Crucial role of *Salmonella* genomic island 1 master activator in the parasitism of IncC plasmids

Romain Durand *, Kévin T. Huguet †, Nicolas Rivard †, Nicolas Carraro *, Sébastien Rodrigue * and Vincent Burrus *

Département de biologie, Faculté des sciences, Université de Sherbrooke, Sherbrooke, QC J1K 2R1, Canada

Received August 28, 2020; Revised March 09, 2021; Editorial Decision March 10, 2021; Accepted March 15, 2021

**ABSTRACT**

IncC conjugative plasmids and the multiple variants of *Salmonella* Genomic Island 1 (SGI1) are two functionally interacting families of mobile genetic elements commonly associated with multidrug resistance in the Gammaproteobacteria. SGI1 and its siblings are specifically mobilised in trans by IncC conjugative plasmids. Conjugative transfer of IncC plasmids is activated by the plasmid-encoded master activator AcaCD. SGI1 carries five AcaCD-responsive promoters that drive the expression of genes involved in its excision, replication, and mobilisation. SGI1 encodes an AcaCD homologue, the transcriptional activator complex SgaCD (also known as FlhDCSGI1) that seems to recognise and activate the same SGI1 promoters. Here, we investigated the relevance of SgaCD in SGI1’s lifecycle. Mating assays revealed the requirement for SgaCD and its IncC-encoded counterpart AcaCD in the mobilisation of SGI1. An integrative approach combining ChIP-exo, Cappable-seq, and RNA-seq confirmed that SgaCD activates each of the 18 AcaCD-responsive promoters driving the expression of the plasmid transfer functions. A comprehensive analysis of the activity of the complete set of AcaCD-responsive promoters of SGI1 and the helper IncC plasmid was performed through reporter assays. qPCR and flow cytometry assays revealed that SgaCD is essential to elicit the excision and replication of SGI1 and destabilise the helper IncC plasmid.

**GRAPHICAL ABSTRACT**

**INTRODUCTION**

Multidrug-resistant bacteria are an economic burden and a global threat to public health (1). Their emergence is being fuelled by diverse mobile genetic elements such as genomic islands and conjugative plasmids (2) that often carry many antibiotic resistance genes. A better understanding of the mechanisms promoting the dissemination of mobile genetic elements is thus urgently needed. Mobilisable Genomic Islands (MGIs), also referred to as Integrated Mobilisable Elements (IMEs), have recently been receiving renewed attention as they are increasingly recognised as key contributors to the propagation of multidrug resistance (3–7). MGIs usually carry diverse gene cargos involved in antibiotic or heavy metal resistance, bacteriocin synthesis, or resistance to phage infection (7–9). *Salmonella* Genomic Is-
land 1 (SGI1) is a 42.4-kb MGI that confers resistance to ampicillin, chloramphenicol, streptomycin, sulfonamides and tetracycline (ACSSuT) (10). Since the initial report of SGI1 in *Salmonella enterica* serovar Typhimurium DT104 two decades ago, a multitude of SGI1 variants has been described in several species of *Gammaproteobacteria*, including human pathogens such as *Salmonella enterica* serovars, *Proteus mirabilis*, *Morganella morganii*, *Providencia stuartii* or *Klebsiella pneumoniae* (11–15). SGI1 and its variants share a conserved set of genes (described below) and often bear a class 1 integron with diverse combinations of antibiotic resistance gene cassettes (16). SGI1 and its relatives are found integrated at the 3′ end of *trmE* (also known as *mmnE* or *thdF*) in the chromosome of their respective hosts and are specifically mobilised in trans by the IncC and closely related IncA conjugative plasmids (17). IncC plasmids are large, broad-host-range, and globally distributed plasmids that contribute to the propagation of multidrug resistance genes. For instance, IncC plasmids are frequently associated with New Delhi metallo-β-lactamase genes (bla*NDM*), that confer resistance against most β-lactams, including carbapenems (18). IncC plasmids have also been key players in the local acquisition of antibiotic resistance genes in African phyletic sub-lineages of the seventh pandemic of cholera caused by *Vibrio cholerae* O1 El Tor (19). Despite a lower epidemiological success compared to IncC plasmids, IncA plasmids have been found responsible for the spread of the carbapenem gene *bla*NDM* in several species of *Enterobacteriales* in hospitalised patients in Italy (20).

Conjugative transfer of IncC plasmids is controlled by two loci, the acr1-acrDC-acr2 region, and the *acab* gene located near *traNc*. *acabDC* encodes the heteromeric complex AcaCD distinctly related to SetCD, the transcriptional activator of transfer genes of the integrative and conjugative elements (ICEs) of the SXT/R391 family (21,22). AcaC and SetCD are distinctly related to FlhCD, the transcriptional activator of flagellar operons in *Escherichia coli* and *Salmonella* (21,23,24). Expression of *acabDC* is driven from the promoter *P* *mexl* that is repressed by Acr1 and Acr2 (21,25). Recently, Hancock *et al.* demonstrated that *P* *mexl* is activated by AcaB, a novel transcriptional activator that exhibits structural similarity to bacterial transcription factors from the ribbon-helix-helix (RHH) superfamily (26). Interestingly, *acab* is part of the AcaCD regulon, wherein AcaC and AcaB promote mutual activation, generating a positive feedback loop that activates conjugation. AcaCD turns on a total of 18 AcaCD-activatable promoters driving the expression of major operons involved in the formation of the mating pore and initiation of conjugal transfer, as well as genes of unknown function (21,27). Furthermore, AcaCD also activates the expression of genes carried by distinct families of MGIs, including five operons in SGI1 (7,27–29). The functions of five AcaCD-activatable genes of SGI1, *xis*, *rep*, *traNc*, *traHc* and *tragS*, have been characterised. *xis* encodes the recombination directionality factor that facilitates the excision of SGI1 from the chromosome catalysed by Int (30). *rep* encodes the replication initiator protein that initiates SGI1 replication at the origin of replication (*oriV*) (31). *rep* contains a RepA_C domain (Pfam PF04796) and is distantly related to RepA of IncN2 plasmids (31,32). *traNc*, *traHc* and *tragS* encode three type IV secretion system (T4SS) subunits that replace their respective counterparts TraNC, TraHC and TraGC in the mating pore encoded by the helper IncC plasmid (33). MGIs are usually thought of as relatively passive mobile genetic elements that await the arrival of a helper self-transmissible element (a conjugative plasmid or an ICE) to escape their quiescent state and ride along with their helper element. To transfer to a new host, MGIs first need to excise from the host’s chromosome, and next to be translocated through the mating pore encoded by their helper element. For both processes, IncC-mobilised MGIs take advantage of the AcaCD regulon (7,21,28). Surprisingly, SGI1 has regularly challenged the presumed passivity of MGIs (3). SGI1 has notably been shown to actively reshape the mating pore encoded by IncC conjugative plasmids to enhance its propagation (33). SGI1 also actively replicates and carries a functional toxin-antitoxin system (*sgaAT*) that enhances its stability when an IncC plasmid is concomitantly present in the host (31,34). Besides, SGI1 destabilises the helper IncA and IncC plasmids (35), a mechanism only recently attributed to the activation of the replicative state of excised SGI1 (31). Remarkably, SGI1-K, an SGI1 variant found in the worldwide *Salmonella* Kentucky clone ST198 (sequence type 198) and *Salmonella* Newport, is unable to destabilise IncC plasmids (35–40). SGI1-K lacks the 3′ half of *traNc* (*traS*) and two upstream genes, *sgaD* (*SGI1*706) and *sgaC* (*SGI1*707), that code for an AcaCD ortholog complex named SgaCD (also known as FlhCDSGI1) (29,41,42). SgaCD contains 79 and 46% identity with AcaC and AcaD, respectively. Although SgaCD was reported to activate all five AcaCD-activatable promoters of SGI1 and complement the deletion of *acabDC* in the IncC plasmid R16a, its biological role remains unclear (29). For instance, deletion of *sgaDC* had no impact on the mobilisation of SGI1-C by the helper IncC plasmid R55 (28). The variant SGI1-C differs from SGI1 by its smaller integron conferring resistance to spectinomycin, streptomycin and sulfonamides (43). Furthermore, deletion of *acabDC* of the helper IncC plasmid R16a abolished both self-transfer and SGI1-C mobilisation (28,29).

Nevertheless, the conservation of *sgaDC* in most members of the SGI1 family suggests an important role in their lifecycle rather than simple redundancy for activation of AcaCD-activatable promoters. In this study, we investigated the relevance of SgaCD. First, we showed using mating assays that SgaCD is important for SGI1 mobilisation by IncC plasmids. Using a combination of ChIP-exo, Cappable-seq and RNA-seq approaches, we compared transcriptional activation by SgaCD and AcaCD of AcaCD-responsive promoters in SGI1 and in a model IncC plasmid. A systematic assessment of promoter activities through β-galactosidase reporter assays allowed us to measure the differential response of these promoters to SgaCD and AcaCD complexes. Finally, we unveiled the crucial role of SgaCD in the excision and replication of SGI1, ultimately leading to the destabilisation of IncC conjugative plasmids.

