Title: Dynamics of the SetCD-regulated integration and excision of Genomic Islands mobilized by Integrating Conjugative Elements of the SXT/R391 family.

Running title: Control of integration and excision of MGIs

Authors and Affiliations: Aurélie Daccord, Mathias Mursell, Dominic Poulin-Laprade and Vincent Burrus

Centre d’Étude et de Valorisation de la Diversité Microbienne (CEVDM), Département de biologie, Université de Sherbrooke, Sherbrooke, Québec, Canada

#Corresponding author:

Mailing address:
Département de biologie, Université de Sherbrooke, 2500 boulevard de l’Université, Sherbrooke, Québec, Canada J1K 2R1

Phone: (819) 821-8000 ext 65223

Fax: (819) 821-8049

E-mail: Vincent.Burrus@USherbrooke.ca
Abstract

Mobilizable Genomic Islands (MGIs) are small genomic islands that are mobilizable by SXT/R391 Integrating and Conjugative Elements (ICEs) due to a similar origin of transfer. Their site-specific integration and excision is catalyzed by the integrase they encode but their conjugative transfer entirely depends upon the conjugative machinery of SXT/R391 ICEs. In this study we report the mechanisms that control the excision and integration processes of MGIs. We found that while the MGI-encoded integrase IntMGI is sufficient to promote MGI’s integration, efficient excision from the host’s chromosome requires the combined action of IntMGI and of a novel recombination directionality factor, RdfM. We determined that both genes are activated by SetCD, the main transcriptional activators of SXT/R391 ICEs. Although they share the same regulators, we found that unlike rdfM, intMGI has a basal level of expression in the absence of SetCD. These findings explain how an MGI can integrate into the chromosome of a new host in the absence of a co-resident ICE and shed new light on the crosstalk that can occur between mobilizable and mobilizing elements that mobilize them, helping to understand part of the rules that dictate horizontal transfer mechanisms.
Introduction

Horizontal gene transfer plays a fundamental role in bacterial evolution (22, 24, 28, 32, 33). By transferring from a bacterial genome to another, mobile genetic elements allow bacteria to acquire new DNA fragments encoding a wide array of new functions. Genomic islands (GIs) are mobile genetic elements that play a fundamental role in horizontal gene transfer (26). GIs are DNA segments (10-550 kb) that are often associated with tRNA genes and exhibit a G+C content usually different from the surrounding chromosome (16, 26). Based upon the functions they encode, GIs are also known as pathogenicity, symbiosis, metabolic, resistance or fitness islands (26).

Integrating conjugative elements (ICEs) are self-transmissible GIs found in many Gram-positive and Gram-negative bacteria (8, 11, 36-38, 42). ICEs confer a variety of functions to their host such as virulence factors, establishment of symbiosis, new metabolic traits, resistance to antibiotics and factors that enhance bacterial fitness (11). ICEs transfer via conjugation in a conjugative plasmid-like manner, and like many temperate bacteriophages they integrate into their host’s chromosome along which they are replicated. The well-studied family of SXT/R391 ICEs includes more than 30 members that are found mostly in clinical and environmental Vibrio strains as well as in several other γ-proteobacterial species (7). SXT/R391 ICEs share a conserved set of 52 genes with nearly half of them encoding proteins necessary for conjugation, integration/excision and regulation (Fig. 1A) (41). They integrate by site-specific recombination into the 5’ end of prfC, a non-essential gene involved in the termination of translation (25). While integration and excision of SXT/R391 ICEs is catalyzed by the site-specific tyrosine recombinase IntSXT, their excision from the chromosome is
facilitated by the recombination directionality factor (RDF) Xis (10). Conjugative transfer of SXT/R391 ICEs is initiated at a *cis*-acting locus called the origin of transfer \((oriT_{SXT})\) by the putative relaxase TraI and the auxiliary mobilization protein MobI that likely form together a nucleoprotein complex called the relaxosome (12). By analogy with conjugative plasmids, translocation of the ICE DNA through the membranes of the donor and the recipient cell is thought to occur as a linear single-stranded DNA molecule covalently bound to TraI (4). Once in the recipient cell, the ICE DNA is recircularized and its complementary strand is synthesized prior to integration into the chromosome. Regulation of excision and transfer of SXT/R391 ICEs is controlled by \(setR\) which encodes a \(\lambda\) CI-like transcriptional repressor that represses the expression of \(setCD\) (4, 5). \(SetCD\) is a transcriptional activator complex that triggers the expression of all the genes involved in integration, excision, and conjugative transfer. \(SetR\) repression of \(setCD\) expression is alleviated by DNA damages (5), allowing \(SetCD\) to activate excision and transfer of the ICE.

Besides ICEs and bacteriophages, the vast majority of GIs does not have any known mechanism of transfer and are therefore considered as non-self-transmissible. However, such GIs typically harbor functional or cryptic genes that encode site-specific recombinases (integrases) or transposases. Their mechanisms of transfer likely involve the participation of mobilizing self-transmissible elements, such as generalized transducing phages, conjugative plasmids or ICEs (6). We have recently identified in several genomes of *Vibrio* a new family of GIs that rely on a unique mechanism for gene transfer (13). These Mobilizable Genomic Islands (MGIs) have a size of less than 25 kb and can be mobilized at high frequency by SXT/R391 ICEs using a *cis*-acting \(oriT\)
sequence that mimics oriT_{SXT}. MGIs integrate into the 3’ end of yicC, a conserved gene encoding a putative stress induced protein (13). MGIs’ integration is catalyzed by the site-specific recombinase Int_{MGI}, a distant relative of Int_{SXT}. Besides int_{MGI} and the oriT_{SXT}-like oriT_{MGI} sequence, all MGIs identified to date share only three conserved genes (Fig. 1B), none of which are predicted to encode components of a conjugative transfer machinery or an RDF. Interestingly, while MGIs’ excision is independent of int_{SXT} and xis, it requires the presence of the ICE-encoded SetCD transcriptional activators (13).

