traN* to the empty recipient divided by the transfer frequency to the recipient expressing sfx_{94}. (C) Heatmap of surface exclusion indices of *sfx* and *traN* pairs from a diverse set of ACL plasmids. Exclusion indices are calculated by dividing the frequency of transfer of $pVCR94^{Sp}$ Δ*traN* to the empty recipient divided by the transfer frequency to the recipient expressing the indicated *sfx* variant. In (B) and (C), *E. coli* GG56 (Nx^r) containing $pVCR94^{Sp}$ Δ*traN* and expressing the indicated *traN* variant from pBAD30 (Ap^r) (donor, D) was mated with *E. coli* CAG18439 (Tc^r) expressing the indicated *sfx* variant from a single-copy, chromosomally integrated pAH56 (Kn^r) (recipient, R). Transconjugants containing $pVCR94^{Sp}$ Δ*traN* were selected as the Tc^r Sp^r colonies.

To investigate the involvement of TraN in Sfx-dependent exclusion, we used in mating assays a *traN* deletion mutant of pVCR94^{Sp} complemented ectopically with either its native $traN_{94}$ or the distantly related homolog $traN_{AQU1}$ [40]. Despite sharing only 66% identity with $TraN_{94}$, $TraN_{AQU1}$ fully complemented the Δ$traN_{94}$ mutation (Fig 4B). Furthermore, whereas the expression of sfx_{94} in the recipient strongly inhibited conjugation supported by $traN_{94}$, it did not impair transfer supported by $traN_{AQU1}$ (Fig 4B). This result implies that TraN is, if not a direct interactant of Sfx, at least a key player in Sfx-mediated exclusion and the determinant of exclusion specificity. To further investigate the diversity of exclusion specificity, we tested the *traN*/*sfx* pairs of pAsa4c and pAhD4-1, two plasmids belonging to distinct entry exclusion groups, and pAQU1 [9,41,42]. The corresponding genes were tested in pairs in all possible combinations to determine exclusion indices. Two separate exclusion groups emerged (Fig 4C and S5 Fig). *sfx* homologs from pVCR94, pAsa4c and pAhD4-1 specifically exerted exclusion against *traN* homologs from these same plasmids. sfx_{AQU1} only excluded $traN_{AQU1}$. In addition, we found that sfx_{Aut1} from the putative IME IE*Vch*Aut1 is functional. sfx_{Aut1} exhibits a surface exclusion pattern resembling sfx_{Asa4c}, excluding transfer mediated by *traN* of pVCR94, pAsa4c and pAhD4-1 but not pAQU1.

Discussion

All known conjugative plasmids encode at least one exclusion factor [13]. An entry exclusion system involving the entry exclusion factor EexC and the VirB6-homolog TraG was previously identified in IncA and IncC conjugative plasmids [9]. Here, we identified a second exclusion factor within a region suspected by Humbert *et al*. to be responsible for additional exclusion activity [9]. A TraDIS experiment pointed to nine candidate genes likely to enhance incoming transfer once disrupted. *sfx* was the only one whose activity in exclusion could be experimentally validated. Together, *eexC* and *sfx* account for the totality of the exclusion phenotype (Fig 1D), akin to the exclusion mediated by *traS* and *traT* of the F plasmid [43]. We

could not detect any impact of abolishing exclusion on cell viability (S2 Fig). This result strongly contrasts with reports on F plasmid exclusion for which double mutants of *traS* and *traT* could not be isolated [43]. This discrepancy could be due to the strong repression of the conjugative machinery of IncC plasmids (Fig 2D), whereas F transfer is constitutive due to the inactivation of F *finO* by IS*3* [7,44].

The translation product of *sfx* is a predicted lipoprotein, and we detected Sfx in the outer membrane (Fig 2A), a trait shared with the surface exclusion factor TraT of the F plasmid. In addition, we showed that exclusion mediated by *sfx* is inhibited or abolished by the substitution of the putative adhesin TraN (Fig 4C), suggesting that exclusion specificity is determined by *sfx*/*traN* pairs. Together, these observations support a mechanism of surface exclusion [14,45].

Our search in the Genbank database using Sfx_{94} failed to reveal any known entry or surface exclusion factor, including F plasmid TraS or TraT, establishing Sfx of IncA and IncC plasmids as an unrelated surface exclusion factor (S3 File). As the ACL plasmids pAsa4c and pAhD4-1 belong to the same surface exclusion group as pVCR94 despite their divergence (Fig 4C), IncA and IncC plasmids are expected to exert surface exclusion against each other, hence contributing alongside entry exclusion to the previously observed exclusion phenotype between the two incompatibility groups [9,34]. Among the Sfx homologs tested in this study, only Sfx_{AQU1}, which failed detection by the NCBI blastp search due to low similarity with Sfx_{94}, belongs to a separate surface exclusion group (Fig 4C). A pairwise comparison illustrates that, like Sfx_{AQU1}, the putative surface exclusion proteins WP_035235290.1 and WP_121638668.1 share minimal homology with all other representatives (Fig 3A and S4 Fig). Therefore, we speculate that neither belongs to the two surface exclusion groups described here and that all other IncC, IncA and close ACL plasmids belong to the Sfx_{94} surface exclusion group. This hypothesis is further supported by the exclusion pattern of sfx_{Aut1} (IE*Vch*Aut1), virtually identical to sfx_{Asa4c}. From our

observations with ACL plasmids, surface exclusion strikingly contrasts with entry exclusion, which seems much more specific. Indeed, pAsa4c and pAhD4-1 each belong to separate entry exclusion groups (EexE and EexD, respectively) distinct from the one encompassing IncC and IncA plasmids (EexC) [9]; yet, all belong to the same surface exclusion group. These observations suggest IncC, IncA and ACL plasmids form a two-tier community, exerting surface and entry exclusion against closely related peers but only surface exclusion against a broader range of distantly related conjugative elements.

This work has identified orphan surface exclusion factors borne by IMEs. IE*Ave*Tha1-like elements belong to the MGI*Vch*Hai6 family of IncC-mobilizable IMEs [4,5], whereas IE*Vch*Aut1-like elements constitute a new family of IMEs likely to be mobilizable by ACL plasmids (Fig 3C). We showed that sfx_{Aut1} of IE*Vch*Aut1 is functional (Fig 4D). Since these IMEs lack a *traN* gene, we propose they engage in surface exclusion as a defensive mechanism, forcing their host to favour mating partners that could enhance both their stability and transmissibility. Hence, these IMEs could modulate the circulation of conjugative elements in bacterial populations without affecting whether surface exclusion from a potential recipient will target transfer mediated by their helper element. Such behaviour would contrast with the way SGI1 manipulates the exclusion system of its helper IncC plasmid. SGI1, which is not known to exert exclusion, evades IncC entry exclusion by substituting TraG in the mating apparatus with a distant VirB6 homolog [11]. Manipulation of exclusion by mobilizable elements is not well known. While homologous ColE1 mobilizable plasmids were reported to inhibit each other's transfer mildly, this phenotype was attributable to the accessory mob gene *mbeD* with no identified target in the donor [20,46,47]. It thus remains unclear whether exclusion is at play in that case. The putative IMEs we reported here may be the first instance of exclusion exerted by mobilizable elements. The close relationship between Sfx proteins encoded by IE*Ave*Tha1 and IE*Vch*Aut1 with those

tested in this study (Sfx_{94}, Sfx_{AhD4-1}, and Sfx_{Asa4c}) suggests they belong to the same exclusion group (Fig 3, S4 Fig). This activity has been corroborated by experimental evidence for Sfx_{Aut1} (Fig 4C, S5 Fig), establishing surface exclusion as an actual barrier to the circulation of mobile genetic elements on par with other bacterial defense mechanisms instead of being limited to preventing redundant transfer and, potentially, mediating mating pair separation [13,39]. Surprisingly, none of the IMEs identified by the TraN search, including SGI1, appears in the Sfx tree, suggesting they do not encode a surface exclusion factor related to Sfx_{94}. Conversely, IMEs identified by the Sfx search are ostensibly absent from the TraN tree as they lack the components of the conjugative apparatus. This observation suggests different survival strategies. Whereas SGI1 maximizes its own dissemination by remodelling the mating apparatus to improve transfer and bypass entry exclusion [11], IE*Ave*Tha1 and IE*Vch*Aut1 may ensure persistence by hindering via surface exclusion the entry of destabilizing, excision-inducing ACL plasmids. Although the mobilization mechanism of IE*Vch*Aut1 remains unknown, its gene content and the presence of AcaCD binding motifs suggest mechanistic parallels with that of SGI1, albeit using totally unrelated mobilization factors. In ACL^+ cells, AcaCD likely promotes IE*Vch*Aut1's excision through the expression of *rdfM* [38]. A predicted homolog of the replication protein P from bacteriophage lambda suggests the excised IME is replicative [48]. IME DNA transfer seems to rely on the IME-encoded TraI, a predicted relaxase of the MOB_F family [49]. While SGI1 encodes its own replication and DNA processing functions [10,50,51], IE*Vch*Aut1 additionally encodes a type IV coupling protein, presumably making it independent from the IncC-encoded TraD and DtrJ (also known as TraJ) [5,52].

IncHI1 EexB and IncF TraT are two other known surface exclusion factors. While EexB seems to have a dual role as an exclusion factor and exclusion target, TraT's mechanism remains elusive as exclusion specificity could not be attributed to direct interaction with pilin or indirect

interaction with the putative adhesin TraN through OmpA [15,27,28,30]. Our results demonstrate that Sfx is only required in the recipient to enable exclusion (Fig 4A), ruling out the self-targeting of Sfx. Instead, we found that Sfx exclusion can be neutralized by substituting TraN in the donor with TraN from another surface exclusion group (Fig 4B and 4C). Distinct exclusion groups within a given family were previously reported in IncI and IncF plasmids and SXT/R391 ICEs [16,23,53–55].

