



UNIL | Université de Lausanne

Unicentre

CH-1015 Lausanne

<http://serval.unil.ch>

Year : 2013

Regulation of horizontal gene transfer of the catabolic island *ICE_{cl}*

Nicolas Pradervand

Nicolas Pradervand 2013 Regulation of horizontal gene transfer of the catabolic *ICE_{cl}*

Originally published at : Thesis, University of Lausanne

Posted at the University of Lausanne Open Archive.
<http://serval.unil.ch>

Droits d'auteur

L'Université de Lausanne attire expressément l'attention des utilisateurs sur le fait que tous les documents publiés dans l'Archive SERVAL sont protégés par le droit d'auteur, conformément à la loi fédérale sur le droit d'auteur et les droits voisins (LDA). A ce titre, il est indispensable d'obtenir le consentement préalable de l'auteur et/ou de l'éditeur avant toute utilisation d'une oeuvre ou d'une partie d'une oeuvre ne relevant pas d'une utilisation à des fins personnelles au sens de la LDA (art. 19, al. 1 lettre a). A défaut, tout contrevenant s'expose aux sanctions prévues par cette loi. Nous déclinons toute responsabilité en la matière.

Copyright

The University of Lausanne expressly draws the attention of users to the fact that all documents published in the SERVAL Archive are protected by copyright in accordance with federal law on copyright and similar rights (LDA). Accordingly it is indispensable to obtain prior consent from the author and/or publisher before any use of a work or part of a work for purposes other than personal use within the meaning of LDA (art. 19, para. 1 letter a). Failure to do so will expose offenders to the sanctions laid down by this law. We accept no liability in this respect.



UNIL | Université de Lausanne

Faculté de biologie
et de médecine

Département de microbiologie fondamentale

**REGULATION OF HORIZONTAL GENE TRANSFER OF THE
CATABOLIC ISLAND ICE*clc***

Thèse de doctorat ès sciences de la vie (PhD)

présentée à la

Faculté de biologie et de médecine
de l'Université de Lausanne

par

Nicolas PRADERVAND

Master de l'Université de Lausanne

Jury

Prof. Alexandre Reymond, Président
Prof. Jan Roelof van der Meer, Directeur de thèse

Prof. Martin Ackermann, expert
Prof. Ulrich Dobrindt, expert

Lausanne 2013



UNIL | Université de Lausanne

Faculté de biologie
et de médecine

Ecole Doctorale

Doctorat ès sciences de la vie

Imprimatur

Vu le rapport présenté par le jury d'examen, composé de

| | | | |
|---------------------------|-----------|------------------|---------------------|
| <i>Président</i> | Monsieurq | Prof. Alexandre | Reymond |
| <i>Directeur de thèse</i> | Monsieur | Prof. Jan Roelof | Van der Meer |
| <i>Experts</i> | Monsieur | Prof. Martin | Ackermann |
| | Monsieur | Prof. Ulrich | Dobrindt |

le Conseil de Faculté autorise l'impression de la thèse de

Monsieur Nicolas Pradervand

Master ès sciences génomique de l'Université de Lausanne

intitulée

**REGULATION OF HORIZONTAL GENE TRANSFER
OF THE CATABOLIC ISLAND ICElc**

Lausanne, le 20 décembre 2013

pour Le Doyen
de la Faculté de Biologie et de Médecine

Prof. Alexandre Reymond

A Stéphanie

TABLE OF CONTENTS

| | |
|--|-----|
| SUMMARY | 2 |
| RÉSUMÉ | 4 |
| CHAPTER 1 : INTRODUCTION TO ICECLC AND OTHER INTEGRATIVE AND CONJUGATIVE ELEMENTS | 7 |
| CHAPTER 2 : TRANSCRIPTOME ANALYSIS OF THE MOBILE GENOME ICECLC IN <i>PSEUDOMONAS KNACKMUSSII</i> B13 | 59 |
| CHAPTER 3 : ICECLC BEHAVIOUR IN DIFFERENT HOST BACTERIA | 103 |
| CHAPTER 4 : CHARACTERIZATION OF A TETR-TYPE REPRESSOR REGULATED MAJOR FACILITATOR SYSTEM ON THE INTEGRATIVE AND CONJUGATIVE ELEMENT ICECLC IN <i>PSEUDOMONAS KNACKMUSSII</i> B13 | 125 |
| CHAPTER 5 : A UNIQUE OPERON OF THREE TRANSCRIPTIONAL REGULATORS CONTROLS ICECLC HORIZONTAL GENE TRANSFER IN <i>PSEUDOMONAS KNACKMUSSII</i> B13 | 159 |
| CHAPTER 6 : GENERAL DISCUSSION | 199 |
| CURRICULUM VITAE | 208 |
| LIST OF PUBLICATIONS | 209 |

Summary

ICE*clc* is a mobile genetic element found in two copies on the chromosome of the bacterium *Pseudomonas knackmussii* B13. ICE*clc* harbors genes encoding metabolic pathways for the degradation of chlorocatechols (CLC) and 2-aminophenol (2AP). At low frequencies, ICE*clc* excises from the chromosome, closes into a circular DNA molecule which can transfer to another bacterium via conjugation. Once in the recipient cell, ICE*clc* can reintegrate into the chromosome by site-specific recombination. This thesis aimed at identifying the regulatory network underlying the decisions for ICE*clc* horizontal transfer (HGT).

The first chapter is an introduction on integrative and conjugative elements (ICEs) more in general, of which ICE*clc* is one example. In particular I emphasized the current knowledge of regulation and conjugation machineries of the different classes of ICE.

In the second chapter, I describe a transcriptional analysis using microarrays and other experiments to understand expression of ICE*clc* in exponential and stationary phase. By overlaying transcriptomic profiles with Northern hybridizations and RT-PCR data, we established a transcription map for the entire core region of ICE*clc*, a region assumed to encode the ICE conjugation process. We also demonstrated how transcription of the ICE*clc* core is maximal in stationary phase, which correlates to expression of reporter genes fused to key ICE*clc* promoters.

In the third chapter, I present a transcriptome analysis of ICE*clc* in a variety of different host species, in order to explore whether there are species-specific differences.

In the fourth chapter, I focus on the role of a curious ICE*clc*-encoded TetR-type transcriptional repressor. We find that this gene, which we name *mfsR*, not only controls its own expression but that of a set of genes for a putative multi-drug efflux pump (*mfsABC*) as well. By using a combination of biochemical and molecular biology techniques, I could show that MfsR specifically binds to operator boxes in two ICE*clc* promoters (P_{mfsR} and P_{mfsA}), inhibiting the transcription of both the *mfsR* and *mfsABC-orf38184* operons. Although we could not detect a clear phenotype of an *mfsABC* deletion, we discuss the implications of pump gene reorganizations in ICE*clc* and close relatives.

In the fifth chapter, we find that *mfsR* not only controls its own expression and that of the *mfsABC* operon, but is also indirectly controlling ICE*clc* transfer. Using gene deletions, microarrays, transfer assays and microscopy-based reporter fusions, we demonstrate that *mfsR* actually controls a small operon of three regulatory genes. The last gene of this *mfsR* operon, *orf17162*, encodes a LysR-type activator that when deleted strongly impairs ICE*clc* transfer. Interestingly, deletion of *mfsR* leads to transfer competence in almost all cells, thereby overruling the bistability process in the wild-type.

In the final sixth chapter, I discuss the relevance of the present thesis and the resulting perspectives for future studies.

Résumé

ICE*clc* est un élément génétique trouvé dans le chromosome de la bactérie *Pseudomonas knackmussii* B13. ICE*clc* contient des gènes codant pour des voies métaboliques auxiliaires et permettant la dégradation de polluants tels que les chlorocatechols (CLC) et le 2-aminophenol (2AP). Lors de certaines conditions, ICE*clc* s'excise du chromosome, se circularise et se transfère à une autre bactérie à la façon d'un plasmide conjugatif. Une fois dans la cellule récipiente, ICE*clc* se réintègre dans le chromosome et confère à son nouvel hôte la capacité de dégrader des CLC et du 2AP. L'objectif de cette thèse de doctorat était d'explorer le système de régulation génétique contrôlant le processus de transfert horizontal d'ICE*clc*.

Le premier chapitre est une introduction générale aux éléments intégratifs et conjugatifs (ICEs), auxquels appartient ICE*clc*. Le point est fait sur l'état actuel des connaissances en matière de régulation et de la machinerie du transfert horizontal des différents modèles d'ICEs.

Dans le second chapitre, des expériences basées sur des microarrays ont été utilisées pour comprendre l'expression des gènes d'ICE*clc* en phases exponentielle et stationnaire. En superposant les profils transcriptomiques avec des expériences antérieures du Dr M. Gaillard, il fut possible d'établir une carte des transcrits du module génétique conservé d'ICE*clc*. Ce même module est supposé coder la plupart des facteurs nécessaires au transfert horizontal de l'ICE*clc*. Simultanément, nous avons démontré que la transcription du module conservé est maximale en phase stationnaire, en même temps que le transfert horizontal est le plus fréquent.

Dans le troisième chapitre, nous présentons les transcriptomes d'ICE*clc* à l'intérieur de différentes espèces bactériennes afin d'explorer les influences de la cellule hôte sur la transcription des gènes.

Dans le quatrième chapitre, nous nous concentrons sur le rôle d'un facteur de transcription de la famille des répresseurs TetR, encodé sur ICE*clc*. Nous démontrons que le gène *mfsR* ne contrôle pas seulement sa propre expression, mais également celle d'une série de gènes codant pour un système de pompes à efflux (*mfsABC*). En utilisant une combinaison de techniques de biologie moléculaire, de biochimie et de technologie génomique, nous démontrons que MfsR se lie spécifiquement à des séquences opératrices de deux promoteurs d'ICE*clc* (P_{mfsR} et P_{mfsA}), inhibant de la sorte les deux opérons *mfsR-orf17984-tciR* et *mfsABC-orf38184*. Bien qu'il ne fut pas possible d'identifier un phénotype clair, nous discutons les implications de la réorganisation des gènes des pompes à efflux dans ICE*clc* et dans des éléments parents.

Dans le cinquième chapitre, nous démontrons que *mfsR* ne contrôle pas uniquement sa propre expression et celle de l'opéron *mfsABC*, mais également celle d'un régulateur majeur du transfert horizontal d'ICE*clc*, *tciR*. Ce gène code pour un activateur responsable pour l'initiation du phénotype de compétence de transfert. Curieusement, une délétion dans *mfsR* a pour conséquence l'activation de la compétence de transfert dans pratiquement toutes les cellules, court-circuitant ainsi le processus de bistabilité de la souche sauvage.

Dans le chapitre final, nous discutons des découvertes faites durant cette thèse de doctorat et les mettons en perspectives.

CHAPTER 1

INTRODUCTION TO ICECLC AND OTHER INTEGRATIVE AND CONJUGATIVE ELEMENTS

Integrative and Conjugative Elements, a new class of mobile genetic elements

The discovery of horizontal gene transfer has raised great interest, partly because it represented a refreshing alternative to vertical heredity-based genetics but also because of growing public health concern over the emergence of antibiotic resistances [1–4]. Not only the variety of different horizontal transmission modes appeared astonishing (i.e., transformation, conjugation or transduction), but also the great variety of selfish mobile genetic elements (MGEs) that mediate horizontal gene transfer. Consequently, elements such as prophages, plasmids and transposons have been extensively studied for decades, both for the auxiliary functions they brought to host cells as well as for their remarkable life-styles [4–6]. Further hybrid genetic entities were identified that were classified as conjugative transposons (CTns), integrative plasmids or (with the expansion of whole genome sequencing) as genomic islands (GIs) [7]. Integrative and conjugative elements (ICE) is a newer nomenclature describing elements with confirmed capability to transfer by conjugation, but which under normal conditions reside in an integrated state in the bacterial chromosome [8–10]. Despite forming a very diverse family of mobile elements, ICE share a common overall life-style with two conformations, (i) the integrated form, in which the ICE lays inserted in the host chromosome (often at one or more site-specific locations), and (ii) the circular form, in which the ICE exists as an extrachromosomal, plasmid-like, double-stranded and covalently closed circular DNA molecule (Figure 1). ICE are thus similar to temperate bacteriophages in that they insert in the host chromosome, but similar to conjugative plasmids in that their excised form can be horizontally transmitted to a new recipient cell via conjugation. In the inserted form the ICE is replicated through chromosome replication and inherited by each of the daughter cells.

The route of horizontal transmission of an ICE thus requires a number of defined steps : (a) a molecular switch leading to activation of the horizontal transmission program, (b) excision and recombination to the circular form, (c) DNA processing and conjugative transfer to a recipient cell, (d) site-specific reintegration in the new host chromosome, and (e) reprogramming of the molecular switch to maintain the integrated state (Figure 1). Interestingly, many ICEs have the propensity to site-specifically integrate into the 3'-ends of tRNA genes, without disrupting the reading frame by carefully reconstituting its end [11–15]. Several ICEs have been well-studied at the molecular level and currently serve as models for our general understanding of ICE life style. Despite showing a relatively common life-cycle, ICEs are not necessarily phylogenetically related, but rather consist of analogous entities with similar challenges for survival and distribution [9,16,17]. All ICEs discovered so far harbour a mosaic-like modular architecture, with modules being dedicated to specific functions either related to (i) the various steps of the life-style or (ii) to auxiliary functions they may carry. The modules linked to life-style tend to be conserved among related ICEs and are termed the "core" region or backbone. The auxiliary functions can greatly vary, even among ICEs of the same family, and are therefore grouped as "variable" or "cargo" modules. These auxiliary functions are largely responsible for the evolutionary success of ICEs, as they enable their hosts with the potential to resist, colonize and compete in new/changing ecological niches, thus theoretically enhancing their fitness [18–21]. Although not all studied ICEs bear clear auxiliary functions, frequently one finds genes for antibiotic or heavy metal resistances, for potential virulence factors, symbiosis traits or additional metabolic functions.

The next few paragraphs summarize the life-style and regulatory mechanisms of the best-studied ICE models, namely ICEBs1, SXT-R391, CTnDOT, Tn916, from which most of the current knowledge on this particular class of mobile genomes is drawn. Additional paragraphs then continue with other very interesting ICEs, such as pSAM2, ICEMISym^{R7A}, ICESt1-3 and HPI. Finally, ICE*clc* and its family of related ICEs is described.

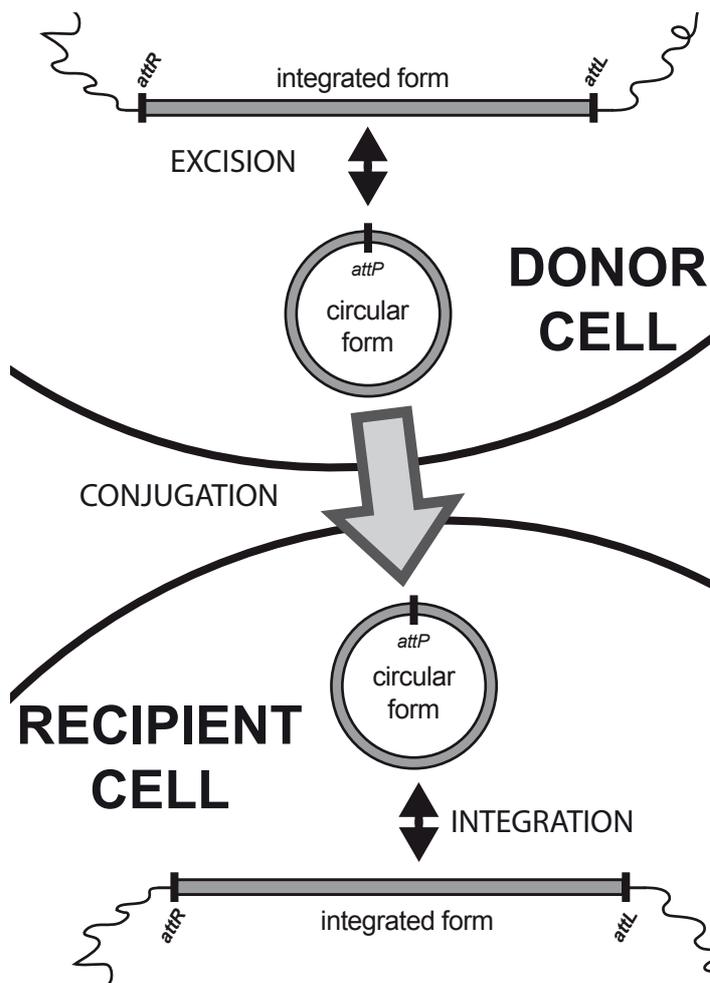


Figure 1. General overview of ICE lifestyle. In a donor cell, an ICE resides in the chromosome of their host as integrated form, delimited by attachment sites *attR* and *attL*. Upon specific conditions, the integrated form is excised and recombined to a circular form. After specific DNA processing, the ICE circular form is translocated to the recipient cell via a conjugation apparatus. In the recipient cell, the circular form is integrated into its integration site (*attB*) to reconstitute the integrated form. In such state, the integrated form is vertically transferred to progeny by chromosome replication and segregation.

ICEBs1

ICEBs1 is a 21 kb long ICE, thought to encode 25 genes, which is originally found in *B. subtilis*, although it can colonize (via laboratory matings) *B. anthracis*, *B. licheniformis* and *L. monocytogenes* [22]. Interestingly, no clear phenotype is conferred by ICEBs1. ICEBs1 is integrated in a gene for tRNA^{Leu}, although secondary sites with the same 17-bp *attB* consensus can also be used when the primary site is deleted [23]. ICEBs1 excision and transfer is intricately regulated as a function of nutrient availability, cell density and the presence of non-ICEBs1-bearing cells [22]. The main factors that orchestrate this play are PhrI, RapI and AbrB (Figure 2A). PhrI is a signal pentapeptide encoded by ICEBs1, which is secreted outside the cell, where it accumulates until it is re-imported by the oligopeptide permease Opp (which is chromosomally-encoded). When the concentration of ICEBs1-bearing cells is high (for example in stationary phase), PhrI reaches high levels and inhibits (somehow) RapI, which is the main global activator for ICEBs1 excision and conjugation. Therefore, transfer is avoided when it is not necessary (i.e. many cells already contain ICEBs1). During exponential phase, PhrI levels are too low to inhibit the HGT activator RapI, but now AbrB (a chromosomal factor) represses the transcription of *rapI*, thus preventing ICEBs1 transfer during cell division (Figure 2A). On the other hand, when nutrients are scarce, cell density is high, but ICEBs1 carrying cells are rare, PhrI and AbrB fail to repress RapI, which can accumulate and initiate excision [22]. RapI enhances ImmA-protease cleavage of ImmR, a repressor of the excisionase gene *xis* [24–26] (Figure 2B). Thus, unchecked RapI ultimately causes the excision of the integrated form ICEBs1 via the action of Xis and the constitutively expressed integrase Int [23]. ImmR also autoregulates its own

expression and that of *immA* and *int*, which are in the same operon on *ICEBs1*, resulting in an apparent constitutive expression [22,24].

Independently from the PhrI-Rapl cascade, the SOS response can also induce excision of *ICEBs1* in a similar manner as known for phage Lambda control [22]. DNA damage such as produced by mitomycin C triggers RecA to enforce (in a yet unknown mechanism) ImmA cleavage of ImmR, freeing *xis* expression (Figure 2B). It was postulated that this alternative HGT induction evolved in order to allow *ICEBs1* to abandon damaged host cells [22].

The excised dsDNA *ICEBs1* molecule is processed for transfer by NickK, an *ICEBs1* relaxase encoded in the *xis* operon. NickK catalyzes (alone) a single-strand nick at the origin of transfer (*oriT*), which partially overlaps with the *nick* gene [27]. The host-encoded helicase PcrA and the *ICEBs1*-encoded helicase HelP each assemble at the nicked *oriT* and cooperatively unwind the ds circular DNA in a single direction [28,29]. A chromosomally-encoded single strand DNA binding protein (Ssb) stabilizes the unwound DNA, forming a nucleoprotein complex together with NickK on the nicked strand. At this stage, two concurring events take place (i) a rolling circle-like replication restoring the dsDNA molecule and increasing its copy number, and presumably (ii), the processing of the nicked strand toward the conjugation apparatus [28,30]. Interestingly, nicked *oriT* serves not only as an origin of transfer but also as an origin of replication for a rolling circle-like replication. Indeed, NickK, HelP and the host-encoded PcrA are essential to support the autonomous replication of *ICEBs1* catalyzed by the host cell machinery, including PolC, DnaN and Ssb [28,29]. *ICEBs1* represented the first ICE model, for which plasmid-like autonomous replication could be demonstrated, a process not expected since replication of the integrated form along with the chromosome was believed the only mechanism for ICE maintenance

[10]. Autonomous replication is not essential for *ICEBs1* horizontal transfer, but guarantees its stability in the donor, especially if growth resumes before reintegration can occur [28].

By analogy to plasmid transfer, it is believed that the nicked single-stranded nucleoprotein complex is directed toward the conjugative machinery, which translocates it into the recipient cell. Bioinformatic analysis predicts that the *ICEBs1*-encoded YdcQ may act as the coupling protein [27], whereas YddB, YddC, YddD, Yddl, YddJ, YddM and YddF may be part of the translocation channel. CwlT may serve as cell wall hydrolase possibly involved in transfer [27]. ConE and ConG are two *ICEBs1* proteins essential for transfer and resemble the mating pore formation proteins (Mpf) VirB4 and VirB6, respectively. ConE is known for localizing mostly at the cell poles, where *ICEBs1* replication foci are also observed [30]. Thus it is believed that both *ICEBs1* conjugation and replication occur predominantly at the cell poles, although transfer can occur from non polar zones of the cell, too [31]. Polar localization of the conjugation machinery is thought to increase the efficiency at which the *ICEBs1* transfers within chains of interconnected bacteria, a phenomenon predominant in biofilms [31]. Once in the recipient cell, the *ICEBs1* single-stranded circular form is supposedly replicated. Since *int* expression is constitutive, Int directs (alone) the integration of the circular form into the extremity of a tRNA^{Leu} site [23]. In the transconjugant, ImmR protects *ICEBs1* against secondary acquisition of *ICEBs1* by the host cell by repressing expression of *int* [24].

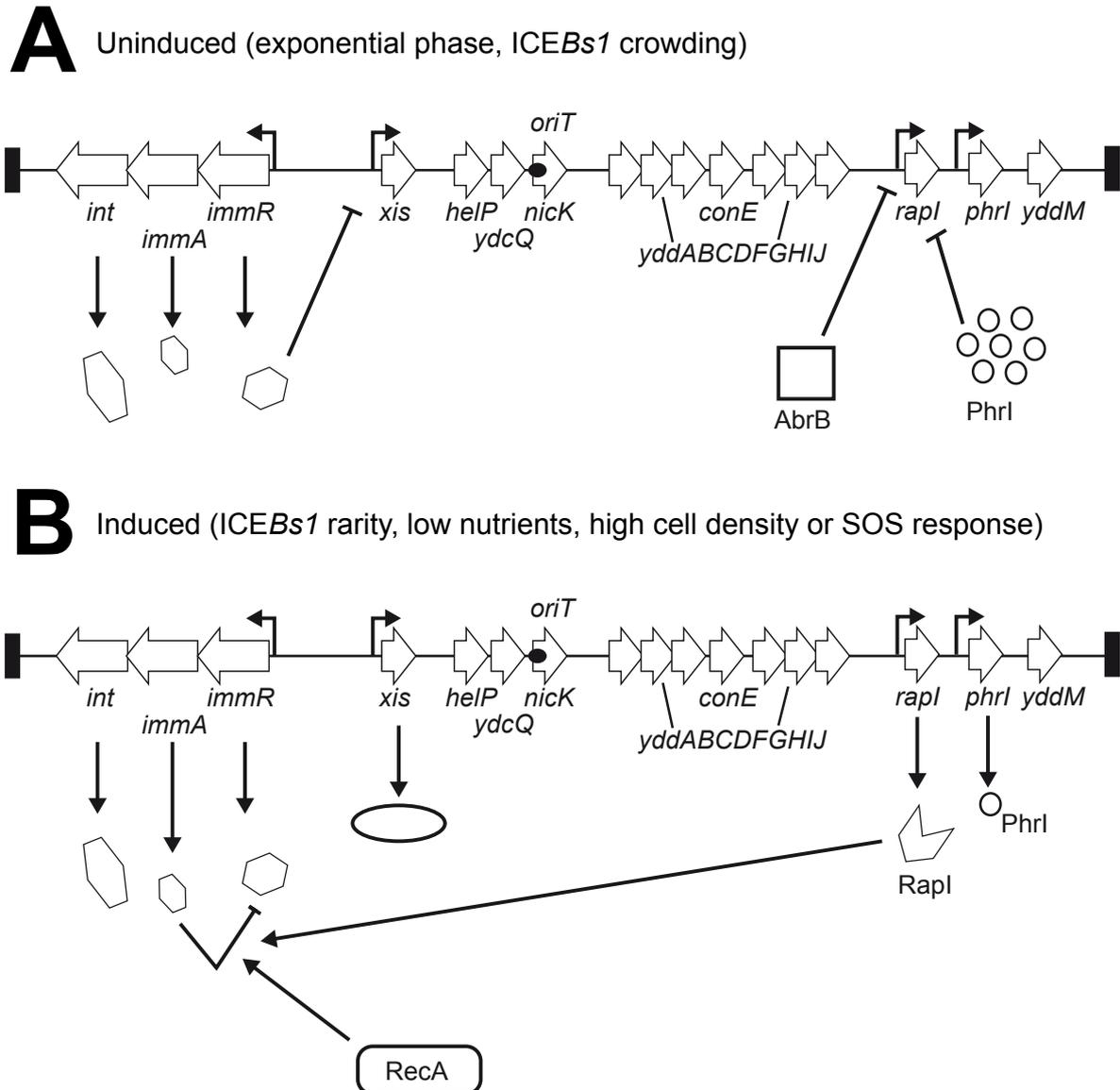


Figure 2. Lifestyle of ICEBs1 of *Bacillus subtilis*. A, Uninduced state. B, HGT induction. Relevant genes are indicated. For explanation, see main text. Modified from Auchtung et al., 2005 [22].

SXT-R391

SXT is a 99.5 kb long genetic element first discovered in *Vibrio cholerae* O139 [32]. SXT confers resistances to chloramphenicol (Cm^{R}), streptomycin (Sm^{R}) and sulfamethoxazole/trimethoprim (SXT^{R} , hence the name). R391 is a closely-related ICE of 89 kb, originally found in *Providencia rettgeri* and conferring resistance against mercury (Hg^{R}) and kanamycin (Km^{R}) [33]. Together, SXT and R391 form a family of

ICEs currently encompassing some 50 known members among a wide spectrum of *γ-proteobacteria* [33].

Interestingly, SXT-R391-type ICEs display an extended array of variability among their cargo regions, but often contain resistance genes, in addition to genes for cyclic-di-GMP biosynthesis or anti-phage restriction/modification [34–36]. Variable regions in all SXT-R391 ICEs are located at defined positions within a common conserved architecture, termed the core backbone [37]. This backbone constitutes some 52 genes resembling (at varying degrees) the core region of plasmids from the IncA-IncC family, equally notorious for bearing antibiotic resistance genes [37,38]. Since the backbone genes are present in all SXT-R391 ICEs and are highly conserved, it is generally assumed that the horizontal transfer of all family members occurs through a similar mechanism as SXT^{MO10} or R391. In the integrated form, SXT (and R391) resides in the extremity of the peptide chain release factor 3 gene (*prfC*). However, secondary integration sites are used when *prfC* is deleted [39,40]. Initiation of the transfer program is induced by the SOS response [41], triggered in *V. cholerae* by DNA-damage through exposure to mitomycin C, ciprofloxacin, or UV radiation. Induction of the SOS response through antibiotic exposure is especially cumbersome in clinical strains of *V. cholerae*, and leads to an increased propagation of SXT-R391 antibiotic resistance types [41]. At the mechanistic level, SOS response triggers RecA- and LexA-dependent proteolysis of the SXT-encoded repressor SetR (Figure 3B). This liberates the *setCD* operon from SetR transcriptional repression [41,42] (Figure 3A). The resulting SetC and SetD then activate the transcription of the integrase gene *int*, the putative T4SS operon *traLEKVAB* and the excisionase gene *xis* [37,39,43] (Figure 3B). Int and Xis mediate SXT excision, yielding a circular form [39]. SXT excision in donors occurs more frequently than successful conjugation into

recipient cells [39], indicating that excision is not necessarily the limiting step in ICE transfer. Presumably, the SXT-encoded putative relaxase *Tral* cuts one of the dsDNA SXT strands at the *oriT* sequence, which is located some 40 kb upstream of *tral* [44,45]. Another SXT protein, *MobI*, encoded in the vicinity of *oriT* and required for transfer initiation, may contribute to *oriT* recognition [45,46]. The nicked *oriT* is then presumed to serve as the starting point for reconstructive replication or for further DNA processing and conjugation.