In this study, we report the identification of the new RDF RdfM which is required for MGIs chromosomal excision. Like int_{MGI}, expression of rdfM is activated by SetCD. Comparison of the regulation of the integration/excision genes of SXT/R391 ICEs and of those of MGIs revealed that they are similarly regulated by SetCD in the donor cells; yet int_{MGI} is expressed independently of SetCD in the recipient cells, allowing MGIs to integrate into the chromosome of a cell lacking an SXT/R391 ICE. To the best of our knowledge, this is the first report of such an intimate interaction between two unrelated families of mobile genetic elements.
98 Materials and Methods
99
100 Bacterial strains and media
The bacterial strains and plasmids used in this study are described in Table 1. The
101 strains were routinely grown in Luria-Bertani (LB) broth at 37°C in an orbital
102 shaker/incubator and were maintained at -80°C in LB broth containing 15% (vol/vol)
103 glycerol. Antibiotics were used at the following concentrations: ampicillin (Ap), 100
104 µg/ml; kanamycin (Kn), 50 µg/ml; rifampicin (Rf), 50 µg/ml; spectinomycin (Sp), 50
105 µg/ml; sulfamethoxazole (Su), 160 µg/ml; tetracycline (Tc), 12 µg/ml; and trimethoprim
106 (Tm), 32 µg/ml. When required, bacterial cultures were supplemented with 0.3mM DL-
107 α,ε-diaminopimelic acid (DAP), 100 ng/ml mitomycin C, or 0.02% L-arabinose.
108
109 Plasmids and strain constructions
110 Plasmids and primers used in this study are described in Table 1 and 2, respectively.
111 Plasmid pVB200 was constructed by sub-cloning of XbaI-flanked digestion product
112 attP_MGI/Vfl
113 Ind1 into the XbaI site of pSW23T. Product attP_MGI/Vfl
114 Ind1 was amplified using
115 genomic DNA of V. fluvialis H-08942 as a template and primer pair attPAD-L1/attPAD-
116 R1-AC for the first round, attPAD-L2/attPAD-R2-AC for the second round and then
117 cloned into vector pCR2.1-TOPO (Invitrogen). Plasmid p9, p8 and p9-8 were constructed
118 by cloning cds9_MGI/VvuTai1, cds8_MGI/VvuTai1 or cds9-8_MGI/Vfl
119 Ind1 into the TA cloning expression
120 vector pBAD-TOPO (Invitrogen) according to the manufacturer’s instructions.
121 cds9_MGI/VvuTai1, cds8_MGI/VvuTai1 and cds9-8_MGI/Vfl
122 Ind1 were amplified by PCR from genomic
123 DNA of V. vulnificus YJ016 or V. fluvialis H-08942 as a template using primers pairs
124 AD4-V-F/AD4-R1, AD5-F/AD5-V-R1 and AD5-A-R1/AD4-A-F, respectively (Table 2).
All deletion mutants were constructed in *E. coli* AD57 using the one-step chromosomal gene inactivation technique (14). All mutations were designed to be non-polar. The Δ*cds4*, Δ*cds8* and Δ*cds9* mutations were introduced in MGI/VflInd1 using primer pairs AD13-WF/AD13-WR, Gene8-WF/Gene8-WR, and AD11-WF/AD11-WR (Table 2) respectively, and pKD13 as the template. All deletion mutations were verified by PCR amplification using primers flanking the deletion.

**Bacterial conjugation**

Conjugation assays were used to transfer SXT, R997, MGI/VflInd1 or plasmids into *E. coli*. Mating assays were performed by mixing equal volumes of overnight cultures of donor and recipient strains. The cells were harvested by centrifugation and resuspended in a 1/20 volume of LB broth. Cell suspensions were poured onto LB agar plates and incubated at 37°C for 6h. The cells were then resuspended in 1 ml of LB medium, and serial dilutions were plated onto appropriate selective media to determine the number of donors, recipients, and exconjugants. Frequency of transfer was expressed as the number of exconjugant cells per recipient cells in the mating mixture at the time of plating. *E. coli* CAG18439, MC4100 λpir or VB112 were used as recipients in conjugation experiments (Table 1). To induce expression of IntMGI from pIntVvu, SetCD from pGG2B, protein 8 from p8, protein 9 (RdfM) from p9, or proteins 9 and 8 from p9-8 (Table 1) in complementation assays, mating experiments were carried out on LB-agar plates supplemented with 0.02% L-arabinose.

**Molecular biology techniques**

All the enzymes were used according to manufacturer’s instructions (New England Biolabs). Plasmid DNA was prepared with a QIAPrep spin mini prep kit (Qiagen), and
chromosomal DNA was prepared with a Wizard Genomic DNA purification kit (Promega) as described in the manufacturer’s instructions.

PCR assays were carried out in 50-µl PCR mixtures with 1U of Taq DNA polymerase (New England Biolabs). The PCR conditions were as follows: (i) 3 min at 94°C; (ii) 30 cycles of 30 s at 94°C, 30 s at a suitable annealing temperature, and 30 s to 60 s at 72°C; and (iii) 5 min at 72°C. When needed, PCR products were purified using a QIAquick PCR purification kit (Qiagen) according to the manufacturer’s instructions. The purified PCR products or inserts of constructed plasmids were sequenced by Centre d’Innovation Génome Québec (McGill University, Montréal, Québec, Canada). DNA sequences were compared with the GenBank DNA sequence database using the BLASTN program (3).

E. coli was transformed by electroporation in 1 mm-gap cuvettes according to Dower et al. (18), using a GenePulser Xcell apparatus (Bio-Rad) set at 25µF, 200 Ω, and 1.8 kV.

**Real-time quantitative PCR assays for relative quantification of attB and rph**

Real-time quantitative PCR assays were used to measure the percentages of cells in a culture that contained unoccupied MGI attB site (the 3’ end of yicC) as described previously (13). Briefly, this corresponds to a comparison of the amount of excised circularized MGI relatively to the amount of chromosome copies deduced from the amplification of rph, a gene located immediately 5’ of yicC. Primer pairs Q-PCR-1F/Q-PCR-1R and Q-PCR-2F/Q-PCR-2R were used for the amplification of attB and rph, respectively (Table 2).

**RNA extraction and cDNA synthesis**

Bacterial cultures were grown at 37°C to early exponential phase (OD600=0.2). Cultures were split in two and induction was initiated by addition of 100 ng/ml...
mitomycin C or 0.02% L-arabinose. Two hours after induction, aliquots of bacterial
cultures were directly mixed with RNA Protect Bacteria Reagent (Qiagen) and treated
according to manufacturer’s instructions. Bacterial RNA was isolated after treating the
cells with lysosyme (Sigma), using the RNeasy Mini Kit (Qiagen). In addition, RNA
samples were treated with DNase (RNase-free DNase set, Qiagen) during purification
and Turbo DNase (Ambion) after purification. RNA purity and concentration were
evaluated with a ND-1000 NanoDrop Spectrophotometer (Thermo Fisher
Scientific/NanoDrop Products). cDNA was prepared using the SuperScript II (Invitrogen)
following the manufacturer’s recommendations. 50 ng of random hexamers (Integrated
DNA Technologies) and 1 µg of total bacterial RNA were used in each reaction. After
synthesis cDNA sample mixtures were purified with the PCR Purification Kit (Qiagen)
and stored at -20°C.

**Reverse transcription quantitative PCR**

The MasterCycler ep realplex4 sequence detection system (Eppendorf) was used to
quantify the increase in fluorescence emission of SYBR Green I during PCR. The
realplex software (version 1.5; Eppendorf) was used for data acquisition and analysis.