Although a direct interaction between Eex and TraG has been suggested in SXT/R391 ICEs and likely extends to other VirB6-targeting entry exclusion factors [26], direct contact between Sfx and TraN is unlikely for surface exclusion. In fact, the binding of outer membrane proteins on donor and recipient cells is hard to reconcile with the idea that surface exclusion physically prevents the formation or disrupts mating aggregates [14]. Instead, surface exclusion factors have been proposed to act by concealing outer membrane receptors involved in mating aggregation [29]. According to Riede and Eschbach, TraT of F directly interacts with the major outer membrane protein OmpA as TraT inhibits the OmpA-specific phage K3 [29]. Yet, despite ample evidence that $TraN_F$ interacts with OmpA during conjugation, it does not seem to determine exclusion specificity as $TraN_F$ substitution with $TraN_{R100\text{-}1}$ does not counter exclusion by TraT of F [30–32].

We found that sfx_{94} is constitutively expressed independently of AcaCD at a much higher level than the entry exclusion factor *eexC* or the adhesin gene *traN* when donor cells are in pure culture (Fig 2C and 2D). The high expression of *sfx* relative to *eexC*, despite similar levels of exclusion exerted by the two genes (Fig 1D, Fig 2D), is consistent with an interaction with a highly abundant surface protein. While few EexC proteins might suffice to target TraG in one mating pore during conjugation, an abundance of Sfx proteins is likely required to conceal most, if not all, TraN receptors on the recipient surface at any time. In pure cultures, expression of

traN remained low, suggesting that the level of Sfx could be sufficient to quench spurious donor-donor mating (Fig 2D). Unlike TraT of F, we have identified TraN as the specificity determinant for Sfx-mediated exclusion between ACL plasmids. Thus, we propose that $TraN_{94}$ and $TraN_{AQU1}$ recognize distinct receptors on the recipient surface and that Sfx prevents mating pair stabilization by specific binding to the receptor of its cognate TraN. Further investigation will be required to identify the receptors involved.

Materials and methods

Bacterial strains and media

The bacterial strains and plasmids used in this study are described in Table 2. Bacterial strains were grown at 37°C on Luria-Bertani (LB) agar or with agitation in LB broth and maintained at −75°C in LB broth containing 20% (vol/vol) glycerol. Antibiotics were used at the following concentrations: ampicillin (Ap), 100 µg/ml; kanamycin (Kn), 50 µg/ml; spectinomycin (Sp), 50 µg/ml; rifampicin (Rf), 50 µg/ml; chloramphenicol (Cm), 20 µg/ml and nalidixic acid (Nx), 40 µg/ml. Media were supplemented with 0.02% L-arabinose or 0.1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) to induce gene expression from pBAD30 and pAH56 constructs, respectively.

Table 2. Strains and plasmids used in this study

Strain, plasmid or element	Relevant genotype or phenotype	Reference
Escherichia coli		
DH5α	F^- φ80*lacZ*ΔM15 Δ(*lacZYA-argF*)U169 *recA1 endA1 hsdR17*(r_K^-, m_K^+) *phoA supE44* λ^- *thi*-1 *gyrA96 relA1*	[56]
GG56	Nx^r derivative of BW25113 (Nx)	[57]
KH95	Rf^r derivative of BW25113 (Rf)	[10]
MFD*pir*	MG1655 RP4-2-Tc::[ΔMu1::*aac(3)IV*-Δ*aphA*-Δ*nic35*-ΔMu2::*zeo*] Δ*dapA*::(*erm-pir*) Δ*recA*	[58]
VB112	Rf^r derivative of MG1655 (Rf)	[59]
CAG18439	MG1655 *lacZU118 lacI42*::Tn*10* (Tc)	[60]
Vibrio cholerae		
A12JL5W90	non-O1/non-O139 isolated from a large Austrian lake, 2012	[61]
Plasmids		
pVCR94	IncC conjugative plasmid, *V. cholerae* O1 El Tor (Su Tm Cm Ap Tc Sm)	[62]
$pVCR94^{Sp}$	Sp^r derivative of pVCR94 (pVCR94ΔX2) (Su Sp)	[11]
$pVCR94^{Kn}$	Kn^r derivative of pVCR94 (pVCR94ΔX3) (Su Kn)	[11]
$pVCR94^{Kn}$ Δ*eexC*::*cat*	*eex* deletion mutant of $pVCR94^{Kn}$ (Su Kn Cm)	This study
pFG036	ori_{pMB1}, *cI857* (Ts) repressor, *tetM* (Tc)	Addgene #137996
pFG051	ori_{R6K}, Tn*5 tnp* under λP_L promoter, $oriT_{RP4}$, Tn*5d-aadA7* (Sp)	Addgene #137997
$pVCR94^{Sp}$ Δ*018-019*	*vcrx018-vcrx019* deletion mutant of $pVCR94^{Sp}$ (Su Sp)	This study

pVCR94^{Sp} Δ*062*	*vcrx062* deletion mutant of pVCR94^{Sp} (Su Sp)	This study
pVCR94^{Sp} Δ*107-108*	*vcrx107-vcrx108* deletion mutant of pVCR94^{Sp} (Su Sp)	This study
pVCR94^{Sp} Δ*111*	*vcrx111* deletion mutant of pVCR94^{Sp} (Su Sp)	This study
pVCR94^{Sp} Δ*137*	*vcrx137* deletion mutant of pVCR94^{Sp} (Su Sp)	This study
pVCR94^{Sp} Δ*acr2*	*vcrx150* deletion mutant of pVCR94^{Sp} (Su Sp)	[7]
pVCR94^{Sp} Δ*eexC*	*vcrx145* deletion mutant of pVCR94^{Sp} (Su Sp)	[9]
pVCR94^{Sp} Δ*sfx*	*vcrx085* deletion mutant of pVCR94^{Sp} (Su Sp)	This study
pVCR94^{Sp} Δ*eexC*Δ*sfx*	*vcrx145-vcrx085* double deletion mutant of pVCR94^{Sp} (Su Sp Cm)	This study
pVCR94^{Kn} Δ*acaCD*	*acaCD* deletion mutant of pVCR94^{Kn} (Su Kn)	This study
pClo	CloDF13::TnAΔEcoRV (pSU4628) (Ap)	[63]
p*acaCD*	pBAD30::*acaCD* (Ap)	[7]
pVCR94^{Sp} Δ*traN*	*traN* deletion mutant of pVCR94^{Sp} (Su Sp)	[11]
pVCR94^{Kn} Δ(*087-traG*)	*vcrx87-traG* deletion mutant of pVCR94^{Kn} (Su Kn)	This study
pAsa4c	Cm^{r} conjugative plasmid from *A. salmonicida* JF2267 (Cm)	[41]
p*traN*$_{94}$	pBAD30::*traN*$_{pVCR94}$ (Ap)	[11]
p*traN*$_{Asa4c}$	pBAD30::*traN*$_{pAsa4c}$ (Ap)	This study
p*traN*$_{AhD4-1}$	pBAD30::*traN*$_{pAhD4-1}$ (Ap)	This study
p*traN*$_{AQU1}$	pBAD30::*traN*$_{pAQU1}$ (Ap)	This study
p*sfx*$_{94}$	pAH56::*sfx*$_{pVCR94}$ (Kn)	This study
p*sfx*$_{Asa4c}$	pAH56::*sfx*$_{pAsa4c}$ (Kn)	This study
p*sfx*$_{AhD4-1}$	pAH56::*sfx*$_{pAhD4-1}$ (Kn)	This study
p*sfx*$_{AQU1}$	pAH56::*sfx*$_{pAQU1}$ (Kn)	This study
p*sfx*$_{Aut1}$	pAH56::*sfx*$_{IEVchAut1}$ (Kn)	This study
p*eexC*	pAH56::*eex*$_{pVCR94}$ (Kn)	This study
p*sfx*$^{3\times Flag}$	pAH56::*sfx*$_{pVCR94}$$^{3\times Flag}$ (Kn)	This study
p*acaDC*$^{3\times Flag}$	pAH56::*acaDC*$^{3\times Flag}$ (Kn)	[7]
pVCR94^{Kn} *eexC*'-'*lacZ*	Translational fusion of the 8th codon of *lacZ* at the 3rd codon of *vcrx145* (Su Kn)	This study
pVCR94^{Kn} *sfx*'-'*lacZ*	Translational fusion of the 8th codon of *lacZ* at the 3rd codon of *vcrx085* (Su Kn)	This study
pVCR94^{Kn} *traN*'-'*lacZ*	Translational fusion of the 8th codon of *lacZ* at the 3rd codon *traN* (Su Kn)	This study
pVCR94^{KnΔCD} *eexC*'-'*lacZ*	*acaCD* deletion mutant of pVCR94^{Kn} *eexC*'-'*lacZ* (Su Kn Cm)	This study
pVCR94^{KnΔCD} *sfx*'-'*lacZ*	*acaCD* deletion mutant of pVCR94^{Kn} *sfx*'-'*lacZ* (Su Kn Cm)	This study
pVCR94^{KnΔCD} *traN*'-'*lacZ*	*acaCD* deletion mutant of pVCR94^{Kn} *traN*'-'*lacZ* (Su Kn Cm)	This study
p*eexC-sfx*	pBeloBAC11::(*sfx-eexC*) from pVCR94^{Kn} Δ(*087-traG*) (Cm)	This study
pBeloBAC11	Single-copy vector derived from the F plasmid (Cm)	New England Biolabs
pAH56	*oriV*$_{R6Kγ}$ *attP*$_{λ}$ *lacI* P_{tac} (Kn)	[64]
pBAD30	*ori*$_{p15A}$ *bla araC* P_{BAD} (Ap)	[65]

pINT-ts	*ori*$_{R101}$ *cI857* *λp*$_R$-*int*$_λ$ (ts, Ap)	[64]
pVI42b	pVI36 BamHI::*P*$_{lac}$-*lacZ* (Sp)	[66]
pKD3	*cat* template for one-step chromosomal gene inactivation (Cm)	[67]
pKD4	*aph* template for one-step chromosomal gene inactivation (Kn)	[67]
pCP20	Flp recombinase thermo-inducible encoding plasmid (ts, Ap Cm)	[68]
pSIM6	λ Red recombination thermo-inducible encoding plasmid (ts, Ap)	[69]