Reconstructive replication (i.e., the formation of ds- from single-stranded DNA) may be accompanied by further multiplication of SXT, since it was found that the ratio between circular SXT *attP* and chromosomal *attB* after excision was slightly above one [39]. In some SXT-R391 type ICEs, failure to restore a stable SXT in the donor is sanctioned by one or two toxin/anti-toxin systems, *MosT/MosA* and *S044/S045* [47,48]. Expression of the *mosAT* operon increases upon SXT excision, which was interpreted as a mechanism to avoid cellular loss of SXT in the excised state [47]. *MosT* is a bacteriostatic toxin, which is countered by the antitoxin *MosA*. Presumably, the deleterious *MosT* is more stable than *MosA* and therefore, accidental loss of SXT in dividing cells would result in a rapid decline in *MosA* concentration, causing the enduring *MosT* to damage the cell. Upon reintegration of SXT, expression of *mosAT* returns to a relative silent state driven by *MosA* autorepression [47]. A second toxin/antitoxin system can be present in some SXT and involves the products of the *s044* and *s045* genes, a growth-inhibiting toxin and its neutralizing antitoxin, respectively [48]. In this case, the toxin gene *s044* is placed in front of a strong promoter, whereas the antitoxin gene *s045* is placed in front of a comparatively weaker promoter. Details of the balance between the two factors remain to be elucidated.

Transfer of the single-stranded SXT molecule is supposed to be mediated by a putative type IV secretion system, encoded by three operons, *traLEKBVA*, *s054/traC/trhF/traWUN* and *traFHG* [36]. These genes display relatively high homology and identical synteny to those of IncA-IncC conjugative plasmids [36]. As mentioned earlier, expression of *traLEKBVA* is induced by SetC and SetD. Additionally, the gene *s063* was found to be required for effective HGT and might thus be part of the conjugation machinery as well [36]. *traJ* and *traD*, encoded on the same operon as *traI*, were found to be essential to transfer and show homologies to coupling proteins [44]. Most interestingly, the entry of SXT and R391 into a new host is subject to an exclusion system driven by the inner membrane proteins TraG and Eex, the donor and recipient exclusion factors, respectively [49,50]. Indeed, members of the SXT-R391 family of ICEs segregates into two exclusion groups : S (for SXT-like) and R (for R391-like). This system prevents a transferring ICE of a given group (S or R) to settle into a recipient cell that already contains an ICE of the same group. The exclusion system is not water tight and can be bypassed as ICEs of the same group can colonize the same cell [49]. In addition, incoming ICE of a group different from that already in the recipient are not blocked [51]. The exact mechanism of exclusion is not so well established, but its believed that TraG (or part of it) travels to the recipient cell along with the ICE, which was concluded from TraG-Eex interactions taking place in the cytoplasm [50]. After the entrance (and acceptance) in the recipient cell, the single-stranded DNA is replicated to a dsDNA, which is reintegrated into the *prfC* gene by the Int [39]. Int-mediated integration and imperfect (or absent) exclusion can lead to the creation of tandem arrays of SXT-R391 family members [52,53], which are generally unstable and lead to deletions through recombination [52]. Additionally, tandem arrays of SXT-R391 ICEs can lead to the

formation of hybrid ICEs, in part due to RecA activity, although these events are described as more likely to occur in the donor rather than in the recipient [54]. In certain SXT-R391 ICEs, two SXT factors Bet and Exo, upregulated by SetC and SetD, can also generate hybrids by their recombination activity, but the exact mechanism of this is still unclear [52,54,55]. Finally, the whole or part of the HGT machinery of SXT-R391 can help to mobilize other resident compatible genetic elements such as plasmids, genomic islands or even large chunks of chromosomes, casting a new light on the evolutionary impact that ICEs can have on bacterial genomes [35,46,56–58].

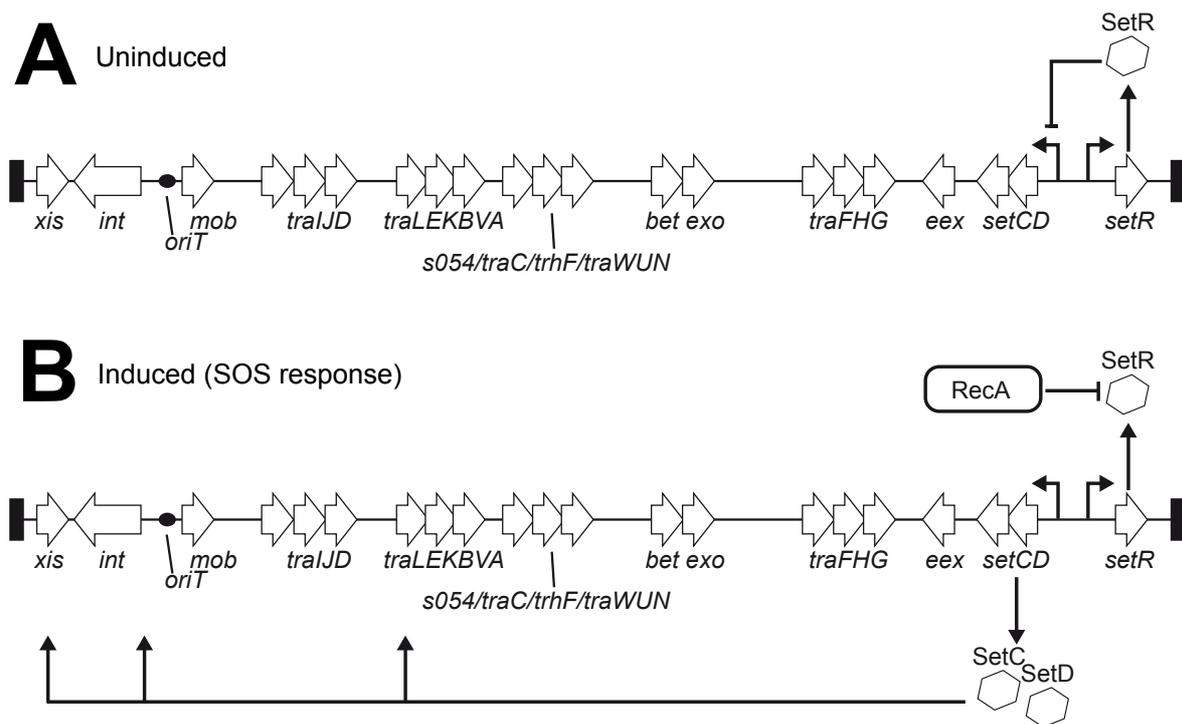


Figure 3. Lifestyle of SXT of *Vibrio cholerae*. A, Uninduced state. B, HGT induction. Relevant genes are indicated. For explanation, see main text. Modified from Wozniak et al., 2009 [36].

CTnDOT

CTnDOT is a 65 kb long ICE integrated at sequence-specific sites in the genome of *Bacteroides thetaiotamicron* [59]. CTnDOT belongs to a wider family of CTn-like ICEs, which are present in all *Bacteroides* species discovered so far [60]. CTnDOT enables its host to resist to tetracycline and erythromycin [59].

Expression from CTnDOT genes is mostly silent in the integrated state, except for three genes that are expressed at constitutive levels: (i) *intDOT*, encoding the IntDOT integrase, (ii) *rteR*, encoding a small RNA repressing the *tra* operon for conjugation [61,62], and (iii) the operon *tetQ-rteA-rteB*, encoding a tetracycline resistance factor (TetQ) and two key HGT regulators (RteA and RteB) [63] (Figure 4A). Interestingly, the *tetQ-rteA-rteB* operon yields a transcript starting with a clear hairpin-forming leader mRNA (downstream of the P_Q promoter but upstream of the *tetQ* start codon) that constitutively produces a leader peptide of 3 amino acids [64,65]. Two possible hairpins have been shown in the leader mRNA, a relatively stable one composed of two sequences named Hp1 and Hp8, and a shorter less frequent one, constituted of Hp1 and another sequence named Hp2. In the absence of induction, the Hp1-Hp8 hairpin prevents the translation of *tetQ*-mRNA by occluding the ribosome binding site for *tetQ*, which is contained within the Hp8 loop [64,65] (Figure 4A). The translation of *rteA* and *rteB* is seemingly affected by the same attenuation mechanism.

Under inducing conditions, identified so far as presence of tetracycline or other ribosome-targeting antibiotics, this translation attenuation is disrupted. Presumably, tetracycline-affected ribosomes tend to stall during leader peptide synthesis, driving the Hp1-Hp2 hairpin to form on the leader mRNA instead of the Hp1-Hp8 hairpin [65] (Figure 4B). The liberated *tetQ* ribosome binding site now enables translation of *tetQ*, *rteA* and *rteB*. RteA phosphorylates RteB, which subsequently binds to promoter P_C,

stimulating the expression of *rteC* [66]. RteC activates the transcription of the *xis2c-2d-exc* operon, by binding its promoter P_E [66–68]. The three gene products of the *xis2c-2d-exc* operon, Xis2c, Xis2d and Exc perform two tasks in CTnDOT transfer: (i) Xis2c, Xis2d and Exc help the IntDOT integrase to catalyze the excision of the ICE and (ii) Xis2c and Xis2d activate the transcription of the *tra* genes and possibly that of the *mob* genes as well [68,69].

Xis2c and Xis2d are essential for the excision of CTnDOT. They bind to *attL* and *attR* and interact with IntDOT, the chromosome-encoded BHF α and Exc to form the so-called excisive intasome complex [70–73]. Exc is a topoisomerase III essential to excision *in vivo* but curiously not *in vitro* [74,75]. Exc enhances the excision reaction possibly by promoting the stability of the intasome at *attR/L* sites [72]. The host factor BHF α , which is required not only for excision but also for integration is a chromosomally-encoded protein that interacts with IntDOT ([76], Ringwald and Gardner unpublished). IntDOT interacts and assembles at the *attL* and *attR* sites with the other factors into the excisive intasome [73,77,78]. IntDOT catalyzes the recombination between *attR* and *attL* asymmetrically, with the unusual formation of a heteroduplex region [72,73,75,77]. The integrated form of CTnDOT is thus converted into a circular form by recombination of *attR* and *attL* into a single *attDOT*.

Besides their essential role in excision, Xis2c and Xis2d additionally promote the transcription of the *tra* genes via an unknown mechanism, which involves Xis2c and Xis2d binding to the promoter P_{tra} [68,69] (Figure 4B). It is believed that the circular form CTnDOT is mobilized and translocated to a recipient cell by the products of CTnDOT-encoded *mob* and *tra* genes, respectively [79–81]. In analogy to other conjugative systems, it is assumed that Mob and Tra proteins nick the CTnDOT circular form at its *oriT*, process and export one of the two resulting single-stranded

DNA molecules into a recipient cell. Reconstructive replication is thought to occur in both the donor and the recipient. The regenerated double-stranded circular form CTnDOT is then integrated into the host chromosome at the *attB* site [70].

The *attB* site is composed of two essential core sites, termed B and B', whereas *attDOT* on CTnDOT contains D and D' core sites [82]. Both B and D cores contain the GTANNTTT sequence, which is recognized by IntDOT. To promote site-specific recombination, it is thought that IntDOT (possibly as 4 monomers) assembles at *attDOT* along with the host factor BHFa (in unknown stoichiometry), to form the integrative intasome nucleoprotein complex. This integrative intasome differs from its excisive counterpart by the absence of Xis2c, Xis2d and Exc. The complex interacts with the *attB* site and recombines both *attDOT-attB* sites, regardless of the heterology of the coupling pairs [76,83–86]. In the new host cell, CTnDOT resumes its silent state (at least concerning HGT genes). Like many other mobile genetic elements, CTnDOT has been shown to mobilize certain transposons in natural conditions via its *mob* and *tra* genes [87–89] and exert an important transcriptional influence on the chromosome, via its RteA/RteB regulators [90]. As for SXT-R391-type ICEs, CTnDOT-like hybrids, such as CTn12256 have been described in *Bacteroides* species [91]. It is very interesting to note that, like ICE*clc* (see below), excision in CTnDOT occurs only in small percentage of donor cells [71].

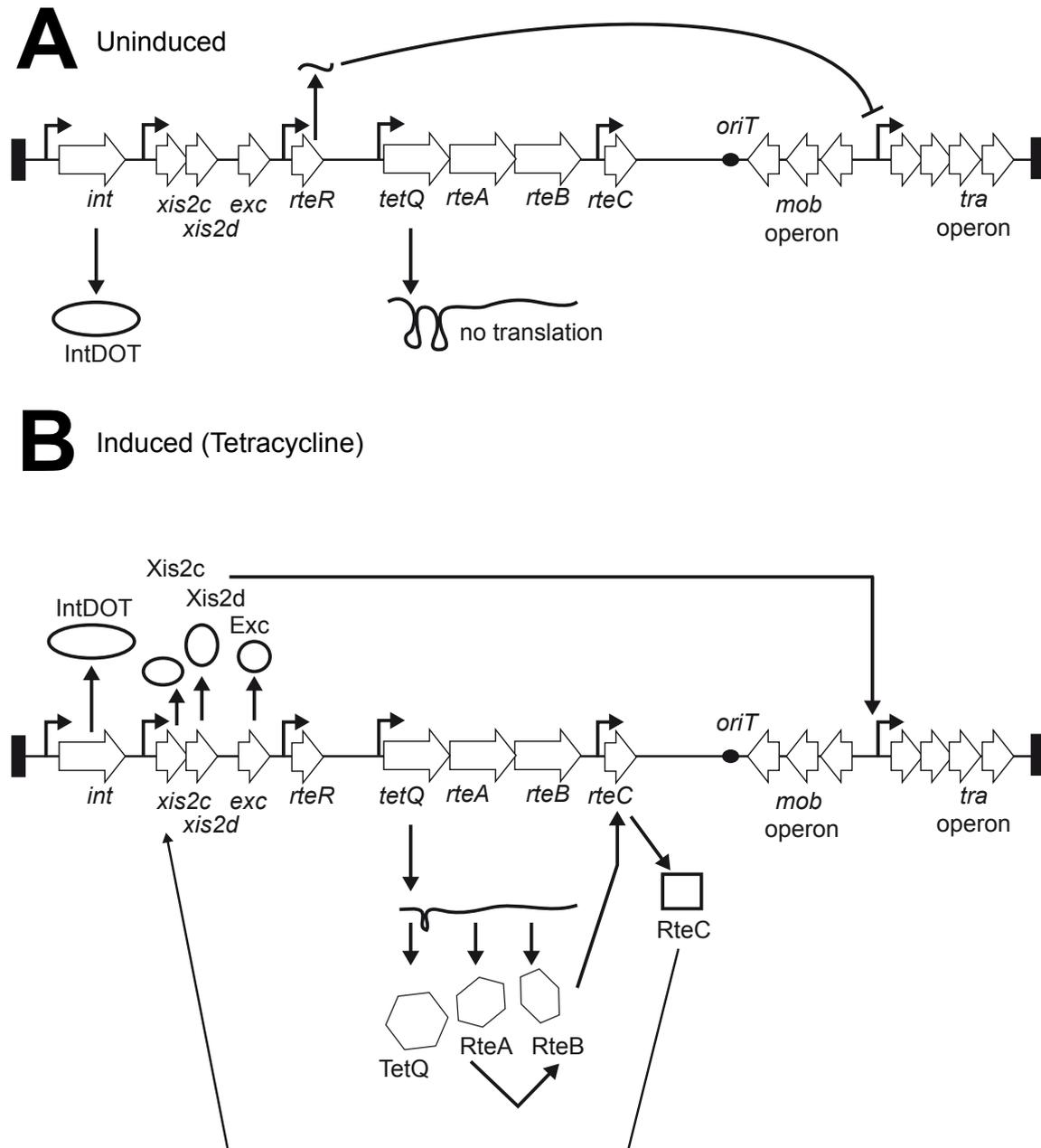


Figure 4. Lifestyle of CTnDOT of *Bacteroides thetaiotamicron*. A, Uninduced state. B, HGT induction. Relevant genes are indicated. For explanation, see main text. Modified from Jeters et al., 2009 [61].

Tn916

Tn916 is an 18 kb long ICE found in *Enterococcus faecalis* DS16 at multiple AT-rich sites on the chromosome [92]. Tn916 belongs to a wide family of ICEs, confers

tetracycline resistance to its host (by the *tetM* gene) and can transfer to several gram-positives and gram-negatives [93,94].

In term of HGT regulation, Tn916 is generally silent when in the integrated state, since the gene product of *orf9* represses two regulator genes, *orf7* and *orf8* by inhibiting activity of the P_{orf7} promoter [95] (Figure 5A). Additionally, expression of the tetracycline resistance gene *tetM* is initiated from the P_{tetM} promoter but ribosomes tend to stall on the leader sequence mRNA, seemingly enabling hairpins to form and terminate transcription prematurely [96].

In presence of tetracycline, affected ribosomes somehow proceed past the mRNA leader sequence, preventing the formation of the hairpin and thus allowing the polymerase to read through to *orf6*, *orf9* (non-coding strand), *orf7* and *orf8* (Figure 5B). The exact mechanism of this attenuation remains elusive but may involve increasing amounts of charged tRNA in the cell. This over-extended transcription yields an antisense *orf9* RNA that inhibits the translation of *orf9*, thus alleviating the repression of Orf9 on P_{orf7}. The produced Orf7 and Orf8 proteins now autoactivate their own expression by stimulating P_{orf7}. Enhanced transcription from P_{orf7} not only produces more Orf7 and Orf8, but also reads through to *xis* and *int* genes. Production of Xis and Int ultimately leads to the excision of Tn916 [95–97] (Figure 5B).

The Int enzyme is a tyrosine recombinase that mediates the excision of Tn916 to creating the circular form Tn916 [98–100]. Int requires the excisionase Xis for efficient Tn916 excision [101–103] and both proteins bind DNA at or close to the *attR* and *attL* sites [104–107]. Interestingly, Xis-*attL* interaction promotes excision, contrary to Xis-*attR* that inhibits it [101]. Similarly to IntDOT of CTnDOT, Int from Tn916 does not require homology between the coupling sequences in *attR* and *attL*

and thus creates heteroduplexes from the coupling sequences [106,108–111]. Since the recombination event is not site-specific but solely requires AT-rich regions, Tn916 has been used to generate random mutations for scientific studies [112]. Similarly to many ICEs, excision is a rare event, but its frequency can be affected by the host background and like CTnDOT or SXT, is not the limiting step of transfer [99,113].

Once the circular form Tn916 is formed, the *tecLKJIHGFEDCBA* operon for the conjugation machinery, which is promoterless in the integrated form, is placed directly downstream of the *xis-int* operon [114,115] (Figure 5C). Thus, only in this circular configuration, P_{xis} and P_{orf7} promote the synthesis of long mRNAs extending down to the *tec* operon by reading through *attP* [95]. Therefore, this mechanism guarantees the transcription of transfer and mobilization genes only once excision and circularization have successfully occurred.

Subsequently, the produced relaxase TechH associates to Int and catalyzes single-strand nicks at the *oriT*, located just in the intergenic region upstream of *techH* [116–118]. Interestingly, Int is necessary to give site- and single-strand DNA specificity to the relaxase, an interaction hypothesized to add another level of control on the HGT of Tn916 [118]. Presumably, a nicked single-stranded Tn916 is translocated into the recipient cell, where it is believed to be reconstituted as a double stranded circular molecule prior to its integration in the host chromosome [114]. The protein ArdA, encoded on the *tec* operon, lures and sequesters the recipient's type I restriction/modification enzymes, thus providing some level of immunity to Tn916 [119].

Contrary to many site-specific integrases, Int can mediate recombination at various AT-rich or bent DNA sites, thus resulting in an unusual multiplicity of diverse insertion sequences [94,106,110,114,120–122]. As for the excision, Int catalyzes

recombination in a homology-independent manner between *attB* and *attP*, creating *attR* and *attL* [100,108,111,114,123]. In order to stay stably embedded in the chromosome, Xis prevents re-excision by binding *attL* in place of Int [101].

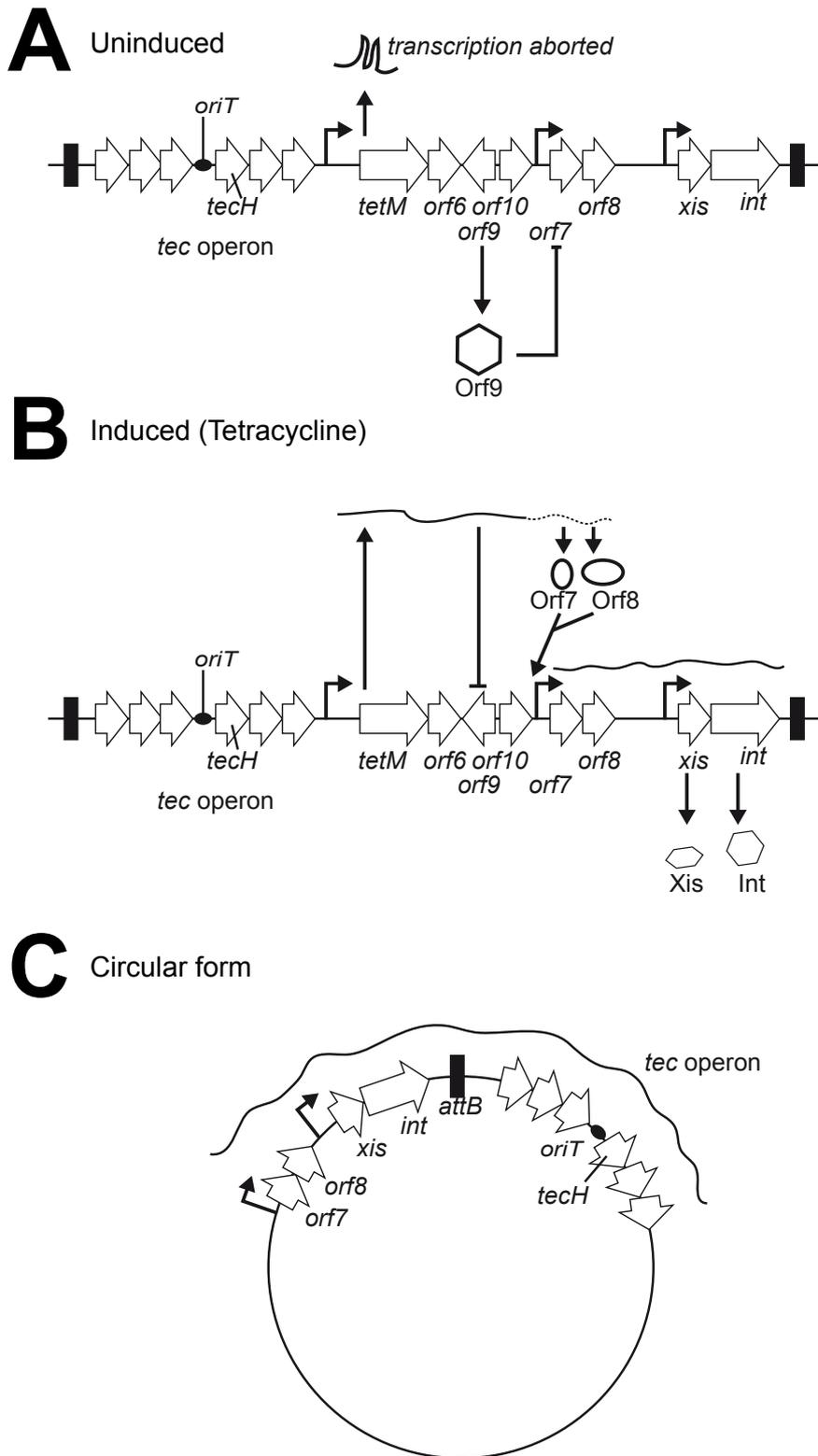


Figure 5. Lifestyle of Tn916 of *Enterococcus faecalis* DS16. A, Uninduced state. B, HGT induction. Relevant genes are indicated. For explanation, see main text. Modified from Roberts and Mullany, 2011 [199].

pSAM2

pSAM2 is a 10.9 kb ICE that resides in *Streptomyces ambofaciens* and is the model for its large family distributed among *Actinomycetes* species, therefore called AICEs (Actinomycetes ICEs). Insofar as known, no specific phenotypes are conferred by pSAM2. pSAM2 distinguishes itself from other ICEs by its exotic DNA translocation mode as well as its clear autonomous replication. A transcriptional repressor KorSA suppresses expression of the *pra* gene as well as its own, by binding specifically to both of these promoters [124] (Figure 6). The *pra* gene encodes the major pSAM2-HGT activator Pra, which is produced only when KorsA is inhibited. This is thought to occur upon acquisition of a (mysterious) signal molecule arising from donor-recipient contacts. The resulting Pra activates the transcription of the *repSA-xis-int* operon that leads to Xis/Int-mediated excision and subsequent RepSA-mediated replication [125,126]. The excisionase Xis and the integrase Int catalyze the recombination between *attR* and *attL*, without the need of further factors [127–132]. RepSA is a replicase that initiates rolling-circle replication by catalyzing a double stranded nick at the so-called origin *dso*, a step essential for HGT [126,133–135]. Subsequently, a copied double-stranded DNA is directly translocated into a recipient cell by a dedicated mating apparatus composed of TraSA (a dsDNA translocase) and the putative transmembrane proteins SpdA, SpdB, SpdC and SpdD [133,136,137]. Whether transfer occurs from the mycelium's edge remains to be demonstrated.

pSAM2 encodes a special immunity factor, Pif, that prevents donor-donor ICE transfer, probably by inhibiting the import of pSAM2 in a strain that already possesses it [138]. Once in the recipient cell, pSAM2 re-circularizes, replicates and reintegrates into a tRNA^{Pro} gene via Int-mediated recombination [128,130,131,135,139,140].

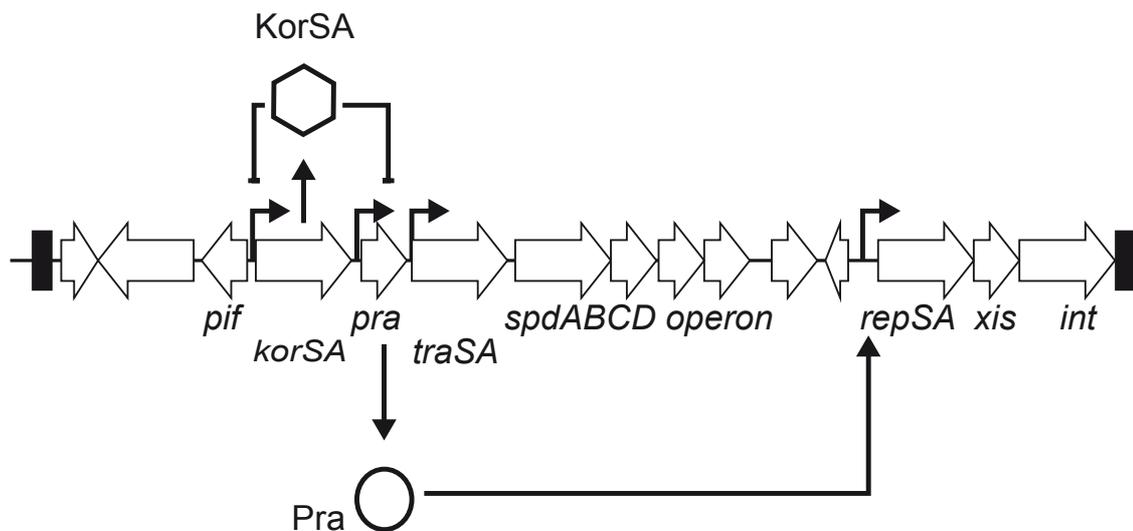


Figure 6. Lifestyle of pSAM2 of *Streptomyces ambofaciens*. Relevant genes are indicated. For explanation, see main text. Modified from Poele et al., 2008 [200].