Each 25-µl reaction mixture contained 12.5 µl of 2× SYBR Green PCR Master Mix
(Qiagen), 1 µM of each primer, and 1 µl of cDNA template. Primer pairs RTgene9-F
b/RTgene9-R b, RTintVf-F/ RTintVf-R, RTrpoZcoli-F/ RTrpoZcoli-R and RTyicC-
F/RTyicC-R were used for the amplification of *cds9*<sup>MGI</sup>*MGI<br>Int1*, *int*<sup>MGI</sup>*Int1*, *rpoZ* and *yicC*,
respectively (Table 2). The PCR conditions were (i) 5 min at 95°C, (ii) 45 cycles of 10 s
at 95°C and 30 s at 60°C, (iii) 15 s at 95°C, (iv) 15 s at 60°C, (v) melting curve from
60°C to 95°C and (vi) 15 s at 95°C. Three reactions were performed for each sample. For
normalization, the rpoZ gene was used and results expressed as relative expression based on the ΔΔCt calculation method. Experiments were carried out three times and combined.
Results

Int\text{MGI} is the only MGI-encoded protein necessary for MGI integration

In a previous study we showed that Int\text{MGI} is required for integration and excision of MGI/V/Ind1 (13). To examine whether Int\text{MGI} is the sole MGI-encoded protein necessary to mediate MGI’s integration into the 3’ end of yicC, Int\text{MGI}-mediated recombination between att\text{P}_{MGI} and attB at yicC was monitored using pVB200, a derivative of the mobilizable Cm’ suicide vector pSW23T harboring att\text{P}_{MGI} (Fig. 2A). Since pSW23T requires the product of pir to replicate, Cm’ exconjugants can be isolated after its conjugative transfer from a pir’ host to a pir’ host only if it has integrated into the chromosome.

As a negative control, we first mobilized empty pSW23T from a mob^+ pir’ donor strain to CAG18439 or CAG18439 harboring plIntVvu (AD208), a plasmid expressing Int\text{MGI} under the control of an arabinose inducible promoter. In both cases, the frequency of exconjugant formation was below 5×10^{-6} exconjugants/recipient, a value we established as our baseline for subsequent experiments (Fig. 2B). The few recovered exconjugants can be attributed to random integration of the plasmid into the recipient chromosome. As a positive control we also mobilized pVB200 from the same donor strain to a pir’ strain (MC4100 λpir) to verify that the constructed plasmid remained mobilizable. We found that up to 39% of the pir’ recipient cells acquired and maintained the plasmid. We then mobilized pVB200 from the same donor strain to CAG18439 or AD208 (Fig. 2B). When Int\text{MGI} was expressed in the pir’ recipient strain, the frequency of exconjugant formation was as high as the positive control, indicating that the plasmid was able to maintain by site-specific integration into the recipient’s chromosome. Thus, we
conclude that Int_{MGI} is the only MGI-encoded protein needed to mediate efficient integration of MGIs.

**MGI's integration does not require activation by SetCD**

In our initial study of MGIs, we showed that the SXT/R391 ICE-encoded transcripational activator SetCD is necessary for MGI’s excision from the chromosome, suggesting that it is required to activate the expression of Int_{MGI} (13). Surprisingly, we also observed that colonies harboring an MGI but devoid of any ICE were also recovered at high frequency in exconjugant populations. This observation is consistent with the natural occurrence of environmental and clinical isolates (*V. cholerae* RC385 and *V. vulnificus* YJ016) having similar configurations (13) and suggests that while SetCD is necessary for excision, it is not required for *de novo* expression of Int_{MGI} in the recipient cells. To test this hypothesis, we mobilized pVB200 into CAG18439 harboring MGI/VflInd1 alone (AD130) or along with either R997 (AD207), an Ap'-conferring ICE of the SXT/R391 family, or pGG2B (AD132), a plasmid expressing SetCD under control of an arabinose inducible promoter (Fig. 2C). Interestingly, exconjugants formed at high frequency in the sole presence of MGI/VflInd1 whereas the presence of R997 or expression of SetCD in the recipient cells did not significantly improve transfer and maintenance of pVB200 supporting the notion that SXT/R391 ICEs and SetCD are necessary for MGI’s excision but not its integration.

**Int_{MGI} alone does not promote efficient excision of MGIs**

Next we examined whether Int_{MGI} alone was able to promote efficient excision of pVB200 integrated into *yicC*. We used a semi-quantitative PCR assay to detect unoccupied attB sites in the cell populations compared to *rph* as a reference target. The
formation of attB site was tested in CAG18439 as a positive control, and in CAG18439
harboring attB::pVB200 along with pIntVvu (AD217), MGI/V/Ind1 (AD130),
MGI/V/Ind1 and R997 (AD207), or MGI/V/Ind1 and pGG2B (AD132). We found that
IntMGI alone did not mediate efficient excision, even when overexpressed (Fig. 2D, lane
1-3). In fact, excision was only detectable in the presence of MGI/V/Ind1 either along
with R997 or upon expression of SetCD from pGG2B (Fig. 2D, lane 4, 5 and 7). These
results led us to suppose that an unidentified MGI-encoded factor likely helps IntMGI to
mediate efficient site-specific excision and that expression of this factor is likely
activated by SetCD.

MGIs encode a putative RDF

Considering that IntMGI is required but not sufficient to promote efficient excision, we
looked at the genes conserved among sequenced MGIs to identify an RDF that could
facilitate the IntMGI-mediated excision of MGIs. RDFs control the directionality of
tyrosine recombinase-mediated site-specific recombination events (30), and are usually
small basic proteins (<100 amino acids) with or without a putative helix-turn-helix
(HTH) DNA-binding motif. Besides intMGI, only 3 genes are common to all MGIs
identified and sequenced to date: cds4 encodes a 214-amino acid protein of unknown
function, cds8 encodes a 580-amino acid putative helicase and cds9 encodes an 80-amino
acid predicted transcriptional regulator (Fig. 1B). Interestingly, the translation product of
cds9 shares 36% identity with Hef encoded by the High Pathogenicity Island (HPI) of
Yersinia pseudotuberculosis and 29% identity with AlpA encoded by E. coli prophage
CP4-57 (Fig. 3). Hef has been reported to act as a RDF (29) whereas AlpA has been
reported to activate the expression of its cognate integrase gene (27). Given the size of
the predicted translation product of *cds9*, and its similarity with Hef, we considered it to be a good candidate for a putative RDF. Yet, given its similarity with the transcriptional regulator AlpA, we could not rule out at this point the possibility that the product of *cds9* could activate the expression of *intMGI*.