Ap, ampicillin; Cm, chloramphenicol; Kn, kanamycin; Nx, nalidixic acid; Rf, rifampicin; Sm, streptomycin; Sp, Spectinomycin; Su, sulfamethoxazole; Tc, tetracycline; Tm, trimethoprim; ts, thermosensitive

Plasmid and strain construction

Plasmid DNA was prepared using the QIAprep Spin Miniprep kit (Qiagen), and genomic DNA was isolated with the QIAamp DNA mini kit (Qiagen) as recommended by the manufacturer. PCR products were purified using the PCR purification kit (Qiagen). All molecular biology manipulations were carried out by standard procedures following the Current Protocols in Molecular Biology [70]. The oligonucleotides used in this study are described in S1 Table. Deletions were constructed using the one-step chromosomal gene inactivation technique (λ Red-mediated mutagenesis) with pKD3 and pKD4 as templates for antibiotic resistance cassettes [67]. The deletions of *vcrx018*-*vcrx019*, *vcrx062*, *vcrx085*, *vcrx107*-*vcrx108*, *vcrx111*, *vcrx137* and *eexC* in pVCR94^{Sp} or pVCR94^{Kn} were done using primer pairs 94delvcrx018.for/94delvcrx019.rev, 94delvcrx062.for/94delvcrx062.rev, 94delvcrx085.for/94delvcrx085.rev, 94delvcrx107.for/94delvcrx108.rev, 94delvcrx111.for/94delvcrx111.rev, 94delvcrx137.for/94delvcrx137.rev and 94deI145.for/94deI145.rev, respectively, and pKD3 or pKD4 as the template. Deletion of *vcrx085* in pVCR94^{Sp} Δ*eexC* was done using primer pair 94delsfx.for/94delsfx.rev and pKD3 as the template. The λRed recombination system was expressed using pSIM6. When appropriate, resistance cassettes were excised from the resulting constructions using the Flp-encoding plasmids pCP20. All deletions were validated by antibiotic profiling and PCR.

The *eexC'-'lacZ*, *sfx'-'lacZ*, and *traN*'-'*lacZ* translational fusions were introduced into pVCR94^{Kn} using primer pairs 94eex-lacZ.f/94eex-lacZ.r, 94sfx-lacZ.f/94sfx-lacZ.r, and 94traN-lacZ.f/94traN-lacZ.r, and pVI42b as the template. In all cases, the third codon of the gene of interest was fused to the eighth codon of *lacZ*. The deletion of *acaCD* in pVCR94^{Kn}, pVCR94^{Kn} *eexC'-'lacZ*, pVCR94^{Kn} *sfx'-'lacZ*, and pVCR94^{Kn} *traN'-'lacZ* was done using primer pair 94DelacaD.for/94DelacaC.rev and pKD3 as the template.

pVCR94 and pAsa4c served as templates to amplify their respective genes of interest. $sfx_{AhD4\text{-}1}$, sfx_{AQU1}, $traN_{AhD4\text{-}1}$ and $traN_{AQU1}$ were chemically synthesized (Bio Basic, Markham). sfx_{94}, sfx_{Asa4c}, $sfx_{AhD4\text{-}1}$, sfx_{AQU1} and *eexC* were amplified using primer pairs 94sfxNdeI.f/94sfxSalI.r, Asa4sfxNdeI.f/Asa4sfxSalI.r, AhD4sfxNdeI.f/AhD4sfxSalI.r, AQU1sfxNdeI.f/AQU1sfxSalI.r, and eexCNdeI.f/eexCSalI.r, respectively. sfx_{Aut1} was amplified using primer pair Aut1sfxNdeI.f/Aut1sfxSalI.r and the genomic DNA of *V. cholerae* A12JL5W90 as the template. Amplicons were digested with NdeI and SalI and cloned into NdeI/SalI-digested pAH56. psfx^{3xFlag} was constructed from psfx_{94} using primer pair pAH56_3xFlag.F/94sfx_3xFlag.R and the Q5 Site-Directed Mutagenesis Kit (New England Biolabs) according to the manufacturer's instructions. The resulting constructs were single-copy integrated into the $attB_{\lambda}$ chromosomal site of CAG18439 using pINT-ts.

To assemble a DNA fragment encompassing sfx_{94} and *eexC* alongside their native promoters, the region spanning from *acaB* to *traG* in pVCR94^{Kn} was deleted using the one-step chromosomal gene inactivation technique with primer pair 94del86acaB.for/94del144traG.rev and pKD3 as the template. After the excision of the resistance cassette using pCP20, the region of interest was amplified using primer pair 94eexBamHI.for/94sfxHindIII.rev. The amplicon was digested with BamHI and HindIII and cloned into a BamHI/HindIII-digested pBeloBAC11.

500 *traN*$_{Asa4c}$, *traN*$_{AhD4-1}$, *traN*$_{AQU1}$ were amplified using primer pairs Asa4traNEcoRI.f/Asa4traNSalI.r,
501 AhD4traNEcoRI.for/AhD4traNEcoRI.rev and AQU1traNEcoRI.for/AQU1traNEcoRI.rev,
502 respectively. Amplicons were digested with EcoRI or EcoRI/SalI and cloned into EcoRI- or
503 EcoRI/SalI-digested pBAD30.
504 All constructs were verified by PCR and DNA sequencing at the Plateforme de Séquençage et
505 de Génotypage du Centre de Recherche du CHUL (Québec, QC, Canada).

506 **Bacterial conjugation and exclusion assays**

507 Bacteria were grown for 16 h in LB broth with the appropriate antibiotics. Mating assays were
508 carried out by mixing 100 µl of donor and recipient cells. Cells were pelleted by centrifugation,
509 then washed once in 1 volume of LB broth and resuspended in 1/20 volume of LB broth.
510 Bacterial mixtures were incubated for 6 h on LB agar plates at 37°C to allow conjugation. Serial
511 dilutions were then plated on selective LB-agar plates with appropriate antibiotics to discriminate
512 between donor, recipient and transconjugant CFUs. Transfer frequencies were calculated by
513 dividing the number of transconjugant CFUs by the number of donor CFUs. Exclusion indices
514 were calculated as the ratio of an element's transfer frequency toward an empty recipient to its
515 transfer frequency toward the tested recipient.

516 **Transposon-directed insertion sequencing (TraDIS)**

517 A conjugation-assisted random transposon mutagenesis experiment was performed on *E. coli*
518 GG56 bearing pVCR94^{Kn} Δ*eexC*::*cat*. The transposition system was composed of *E. coli* MFD*pir*
519 carrying pFG036 (a plasmid coding for the thermosensitive cI857 transcriptional repressor) and
520 pFG051 (a Pi-dependent RP4-mobilizable plasmid coding for the Tn*5* transposition machinery
521 repressed by cI857, and carrying a mini-Tn*5* (Sp) transposon). This diaminopimelate (DAP)-
522 auxotrophic strain contains a chromosomal RP4 conjugative machinery and expresses the Pi
523 protein required for pFG051 replication. The TraDIS experiment was performed in several

successive steps. First, pFG051 was transferred by conjugation from MFD*pir* to GG56 bearing pVCR94^{Kn} Δ*eexC*::*cat* in a 2-h mating experiment at 30°C on LB agar plates supplemented with DAP in duplicates. Once in the recipient strain that lacks cI, the constitutively expressed Tn*5* machinery of pFG051 mediates random mini-Tn*5* (Sp) insertions in the genome. The mating mixture was then entirely spread onto 40 large LB agar plates (150 mm) supplemented with Cm, Kn, Nx, and Sp. Plates were incubated until near confluence (40k to 60k CFUs) to select clones carrying mini-Tn*5* (Sp) insertions. After overnight incubation at 37°C, Cm Nx Kn Sp-resistant colonies were collected using a cell scraper and resuspended in LB broth. The collected sample, designated as the 'input library' was washed, then resuspended in 4.5 ml of LB broth and cryopreserved. The total DNA of a 1.5 ml aliquot of the input library was extracted for sequencing. Another 1 ml aliquot of input library was used to inoculate 50 ml of LB broth supplemented with Cm, Kn, Nx and Sp, which was then grown overnight at 37°C. The resulting culture was used as the recipient in a mating assay, mixed in equal volumes with *E. coli* KH95 carrying pVCR94^{Kn} and pClo used as the donor. After 2-h incubation at 37°C, mating mixtures were spread onto 20 large LB agar plates (150 mm) supplemented with Nx, Kn, Sp, Cm, and Ap to select for mini-Tn*5* (Sp) insertions that allowed the entry and replication of pClo into the recipients. After overnight incubation at 37°C, Nx, Kn, Sp, Cm, and Ap-resistant colonies were collected and subsequently resuspended in LB broth, washed and resuspended in 4.5 ml of LB broth and cryopreserved. These samples were designated as 'output libraries'. The total DNA of a 1.5 ml aliquot of the output libraries was extracted and used for sequencing.

Preparation of TraDIS libraries and Illumina sequencing

For each library, a 1.5 ml frozen stock aliquot was thawed on ice for 15 min and used to prepare sequencing libraries as described previously [71]. Mutant libraries were then pooled and sequenced by Illumina using the NextSeq® 500/550 High Output Kit v2 at the RNomics platform

of the Laboratoire de Génomique Fonctionnelle de l'Université de Sherbrooke (https://rnomics.med.usherbrooke.ca) (Sherbrooke, QC, Canada). The transposon data analysis was carried out as described previously [71].