ICEMISym^{R7A}

ICEMISym^{R7A} is a 502 kb ICE residing in the chromosome of *Mesorhizobium loti* R7A, in the single tRNA^{Phe} gene [141,142]. The ICE can be transferred to several other bacteria species in the soil and contains genes enabling nitrogen fixation and nodule formation in mutualistic interactions with plants [143,144].

The HGT of ICEMISym^{R7A} demonstrates an interesting regulatory system. The balance between integrated and excised state for ICEMISym^{R7A} is controlled by the regulators QseC and QseM (Figure 7A). QseC is a transcriptional regulator which controls both its own and *qseM* expression. The *qseC* and *qseM* are adjacent but divergently oriented and control is exerted through binding of QseC to two operators (O_R and O_L) [145]. QseM is an inhibitor of the quorum-sensor regulator TraR, which induces the ICEMISym^{R7A} transfer genes. At low population density, excision is prevented because QseC preferentially binds the O_L operator, which at low QseC

concentrations in the cell stimulates *qseC* expression but does not impede *qseM* expression (Figure 7A). At high population density and stationary phase, QseC accumulates in cells, which at the highest levels - assumed to arise only in a small fraction of cells (6%), cause QseC to bind simultaneously to both operators O_L and O_R . This configuration represses the transcription of both *qseC* and *qseM* genes (Figure 7B). Absence of the allosteric inhibitor QseM permits the quorum-sensor regulator TraR to respond to HGT-inducing AHL signals, which are produced by TraI. TraR induction stimulates both *traI* expression and that of the *msi171* and *ms172* genes. Ms171 and Ms172 are essential for ICEMISym^{R7A} excision and transfer [145–147]. ICEMISym^{R7A} excision is mediated by the integrase IntS with the support of the excisionase RdfS [146]. ICEMISym^{R7A} transfer requires a T4SS (*trb* operon) and the putative relaxase Rlx. Interestingly, Rlx is also essential for the maintenance of ICEMISym^{R7A}, although its exact function in this is unclear [146,147].

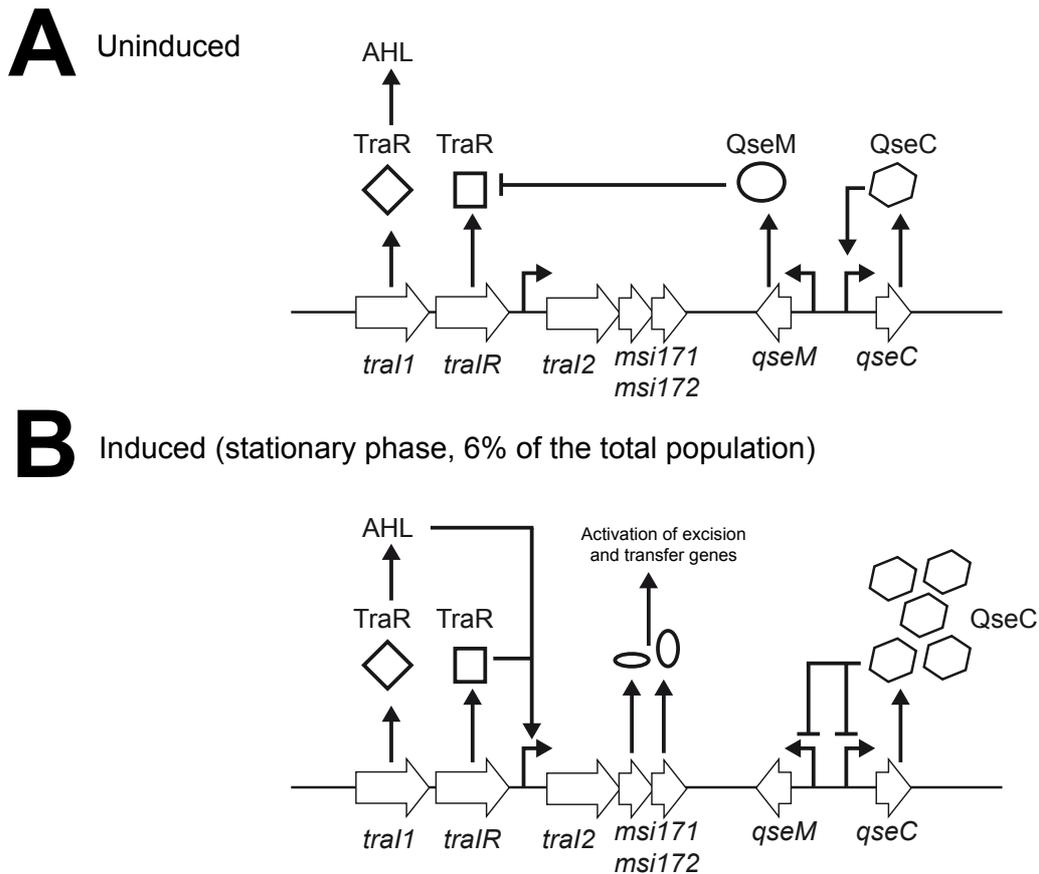


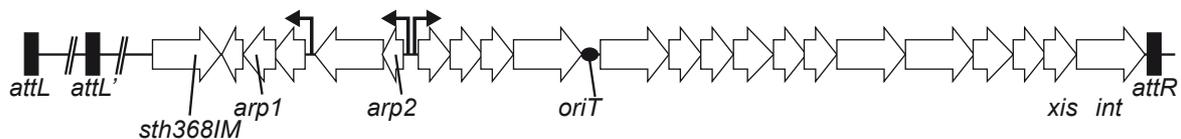
Figure 7. Lifestyle of ICEM/Sym^{R7A} of *Mesorhizobium loti* R7A. A, Uninduced state. B, HGT induction. Relevant genes are indicated. For explanation, see main text. Modified from Ramsay et al., 2013 [145].

ICESt1 and ICESt3

ICESt1 is a 34.7 kb ICE integrated into the genome of *Streptococcus thermophilus* CNRZ368 [148][149]. It is inserted at the end of the *fda* gene, the same insertion site that is used by its close relative ICESt3 [150]. Curiously, ICESt1 encompasses as secondary *attL* site, *attL'* that can be used by the excision machinery with *attR* and therefore yields a truncated alternative ICE, termed ICESt2, instead of the larger ICESt1 [150] (Figure 8). Similarly to other ICEs, ICESt1 can mobilize other ICESt-like GIs [149,150]. Induction of the HGT program of both ICESt1 and ICESt3 involves the SOS response (typically mediated by mitomycin C) and/or stationary phase [151]. Mitomycin treatment represses two phage like repressors, Arp1 and ImmR (Arp2),

which show homology to the phage λ *cl* repressor and *cl*-like repressor, respectively [151–153]. *Arp1* and *ImmR* are encoded on both *ICESt1* and *ICESt3*, and repress the excision and transfer of both ICEs (Figure 8). However, upon *RecA* stimulation, *Arp1* and *ImmR* undergo self-cleavage. The activation cascade is not known, but since *ICESt1* and *ICESt3* have different regulation modules, despite identical core regions, it is thought to be different [154]. Indeed, *ICESt1* encompasses two regulatory operons, whereas *ICESt3* possesses only one regulatory operon [154]. Excision requires both the integrase *Int* and the excisionase *Xis*, but *Int* is sufficient for integration [153]. Transfer has been demonstrated between *Streptococcus thermophilus* and *Enterococcus faecalis* or *Enterococcus lactis*. Interestingly, the copy number of *ICESt3* dramatically increases after mitomycin treatment, thus suggesting active replication [154]. Additionally, *ICESt1* encodes a novel type II restriction/modification system, named *Sth368I*. The system is composed of *Sth368IR* and *Sth368IM*, which are responsible for the immunity to Φ ST84. They mediate restriction cleavage at GATC sites, as well as confer protection by methylation of the terminal C of this same sequence in native DNA [119].

ICESt1



ICESt3

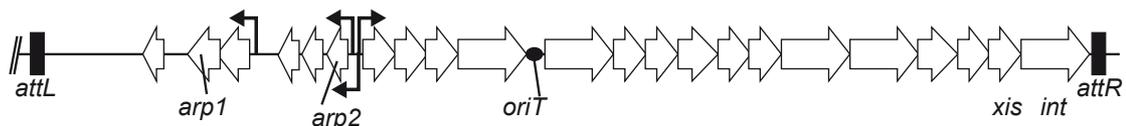


Figure 8. Lifestyle of *ICESt1* and *ICESt3* of *Streptococcus thermophilus* CNRZ368. Relevant genes are indicated. For explanation, see main text. Modified from Carraro et al., 2011 [154].

HPI

High Pathogenicity Island (HPI) is a 102 kb long ICE originally found in *Yersinia pestis* [155,156]. It encodes siderophore (yersiniabactin) biosynthesis and transport allowing iron-scavenging in human environment, thus conferring its enhanced pathogenicity phenotype [157]. HPI can be transferred to other *Yersinia* species or to other *Enterobacteriaceae* [158,159]. HPI encodes Int, a recombinase of P4-phage family, which mediates integration into the *attB* site in tRNA^{Asn} genes. Int requires an auxiliary Xis (equally referred to as Hef) for excision to occur [160–165]. Not more is known about the mechanisms of HPI transfer.

ICE*clc*

Previously referred to as the "CLC element", ICE*clc* is a 103 kb-long element that resides in two copies in its original host, *Pseudomonas knackmussii* B13 [166,167]. ICE*clc* is composed of different genetic modules, (i) the cargo module, (ii) the site-specific recombinase gene *intB13* and (iii) a core module (Figure 9A).

The cargo module in ICE*clc* is defined as a roughly 50 kb region delimited between the recombinase gene *intB13* at its left and the core region at its right. This cargo or "variable" module contains various gene clusters, among which two operons encoding the metabolic pathways for 3- or 4-chlorocatechol (CLC) degradation (*clc* genes) and for 2-aminophenol (2AP, *amn* genes) degradation [168] (Figure 9A).

Pseudomonas knackmussii B13 was originally isolated from sewage sludge by its propensity to mineralize 3-chlorobenzoate (3CBA) as sole carbon and energy source [166]. 3CBA metabolism requires the combination of genes on the host chromosome (i.e. oxidation of 3CBA to 3CLC), and the *clc* genes of ICE*clc*. This so-called lower

pathway (i.e. from 3CLC to succinyl-CoA) is catalyzed by enzymes encoded in the *clcEDBA* operon [168]. Similarly, complete oxidation of chlorobenzene (MCB) or 1,4-dichlorobenzene (1,4-DCB) is rendered possible by the introduction of ICE*clc*'s CLC pathway in bacteria expressing the toluene degradation *tod* genes [167,169]. Thus, in the appropriate host background (i.e. containing *cba* or *tod* genes), ICE*clc* confers the ability to use 3CBA or MCB/1,4-DCB as sole carbon and energy sources [167,168,170]. Since CLCs are the intermediate products in metabolism of other chlorinated aromatic compounds, such as chlorophenols or chloroaniline, the biochemical spectrum enhanced by ICE*clc* acquisition is suspected to be even larger [168,171].

The second clearly identified and confirmed catabolic operon is the *amnRBACDFEHG*, whose products catalyze 2AP-degradation via aromatic ring *meta* cleavage, thus allowing growth on 2AP as sole carbon and energy sources [168,170]. Additionally, ICE*clc*'s cargo module is predicted to encode further metabolic functions. This includes putative efflux pumps, a putative anthranilate-like/salicylate-like dioxygenase, and a putative nitrilase, but these have not further been experimentally investigated [168].

The cargo region of ICE*clc* is presenting an ecological advantage to the host cell, since it enables the use of at least two new substrates, CLC and 2AP. This ecological advantage has not only been documented under laboratory conditions but has turned out to be evolutionary successful, since other reports have demonstrated the propagation of ICE*clc* in engineered communities and even further gene acquisition on natural ICE*clc*-like elements. As example, inoculation of a waste-water bioreactor with *P. knackmussii* B13 showed extensive ICE*clc* transfer to native *Ralstonia* species, which colonized the reactors and carried out effective 3CBA degradation

[172]. Natural ICE*clc*-derivatives were found in *Ralstonia* sp. strain JS705, a MCB-degrader isolated from contaminated groundwater carrying a 12 kb insertion into ICE*clc*. This gene fragment contained genes for a multicomponent toluene dioxygenase (*mcb*) [173]. *Burkholderia xenovorans* LB400, a chlorobiphenyl-degrading organism, carries an ICE*clc* with two further gene insertions, one of which is 20 kb long and encodes an auxiliary *ortho*-halobenzoate 1, 2-dioxygenase (*ohbAB*) and proteins putatively involved in secretion/transport [168,174]. Also, an ICE*clc*-like element is found in *Bordatella petrii* DSM12804, termed GI-3, which contains the *clc* operon and the putative anthranilate dioxygenase genes but not the *amn* operon [175,176]. Taken together these observations show that ICE*clc* elements have rather flexible composition of their "cargo gene zone", which enables an ecological advantage to the host cells in anthropogenic conditions, like polluted groundwater or waste water.

Flexibility of the cargo zone is further evident from analysis of other ICE*clc*-like elements. Three related other ICEs in *B. petrii* DSM12804, GI-1, GI-2 and GI-6 have very different sets of cargo genes [175,176]. As examples, GI-1's cargo module encodes the degradation of an unknown aromatic compound, whereas GI-2's cargo genes encode the metabolism of benzoate, benzylalcohol and 3-hydroxybenzoate, and GI-6 is predicted to encode an iron transport system and multidrug efflux pumps [175,176]. In *Acidovorax* JS42 and *Cupriavidus metallidurans* CH34, ICE related to ICE*clc* are found which contain heavy metal resistance genes [177–179]. The ICE*clc*-like element PAGI-2 in *Pseudomonas aeruginosa* strain C, has a cargo module encoding heavy metal complexation, transport and resistance functions, whereas PAGI-3 bears antibiotic resistance genes and metabolism traits [180–183]. Similarly, the ICE*clc*-like "pathogenicity" elements pKLC102 and PAPI-1 of *P. aeruginosa* strain

C and PA14, respectively, encode virulence factors involved in biofilm formation and adhesion to eukaryotic cells [184–187]. *Herminiimonas arsenicoxydans* encompasses an ICE c/c -like element, whose cargo module encodes arsenic resistance/detoxification genes and formaldehyde oxidation enzymes [188]. Finally, the ICE $Hin1056$ -subfamily of ICEs in *Haemophilus influenzae* has cargo modules dedicated to various antibiotics resistances [189,190].

The recombinase gene *intB13* is the first gene encoded on the integrated form ICE c/c and lays immediately downstream of the attachment site *attR*. The enzyme IntB13 is an atypically long site-specific tyrosine recombinase of the phage P4 family [13,191]. IntB13 mediates the recombination between the integrated form *attL* and *attR* (i.e. excision), as well as that between the circular form *attP* and the chromosomal *attB* (i.e. integration) [192]. Integration takes place in the 3'-end of genes for tRNA^{Gly}, which contains an 18-bp sequence identical to that in the *attP* region [191,192]. ICE c/c can integrate in several tRNA^{Gly}-genes, of which there are usually multiple copies, which leads to the occasional formation of multiple insertions, including tandems like for SXT [52,192]. Whether an excisionase factor (Xis) is required for guiding the Int-mediated excision, like in many ICE models, is not known for ICE c/c . Expression of *intB13* is directed by either one of two conformational promoters, P_{int} and P_{circ} [193,194]. In the integrated form of ICE c/c , P_{int} controls *intB13* expression as it is placed between the *attR* and *intB13* (Figure 9A). A strong constitutive promoter named P_{circ} is positioned at the opposite end of ICE c/c , upstream of *attL* (Figure 9A). In the circular form, *attL* and *attR* are recombined, plugging P_{circ} directly in front of P_{int} and *intB13* (Figure 9B). In such conformation, the strong promoter P_{circ} overrides *attP* and P_{int}, therefore constitutively expressing *intB13*. This regulation of *intB13*

expression in two ICE*c/c* states is somewhat reminiscent of the Tn916 *tec* operon expression (see above). The high expression of IntB13 from the circular form is assumed to guarantee the efficient reintegration of ICE*c/c* in recipient and (perhaps also) in donor cells [192].

Expression from P_{int} is rather curious because it becomes active in stationary phase cells but only in a small fraction of the total cell population (3-5%), which have been named "transfer competent (tc) cells" [193,195]. Direct visualization experiments demonstrated that ICE*c/c* HGT only occurred from cells expressing a fluorescent reporter fused to P_{int} , suggesting that ICE*c/c* can only transfer from such tc cells [195]. Activation of the tc pathway was proposed to lead to a bistable state of cells; the majority, in which ICE*c/c* remains silent and a minority, in which ICE*c/c* prepares for its transfer. Regulation of the switch between integrated and tc state is not yet well understood, although several factors contributing to it have been identified. Activation of P_{int} has been shown to be dependent on a gene named *inrR*, which is encoded on ICE*c/c* as well, and cells expressing P_{int} equally express the promoter of the *inrR* operon, named P_{inR} [196]. The role of InrR is not well understood, except that it does not directly bind to the P_{int} promoter. Expression of P_{int} and P_{inR} is optimal with the starvation sigma factor RpoS, but, intriguingly, two copies of *rpoS* in cells lead to higher levels of cells expressing P_{int} and P_{inR} promoters [197]. Correlating RpoS-mCherry fusion levels in cells with P_{int} and P_{inR} promoter activity suggested that only cells with highest RpoS-levels are "prone" to activate ICE*c/c* [197]. Contrary to other ICE, SOS-response stimuli did not induce P_{int} expression or ICE*c/c* excision [193].

The core module of ICE*clc*, roughly the second half 50 kb of ICE*clc*, is widely conserved among a large class of related ICE in other bacteria. It is supposed to encode the conjugative transfer system and further regulatory aspects, based on homologies to conjugative plasmids and other ICEs [168] (Figure 9A).

The first gene in the core region (*orf50240*) encodes a relaxase that catalyzes alone the single-strand nick at two independent origins of transfer in ICE*clc* (*oriTs*) [198]. These unique dual *oriTs* have the same efficiency and both contain (dissimilar) sequence repeat motifs, but are possibly differently processed by the relaxase. The first *oriT* locates in the intergenic region upstream of the relaxase gene, whereas the second one lays within the *inrR* gene. It has been demonstrated that the presence of two *oriTs* synergistically enhances transfer frequencies of ICE*clc* [198]. Sequence homologies suggest that this aspect may be conserved among close ICE*clc* relatives [198].

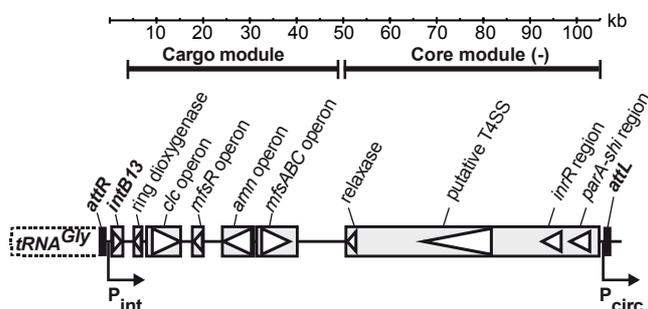
The central core module encodes putative functions with (poor) homologies to known type IV secretion systems (T4SS), like homologues to the channel-energizing VirB4 protein (*orf59888*), the coupling protein VirD4-TraG (*68987*), the pili component PilL (*orf73676*), the relaxosome-related helicase (*orf75419*), the topoisomerase III (*orf91884*) and Ssb (*orf94175*) [168]. However, no experimental data are available for these factors.

Beside the integrase activator *inrR* already mentioned earlier, the end of the core region encodes several genes recently investigated [195]. Genes with homologies to *parA-soj* and *parB* as well as the small orf *shi* in between are involved in formation of the so-called "tc cell microcolonies" or TCM. TCM is a phenotype observed in certain tc cells that result in a poorly growing microcolony following re-exposure to nutrients after prolonged starvation. tc cells in TCM divide only a restricted number of times,

and tend to adopt elongated shapes and lyse more frequently than non-tc cells. All daughter cells in a TCM are "locked" in this state, yielding microcolonies with no more than a dozen of cells at best [195]. Heterologous expression of *shi* and part of *parA* caused slow cell division and formation of elongated cells in *P. putida* without ICE*clc*. Specific deletion of *shi* abolished TCM formation and slightly affected transfer rates, from which it was concluded that limited donor cell division may increase the chances of meeting a potential recipient [195]. Simultaneously, the higher frequency of cell death (lysis) observed in tc cells might have evolved to doom any ICE*clc* donor in which the risk of ICE*clc* loss is greatest. This way, ICE*clc* would guarantee its own maintenance by preventing the emergence of ICE*clc*-free cells. Since this process is detrimental to the fate of tc cells, the appearance of tc cells is limited to a small fraction of the total population of ICE*clc* host, or else the survival of the population as a whole would be jeopardized.

Independently of the advantages brought by the cargo module, ICE*clc* has been shown to impose very little metabolic burden on its host cell, thus limiting fitness cost [170]. This capacity to render itself "stealthy" is probably mediated by the regulation of ICE*clc* and thus contributing to its evolutionary success.

A ICE*clc* integrated form



B ICE*clc* circular form

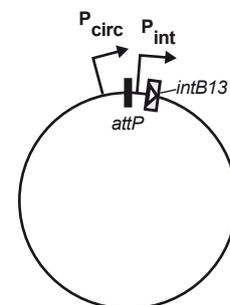


Figure 9. Lifestyle of ICE*clc* of *Pseudomonas knackmussii* B13. A, integrated form. B, circular form. Relevant genes are indicated. For explanation, see main text. Modified from Gaillard et al., 2010 [201].

Aim of this thesis

At the onset of this thesis work, the identification of the ICE*clc* transfer machinery besides the recombinase IntB13, was solely based on homology predictions in the recently sequenced ICE*clc*. Moreover, the regulatory network controlling ICE*clc* HGT was largely unexplored. The fascinating existence of the two conformation-dependent integrase promoters, P_{circ} and P_{int} was known since a decade and we had already discovered that only a small fraction of the total cell population activated P_{int} [167,191,193,194]. Emphasis on this bistability phenomenon was beginning with the discovery of the role and expression of *inrR*, an *intB13* activator [196]. No other regulators had been identified, although a *parB*- and *alpA*-homolog were suspected to be involved in repression of transfer. In other ICE models, major HGT repressors/activators were known, such as RapI/PhrI in ICEBs1, SetR/SetCD in SXT, KorSA/Pra in pSAM2 [22,39,124,125]. Additionally, tetracycline-based attenuation mechanisms were known to control the activator cascades RteA-RteB-RteC-Exc in CTnDOT and Orf7-8 in Tn916 [65,96]. Nothing similar was available for ICE*clc*, except that stationary phase starvation preceded by 3CBA metabolism was required for optimal ICE*clc* transfer to arise.

The major aims of this work were thus to discover the factors involved in the regulation of the ICE*clc* transfer and if possible to explain their mechanisms of regulation.

In the next chapter, microarray-based experiments were used to understand expression of ICE*clc* in exponential and stationary phase. By overlaying transcriptomic profiles with previous experimental data from fellow researcher Dr M. Gaillard, it was possible to establish a transcription map for the entire core region of

ICE*clc*, a region known to encode most predicted HGT functions. Simultaneously, we also demonstrated how transcription of the ICE*clc* core is maximal in stationary phase, in the time frame where we knew that expression of key ICE*clc* promoters occurs.

In the third chapter, we present the transcriptomes of ICE*clc* in a variety of different host species, in order to explore whether there are species-specific differences.

In the fourth chapter, we focus on the role of a curious ICE*clc*-encoded TetR-type transcriptional repressor. We find that this gene, which we name *mfsR*, not only controls its own expression but that of a set of genes for a multi-drug efflux pump (*mfsABC*). By using a combination of techniques from molecular biology, biochemistry and 'omics technology, we could show that MfsR specifically binds to operator boxes in two ICE*clc* promoters (P_{mfsR} and P_{mfsA}), inhibiting the transcription of the two *mfsR* and *mfsABC-orf38184* operons. Although we could not detect a clear phenotype of an *mfsABC* deletion, we discuss the implications of pump genes reorganizations in ICE*clc* and close relatives.

In the fifth chapter, we find that *mfsR* not only controls its own expression and that of the *mfsABC* operon, but also encodes a major factor in control of ICE*clc* transfer. Using gene deletions, microarrays, transfer assays and microscopy-based reporter fusions, we demonstrate that *mfsR* actually controls a small operon of three regulatory genes. The last gene of this *mfsR* operon, *orf17162*, encodes a major activator that enables initiation of the transfer-competence phenotype. Interestingly,

deletion of *mfsR* leads to transfer competence in almost all cells, thereby overruling the bistability process in the wild-type.

REFERENCES

1. Ochman H, Lawrence JG, Groisman EA (2000) *Lateral gene transfer and the nature of bacterial innovation*. *Nature* 405: 299–304.
2. Andam CP, Gogarten JP (2011) *Biased gene transfer in microbial evolution*. *Nature Reviews Microbiology* 9: 543–555.
3. Beiko RG, Harlow TJ, Ragan MA (2005) *Highways of gene sharing in prokaryotes*. *Proceedings of the National Academy of Sciences of the United States of America* 102: 14332–14337.
4. Gogarten JP, Townsend JP (2005) *Horizontal gene transfer, genome innovation and evolution*. *Nature Reviews Microbiology* 3: 679–687.
5. Frost LS, Leplae R, Summers AO, Toussaint A (2005) *Mobile genetic elements: the agents of open source evolution*. *Nature Reviews Microbiology* 3: 722–732.
6. Thomas CM, Nielsen KM (2005) *Mechanisms of, and barriers to, horizontal gene transfer between bacteria*. *Nature Reviews Microbiology* 3: 711–721.
7. Osborn M, Bol (2002) *When phage, plasmids, and transposons collide: genomic islands, and conjugative- and mobilizable-transposons as a mosaic continuum*. *Plasmid* 48: 202–212.
8. Burrus V, Pavlovic G, Decaris B, Guédon G (2002) *Conjugative transposons: the tip of the iceberg*. *Molecular microbiology* 46: 601–610.
9. Wozniak RAF, Waldor MK (2010) *Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow*. *Nature Reviews Microbiology* 8: 552–563.
10. Burrus V, Waldor MK (2004) *Shaping bacterial genomes with integrative and conjugative elements*. *Research in Microbiology* 155: 376–386.
11. Rajeev L, Malanowska K, Gardner JF (2009) *Challenging a paradigm: the role of DNA homology in Tyrosine recombinase reactions*. *Microbiology and Molecular Biology Reviews* 73: 300–309.
12. Van Houdt R, Leplae R, Lima-Mendez G, Mergeay M, Toussaint A (2012) *Towards a more accurate annotation of tyrosine-based site-specific recombinases in bacterial genomes*. *Mobile DNA* 3: 1–11.
13. Roelof van der Meer J, Ravatn R, Sentchilo V (2001) *The *clc* element of *Pseudomonas* sp. strain B13 and other mobile degradative elements employing phage-like integrases*. *Archives of microbiology* 175: 79–85.
14. Mullany P, Roberts AP, Wang H (2002) *Mechanism of integration and excision in conjugative transposons*. *Cellular and Molecular Life Sciences* 59: 2017–2022.