**MATERIALS AND METHODS**

**Bacterial strains and media**

Bacterial strains, plasmids and genomic islands used in this study are described in Table 1. Strains were rou-
tinely grown in lysogeny broth (LB) at 37°C in an orbital shaker/incubator and were preserved at −75°C in LB broth containing 20% (vol/vol) glycerol. Antibiotics were used at the following concentrations: ampicillin (Ap), 100 μg/ml; chloramphenicol (Cm), 20 μg/ml; kanamycin (Km), 50 μg/ml or 10 μg/ml for single copy integrants of pOPlacZ; nalidixic acid (Nxa), 40 μg/ml; spectinomycin (Sp), 50 μg/ml; tetracycline (Tc), 12 μg/ml; rifampicin (Rf), 50 μg/ml. To induce expression from pBAD30 and pAH56, LB medium was supplemented with 0.2% L-arabinose or 0.1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), respectively. Conjugation assays were performed as described previously (33).

Molecular biology

Plasmid DNA was prepared using either the EZ-10 Spin Column Plasmid DNA Miniprep Kit (Bio Basic) or the QIAprep Spin Miniprep Kit (Qiagen), according to manufacturer’s instructions. Genomic DNA was prepared using the QIAamp DNA Mini Kit (Qiagen), according to manufacturer’s instructions. Restriction enzymes used in this study were purchased from New England Biolabs. Several DNA polymerases were used: Q5 (New England Biolabs), Taq (New England Biolabs) and Easy Taq (Civic Bioscience). PCR products were purified using either the EZ-10 Spin Column PCR Products Purification Kit (Bio Basic) or the QIAquick PCR Purification Kit (Qiagen), according to manufacturer’s instructions. E. coli was transformed by electroporation as described by Dower et al. (44) in a Bio-Rad GenePulser Xcell apparatus set at 25 kV using 1-mm gap electroporation cuvettes. Sanger sequencing reactions were performed by the Plateforme de Séquençage et de Génotypage du Centre de Recherche du CHUL (Québec, QC, Canada).

Plasmids and strains constructions

Oligonucleotides used in this study are listed in Supplementary Table S1. New sequencing data obtained from pVCR94 (NZ_CP033514.1) revealed that pVCR94Δ [pVCR94Δ(EEL44_05995-EEL44_06400)] (21,45) was missing eight genes of unknown function that are conserved in IncC plasmids. Therefore a new derivative, pVCR94Δκ [pVCR94Δ(EEL44_06035-EEL44_06400)], was constructed using the one-step chromosomal gene inactivation technique with pMS1, primer pair pVCR94delY2.f/94DelXnoFRT.rev and pVI36 as the template (46). SG1-C1κm and SG1-C2κm were constructed using the same technique with pMS6, primer pair SG11In104cm2.f/SG11In104cm2.r and pKD3 as the template. Deletion mutants of pVCR94Δκ, SG11κκ and SG11κκκ were constructed in a similar fashion. Deletion of acaDC in pVCR94Δκ was obtained using pMS6, primer pair 94DelacaD.f/94DelacaD.r and pKD4 as the template. Deletions of xis, sgaD, sgaC and sgaDC in SG11κκ were obtained using pMS6, primer pairs SG11delxis.f/SG11delxis.rev, SG11delsgaD07.for/SG11delsgaD07.rev, SG11delsgaC06.for/SG11delsgaC06.rev and SG11delsgaD07.for/SG11delsgaC06.rev, respectively, and pKD3 as the template. Deletion of sgaDC in SG11κκκ was obtained using pMS1, primer pair SG11delsgaD07.for/SG11delsgaC06.rev and pKD4 as the template. When possible, the antibiotic resistance cassette was removed from the resulting construction by Flp-catalysed excision using pCP20. All deletions were verified by PCR and antibiotic resistance profiling. sgaDC was amplified using primer pair SG11sgaDEC01.f/SG11sgaDEC01.r and genomic DNA of E. coli V6133 containing SG1κκκ as the template. The amplicon was then digested with EcoRI and cloned into EcoRI-digested pBAD30 using T4 DNA ligase (NEB), generating pBAD-sgaDC. psgaDCκκκ was derived from pacaDCκκκ. sgaDC was amplified using primer pair pAH56sgaCDinsF/pAH56sgaCDinsR and pacaDCκκκ was linearized using primer pair pAH56sgaCDvecF/pAH56sgaCDvecR. These four primers were designed using NEBuilder® Assembly Tool (NEB). psgaDCκκκ was obtained by ligating both amploncs using the Gibson Assembly@® Cloning Kit (NEB), to replace acaDC with sgaDC. PCR fragments containing the promoter region upstream of verx012, verx035–36, verx059, verx068, traA, verx076, dsbC, traC, acab-verx087, verx098, verx114, verx128, verx140, S004 and S018 were amplified using the corresponding primer pairs listed in Supplementary Table S1, and cloned into pOPlacZ using either PstI or PstI and XhoI to produce the corresponding pOPlacZ derivatives (Supplementary Figure S1). These vectors were ultimately integrated in single copy into the chromosomal site attBκ of E. coli BW25113 using pINT-Ts (47).

ChIP-exo assays

LB medium supplemented with 50 μg/ml rifampicin, 10 μg/ml kanamycin, 50 μg/ml spectinomycin and, when necessary, 20 μg/ml chloramphenicol was inoculated with E. coli MG1655 Rf bearing psgaDCκκκ, pVCR94Δκ ΔacaDC and SG11κκκ κΔsgaDC. Induction of sgaDCκκκ expression was done by adding 0.1 mM IPTG to cultures grown to an OD600 of 0.2, followed by a 1-h incubation at 37°C with shaking. 10 ml of culture was used for the ChIP-exo experiment, which was carried out as described previously (21), except for a shorter blocking step of magnetic beads used for immunoprecipitation: Dynabeads® Protein A (Invitrogen®) were blocked by washing three times for 5 min each, in PBS/BSA (5 mg/ml). Replicates are detailed in Supplementary Table S2.

Cappable-seq and RNA-seq assays

RNA was extracted from 5 ml cultures prepared as detailed above. Cultures were centrifuged for 10 min at 3700 g, and the pellet was thoroughly resuspended in 1 mL TRI Reagent® (Sigma-Aldrich). Total RNA was then extracted using the Direct-zol RNA MiniPrep Kit (Zymo Research) with the recommended DNase I treatment, according to manufacturer’s instructions. Cappable-seq was carried out as described elsewhere (48), with the following modifications. Approximately 10 μg of clean RNA was used for each assay. The very first RNA clean-up step was performed using Agencourt® RNAClean® XP Beads (Beckman) to en-
Table 1. Strains and elements used in this study