Fractionation of cellular proteins

Cell fractionation was carried out using a protocol derived from Sandrini *et al.* [72]. Briefly, *E. coli* DH5α λ*pir* bearing p*acaDC*$^{3\times Flag}$ or p*sfx*^{3xFlag} was grown overnight with or without 0.1 mM IPTG. The cells were pelleted at 3,900 *g* for 20 min at 4°C. Cell pellets were washed twice with 10 mM Tris buffer (pH 7.5), resuspended in 10 ml of the same buffer, and frozen for 2 h at -80°C. Samples were lysed by ultrasonication (Qsonica Q125) in an ice bath using five cycles of 30 s each at 70% amplitude, followed by 30 s of cooling. The lysates were treated with DNAse I (1 mg/ml). Cell debris were removed by centrifugation (3,900 *g* for 30 min at 4°C). The proteins in the supernatant were separated by ultracentrifugation (185,000 *g* for 30 min at 4°C). The supernatant (cytoplasmic proteins) was stored at -20°C. The pellet (total membrane proteins) was washed three times with 10 mM Tris buffer (pH 7.5), resuspended in 10 mM Tris buffer with 2% (v/v) Triton X-100, and incubated for 30 min at room temperature. The mixture was centrifuged at 185,000 *g* for 30 min at 4°C. The supernatant (inner membrane proteins) was concentrated using Pierce™ Protein Concentrators PES 3K, following the manufacturer's instructions (Thermo Scientific, cat.# 88515) and stored at -20°C. The pellet (outer membrane proteins) was washed three times with 10 mM Tris buffer (pH 7.5) and stored at -20°C.

SDS-page and Immunoblot analysis

Bradford quantification was used to determine the quantity of fractionated proteins and normalize protein samples. Each normalized fraction was resuspended in NuPAGE LDS Sample Buffer (Invitrogen) supplemented with NuPAGE Sample Reducing agent (Invitrogen), incubated at 70°C for 10 min, and separated by 12% SDS-PAGE (120 mV, 2 h). Proteins were detected by

silver staining (Pierce Silver Stain Kit) following the manufacturer's instructions (Thermo Scientific, cat.# 24612) or transferred onto a polyvinylidene fluoride (PVDF) membrane. Western blotting was performed according to standard procedures. The membranes were probed using 1:300 rabbit OMPC polyclonal primary antibody (ThermoFisher, cat # BS-20213R) or 1:3000 rabbit anti-FLAG primary antibody (Sigma-Aldrich, cat. # F7425), then with 1:3000 HRP goat anti-rabbit IgG HRP-conjugated secondary antibody (Invitrogen, cat. # G-21234). Detection was carried out with the Clarity western ECL substrate (BioRad, cat.# 170-5061) using the ChemiDoc™ MP imaging system (BioRad).

Molecular biology

Plasmid DNA was extracted using the EZ-10 Spin Column Plasmid DNA Minipreps Kit (Bio Basic) following the manufacturer's instructions. Enzymes used in this study were purchased from New England Biolabs. PCR assays were performed with primers listed in Table S1. PCR conditions were as follows: (i) 3 min at 94°C; (ii) 30 cycles of 30 s at 94°C, 30 s at the appropriate annealing temperature, and 1 minute/kb at 68°C; and (iii) 5 min at 68°C. When required, the resulting products were purified using the EZ-10 Spin Column PCR Products Purification Kit (Bio Basic) following the manufacturer's instructions. *E. coli* strains were transformed by electroporation as described previously [73] in a Bio-Rad Gene Pulser Xcell device set at 25 µF, 200 V and 1.8 kV using 1-mm gap electroporation cuvettes.

RNA extraction and cDNA synthesis

RNA extractions were performed as follows. *E. coli* GG56 containing pVCR94^{Kn} with or without p*acaCD* was grown at 37°C for 16 h in LB broth containing the appropriate antibiotics. The cultures were diluted 1:200 in fresh medium containing the appropriate antibiotics and grown to an OD_{600} of 0.2 before being diluted 1:10 again in fresh medium containing the appropriate antibiotics and supplemented with 0.02% arabinose when needed. After a 2-h incubation period,

4 ml of the cultures were used for total RNA extraction using the Direct-zol RNA extraction kit (Zymo Research) and TRI Reagent (Sigma-Aldrich) according to the manufacturer's instructions. Once purified, the RNA samples were treated using 2 units of DNase I (New England Biolabs) according to the manufacturer's instructions to eliminate residual gDNA. cDNA was synthesized from 0.5 ng of RNA and 1 pmol of gene-specific primer 85RT (Integrated DNA Technologies) using the reverse transcriptase SuperScript IV (Invitrogen), according to the manufacturer's instructions. Control reactions without reverse transcriptase treatment ('RNA') were performed for each sample. PCR reactions aiming at amplifying sfx_{94} and *acaD* were carried out using 50 pg of RNA or the corresponding amount of cDNA, or 0.5 ng of gDNA as the template, and primer pairs sfxF/sfxR and 086F2/086R2, respectively.

β-galactosidase assays

Quantitative β-galactosidase assays were performed with 2-nitrophenyl-β-D-galactopyranoside (ONPG) according to a protocol adapted from Miller [74]. After overnight incubation at 37°C with appropriate antibiotics, cultures were diluted 1:100 in 50ml LB broth supplemented with 50 µg/ml ampicillin and grown until an OD_{600} of 0.2 was reached. Three series of 1/10 dilutions were then prepared in total volumes of 5 ml LB broth supplemented with 50 µg/ml ampicillin with or without 0.2% arabinose and incubated for 2 hours at 37°C.

Statistical analyses

Numerical data presented in graphs are available in S1 File. All analyses were performed using R Statistical Software (v4.2.1) and RStudio (v2022.12.0.252) [75]. We assumed a normal distribution of the frequencies of transconjugant formation and tested the homogeneity of variance of the log10-transformed frequencies using Bartlett's tests. Dunnett and Tukey-Kramer tests and compact letter display were performed via the multcomp R package (v1.4-20) [76].

Details of statistical tests are provided in S2 File. Graphics were rendered via the ggplot2 R package (v3.3.6) [77]. All figures were prepared using Inkscape 0.92 (https://inkscape.org/).

Phylogenetic analyses

The primary sequence of Sfx_{94} and $TraN_{94}$ homologs were obtained using the NCBI blastp algorithm [78] against the nr/nt database restricted to Gammaproteobacteria (taxid: 1236). Primary sequences sharing less than 45% identity and under 85% minimum coverage were filtered out of subsequent analyses. The distantly related Sfx_{AQU1} from pAQU1 was added manually to the Sfx dataset as an outgroup. The datasets were clustered with CD-HIT [79] to the best cluster that met the 0.90 identity cut-off before alignment. Where applicable, representative sequences ("seeds") were manually substituted with sequences of interest within the same clusters for downstream analyses. The resulting datasets were then aligned with MUSCLE [80], and poorly aligned regions were discarded from the resulting amino acid alignments using TrimAl v1.3 software with the automated heuristic approach [81]. Evolutionary analyses were performed within MEGA11 (v 11.0.13) [82] using the Maximum Likelihood method (PhyML) [83] and the Le and Gascuel matrix-based model [84] with gamma distribution (LG + G). Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model and then selecting the topology with superior log likelihood value. Percent identity values were retrieved from the MUSCLE output [80].

Bioinformatic predictions

The prediction of signal peptides and cleavage sites in Sfx_{94} and $TraN_{94}$ was performed using SignalP 6.0 with the slow model mode against other organisms [85]. Terminators were predicted on both strands using ARNold [86,87]. *oriT* loci and mobilization proteins in putative IMEs were predicted using oriTfinder [88].

Data availability

Raw sequencing data of TraDIS experiments were submitted to Genbank Sequence Read Archive (SRA) under BioProject accession number PRJNA857050 with the following BioSample accession numbers: SAMN32907975 and SAMN32907976 for *E. coli* GG56 bearing pVCR94^{Kn} Δ*eexC*::*cat* as the recipient strain (input); SAMN32907978 and SAMN32907979 for *E. coli* GG56 transconjugants bearing pVCR94^{Kn} Δ*eexC*::*cat* and pSU4628 (output). Complete data from aligned reads can also be visualized using the UCSC genome browser [89] at https://genome.ucsc.edu/cgi-bin/hgHubConnect?hgHub_do_redirect=on&hgHubConnect.remakeTrackHub=on&hgHub_do_firstDb=1&hubUrl=https://datahub-101-cw2.p.genap.ca/Rivard_2023/pVCR94dX3deexC.hub.txt.

Acknowledgement

We are grateful to Emilie Boulerne, Colin Poirier, Simon Jeanneau, and Frédéric Grenier for technical assistance. We thank Dr. Alexander Kirschner, Medical University of Vienna, for the kind gift of *V. cholerae* A12JL5W90.

References

1. Harmer CJ, Hall RM. The A to Z of A/C plasmids. Plasmid. 2015;80: 63–82. doi:10.1016/j.plasmid.2015.04.003

2. Doublet B, Boyd D, Mulvey MR, Cloeckaert A. The *Salmonella* genomic island 1 is an integrative mobilizable element. Molecular Microbiology. 2005;55: 1911–1924. doi:10.1111/j.1365-2958.2005.04520.x

3. Douard G, Praud K, Cloeckaert A, Doublet B. The *Salmonella* Genomic Island 1 Is Specifically Mobilized *In Trans* by the IncA/C Multidrug Resistance Plasmid Family. PLoS One. 2010;5. doi:10.1371/journal.pone.0015302

4. Carraro N, Rivard N, Ceccarelli D, Colwell RR, Burrus V. IncA/C Conjugative Plasmids Mobilize a New Family of Multidrug Resistance Islands in Clinical *Vibrio cholerae* Non-O1/Non-O139 Isolates from Haiti. mBio. 2016;7: e00509-16. doi:10.1128/mBio.00509-16

5. Rivard N, Colwell RR, Burrus V. Antibiotic Resistance in *Vibrio cholerae*: Mechanistic Insights from IncC Plasmid-Mediated Dissemination of a Novel Family of Genomic Islands Inserted at *trmE*. mSphere. 2020;5. doi:10.1128/mSphere.00748-20

6. Siebor E, Neuwirth C. Overview of *Salmonella* Genomic Island 1-Related Elements Among Gamma-Proteobacteria Reveals Their Wide Distribution Among Environmental Species. Front Microbiol. 2022;13: 857492. doi:10.3389/fmicb.2022.857492

7. Carraro N, Matteau D, Luo P, Rodrigue S, Burrus V. The Master Activator of IncA/C Conjugative Plasmids Stimulates Genomic Islands and Multidrug Resistance Dissemination. PLoS Genet. 2014;10. doi:10.1371/journal.pgen.1004714

8. Poulin-Laprade D, Carraro N, Burrus V. The extended regulatory networks of SXT/R391 integrative and conjugative elements and IncA/C conjugative plasmids. Front Microbiol. 2015;6. doi:10.3389/fmicb.2015.00837