15. Williams KP (2002) *Integration sites for genetic elements in prokaryotic tRNA and tmRNA genes: sublocation preference of integrase subfamilies*. *Nucleic Acids Res* 30: 866–875.
16. Guglielmini J, Quintais L, Garcillán-Barcia MP, de la Cruz F, Rocha EPC (2011) *The repertoire of ICE in prokaryotes underscores the unity, diversity, and ubiquity of conjugation*. *PLoS Genetics* 7:8 e1002222.
17. Matic I, Taddei F, Radman M (1996) *Genetic barriers among bacteria*. *Trends in microbiology* 4: 69–73.
18. Hacker J, Carniel E (2001) *Ecological fitness, genomic islands and bacterial pathogenicity*. *EMBO reports* 2: 376–381.
19. Van der Meer JR, Sentchilo V (2003) *Genomic islands and the evolution of catabolic pathways in bacteria*. *Current opinion in biotechnology* 14: 248–254.
20. Stokes HW, Gillings MR (2011) *Gene flow, mobile genetic elements and the recruitment of antibiotic resistance genes into Gram-negative pathogens*. *FEMS Microbiology Reviews* 35: 790–819.
21. Top EM, Springael D (2003) *The role of mobile genetic elements in bacterial adaptation to xenobiotic organic compounds*. *Current Opinion in Biotechnology* 14: 262–269.
22. Auchtung JM, Lee CA, Monson RE, Lehman AP, Grossman AD (2005) *Regulation of a Bacillus subtilis mobile genetic element by intercellular signaling and the global DNA damage response*. *Proceedings of the National Academy of Sciences of the United States of America* 102: 12554–12559.
23. Lee CA, Auchtung JM, Monson RE, Grossman AD (2007) *Identification and characterization of int (integrase), xis (excisionase) and chromosomal attachment sites of the integrative and conjugative element ICEBs1 of Bacillus subtilis*. *Molecular microbiology* 66: 1356–1369.
24. Auchtung JM, Lee CA, Garrison KL, Grossman AD (2007) *Identification and characterization of the immunity repressor (ImmR) that controls the mobile genetic element ICEBs1 of Bacillus subtilis*. *Molecular microbiology* 64: 1515–1528.
25. Bose B, Grossman AD (2011) *Regulation of horizontal gene transfer in Bacillus subtilis by activation of a conserved site-specific protease*. *Journal of Bacteriology* 193: 22–29.
26. Bose B, Auchtung JM, Lee CA, Grossman AD (2008) *A conserved anti-repressor controls horizontal gene transfer by proteolysis*. *Molecular microbiology* 70: 570–582.
27. Lee CA, Grossman AD (2007) *Identification of the origin of transfer (oriT) and DNA relaxase required for conjugation of the integrative and conjugative element ICEBs1 of Bacillus subtilis*. *Journal of Bacteriology* 189: 7254–7261.

28. Lee CA, Babic A, Grossman AD (2010) *Autonomous plasmid-like replication of a conjugative transposon*. *Molecular microbiology* 75: 268–279.
29. Thomas J, Lee CA, Grossman AD (2013) *A conserved helicase processivity factor is needed for conjugation and replication of an integrative and conjugative element*. *PLoS Genetics* 9:1 e1003198.
30. Berkmen MB, Lee CA, Loveday EK, Grossman AD (2010) *Polar positioning of a conjugation protein from the integrative and conjugative element ICEBs1 of Bacillus subtilis*. *Journal of Bacteriology* 192: 38–45.
31. Babic A, Berkmen MB, Lee CA, Grossman AD (2011) *Efficient gene transfer in bacterial cell chains*. *MBio* 15:2(2).
32. Waldor MK, Tschäpe H, Mekalanos JJ (1996) *A new type of conjugative transposon encodes resistance to sulfamethoxazole, trimethoprim, and streptomycin in Vibrio cholerae O139*. *Journal of Bacteriology* 178: 4157–4165.
33. Burrus V, Marrero J, Waldor MK (2006) *The current ICE age: Biology and evolution of SXT-related integrating conjugative elements*. *Plasmid* 55: 173–183.
34. Bordeleau E, Brouillette E, Robichaud N, Burrus V (2010) *Beyond antibiotic resistance: integrating conjugative elements of the SXT/R391 family that encode novel diguanylate cyclases participate to c-di-GMP signalling in Vibrio cholerae*. *Environmental Microbiology* 12: 510–523.
35. Osorio CR, Marrero J, Wozniak RAF, Lemos ML, Burrus V, et al. (2008) *Genomic and functional analysis of ICEPdaSpa1, a fish-pathogen-derived SXT-related integrating conjugative element that can mobilize a virulence plasmid*. *Journal of Bacteriology* 190: 3353–3361.
36. Wozniak RAF, Fouts DE, Spagnoletti M, Colombo MM, Ceccarelli D, et al. (2009) *Comparative ICE genomics: insights into the evolution of the SXT/R391 family of ICEs*. *PLoS genetics* 5:12 e1000786.
37. Beaber JW, Burrus V, Hochhut B, Waldor MK (2002) *Comparison of SXT and R391, two conjugative integrating elements: definition of a genetic backbone for the mobilization of resistance determinants*. *Cellular and Molecular Life Sciences* 59: 2065–2070.
38. Böltner D, Osborn AM (2004) *Structural comparison of the integrative and conjugative elements R391, pMERPH, R997, and SXT*. *Plasmid* 51: 12–23.
39. Burrus V, Waldor MK (2003) *Control of SXT integration and excision*. *Journal of Bacteriology* 185: 5045–5054.
40. Hochhut B, Waldor MK (1999) *Site-specific integration of the conjugal Vibrio cholerae SXT element into prfC*. *Molecular microbiology* 32: 99–110.
41. Beaber JW, Hochhut B, Waldor MK (2004) *SOS response promotes horizontal dissemination of antibiotic resistance genes*. *Nature* 427: 72–74.

42. Beaber JW, Waldor MK (2004) *Identification of operators and promoters that control SXT conjugative transfer*. *Journal of Bacteriology* 186: 5945–5949.
43. O'Halloran JA, McGrath BM, Pembroke JT (2007) *The orf4 gene of the enterobacterial ICE, R391, encodes a novel UV-inducible recombination directionality factor, Jef, involved in excision and transfer of the ICE*. *FEMS Microbiology Letters* 272: 99–105.
44. Beaber JW, Hochhut B, Waldor MK (2002) *Genomic and functional analyses of SXT, an integrating antibiotic resistance gene transfer element derived from Vibrio cholerae*. *Journal of Bacteriology* 184: 4259–4269.
45. Ceccarelli D, Daccord A, Rene M, Burrus V (2008) *Identification of the origin of transfer (oriT) and a new gene required for mobilization of the SXT/R391 family of integrating conjugative elements*. *Journal of Bacteriology* 190: 5328–5338.
46. Daccord A, Ceccarelli D, Burrus V (2010) *Integrating conjugative elements of the SXT/R391 family trigger the excision and drive the mobilization of a new class of Vibrio genomic islands*. *Molecular Microbiology* 78: 576–588.
47. Wozniak RAF, Waldor MK (2009) *A toxin–antitoxin system promotes the maintenance of an integrative conjugative element*. *PLoS Genetics* 5: e1000439.
48. Dziewit L, Jazurek M, Drewniak L, Baj J, Bartosik D (2006) *The SXT conjugative element and linear prophage N15 encode toxin-antitoxin-stabilizing systems homologous to the tad-ata module of the Paracoccus aminophilus plasmid pAMI2*. *Journal of Bacteriology* 189: 1983–1997.
49. Marrero J, Waldor MK (2005) *Interactions between inner membrane proteins in donor and recipient cells limit conjugal DNA transfer*. *Developmental Cell* 8: 963–970.
50. Marrero J, Waldor MK (2007) *Determinants of entry exclusion within Eex and TraG are cytoplasmic*. *Journal of Bacteriology* 189: 6469–6473.
51. Marrero J, Waldor MK (2007) *The SXT/R391 Family of Integrative Conjugative Elements is composed of two exclusion groups*. *Journal of Bacteriology* 189: 3302–3305.
52. Burrus V, Waldor MK (2004) *Formation of SXT tandem arrays and SXT-R391 hybrids*. *Journal of Bacteriology* 186: 2636–2645.
53. Hochhut B, Beaber JW, Woodgate R, Waldor MK (2001) *Formation of chromosomal tandem arrays of the SXT element and R391, two conjugative chromosomally integrating elements that share an attachment site*. *Journal of Bacteriology* 183: 1124–1132.
54. Garriss G, Waldor MK, Burrus V (2009) *Mobile antibiotic resistance encoding elements promote their own diversity*. *PLoS Genetics* 5: e1000775.
55. Garriss G, Poulin-Laprade D, Burrus V (2013) *DNA-damaging agents induce*

- the RecA-independent homologous recombination functions of integrating conjugative elements of the SXT/R391 family.* Journal of Bacteriology 195: 1991–2003.
56. Daccord A, Mursell M, Poulin-Laprade D, Burrus V (2012) *Dynamics of the SetCD-regulated integration and excision of genomic islands mobilized by integrating conjugative elements of the SXT/R391 family.* Journal of Bacteriology 194: 5794–5802.
57. Daccord A, Ceccarelli D, Rodrigue S, Burrus V (2012) *Comparative analysis of mobilizable genomic islands.* Journal of Bacteriology 195: 606–614.
58. Hochhut B, Marrero J, Waldor MK (2000) *Mobilization of plasmids and chromosomal DNA mediated by the SXT element, a *constin* found in *Vibrio cholerae* O139.* Journal of Bacteriology 182: 2043–2047.
59. Shoemaker NB, Barber RD, Salyers AA (1989) *Cloning and characterization of a *Bacteroides conjugal tetracycline-erythromycin* resistance element by using a shuttle cosmid vector.* Journal of Bacteriology 171: 1294–1302.
60. Shoemaker NB, Vlamakis H, Hayes K, Salyers AA (2001) *Evidence for extensive resistance gene transfer among *Bacteroides* spp. and among *Bacteroides* and other genera in the human colon.* Applied and Environmental Microbiology 67: 561–568.
61. Jetters RT, Wang G-R, Moon K, Shoemaker NB, Salyers AA (2009) *Tetracycline-associated transcriptional regulation of transfer genes of the *Bacteroides conjugative transposon CTnDOT*.* Journal of Bacteriology 191: 6374–6382.
62. Waters JL, Salyers AA (2012) *The small RNA *RteR* inhibits transfer of the *Bacteroides conjugative transposon CTnDOT*.* Journal of Bacteriology 194: 5228–5236.
63. Stevens AM, Shoemaker NB, Li L-Y, Salyers AA (1993) *Tetracycline regulation of genes on *Bacteroides conjugative transposons*.* Journal of Bacteriology 175: 6134–6141.
64. Wang Y, Shoemaker NB, Salyers AA (2004) *Regulation of a *Bacteroides operon* that controls excision and transfer of the conjugative transposon *CTnDOT*.* Journal of Bacteriology 186: 2548–2557.
65. Wang Y, Rotman ER, Shoemaker NB, Salyers AA (2005) *Translational control of tetracycline resistance and conjugation in the *Bacteroides conjugative transposon CTnDOT*.* Journal of Bacteriology 187: 2673–2680.
66. Moon K, Shoemaker NB, Gardner JF, Salyers AA (2005) *Regulation of excision genes of the *Bacteroides conjugative transposon CTnDOT*.* Journal of Bacteriology 187: 5732–5741.
67. Park J, Salyers AA (2010) *Characterization of the *Bacteroides CTnDOT* regulatory protein *RteC*.* Journal of Bacteriology 193: 91–97.

68. Whittle G, Shoemaker NB, Salyers AA (2002) *Characterization of genes involved in modulation of conjugal transfer of the Bacteroides conjugative transposon CTnDOT*. Journal of Bacteriology 184: 3839–3847.
69. Keeton CM, Park J, Wang G-R, Hopp CM, Shoemaker NB, et al. (2013) *The excision proteins of CTnDOT positively regulate the transfer operon*. Plasmid 69: 172–179.
70. Cheng Q, Paszkiet BJ, Shoemaker NB, Gardner JF, Salyers AA (2000) *Integration and excision of a Bacteroides conjugative transposon, CTnDOT*. Journal of Bacteriology 182: 4035–4043.
71. Cheng Q, Sutanto Y, Shoemaker NB, Gardner JF, Salyers AA (2001) *Identification of genes required for excision of CTnDOT, a Bacteroides conjugative transposon*. Molecular microbiology 41: 625–632.
72. Keeton CM, Gardner JF (2012) *Roles of Exc protein and DNA homology in the CTnDOT excision reaction*. Journal of Bacteriology 194: 3368–3376.
73. Keeton CM, Hopp CM, Yoneji S, Gardner JF (2013) *Interactions of the excision proteins of CTnDOT in the attR intasome*. Plasmid.
74. Sutanto Y, Shoemaker NB, Gardner JF, Salyers AA (2002) *Characterization of Exc, a novel protein required for the excision of Bacteroides conjugative transposon*. Molecular microbiology 46: 1239–1246.
75. Sutanto Y, DiChiara JM, Shoemaker NB, Gardner JF, Salyers AA (2004) *Factors required in vitro for excision of the Bacteroides conjugative transposon, CTnDOT*. Plasmid 52: 119–130.
76. Cheng Q, Wesslund N, Shoemaker NB, Salyers AA, Gardner JF (2002) *Development of an in vitro integration assay for the Bacteroides conjugative transposon CTnDOT*. Journal of Bacteriology 184: 4829–4837.
77. DiChiara JM, Salyers AA, Gardner JF (2005) *In vitro analysis of sequence requirements for the excision reaction of the Bacteroides conjugative transposon, CTnDOT*. Molecular Microbiology 56: 1035–1048.
78. DiChiara JM, Mattis AN, Gardner JF (2007) *IntDOT interactions with core- and arm-type sites of the conjugative transposon CTnDOT*. Journal of Bacteriology 189: 2692–2701.
79. Bonheyo G, Graham D, Shoemaker NB, Salyers AA (2001) *Transfer region of a Bacteroides conjugative transposon, CTnDOT*. Plasmid 45: 41–51.
80. Bonheyo GT, Hund BD, Shoemaker NB, Salyers AA (2001) *Transfer region of a Bacteroides conjugative transposon contains regulatory as well as structural genes*. Plasmid 46: 202–209.
81. Peed L, Parker AC, Smith CJ (2010) *Genetic and functional analyses of the mob operon on conjugative transposon CTn341 from Bacteroides spp*. Journal of Bacteriology 192: 4643–4650.

82. Laprise J, Yoneji S, Gardner JF (2013) *IntDOT interactions with core sites during integrative recombination*. *Journal of Bacteriology* 195: 1883–1891.
83. Malanowska K, Salyers AA, Gardner JF (2006) *Characterization of a conjugative transposon integrase, IntDOT*. *Molecular Microbiology* 60: 1228–1240.
84. Malanowska K, Yoneji S, Salyers AA, Gardner JF (2007) *CTnDOT integrase performs ordered homology-dependent and homology-independent strand exchanges*. *Nucleic Acids Research* 35: 5861–5873.
85. Wood MM, DiChiara JM, Yoneji S, Gardner JF (2010) *CTnDOT integrase interactions with attachment site DNA and control of directionality of the recombination reaction*. *Journal of Bacteriology* 192: 3934–3943.
86. Malanowska K, Cioni J, Swalla BM, Salyers A, Gardner JF (2009) *Mutational analysis and homology-based modeling of the IntDOT core-binding domain*. *Journal of Bacteriology* 191: 2330–2339.
87. Shoemaker NB, Wang G-R, Stevens AM, Salyers AA (1993) *Excision, transfer, and integration of NBU1, a mobilizable site-selective insertion element*. *Journal of Bacteriology* 175: 6578–6587.
88. Shoemaker NB, Wang G-R, Salyers AA (1996) *The Bacteroides mobilizable insertion element, NBU1, integrates into the 3' end of a Leu-tRNA gene and has an integrase that is a member of the lambda integrase family*. *Journal of Bacteriology* 178: 3594–3600.
89. Shoemaker NB, Wang G-R, Salyers AA (1996) *NBU1, a mobilizable site-specific integrated element from Bacteroides spp., can integrate nonspecifically in Escherichia coli*. *Journal of Bacteriology* 178: 3601–3607.
90. Moon K, Sonnenburg J, Salyers AA (2007) *Unexpected effect of a Bacteroides conjugative transposon, CTnDOT, on chromosomal gene expression in its bacterial host*. *Molecular Microbiology* 64: 1562–1571.
91. Wang G-R, Shoemaker NB, Jeters RT, Salyers AA (2011) *CTn12256, a chimeric Bacteroides conjugative transposon that consists of two independently active mobile elements*. *Plasmid* 66: 93–105.
92. Franke AE, Clewell DB (1981) *Evidence for a chromosome-borne resistance transposon (Tn916) in Streptococcus faecalis that is capable of "conjugal" transfer in the absence of a conjugative plasmid*. *Journal of Bacteriology* 145: 494–502.
93. Poyart C, Celli J, Trieu-Cuot P (1995) *Conjugative transposition of Tn916-related elements from Enterococcus faecalis to Escherichia coli and Pseudomonas fluorescens*. *Antimicrobial Agents and Chemotherapy* 39: 500–506.
94. Roberts AP, Mullany P (2009) *A modular master on the move: the Tn916 family of mobile genetic elements*. *Trends in Microbiology* 17: 251–258.

95. Celli J, Trieu-Cuot P (1998) *Circularization of Tn916 is required for expression of the transposon-encoded transfer functions: characterization of long tetracycline-inducible transcripts reading through the attachment site*. *Molecular microbiology* 28: 103–117.
96. Su YA, He P, Clewell DB (1992) *Characterization of the tet(M) determinant of Tn916: evidence for regulation by transcription attenuation*. *Antimicrobial Agents and Chemotherapy* 36: 769–778.
97. Celli J, Poyart C, Trieu-Cuot P (1997) *Use of an excision reporter plasmid to study the intracellular mobility of the conjugative transposon Tn916 in gram-positive bacteria*. *Microbiology* 143: 1253–1261.
98. Bringel F, Van Alstine GL, Scott JR (1992) *Conjugative transposition of Tn916: the transposon int gene is required only in the donor*. *Journal of Bacteriology* 174: 4036–4041.
99. Scott JR, Kirchman PA, Caparon MG (1988) *An intermediate in transposition of the conjugative transposon Tn916*. *Proceedings of the National Academy of Sciences* 85: 4809–4813.
100. Storrs MJ, Poyart-Salmeron C, Trieu-Cuot P, Courvalin P (1991) *Conjugative transposition of Tn916 requires the excisive and integrative activities of the transposon-encoded integrase*. *Journal of Bacteriology* 173: 4347–4352.
101. Hinerfeld D, Churchward G (2001) *Xis protein of the conjugative transposon Tn916 plays dual opposing roles in transposon excision*. *Molecular microbiology* 41: 1459–1467.
102. Marra D, Scott JR (1999) *Regulation of excision of the conjugative transposon Tn916*. *Molecular microbiology* 31: 609–621.
103. Rudy C, Taylor KL, Hinerfeld D, Scott JR, Churchward G (1997) *Excision of a conjugative transposon in vitro by the Int and Xis proteins of Tn916*. *Nucleic acids research* 25: 4061–4066.
104. Connolly KM, Iwahara M, Clubb RT (2002) *Xis protein binding to the left arm stimulates excision of conjugative transposon Tn916*. *Journal of Bacteriology* 184: 2088–2099.
105. Jia Y, Churchward G (1999) *Interactions of the integrase protein of the conjugative transposon Tn916 with its specific DNA binding sites*. *Journal of Bacteriology* 181: 6114–6123.
106. Lu F, Churchward G (1994) *Conjugative transposition: Tn916 integrase contains two independent DNA binding domains that recognize different DNA sequences*. *The EMBO journal* 13: 1541.
107. Rudy CK, Scott JR, Churchward G (1997) *DNA binding by the Xis protein of the conjugative transposon Tn916*. *Journal of Bacteriology* 179: 2567–2572.
108. Caparon MG, Scott JR (1989) *Excision and insertion of the conjugative*

- transposon Tn916 involves a novel recombination mechanism.* Cell 59: 1027–1034.
109. Jaworski DD, Clewell DB (1994) *Evidence that coupling sequences play a frequency-determining role in conjugative transposition of Tn916 in Enterococcus faecalis.* Journal of Bacteriology 176: 3328–3335.
 110. Lu F, Churchward G (1995) *Tn916 target DNA sequences bind the C-terminal domain of integrase protein with different affinities that correlate with transposon insertion frequency.* Journal of Bacteriology 177: 1938–1946.
 111. Taylor KL, Churchward G (1997) *Specific DNA cleavage mediated by the integrase of conjugative transposon Tn916.* Journal of Bacteriology 179: 1117–1125.
 112. Smidt H, Song D, van der Oost J, de Vos WM (1999) *Random transposition by Tn916 in Desulfitobacterium dehalogenans allows for isolation and characterization of halorespiration-deficient mutants.* Journal of Bacteriology 181: 6882–6888.
 113. Marra D, Pethel B, Churchward GG, Scott JR (1999) *The frequency of conjugative transposition of Tn916 is not determined by the frequency of excision.* Journal of Bacteriology 181: 5414–5418.
 114. Scott JR, Bringel F, Marra D, Van Alstine GL, Rudy CK (1994) *Conjugative transposition of Tn916: preferred targets and evidence for conjugative transfer of a single strand and for a double-stranded circular intermediate.* Molecular Microbiology 11: 1099–1108.
 115. Senghas E, Jones JM, Yamamoto M, Gawron-Burke C, Clewell DB (1988) *Genetic organization of the bacterial conjugative transposon Tn916.* Journal of Bacteriology 170: 245–249.
 116. Hinerfeld D, Churchward G (2001) *Specific binding of integrase to the origin of transfer (oriT) of the conjugative transposon Tn916.* Journal of Bacteriology 183: 2947–2951.
 117. Jaworski DD, Clewell DB (1995) *A functional origin of transfer (oriT) on the conjugative transposon Tn916.* Journal of Bacteriology 177: 6644–6651.
 118. Rocco JM, Churchward G (2006) *The integrase of the conjugative transposon Tn916 directs strand- and sequence-specific cleavage of the origin of conjugal transfer, oriT, by the endonuclease Orf20.* Journal of Bacteriology 188: 2207–2213.
 119. Serfiotis-Mitsa D, Roberts GA, Cooper LP, White JH, Nutley M, et al. (2008) *The orf18 gene product from conjugative transposon Tn916 is an ArdA antirestriction protein that inhibits type I DNA restriction–modification systems.* Journal of Molecular Biology 383: 970–981.
 120. Cookson AL, Noel S, Hussein H, Perry R, Sang C, et al. (2011) *Transposition of Tn916 in the four replicons of the Butyrivibrio proteoclasticus B316T genome.*

- FEMS Microbiology Letters 316: 144–151.
121. Mullany P, Williams R, Langridge GC, Turner DJ, Whalan R, et al. (2012) *Behavior and target site selection of conjugative transposon Tn916 in two different strains of toxigenic Clostridium difficile*. Applied and Environmental Microbiology 78: 2147–2153.
 122. Roberts AP, Johanesen PA, Lyras D, Mullany P, Rood JI (2001) *Comparison of Tn5397 from Clostridium difficile, Tn916 from Enterococcus faecalis and the CW459tet (M) element from Clostridium perfringens shows that they have similar conjugation regions but different insertion and excision modules*. Microbiology 147: 1243–1251.
 123. Rudy CK, Scott JR (1994) *Length of the coupling sequence of Tn916*. Journal of Bacteriology 176: 3386–3388.
 124. Sezonov G, Possoz C, Friedmann A, Pernodet J-L, Guerineau M (2000) *KorSA from the Streptomyces integrative element pSAM2 is a central transcriptional repressor: target genes and binding sites*. Journal of Bacteriology 182: 1243–1250.
 125. Sezonov G, Hagège J, Friedmann A, Guerineau M (1995) *Characterization of pra, a gene for replication control in pSAM2, the integrative element of Streptomyces ambofaciens*. Molecular Microbiology 17: 533–544.
 126. Sezonov G, Duchêne A-M, Friedmann A, Guérineau M, Pernodet J-L (1998) *Replicase, excisionase, and integrase genes of the Streptomyces element pSAM2 constitute an operon positively regulated by the pra gene*. Journal of Bacteriology 180: 3056–3061.
 127. Boccard F, Smokvina T, Pernodet JL, Friedmann A, Guerineau M (1989) *The integrated conjugative plasmid pSAM2 of Streptomyces ambofaciens is related to temperate bacteriophages*. The EMBO journal 8: 973.
 128. Boccard F, Smokvina T, Pernodet J, Friedmann A, Guerineau M (1989) *Structural analysis of loci involved in pSAM2 site-specific integration in Streptomyces*. Plasmid 21: 59–70.
 129. Kuhstoss S, Richardson MA, Rao RN (1989) *Site-specific integration in Streptomyces ambofaciens: localization of integration functions in S. ambofaciens plasmid pSAM2*. Journal of Bacteriology 171: 16–23.
 130. Mazodier P, Thompson C, Boccard F (1990) *The chromosomal integration site of the Streptomyces element pSAM2 overlaps a putative tRNA gene conserved among actinomycetes*. Molecular & General Genetics 222: 431–434.
 131. Raynal A, Tuphile K, Gerbaud C, Luther T, Guérineau M, et al. (1998) *Structure of the chromosomal insertion site for pSAM2: functional analysis in Escherichia coli*. Molecular microbiology 28: 333–342.
 132. Raynal A, Friedmann A, Tuphile K, Guerineau M, Pernodet J-L (2002) *Characterization of the attP site of the integrative element pSAM2 from*

- Streptomyces ambofaciens*. Microbiology 148: 61–67.
133. Hagège J, Pernodet J, Friedmann A, Guérineau M (1993) *Mode and origin of replication of pSAM2, a conjugative integrating element of Streptomyces ambofaciens*. Molecular Microbiology 10: 799–812.
 134. Hagège J, Bocard F, Smokvina T, Pernodet J, Friedmann A, et al. (1994) *Identification of a gene encoding the replication initiator protein of the Streptomyces integrating element, pSAM2*. Plasmid 31: 166–183.
 135. Smokvina T, Bocard F, Pernodet J, Friedmann A, Guérineau M (1991) *Functional analysis of the Streptomyces ambofaciens element pSAM2*. Plasmid 25: 40–52.
 136. Hagège J, Pernodet JL, Sezonov G, Gerbaud C, Friedmann A, et al. (1993) *Transfer functions of the conjugative integrating element pSAM2 from Streptomyces ambofaciens: characterization of a kil-kor system associated with transfer*. Journal of Bacteriology 175: 5529–5538.
 137. Possoz C, Ribard C, Gagnat J, Pernodet J-L, Guérineau M (2001) *The integrative element pSAM2 from Streptomyces: kinetics and mode of conjugal transfer*. Molecular microbiology 42: 159–166.
 138. Possoz C, Gagnat J, Sezonov G, Guérineau M, Pernodet J (2003) *Conjugal immunity of Streptomyces strains carrying the integrative element pSAM2 is due to the pif gene (pSAM2 immunity factor)*. Molecular microbiology 47: 1385–1393.
 139. Martin C, Mazodier P, Mediola M, Gicquel B, Smokvina T, et al. (1991) *Site-specific integration of the Streptomyces plasmid pSAM2 in Mycobacterium smegmatis*. Molecular Microbiology 5: 2499–2502.
 140. Pernodet J, Simonet J, Guérineau M (1984) *Plasmids in different strains of Streptomyces ambofaciens: free and integrated form of plasmid pSAM2*. Molecular & General Genetics 198: 35–41.
 141. Sullivan JT, Trzebiatowski JR, Cruickshank RW, Gouzy J, Brown SD, et al. (2002) *Comparative sequence analysis of the symbiosis island of Mesorhizobium loti strain R7A*. Journal of Bacteriology 184: 3086–3095.
 142. Sullivan JT, Ronson CW (1998) *Evolution of rhizobia by acquisition of a 500-kb symbiosis island that integrates into a phe-tRNA gene*. Proceedings of the National Academy of Sciences 95: 5145–5149.
 143. Sullivan JT, Brown SD, Ronson CW (2013) *The NifA-RpoN regulon of Mesorhizobium loti strain R7A and its symbiotic activation by a novel LacI/GalR-family regulator*. PLoS ONE 8: e53762.
 144. Sullivan JT, Patrick HN, Lowther WL, Scott DB, Ronson CW (1995) *Nodulating strains of Rhizobium loti arise through chromosomal symbiotic gene transfer in the environment*. Proceedings of the National Academy of Sciences 92: 8985–8989.