Deletion of *cds9* dramatically affects MGI’s excision and transfer

To assess whether one of the 3 MGI’s conserved genes could act as an RDF we first constructed deletion mutants of each gene in MGI/ffInd1. We tested the ability of each mutant to be mobilized by ICE/ffInd1. While deletions of *cds4* or *cds8* had virtually no impact, deletion of *cds9* led to a dramatic reduction of the MGI frequency of transfer (Fig. 4). Mobilization of MGI/ffInd1 Δ*cds9* could be partially restored when *cds9* was expressed *in trans* from an inducible promoter. Complementation with *cds8* did not restore MGI transfer, whereas complementation with *cds9-cds8* expressed from the same inducible promoter restored mobilization to the same level as *cds9* alone, ruling out the possible polar effects of the Δ*cds9* deletion on *cds8* that could have explained the partial complementation phenotype observed upon *cds9* overexpression.

To investigate whether the reduction of MGI/ffInd1 transfer was a consequence of reduced or abolished excision caused by deletion of *cds9*, we conducted real-time quantitative PCR assays to measure the percentage of cells in a culture containing unoccupied *attB* sites and found that excision of MGI/ffInd1 Δ*cds9* was undetectable (Fig. 4). On the opposite, excision of the same mutant was dramatically enhanced (50-fold wild-type level) when *cds9* was expressed *in trans*. Interestingly, the rate of excision of the mutant was restored to wild-type level when *cds9* and *cds8* were expressed together *in trans*, suggesting a possible regulatory activity of the protein encoded by
The product of *cds9* acts as an RDF, not as a transcriptional activator

AlpA was shown to activate the expression of the integrase gene of the cryptic prophase CP4-57 in *E. coli* (27). This ability prompted us to investigate whether the product of *cds9* could act as a transcriptional activator of *int*MGI. Expression of *int*MGI was measured by reverse transcription real-time quantitative PCR in the presence or absence of mitomycin C and in different genetic backgrounds, including cells devoid of ICE (AD130), cells containing an ICE and a wild-type (AD72) or Δ*cds9* (AD167) copy of MGI/VflInd1, or cells containing both ICE/VflInd1 and MGI/VflInd1 and expressing *cds9* from an inducible promoter (Fig. 5A). We found that *int*MGI expression was strongly stimulated by the addition of the DNA-damaging agent mitomycin C, but only in the presence of an SXT/R391 ICE. The absence of *cds9* had no effect on the mitomycin C-induced activation of *int*MGI expression. Overexpression of *cds9* in the absence of mitomycin C induction led to a slight yet non-significant activation of *int*MGI expression. This barely detectable level of activation is probably not dependent upon *cds9* expression, but rather the result of the expression of SetCD in a subpopulation of cells inherently expressing the SOS response, as reported by McCool *et al.* (31). This phenomenon also explains the constitutive basal level of transfer of SXT/R391 ICE in the absence of DNA-damaging agents. These results combined with our previous observations on SetCD-mediated activation of MGI excision indicate that the product of *cds9* acts as a RDF rather than an activator of *int*MGI expression. From now on, *cds9* will therefore be referred to as *rdfM* for “recombination directionality factor of MGI”.
Expression of both \textit{int}_{MGI} and \textit{rdfM} is activated by SetCD

Given that SetCD activates MGI excision (13) and that DNA-damaging agents stimulate \textit{int}_{MGI} expression, we hypothesized that SetCD acts as transcriptional activators of both \textit{int}_{MGI} and \textit{rdfM}. To verify this hypothesis, we measured the expression of \textit{int}_{MGI} and \textit{rdfM} in \textit{E. coli} cells carrying MGI/VflInd1 and SXT (AD72), SXT Δ\textit{setCD} (AD133) or expressing SetCD from pGG2B (AD132). We also measured the expression of \textit{yicC} in the same cells since it is described in Genbank as a gene coding for a putative stress-induced protein and the relative position and orientation of \textit{yicC} and \textit{int}_{MGI} suggest that both genes may be co-transcribed. Induction was carried out with either mitomycin C (AD72 and AD133) or L-arabinose (AD132). First, we observed that neither SetCD nor mitomycin C modulates the expression of \textit{yicC}, ruling out the possible expression of \textit{int}_{MGI} from the promoter of \textit{yicC} when the MGI is integrated into its 3’ end (Fig. 5B). In contrast, in the presence of wild-type SXT, mitomycin C was found to stimulate the expression of both \textit{int}_{MGI} and \textit{rdfM} (24- and 44-fold increase, respectively), whereas it had no effect in the presence of the SXT Δ\textit{setCD} mutant. This stimulation of expression is attributable to the increased expression of SetCD from SXT, as the expression of SetCD from pGG2B in a strain lacking SXT resulted in an increase of ~300 and ~2,000 fold of the transcript levels of \textit{int}_{MGI} and \textit{rdfM}, respectively.

\textit{int}_{MGI} is constitutively expressed, allowing MGI integration in a strain devoid of an SXT/R391 ICE

We previously reported that upon mating an \textit{E. coli} donor strain harboring MGI/VflInd1 and ICE/VflInd1 ~98% of the isolated exconjugant colonies selected for the MGI were devoid of ICE/VflInd1, highlighting the independence of MGIs from ICEs for...
their integration into the chromosome (13). This result is supported by naturally occurring isolates containing MGIs but devoid of SXT/R391 ICEs (13). However, it contrasts with our aforementioned expression results indicating that SetCD activates the expression of intMGI. To explain how MGIs integrates into yicC in the absence of ICE-encoded SetCD, we had a closer look at intMGI expression data in non-induced conditions which revealed that intMGI has a low-level constitutive expression. In the presence of SXT (AD72), intMGI and rdfM exhibit detectable $2^{\Delta Ct}$ values of 0.041 and 0.01 relatively to rpoZ, respectively (Fig. 6). This level of expression is most likely a consequence of spontaneous induction of the SOS response (31). When the same experiment was carried out using SXT ΔsetCD (AD133), rdfM expression was below the limit of detection whereas expression of intMGI was only reduced by 36% (Fig. 6). This result, which is also supported by the high rate of exconjugant formation upon mobilization of pVI200 to a strain containing MGI/VflInd1 but lacking an ICE (Fig. 2C), indicates that intMGI is constitutively expressed at a low level in the absence of SetCD. This basal level of expression is necessary and sufficient to promote integration of MGIs into the chromosome of a new host in the absence of a helper SXT/R391 ICE.
In this study we investigated the integration and excision dynamics of MGIs. We found that while the MGI-encoded integrase alone is sufficient to promote efficient MGI’s integration into the chromosome, it also requires the MGI-encoded RDF RDFM to promote efficient MGI’s excision. We found that both intMGi and rdfM are activated by the SXT/R391 ICE-encoded transcriptional regulators SetCD. However, the expression of intMGi does not strictly require SetCD whereas the expression of rdfM does. These findings help to establish how an MGI cannot excise from the chromosome of a cell devoid of an SXT/R391 ICE but can integrate into the chromosome of such a cell. Accordingly, we propose a model of the regulation pathways responsible for the excision and integration processes of MGIs in the donor and the recipient cells (Fig. 7).