| Strains, plasmids or elements | Relevant genotype or phenotypea | Source or reference |
|-------------------------------|--------------------------------|---------------------|
| **E. coli**                   |                                |                     |
| BW2511                        | F- Δ araD-araB567, ΔlacZΔM715::Tn9, Δ (rhaD-rhaB)568, hsdR514 | (46)               |
| VBI13                         | Ns-derivative of BW2511        | (66, 79)            |
| KH95                          |                                 | (31)                |
| CAG1843                       | MG1655 lacZU18 lacI2::Tn10 (Tc) | (80)                |
| **MG1655 Rf**                 | Rf-derivative of MG1655        | (81)                |
| **Plasmids**                  |                                |                     |
| pVCR94Δ6                      | ΔEEL44::05995-EEL44, 06400) mutant of pVCR94 (Sp Su) | (21, 45)           |
| pVCR94Δ62                     | ΔEEL44::06035-EEL44, 06400) mutant of pVCR94 (Sp Su) | This study         |
| pVCR94Δ6 ΔacaDC               | ΔacaDC mutant of pVCR94Δ62     | (21)                |
| pVCR94Δ6 ΔacaDC               | ΔacaDC mutant of pVCR94Δ62     | This study         |
| pVCR4GreenSp                  | pVCR4Δ6 traG, PBAD-mNeonGreen-FRT (Su Sp) | (31)                |
| pMS1                          | pSIM5 Δcat::gen; Thermo-inducible expression of ARed recombination (Ts, Gn) | (66)                |
| pSIM6                         | Thermo-inducible expression of ARed recombination (Ts Ap) | (82)                |
| pV136                         | SpR PCR template for one-step chromosomal gene inactivation | (81)                |
| pKD3                          | CmR PCR template for one-step chromosomal gene inactivation | (46)                |
| pKD4                          | KnR PCR template for one-step chromosomal gene inactivation | (46)                |
| pCP20                         | Thermo-inducible expression of Flp recombinase (Ts Ap Cm) | (83)                |
| pBAD30                        | oriR101, lacZ4787, hPr- int8 (Ap Ts) | (21)                |
| pBAD-acaDC                    |                                |                     |
| pBAD30::sgaDC                 |                                |                     |
| pACaDC-3FLAG                  | pAH56::acaDC-3FLAG (Kn)       | (21)                |
| psigaDC-3FLAG                 | pAH56::sgaDC-3FLAG (Kn)       | This study         |
| pPLOlacZ                      | pAH56 lacZ (Kn)               | (21)                |
| pPromercl1                    | pPLOlacZ pArec-lacZ (Kn)      | (21)                |
| pINT-TS                       | oriR101; cB857; hPr-int8 (Ap Ts) | (47)                |
| **Genomic Islands**           |                                |                     |
| SGI1Δxis                      |                               |                     |
| SGI1Δxis                      | Δxis mutant of SGI1Δxis        | This study         |
| SGI1ΔsgaD                     | ΔsgaD mutant of SGI1ΔsgaD     | This study         |
| SGI1ΔsgaC                     | ΔsgaC mutant of SGI1ΔsgaC     | This study         |
| SGI1ΔsgaDC                    | ΔsgaDC mutant of SGI1ΔsgaDC   | This study         |
| SGI1Red                       |                                |                     |
| SGI1Red                      |                                |                     |
| SGI1Red·ΔsgaDC                | ΔsgaDC::aph mutant of SGI1RedΔsgaDC (Cm) | (31)                |
| SGI1·C0047SA97                | SGI1-C variant from Salmonella Agona 0047SA97 (Sm Sp Su) | (43)                |
| SGI1·C0047SA97                | SGI1-C variant from Salmonella Typhimurium DT104 S/954435 (Sm Sp Su) | (43)                |
| SGI1·CΔ6                      | ΔIns::cat mutant of SGI1·CΔ6 lacking the integron (Cm) | This study         |
| SGI1·CΔ6                      | ΔIns::cat mutant of SGI1·CΔ6 lacking the integron (Cm) | This study         |

*a*Ap, ampicillin; Cm, chloramphenicol; Gn, gentamycin; Kn, kanamycin; Rf, rifampicin; Sp, spectinomycin; Sm, Streptomycin; Su, sulfamethoxazole; Tc, tetracycline; Ts, thermosensitive.

...maximum elimination of unincorporated DTB-GTP. The removal of 3’ phosphates from fragmented RNA was performed using the Thermo Scientific® T4 Polynucleotide Kinase (Thermo Fisher Scientific) and its supplied ATP-free buffer. For RNA-seq samples, 800 ng of total RNA was fragmented in 5× RNA Fragmentation Buffer (200 mM Tris-acetate, pH 8.1, 500 mM KOAc, 150 mM MgOA) by incubating at 95°C for 7 min and quenching immediately on ice. RNA was then purified using the RNA Clean & Concentrator-5 Kit (Zymo Research), according to manufacturer’s instructions. The quality and concentration of RNA before and after fragmentation were evaluated using a 2100 Bioanalyzer instrument (Agilent Technologies). Replicates are detailed in Supplementary Table S2.

**Illumina sequencing library preparation**

ChIP-exo libraries were prepared as described previously (21), except for the second strand synthesis step that was performed using the Bst X DNA polymerase (Enzymatics) with ThermoPol® Buffer (NEB). Capable-seq and RNA-seq libraries were prepared using the NEBNext® Small RNA Library Prep Set for Illumina® (NEB), according to manufacturer’s instructions, except NEBNext 5’ SR Adaptor for Illumina was replaced by previously described 5’-hybrid-A0 oligo (21). DNA molecules corresponding to the rRNA transcript were depleted using the duplex-specific nuclease (Evrogen), as described elsewhere (49). All libraries were amplified and checked as previously described (21), then ultimately pooled. Illumina sequencing was performed in two different sequencing runs on a NextSeq® 500/550 High Output system at the Plateforme Rnomique de l’Université de Sherbrooke (Sherbrooke, QC, Canada).

**Bioinformatic analyses**

Reads were trimmed with Trimmomatic (50) to discard nucleotides with a quality score below 30 and reads with a length below 36 bp. Quality was assessed before and after using FastQC (51). Trimmed reads were aligned on the *E. coli* MG1655 genome (NC_000913), pVCR94Δ62 ΔacaDC...
and SG11Red ΔsgaDC using Bowtie 2 (52). Alignment quality was assessed using SAMStat (53), and reads with a quality score below 10 were discarded using SAMTools view (54). Information on the number of reads before and after trimming, the quality of mapping and coverage for each sample, is detailed in Supplementary Table S2. Reads were then compressed, sorted and indexed using SAMTools (54). ChIP-exo and Cappable-seq reads were chopped to their first nucleotide and density was calculated separately for each DNA strand using BEDTools genomcview (55). Density files were ultimately compressed to BigWig format and visualised on the UCSC Genome Browser.

The footprint profile for each transcription start site of interest was analysed using the Versatile Aggregate Profiler (VAP) (56). Three pairs of divergent AcaCD-dependent promoters (verx035–036, verx059-tral and acaB-verx087) were excluded from the analysis to prevent potentially overlapping ChIP-exo signals on the positive and negative strands. Replicates of a given condition were pooled using SAMTools and reads were treated as described above. Densities for any condition and strand were normalised by the total signal obtained for a given reference (pVCR94Sp2 ΔacaDC or SG11Red ΔsgaDC). The start and end coordinates of each transcription start site of interest were used as reference points. The signal was reported for each base pair and represented as median or first to ninth deciles or full range without smoothing. RPKM values were calculated for each DNA strand separately, using a script adapted from EDGE-pro (57). Differential expression analysis was performed using DESeq2 (58). Signal was also calculated per chromosome separately, using a script adapted from EDGE-pro (57). The data were analysed for the variability between replicates (Supplementary Figure S2).

The KEGG pathway enrichment analysis was performed in RStudio v1.3. DESeq2 output was filtered to discard MG1655 genes displaying less than 10 accumulated reads. Log2 of fold-change values were analysed using GAGE to estimate up- and down-regulated pathways (59). Pathways were considered to be significantly enriched or depleted when the adjusted P-value was inferior to 0.1. Search for SgaC homologues was carried out using Blastp (60) against the Genbank non-redundant protein sequence (nr) database restricted to the Enterobacteriaceae (taxid:543).

β-Galactosidase assays

The assays were carried out as described previously, using o-nitrophenyl-β-D-galactopyranoside (ONPG) as substrate (61). Cultures were prepared with LB medium supplemented with 10 μg/ml kanamycin to select the strain and 50 μg/ml ampicillin to maintain pBAD-acaDC or pBAD-sgaDC. Induction of acaDC or sgaDC was done by adding 0.2% arabinose to a refreshed culture grown to an OD600 of 0.2, followed by a 2-h incubation at 37°C with shaking prior to cell sampling.

qPCR assays

Genomic DNA was obtained from 1 ml of cell cultures of E. coli VB113 bearing pVCR94Sp, SG11Kn or their mutants, grown for 16 h in LB medium supplemented with 40 μg/ml nalidixic acid, 50 μg/ml spectinomycin and 50 μg/ml kanamycin. Genomic DNA purity and concentration were measured with an ND-1000 NanoDrop spectrophotometer (Thermo Fisher Scientific). qPCR experiments were performed in technical triplicate for three biological replicates at the Plateforme Rnomique de l’Université de Sherbrooke (Sherbrooke, QC, Canada). All forward and reverse primers were individually resuspended to 20–100 μM in Tris-EDTA buffer (IDT) and diluted as primer pairs to 1 μM in RNase DNase-free water (IDT). Amplicons were analysed by automated chip-based microcapillary electrophoresis on Labchip GX Touch HT instruments (Perkin Elmer). Quantitative PCR (qPCR) reactions were performed in 10 μl in 384-well plates on a CFX-384 thermocycler (BioRad) with 5 μl of 2X PerfeCTa SYBR® Green Supermix (QuantaBio), 10 ng (3 μl) cDNA, and 200 nM final (2 μl) primer pair solutions. The following cycling conditions were used: 3 min at 95°C; 50 cycles: 15 sec at 95°C, 30 s at 60°C, 30 s at 72°C. attrB (236 bp), sgiA (5026, 234 bp) from SG11Kn, repA (237 bp) from pVCR94Sp, as well as chromosomal reference genes dnaB (235 bp), hicB (235 bp) and trmA (238 bp) from E. coli, were quantified using primer pairs qAttBFw/qAttBRv, qS026Fw/qS026Rv, qFwpVCR/qRvpVCR, qdnaBFw/qdnaBRv, qhicBFw/qhicBRv and qthdBFw/qthdFRv, respectively (Supplementary Table S1). The data were analysed and normalised using all three chromosomal genes dnaB, hicB and trMA as references and the qBase framework (62). Excision of SG11Kn was calculated as the ratio of free attrB site per chromosome, SG11Kn and pVCR94Sp copy numbers were calculated as the ratios of sgiA and repA per chromosome, respectively. For clarity, E. coli VB113 bearing only pVCR94Sp was considered to exhibit a 100% excision rate. VB113 bearing either pVCR94Sp or SG11Kn were presumed to contain only one copy of a given element.