145. Ramsay JP, Major AS, Komarovskiy VM, Sullivan JT, Dy RL, et al. (2013) *A widely conserved molecular switch controls quorum sensing and symbiosis island transfer in Mesorhizobium loti through expression of a novel antiactivator*. *Molecular Microbiology* 87: 1–13.
146. Ramsay JP, Sullivan JT, Stuart GS, Lamont IL, Ronson CW (2006) *Excision and transfer of the Mesorhizobium loti R7A symbiosis island requires an integrase IntS, a novel recombination directionality factor RdfS, and a putative relaxase RlxS*. *Molecular Microbiology* 62: 723–734.
147. Ramsay JP, Sullivan JT, Jambari N, Ortori CA, Heeb S, et al. (2009) *A LuxRI-family regulatory system controls excision and transfer of the Mesorhizobium loti strain R7A symbiosis island by activating expression of two conserved hypothetical genes*. *Molecular Microbiology* 73: 1141–1155.
148. Burrus V, Roussel Y, Decaris B, Guédon G (2000) *Characterization of a novel integrative element, ICESt1, in the lactic acid bacterium Streptococcus thermophilus*. *Applied and Environmental Microbiology* 66: 1749–1753.
149. Burrus V, Pavlovic G, Decaris B, Guédon G (2002) *The ICESt1 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.
150. Pavlovic G (2004) *Evolution of genomic islands by deletion and tandem accretion by site-specific recombination: ICESt1-related elements from Streptococcus thermophilus*. *Microbiology* 150: 759–774.
151. Bellanger X, Morel C, Decaris B, Guédon G (2006) *Derepression of excision of integrative and potentially conjugative elements from Streptococcus thermophilus by DNA damage response: implication of a cl-related repressor*. *Journal of Bacteriology* 189: 1478–1481.
152. Bellanger X, Morel C, Decaris B, Guédon G (2008) *Regulation of excision of integrative and potentially conjugative elements from Streptococcus thermophilus: role of the arp1 repressor*. *Journal of Molecular Microbiology and Biotechnology* 14: 16–21.
153. Bellanger X, Roberts AP, Morel C, Choulet F, Pavlovic G, et al. (2009) *Conjugative transfer of the integrative conjugative elements ICESt1 and ICESt3 from Streptococcus thermophilus*. *Journal of Bacteriology* 191: 2764–2775.
154. Carraro N, Libante V, Morel C, Decaris B, Charron-Bourgoin F, et al. (2011) *Differential regulation of two closely related integrative and conjugative elements from Streptococcus thermophilus*. *BMC microbiology* 11: 238.
155. Buchrieser C, Prentice M, Carniel E (1998) *The 102-kilobase unstable region of Yersinia pestis comprises a high-pathogenicity island linked to a pigmentation segment which undergoes internal rearrangement*. *Journal of Bacteriology* 180: 2321–2329.
156. Schubert S (2004) *The high-pathogenicity island (HPI): evolutionary and*

- functional aspects*. International Journal of Medical Microbiology 294: 83–94.
157. Carniel E (2001) *The Yersinia high-pathogenicity island: an iron-uptake island*. Microbes and infection 3: 561–569.
 158. Benedek O, Schubert S (2007) *Mobility of the Yersinia high-pathogenicity island (HPI): Transfer mechanisms of pathogenicity islands (PAIs) revisited*. Acta Microbiologica et Immunologica Hungarica 54: 89–105.
 159. Soler Bistué AJC, Birshan D, Tomaras AP, Dandekar M, Tran T, et al. (2008) *Klebsiella pneumoniae multiresistance plasmid pMET1: similarity with the Yersinia pestis plasmid pCRY and integrative conjugative elements*. PLoS ONE 3: e1800.
 160. Szwagierczak A, Antonenka U, Popowicz GM, Sitar T, Holak TA, et al. (2009) *Structures of the arm-type binding domains of HPI and HAI7 integrases*. Journal of Biological Chemistry 284: 31664–31671.
 161. Buchrieser C, Brosch R, Bach S, Guiyoule A, Carniel E (1998) *The high-pathogenicity island of Yersinia pseudotuberculosis can be inserted into any of the three chromosomal asn tRNA genes*. Molecular microbiology 30: 965–978.
 162. Lesic B, Bach S, Ghigo J-M, Dobrindt U, Hacker J, et al. (2004) *Excision of the high-pathogenicity island of Yersinia pseudotuberculosis requires the combined actions of its cognate integrase and Hef, a new recombination directionality factor*. Molecular Microbiology 52: 1337–1348.
 163. Lesic B, Carniel E (2005) *Horizontal transfer of the high-pathogenicity island of Yersinia pseudotuberculosis*. Journal of Bacteriology 187: 3352–3358.
 164. Rakin A, Noelting C, Schropp P, Heesemann J (2001) *Integrative module of the high-pathogenicity island of Yersinia*. Molecular microbiology 39: 407–416.
 165. Antonenka U, Nölting C, Heesemann J, Rakin A (2006) *Independent acquisition of site-specific recombination factors by asn tRNA gene-targeting genomic islands*. International Journal of Medical Microbiology 296: 341–352.
 166. Dorn E, Hellwig M, Reineke W, Knackmuss H (1974) *Isolation and characterization of a 3-chlorobenzoate degrading pseudomonad*. Arch Microbiol 99: 61–70.
 167. Ravatn R, Studer S, Springael D, Zehnder AJB, van der Meer JR (1998) *Chromosomal integration, tandem amplification, and deamplification in Pseudomonas putida F1 of a 105-kilobase genetic element containing the chlorocatechol degradative genes from Pseudomonas sp. strain B13*. Journal of Bacteriology 180: 4360–4369.
 168. Gaillard M, Vallaëys T, Vorhölter FJ, Minoia M, Werlen C, et al. (2006) *The clc element of Pseudomonas sp. strain B13, a genomic island with various catabolic properties*. Journal of Bacteriology 188: 1999–2013.
 169. Ravatn R, Zehnder AJB, van der Meer JR (1998) *Low-frequency horizontal*

- transfer of an element containing the chlorocatechol degradation genes from Pseudomonas sp. strain B13 to Pseudomonas putida F1 and to indigenous bacteria in laboratory-scale activated-sludge microcosms. Applied and environmental microbiology 64: 2126–2132.*
170. Gaillard M, Pernet N, Vogne C, Hagenbüchle O, van der Meer JR (2008) *Host and invader impact of transfer of the clc genomic island into Pseudomonas aeruginosa PAO1.* Proceedings of the National Academy of Sciences 105: 7058–7063.
 171. Gao J, Ellis LBM, Wackett LP (2009) *The University of Minnesota biocatalysis/biodegradation database: improving public access.* Nucleic Acids Research 38: D488–D491.
 172. Springael D, Peys K, Ryngaert A, Roy SV, Hooyberghs L, et al. (2002) *Community shifts in a seeded 3-chlorobenzoate degrading membrane biofilm reactor: indications for involvement of in situ horizontal transfer of the clc-element from inoculum to contaminant bacteria.* Environmental microbiology 4: 70–80.
 173. Müller TA, Werlen C, Spain J, Van Der Meer JR (2003) *Evolution of a chlorobenzene degradative pathway among bacteria in a contaminated groundwater mediated by a genomic island in Ralstonia.* Environmental microbiology 5: 163–173.
 174. Chain PS, Deneff VJ, Konstantinidis KT, Vergez LM, Agulló L, et al. (2006) *Burkholderia xenovorans LB400 harbors a multi-replicon, 9.73-Mbp genome shaped for versatility.* Proceedings of the National Academy of Sciences 103: 15280–15287.
 175. Lechner M, Schmitt K, Bauer S, Hot D, Hubans C, et al. (2009) *Genomic island excisions in Bordetella petrii.* BMC Microbiology 9: 141.
 176. Gross R, Guzman CA, Sebahia M, Martins dos Santos VA, Pieper DH, et al. (2008) *The missing link: Bordetella petrii is endowed with both the metabolic versatility of environmental bacteria and virulence traits of pathogenic Bordetellae.* BMC Genomics 9: 449.
 177. Van Houdt R, Monchy S, Leys N, Mergeay M (2009) *New mobile genetic elements in Cupriavidus metallidurans CH34, their possible roles and occurrence in other bacteria.* Antonie van Leeuwenhoek 96: 205–226.
 178. Janssen PJ, Van Houdt R, Moors H, Monsieurs P, Morin N, et al. (2010) *The complete genome sequence of Cupriavidus metallidurans strain CH34, a master survivalist in harsh and anthropogenic environments.* PLoS ONE 5: e10433.
 179. Van Houdt R, Monsieurs P, Mijndonckx K, Provoost A, Janssen A, et al. (2012) *Variation in genomic islands contribute to genome plasticity in cupriavidus metallidurans.* BMC genomics 13: 111.
 180. Klockgether J, Wurdemann D, Wiehlmann L, Tumbler B (2008) *Transcript*

- profiling of the Pseudomonas aeruginosa genomic islands PAGI-2 and pKLC102*. Microbiology 154: 1599–1604.
181. Klockgether J, Wurdemann D, Reva O, Wiehlmann L, Tummler B (2006) *Diversity of the abundant pKLC102/PAGI-2 family of genomic islands in Pseudomonas aeruginosa*. Journal of Bacteriology 189: 2443–2459.
182. Würdemann D, Tümmler B (2007) *In silico comparison of pKLC102-like genomic islands of Pseudomonas aeruginosa*. FEMS Microbiology Letters 275: 244–249.
183. Larbig KD, Christmann A, Johann A, Klockgether J, Hartsch T, et al. (2002) *Gene islands integrated into tRNAGly genes confer genome diversity on a Pseudomonas aeruginosa clone*. Journal of Bacteriology 184: 6665–6680.
184. He J, Baldini RL, Déziel E, Saucier M, Zhang Q, et al. (2004) *The broad host range pathogen Pseudomonas aeruginosa strain PA14 carries two pathogenicity islands harboring plant and animal virulence genes*. Proceedings of the National Academy of Sciences of the United States of America 101: 2530–2535.
185. Harrison EM, Carter MEK, Luck S, Ou H-Y, He X, et al. (2010) *Pathogenicity islands PAPI-1 and PAPI-2 contribute individually and synergistically to the virulence of Pseudomonas aeruginosa strain PA14*. Infection and Immunity 78: 1437–1446.
186. Qiu X, Gurkar AU, Lory S (2006) *Interstrain transfer of the large pathogenicity island (PAPI-1) of Pseudomonas aeruginosa*. Proceedings of the National Academy of Sciences 103: 19830–19835.
187. Klockgether J, Reva O, Larbig K, Tummler B (2003) *Sequence analysis of the mobile genome island pKLC102 of Pseudomonas aeruginosa C*. Journal of Bacteriology 186: 518–534.
188. Muller D, Médigue C, Koechler S, Barbe V, Barakat M, et al. (2007) *A tale of two oxidation states: bacterial colonization of arsenic-rich environments*. PLoS Genetics 3: e53.
189. Mohd-Zain Z, Turner SL, Cerdeno-Tarraga AM, Lilley AK, Inzana TJ, et al. (2004) *Transferable antibiotic resistance elements in Haemophilus influenzae share a common evolutionary origin with a diverse family of syntenic genomic islands*. Journal of Bacteriology 186: 8114–8122.
190. Juhas M, Power PM, Harding RM, Ferguson DJ, Dimopoulou ID, et al. (2007) *Sequence and functional analyses of Haemophilus spp. genomic islands*. Genome Biol 8: R237.
191. Ravatn R, Studer S, Zehnder AJB, Roelof van der Meer J (1998) *Int-B13, an unusual site-specific recombinase of the bacteriophage P4 integrase family, is responsible for chromosomal insertion of the 105-kilobase clc element of Pseudomonas sp. strain B13*. Journal of bacteriology 180: 5505–5514.

192. Sentchilo V, Czechowska K, Pradervand N, Minoia M, Miyazaki R, et al. (2009) *Intracellular excision and reintegration dynamics of the ICE_{clc} genomic island of Pseudomonas knackmussii sp. strain B13*. *Molecular microbiology* 72: 1293–1306.
193. Sentchilo V, Ravatn R, Werlen C, Zehnder AJB, Van Der Meer JR (2003) *Unusual integrase gene expression on the clc genomic island in Pseudomonas sp. strain B13*. *Journal of Bacteriology* 185: 4530–4538.
194. Sentchilo V, Zehnder AJB, Van Der Meer JR (2003) *Characterization of two alternative promoters for integrase expression in the clc genomic island of Pseudomonas sp. strain B13*. *Molecular Microbiology* 49: 93–104.
195. Reinhard F, Miyazaki R, Pradervand N, van der Meer JR (2013) *Cell differentiation to “mating bodies” induced by an integrating and conjugative element in free-living bacteria*. *Current Biology* 23: 255–259.
196. Minoia M, Gaillard M, Reinhard F, Stojanov M, Sentchilo V, et al. (2008) *Stochasticity and bistability in horizontal transfer control of a genomic island in Pseudomonas*. *Proceedings of the National Academy of Sciences* 105: 20792–20797.
197. Miyazaki R, Minoia M, Pradervand N, Sulser S, Reinhard F, et al. (2012) *Cellular variability of RpoS expression underlies subpopulation activation of an integrative and conjugative element*. *PLoS Genetics* 8: e1002818.
198. Miyazaki R, van der Meer JR (2011) *A dual functional origin of transfer in the ICE_{clc} genomic island of Pseudomonas knackmussii B13*. *Molecular Microbiology* 79: 743–758.
199. Roberts AP, Mullany P (2011) *Tn916-like genetic elements: a diverse group of modular mobile elements conferring antibiotic resistance*. *FEMS Microbiology Reviews* 35: 856–871.
200. Poele EM, Bolhuis H, Dijkhuizen L (2008) *Actinomycete integrative and conjugative elements*. *Antonie van Leeuwenhoek* 94: 127–143.
201. Gaillard M, Pradervand N, Minoia M, Sentchilo V, Johnson DR, et al. (2010) *Transcriptome analysis of the mobile genome ICE_{clc} in Pseudomonas knackmussii B13*. *BMC microbiology* 10: 153.

CHAPTER 2

TRANSCRIPTOME ANALYSIS OF THE MOBILE GENOME *ICECLC* IN *PSEUDOMONAS KNACKMUSSII* B13

Published previously in BMC Microbiology 2010 May 26;10:153. (doi: 10.1186/1471-2180-10-153), by Muriel Gaillard[°], Nicolas Pradervand[°], Marco Minoia, Vladimir Sentchilo, David R. Johnson⁺ and Jan Roelof van der Meer.

[°]These authors contributed equally and should therefore be considered first author.

ABSTRACT

Background

Integrative and conjugative elements (ICE) form a diverse group of DNA elements that are integrated in the chromosome of the bacterial host, but can occasionally excise and horizontally transfer to a new host cell. ICE come in different families, typically with a conserved core for functions controlling the element's behavior and a variable region providing auxiliary functions to the host. The ICE*c/c* element of *Pseudomonas knackmussii* strain B13 is representative for a large family of chromosomal islands detected by genome sequencing approaches. It provides the host with the capacity to degrade chloroaromatics and 2-aminophenol.

Results

Here we study the transcriptional organization of the ICE*c/c* core region. By northern hybridizations, reverse-transcriptase polymerase chain reaction (RT-PCR) and Rapid Amplification of cDNA Ends (5'-RACE) fifteen transcripts were mapped in the core region. The occurrence and location of those transcripts were further confirmed by hybridizing labeled cDNA to a semi-tiling micro-array probing both strands of the ICE*c/c* core region. Dot blot and semi-tiling array hybridizations demonstrated most of the core transcripts to be upregulated during stationary phase on 3-chlorobenzoate, but not on succinate or glucose.

Conclusions

The transcription analysis of the ICE*c/c* core region provides detailed insights in the mode of regulatory organization and will help to further understand the complex mode of behavior of this class of mobile elements. We conclude that ICE*c/c* core transcription is concerted at a global level, more reminiscent of a phage program than of plasmid conjugation.

INTRODUCTION

The acquisition of horizontally transferred genes plays an important role in prokaryotic evolution [1]. The colonization of new ecological niches is often enabled by the acquisition of foreign genes, which can be transmitted by a large variety of mobile genetic elements (MGE) present in individual members of the microbial community. In terms of evolutionary success, it is thus interesting to understand how different mobile DNA elements control their mobility and may adapt to their bacterial host [2].

Various classes of MGE are known, the most well-studied of which are plasmids and bacteriophages [3, 4]. Plasmids, apart from certain exceptions such as the F-episome in *Escherichia coli*, generally occur as extrachromosomal DNA in the bacterial cell. An important aspect of their life-style, therefore, is to ensure replication, stability and maintenance in the host cell [5], and a variety of control mechanisms have evolved hereto [6]. Conjugative plasmids encode and orchestrate specific machineries to produce the transfer system dedicated to their own distribution (e.g., type IV secretion system) [7]. By contrast, temperate bacteriophages insert into the host's chromosome, where they can remain silent and are co-replicated with the host's DNA for many generations, or are eventually genetically defunctionalized. Feedback regulatory systems silence phage behaviour in the temperate form, but can very rapidly induce the lytic phase (e.g., upon SOS response), upon which thousands of phage particles are produced to commence a new infection cycle [8] [9].

More recently, a large new class of DNA elements has been recognized that contributes importantly to bacterial genome evolution via horizontal gene transfer. Most of these have been detected by comparative genome sequencing and have in general been named 'genomic islands' (GEI) to portray their foreign character within the host genome [10]. Often, according to the functions encoded by the GEI, they were classified as pathogenicity, symbiosis, metabolic, secretion or resistance islands [11 , 12]. Recently we and others proposed that GEI should be considered an overarching group of elements comprising both phage-like, Integrative and Conjugative Elements (ICE), as well as conjugative transposons [10]. Although GEI are assumed to have been acquired via horizontal gene transfer, for most of them self-transfer has not been tested under experimental conditions. In some cases only GEI excision from its chromosomal location has been observed, which is presumed to be the first step in horizontal transfer [13]. A self-transferable GEI (e.g., ICE, conjugative transposons and other types) can move its excised DNA to a new host, where it can reintegrate with the help of an integrase enzyme at one or more specific insertion sites. GEI transfer can be mediated by conjugation or transduction, either by the element itself or via mobilization by another MGE. For some GEI the conjugation machinery closely resembles that of known plasmid-types, such as that of the SXT element of *Vibrio cholerae* [14] or the ICEMISymR7A element of *Mesorhizobium loti* [15]. For others it is very distantly related to known plasmid conjugative systems, like for ICEHin1056 of *Haemophilus influenzae*, suggesting them to be evolutionary ancient elements [16]. The findings that many GEI resemble phages by their integrase, but plasmids by their conjugative system [10], suggests they are evolutionary hybrids, which may have global control mechanisms reminiscent of both phages and plasmids. To better understand the global control of such evolutionary

hybrid elements and the consequences of the element's behavior for its bacterial host, it would be helpful to have detailed information on their transcriptional organization and regulation, which is presently still very fragmented. The SXT-element, for example, displays a key regulator (SetR) similar to the phage λ CI repressor that is autocleaved upon SOS response, after which SXT transfer becomes strongly induced [17, 18]. Preliminary regulation studies were also performed on ICEHin1056 [16] and the *Pseudomonas aeruginosa* elements pKLC102 and PAGI-2 [19], but without attaining a global level.

Our group has been studying a mobile GEI in *Pseudomonas*, *Ralstonia* and *Burkholderia*, called the *clc* element or ICE*clc* [20]. ICE*clc* has a size of 103 kilobase-pairs (kbp) and is integrated into the chromosome at the 3' 18-bp extremity of one or more *tRNA^{Gly}* genes by the help of an unusually long P4-type integrase [21, 22, 23]. The first half of ICE*clc* encodes two catabolic pathways involved in chlorocatechol (*clc* genes) and 2-aminophenol (*amn* genes) degradation [20] (Fig. 1A). The second half contains a large set of syntenic genes that were defined as life-style 'core' for sixteen GEIs originating from different *Beta*- and *Gammaproteobacteria* [24]. Among other things, this core has been proposed to encode a type IV conjugative secretion system distantly related to that of ICEHin1056 [16]. In addition, this part of ICE*clc* is assumed to encode the relaxosome complex needed for conjugation and was shown to bear a regulatory factor controlling excision and transfer [25, 26]. ICE*clc* is transferred from *P. knackmussii* B13 as donor to e.g., *Pseudomonas putida* as recipient almost exclusively in stationary phase cultures with frequencies of self-transfer $\approx 10^{-2}$ per donor. Self-transfer rates are highest in stationary phase cells grown with 3-chlorobenzoate and lower with fructose [27]. In line with this, expression

of the promoter for the integrase is highest after growth on 3-chlorobenzoate, lower on fructose and essentially absent on glucose [26]. Because of the conservation of the ICE*clc* core region among different GEIs we were interested to study its transcriptional organization, as a further step towards the understanding of the life-style program of this class of mobile elements.

In order to resolve the global transcription network of ICE*clc* in *P. knackmussii* B13, we carried out a combined approach of Northern hybridizations, reverse-transcriptase polymerase chain reaction (RT-PCR), semi-tiling array hybridization and Rapid Amplification of cDNA Ends (5'-RACE). We detected fifteen transcripts, some of which were expressed to high levels in stationary phase cultures, but – interestingly, not with all carbon sources.

RESULTS

Transcriptional organization of the ICE*clc* core region. In order to analyze the transcriptional organization of the core region of ICE*clc*, we used a combination of conventional molecular techniques and semi-tiling micro-array analyses. The ICE*clc* core spans the region between nucleotide 50,000 until the left end of the element (position 102,843; ICE*clc* numbering, GenBank Accession Number AJ617740), and comprises the most conserved stretch among a number of closely related GEI [24, 26]. Furthermore, it includes the integrase gene at the other side of ICE*clc* (Fig. 1A). Figure 1 schematically presents the analysis of intergenic regions in the ICE*clc* core region, whilst combined RT-PCR results are shown in Figure 2. RT-PCR provided a first view of potentially linked polycistronic mRNAs. In a number of cases the clear absence of RT-PCR amplicons coincided with distinct large intergenic regions predicted from the nucleotide sequence (Fig. 1B), whereas in other cases gaps were rather unexpected with respect to the close gene organization and could point to mRNA cleavage sites rather than new transcription start sites. This occurred, for example, in the regions between ORFs 62755-63176 (overlapping ORFs), ORFs 66202-66625 (12 bp intergenic region) and ORFs 73676-74436 (139 bp intergenic region, Fig. 1, 2).

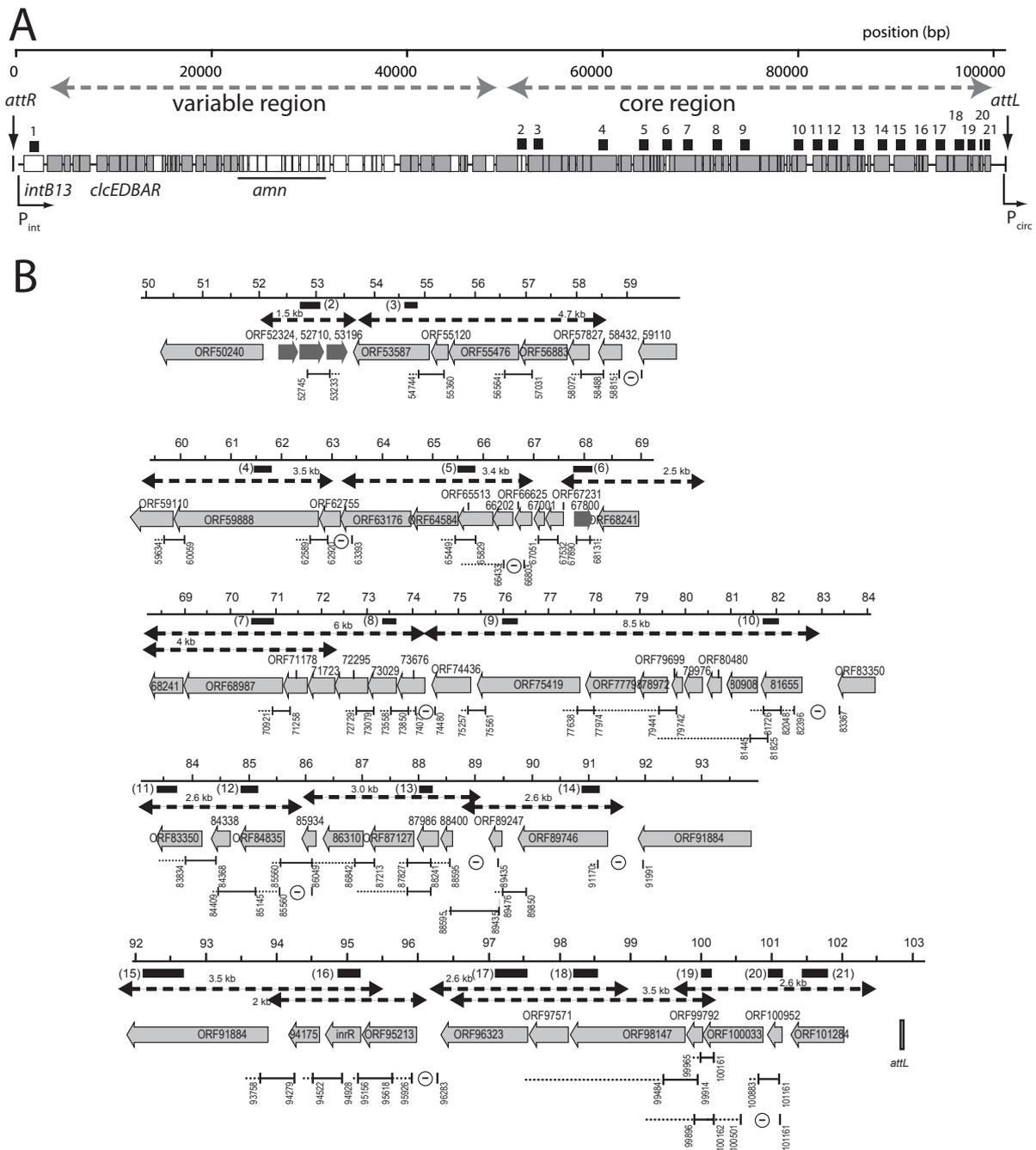


Figure 1. Global gene organization of ICE*clc* and strategy for analysis of the core region transcriptional units. A) Approximate locations of the ICE*clc* variable and core regions, with indication of gene functions known so far. Open reading frames are indicated by open (plus strand) or grey boxes (minus strand). Small numbered black stripes above point to the location of the probes used for macroblot hybridizations. B) Detailed gene structure of the core region with positions and results of RT-PCR analysis, and placement of transcript lengths (dashed lines) revealed by Northern analysis using the probes indicated as black numbered bars below the scale bar. RT-PCR indications are the following: stippled line indicates reverse transcribed regions. Solid line with two upright ends indicates the amplified region. A ‘minus’ within a circle indicates that no amplicon was obtained for that region. ORF numbering for ICE*clc* as in Genbank AJ617740.

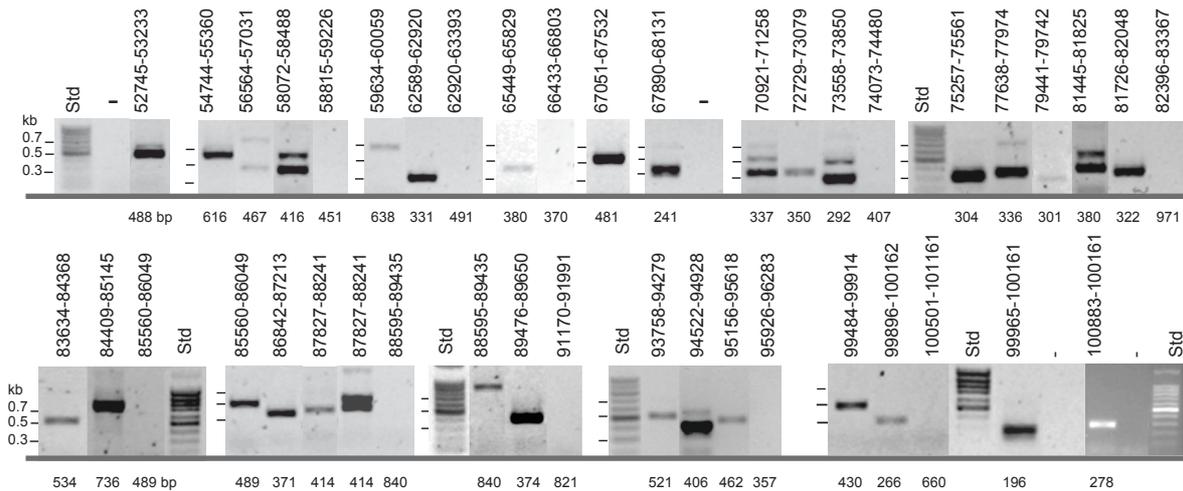


Figure 2. Reverse transcriptase-PCR amplifications of the analyzed transcript connections indicated in Figure 1. Numbers above amplicons indicate the examined region in ICEc/c numbering; numbers below the calculated amplicon size. ‘Minuses’ are negative control reactions with PCR only without reverse-transcriptase step to verify DNA contamination. Different panels are reactions run on the same gel but not necessarily in consecutive lanes. Electronic images were auto-leveled and relevant lanes were placed side-by-side using ADOBE PHOTOSHOP CS3. Std, DNA size standard (in kilobase-pairs, kb). At least one negative control was performed on every batch of purified RNA.