Integration and excision are critical steps in the maintenance and dissemination of an integrative mobile genetic element, whether it is a temperate bacteriophage, an ICE or a mobile genomic island. Site-specific integration typically requires the action of a single mobile element-encoded site-specific recombinase and may require the help of host-encoded nucleoid proteins, such as the integration host factor (IHF) and the factor for inversion stimulation Fis (reviewed in references 20, 21). In contrast, the reverse recombination event, the site-specific excision usually requires an additional genetic element-encoded protein, the RDF, aka Xis/excisionase although it usually lacks a catalytic activity per se. RDF are usually small basic proteins that play architectural roles in the recombination events catalyzed by their cognate tyrosine or serine recombinase. Given their small size RDF-encoding genes can be difficult to identify and since a subset...
of RDFs harbor putative helix-turn-helix domains they are often annotated as putative transcriptional regulators.

Lewis and Hatfull divided RDFs in 11 subgroups based on sequence similarity (30). We showed here that RdfM belongs to the SLP1 subgroup of RDFs and as such has a putative conserved N-terminal HTH DNA binding motif (HTH1) (Fig. 3). Peculiarly, unlike other members of this subgroup, RdfM is predicted to contain a second C-terminal HTH DNA-binding motif (HTH2) (Fig. 3), the role of which remains to be determined as our results indicate that RdfM is not a transcriptional regulator of expression of \textit{int}_{\text{MGi}}.

This strongly contrasts with observations reported for AlpA of CP4-57, which activates the expression of \textit{slpA}, the integrase gene of this cryptic prophage, and for which the role as an actual RDF remains unclear (27). Similarly, in addition to its RDF function, Vis of satellite prophage P4 has been shown to negatively regulate the expression of the P4 integrase by acting as an RNA-binding protein that post-transcriptionally regulates \textit{int} expression (35). To date, all the RDFs belonging to the SLP1 subgroup have been found associated with P4-type integrases, as it is also the case for RdfM (2, 10, 29, 30).

Given their importance for transfer and stability of integrating mobile elements, excision and integration must be tightly regulated. In many temperate bacteriophages and ICEs, the genes coding for the integration and excision functions are organized as a functional module in which the gene encoding the RDF precedes the gene coding for the recombinase, and both genes are organized in an operon-like structure (1, 9, 23, 40). In contrast, the organization of the genes coding for the integration and excision functions of MGIs is atypical: \textit{int}_{\text{MGi}} is located immediately adjacent to \textit{attL} and \textit{rdfM} is located on the opposite side, near the \textit{attR} site (Fig. 1B). The integrase gene \textit{intV2} (VC1758) and the
RDF gene vefA (VC1809) of the Vibrio Pathogenicity Island-2 (VPI-2) are organized in a similar manner (2). This organization also resembles to the KplE1 prophage, in which the attL site overlaps with the promoter of the intS gene, and the gene coding for the RDF TorI is remotely located, near the attR site (34). Yet while the recombination function of MGIs and KplE1 seem to be organized alike, they are functionally unrelated. In KplE1, intS gene is tightly regulated by its own product as well as by the TorI protein (34). On the contrary, we showed that in MGIs, rdfM is not a transcriptional regulator of intMGI. In fact, the uncoupled transcription of rdfM and intMGI is likely a feature selected for by the opportunistic behavior of MGIs as they rely on the SXT/R391 ICE-encoded transcriptional regulator SetCD to excise from their host chromosome. Unlike MGI excision, MGI integration does not require activation by SetCD. Once in a recipient cell the MGI integrates site-specifically through the action of IntMGI expressed at low level in the absence of SetCD. This process allows an MGI to establish itself in the host cell and be maintained in its progeny even in the absence of an SXT/R391 ICE. Such regulatory mechanism might have been selected to prevent MGI loss due to unproductive excision in the absence of a potentially mobilizing ICE of the SXT/R391 family. As a consequence, like for these ICEs, MGI excision and transfer is triggered by any physical or chemical agents that will damage DNA and stimulate the bacterial SOS response, including UV-light irradiation or exposure to mitomycin C and antibiotics such as ciprofloxacin (5, 13). Almagro-Moreno et al (2) showed that Vibrio Pathogenicity Island-2 (VPI-2) excise from chromosome I of V. cholerae N16961 after sublethal UV-light irradiation of the cells. Increased excision of VPI-2 was correlated to increased expression of intV2 and vefA. However, since V. cholerae N16961 is devoid of...
SXT/R391 ICE, induction of VPI-2 excision cannot be controlled by the SetR/SetCD pathway, suggesting that it may instead rely on the SOS response regulon repressor lexA.

This study brings a new understanding of the dynamics of excision, transfer and integration of mobile genetic elements, giving a better insight on the rules that direct their mobility. We also show a novel interaction between two phylogenetically unrelated families of GIs and show how a small non self-transmissible GI with no conjugative functions, can take advantage of the conjugative machinery and regulatory pathways of a self-transmissible GI in order to transfer from one cell to another.
Acknowledgments

This work was supported by a Discovery Grant and Discovery Acceleration Supplement from the Natural Sciences and Engineering Council of Canada (V.B). D.P.L holds a Fonds de recherche du Québec doctoral fellowship. M.M. was supported by the Deutscher Akademischer Austausch Dienst RISE program. V.B. holds a Canada Research Chair in molecular bacterial genetics and is a member of the FRSQ-funded Centre de Recherche Clinique Étienne-Le Bel.

We are grateful to D. Mazel for kindly providing us with E. coli β2163 and plasmid pSW23T. We would like to thank E. Bordeleau, G. Garriss, N. Carraro and A. Lavigne for helpful comments and critical reading of the manuscript.
References


9. Burrus V, Pavlovic G, Decaris B, Guedon G. 2002. The ICEStl element of Streptococcus thermophilus belongs to a large family of integrative and
conjugative elements that exchange modules and change their specificity of integration. Plasmid 48:77-97.