9. Humbert M, Huguet KT, Coulombe F, Burrus V. Entry Exclusion of Conjugative Plasmids of the IncA, IncC, and Related Untyped Incompatibility Groups. J Bacteriol. 2019;201: e00731-18. doi:10.1128/JB.00731-18

10. Huguet KT, Rivard N, Garneau D, Palanee J, Burrus V. Replication of the *Salmonella* Genomic Island 1 (SGI1) triggered by helper IncC conjugative plasmids promotes incompatibility and plasmid loss. PLoS Genet. 2020;16. doi:10.1371/journal.pgen.1008965

11. Carraro N, Durand R, Rivard N, Anquetil C, Barrette C, Humbert M, et al. *Salmonella* genomic island 1 (SGI1) reshapes the mating apparatus of IncC conjugative plasmids to promote self-propagation. PLoS Genet. 2017;13. doi:10.1371/journal.pgen.1006705

12. Novick RP. Plasmid incompatibility. Microbiol Rev. 1987;51: 381–395.

13. Garcillán-Barcia MP, de la Cruz F. Why is entry exclusion an essential feature of conjugative plasmids? Plasmid. 2008;60: 1–18. doi:10.1016/j.plasmid.2008.03.002

14. Achtman M, Kennedy N, Skurray R. Cell-cell interactions in conjugating *Escherichia coli*: role of *traT* protein in surface exclusion. Proc Natl Acad Sci U S A. 1977;74: 5104–5108.

15. Gunton JE, Ussher JER, Rooker MM, Wetsch NM, Alonso G, Taylor DE. Entry exclusion in the IncHI1 plasmid R27 is mediated by EexA and EexB. Plasmid. 2008;59: 86–101. doi:10.1016/j.plasmid.2007.11.004

16. Sakuma T, Tazumi S, Furuya N, Komano T. ExcA proteins of IncI1 plasmid R64 and IncIγ plasmid R621a recognize different segments of their cognate TraY proteins in entry exclusion. Plasmid. 2013;69: 138–145. doi:10.1016/j.plasmid.2012.11.004

17. Pohlman RF, Genetti HD, Winans SC. Entry Exclusion of the IncN Plasmid pKM101 Is Mediated by a Single Hydrophilic Protein Containing a Lipid Attachment Motif. Plasmid. 1994;31: 158–165. doi:10.1006/plas.1994.1017

18. Haase J, Kalkum M, Lanka E. TrbK, a small cytoplasmic membrane lipoprotein, functions in entry exclusion of the IncP alpha plasmid RP4. J Bacteriol. 1996;178: 6720–6729.

19. Bolland S, Llosa M, Avila P, de la Cruz F. General organization of the conjugal transfer genes of the IncW plasmid R388 and interactions between R388 and IncN and IncP plasmids. J Bacteriol. 1990;172: 5795–5802.

20. Yamada Y, Yamada M, Nakazawa A. A ColE1-encoded gene directs entry exclusion of the plasmid. J Bacteriol. 1995;177: 6064–6068.

21. Kamruzzaman M, Mathers AJ, Iredell JR. A Novel Plasmid Entry Exclusion System in pKPC_UVA01, a Promiscuous Conjugative Plasmid Carrying the bla_{KPC} Carbapenemase Gene. Antimicrob Agents Chemother. 66: e02322-21. doi:10.1128/aac.02322-21

22. Schmitt A, Hirt H, Järvå MA, Sun W-S, ter Beek J, Dunny GM, et al. Enterococcal PrgA Extends Far Outside the Cell and Provides Surface Exclusion to Protect against Unwanted Conjugation. Journal of Molecular Biology. 2020;432: 5681–5695. doi:10.1016/j.jmb.2020.08.018

23. Marrero J, Waldor MK. The SXT/R391 Family of Integrative Conjugative Elements Is Composed of Two Exclusion Groups. J Bacteriol. 2007;189: 3302–3305. doi:10.1128/JB.01902-06

24. Avello M, Davis KP, Grossman AD. Identification, characterization, and benefits of an exclusion system in an integrative and conjugative element of *Bacillus subtilis*. Mol Microbiol. 2019;112: 1066–1082. doi:10.1111/mmi.14359

25. Meynell E, Ewins A. Effect on Exclusion of Alterations to the Sex Pilus. J Bacteriol. 1973;113: 71–75.

26. Marrero J, Waldor MK. Determinants of Entry Exclusion within Eex and TraG Are Cytoplasmic. J Bacteriol. 2007;189: 6469–6473. doi:10.1128/JB.00522-07

27. Harrison JL, Taylor IM, Platt K, O'Connor CD. Surface exclusion specificity of the TraT lipoprotein is determined by single alterations in a five-amino-acid region of the protein. Molecular Microbiology. 1992;6: 2825–2832. doi:10.1111/j.1365-2958.1992.tb01462.x

28. Anthony KG, Sherburne C, Sherburne R, Frost LS. The role of the pilus in recipient cell recognition during bacterial conjugation mediated by F-like plasmids. Molecular Microbiology. 1994;13: 939–953. doi:10.1111/j.1365-2958.1994.tb00486.x

29. Riede I, Eschbach M-L. Evidence that TraT interacts with OmpA of *Escherichia coli*. FEBS Letters. 1986;205: 241–245. doi:10.1016/0014-5793(86)80905-X

30. Klimke WA, Frost LS. Genetic Analysis of the Role of the Transfer Gene, *traN*, of the F and R100-1 Plasmids in Mating Pair Stabilization during Conjugation. J Bacteriol. 1998;180: 4036–4043.

31. Klimke WA, Rypien CD, Klinger B, Kennedy RA, Rodriguez-Maillard JM, Frost LS. The mating pair stabilization protein, TraN, of the F plasmid is an outer-membrane protein with two regions that are important for its function in conjugation. Microbiology. 2005;151: 3527–3540. doi:10.1099/mic.0.28025-0

32. Low WW, Wong JLC, Beltran LC, Seddon C, David S, Kwong H-S, et al. Mating pair stabilization mediates bacterial conjugation species specificity. Nat Microbiol. 2022;7: 1016–1027. doi:10.1038/s41564-022-01146-4

33. Low WW, Seddon C, Beis K, Frankel G. The Interaction of the F-Like Plasmid-Encoded TraN Isoforms with Their Cognate Outer Membrane Receptors. Journal of Bacteriology. 2023;205: e00061-23. doi:10.1128/jb.00061-23

34. Ambrose SJ, Harmer CJ, Hall RM. Compatibility and entry exclusion of IncA and IncC plasmids revisited: IncA and IncC plasmids are compatible. Plasmid. 2018. doi:10.1016/j.plasmid.2018.02.002

35. Nivaskumar M, Francetic O. Type II secretion system: A magic beanstalk or a protein escalator. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research. 2014;1843: 1568–1577. doi:10.1016/j.bbamcr.2013.12.020

36. Rocque WJ, McGroarty EJ. Isolation and preliminary characterization of wild-type OmpC porin dimers from *Escherichia coli* K-12. Biochemistry. 1989;28: 3738–3743. doi:10.1021/bi00435a017

37. Hancock SJ, Phan M-D, Luo Z, Lo AW, Peters KM, Nhu NTK, et al. Comprehensive analysis of IncC plasmid conjugation identifies a crucial role for the transcriptional regulator AcaB. Nature Microbiology. 2020; 1–9. doi:10.1038/s41564-020-0775-0

38. Daccord A, Mursell M, Poulin-Laprade D, Burrus V. Dynamics of the SetCD-Regulated Integration and Excision of Genomic Islands Mobilized by Integrating Conjugative Elements of the SXT/R391 Family. J Bacteriol. 2012;194: 5794–5802. doi:10.1128/JB.01093-12

39. Lawley TD, Klimke WA, Gubbins MJ, Frost LS. F factor conjugation is a true type IV secretion system. FEMS Microbiology Letters. 2003;224: 1–15. doi:10.1016/S0378-1097(03)00430-0

40. Nonaka L, Maruyama F, Miyamoto M, Miyakoshi M, Kurokawa K, Masuda M. Novel Conjugative Transferable Multiple Drug Resistance Plasmid pAQU1 from *Photobacterium damselae* subsp. *damselae* Isolated from Marine Aquaculture Environment. Microbes Environ. 2012;27: 263–272. doi:10.1264/jsme2.ME11338

41. Tanaka KH, Vincent AT, Trudel MV, Paquet VE, Frenette M, Charette SJ. The mosaic architecture of *Aeromonas salmonicida* subsp. *salmonicida* pAsa4 plasmid and its consequences on antibiotic resistance. PeerJ. 2016;4: e2595. doi:10.7717/peerj.2595

42. Zhu L, Zheng J-S, Wang W-M, Luo Y. Complete Genome Sequence of Highly Virulent *Aeromonas hydrophila* Strain D4, Isolated from a Diseased Blunt-Snout Bream in China. Microbiol Resour Announc. 2019;8: e01035-18. doi:10.1128/MRA.01035-18

43. Achtman M, Manning PA, Kusecek B, Schwuchow S, Neil W. Genetic analysis of F sex factor cistrons needed for surface exclusion in *Escherichia coli*. Journal of Molecular Biology. 1980;138: 779–795. doi:10.1016/0022-2836(80)90065-0

44. Yoshioka Y, Ohtsubo H, Ohtsubo E. Repressor gene *finO* in plasmids R100 and F: constitutive transfer of plasmid F is caused by insertion of IS*3* into F *finO*. J Bacteriol. 1987;169: 619–623. doi:10.1128/jb.169.2.619-623.1987

45. Perumal NB, Minkley EG. The product of the F sex factor *traT* surface exclusion gene is a lipoprotein. Journal of Biological Chemistry. 1984;259: 5357–5360. doi:10.1016/S0021-9258(18)91011-2

46. Inselburg J. Incompatibility Exhibited by Colicin Plasmids E1, E2, and E3 in *Escherichia coli*. J Bacteriol. 1974;119: 478–483.