RNA isolation and quantitative reverse transcription qRT-PCR

Total RNA was extracted as described above. RNA purity and concentration were evaluated with an ND-1000 NanoDrop spectrophotometer (Thermo Fisher Scientific). qRT-PCR assays were performed in technical triplicate for three biological replicates at the Plateforme Rnomique de l’Université de Sherbrooke (Sherbrooke, QC, Canada). RNA integrity was verified using a 2100 Bioanalyzer instrument (Agilent Technologies). cDNA was prepared from 1 μg of total RNA using 10 units of Transcriptor reverse transcriptase (Roche), 0.08 μl of random hexamers (IDT), dNTPs (Roche) and 10 units of RNaseOUT (Invitrogen Life Technologies) according to the manufacturer’s recommendations in a total volume of 20 μl. Quantitative amplification of acaD and sgaD was carried out as described above with primer pairs acaDFqPCR/acaDR/qPCR and sgaDFqPCR/sgaDRqPCR, respectively. The stably expressed chromosomal gene rpoZ was amplified with primer rpoZMG1655F/rpoZMG1655R. Relative expression levels were calculated using the qBase framework (62) and rpoZ as the reference.
Cohabitation assays

The assays were carried out as described previously (31). Culture samples were diluted 1:1000 in 1 ml of PBS. Fluorescence intensity of mNeonGreen and mCherry in cells was monitored by flow cytometry analysis on a BD FACSCalibur (BD Biosciences), and data were acquired with the BD FACSortware. mNeonGreen and mCherry were excited with 488 and 561 nm solid-state lasers, and their emission was detected using 513/17 and 610/20 nm emission filters, respectively. For each sample, fluorescence of 20,000 cells was captured, and the data was analysed using FCS Express 7 (De Novo Software).

Statistical analyses and figures

Prism 8 (GraphPad Software) was used to plot graphics and to carry out statistical analyses. All figures were prepared using Inkscape 1.0 (https://inkscape.org/) or BioRender (https://biorender.com).

RESULTS

SGI1 rescues the transfer of an IncC plasmid lacking its master activator of transfer

Given the similarity between AcaCD and SgaCD, and the essentiality of AcaCD for IncC plasmid transfer activation, we first assessed the importance of sgaDC. Transfer rates of SGI1Red and the coresident IncC helper plasmid pVCR94Sp2 were assessed in mating assays using combinations of wildtype and ΔDC mutants of both elements. Deletion of acaDC had no impact on the transfer of the helper plasmid, SGI1Red or the cotransfer of both elements, confirming that SGI1 can complement the loss of acaDC (Figure 1A). In contrast, deletion of sgaDC reduced the transfer of SGI1Red nearly 2000-fold compared to the wildtype, whereas it had no impact on the transfer of the helper plasmid or the cotransfer of both elements, showing that sgaDC is important for SGI1 transfer and/or stability (Figure 1A). Deletion of both acaDC and sgaDC abolished the transfer of both elements. Finally, overexpression of a single chromosomal copy of sgaDC^{3XFLAG} under the control of the IPTG-inducible Ptac promoter activated the transfer of both elements beyond wildtype levels (Figure 1A). To confirm that this phenotype was exclusively attributable to SgaCD, trans-complementation with chromosomal Ptac^{sgaDC^{3XFLAG}} was also performed using donors lacking SGI1. While the deletion of acaDC abolished pVCR94Sp2 transfer in this context, it was fully restored by sgaDC^{3XFLAG} overexpression (Figure 1B).

SgaCD activates all AcaCD-activatable promoters and generates the same binding footprint as AcaCD

The SgaCD regulon was characterised using a combination of ChIP-exo, Cappable-seq and RNA-seq approaches to identify the binding sites, transcriptional start sites (TSS) and quantify mRNA transcript levels, respectively. In these experiments, either sgaDC^{3XFLAG} or acaDC^{3XFLAG} was expressed from Ptac instead of the native promoter to ensure homogenous expression in cell populations. Otherwise isogenic cells lacking sgaDC^{3XFLAG} and acaDC^{3XFLAG} were used as negative controls. We also used ΔDC mutants of pVCR94Sp2 and SGI1Red to prevent interference from the native unlabelled complexes. The ChIP-exo experiment revealed that SgaCD binds all known AcaCD-activatable promoters on pVCR94Sp2 and SGI1 (Figure 2). SgaCD notably activates the promoter of acaB whose translation product activates acaDC expression (26). No additional promoter was found to be bound by SgaCD, showing that AcaCD and SgaCD complexes have similar activities and recognise identical sequence motifs. ChIP-exo peaks usually paired with those of Cappable-seq on the positive DNA strand, except for the control condition, confirming that SgaCD binding correlates with transcription of each promoter (Figure 2). Transcriptomic data obtained from RNA-seq assays with pVCR94Sp2 confirmed that sgaDC expression leads to an overall increase of mRNA levels, especially for operons associated with conjugative transfer functions (Figure 2A). Likewise, expression of most SGI1 genes increased upon expression of sgaDC or acaDC (Figure 2B).

To compare the binding footprint of both complexes, two aggregated profiles were built by compiling ChIP-exo and Cappable-seq signals of pVCR94Sp2 and SGI1 promoters. No major difference could be observed between the SgaCD and AcaCD profiles (Figure 3, upper panels). Both complexes bind to the same site located -64 to -37 bp upstream of the transcription start site (TSS) (21). The large protected sequence starts -42 bp upstream of the TSS, that is 4 bp downstream of the GGCGCGWWGGGC palindromic motif, and ends 22 bp after the TSS. This protection is compatible with the promoter-bound RNA polymerase holoenzyme complex, as previously reported for class II activation where the bound activator complex abuts the promoter –35 element (Figure 3, upper panels and Supplementary Figure S3A–D) (21,22,63,64). The sequence overlapping the binding motif, where two peaks are observed immediately after GCCCG and GGCG, corresponds to the AcaCD/SgaCD footprint. The AcaCD profile displayed sharper boundaries than that of SgaCD with a stronger signal at these positions (Figure 3, upper panels and Supplementary Figure S3A–D). Assuming comparable crosslink efficiencies, this observation suggests better recruitment and tighter binding of AcaCD compared to SgaCD. Remarkably, the Cappable-seq signal of SgaCD on SGI1 promoters was weaker than that of AcaCD. Nevertheless, these results show that SgaCD and AcaCD, though encoded by two unrelated mobile genetic elements, recognise and bind identical sequence motifs.

Both complexes appeared to bind better to P_{trans} relatively to other AcaCD-activatable promoters (Figure 3 lower panels). Comparison of individual profiles obtained for P_{trans} and P_{5000} promoters with AcaCD and SgaCD, revealed clear differences. AcaCD and SgaCD profiles at P_{trans} were similar to the aggregated profiles (Supplementary Figure S3E, F), supporting the hypothesis that SgaCD binds less efficiently than AcaCD, leading to a weaker transcription initiation, as suggested by the weaker Cappable-seq signal. In strong contrast, the individual profiles obtained with SGI1-borne P_{5000} promoter were unique (Supplementary Figure S3G, H). Comparison of ChIP-exo signals revealed efficient binding of SgaCD but not AcaCD. Detection of a strong AcaCD ChIP-exo signal on the right border shortly after the bind-
ing site (at position –26) suggests premature detachment or poor recruitment of the transcriptional activator (Supplementary Figure S3H). The extremely weak Cappable-seq signal supports the hypothesis of a faulty transcriptional initiation by AcaCD. Except for a stretch of five Gs located between the binding site and –10 element, no specific feature could be found that could explain the weak activation of $P_{S004}$ by AcaCD (Supplementary Figure S1). In contrast, the Cappable-seq signal obtained with SgaCD was higher and displayed 3 peaks, at the predicted TSS, and 5 and 8 bp after the TSS. In summary, SgaCD seems to activate $P_{S004}$ with much higher efficiency than its IncC-encoded homologue.