On top of the RT-PCR analysis we mapped the length of detectable transcripts by Northern hybridizations of RNA isolated from *P. knackmussii* B13 cultures grown to stationary phase on 3-chlorobenzoate (Fig. 3). Arguably, Northern hybridizations do not always produce clear-cut signals and often show multiple bands indicative for mRNA degradation or processing, but for most of the transcript sizes and positions proposed by RT-PCR analysis supporting evidence was provided by Northern (Fig. 1, 3). Even the breakpoints detected between ORFs 62755-63176 coincided with two detectable transcripts of around 3.5 kb that could be positioned around the gap (Fig. 1). The longest detected transcript seems to be formed by an estimated 8.5 kb polycistronic mRNA that would start upstream of ORF81655 and ending at ORF74436. It is possible, as we will argue below, that this transcript is actually synthesized as a much longer one, but cleaved somewhere in the area of the gap

identified by RT-PCR between ORF73676 and 74436. The downstream part would be formed by a 6 kb mRNA that was detectable by probes for the ORFs 68987 and 73029 (Fig. 3). Although a -10 promoter region was predicted upstream of ORF73676 by bioinformatic analysis, several others were predicted in this 8.5 kb region as well (see below and Table S1). Therefore, promoter prediction was not sufficiently accurate to support or refute the hypothesis for the 8.5 and 6 kb regions being transcribed as a single polycistronic mRNA.

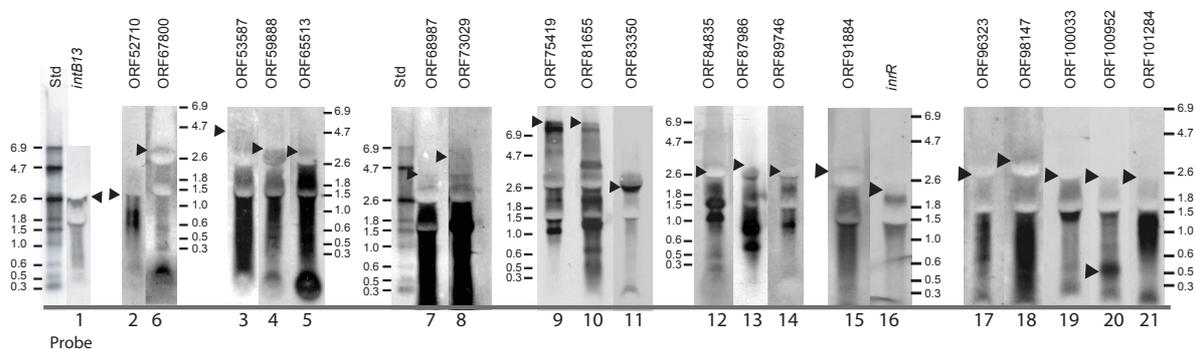


Figure 3. Compiled Northern analysis of transcript sizes in the ICEcIc core region on RNA isolated from cells grown to stationary phase on 3-chlorobenzoate. Probe used in hybridization for a respective panel is indicated as the ORF number above and the probe number below, corresponding to the indications in Figure 1. Black triangles point to the largest size determined for the hybridizing transcript. Size indications for each panel in kilobases, based on an RNA size ladder. Electronic images were auto-leveled and relevant lanes were placed side-by-side using ADOBE PHOTOSHOP CS3.

Micro-array analysis assisted transcript mapping. To complement the RT-PCR and Northern analyses, we hybridized Cy3-labeled cDNA synthesized from total RNA isolated from *P. knackmussii* B13 cultures during exponential growth on 3-chlorobenzoate and during the following stationary phase, to custom-designed semi-tiling microarrays for ICEcIc. The semi-tiling array contained a 50-mer probe at approximately every 200 bases over the whole length of ICEcIc and for both strands, each in sixfold replicate on the array. We expected that a semi-tiling array format would permit us to map the position of ICEcIc transcripts in a complementary way to

the conventional molecular analysis, which would help to reinforce the conclusions drawn on the transcriptional organization of the *ICEc/c* core. Figure 4 shows an overlay of the core gene organization and RT-PCR plus Northern derived transcriptional organization with the average micro-array hybridization signals per probe on the plus- and the minus-strand of the *ICEc/c* core region, whilst Table 1 summarizes the transcript details across all three methods. Very strikingly, most of the predicted transcripts follow a clear 5'-3' decrease in signal intensity, the slopes of which were different for each transcript region (see, for example, the region for the long transcript proposed between position 82,000 and 68,000). We think the 5'-3' decrease in intensity may partially be caused by the fact that more transcripts are formed near the transcription start, which perhaps are incompletely finished, or by preferential 3'-end degradation. This effect has been noted by others using tiling approaches for transcript determination [28]. Different slopes may be the result of varying mRNA stability and processing speed.

Semi-tiling array hybridizations confirmed most of the proposed transcripts, including breakpoints, where the slope of the decrease in hybridization intensity as a function of probe position changed abruptly (e.g., regions around position 63,000 and 86,000). An exception here was the RT-PCR detected breakpoint in between ORFs 73676 and 74436, where micro-array hybridizations did not show any aberrant change in slope of signal decrease. From this, therefore, we conclude that the long transcripts of 8.5 and 6 kb mentioned above actually originate from one 14.5 kb-long polycistronic mRNA starting at ORF81655 and ending downstream of ORF68241. This transcript would then be rapidly processed in the indicated breakpoint area, although this should be confirmed by alternative techniques.

For one other region the pattern of 5'-3' decreasing slope did not match the

hypothesis of a single transcript predicted from RT-PCR and Northern. This occurred in the area around 92,000 to 96,000 where RT-PCR had predicted a continuing transcript covering a four-gene cluster including ORF91884 (putatively encoding a DNA topoisomerase) [20], ORF94175 (putative single-strand DNA binding protein), *intR* (the proposed IntB13 activator) [26] and ORF95213 (hypothetical protein). Indeed, Northern blots had already suggested two transcripts here, not completely covering the whole region (Fig. 1 and 3), and also tiling array hybridizations showed two or even three differently 'sloped' hybridization patterns. Therefore, it might be that there is read-through from ORF94175 into ORF91884, producing the detected RT-PCR connection, but an additional promoter upstream of ORF91884 does not seem unlikely (Table S1).

Whereas most of the genes in the ICE c lc core region are organized on the minus strand (with respect to the *intB13* gene, Fig. 1), four genes are oriented on the plus strand. In general, hybridization signal intensities on micro-array for plus strand probes were lower than for the minus strand (Fig. 4), but the signals coming from the small cluster of three genes at around 53,000 and the single ORF67800 can be discerned on micro-array as being significantly above local background.

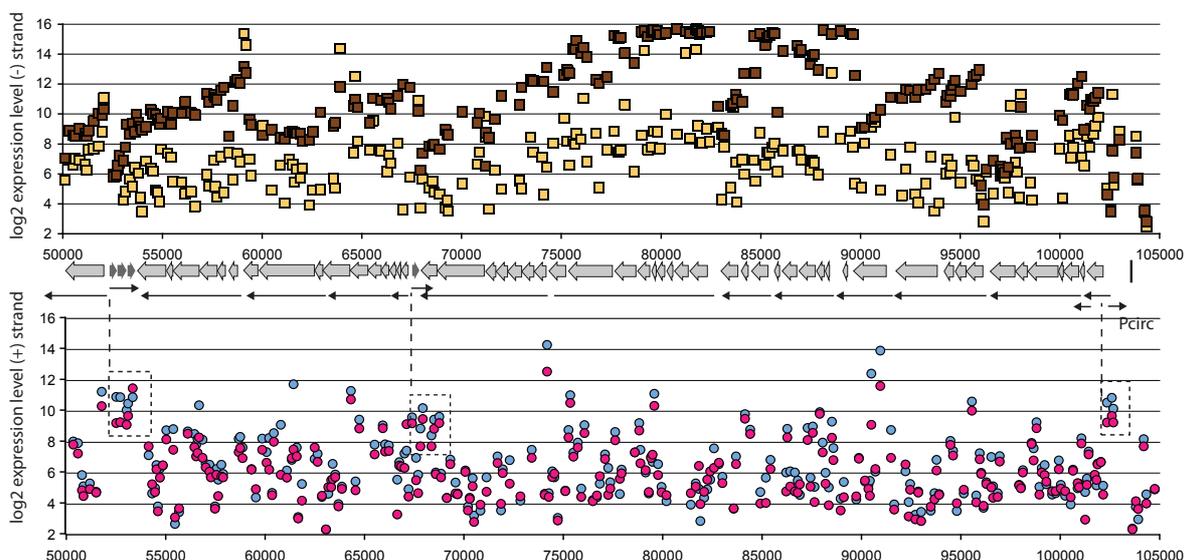


Figure 4. Transcriptome of the ICE*clc* core region. Shown is a compilation of micro-array hybridizations with minus- (top image) and plus-strand located probes (bottom image), both for exponential (yellow squares and blue circles) and stationary phase cultures (dark squares and pink circles). Data points are mean hybridization signals (on log₂-scale) from six replicate probes per array, averaged over three replicate arrays.). X-axes, position numbering on ICE*clc*. Middle part, representation of the gene locations in the ICE*clc* core region (block arrows), and the size and position of the transcripts concluded from RT-PCR and Northern (Fig. 1-3).

Table 1. Summary of ICE*clc* core transcripts.

| Transcript | Strand ^a | Size on Northern ^b | RT-PCR ^c | Promoter ^d | Log ₂ Stat-Expo Ratio ^e |
|---------------|---------------------|-------------------------------|---------------------|-----------------------|---|
| <i>intB13</i> | + | 2.5 | + | 102,729 (Pcirc) | 3.1 ± 1.0 |
| ORF50240 | - | ND (1.8) | ND | | 1.6 ± 0.6 |
| 52324-53196 | + | 1.5 (1.2) | + | 51,218 | -1.2 ± 0.4 |
| 53587-58432 | - | 4.7 (5.3) | + | 58,771 | 4.2 ± 1.4 |
| 59110-62755 | - | 3.5 (4.0) | + | 63,191 | 2.6 ± 1.3 |
| 63176-66202 | - | 3.5 (3.4) | + | 66,976 | 2.3 ± 1.7 |
| 66625-67231 | - | ND (1.0) | + | 67,610 | 5.8 ± 2.2 |
| 67800 | + | 2.6 (0.5) | + | 67,817 | -0.9 ± 0.2 |
| 68241-81655 | - | 4-6 | + | | 4.0 ± 1.7 |
| | | 8.5 (14.3) | (exc. 73676-74436) | | 5.7 ± 1.6 |
| 83350-84835 | - | 2.6 (2.3) | + | | 6.3 ± 1.6 |
| 85934-88400 | - | 3.0 (2.7) | + | 89,109 | 6.5 ± 0.8 |
| 89247-89746 | - | 2.5 (2.1) | + | | 2.2 ± 1.9 |
| 91884-95213 | - | 3.5/2 (4.1) | + | 96,204 (RACE) | 5.6 ± 1.5 |
| 96323-100033 | - | 2.5-3.5 (4.5) | | | 2.1 ± 1.6 |
| 100952 | - | 0.5 | + | | ND |
| 100033-101284 | - | 2.6 (2.0) | + | 102,270 (RACE) | 2.0 ± 0.2 |

a) plus strand is same orientation as *intB13*.

b) in kilobase observed; within brackets, size calculated from sequence.

c) ORF connections detected by reverse-transcriptase PCR on RNA from strain B13 during stationary phase after growth on 3-chlorobenzoate.

d) predicted location from bioinformatic analysis or observed by 5'RACE. Position according to numbering of AJ617740.

e) log₂-average ratio of hybridization intensities over all microarray probes covering the presumed transcript during stationary phase versus exponential phase on 3-chlorobenzoate.

Table 2. ICEcI core gene transcript abundance in *P. knackmussii* B13 cultures grown with 3-chlorobenzoate as a function of growth phase as quantified by macroblot hybridization.

| Probes | Probe number | expo | | e-stat | | 12 h | | 24 h | | 36 h | | 48 h | | 72 h | |
|-----------------------|--------------|-------------------|--------------------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| | | mRNA ^a | Std Dev ^b (%) | mRNA | Std Dev (%) |
| <i>intB13</i> | 1 | 4.5 | 11.2 | 4.3 | 12.9 | 4.6 | 15.6 | 5.1 | 28.5 | 3.2 | 5.3 | 3.4 | 0.9 | 3.5 | 14.6 |
| ORF52710 | 2 | 21.3 | 46.5 | 29.6 | 8.7 | 17.8 | 3.4 | 9.3 | 39.9 | 7.8 | 53.8 | 12.6 | 18.6 | 6.4 | 41.8 |
| ORF53587 | 3 | 4.2 | 30.2 | 2.9 | 25.9 | 2.6 | 27.1 | 1.7 | 37.3 | 3.4 | 11.9 | 3.1 | 20.4 | 1.4 | 12.2 |
| ORF59888 | 4 | 18.6 | 33 | 20 | 7.5 | 14.7 | 18.9 | 8.4 | 32.3 | 16.8 | 23.9 | 22.4 | 9.3 | 14.6 | 43.4 |
| ORF65513 | 5 | 17.3 | 19.4 | 17.1 | 0.8 | 16.7 | 10.3 | 13.4 | 9.9 | 11.8 | 9.5 | 13.5 | 2.4 | 12.4 | 10.7 |
| ORF67800 | 6 | 16.6 | 2.7 | 12.4 | 26.1 | 10.1 | 11.6 | 8 | 12.9 | 14.6 | 4.3 | 12.6 | 10.7 | 8.5 | 16.6 |
| ORF68987 ^c | 7 | 2.1 | 4.3 | 1.7 | 8.2 | 1.2 | 30.1 | 0.8 | 12.9 | 1.7 | 6.5 | 1.5 | 4.3 | 1 | 22.3 |
| ORF73029 | 8 | 2.5 | 20.8 | 1.4 | 15 | 2.1 | 18.5 | 2.6 | 15 | 2 | 14.6 | 2.2 | 2.3 | 1.5 | 10.4 |
| ORF75419 | 9 | 7.5 | 18.1 | 4.5 | 7.6 | 8.7 | 0.4 | 11.1 | 32 | 14 | 27.1 | 20.5 | 9.4 | 28 | 31.6 |
| ORF81655 | 10 | 10.2 | 30.1 | 6.4 | 35.8 | 104 | 4.8 | 168 | 24.5 | 113 | 24.3 | 191 | 14.5 | 177 | 10.9 |
| ORF83350 | 11 | 3.3 | 18.9 | 1.7 | 7 | 0.9 | 17.8 | 0.9 | 26.1 | 0.9 | 3.5 | 0.9 | 5 | 0.9 | 5.6 |
| ORF84835 | 12 | 0.4 | 14.4 | 0.3 | 38.3 | 7.4 | 53.2 | 9.5 | 7.7 | 20.6 | 28.9 | 28.1 | 29.5 | 29.8 | 22.8 |
| ORF87986 | 13 | 5 | 1 | 4.6 | 39.3 | 41.1 | 16.1 | 64.5 | 7.2 | 41.3 | 12.9 | 54.9 | 22.1 | 100 | 11.5 |
| ORF89746 | 14 | 12.9 | 34.1 | 12.9 | 10.4 | 4.9 | 12 | 2.2 | 41.2 | 1.3 | 44.9 | 3.4 | 28.9 | 3.2 | 19.8 |
| ORF91884 | 15 | 3.3 | 11.7 | 2.1 | 25.6 | 2.4 | 14.7 | 3 | 32.4 | 3.1 | 14 | 3.2 | 11.3 | 4.1 | 5.2 |
| <i>inrR</i> | 16 | 8.3 | 11.9 | 6.4 | 10.5 | 7.4 | 13.1 | 4.5 | 11.6 | 5.1 | 18.8 | 7.2 | 0.6 | 5.9 | 12 |
| ORF96323 | 17 | 3 | 13 | 1.7 | 13.3 | 1.1 | 5.5 | 1 | 2.8 | 1.3 | 12.8 | 1.1 | 14.4 | 1.2 | 13.2 |
| ORF98147 | 18 | 1.1 | 10.7 | 0.5 | 14.2 | 0.4 | 4.6 | 0.5 | 5.9 | 0.5 | 8.3 | 0.4 | 2.4 | 0.4 | 0.9 |
| ORF100033 | 19 | 30.6 | 4 | 23 | 4.3 | 26 | 8.7 | 12.3 | 16.7 | 19.6 | 14.3 | 20.4 | 22.4 | 21.5 | 16.3 |
| ORF100952 | 20 | 1.4 | 13.2 | 3.7 | 31.2 | 2.8 | 4.9 | 1.8 | 3.1 | 1.7 | 58.5 | 3.5 | 30.8 | 6.7 | 13.4 |
| ORF101284 | 21 | 3.7 | 18.9 | 4.5 | 10.1 | 2.1 | 5.8 | 1 | 23.1 | 1.9 | 9.7 | 1.5 | 11 | 1.3 | 11.5 |

a) mRNA is the average calculated amount of mRNA copies $\times 10^8$ per μg of total RNA from triplicate determinations. This value corresponds approximately to the calculated number of mRNA of this transcript per cell. Expo, exponential phase; e-stat, early stationary phase (6 h).

b) standard deviation from the average mRNA, expressed in percentage.

c) ORFs surrounded by a common border are on the same transcript.

Table 3. Quantification of ICEc/c core gene expression by dot-blot hybridization in strain B13 grown on different carbon substrates.

| Probe number and probe | Exponential phase | | | | After 24 h at stationary phase | | | | | | | |
|------------------------|-------------------|-----------------------------|-------------|----------------|--------------------------------|----------------|-------------------------|----------------|--------------|----------------|-------------|----------------|
| | 3-chlorobenzoate | | succinate | | 3-chlorobenzoate | | succinate | | fructose | | glucose | |
| | mRNA ^a | Std Dev ^b (%) | mRNA | Std Dev (%) | mRNA | Std Dev (%) | mRNA | Std Dev (%) | mRNA | Std Dev (%) | mRNA | Std Dev (%) |
| 1) <i>intB13</i> | 4.5 | 11.2 | 7.3 | 13.1 | 5.1 | 28.5 | 4.1 | 11.2 | 4.4 | 51.7 | 3.2 | 8.1 |
| 2) ORF52710 | 21.3 | 46.5 | 19.7 | 16.9 | 9.3 | 39.9 | 9.6 | 8 | 5.1 | 42.7 | 9.4 | 30.6 |
| 3) ORF53587 | 4.2 | 30.2 | 3.6 | 0.1 | 1.7 | 37.3 | 1.7 | 21.1 | 2 | 0.4 | 1.9 | 2.6 |
| 4) ORF59888 | 18.6 | 33 | 16.9 | 2.3 | 8.4 | 32.3 | 12.9 | 18.6 | 16.8 | 7.3 | 23.8 | 15.9 |
| 5) ORF65513 | 17.3 | 19.4 | 19.5 | 2.8 | 13.4 | 9.9 | 12.7[‡] | 5.3 | 13.8 | 7.6 | 13.8 | 11 |
| 6) ORF67800 | 16.6 | 2.7 | 16.6 | 5.5 | 8 | 12.9 | 12.7 | 18.3 | 11.6 | 33.7 | 17.9 | 38.6 |
| 7) ORF68987 | 2.1 | 4.3 | 2.1 | 11 | 0.8 | 12.9 | 0.8 | 0.2 | 1.3 | 13.8 | 1.1 | 11.7 |
| 8) ORF73029 | 2.5 | 20.8 | 2.9 | 12.6 | 2.6* | 15 | 0.9[‡] | 18.2 | 1.4 | 6.7 | 1.1 | 4.2 |
| 9) ORF75419 | 7.5 | 18.1 | 7.3 | 6.8 | 11.1 | 32 | 3[‡] | 3.9 | 3.9 | 3.3 | 2.8 | 5.4 |
| 10) ORF81655 | 10.2 | 30.1 | 18.7 | 36.6 | 168* | 24.5 | 6.3 | 2.7 | 45.7* | 3.6 | 9.2 | 27 |
| 11) ORF83350 | 3.3 | 18.9 | 2.8 | 16.5 | 0.9 | 26.1 | 0.5 | 37.3 | 0.5 | 14.5 | 0.4 | 10.2 |
| 12) ORF84835 | 0.4 | 14.4 | 0.3 | 16.3 | 9.5* | 7.7 | 0.3 | 25.8 | 1.7 | 16.1 | 0.3 | 1.5 |
| 13) ORF87986 | 5 | 1 | 5.2 | 0.1 | 64.5* | 7.2 | 5.5[‡] | 0.4 | 14.2 | 26.9 | 5.5 | 2 |
| 14) ORF89746 | 12.9 | 34.1 | 24.4 | 19.8 | 2.2 | 41.2 | 2.1 | 17.3 | 2.1 | 36.7 | 0.5 | 15.4 |
| 15) ORF91884 | 3.3 | 11.7 | 4.5 | 3 | 3 | 32.4 | 1.6[‡] | 3.2 | 2.3 | 33.1 | 1.1 | 5.9 |
| 16) <i>inrR</i> | 8.3 | 11.9 | 8.2 | 21.3 | 4.5 | 11.6 | 4 | 7.5 | 6.4 | 8.1 | 4.9 | 39.7 |
| 17) ORF96323 | 3 | 13 | 5.3 | 27.1 | 1 | 2.8 | 1.8 | 35.6 | 0.9 | 53.2 | 1.1 | 31.8 |
| 18) ORF98147 | 1.1 | 10.7 | 1.5 | 5.1 | 0.5 | 5.9 | 0.4[‡] | 7 | 0.4 | 3.7 | 0.4 | 1.7 |
| 19) ORF100033 | 30.6 | 4 | 40.4 | 20.2 | 12.3 | 16.7 | 17.6 | 18 | 22.9 | 6.4 | 22.2 | 30.2 |
| 20) ORF100952 | 1.4 | 13.2 | 2.2 | 22.2 | 1.8[§] | 3.1 | 0.9 | 1.7 | 1.9 | 7.9 | 0.9 | 29.4 |
| 21) ORF101284 | 3.7 | 18.9 | 3.2 | 6.9 | 1 | 23.1 | 1[‡] | 7.9 | 1.1 | 1.9 | 1.2 | 14.5 |

a) mRNA is the average calculated amount of mRNA copies $\times 10^3$ per μg of total RNA from triplicate determinations. This value corresponds approximately to the calculated number of mRNA of this transcript per cell.

b) standard deviation from the average mRNA, expressed in percentage.

c) *, values statistically significantly different to the cells cultivated with other carbon sources. §, values statistically significantly different to the cells cultivated with succinate or glucose. ‡, values statistically significantly different between exponential phase and stationary phase. $P \leq 0.05$ in pairwise Student's T test

Carbon substrate dependent expression of ICEc/c core genes.

Micro-array hybridizations clearly demonstrated that most of the core genes on the minus strand are upregulated in stationary phase conditions (Table 1, Fig. 4), with \log_2 -fold changes ranging from 2^2 (e.g., for ORF50240 or the cluster of genes between 96,000 and 100,000) to 2^7 (e.g., ORF81655). RNAs from a larger number of different growth conditions were hybridized in dot-blot format using digoxigenin-labeled probes representative for all proposed transcripts (Tables 2 and 3). This showed that the expression of the highly abundant core transcripts represented by ORF81655, ORF87986 and ORF84835 (Table 2) actually started in the first twelve hours after reaching stationary phase and then increased continuously further up to 72 h. In contrast, transcription from the three plus strand ORFs 52324-53196 seemed to 'peak' in very early stationary phase, but then successively decreased (Table 2). Hybridizing blotted RNAs from *P. knackmussii* B13 grown to stationary phase on different carbon substrates showed, interestingly, that the three transcripts 68241-81655 (represented by probes 7, 8, 9 and 10), 83350-84835 (probes 11 and 12), and 85934-88400 (probe 13) were highly induced only in stationary phase cells that had been cultured with 3-chlorobenzoate or fructose, but not at all with succinate or glucose (Table 3). Highest induction of the ICEc/c core region genes in stationary phase cells grown with 3-chlorobenzoate is in agreement with previous experiments that showed the highest proportion of excised ICEc/c and highest ICEc/c transfer rates in cells cultured on 3-chlorobenzoate to stationary phase [26, 27].

Another gene that produced relatively high signals in dot-blot hybridizations was ORF100033, which urged us to analyze its expression more

conspicuously by RT-PCR. Contrary to RNA isolated in stationary phase from 3-chlorobenzoate or fructose-grown cultures, consistently no RT-PCR product was obtained for the intergenic region between ORF100952 and ORF101284 on RNA from cells that had been cultivated with glucose (Fig. 5, panels d and e). RNA isolated from all three substrate conditions did produce a smaller RT-PCR fragment directly upstream of ORF100952 (Fig. 5B panel b), suggesting that an additional promoter exists that produces a transcript covering ORF100952 only. In fact, Northern hybridizations with a probe for ORF100952 produced an additional band of 0.5 kb length (Fig. 3). The promoter located in front of ORF101284 might thus be specifically repressed after growth on glucose (and perhaps succinate), or specifically activated after growth on 3-chlorobenzoate and fructose.

Promoter analysis. Results from 5'-RACE were not as conclusive as expected. Although various amplicons were produced from cDNA ends, only few matched the start region for transcripts detected by RT-PCR, Northern and micro-array. In contrast, the start site for the transcript covering *inrR* could be mapped in the region upstream of ORF95213 to a thymine located 25 nt upstream of the ORF95213 start codon. Interestingly, the corresponding -10 box (TGTCGATCCT) and -35 (TTGACT) are close to the proposed consensus sequence of σ^s and not σ^{70} , suggesting it is controlled by RpoS [26]. This could explain a higher abundance of this transcript during stationary compared to exponential phase as seen on micro-array (Fig. 4). 5'-RACE also produced a clear transcription start upstream of ORF101284 (position 102,270), which is in agreement to the other data and suggests a promoter

being present at positions 102,283 to 102,288 (-10, TATTAC) and 102,300 to 102,305 (-35 region, TTGCAG). On the contrary, no 5'RACE product but a very weak product was obtained by primer extension in the region upstream of ORF81655, which located at ~250 bp upstream of the start codon (results not shown), even though this transcript was among the most abundant ones of the ICE*c* core region (Fig. 4). In a few other cases, bioinformatic searches identified promoter signatures which locate in regions where transcripts were deemed to start (Table 1, S1), but their nature remains to be experimentally verified.

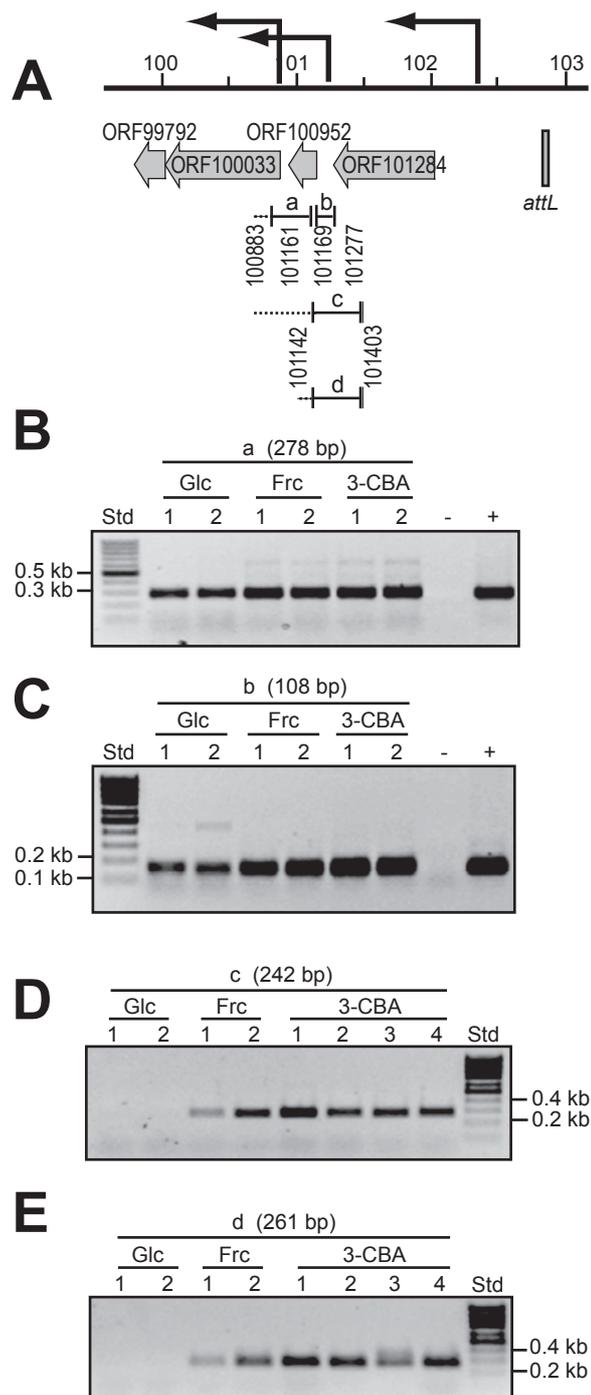


Figure 5. Carbon substrate-dependent transcript linkage in the region at the outermost ICE*clc* left end. A) Gene organization, reverse-transcribed regions and PCR amplicons. Arrows to the left point to inferred promoters. B) RT-PCR results for amplicon (a). C) idem for amplicon (b). D) Amplicon (c). E) Amplicon (d). All RNAs sampled from cultures during stationary phase after growth on the indicated carbon source. Glc, glucose. Frc, fructose. 3-CBA, 3-chlorobenzoate. Numbers below point to independent replicate reactions. -, PCR but without RT-step.+, PCR on B13 genomic DNA as template.