47. Chan PT, Ohmori H, Tomizawa J, Lebowitz J. Nucleotide sequence and gene organization of ColE1 DNA. Journal of Biological Chemistry. 1985;260: 8925–8935. doi:10.1016/S0021-9258(17)39438-3

48. Friedman DI, Olson ER, Georgopoulos C, Tilly K, Herskowitz I, Banuett F. Interactions of bacteriophage and host macromolecules in the growth of bacteriophage lambda. Microbiol Rev. 1984;48: 299–325.

49. Garcillán-Barcia MP, Francia MV, De La Cruz F. The diversity of conjugative relaxases and its application in plasmid classification. FEMS Microbiology Reviews. 2009;33: 657–687. doi:10.1111/j.1574-6976.2009.00168.x

50. Szabó M, Murányi G, Kiss J. IncC helper dependent plasmid-like replication of *Salmonella* Genomic Island 1. Nucleic Acids Res. 2021;49: 832–846. doi:10.1093/nar/gkaa1257

51. Kiss J, Szabó M, Hegyi A, Douard G, Praud K, Nagy I, et al. Identification and Characterization of *oriT* and Two Mobilization Genes Required for Conjugative Transfer of *Salmonella* Genomic Island 1. Front Microbiol. 2019;10. doi:10.3389/fmicb.2019.00457

52. Hancock SJ, Phan M-D, Roberts LW, Vu TNM, Harris PNA, Beatson SA, et al. Characterization of DtrJ as an IncC plasmid conjugative DNA transfer component. Molecular Microbiology. 2021;116: 154–167. doi:https://doi.org/10.1111/mmi.14697

53. Hedges RW, Datta N. Plasmids Determining I Pili Constitute a Compatibility Complex. Microbiology. 1973;77: 19–25. doi:10.1099/00221287-77-1-19

54. Willetts N, Maule J. Interactions between the surface exclusion systems of some F-like plasmids. Genetics Research. 1974;24: 81–89. doi:10.1017/S0016672300015093

55. Willetts N, Maule J. Specificities of IncF plasmid conjugation genes. Genet Res. 1986;47: 1–11. doi:10.1017/s0016672300024447

56. Woodcock DM, Crowther PJ, Doherty J, Jefferson S, DeCruz E, Noyer-Weidner M, et al. Quantitative evaluation of *Escherichia coli* host strains for tolerance to cytosine methylation in plasmid and phage recombinants. Nucl Acids Res. 1989;17: 3469–3478. doi:10.1093/nar/17.9.3469

57. Grenier F, Matteau D, Baby V, Rodrigue S. Complete Genome Sequence of *Escherichia coli* BW25113. Genome Announc. 2014;2: e01038-14. doi:10.1128/genomeA.01038-14

58. Ferrières L, Hémery G, Nham T, Guérout A-M, Mazel D, Beloin C, et al. Silent mischief: bacteriophage Mu insertions contaminate products of *Escherichia coli* random mutagenesis performed using suicidal transposon delivery plasmids mobilized by broad-host-range RP4 conjugative machinery. J Bacteriol. 2010;192: 6418–6427. doi:10.1128/JB.00621-10

59. Ceccarelli D, Daccord A, René M, Burrus V. Identification of the Origin of Transfer (*oriT*) and a New Gene Required for Mobilization of the SXT/R391 Family of Integrating Conjugative Elements. J Bacteriol. 2008;190: 5328–5338. doi:10.1128/JB.00150-08

60. Singer M, Baker TA, Schnitzler G, Deischel SM, Goel M, Dove W, et al. A collection of strains containing genetically linked alternating antibiotic resistance elements for genetic mapping of *Escherichia coli*. Microbiol Rev. 1989;53: 1–24.

61. Lepuschitz S, Baron S, Larvor E, Granier SA, Pretzer C, Mach RL, et al. Phenotypic and Genotypic Antimicrobial Resistance Traits of *Vibrio cholerae* Non-O1/Non-O139 Isolated From a Large Austrian Lake Frequently Associated With Cases of Human Infection. Front Microbiol. 2019;10: 2600. doi:10.3389/fmicb.2019.02600

62. Carraro N, Sauvé M, Matteau D, Lauzon G, Rodrigue S, Burrus V. Development of pVCR94ΔX from *Vibrio cholerae*, a prototype for studying multidrug resistant IncA/C conjugative plasmids. Front Microbiol. 2014;5. doi:10.3389/fmicb.2014.00044

63. Cabezón E, Sastre JI, de la Cruz F. Genetic evidence of a coupling role for the TraG protein family in bacterial conjugation. Mol Gen Genet. 1997;254: 400–406.

64. Haldimann A, Wanner BL. Conditional-Replication, Integration, Excision, and Retrieval Plasmid-Host Systems for Gene Structure-Function Studies of Bacteria. J Bacteriol. 2001;183: 6384–6393. doi:10.1128/JB.183.21.6384-6393.2001

65. Guzman LM, Belin D, Carson MJ, Beckwith J. Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter. Journal of Bacteriology. 1995;177: 4121–4130. doi:10.1128/jb.177.14.4121-4130.1995

66. Garriss G, Waldor MK, Burrus V. Mobile Antibiotic Resistance Encoding Elements Promote Their Own Diversity. PLoS Genet. 2009;5. doi:10.1371/journal.pgen.1000775

67. Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci U S A. 2000;97: 6640–6645.

68. Cherepanov PP, Wackernagel W. Gene disruption in *Escherichia coli*: Tc^R and Km^R cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. Gene. 1995;158: 9–14. doi:10.1016/0378-1119(95)00193-A

69. Datta S, Costantino N, Court DL. A set of recombineering plasmids for gram-negative bacteria. Gene. 2006;379: 109–115. doi:10.1016/j.gene.2006.04.018

70. Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, et al. Current Protocols in Molecular Biology. Hoboken: John Wiley & Sons; 2003.

71. Roy D, Huguet KT, Grenier F, Burrus V. IncC conjugative plasmids and SXT/R391 elements repair double-strand breaks caused by CRISPR-Cas during conjugation. Nucleic Acids Res. 2020;48: 8815–8827. doi:10.1093/nar/gkaa518

72. Sandrini S, Haigh R, Freestone P. Fractionation by Ultracentrifugation of Gram Negative Cytoplasmic and Membrane Proteins. BIO-PROTOCOL. 2014;4. doi:10.21769/BioProtoc.1287

73. Dower WJ, Miller JF, Ragsdale CW. High efficiency transformation of *E. coli* by high voltage electroporation. Nucleic Acids Res. 1988;16: 6127–6145. doi:10.1093/nar/16.13.6127

74. Miller JF. A short course in bacterial genetics. Plainview, NY: Cold Spring Harbor Laboratory Press; 1992.

75. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2021. Available: https://www.R-project.org/

76. Hothorn T, Bretz F, Westfall P. Simultaneous inference in general parametric models. Biom J. 2008;50: 346–363. doi:10.1002/bimj.200810425

77. Wickham H. ggplot2: Elegant Graphics for Data Analysis. New York, NY: Springer-Verlag; 2016.

78. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215: 403–410. doi:10.1016/S0022-2836(05)80360-2

79. Li W, Godzik A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics. 2006;22: 1658–1659. doi:10.1093/bioinformatics/btl158

80. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32: 1792–1797. doi:10.1093/nar/gkh340

81. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009;25: 1972–1973. doi:10.1093/bioinformatics/btp348

82. Tamura K, Stecher G, Kumar S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. Mol Biol Evol. 2021;38: 3022–3027. doi:10.1093/molbev/msab120

83. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol. 2010;59: 307–321. doi:10.1093/sysbio/syq010

84. Le SQ, Gascuel O. An Improved General Amino Acid Replacement Matrix. Molecular Biology and Evolution. 2008;25: 1307–1320. doi:10.1093/molbev/msn067

85. Teufel F, Almagro Armenteros JJ, Johansen AR, Gíslason MH, Pihl SI, Tsirigos KD, et al. SignalP 6.0 predicts all five types of signal peptides using protein language models. Nat Biotechnol. 2022;40: 1023–1025. doi:10.1038/s41587-021-01156-3

86. Macke TJ, Ecker DJ, Gutell RR, Gautheret D, Case DA, Sampath R. RNAMotif, an RNA secondary structure definition and search algorithm. Nucleic Acids Research. 2001;29: 4724–4735. doi:10.1093/nar/29.22.4724

87. Gautheret D, Lambert A. Direct RNA motif definition and identification from multiple sequence alignments using secondary structure profiles. Journal of Molecular Biology. 2001;313: 1003–1011. doi:10.1006/jmbi.2001.5102

88. Li X, Xie Y, Liu M, Tai C, Sun J, Deng Z, et al. oriTfinder: a web-based tool for the identification of origin of transfers in DNA sequences of bacterial mobile genetic elements. Nucleic Acids Res. 2018;46: W229–W234. doi:10.1093/nar/gky352

89. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The Human Genome Browser at UCSC. Genome Res. 2002;12: 996–1006. doi:10.1101/gr.229102

Supporting information

S1 Fig. Identification of an additional IncC exclusion factor by TraDIS. The tracks plot the number of reads from two independent replicates as a function of position in pVCR94^{Kn} Δ*eexC*::*cat* for both the input and output libraries. ORFs with similar functions are colour-coded as indicated in the figure. This figure was created using the UCSC Genome Browser (http://genome.ucsc.edu).

S2 Fig. Impact of exclusion on the fitness of IncC^{+} cells. Colony forming units (CFUs) were obtained from mating assays conducted in Fig 1D. Crossbars show the mean and standard error of the mean of three independent experiments. One-way ANOVA (p=3.9e-6) with a Tukey-Kramer post-test was used on the $\log_{10}$-transformed values to compare the means. Statistical significance (S2 File) is shown as a compact letter display for pairwise comparisons where means grouped under identical letters are not statistically different.

S3 Fig. Sfx is a predicted lipoprotein. (A) Schematic representation of the predicted translation products of sfx_{94} and $traN_{94}$ and their predicted signal peptide sequences. (B) Predicted Rho-independent terminators in the proximity of *sfx* in the IncC plasmid pVCR94. Terminator sequences are provided with the stems and loops in blue and red, respectively. Their position is represented by hairpins on the schematic representation of the region. Predicted ΔG values are indicated in kcal/mol.