**Overexpression of sgaDC significantly affects the transcriptome**

A differential expression analysis was performed on transcriptomic data to identify up- and down-regulated genes following sgaDC overexpression. This analysis considers the different sequencing depths of each condition and provides statistical metrics. A vast majority of differentially expressed genes were found to be activated by SgaCD (Figure 4A). For instance, expression of the IncC plasmid genes *traL* or *traK* that encode predicted inner and outer membrane subunits of the conjugal T4SS, respectively, increased 500 to 1000-fold (Figure 4A). In contrast, few plasmid genes were found to be downregulated upon sgaDC overexpression. The most significant reduction of expression was observed for *verxI19* and *verxI20*, two genes of unknown function. Since no binding signal was detected upstream of these two genes, repression by SgaCD is likely indirect (Figure 2). Furthermore, expression of the mobilisation gene *mobI*, which encodes a key factor for initiation of conjugative transfer at oriT, remained unchanged, confirming its independence from AcaCD and SgaCD (Figure 4A) (21,65,66). Finally, none of the eight genes of pVCR94Sp2 that are missing in pVCR94Sp were expressed under the control of an AcaCD-activatable promoter (Figures 2 and 4A).

916 chromosomal genes were differentially expressed upon sgaDC overexpression in the presence of pVCR94Sp or AcaCD (Supplementary Figure S4A and Supplementary Table S3). Several of the most up-regulated genes have been shown to be involved in resistance to stress and membrane transport. Down-regulated genes are involved in cysteine and enterobactin synthesis, outer membrane transport and anaerobic respiration. A pathway enrichment analysis showed that sgaDC overexpression down-regulated several pathways. In virtually all tested conditions, aminoacyl-tRNA biosynthesis, amino acids and secondary metabolite pathways were significantly impacted (Supplementary Table S4). Interestingly, flagellar assembly was also significantly depleted by overexpression of either sgaDC or acaDC. As no SgaCD binding could be detected by ChIP-exo in the promoter region of the affected genes, SgaCD influence on chromosomal gene expression was likely indirect. Although depletion of these pathways correlates with the activation of transfer of IncC conjugative plasmids and SGII, it is not clear whether it results from the burden caused by the conjugal transfer or from an active redirection of metabolic pathways to favour this energy-consuming process.

Four additional differential expression analyses were conducted to evaluate the impact of SGII and to compare activation by SgaCD and AcaCD. Expression of *mobI* was repressed upon overexpression of sgaDC in the presence of SGII**Red** ΔsgaDC (Figures 4B and D). Since expression of *mobI* is independent of SgaCD, such repression likely results from expression of an SGII-encoded factor. Few genes were found to be differentially expressed when overexpressing sgaDC compared to acaDC (Figures 4C and F). Therefore, a weaker transcriptional initiation as suggested by Cappable-seq does not necessarily lead to lower levels
of transcripts in the artificial context of overexpression. Finally, we confirmed that most SGI1 genes are up-regulated by SgaCD, including $\texttt{xis}$ and $\texttt{rep}$, with $\texttt{S004}$ and $\texttt{traH_S}$ presenting the highest fold-change values (Figure 4E).

**Activation of gene expression by SgaCD is comparable to AcaCD**

Since transcriptomic data only show the outcome of potentially multiple regulatory processes, ChIP-exo and Cappable-seq results were confirmed using an expression reporter assay based on transcriptional $\texttt{lacZ}$ fusions to each of the 23 AcaCD-activatable promoters inserted in single copy into the chromosome. $\beta$-galactosidase assays were carried out with both AcaCD and SgaCD to quantify the relative expression level with each activator complex (Figure 5 and Supplementary Figure S1). While $\texttt{verx035}$, $\texttt{verx036}$ and $\texttt{verx087}$ were constitutively expressed, virtually all other promoters were directly activated by AcaCD and SgaCD (Figure 5A–C). Induction ratios of SgaCD were almost invariably lower than those of AcaCD, confirming a previous report with the five AcaCD-activatable promoters of SGI1 (Figure 5D) (29). However, the difference was not statistically significant for 15 out of 23 promoters. The gene of unknown function $\texttt{verx087}$ was expressed both constitutively,

---

**Figure 2.** In-depth analysis of the AcaCD/SgaCD regulon. Results of ChIP-exo, Cappable-seq and RNA-seq experiments on *E. coli* MG1655 Rf carrying both pVCR94*Δ* $\texttt{acaDC}$ (A) and SGI1*Δ* $\texttt{sgaDC}$ (B) with or without a single chromosomal copy of either psgaDC$\texttt{3XFLAG}$ expressing the native SgaD subunit along with a C-terminal 3XFLAG-tagged SgaC subunit induced by IPTG, or pacDC$\texttt{3XFLAG}$. The circular map of the plasmid was linearised after the $\texttt{repA}$ gene. The location and orientation of ORFs are depicted by arrowed boxes, which are color-coded by function as indicated. Green arrows indicate previously identified (A) or predicted (B) AcaCD-dependent promoters. Each track (one representative replicate per condition as detailed in Supplementary Table S2) plots the number of mapped reads as a function of the position in each element. Read densities are displayed as black bars, or blue and orange bars for Cappable-seq densities on the positive and negative DNA strands (linear scale for ChIP-exo and Cappable-seq densities, log scale for RNA-seq densities). Pink dots at the summit of peaks indicate a signal beyond the represented y-axis maximal value.
which can be explained by the presence of nearly canonical $\sigma^{35}$ and $-10$ elements in its promoter sequence, and under the control of AcaCD/SgaCD (Figure 5B and Supplementary Figure S1). In contrast, although the promoters driving the expression of \textit{vcrx035} and \textit{vcrx036} seemed constitutively active, both lacked canonical $-35$ and $-10$ elements (Supplementary Figure S1).

SgaCD is essential for SGI1 replication

To investigate the role of SgaCD in the lifecycle of SGI1, we conducted qPCR assays using different mutants of pVCR94Sp and SGI1Kn. As shown previously (21,30), SGI1 excision was undetectable in the absence of the helper plasmid or when a $\Delta$xis mutant was used (Figure 6A). Surprisingly, deletion of \textit{acaDC} had a somewhat limited, statistically non-significant impact on the excision rate of SGI1Kn. On the contrary, deletion of \textit{sgaDC} reduced the excision rate nearly 3000-fold compared to the wild-type level, supporting a crucial role of SgaCD compared to AcaCD in SGI1’s lifecycle. Deletion of both abolished the excision, indicating that residual excision of SGI1Kn $\Delta$sgaDC is triggered by the helper plasmid-encoded AcaCD. A comparable phenotype was observed when measuring the copy number of SGI1Kn (Figure 6B). Deletion of \textit{sgaDC} abolished SGI1 replication, whereas deletion of \textit{acaDC} reduced it only 2.7-fold (Figure 6B). None of the deletions had any statistically significant impact on pVCR94Sp copy number (Figure 6C). Together, these results account for the reduced mobilisation of SGI1Kn $\Delta$sgaDC by pVCR94Sp (Figure 1A). Besides, residual excision and replication of SGI1Kn observed in the presence of pVCR94Sp $\Delta$acaDC but not in cells lacking the IncC plasmid suggest that an unidentified plasmid-encoded factor triggers or derepresses the expression of \textit{sgaDC}.

Expression of \textit{sgaDC} and \textit{acaDC} increases in IncC$^+$ SGI1$^+$ cells

\textit{sgaDC} was reported to be expressed constitutively at a low level (28). We measured the impact of the coexistence of
SGI1 and its helper IncC plasmid on the relative expression of *sgaDC* and *acaDC* using RT-qPCR. Cells bearing either pVCR94^Sp^ or SGI1^Kn^ were used as controls (relative expression level of 1). The presence of SGI1 resulted in a ∼8-fold increase of *acaDC* expression, whereas the IncC plasmid led to a ∼4.5-fold increase of *sgaDC* expression (Figure 6D). These results show that the presence of an IncC plasmid elicits *sgaDC* expression. Conversely, SGI1 augments the expression of *acaDC*, likely via activation of *acaB* expression.