**Table 1.** Strains and plasmids used in this study

<table>
<thead>
<tr>
<th>Strains or plasmids</th>
<th>Relevant genotype or phenotype</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Escherichia coli</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β2163</td>
<td>(F) RP4-2: Tc::Mu ΔdapA::(erm-pir) (Kn’ Em’)</td>
<td>(15)</td>
</tr>
<tr>
<td>MC4100 λpir</td>
<td>F’ araD139 (argF-lac) U169 rpsL150 (Sm’ tetrA flbB5301 deoC1 ptsF25 rbsR Δλpir)</td>
<td>(19)</td>
</tr>
<tr>
<td>CAG18439</td>
<td>MG1655 lacZU118 lacI42::Tn10 (Tc r)</td>
<td>(39)</td>
</tr>
<tr>
<td>VB112</td>
<td>MG1655 Rf’</td>
<td>(12)</td>
</tr>
<tr>
<td>AD57</td>
<td>CAG18439 prfC::ICE/Ind1 yicC::MGI/Ind1 (Tc’ Su’ Tm’ Kn’ rpsL150 (Sm’ tetrA flbB5301 deoC1 ptsF25 rbsR Δλpir))</td>
<td>(13)</td>
</tr>
<tr>
<td>AD63</td>
<td>CAG18439 prfC::ICE/Ind1 yicC::MGI/Ind1::aph (Tc’ Su’ Tm’ Kn’ rpsL150 (Sm’ tetrA flbB5301 deoC1 ptsF25 rbsR Δλpir))</td>
<td>(13)</td>
</tr>
<tr>
<td>AD72</td>
<td>CAG18439 prfC::SXT yicC::MGI/Ind1::aph (Tc’ Su’ Tm’ Kn’ rpsL150 (Sm’ tetrA flbB5301 deoC1 ptsF25 rbsR Δλpir))</td>
<td>(13)</td>
</tr>
<tr>
<td>AD81</td>
<td>CAG18439 prfC::ICE/Ind1 yicC::MGI/Ind1 Δint::aad7 (Tc’ Su’ Tm’ Sp’ rpsL150 (Sm’ tetrA flbB5301 deoC1 ptsF25 rbsR Δλpir))</td>
<td>(13)</td>
</tr>
<tr>
<td>AD130</td>
<td>CAG18439 yicC::MGI/Ind1::aph (Tc’ Kn’ rpsL150 (Sm’ tetrA flbB5301 deoC1 ptsF25 rbsR Δλpir))</td>
<td>This study</td>
</tr>
<tr>
<td>AD132</td>
<td>CAG18439 yicC::MGI/Ind1::aph pG2B (Tc’ Kn’ Ap’ rpsL150 (Sm’ tetrA flbB5301 deoC1 ptsF25 rbsR Δλpir))</td>
<td>This study</td>
</tr>
<tr>
<td>AD133</td>
<td>CAG18439 prfC::SXT ΔsetCD yicC::MGI/Ind1::aph (Tc’ Su’ Tm’ Kn’ rpsL150 (Sm’ tetrA flbB5301 deoC1 ptsF25 rbsR Δλpir))</td>
<td>This study</td>
</tr>
<tr>
<td>AD167</td>
<td>CAG18439 prfC::ICE/Ind1 yicC::MGI/Ind1 Δads9::aph (Tc’ Su’ Tm’ Kn’ rpsL150 (Sm’ tetrA flbB5301 deoC1 ptsF25 rbsR Δλpir))</td>
<td>This study</td>
</tr>
<tr>
<td>AD169</td>
<td>CAG18439 prfC::ICE/Ind1 yicC::MGI/Ind1 Δads4::aph (Tc’ Su’ Tm’ Kn’ rpsL150 (Sm’ tetrA flbB5301 deoC1 ptsF25 rbsR Δλpir))</td>
<td>This study</td>
</tr>
<tr>
<td>AD192</td>
<td>CAG18439 prfC::ICE/Ind1 yicC::MGI/Ind1 Δads8::aph (Tc’ Su’ Tm’ Kn’ rpsL150 (Sm’ tetrA flbB5301 deoC1 ptsF25 rbsR Δλpir))</td>
<td>This study</td>
</tr>
<tr>
<td>AD207</td>
<td>CAG18439 prfC::R997 yicC::MGI/Ind1::aph (Tc’ Ap’ Kn’ rpsL150 (Sm’ tetrA flbB5301 deoC1 ptsF25 rbsR Δλpir))</td>
<td>This study</td>
</tr>
<tr>
<td>AD208</td>
<td>CAG18439 plntVvu (Tc’ Ap’ rpsL150 (Sm’ tetrA flbB5301 deoC1 ptsF25 rbsR Δλpir))</td>
<td>This study</td>
</tr>
<tr>
<td>AD210</td>
<td>β2163 pVB200 (Em’ Cm’ rpsL150 (Sm’ tetrA flbB5301 deoC1 ptsF25 rbsR Δλpir))</td>
<td>This study</td>
</tr>
<tr>
<td>AD217</td>
<td>CAG18439 yicC::pVB200 plntVvu (Tc’ Cm’ Ap’ rpsL150 (Sm’ tetrA flbB5301 deoC1 ptsF25 rbsR Δλpir))</td>
<td>This study</td>
</tr>
<tr>
<td>AD232</td>
<td>β2163 pSW23T (Em’ Cm’ rpsL150 (Sm’ tetrA flbB5301 deoC1 ptsF25 rbsR Δλpir))</td>
<td>This study</td>
</tr>
</tbody>
</table>
Plasmids

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>pIntVvu</td>
<td>pBAD-TOPO <em>int</em>&lt;sub&gt;MGI&lt;/sub&gt;\textit{vvuTai1} (Ap&lt;sup&gt;r&lt;/sup&gt;) (13)</td>
</tr>
<tr>
<td>pGG2B</td>
<td>pBAD30 setCD (Ap&lt;sup&gt;r&lt;/sup&gt;) G. Garriss</td>
</tr>
<tr>
<td>pSW23T</td>
<td><em>ori</em>T&lt;sub&gt;RP4&lt;/sub&gt;; <em>ori</em>V&lt;sub&gt;R6K&lt;/sub&gt; (Cm&lt;sup&gt;r&lt;/sup&gt;) (15)</td>
</tr>
<tr>
<td>pVB200</td>
<td>pSW23T <em>attP</em>&lt;sub&gt;MGI&lt;/sub&gt;/\textit{vflind} (Cm&lt;sup&gt;r&lt;/sup&gt;) This study</td>
</tr>
<tr>
<td>p8</td>
<td>pBAD-TOPO <em>cds9</em>&lt;sub&gt;MGI&lt;/sub&gt;/\textit{vvuTai1} (Ap&lt;sup&gt;r&lt;/sup&gt;) This study</td>
</tr>
<tr>
<td>p9</td>
<td>pBAD-TOPO <em>cds9</em>&lt;sub&gt;MGI&lt;/sub&gt;/\textit{vvuTai1} (Ap&lt;sup&gt;r&lt;/sup&gt;) This study</td>
</tr>
<tr>
<td>p9-8</td>
<td>pBAD-TOPO <em>cds9-8</em>&lt;sub&gt;MGI&lt;/sub&gt;/\textit{vflind} (Ap&lt;sup&gt;r&lt;/sup&gt;) This study</td>
</tr>
<tr>
<td>pKD13</td>
<td>PCR template for one-step chromosomal gene inactivation (Kan&lt;sup&gt;r&lt;/sup&gt;) (14)</td>
</tr>
</tbody>
</table>