S4 Fig. Sequence identity heatmap showing pairwise percentage identity between 18 selected Sfx homologs. Panels are shaded as indicated in the Figure according to the identity calculated for each pair.

S5 Fig. TraN directs surface exclusion specificity. *E. coli* GG56 (Nx^{r}) containing pVCR94^{Sp} Δ*traN* and expressing the indicated *traN* variants from pBAD30 (Ap^{r}) served as the donor, and *E. coli* CAG18439 (Tc^{r}) expressing the indicated *sfx* variants from a single-copy, chromosomally

integrated pAH56 (Kn^r) was used as the recipient. Transconjugants containing pVCR94^{Sp} Δ*traN* were selected as the Tc^r Sp^r colonies. Crossbars show the mean and standard error of the mean of three independent experiments. One-way ANOVA (*traN* pVCR94, p=2.21e-8; *traN* pAsa4c, p=4.84e-10; *traN* pAhD4-1, p=2.20e-8; *traN* pAQU1, p=1.81e-9) with a Tukey-Kramer post-test was used on the log_{10}-transformed values to compare the means. Exclusion indices, shown at the bottom of each crossbar, are calculated by dividing the frequency of transfer of pVCR94^{Sp} Δ*traN* to the empty recipient divided by the transfer frequency to the recipient expressing the indicated *sfx* variant. Statistical significance (S2 File) is shown as a compact letter display for pairwise comparisons where means grouped under identical letters are not statistically different.

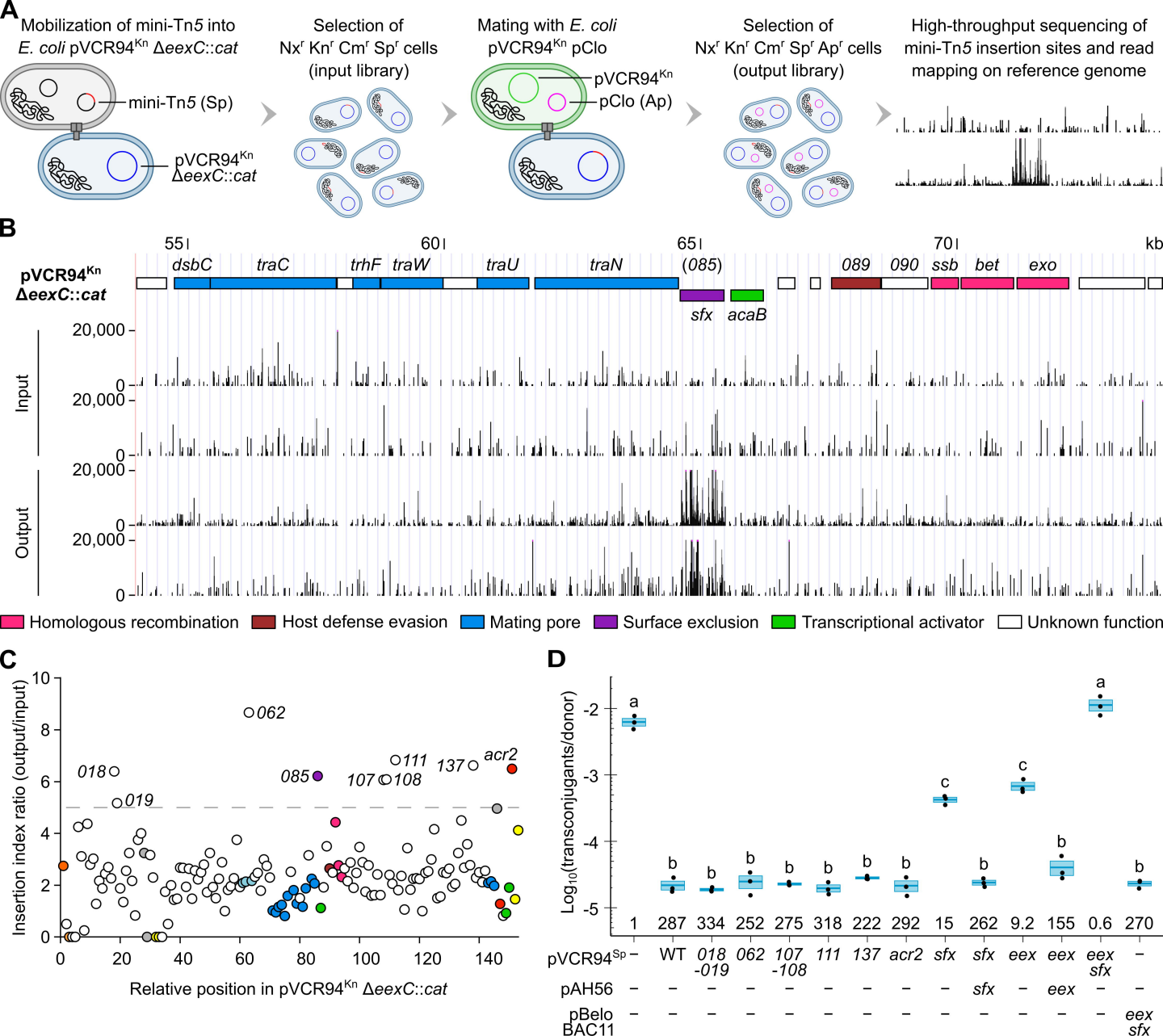
S1 Table. Primers used in this study.

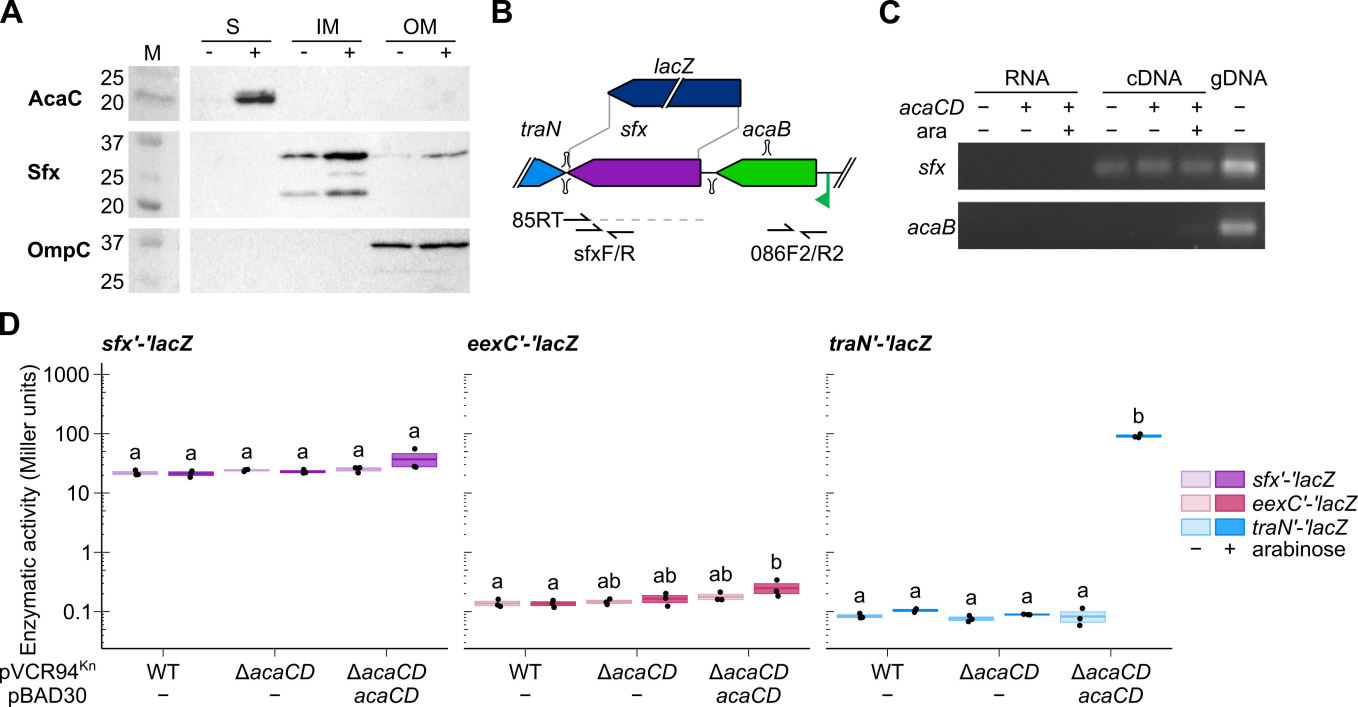
S2 Table. Predicted AcaCD binding sites.