SgaCD-activated replication of SGI1 destabilises the helper IncC plasmid

Incompatibility between SGI1 and a co-resident helper IncC plasmid has been shown to be linked to the replicative
Figure 5. Systematic analysis of AcaCD/SgaCD-dependent promoters. Activity of all AcaCD-dependent promoters on IncC plasmids and SGI1 was monitored from single-copy, chromosomally integrated transcriptional lacZ fusions in E. coli BW25113. (A) Qualitative assay on LB medium supplemented with X-Gal. Promoters are identified by the first gene of the corresponding operon. AcaCD-independent promoters are labelled in black. (B, C) Miller units. (D) Induction ratios. β-galactosidase assays were carried out in LB medium supplemented with (grey bars in A, B) or without (white bars in A, B) arabinose to express acaDC or sgaDC from PBAD on pBAD-acaDC (A) or pBAD-sgaDC (B). The bars represent the mean and standard error of the mean obtained from a biological triplicate. In panel D, the x axis crosses the y axis at y = 1, indicating no induction. For panels B-D, three independent two-way ANOVA with Sidak’s multiple comparison test were performed on the logarithm of the values to compare the pair of bars for each promoter. Statistical significance is indicated as follows: **** P < 0.0001; *** P < 0.001; ** P < 0.01; * P < 0.05; ns, not significant.
state of SGI1 (31). To test whether sgaDC had any role to play in SGI1 replication and incompatibility, we used a previously designed flow cytometry assay aimed at assessing the evolution of a population of cells carrying red fluorescence-producing SGI1Red and green fluorescence-producing pVCR94GreenSp over time (31). Cells producing red fluorescence segregate into two populations. Low-intensity red-fluorescent cells carry a single copy of SGI1 usually integrated in the chromosome, whereas cells producing high-intensity red fluorescence contain excised replicating SGI1. pVCR94GreenSp alone remained relatively stable and was retained in 71% of cells at T48 (Figure 7A). Cells carrying SGI1Red or its ΔsgaDC mutant produced mostly low-intensity red fluorescence and remained steady (Figure 7B, C). Overexpression of acaDC was sufficient to promote SGI1 excision and replication in the absence of the helper IncC plasmid as 71% of the cells produced high-intensity red fluorescence at T0 (Figure 7D). However, SGI1Red ΔsgaDC was rapidly lost as most cells failed to produce any fluorescence at T24. Overexpression of sgaDC resulted in a more widespread activation of SGI1 replication as >90% of cells were highly red fluorescent at T0 (Figure 7E). SGI1 loss was delayed with sgaDC compared to acaDC, with >6% of cells retaining actively replicating SGI1 at T48. Together, these results show that SGI1 excision and replication can occur in the absence of the helper IncC plasmid when provided with AcaCD or SgaCD, although the latter seemed to promote stronger activation of rep expression. However, after excision, replication of SGI1 was insufficient to ensure its inheritance in the cell population in the absence of the IncC plasmid, despite the presence of the sgiAT toxin-antitoxin system (34).

When SGI1Red was in the presence of pVCR94GreenSp, most cells produced high red fluorescence at T0, confirming SGI1 replication (Figure 7F). At T48, only 16% of the cells produced high red fluorescence, a striking decrease that correlated with the loss of pVCR94GreenSp. Conversely, low red fluorescence signal increased, consistent with chromosomally integrated SGI1Red. In contrast, when SGI1Red ΔsgaDC was used, less than 4% of cells exhibited a high red fluorescence signal at T0, showing reduced SGI1 replication (Figure 7G). In addition, green and low red fluorescence remained steady throughout the experiment confirming that SGI1Red remained chromosomally integrated and that pVCR94GreenSp persisted in the cells. Remarkably, and consistent with the qPCR assays (Figure 6B), the presence of acaDC on the helper plasmid failed to trigger SGI1 replication in this context, confirming the importance of sgaDC in SGI1’s lifecycle. Complementation of ΔsgaDC using pBAD-sgaDC restored SGI1Red replication and a strong incompatibility phenotype (Figure 7H). Remarkably, IncC plasmid loss correlated with decreased SGI1 replication and increased SGI1 integration, confirming that plasmid instability is caused by the rescue of sgaDC expression (Figure 7H). Finally, complementation using pBAD-acaDC triggered SGI1 excision but failed to restore SGI1 replication, resulting in progressive loss of both SGI1Red and pVCR94GreenSp (Figure 7I). Altogether, these data indicate that sgaDC, not acaDC, controls SGI1 excision and replication, and the subsequent destabilisation of the helper IncC plasmid.

A rare single nucleotide polymorphism (SNP) in sgaC of several SGI1 variants likely affects SgaCD activity

Contrasting with our observations with SGI1, deletion of sgaDC was shown to have no impact on the mobilisation of SGI1-C, whereas deletion of acaDC completely abolished both plasmid self-transfer and SGI1-C mobilisation (28,29). Moreover, excision and replication of SGI1-C are undetectable in the presence of the ΔacaDC mutant of the helper IncC plasmid (32). Hence, unlike sgaDC of SGI1, sgaDC of SGI1-C seems to be unable to complement an acaDC null mutant of its helper plasmid or is perhaps not even functional. To identify the cause of these discrepancies, we sought to align S008-sgaDC genes and upstream sequences of SGI1 and SGI1-C. Unfortunately, the sequence of SGI1-C used by Kiss et al. (28) to test the role of sgaDC is not available, and its original host not associated with a specific strain of S. enterica serovar Typhimurium.
Figure 7. Effect of sgaDC deletion on incompatibility between SGI1 and IncC plasmids in the absence of selective pressure. Evolution of the percentage of E. coli KH95 cells bearing (A) pVCR94GreenSp (IncC) or (B) SGI1Red (IncI) or (C–E) SGI1Red ΔsgaDC (SGI1) or the indicated combination of both elements (F–I), in the absence or presence of pBAD-sgaDC (D,H) or pBAD-acaDC (E,I), over 48 h in the absence of antibiotics (except for ampicillin for maintenance of pBAD derivatives) as monitored by flow cytometry. Plots show the mean and standard error of the mean values obtained from a biological triplicate.

The Genbank nucleotide database contains only four sequences identified as SGI1-C in clinical strains of P. mirabilis from China, including strain CA150323 (Genbank MH990679) used here as the reference (68). SGI1-PmBRI, identified in the clinical isolate P. mirabilis PmBRI from France (Genbank JX089582.1), is strikingly similar to SGI1-C from China (14). We found only two SNPs in SGI1-C of CA150323 at positions 6455 (C to A) and 6,510 bp (G to A) relative to SGI1 but not in SGI1-PmBRI. While the second SNP is a silent mutation in sgaC, the first SNP changes the CTG codon to an ATG codon, resulting in the L139M substitution in SgaC C-terminal moiety immediately upstream of the predicted Zinc finger (Supplementary Figure S5F). A Blastp search of SgaC homologues revealed either L or M amino acid residues at position 139. However, the M residue at this position is extremely rare in SGI1 variants (3 out of the first 100 hits), with the notable exception of SGI2, formerly known as SGI1-J, of S. enterica serovar Emek (Genbank AY963803) from the UK (1999) and SGI1-F of S. enterica serovars Albany and Cerro (GenBank KU847976) (35,69–71). Furthermore, distantly related C subunits of other conjugative elements also have an L residue at the corresponding position (Supplementary Figure S5F), suggesting that the L139M substitution could be detrimental to their activity. FlhC, the most distant homologue, has a V residue at the corresponding position with a similarly hydrophobic yet shorter side chain than L.

To test whether the L139M substitution in SgaC reduces the activity of SgaDC, we introduced SGI1-CCm, a chloramphenicol-resistant derivative of the prototypical SGI1-C variant from S. enterica Agona strain 0047SA97 and S. enterica Typhimurium DT104 strain S/954435 (43), into E. coli bearing pVCR94 ΔacaDC. To our surprise, SGI1-Cm complemented the ΔacaDC helper plasmid to levels that were comparable to SGI1Cm (Supplementary Figure S6). Furthermore, sequencing of sgaC from SGI1-
DISCUSSION

SGI1 encodes the transcriptional activator complex SgaCD, whose biological relevance remained unclear until now (29). The similarity of SgaCD to other activator complexes, such as AcaCD encoded by IncC conjugative plasmids and SetCD encoded by SXT/R391 ICEs, suggested an important, yet perhaps redundant, role in SGI1 mobilisation (28,41). Here, we characterised the IncC regulon using ChIP-exo experiments, transcriptomic analyses, and β-galactosidase reporter assays. ChIP-exo experiments showed that SgaCD recognises and binds to the same sites as AcaCD both on the IncC plasmid and on SGI1, leading to activation of the transfer genes and operons (Figures 3 and 5). We also showed that despite the evolutionary distance, both complexes bind the same DNA motifs. However, sgaDC and acaDC are not exchangeable in their natural context, as deletion of either gene set leads to drastically different phenotypes. While acaDC is essential to activate IncC plasmid transfer in the absence of SGI1, it becomes dispensable in SGI1+ cells. In contrast, suppression of sgaDC of SGI1 in the presence of the IncC plasmid abolishes SG11 replication and allows peaceful coexistence of both elements despite the presence of a fully functional copy of acaDC. During revision of this manuscript, others have confirmed the importance of SgaCD in SGI1 replication (32). Consistent with our results, SGI1-K that lacks sgaDC due to a 2779-bp deletion extending from S008 to the 5′ half of traN5 (S005) is compatible with IncC plasmids (35,42). Interestingly, SGI1-C cannot excise and replicate in cells bearing a ΔacaDC helper IncC plasmid, and fails to complement the transfer of such a mutant (28,32). Furthermore, the sgaDC deletion has no impact on SGI1-C mobilisation, excision or replication in the presence of a wild-type IncC plasmid, suggesting that it is not functional. We propose here that the single rare mutation L139M in SgaC encoded by a subset of SGI1-C variants and by SGI1-F and SGI2 could be responsible for this phenotype, likely rendering SgaCD unable to act as a transcriptional activator at its physiological expression level.