DISCUSSION

By using a combination of semi-tiling micro-array hybridization and conventional techniques for transcription analysis, we obtained a highly detailed picture on the transcriptional organization of the ICE*lc* core region. To our knowledge, this is one of the first examples of tiling micro-array in combination with RT-PCR and Northern hybridizations to study transcriptional organization of mobile DNA elements, the only other one currently being a study on the plasmid pCAR1 of *P. resinovorans* [29]. We conclude from our results that such a combined approach can give excellent complementary data and retrieve details that either one of the typical transcription approaches alone cannot obtain. Except for a few locations, the results from all approaches on ICE*lc*'s transcriptome were mostly in agreement with each other, or critically supported omissions in each of them individually. Fourteen transcripts were detected by RT-PCR and Northern; one more was inferred from micro-array hybridization (ORF50240). Some transcripts seem clearly part of one larger but rapidly cleaved polycistronic mRNA (e.g, ORF68241-81655), whereas in one case (ORF59110-67231) three transcripts were consistently detected but gene organization suggests close linkage.

The importance of the ICE*lc* core gene region lays in its proposed control of the element's behavior: excision, self-transfer, maintenance and reintegration. Even though still only few ICE*lc* core genes have clear identifiable homology to known proteins (Additional file 1, Table S1), the region as a whole is largely conserved in a large collection of other GEI, underscoring its functional

importance for life-style [23, 24]. The 14 or 15 transcripts in the ICE*clc* core region, including a long 14.5 kb transcript (Fig. 1, 4), is in the order of transcript numbers typically found for self-transfer and maintenance functions of conjugative plasmids (e.g., eight for R27 in *E. coli* [30], 14 for pCAR1 in *P. resinovorans* [29]). Four of the core transcripts (between ORF53587 and ORF73676) might code for a type IV secretion system (mating pair formation complex) similar to that of ICE*Hin1056* from *H. influenza* (Fig. 6, Additional file 1, Table S1) [16]. Furthermore, the ORF50240 product has weak homology to a relaxase and ORF91884 has a DNA topoisomerase domain, suggesting possible implication in a relaxosome complex involving the excised circular ICE*clc* form (Additional file 1, Table S1). A comparison with the ICE*Hin1056* transcriptional organization in this area shows a number of differences, which are likely due to extensive gene arrangements during evolutionary divergence between the two elements (Fig. 6). For example, the long ICE*Hin1056* transcript covering the mating pair complex (*PilL*, *TraB*, *TraD* etc.), is interrupted on ICE*clc* by the reversely oriented ORF67800. The transcript containing ORF73676 (the presumed *pilL*) is not the start, but part of a much longer transcript starting at ORF81655 on ICE*clc*. Second difference between ICE*clc* and ICE*Hin1056* relates to the large inversion of the genes *tfc21* to *tfc24* (Fig. 6). ICE*Hin1056* data suggested two transcripts in this region, with one being formed by the presumed regulatory gene *tfc24* [16]. In contrast, on ICE*clc* ORF57827 (the homologue of *tfc24* on ICE*clc*, Fig. 6) is apparently the second gene of a six-gene transcript.

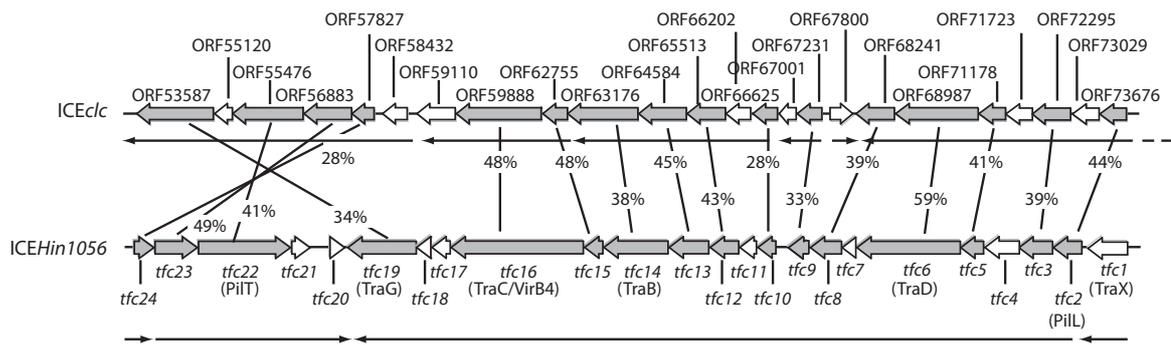


Figure 6. Comparison of the *tfc*-like gene region on ICE*clc* with ICE*Hin1056* from *H. influenzae*. Lines indicate percentage amino acid similarity between common genes (grey-shaded). Genes indicated in open arrows have no significant homologies among the two ICE. Arrows underneath point to the transcriptional organization in this region. Data on ICE*Hin1056* redrawn from [16].

The relative abundance of transcripts in the region ORF50240 to ORF81655 of ICE*clc* was up to 64-fold (microarray) different between stationary and exponential phase (Fig. 2 and 3, Table 1). If the postulate is correct that these genes would encode part of the type IV secretion system necessary for ICE*clc* transfer (i.e., the equivalent of the Mating Pair Formation or *mpf* complex in conjugative plasmids [6]), their induction would be much more pronounced than what is usual for plasmid conjugative systems. In most cases, the *mpf* genes are either weakly expressed or tightly regulated and inducible [6], the reason presumably being that expression of the conjugative apparatus is energy costly and could favor male-type specific phage infection. Tight control of the transfer genes of plasmids is often achieved by autoregulatory loops, such as the IncP-9 pWW0 plasmid *traA* and *mpfR* genes that control the relaxosome complex and *mpf* operons, respectively [31]. Also, the presumed genes involved in conjugative transfer of the IncP-7 plasmid pCAR1 in *Pseudomonas putida* and *P. resinovorans* are expressed at low and similar

transcriptional level (without further specification) during growth on succinate or carbazole [29]. Induction of the putative conjugative system of ICE*clc* would thus be more similar to the type of induction found in the SXT element [18], which is a hybrid between phage-lambda type control and plasmid-like conjugation. However, none of the ICE*clc* functions has any significant sequence similarity to the SetR – SetC – SetD regulators of SXT, nor to the CI repressor from *λ*. The type of program that seems to be followed by ICE*clc* (integrase induction, excision, and global response of the core region) is indeed reminiscent of a phage program [9], but we conclude that control layers are different from phage type. The finding that basically eight of fifteen identified transcripts in the ICE*clc* core region are upregulated during stationary phase, suggests a coordinated global control mechanism, which is perhaps assisted by the stationary phase sigma factor RpoS. Indeed, some evidence for RpoS control was obtained from sequence motifs in the *inrR* promoter. It is interesting to speculate as to what would be the ecological or physiological advantage for ICE*clc* to become active during stationary phase. One hypothesis is that of the ‘sinking ship’ : the element senses that its host survival (and therefore that of itself) is endangered and tries to escape to a more favorable host cell (even though this must be in its immediate vicinity). Even more intriguing is perhaps the carbon substrate-specific upregulation of ICE*clc* activity, which is highest after growth on 3-chlorobenzoate, less with fructose and very low with glucose or succinate as carbon sources. Upregulation of the ICE*clc* core region expression in stationary phase cells grown with 3-chlorobenzoate is in agreement with previous results showing increased activity of the integrase promoter [26], increased proportion of

ICE*clc* excised DNA and increased ICE*clc* transfer rates [27]. Since it is assumed that during stationary phase cells have depleted their carbon source, the carbon source can no longer be directly be responsible for the activation, but somehow must have generated a 'memory' effect which triggers ICE*clc* response. In this light, the repression seen for transcription read-through from ORF101284 with glucose and succinate might point to a Crc-type regulation of catabolite repression in *Pseudomonas* [32, 33], although for the time being no specific Crc binding motifs were detected in the ICE*clc* core region.

CONCLUSIONS

In conclusion, we have identified fifteen transcripts covering the presumed core region for behavioral functions of ICE*clc*. Eight of those are concertedly upregulated during stationary phase, but only after previous growth of the cells on 3-chlorobenzoate or fructose, which explains previous results that have seen highest ICE*clc* transfer rates under such conditions [27]. The number and lengths of ICE*clc* transcripts are similar to that found for typical conjugative plasmid systems, yet the mode of global transcription control is more reminiscent for phage-type control. We thus conclude that the hybrid transcriptional control mode comprising both conjugative plasmid and phage strategies has been selected in mobile elements of the ICE*clc* group.

METHODS

Growth conditions and harvesting. *P. knackmussii* B13, the original host of ICE*cl*, was cultivated in minimal medium (MM) based on the type 21C medium [34]. This MM was complemented with 3-chlorobenzoate, fructose, glucose (all at 10 mM) or 15 mM succinate, and the bacteria were grown at 30°C. Bacterial growth was assessed from culture turbidity at 600 nm (OD₆₀₀). Cells were recovered during exponential phase (OD₆₀₀ of 0.4) or early stationary phase (OD₆₀₀ = 1.2), which was defined as the point where growth began to cease plus one period equivalent to the shortest generation time on that substrate. Bacteria were also recovered 12, 24, 36, 48 or 72 h after the beginning of the stationary phase. For RNA isolation, 100 ml of culture was immediately harvested by centrifugation (at 15,000 × *g* for 1 min at 4°C) and the supernatant was decanted. Cell pellets were resuspended in 4 ml RNeasy Protect Bacteria Reagent (QIAGEN GmbH). After 5 min incubation, the suspensions were centrifuged again (at 5,000 × *g* for 5 min at room temperature); the supernatant was discarded and pellets were stored at -80°C.

RNA isolation. Prior to RNA extraction, pellets were slowly thawed, then resuspended in 0.5 ml TES buffer [10 mM Tris-HCl (pH 8.0), 1 mM EDTA, 100 mM NaCl], followed by addition of and mixing with 0.25 ml lysis solution [20 mM sodium acetate (pH 5.5), 1 mM EDTA, 0.5% SDS]. After that, the total RNA was further purified by the hot acid-phenol method as described previously [35]. RNA samples were purified from contaminating DNA by treatment with 50 U of DNase I (RNase free; Roche) during 1 h at 37°C.

Finally, the RNA was dissolved in 50 μ l diethylpyrocarbonate (DEPC)-treated water and quantified by absorbance at 260 and 280 nm on a NanoDrop spectrophotometer (Witec AG). The integrity of RNA was determined by agarose gel electrophoresis and the absence of DNA was verified by PCR.

Reverse transcription PCR (RT-PCR). Reverse transcription was made on RNA isolated from cultures grown with 3-chlorobenzoate, glucose or fructose, and harvested 24 h after the beginning of stationary phase. 0.5 mg of total RNA was denatured by heating at 65°C and reverse transcribed using the Omniscript RT kit (QIAGEN GmbH) following the instructions of the manufacturer, using primers listed in Additional file 1, Table S2. Primer designations refer to their exact position on *ICEc/c* according to the numbering in AJ617740 (Genbank Accession number). 30 cycles of PCR amplification with the produced cDNA templates was performed with the HotStarTaq Master Mix kit (QIAGEN GmbH), using one tenth of volume from the reverse transcription reaction and 10 μ M of a pair of specific primers (Additional file 1, Table S2). Amplification of regions between ORF94175 and *inrR* known to be co-transcribed served as positive control for the quality of the RT-PCR reaction. Finally, for each RNA sample, a PCR was performed without reverse transcriptase step, in order to control for the absence of DNA contamination.

Mapping of transcriptional start sites. The 5' end of the transcript including *inrR* was mapped with the SMART RACE cDNA Amplification Kit (Clontech Laboratories, Inc.) according to the manufacturer's protocol. cDNA

was synthesized from 0.5 μg RNA with the primer 95,129rv, which is located within the *inrR* region. After generation of RACE-Ready cDNA, a PCR and a nested PCR were performed by using the *inrR*-specific primer 95,156rv plus the Universal Primer A (UPM, Clontech), and the *inrR* primer 95,677rv plus the Nested Universal Primer A (NUP), respectively. Both PCR products were sequenced using a further *inrR* specific primer 95,790rv in the BigDye Terminator v3.1 cycle sequencing kit (Applied Biosystems), and were separated on ABI PRISM 3100 Genetic Analyzer (Applied Biosystems). A further successful mapping was deployed with 5'RACE on the transcript starting upstream of the most distal ICE*clc* ORF101284. 5'RACE reactions for the regions upstream of ORFs 58432, 66202, 73676, 81655, 88400, and 89746 did not produce specific fragments.

Digoxigenin-labeled probe synthesis. DNA regions of between 126 and 560 bp of 21 selected ORFs from the *clc* element's core region (Fig. 1) were amplified by PCR for probe synthesis (Additional file 1, Table S3). One of the PCR primers (reverse complementary to the targeted ORF) included the sequence for the promoter region of T7 RNA polymerase. Antisense digoxigenin-labeled RNA probes were then synthesized from ~ 1 μg of purified PCR product by using T7 RNA polymerase according to instructions of the suppliers (Roche Applied Science).

Northern hybridization. 20 μg of total RNA were incubated in 20 μl (total volume) of denaturation buffer (containing 1 M glyoxal, 25% v/v dimethylsulfoxide, 10 mM sodium phosphate, pH 7.0) for 1 h at 50°C. 100 ng

of a digoxigenin-labeled RNA molecular weight marker I (0.3 – 6.9 kb, Roche Diagnostics) was treated similarly. A volume of 0.2 µl of a 10 mg/ml ethidium bromide solution and 1 µl loading buffer (containing 50% sucrose, 15 mg/ml bromophenol blue in DEPC-treated H₂O) were added to the samples at the end of the incubation period and mixed. Fragments were separated at 50 V on a 1% agarose gel in 10 mM sodium phosphate buffer (pH 7.0). RNA was subsequently transferred from gel onto Hybond N⁺ nylon membrane (Amersham Biosciences) in 10 × concentrated SSC solution (containing 3 M NaCl and 0.3 M sodium citrate dissolved in demineralized H₂O) with the help of the VacuGene XL system (Amersham Biosciences) for 3.5 h at a vacuum of 50 mbar. After transfer, RNA was fixed to the membrane with a UV crosslinker (CX-2000, UVP) at a dose of 0.3 J per cm². Immediately before hybridization, the membrane was rinsed with 20 mM Tris-HCl (pH 8.0) at 65°C for 10 min to remove glyoxal. The hybridization was performed in DIG Hybridization buffer (Roche Diagnostics) for 15 h at 68°C. The washing steps and the immuno-chemiluminescent detection were done according to the supplier's instructions (Roche Diagnostics) using alkaline-phosphatase-conjugated anti-digoxigenin Fab fragments and CSPD as reagent for the chemiluminescence reaction. Light emission was detected on Hyperfilm ECL (Amersham Biosciences).

Dot blot hybridization and relative mRNA quantification. Total RNA extracted from different growth phases and substrates were further analyzed for expression of genes from the core region of *ICEc/c* in 96-well format dot-blot hybridization. RNA was isolated from three independent cultures of strain

B13 grown with 3-chlorobenzoate at exponential phase, early-stationary phase, as well as at 12, 24, 36, 48 and 72 h after the beginning of stationary phase. Furthermore, duplicate cultures of B13 grown with glucose, fructose and succinate harvested after 24 h, and duplicate cultures grown on succinate in exponential phase were used for RNA purification as well. 15 μ l Aliquots of dilutions containing 1, 0.3, and 0.1 μ g denatured total RNA were dot-blotted using a 96-well manifold (Gibco Life Technologies) onto positively charged nylon transfer membranes (Hybond-N⁺, Amersham Biosciences AG). Different concentrations of denatured PCR products (2.5, 1, 0.5, 0.25, 0.1, 0.05, 0.025 and 0.01 ng) comprising the respective targeted ORF were included on the same blot. RNA was fixed to the membrane with a UV crosslinker before hybridization as described above. Films were scanned and spot intensities were calculated by densitometry using the Image Quant TL program (v2005, Molecular Dynamics, Sunnyville, USA) as grey intensity per standardized surface. The signal intensity of each spot was then compared to the standard curve of DNA dilutions on the same blot to calculate an 'equivalent number of DNA copies', and divided by the total amount of RNA in the spot to normalize to a value of 'equivalent number of copies per μ g RNA'.

Microarray design. A series of 950 non-overlapping 50-mer probes was designed to cover both coding and non-coding regions of the ICElc sequence (Acc. No. AJ617740) at approximate distances of 200 bp. Probes were designed using the program OLIGOARRAY version 2.1 [36] with a melting temperature range of 92 to 99°C and a probe GC content range of 52 to 72%. Probes were further designed to not cross-hybridize with gene products from

the following potential host strains of the ICE clc element: *Burkholderia xenovorans* LB400 (Acc. No. CP000270-CP000272), *P. putida* F1 (Acc. No. CP000712), *P. putida* KT2440 (Acc. No. AE015451), *P. aeruginosa* PAO1 (Acc. No. AE004091), *Cupriavidus necator* JMP134 (Acc. No. CP000090-CP000093), and *Ralstonia metallidurans* CH34 (Acc. No. CP000352-CP000355). An additional 93 probes were designed to target housekeeping genes from the potential host strains and 8 probes were designed to target positive/negative controls (GFP, luciferase, and mCherry [37] transcripts). The microarray was manufactured by Agilent Technologies (Santa Clara, CA) in the 8 × 15,000 probe format and each unique probe was synthesized at six randomized spatial locations on the array. The microarray design has been deposited in the NCBI Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo>) under accession number GSE20461.

Microarray hybridization and analysis. Total RNA was isolated and purified from *P. knackmussii* B13 cultures during exponential growth on 3-chlorobenzoate and during subsequent stationary phase after 24 h as described above. For microarray hybridizations, cDNA was synthesized from total RNA and directly labeled with cyanine-3-dCTP using a modification of a protocol described elsewhere [38]. Briefly, each 50- μ L reaction contained 10 μ g of total RNA, 1.25 μ g of random hexanucleotide primers (Promega), 100 μ M each of unlabeled dATP, dGTP, and dTTP (Invitrogen), 25 μ M of unlabeled dCTP (Invitrogen), 25 μ M of cyanine-3-labeled dCTP (Perkin-Elmer), 25 U SUPERase•In (Ambion), and 400 U Superscript II reverse transcriptase (Invitrogen). Reactions were performed by heating at 42°C for 2

hours followed by 70°C for 10 min. RNA was then removed by adding 100 mM NaOH, heating to 65°C for 20 min, and neutralizing with 100 mM HCl and 300 mM sodium acetate (pH 5.2). Labeled cDNA products were purified using the MinElute PCR purification kit (Qiagen) and the quantity and incorporation frequency of cyanine-3-labeled dCTP were calculated using the MICROARRAY function on a NanoDrop Spectrophotometer. Sixty ng of labeled cDNA was then loaded onto each microarray, hybridized for 17 hours at 65°C, and washed and scanned as described for labeled cRNA in the One-Color Microarray-Based Gene Expression Analysis Manual (Agilent). The fragmentation step (heating to 60°C for 30 minutes) was omitted.

Hybridization signal intensities were quantified from microarray image scans using AGILENT FEATURE EXTRACTION software version 9.5.3 (Agilent). Microarray data were normalized and globally scaled over the array using GENESPRING GX software with the RMA algorithm and quantile normalization [39, 40]. Mean probe signals were calculated for each of the three biological replicates and were plotted against their position on the *ICEc/c* sequence for both strands and for RNAs isolated during exponential and stationary phases. All microarray data have been deposited in the NCBI Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo>) under accession number GSE20461.

Bioinformatic tools. Putative promoters, terminators and transcription factor binding sites were predicted by using the BPROM and FindTerm programs on <http://www.Softberry.com>. The map of *ICEc/c* was designed from SeqBuilder of the Lasergene software package (version 6.1.4, DNASTAR, Inc).

AUTHOR'S CONTRIBUTIONS

MG designed and performed transcription analysis. NP and MM performed microarray experiments. DJ designed probes for microarray and developed labeling and hybridization protocol. MG and VS carried out 5'RACE analysis. JvdM designed experiments and wrote the manuscript. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

The work of MG, MM and JvdM was supported by grants 3100A-108199 and 3100-67229 from the Swiss National Science Foundation. NP is supported by a fellowship from the Faculty of Biology and Medicine of the University of Lausanne.

REFERENCES

1. Gogarten JP, Townsend JP. *Horizontal gene transfer, genome innovation and evolution*. Nat Rev Microbiol 2005, 3:679-682.
2. Gaillard M, Pernet N, Vogne C, Hagenbüchle O, van der Meer JR. *Host and invader impact of transfer of the clc genomic island into Pseudomonas aeruginosa PAO1*. Proc Natl Acad Sci U S A 2008, 105:7058-7063.
3. Frost LS, Leplae R, Summers AO, Toussaint A. *Mobile genetic elements: the agents of open source evolution*. Nat Rev Microbiol 2005, 3:722-732.
4. Toussaint A, Merlin C. *Mobile elements as a combination of functional modules*. Plasmid 2002, 47:26-35.
5. Adamczyk M, Jagura-Burdzy G. *Spread and survival of promiscuous IncP-1 plasmids*. Acta Biochim Pol 2003, 50:425-453.
6. Thomas CM. *Transcription regulatory circuits in bacterial plasmids*. Biochem Soc Trans 2006, 34:1072-1074.
7. Christie PJ, Atmakuri K, Krishnamoorthy V, Jakubowski S, Cascales E. *Biogenesis, architecture, and function of bacterial type IV secretion systems*. Annu Rev Microbiol 2005, 59:451-485.
8. Ptashne M. *A Genetic Switch: Phage Lambda Revisited* Third Edition edn. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press; 2004.
9. Osterhout RE, Figueroa IA, Keasling JD, Arkin AP. *Global analysis of host response to induction of a latent bacteriophage*. BMC Microbiol 2007, 7:82.
10. Juhas M, van der Meer JR, Gaillard M, Harding RM, Hood DW, Crook DW. *Genomic islands: tools of bacterial horizontal gene transfer and evolution*. FEMS Microbiol Rev 2009, 33:376-393.
11. Hacker J, Carniel E. *Ecological fitness, genomic islands and bacterial pathogenicity: A Darwinian view of the evolution of microbes*. EMBO Rep 2001, 2:376-381.
12. Dobrindt U, Hochhut B, Hentschel U, Hacker J. *Genomic islands in pathogenic and environmental microorganisms*. Nat Rev Microbiol 2004, 2:414-424.
13. Mathee K, Narasimhan G, Valdes C, Qiu X, Matewish JM, Koehrsen M, Rokas A, Yandava CN, Engels R, Zeng E, Olavarietta R, Doud M, Smith RS, Montgomery P, White JR, Godfrey PA, Kodira C, Birren B, Galagan

- JE, Lory S. *Dynamics of Pseudomonas aeruginosa genome evolution*. Proc Natl Acad Sci U S A 2008, 105:3100-3105.
14. Beaber JW, Hochhut B, Waldor MK. *Genomic and functional analyses of SXT, an integrating antibiotic resistance gene transfer element derived from Vibrio cholerae*. J Bacteriol 2002, 184:4259-4269.
 15. Ramsay JP, Sullivan JT, Stuart GS, Lamont IL, Ronson CW. *Excision and transfer of the Mesorhizobium loti R7A symbiosis island requires an integrase IntS, a novel recombination directionality factor RdfS, and a putative relaxase RlxS*. Mol Microbiol 2006, 62:723-734.
 16. Juhas M, Crook DW, Dimopoulou ID, Lunter G, Harding RM, Ferguson DJ, Hood DW. *Novel type IV secretion system involved in propagation of genomic islands*. J Bacteriol 2007, 189:761-771.
 17. Burrus V, Waldor MK. *Control of SXT integration and excision*. J Bacteriol 2003, 185:5045-5054.
 18. Beaber JW, Hochhut B, Waldor MK. *SOS response promotes horizontal dissemination of antibiotic resistance genes*. Nature 2004, 427:72-74.
 19. Klockgether J, Wurdemann D, Wiehlmann L, Tummler B. *Transcript profiling of the Pseudomonas aeruginosa genomic islands PAGI-2 and pKLC102*. Microbiology 2008, 154:1599-1604.
 20. Gaillard M, Vallaeyts T, Vorholter FJ, Minoia M, Werlen C, Sentchilo V, Puhler A, van der Meer JR. *The clc element of Pseudomonas sp. strain B13, a genomic island with various catabolic properties*. J Bacteriol 2006, 188:1999-2013.
 21. Ravatn R, Studer S, Springael D, Zehnder AJB, van der Meer JR. *Chromosomal integration, tandem amplification, and deamplification in Pseudomonas putida F1 of a 105-kilobase genetic element containing the chlorocatechol degradative genes from Pseudomonas sp. strain B13*. J Bacteriol 1998, 180:4360-4369.
 22. Ravatn R, Studer S, Zehnder AJB, van der Meer JR. *Int-B13, an unusual site-specific recombinase of the bacteriophage P4 integrase family, is responsible for chromosomal insertion of the 105-kilobase clc element of Pseudomonas sp. strain B13*. J Bacteriol 1998, 180:5505-5514.
 23. Sentchilo V, Czechowska K, Pradervand N, Minoia M, Miyazaki R, van der Meer JR. *Intracellular excision and reintegration dynamics of the ICE_{clc} genomic island of Pseudomonas knackmussii sp. strain B13*. Mol Microbiol 2009, 72:1293-1306.
 24. Mohd-Zain Z, Turner SL, Cerdeño-Tárraga AM, Lilley AK, Inzana TJ, Duncan AJ, Harding RM, Hood DW, Peto TE, Crook DW. *Transferable antibiotic resistance elements in Haemophilus influenzae share a common evolutionary origin with a diverse family of syntenic genomic*

- islands*. J Bacteriol 2004, 186:8114-8122.
25. Sentchilo VS, Zehnder AJB, van der Meer JR. *Characterization of two alternative promoters and a transcription regulator for integrase expression in the clc catabolic genomic island of Pseudomonas sp. strain B13*. Mol Microbiol 2003, 49:93-104.
 26. Minoia M, Gaillard M, Reinhard F, Stojanov M, Sentchilo V, van der Meer JR. *Stochasticity and bistability in horizontal transfer control of a genomic island in Pseudomonas*. Proc Natl Acad Sci U S A 2008, 105:20792-20797.
 27. Sentchilo VS, Ravatn R, Werlen C, Zehnder AJB, van der Meer JR. *Unusual integrase gene expression on the clc genomic island of Pseudomonas sp. strain B13*. J Bacteriol 2003, 185:4530-4538.
 28. Guell M, van Noort V, Yus E, Chen WH, Leigh-Bell J, Michalodimitrakis K, Yamada T, Arumugam M, Doerks T, Kuhner S, Rode M, Suyama M, Schmidt S, Gavin AC, Bork P, Serrano L. *Transcriptome complexity in a genome-reduced bacterium*. Science 2009, 326:1268-1271.
 29. Miyakoshi M, Nishida H, Shintani M, Yamane H, Nojiri H. *High-resolution mapping of plasmid transcriptomes in different host bacteria*. BMC Genomics 2009, 10:12.
 30. Alonso S, Bartolome-Martín D, del Alamo M, Diaz E, Garcia JL, Pérera J. *Genetic characterization of the styrene lower catabolic pathway of Pseudomonas sp. strain Y2*. Gene 2003, 319:71-83.
 31. Lambertsen L, Molin S, Kroer N, Thomas CM. *Transcriptional regulation of pWW0 transfer genes in Pseudomonas putida KT2440*. Plasmid 2004, 52:169-181.
 32. Moreno R, Marzi S, Romby P, Rojo F. *The Crc global regulator binds to an unpaired A-rich motif at the Pseudomonas putida alkS mRNA coding sequence and inhibits translation initiation*. Nucleic Acids Res 2009, 37:7678-7690.
 33. Sonnleitner E, Abdou L, Haas D. *Small RNA as global regulator of carbon catabolite repression in Pseudomonas aeruginosa*. Proc Natl Acad Sci U S A 2009, 106:21866-21871.
 34. Gerhardt P, Murray RGE, Costilow RN, Nester EW, Wood WA, Krieg NR, Briggs Phillips G (eds.). *Manual of methods for general bacteriology*. Washington, D.C.: American Society for Microbiology; 1981.
 35. Baumann B, Snozzi M, Zehnder AJB, van der Meer JR. *Dynamics of denitrification activity of Paracoccus denitrificans in continuous culture during aerobic-anaerobic changes*. J Bacteriol 1996, 178:4367-4374.
 36. Rouillard JM, Zuker M, Gulari E. *OligoArray 2.0: design of oligonucleotide probes for DNA microarrays using a thermodynamic*

- approach*. Nucleic Acids Res 2003, 31:3057-3062.
37. Shaner NC, Campbell RE, Steinbach PA, Giepmans BN, Palmer AE, Tsien RY. *Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein*. Nat Biotechnol 2004, 22:1567-1572.
 38. Charbonnier Y, Gettler B, Francois P, Bento M, Renzoni A, Vaudaux P, Schlegel W, Schrenzel J. *A generic approach for the design of whole-genome oligoarrays, validated for genotyping, deletion mapping and gene expression analysis on Staphylococcus aureus*. BMC Genomics 2005, 6:95.
 39. Bolstad BM, Irizarry RA, Astrand M, Speed TP. *A comparison of normalization methods for high density oligonucleotide arrays based on bias and variance*. Bioinformatics 2003, 19:185-193.
 40. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP. *Exploration, normalization, and summaries of high density oligonucleotide array probe level data*. Biostatistics 2003, 4:249-264.