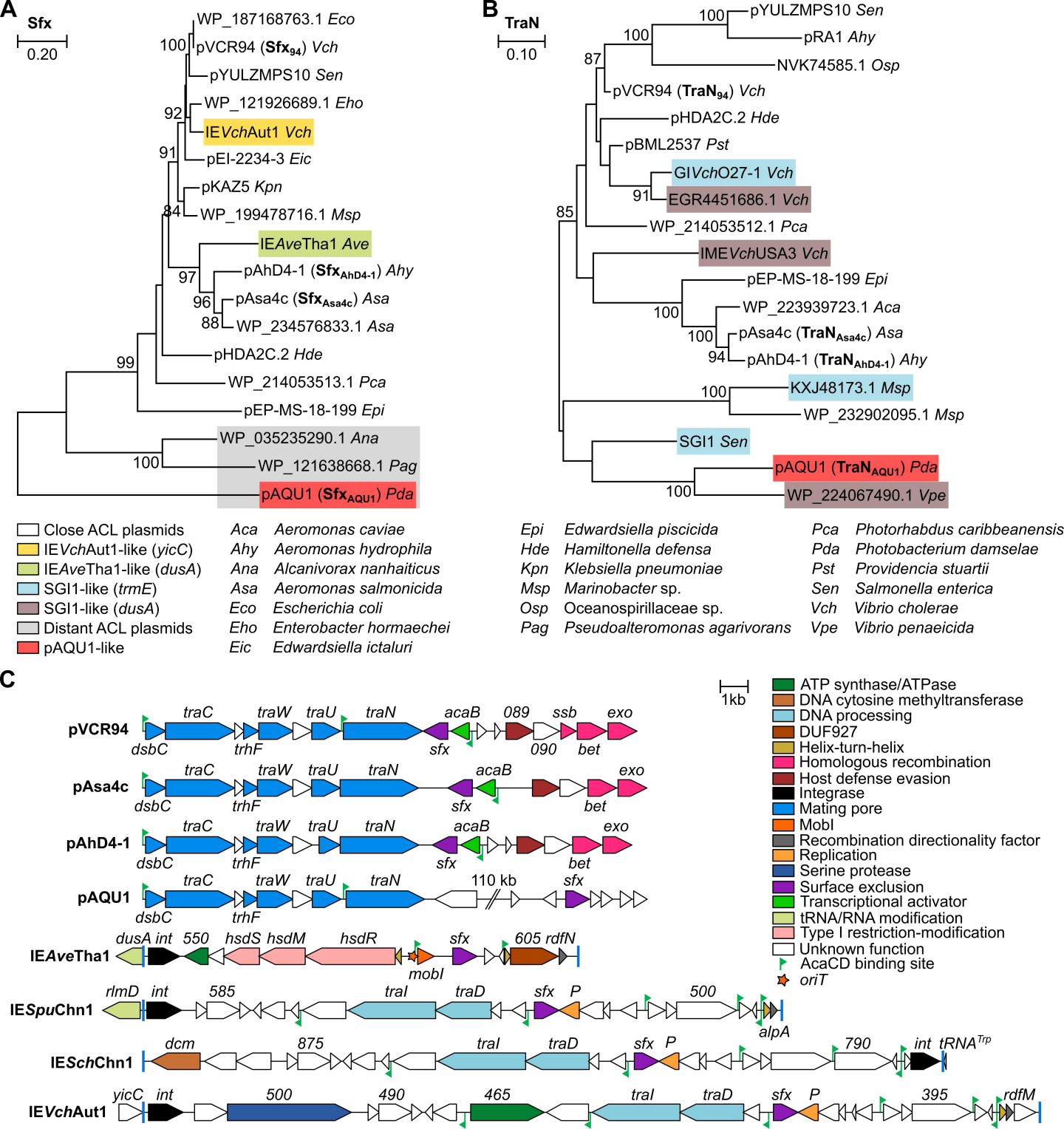
S1 File. Complete data sets.

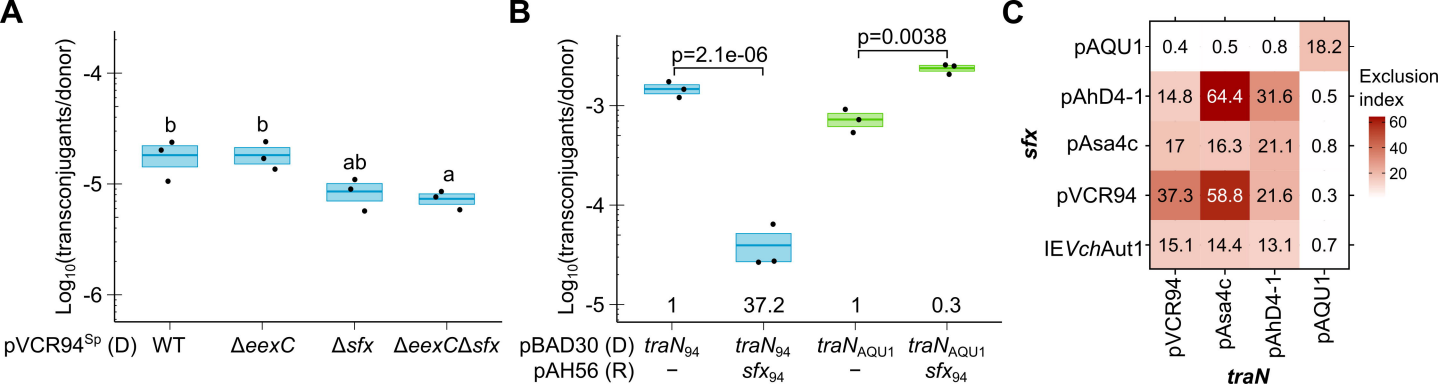
S2 File. Statistical analyses.

S3 File. Sfx and TraN homologs.


A
Mobilization of mini-Tn5 into E. coli pVCR94Kn ΔeexC::cat
mini-Tn5 (Sp)
pVCR94Kn ΔeexC::cat
Selection of Nxr Knr Cmr Spr cells (input library)
Mating with E. coli pVCR94Kn pClo
pVCR94Kn
pClo (Ap)
Selection of Nxr Knr Cmr Spr Apr cells (output library)
High-throughput sequencing of mini-Tn5 insertion sites and read mapping on reference genome
B
pVCR94Kn ΔeexC::cat
55
60
65
70
kb
dsbC
traC
trhF
traW
traU
traN
(085)
089
090
ssb
bet
exo
sfx
acaB
Input
Output
20,000
0
Homologous recombination
Host defense evasion
Mating pore
Surface exclusion
Transcriptional activator
Unknown function
C
Insertion index ratio (output/input)
Relative position in pVCR94Kn ΔeexC::cat
062
018
019
085
107
108
111
137
acr2
D
Log10(transconjugants/donor)
1
287
334
252
275
318
222
292
15
262
9.2
155
0.6
270
pVCR94Sp
WT
018 -019
062
107 -108
111
137
acr2
sfx
sfx
eex
eex
eex sfx
pAH56
sfx
eex
pBelo BAC11
eex sfx


A
M
S
IM
OM
-
+
AcaC
Sfx
OmpC
25
20
37
B
lacZ
traN
sfx
acaB
85RT
sfxF/R
086F2/R2
C
RNA
cDNA
gDNA
acaCD
ara
sfx
acaB
D
sfx'-'lacZ
eexC'-'lacZ
traN'-'lacZ
Enzymatic activity (Miller units)
1000
100
10
1
0.1
a
ab
b
pVCR94Kn
pBAD30
WT
ΔacaCD
ΔacaCD acaCD
sfx'-'lacZ
eexC'-'lacZ
traN'-'lacZ
arabinose


A
Sfx
0.20
WP_187168763.1 Eco
pVCR94 (Sfx94) Vch
pYULZMPS10 Sen
WP_121926689.1 Eho
IEVchAut1 Vch
pEI-2234-3 Eic
pKAZ5 Kpn
WP_199478716.1 Msp
IEAveTha1 Ave
pAhD4-1 (SfxAhD4-1) Ahy
pAsa4c (SfxAsa4c) Asa
WP_234576833.1 Asa
pHDA2C.2 Hde
WP_214053513.1 Pca
pEP-MS-18-199 Epi
WP_035235290.1 Ana
WP_121638668.1 Pag
pAQU1 (SfxAQU1) Pda
B
TraN
0.10
pYULZMPS10 Sen
pRA1 Ahy
NVK74585.1 Osp
pVCR94 (TraN94) Vch
pHDA2C.2 Hde
pBML2537 Pst
GIVchO27-1 Vch
EGR4451686.1 Vch
WP_214053512.1 Pca
IMEVchUSA3 Vch
pEP-MS-18-199 Epi
WP_223939723.1 Aca
pAsa4c (TraNAsa4c) Asa
pAhD4-1 (TraNAhD4-1) Ahy
KXJ48173.1 Msp
WP_232902095.1 Msp
SGI1 Sen
pAQU1 (TraNAQU1) Pda
WP_224067490.1 Vpe
Close ACL plasmids
IEVchAut1-like (yicC)
IEAveTha1-like (dusA)
SGI1-like (trmE)
SGI1-like (dusA)
Distant ACL plasmids
pAQU1-like
Aca Aeromonas caviae
Ahy Aeromonas hydrophila
Ana Alcanivorax nanhaiticus
Asa Aeromonas salmonicida
Eco Escherichia coli
Eho Enterobacter hormaechei
Eic Edwardsiella ictaluri
Epi Edwardsiella piscicida
Hde Hamiltonella defensa
Kpn Klebsiella pneumoniae
Msp Marinobacter sp.
Osp Oceanospirillaceae sp.
Pag Pseudoalteromonas agarivorans
Pca Photorhabdus caribbeanensis
Pda Photobacterium damselae
Pst Providencia stuartii
Sen Salmonella enterica
Vch Vibrio cholerae
Vpe Vibrio penaeicida
C
1kb
ATP synthase/ATPase
DNA cytosine methyltransferase
DNA processing
DUF927
Helix-turn-helix
Homologous recombination
Host defense evasion
Integrase
Mating pore
MobI
Recombination directionality factor
Replication
Serine protease
Surface exclusion
Transcriptional activator
tRNA/RNA modification
Type I restriction-modification
Unknown function
AcaCD binding site
oriT
pVCR94
pAsa4c
pAhD4-1
pAQU1
IEAveTha1
IESpuChn1
IESchChn1
IEVchAut1
traC traW traU traN acaB 089 ssb exo
dsbC trhF sfx 090 bet
110 kb
dusA int 550 hsdS hsdM hsdR sfx 605 rdfN
mobI
rlmD int 585 traI traD sfx P 500
alpA
dcm 875 traI traD sfx P 790 int tRNATrp
yicC int 500 490 465 traI traD sfx P 395 rdfM


A
Log10(transconjugants/donor)
-4
-5
-6
b
b
ab
a
pVCR94Sp (D) WT ΔeexC Δsfx ΔeexCΔsfx
B
Log10(transconjugants/donor)
-3
-4
-5
p=2.1e-06
p=0.0038
1
37.2
1
0.3
pBAD30 (D) traN94 traN94 traNAQU1 traNAQU1
pAH56 (R) – sfx94 – sfx94
C
pAQU1 0.4 0.5 0.8 18.2
pAhD4-1 14.8 64.4 31.6 0.5
pAsa4c 17 16.3 21.1 0.8
pVCR94 37.3 58.8 21.6 0.3
IEVchAut1 15.1 14.4 13.1 0.7
sfx
pVCR94 pAsa4c pAhD4-1 pAQU1
traN
Exclusion index
60
40
20