The ability of SGI1 to complement an acaDC null IncC plasmid could have important epidemiological consequences. Occurrence of naturally acaDC-defective IncC plasmids has previously been reported (21). Although probably unable to activate self-transfer, entry of SGI1 in the host could resuscitate such ‘zombie’ plasmids that would transiently regain their capacity to transfer, and mediate mobilisation of SGI1 (Supplementary Figure S7). This process would be facilitated by the ability of SGI1 to escape entry exclusion exerted by IncA and IncC plasmids that normally prevents or strongly reduces redundant transfer between cells that contain plasmids of the same entry exclusion group (33,45).

The pathway allowing SGI1 to complement an acaDC null IncC plasmid is unclear, and could involve previously reported low-level, constitutive expression of sgaDC (28,72). Rare spontaneous excision of SGI1 suggests that the promoter of sgaDC is either mostly repressed under normal conditions or drastically activated by the presence of an IncA or IncC plasmid (67). Low SgaCD level produced by integrated SGI1 could be unable to switch on xis expression, preventing SGI1 excision and replication in the absence of the helper plasmid. Likewise, low SgaCD level is probably insufficient to trigger expression of the entire IncC plasmid tra gene set and initiate transfer. However, since the rates of mobilisation of SGI1 and self-transfer of the ΔacaDC helper plasmid are comparable to wild-type (Figure 1A), we propose that an IncC plasmid-encoded factor, triggers a positive feed-back loop in response to low levels of SgaCD likely via activation of sgaDC expression. Alternatively, plasmid entry itself could eventually act as a trigger for activation of sgaDC expression (Figure 8), perhaps via activation of the SOS response by the invading single DNA strand during conjugation (73). Recent discovery of AcaB added an important piece to the regulatory switch that controls the ‘On/Off state’ of conjugative transfer of IncA and IncC plasmids (26). Mutual activation of acaB and acaDC has been proposed to be the trigger or amplifier of the conjugative transfer ‘On state’. We showed here that, like AcaCD, SgaCD activates expression of AcaB (Figures 2 and 5), hence low level of SgaCD could initiate derepression of Pacro via AcaB activation. This initial gentle push would then be amplified by AcaCD, promoting excision and replication of SGI1, which in return would raise SgaCD levels. However, although AcaB activates acaDC expression, it probably does not activate sgaDC expression, since no AcaC- or AcaB-binding site was found upstream of sgaDC (21,26). Yet another IncC-encoded factor is likely at play. Consistent with this hypothesis, SGI1’s rep gene was shown to be slightly expressed in the presence of pVCR948p ΔacaDC, but not in its absence (31). Given the crucial role of sgaDC for SGI1 excision, replication, mobilisation and incompatibility with IncC plasmids (Figures 1A, 6A, B and 7G), this observation supports the existence of a plasmid-encoded activator of sgaDC expression. Hence, the promoter of sgaDC instead of Pacro as previously suggested (28) would act as the sensor for IncC plasmid entry. Identification of the putative factor promoting such a feedback loop is on-going.

By analogy with the flagellar activator FlhCD to which they are distantly related, the C subunit of AcaCD and SgaCD likely binds to DNA immediately upstream of the −35 element, whereas the D subunit would stabilise the transcriptional complex (74). In support of this hypothesis, the primary sequence of D subunits is more divergent than the one of C subunits (28). Furthermore, AcaC and SgaC also share predicted tertiary structures that are remarkably similar to that of FlhC, including the zinc-finger domain and its four conserved cysteine residues (Supplementary Figure S5). However, the region located between the two inner cysteine residues differs greatly in FlhC compared to AcaC and SgaC. The similarity of SgaC and AcaC in this region likely accounts for the cross recognition of the same set of binding sites and the difference with FlhC binding sites.
Figure 8. Model of regulation of IncC plasmids and SGI1 gene expression. Major genes and operons involved in conjugative transfer and regulation are depicted as color-coded arrowed boxes (or a star for IncC origin of transfer oriT) based on the function: green, purple or grey, transcriptional activation; red, transcriptional repression; blue, type IV secretion system; orange, relaxosome; black, site-specific recombination; yellow, replication. Promoters are depicted by angled arrows and color-coded based on the corresponding activator: grey, unknown; green, AcaCD/SgaCD-activatable; purple, AcaB-activatable. Activation is represented by faded arrows. Repression is represented by red blocked dashed arrows. A bidirectional grey arrow indicates potential interactions. The chromosome is depicted as a large knotted DNA structure. SGI1 is represented in its initial integrated form, with its attachment sites depicted as two black lines. We propose a model in six main steps: (i) activation of sgaDC expression, (ii) activation of SGI1 excision, (iii) activation of SGI1 replication, (iv) activation of acaB, (v) activation of acaDC and (vi) activation of all the AcaCD/SgaCD-dependent promoters. For clarity, the activation of only two IncC transfer operons is depicted. Timing of the alleviation of repression by Acr1 and Acr2 is unknown. Created with BioRender.com.

(75,76). Considering that SgaC and AcaC complexes derived from a common ancestor, the apparent weaker promoter activation by SgaCD could have evolved during its domestication by SGI1-like elements to reduce the risk of futile excision in the absence of a helper plasmid, enhancing SGI1 stability. We initially hypothesized that weak promoter activation could be compensated by the ability of SGI1 to replicate. IncC plasmids maintain as a single-copy replicon per cell, whereas the replicative cycle of SGI1 in IncC cells generates over seven SGI1 copies per cell, which could boost sgaDC expression (Figure 6B) (31). Indeed, we showed that sgaDC expression is enhanced in the presence of the IncC plasmid (Figure 6D). Likewise, expression of acaDC on the IncC plasmid was also enhanced by the presence of SGI1, likely through direct activation of acaB expression by SgaCD, which consecutively activates acaDC expression (26). Nevertheless, several properties of sgaDC and acaDC expression remain to be characterised to fully understand the crosstalk between SGI1 and its helper plasmid, including the conditions of activation and relative strength of sgaDC and acaDC promoters, the half-life of the corresponding mRNA transcripts, their translation rates, and the relative stability of each activator complex. Although expression of both acaDC and sgcDC increases in IncC SGI1 cells, AcaCD’s role could become negligible and even dispensable if sgaDC mRNA is more stable and translated more efficiently, or if the SgaCD complex turnover is slower than that of AcaCD. Furthermore, we cannot rule out that chimerical activator complexes SgaC-AcaD or AcaC-SgaD also form and play a role in the regulation of gene activation on the helper plasmid and on SGI1. If such chimera exist, and if SgaCD and AcaCD tend to form heterohexamers like E. coli FlhD4C2 (74), then a heterogeneous bestiary of activator complexes regulates gene expression when SGI1 and its helper plasmid occupy the same cell.

Incompatibility between SGI1 and IncC plasmids could have emerged as a defence mechanism deployed by SGI1 to prevent excessive activation of excision and replication, which ultimately results in SGI1 loss in the absence of selective pressure (Figure 7). Destabilisation of the plasmid would allow the island to revert to its quiescent state after transferring to a new host, ensuring its stability by residing integrated in the host’s chromosome. Although the exact mechanism lying underneath the destabilisation is still unknown, we recently suggested titration of endogenous replication proteins (31). Alternatively, SGI1 could be interfering with the partitioning of the IncC plasmid in daughter cells. Being single-copy large plasmids (Figure 6C), IncC replicons are likely extremely susceptible to perturbation of their partition process. In fact, IncC plasmids encode two partitioning systems, a type I parABS system that has been shown to be essential for plasmid maintenance and a putative type II parMRC-like partitioning system dis-
stantly related to srpRMC of SXT/R391 ICEs and encoded by vcrx151–152 (77,78). We show here that expression of vcrx151–152 is activated by AcaCD and SgaCD. Remarkably, forced excision and replication of SG11 via overexpression of acaDC or sgaDC in the absence of an IncC plasmid resulted in rapid elimination of SG11 from the cell population (Figure 7D, E). In stark contrast, when the plasmid was present, SG11 persisted and reintegrated into the chromosome while promoting IncC plasmid loss (Figure 7H). This observation hints at SG11 taking control of the IncC partitioning system, perhaps to enhance its equal segregation into daughter cells during cell division. These lines of inquiry will need to be examined in detail to better understand the processes at stake.