*a* Ap<sup>r</sup>, ampicillin resistant; Cm<sup>r</sup>, chloramphenicol resistant; Kan<sup>r</sup>, kanamycin resistant; N<sup>x</sup>r, nalidixic acid resistant; Rf<sup>r</sup>, rifampicin resistant; Sm<sup>r</sup>, streptomycin resistant; Sp<sup>r</sup>, spectinomycin resistant; Su<sup>r</sup>, sulfamethoxazole resistant; Tc<sup>r</sup>, tetracycline resistant; Tm<sup>r</sup>, trimethoprim resistant
Table 2. DNA sequences of the primers used in this study

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Nucleotide sequence (5' to 3')</th>
<th>Use in the study</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD4-V-F</td>
<td>TAGCAGTGAGGAGAACACATGG</td>
<td>Amplification of ( \text{cds9}_{\text{MGIVvlInd1}} )</td>
</tr>
<tr>
<td>AD4-R1</td>
<td>TTATCCACGGCATAAGCA</td>
<td>Amplification of ( \text{cds9}_{\text{MGIVvlInd1}} )</td>
</tr>
<tr>
<td>AD5-F</td>
<td>GCCGTGATAAACCCTGAGCA</td>
<td>Amplification of ( \text{cds8}_{\text{MGIVvlInd1}} )</td>
</tr>
<tr>
<td>AD5-V-R1</td>
<td>TTAGTCTACAAAATGCGCTTTT</td>
<td>Amplification of ( \text{cds9-8}_{\text{MGIVvlInd1}} )</td>
</tr>
<tr>
<td>AD5-A-R1</td>
<td>TTAATCCACGAGTCTGACCTTTT</td>
<td>Amplification of ( \text{cds9-8}_{\text{MGIVvlInd1}} )</td>
</tr>
<tr>
<td>AD4-A-F</td>
<td>TAGCCGATTAGTACTGGCAAACTCC</td>
<td>Deletion of ( \text{cds9} ) in MGIV/Ind1</td>
</tr>
<tr>
<td>AD11-WF</td>
<td>CAGCCACCGGACCACCCCAATAGGAATGGACGC</td>
<td>Deletion of ( \text{cds9} ) in MGIV/Ind1</td>
</tr>
<tr>
<td>AD11-WR</td>
<td>ATGAACTACACTACAATCAATCCATCCAAAATGAAGCHARCAGA</td>
<td>Deletion of ( \text{cds9} ) in MGIV/Ind1</td>
</tr>
<tr>
<td>AD13-WF</td>
<td>AGTTGTCATCCACGAGTGAATGGATAAGAATGACGA</td>
<td>Deletion of ( \text{cds4} ) in MGIV/Ind1</td>
</tr>
<tr>
<td>AD13-WR</td>
<td>CATCCACGAGTGAATCCATGCCATCTCTCTTTTGACGC</td>
<td>Deletion of ( \text{cds4} ) in MGIV/Ind1</td>
</tr>
<tr>
<td>attPAD-L1</td>
<td>TCGGCTTTGCTGTGCAATTA</td>
<td>Amplification of ( \text{attP}_{\text{MGIVvlInd1}} ) 1st round</td>
</tr>
<tr>
<td>attPAD-R1-AC</td>
<td>TCTGCCTAGCAACAGCAAT</td>
<td>Amplification of ( \text{attP}_{\text{MGIVvlInd1}} ) 1st round</td>
</tr>
<tr>
<td>attPAD-L2</td>
<td>GAGTTTCCCATGTTTACTCCATA</td>
<td>Amplification of ( \text{attP}_{\text{MGIVvlInd1}} ) 2nd round</td>
</tr>
<tr>
<td>attPAD-R2-AC</td>
<td>GTCACAGCTTTTGCCTGCTTT</td>
<td>Amplification of ( \text{attP}_{\text{MGIVvlInd1}} ) 2nd round</td>
</tr>
<tr>
<td>Gene8-WF</td>
<td>TATGCTAGCAACATGCAAAATTACCAGCTGTGTTCTG</td>
<td>Deletion of ( \text{cds8} ) in MGIV/Ind1</td>
</tr>
<tr>
<td>Gene8-WR</td>
<td>TATGCTAGCAACATGCAAAATTACCAGCTGTGTTCTG</td>
<td>Deletion of ( \text{cds8} ) in MGIV/Ind1</td>
</tr>
<tr>
<td>Q-PCR-1F</td>
<td>AAGTGACAAACATCCGCAAC</td>
<td>Amplification of ( \text{attB} ) in ( \text{E. coli} )</td>
</tr>
<tr>
<td>Q-PCR-1R</td>
<td>GCAAGCAGAAGCAAACTGGA</td>
<td>Amplification of ( \text{attB} ) in ( \text{E. coli} )</td>
</tr>
<tr>
<td>Q-PCR-2F</td>
<td>GAACAGCAAGCAAGCAAAAC</td>
<td>Amplification of ( \text{rph} ) in ( \text{E. coli} )</td>
</tr>
<tr>
<td>Q-PCR-2R</td>
<td>GTGCCCTGAGACTCCATGAT</td>
<td>Amplification of ( \text{rph} ) in ( \text{E. coli} )</td>
</tr>
<tr>
<td>RTgene9-F b</td>
<td>TTATCCACGAGCATAAGCA</td>
<td>Amplification of ( \text{cds9}_{\text{MGIVvlInd1}} )</td>
</tr>
<tr>
<td>RTgene9-R b</td>
<td>AGGTAAAGGCAACACACTGAGGCTTTTCTT</td>
<td>Amplification of ( \text{cds9}_{\text{MGIVvlInd1}} )</td>
</tr>
<tr>
<td>RTintVf-F</td>
<td>CCTATCGCAGTTTACCACAA</td>
<td>Amplification of ( \text{int}_{\text{MGIVvlInd1}} )</td>
</tr>
<tr>
<td>RTintVf-R</td>
<td>TTATCCACGAGCATAAGCA</td>
<td>Amplification of ( \text{int}_{\text{MGIVvlInd1}} )</td>
</tr>
<tr>
<td>RTgeoZcoli-F</td>
<td>GCTCGTCAGATGCGAGGAGG</td>
<td>Amplification of ( \text{rpoZ} )</td>
</tr>
<tr>
<td>RTgeoZcoli-R</td>
<td>GCTGTAATTCCAGCGGTCTT</td>
<td>Amplification of ( \text{rpoZ} )</td>
</tr>
<tr>
<td>RTyicC-F</td>
<td>GAGTGACAGAAGGGAACATCA</td>
<td>Amplification of ( \text{yicC} )</td>
</tr>
<tr>
<td>RTyicC-R</td>
<td>GCTTTTCGTGCACCTTCTT</td>
<td>Amplification of ( \text{yicC} )</td>
</tr>
</tbody>
</table>
Figure legends

**Figure 1.** Schematic representation of the core sets of conserved genes of SXT/R391 ICEs (A) and MGIs (B). Vertically hatched ORFs indicate the integration sites of the elements (prfC for SXT/R391 ICEs and yicC for MGIs). Black ORFs represent genes involved in site-specific excision and integration. Light grey ORFs represent genes encoding the conjugative transfer machinery. Dark grey ORFs correspond to genes involved in regulation (setCDR), and white ORFs represent genes of unknown function. oriTs are represented by horizontal grey arrowheads. Hotspots for insertion of variable DNA are indicated by black arrowheads pointing upward.

**Figure 2.** Genetic requirements for integration and excision of a replication-deficient plasmid containing the attP site of MGI/V/F/Indl. A. Schematic map of pVB200. B and C. Mobilization assays of pSW23T and pVB200 performed to assess plasmid integration into the 3’ end of yicC (attB). Conjugation assays were carried out using *E. coli* β2163 (pir+) as a donor and MC4100 λpir (pir+) or CAG18439 variants (pir-) as recipient strains. The genetic background of each recipient strain is indicated on the left side of the panels. R997 is an Ap'-conferring ICE of the SXT/R391 family. To induce expression of IntMGI from pIntVvu or SetCD from pGG2B, the conjugation assays were carried out on media supplemented with 0.02% arabinose. The frequency of exconjugant formation was obtained by dividing the number of exconjugants (Tc' Cm' CFU for CAG18439 or Sm' Cm' for MC4100 λpir) by the number of recipients (Tc' or Sm' CFU, respectively). The
bars indicate the mean values and standard deviations obtained from three independent experiments. D. Analysis of excision of pVB200 integrated into yicC (attB). Ethidium bromide-stained 2% agarose gel of attB and rph fragments amplified by semi-quantitative PCR. Lanes: M, 2-Log molecular size marker; +, CAG18439; 1 and 2, CAG18439 yicC::pVB200 plntVvu; 3, CAG18439 yicC::pVB200-MGI/Ind1; 4 and 5, CAG18439 yicC::pVB200-MGI/Ind1 prfC::R997; 6 and 7 CAG18439 yicC::pVB200-MGI/Ind1 pGG2B. Lanes 2 and 7, cultures were induced with 0.02% arabinose; Lane 5, culture was induced with 100 ng/ml mitomycin C.

Figure 3. Sequence alignment of the translation products of cds9_{MGIVVuTail} and cds9_{MGIVVUSA1} with related RDFs. The primary sequences of RDFs encoded by two sequenced MGIs were aligned using MUSCLE with the transcriptional regulator AlpA from CP4-57 prophage (NP_417113) and the RDFs Xis encoded by ICEs of the SXT/R391 family (ACV96240), Hef of HPI from Y. pseudotuberculosis (CAB46594), VefA of VPI-2 from V. cholera N16961 (NP_231420), SLP1 of plasmid SLP1 from S. coelicolor, and Vis of the satellite bacteriophage P4 (NP_042041). Amino acid residues that are identical or similar (BLOSUM62 substitution matrix) in at least 60% of the sequences are indicated by a black or grey background, respectively. The solid bar indicates a helix-turn-helix (HTH) DNA-binding motif predicted in all proteins based on the Dodd-Egan method (17) (Dodd-Egan scores of 2.36 or higher), whereas the dashed
bar highlights a secondary HTH motif exclusively predicted in RDFs encoded by MGIs (Dodd-Egan score of 3.07). The length of each protein is indicated in the right column.

**Figure 4.** Genetic requirements for excision and transfer of MGI/VflInd1. Mobilization assay of MGI/VflInd1::aph or its Δcds4, Δcds8 or Δcds9 mutants by ICE/VflInd1 were carried out using *E. coli* CAG18439 containing ICE/VflInd1 and MGI/VflInd1 mutants as donors. When indicated, the donor expressed cds9, 8 or 9 and 8 from p9, p8 or p9-8, respectively. *E. coli* VB112 (Rf') was used as the recipient strain. The frequency of exconjugant formation was calculated by dividing the number of exconjugants (Rf' Kn') CFU by the number of donors (Tc' CFU). Real-time quantitative PCR was used to determine the percentage of unoccupied attB sites resulting from the circularization of MGI/VflInd1 or of its Δcds9 mutant in *E. coli* CAG18439 harboring ICE/VflInd1, p9, p8 or p9-8. The bars indicate the mean values and standard deviations obtained from three independent experiments. ND, not determined. Asterisks indicate that the frequency of exconjugant formation or the percentage of attB sites were below the limit of detection of the assays (<1×10⁻⁸ or 0.0004%, respectively).

**Figure 5.** Regulation of the expression of integration and excision genes of MGIs and SXT/R391 ICEs. A. Impact of protein 9 on the expression of *int*<sub>MGI</sub>. The expression of *int*<sub>MGI</sub> was measured by quantitative PCR upon induction of the SOS response (+MC) in AD72, AD130 and AD167, or overexpression of *cds*9 from p9 in AD167 p9 (+ara). One-
way ANOVA with a Dunnett post-test was used to compare the means of relative expression of \(\text{int}_{MGI}\) against the strain devoid of ICE. The confidence interval for the comparisons was \(P < 0.01\) (***). ns, non-significant. B. Impact of SetCD on \(\text{int}_{MGI}\) and \(\text{rdfM}\) expression. The expression of \(\text{yicC}, \text{int}_{MGI}\) and \(\text{rdfM}\) was measured upon induction of the SOS response (+MC) in AD72, and AD133, or upon overexpression of \(\text{setCD}\) from pGG2B (+ara) in AD132.

**Figure 6.** \(\text{int}_{MGI}\) has a basal level of expression in absence of SetCD. The expression of \(\text{int}_{MGI}\) was measured by quantitative PCR in AD72 and AD133. The graph shows differential gene expression values (\(\Delta C_t\)) compared with the housekeeping gene \(\text{rpoZ}\). Results are expressed as the mean of three independent biological replicates. The asterisk indicates that expression was below the detection limit.

**Figure 7.** Integration and excision dynamics of MGIs in donor and recipient cells. A. Excision from the donor cell chromosome. (1) SOS response is activated by DNA damages, alleviating the SetR-mediated repression of \(\text{setCD}\). (2) The transcriptional activator SetCD activates the expression of \(\text{int}_{MGI}\) and \(\text{rdfM}\). (3) Int\(_{MGI}\) and RdfM mediate the excision of the MGI. B. Integration into the recipient cell chromosome. Int\(_{MGI}\) is expressed at low level and allows MGI’s integration regardless of the presence or absence of an SXT/R391 ICE.
A  Conserved core genes of SXT/R391 ICEs

B  Conserved core genes of MGIs
A  Donor cell

DNA damages

(1)

setCD setR

ICE

(2)

MGI

attL

yicC

int

ori

attR

rdfM

MGI excision

(3)

B  Recipient cell

MGI

attP

rdfM

int

MGI integration

ori