Despite the divergence of the primary sequences of AcaC and SgaC, we showed that both transcriptional activator complexes bind to the same DNA motif. Other conjugative plasmids that confer multidrug resistance to marine-dwelling bacteria and have not yet been ascribed to an incompatibility group also encode AcaCD orthologues. The C-terminus of the C subunits encoded by these plasmids (e.g. AqaC or AsaC) diverge significantly from AcaC and SgaC. Nevertheless, their Zinc-finger domain is relatively well conserved (Supplementary Figure 5F-G). Unsurprisingly, these plasmids contain AcaCD-like binding sites upstream genes involved in conjugal transfer (Supplementary Table S5). Based on these observations, the breadth of IncC/SG11-like interactions is likely broader than anticipated, with SG11 and its variants being possibly activated and mobilized by such plasmids that do not belong to the IncA nor the IncC group (17). Furthermore, since AcaCD-responsive promoters have recently been found in MGIs that integrate at trmE, vicC and dusA in the chromosome of several species of Gammaproteobacteria (9), our observations support a complex network of mobilisation events involving diverse families of MGIs and conjugative plasmids of multiple incompatibility groups (IncA, IncC and untyped) encoding AcaCD-like transcriptional activators with identical DNA-binding motifs.

Interactions between members of the SG11 family and their helper plasmids are undoubtedly complex. The use of naturally occurring variants such as SG11-I, SG11-F, SG11-I, SG11-K or SG12, with diverse naturally occurring IncC and even IncA plasmids, while providing interesting hints and clues, also renders comparisons between studies challenging and could lead to erroneous conclusions. These shortcomings result from SNPs that may affect the expression of key effectors that are important for the regulation, replication, stability, and conjugative transfer of the two interacting partners. Considering the discrepancies between others’ results and ours, we urge the establishment of a reliable, robust system based on a selected subset of model SG11 variants and helper plasmids to properly decipher the complex biology and interactions between SG11 and IncC plasmids.

DATA AVAILABILITY

Complete data from aligned reads for ChIP-Exo, Cappable-Seq and RNA-Seq experiments can also be visualized using the UCSC genome browser at http://bioinfo.ccs.usherbrooke.ca/sgaCD.html.

Raw sequencing data were submitted to Genbank under Bioproject accession number PRJNA648047 with the following Biosample accession numbers: for ChiP-exo assays, from SAMN15617565 to SAMN15617572, respectively; for Cappable-Seq assays, from SAMN15617573 to SAMN15617580; for RNA-Seq assays, from SAMN15617581 to SAMN15617604.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Dr. Michael Mulvey (National Microbiology Laboratory, Public Health Agency of Canada) for providing us with S. enterica Agona 0047SA97 and Typhimurium DT104 S/954435. We thank Calcul Quebec (www.calculquebec.ca) and Compute Canada (www.computecanada.ca) for access to bioinformatics resources and support, as well as Jean-François Lucier for his help with bioinformatics analyses. We also thank the Centre de calcul scientifique de l’Université de Sherbrooke for technical assistance and the Université de Sherbrooke RNomics Plateform for assistance with Illumina sequencing. We are grateful to Alain Laviguer for his insightful comments on the manuscript and to Dominick Matteau for thoughtful discussions and advice.

FUNDING

Discovery Grant [2016-04365 to V.B.] from the Natural Sciences and Engineering Research Council of Canada (NSERC); Project Grant [PJT-153071 to V.B.] from the Canadian Institutes of Health Research (CIHR); R.D. is the recipient of a Fonds de recherche du Québec-Nature et Technologies (FRQNT) doctoral fellowship; K.T.H. was supported by a postdoctoral fellowship [SPE20170336797] from the Fondation de la Recherche Médicale (FRM, France); N.R. is the recipient of a Fonds de recherche du Québec-Santé (FRQS) MSc scholarship; Alexander Graham Bell Canada Graduate Scholarship from the NSERC. Funding for open access charge: CIHR [PJT-153071].

Conflict of interest statement. None declared.

REFERENCES

1. CDC (2019) In: Antibiotic Resistance Threats in the United States, 2019. U.S. Department of Health and Human Services, CDC, Atlanta, GA.
2. MacLean, R.C. and Millan, A.S. (2019) The evolution of antibiotic resistance. Science, 365, 1082–1083.
3. Guédon, G., Libante, V., Coluzzi, C., Payot, S. and Leblond-Bourget, N. (2017) The obscure world of integrative and mobilizable elements, highly widespread elements that pirate bacterial conjugative systems. Genes, 8, 337.
4. Bellanger, X., Payot, S., Leblond-Bourget, N. and Guédon, G. (2014) Conjugative and mobilizable genomic islands in bacteria: evolution and diversity. FEMS Microbiol. Rev., 38, 720–760.
5. Curraro, N., Rivard, N., Burrus, V. and Cecarelli, D. (2017) Mobilizable genomic islands, different strategies for the dissemination of multidrug resistance and other adaptive traits. Mob Genet Elements, 7, 1–6.
resistance regions from serovars Typhimurium DT104 and Agona. *Antimicrob. Agents Chemother.*, 46, 1714–1722.

44. Dower, W.J., Miller, J.F. and Ragsdale, C.W. (1988) High efficiency transformation of E. coli by high voltage electroporation. *Nucleic Acids Res.*, 16, 6127–6145.

45. Humbert, M., Huguet, K.T., Coulombe, F. and Burrys, V. (2019) Entry exclusion of conjugative plasmids of the IncA, IncC, and related untypied incompatibility groups. *J. Bacteriol.*, 201, e00731-18.

46. Datsenko, K.A. and Wanner, B.L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U.S.A.*, 97, 6640–6645.

47. Hallidman, A. and Wanner, B.L. (2001) Conditional-replication, integration, excision, and retrieval plasmid-host systems for gene structure-function studies of bacteria. *J. Bacteriol.*, 183, 6384–6393.

48. Ettwiller, L., Buswell, J., Yigit, E. and Schildkraut, I. (2016) A novel enrichment strategy reveals unprecedented number of novel transcription start sites at single base resolution in a model prokaryote and the gut microbiome. *BMC Genomics*, 17, 199.

49. Christodoulou, D.C., Gorham, J.M., Herman, D.S. and Seidman, J.G. (2011) Construction of normalized RNA-seq libraries for next-generation sequencing using the crab duplex-specific nuclease. *Curr. Protoc. Mol. Biol.*, doi:10.1002/0471142727.mb041294.

50. Bolger, A.M., Lohse, M. and Usadel, B. (2014) Trimomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30, 2114–2120.

51. Andrews, S. (2010) FastQC: a quality control tool for high throughput sequence data.

52. Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods*, 9, 375–379.

53. Lassmann, T., Hayashizaki, Y. and Dauh,C.O. (2011) SAMStat: monitoring biases in next generation sequencing data. *Bioinformatics*, 27, 130–131.

54. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. and Durbin, R. (2009) The sequence alignment/map format and SAMtools. *Bioinformatics*, 25, 2078–2079.

55. Quinlan, A.R. and Hall, I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, 26, 841–842.

56. Coulombe, C., Pottraz, C., Nordell-Markovits, A., Brunelle, M., Lavoie, M.-A., Robert, F. and Jacques, P.-E. (2014) VAP: a versatile aggregate profiler for efficient genome-wide data representation and discovery. *Nucleic Acids Res.*, 42, W485–W493.

57. Magoc, T., Wood, D. and Salzberg, S.L. (2013) EDGE-pro: estimated degree of gene expression in prokaryotic genomes. *Evol. Bioinform. Online*, 9, 127–136.

58. Love, M.I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, 15, 50.

59. Luo, W., Friedman, M.S., Shedden, K., Hankenson, K.D. and Woolf, P.J. (2009) GAGE: generally applicable gene set enrichment for pathway analysis. *BMC Bioinformatics*, 10, 161.

60. Aßchul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.*, 215, 403–410.

61. Miller, J.H. (1992) In: *A Short Course in Bacterial Genetics: A Laboratory Manual and Handbook for Escherichia coli and Related Bacteria*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

62. Hellemans, J., Mortier, G., De Paepe, A., Speleman, F. and Vandesompele, J. (2007) qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol.*, 8, R19.

63. Browning, D.F. and Busby, S.J. (2004) The regulation of bacterial transcription initiation. *Nat. Rev. Microbiol.*, 2, 57–65.

64. Latif, H., Federowicz, S., Ebrahim, A., Tarasova, J., Szubin, R., Utrilla, J., Zengler, K. and Palsson, B.O. (2018) ChiP-exo interrogation of Crp, DNA, and RNAP holoenzyme interactions. *PLoS One*, 13, e0197272.

65. Hegyi, A., Szabó, M., Olasz, F. and Kiss, J. (2017) Identification of oriT and a recombination hot spot in the IncA/C plasmid backbone. *Sci. Rep.*, 7, 10595.

66. Carraro, N., Sayvü, M., Matteu, D., Lauzon, G., Rodrigue, S. and Burrys, V. (2014) Development of pVCR94ΔX from *Vibrio cholerae*, a prototype for studying multidrug resistant IncA/C conjugative plasmids. *Front. Microbiol.*, 5, 44.