SUPPLEMENTARY INFORMATION

Table S1: Location of ORFs in the ICE*c/c* core region and bioinformatic predictions of transcription features.

| ICE <i>c/c</i> gene name | Gene name, product and % amino acid identity to ICEHin1056 | | | Location on ICE <i>c/c</i> | Transcript | Predicted -10 box | Predicted terminator | 5'RACE |
|--------------------------|--|----------------|----|----------------------------|------------|-------------------|----------------------|--------|
| (attR) | | | | 60 - 77 | | | | |
| intB13 | | | | 262 – 2235 | | 148 | 1680 | |
| ORF50240 | (putative relaxase) | | | 52087 – 50240 | I | 52361 | | |
| ORF52324 | | | | 52324 - 52710 | II | 51218 | 52556 | |
| ORF52710 | | | | 52710 - 53168 | | 52365 | | |
| ORF53196 | | | | 53196 - 53573 | | | | |
| ORF53587 | <i>tfc19</i> | TraG | 34 | 53587 - 55104 | III | | | |
| ORF55120 | | | | 55120 - 55479 | | | | |
| ORF55476 | <i>tfc22</i> | PilT | 41 | 55476 - 56873 | | | | |
| ORF56883 | <i>tfc23</i> | | 49 | 56883 - 57830 | | | | |
| ORF57827 | <i>tfc24</i> | | 28 | 57827 - 58273 | | | | |
| ORF58432 | | | | 58432 - 58926 | | 58771 | 58471 | |
| ORF59110 | | | | 59110 - 59874 | IV | | | |
| ORF59888 | <i>tfc16</i> | TraC/ VirB4 | 48 | 59888 - 62755 | | 63191 | 62585 | |
| ORF62755 | <i>tfc15</i> | | 48 | 62755 - 63195 | | 63855 | | |
| ORF63176 | <i>tfc14</i> | TraB | 38 | 63176 - 64594 | V | | 64152 | |
| ORF64584 | <i>tfc13</i> | | 45 | 64584 - 65516 | | | | |
| ORF65513 | <i>tfc12</i> | | 43 | 65513 – 66205 | | | | |
| ORF66202 | | | | 66202 – 66612 | | | | |
| ORF66625 | <i>tfc10</i> | | 28 | 66625 - 66984 | VI | 66976 | 66902 | |

Chapter 2

| | | | | | | | |
|----------|-------------|------|----|---------------|------|-------|--------|
| ORF67001 | | | | 67001 - 67234 | | | |
| ORF67231 | <i>tfc9</i> | | 33 | 67231 - 67614 | | 67610 | 67501 |
| ORF67800 | | | | 67800 - 68204 | VII | 67817 | |
| ORF68241 | <i>tfc8</i> | | 39 | 68241 - 68990 | VIII | | |
| ORF68987 | <i>tfc6</i> | TraD | 59 | 68987 - 71173 | | 70182 | 70286 |
| ORF71178 | <i>tfc5</i> | | 41 | 71178 - 71726 | | 70851 | |
| ORF71723 | | | | 71723 - 72313 | | 72183 | |
| ORF72295 | <i>tfc3</i> | | 39 | 72295 - 73014 | | 72826 | |
| ORF73029 | <i>tfc2</i> | PilL | 44 | 73029 - 73679 | | | |
| ORF73676 | | | | 73676 - 74296 | | 74408 | 73693 |
| ORF74436 | | | | 74436 - 75305 | | | 74710 |
| ORF75419 | | | | 75419 - 77698 | | 75373 | 76133 |
| ORF77798 | | | | 77798 - 78907 | | 78621 | 77243 |
| ORF78972 | | | | 78972 - 79622 | | | |
| ORF79699 | | | | 79699 - 79959 | | | |
| ORF79976 | | | | 79976 - 80383 | | 80141 | 80249 |
| ORF80480 | | | | 80480 - 80812 | | 80490 | |
| ORF80908 | | | | 80908 - 81597 | | 80811 | 81030 |
| ORF81655 | | | | 81655 - 82572 | | 81326 | ~82822 |
| ORF83350 | | | | 83350 - 84192 | IX | 81845 | 83783 |
| ORF84338 | | | | 84338 - 84691 | | 84712 | 84674 |
| ORF84835 | | | | 84835 - 85647 | | | 84825 |
| ORF85934 | | | | 85934 - 86212 | X | | 85850 |
| ORF86310 | | | | 86310 - 87047 | | 87066 | |
| ORF87127 | | | | 87127 - 87939 | | | |
| ORF87986 | | | | 87986 - 88378 | | | 88085 |
| ORF88400 | | | | 88400 - 88612 | | 89109 | |
| ORF89247 | | | | 89247 - 89501 | XI | 89491 | |
| ORF89746 | | | | 89746 - 91347 | | 90548 | 91107 |

Chapter 2

| | | | | | | |
|-------------|--|-----------------|------|----------------|--------|--------|
| ORF91884 | (putative DNA topoisomerase) | 91884 - 93896 | XII | 93515 | 92957 | |
| ORF94175 | (putative single stranded DNA binding protein) | 94175 - 94615 | | 93941 | 94002 | |
| <i>inrR</i> | | 94689 - 95216 | | 94614 | | |
| ORF95213 | | 95213 - 95992 | | 96024 | 95862 | 96017 |
| ORF96323 | | 96323 - 97567 | XIII | 97420 | 97156 | |
| ORF97571 | | 97571 - 98131 | | 99656 | | |
| ORF98147 | | 98147 - 99799 | | 100105 | | |
| ORF99792 | | 99792 - 100049 | | 100552 | | |
| ORF100033 | | 100033 - 100908 | | 101378 | 100199 | |
| ORF100952 | | 100952 - 101164 | XIV | | | |
| ORF101284 | | 101284 - 102039 | XV | 102283 | | 102270 |
| <i>attL</i> | | 102826 - 102843 | | 102757 (Pcirc) | | |

Table S2: Primers used in this study

| primer | sequence | location and position |
|--|---------------------------|-----------------------------|
| <u>Primers used for reverse transcription:</u> | | |
| 53,360rv | 5'-GACCAGCTTCACTACCCAT | reverse primer in ORF53196 |
| 54,608rv | 5'-GCATCGATCTCCATACGCA | reverse primer in ORF53587 |
| 56,105rv | 5'-AGAGGCTGCTACAACCTGGT | reverse primer in ORF55476 |
| 57,883rv | 5'-GCTCGCCATAGACCACGTA | reverse primer in ORF57827 |
| 58,625rv | 5'-ACACGCTGTACGACGACAC | reverse primer in ORF58432 |
| 59,421rv | 5'-GTGTGTGAGCATAGACCCA | reverse primer in ORF59110 |
| 62,351rv | 5'-GTAGAGCTGAAGCACCCAG | reverse primer in ORF59888 |
| 65,306rv | 5'-TCTGCACCCGTAGGCGTT | reverse primer in ORF64584 |
| 65,514rv | 5'-CATGGGGTGCTTCCTCCTT | reverse primer in ORF65513 |
| 67,009rv | 5'-GGAGGAAGAAACTCAGGAC | reverse primer in ORF67001 |
| 68,203rv | 5'-CAGACCTTCAACCGTGCGT | reverse primer in ORF67800 |
| 70,762rv | 5'-ACTCCAGCCGTTCCCTCCA | reverse primer in ORF68987 |
| 72,557rv | 5'-TTGACGAACACCGCCATAC | reverse primer in ORF72295 |
| 73,362rv | 5'-GTCGTGCGCAGCGGCATA | reverse primer in ORF73029 |
| 75,187rv | 5'-CACATGCGTTCCAGCCGA | reverse primer in ORF74436 |
| 77,206rv | 5'-GCTTGATGAGCTTGACCAG | reverse primer in ORF75419 |
| 79,441rv | 5'-TCCAGTCCCTGCTCGTCCG | reverse primer in ORF78972 |
| 81,445rv | 5'-CAGCTCCGCGTCCCTTGAC | reverse primer in ORF80908 |
| 83,358rv | 5'-CGCAGAAGCAGGCAATGG | reverse primer in ORF83350 |
| 84,350rv | 5'-CACGACCTCCAACGACAAG | reverse primer in ORF84338 |
| 85,456rv | 5'-GAGCTTGACGACGATGGCA | reverse primer in ORF84835 |
| 86,777rv | 5'-CACCAACTACGCCTTCGC | reverse primer in ORF86310 |
| 87,711rv | 5'-GATAGCTTTCATCCAATGCG | reverse primer in ORF87127 |
| 88,454rv | 5'-GGCATAGCTGCGGAAGTA | reverse primer in ORF88400 |
| 89,417rv | 5'-TGAAGCACATCGTTTCCTC | reverse primer in ORF89247 |
| 91,054rv | 5'-GGCTACGCTGAACGACTG | reverse primer in ORF89746 |
| 95,129rv | 5'-CTTTTAGAGAGGCCGCGTGTCC | reverse primer in inrR |
| 97,539rv | 5'-GCTCGTGGTGCTGAGTCA | reverse primer in ORF96323 |
| 99,233rv | 5'-GCTGATGTGCGACTGCTGTA | reverse primer in ORF98147 |
| 99,896rv | 5'-CGGTCGAGCTGGTGTAGAT | reverse primer in ORF99792 |
| 100,843rv | 5'-GAGGTTGGCTGCGGTGGT | reverse primer in ORF100033 |
| 101,054rv | 5'-GGCACGCATCAGGTTGTAG | reverse primer in ORF100952 |
| <u>Primers used for PCR:</u> | | |
| 52,745fw | 5'-CTGTGTGATCGAGCAGTTG | forward primer in ORF52710 |
| 53,233rv | 5'-AGGCGTTCTGCGGTGCTG | reverse primer in ORF53196 |
| 54,744rv | 5'-TGCTGAAGGACTGCGACC | reverse primer in ORF53587 |
| 55,360fw | 5'-CATCTACTTGGTCGGCAGT | forward primer in ORF55120 |
| 56,564rv | 5'-AGCCGTTCTGTGATGCCGT | reverse primer in ORF55476 |
| 57,031fw | 5'-CCGTCCTGTCATCGTCCT | forward primer in ORF56883 |
| 58,072rv | 5'-GCTCCGCTTCAATCCGATG | reverse primer in ORF57827 |
| 58,488fw | 5'-CACATCGTCGTCGGTCTG | forward primer in ORF58432 |
| 58,815rv | 5'-TGAAGATCCGTCTCTCCAG | reverse primer in ORF58432 |
| 59,266fw | 5'-CCTGCTTGATCGCCAGAC | forward primer in ORF59110 |
| 59,634rv | 5'-CATAGAGCGTCAGCGTGAA | reverse primer in ORF59110 |
| 60,059fw | 5'-TGTGCCGCCAAGCCTCTA | forward primer in ORF59888 |
| 62,589rv | 5'-GCCTGCTCGTCTGCCAAG | reverse primer in ORF59888 |
| 62,902rv | 5'-CTGGAAGTGGCGATAGACC | reverse primer in ORF62755 |
| 62,920fw | 5'-GGTCTATCGCCAGTTCCAG | forward primer in ORF62755 |
| 63,393fw | 5'-CCAGCAACGAAGCCGTAG | forward primer in ORF63176 |

| | | |
|-----------|---------------------------------|---|
| 65,449rv | 5'-GCGGAGGACATGACCACT | reverse primer in ORF64584 |
| 65,829fw | 5'-GCG GCA TCT ACG AAA TCC C | forward primer in ORF65513 |
| 66,433rv | 5'-GATCGTGCGGAACACCCA | reverse primer in ORF66202 |
| 66,803fw | 5'-TGGCATTGCTGGTTGTGGCGTC | forward primer in ORF66625 |
| 67,051rv | 5'-TGCAGACGCCGAGGAACT | reverse primer in ORF67001 |
| 67,532fw | 5'-CGATGGCGTTGGGCGAGT | forward primer in ORF67231 |
| 67,890fw | 5'-GACAGTCTCGTTGAGTGGT | forward primer in ORF67800 |
| 68,131rv | 5'-CTGATCAGTTCGGTCATCTC | reverse primer in ORF67800 |
| 70,921rv | 5'-CGTCCTTGCTGGTCATCAC | reverse primer in ORF68987 |
| 71,258fw | 5'-GTGTCCGGTGACGACTTG | forward primer in ORF71178 |
| 72,729rv | 5'-AGCGCCGCCGTTCTTCAT | reverse primer in ORF72295 |
| 73,079fw | 5'-ACCGCCGTGTTCCAGTCC | forward primer in ORF73029 |
| 73,558rv | 5'-TGGTCGTTTCATGGTCTGGT | reverse primer in ORF73029 |
| 73,850fw | 5'-GCTGCTCACACTGGCTGG | forward primer in ORF73676 |
| 74,073rv | 5'-AGCGTGTAGCGACCGTAG | reverse primer in ORF73676 |
| 74,480fw | 5'-CAACTGGCACAGGTTCGATC | forward primer in ORF74436 |
| 75,257rv | 5'-TCAGGCGAGCATTGACGGT | reverse primer in ORF74436 |
| 75,561fw | 5'-TGACCTGCCTTGAAGTATG | forward primer in ORF75419 |
| 77,638rv | 5'-CGAGCAGTTCGCCCTGTA | reverse primer in ORF75419 |
| 77,974fw | 5'-CTCGTCGTCAAAGGTGACA | forward primer in ORF77798 |
| 79,441rv | 5'-TCCAGTCCCTGCTCGTCG | reverse primer in ORF78972 |
| 79,742fw | 5'-CACTTCACTTCAACCCGCA | forward primer in ORF79699 |
| 81,445rv | 5'-CAGCTCCGCGTCTTGAC | reverse primer in ORF80908 |
| 81,726rv | 5'-CAGATTGGTCGGCAGCGAT | reverse primer in ORF81655 |
| 81,825fw | 5'-GTCTATCCGACGAGCCAG | forward primer in ORF81655 |
| 82,048fw | 5'-TCGGTACTTCGACACCATC | forward primer in ORF81655 |
| 82,396rv | 5'-GTCGAAGTAGCGGTAAGT | reverse primer in ORF81655 |
| 83,367fw | 5'-TGCTTCTGCGGCGACTGA | forward primer in ORF83350 |
| 83,834rv | 5'-CGAATCTTCAGGCTGGTC | reverse primer in ORF83350 |
| 84,368fw | 5'-CTTGTCGTTGGAGTCTGTG | forward primer in ORF84338 |
| 84,409rv | 5'-GCAAGATGCGGGTGAATAC | reverse primer in ORF84338 |
| 85,145fw | 5'-CAGGTGCACGGGATTTAC | forward primer in ORF84835 |
| 85,560rv | 5'-ACACGCGGTACTTGAACGA | reverse primer in ORF84835 |
| 86,049fw | 5'-GCGACATTGACGACTACGAA | forward primer in ORF85934 |
| 86,842rv | 5'-ACAGGTTCAAGTCCATTCC | reverse primer in ORF86310 |
| 87,213fw | 5'-GGCATCAACGACGAGCATT | forward primer in ORF87127 |
| 87,827rv | 5'-GAGAAGTGAGCGGAGCCA | reverse primer in ORF87127 |
| 88,241fw | 5'-CACCCGTGAAGGTCTGAAC | forward primer in ORF87986 |
| 88,595rv | 5'-AGAGGCTTGCTGGGACAT | reverse primer in ORF88400 |
| 89,435fw | 5'-GAGGAAACGATGTGCTTCA | forward primer in ORF89247 |
| 89,476rv | 5'-CGGTCTTGCTTGGTGATGT | reverse primer in ORF89247 |
| 89,850fw | 5'-CGCTACAAGGCAATCGGC | forward primer in ORF89746 |
| 91,170rv | 5'-GGTAACGATGGGCGAGCA | reverse primer in ORF89746 |
| 91,991fw | 5'-TGGTCGTGCAGTCGTTAC | forward primer in ORF91884 |
| 95,156rv | 5'-GCATGAAGGCTTCTGTTGCAGG | reverse primer in inrR |
| 95,677rv | 5'-CCAGGCTCAATGCTGCGGGTA | reverse primer in ORF95213 |
| 95,790rv | 5'-TCGCTGTACGGGTCGTCTTGCTC | reverse primer in ORF95213 |
| 99,484rv | 5'-TCTCCGACCAGAGTTCACGC | reverse primer in ORF98147 |
| 99,896rv | 5'-CGGTCGAGCTGGTGTAGAT | reverse primer in ORF99792 |
| 99,914fw | 5'-ATCTACACCAGCTCGACCG | forward primer in ORF99792 |
| 99,965rv | 5'-TGACGTGCCAGCCTTCGC | reverse primer in ORF99792 |
| 100,162fw | 5'-AGCCGCTCTCGACACCATG | forward primer in ORF100033 |
| 100,501rv | 5'-GGCCATCTCCAGCAAGAC | reverse primer in ORF100033 |
| 100,883rv | 5'-GTCGAGATGATGGATACAAC | reverse primer in ORF100033 |
| 101,142rv | 5'-GGTAGTACAGGTATCTGCGA | reverse primer in ORF100952 |
| 101,161fw | 5'-TCGCAGATACCTGTACTACC | forward primer in ORF100952 |
| 101,169rv | 5'-TTTTGGATCCGTCACCTCTGAACGCTCA | reverse primer in intergenic region between ORF100952 and ORF101284 |

Chapter 2

| | | |
|-----------|----------------------------------|---|
| 101,277fw | 5'-TTTTTCTGCAGCAAGCCCGAATCTGCGGT | forward primer in intergenic region between ORF100952 and ORF101284 |
| 101,403fw | 5'-CAGGAGGCTATCGCTACC | forward primer in ORF101284 |

Table S3. Description of RNA probes for Dot Blot hybridization

| Probe Name | Localisation | Size | Primer Name and Sequence |
|-------------|-------------------|------|---|
| integrase | 1,887 - 2,173 | 286 | 2,173rv-T7P: 5'-CTAATACGACTCACTATAGGGAGA / 1,887fw: |
| ORF52710 | 52,745 - 53,082 | 337 | 53,082rv-T7P: 5'-CTAATACGACTCACTATAGGGAGA / 52,745fw: |
| ORF53587 | 54,608 - 54,840 | 232 | 54,608rv-T7P: 5'-CTAATACGACTCACTATAGGGAGA / 54,840fw |
| ORF59888 | 61,503 - 61,820 | 317 | 61,503rv-T7P: 5'-CTAATACGACTCACTATAGGGAGA / 61,820fw |
| ORF65513 | 65,514 - 65,829 | 315 | 65,514rv-T7P: 5'-CTAATACGACTCACTATAGGGAGA / 65,829fw |
| ORF67800 | 67,842 - 68,203 | 361 | 68,203rv-T7P: 5'-CTAATACGACTCACTATAGGGAGA / 67,842fw |
| ORF68987 | 70,487 - 70,944 | 457 | 70,487rv-T7P: 5'-CTAATACGACTCACTATAGGGAGA / 70,944fw |
| ORF73029 | 73,362 - 73,640 | 278 | 73,362rv-T7P: 5'-CTAATACGACTCACTATAGGGAGA / 73,640fw |
| ORF75419 | 75,961 - 76,272 | 311 | 75,961rv-T7P: 5'-CTAATACGACTCACTATAGGGAGA / 76,272fw |
| ORF81655 | 81,726 - 82,048 | 322 | 81,726rv-T7P: 5'-CTAATACGACTCACTATAGGGAGA / 82,048fw |
| ORF83350 | 83,358 - 83,717 | 359 | 83,358rv-T7P: 5'-CTAATACGACTCACTATAGGGAGA / 83,717fw |
| ORF84835 | 84,867 - 85,145 | 278 | 84,867rv-T7P: 5'-CTAATACGACTCACTATAGGGAGA / 85,145fw |
| ORF87986 | 88,033 - 88,241 | 208 | 88,033rv-T7P: 5'-CTAATACGACTCACTATAGGGAGA / 88,241fw |
| ORF89746 | 90,898 - 91,187 | 289 | 90,898rv-T7P: 5'-CTAATACGACTCACTATAGGGAGA / 91,187fw |
| ORF91884 | 92,115 - 92,675 | 560 | 92,115rv-T7P: 5'-CTAATACGACTCACTATAGGGAGA / 92,675fw |
| <i>inrR</i> | 94,886 - 95,204 | 318 | 94,886rv-T7P: 5'-CTAATACGACTCACTATAGGGAGA / 95,204fw |
| ORF96323 | 97,097 - 97,556 | 459 | 97,097rv-T7P: 5'-CTAATACGACTCACTATAGGGAGA / 97,556fw |
| ORF98147 | 98,203 - 98,528 | 325 | 98,203rv-T7P: 5'-CTAATACGACTCACTATAGGGAGA / 98,528fw |
| ORF100033 | 100,036 - 100,162 | 126 | 100,036rv-T7P: 5'-CTAATACGACTCACTATAGGGAGA / 100,162fw |
| ORF100952 | 100,952 - 101'161 | 209 | 100,952rv-T7P: 5'-CTAATACGACTCACTATAGGGAGA / 101,161fw |
| ORF101284 | 101'444 - 101'788 | 344 | 101,444rv-T7P: 5'-CTAATACGACTCACTATAGGGAGA / 101,788fw: |

CHAPTER 3

ICE_{CLC} BEHAVIOUR IN DIFFERENT HOST BACTERIA

ABSTRACT

The integrative and conjugative element ICE*clc* is a mobile genome that (i) can self-transfer horizontally to several bacterial species and (ii) confers the ability to use 3-chlorobenzoate (3CBA) or 2-aminophenol (2AP) as sole carbon and energy sources. In this study, we examine ICE*clc*-specific gene expression in four different host bacteria during exponential growth on 3CBA and subsequent stationary phase. The host strains with integrated ICE*clc* included *Pseudomonas knackmussii* B13, *Pseudomonas putida* UWC1, *Pseudomonas aeruginosa* PAO1 and *Cupriavidus necator* JMP289. Results from micro-array hybridizations demonstrate that the host has an impact on (i) the expression level of ICE*clc* located genes and (ii) on the timing of expression, notably of the genes for catabolism of 3CBA and 2AP, and of the genes implicated in ICE transfer. By comparing ICE*clc* expression in *P. aeruginosa* PAO1 during exponential growth and stationary phase on succinate with that on 3CBA, we further reveal that the carbon source has an impact on the behavior of ICE*clc*.

INTRODUCTION

ICE*clc* can be transferred from *Pseudomonas knackmussii* B13 to members of both β - and γ -*proteobacteria*, such as *Pseudomonas putida*, *Pseudomonas aeruginosa*, *Cupriavidus necator* or *Ralstonia* [1–3]. Because of the spectrum of different recipient cells for ICE*clc*, we would like to understand whether ICE*clc* expression is host-dependent and, if so, which host or ICE*clc* factors control this. As a first step to study this we study ICE*clc* global expression in three different hosts other than strain B13, under exponential growth with 3-chlorobenzoate (3CBA) as sole carbon and energy source, and during subsequent stationary phase conditions.

P. aeruginosa PAO1 and other environmental isolates are hosts for ICE*clc* [4]. Moreover, various isolates of *P. aeruginosa* carry GEI related to ICE*clc* [5]. Interestingly, however, ICE*clc* has been shown to behave differently among *P. aeruginosa* isolates and compared to strain B13. For example, ICE*clc* transfers with 100- to 1000-fold lower frequencies from PAO1 or other environmental isolates as donors than from B13 [4]. In contrast to B13, PAO1 with ICE*clc* could grow both on 2AP or 3CBA as sole carbon and energy source, whereas the original host B13 can only use 3CBA.

C. necator JMP289 is a β -*Proteobacterium* of the Burkholderiales family. It has been shown to be a suitable host for ICE*clc*, which confers on it the ability to catabolize 3CBA but not 2AP [1]. *C. necator* ICE*clc* is also itself again a donor and transfers the element with a frequency of 10^{-3} per donor. *C. necator* is so far the more phylogenetically distant host known of ICE*clc*. Interestingly,

an element almost perfectly identical to ICE*clc* occur in *Burkholderia xenovorans*, but its transfer rates have not been studied (Chain, 2006).

Also *P. putida* UWC1 has been used as a host to study ICE*clc* behaviour and genetics. *P. putida* UWC1 derivatives have been isolated which carry one or two ICE*clc* copies, in each of four different chromosomal insertion sites [1]. *P. putida* UWC1 (ICE*clc*) can use both 2AP and 3CBA as sole carbon and energy sources, and it can transfer ICE*clc* to other hosts with a frequency of transfer comparable to B13 (10^{-2} per donor, [6]). Finally, excellent tools for genetic and genomic manipulation are available for *P. putida* UWC1.

Here we study the ICE*clc* specific gene expression in four host species carrying ICE*clc*, in order to reveal whether host-specific differences exist that may help to explain the different ICE*clc* behavior. The ICE*clc* transcriptome was captured in cells growing exponentially on 3CBA as sole carbon and energy sources as well as in the subsequent stationary phase. In addition, we use *P. aeruginosa* PAO1 (ICE*clc*) to further study the effect of carbon source on ICE*clc* specific gene expression. We find indeed different host-dependent expression levels, and demonstrate that the timing of ICE*clc* expression is host-dependent. Moreover, we show that growth on 3CBA leads to a much stronger ICE*clc* gene expression than succinate.

RESULTS

Global ICE*clc* expression in different host bacteria. All ICE*clc* host species investigated here contain one copy of ICE*clc*, but for all of them transconjugants can be obtained that have two ICE*clc* copies integrated at different insertion positions (not shown). ICE*clc* global expression was determined in three hosts and compared to that of *P. knackmussii* B13 established earlier [5]. The three hosts were: *P. aeruginosa* PAO1, *P. putida* UWC1 and *C. necator* JMP289, both in exponentially growing cells and in cells taken from stationary phase.

ICE*clc* expression in *P. aeruginosa* PAO1. Thirty-two genes were statistically significantly differentially expressed when comparing exponential and stationary phase transcriptomes of PAO1-ICE*clc* cells grown on 10 mM 3CBA. This amount is much lower than the 103 genes of ICE*clc* expressed differentially in B13 [7], and it thus seemed as if overall gene expression from ICE*clc* in PAO1 was more silent than in B13. This we deduced from absolute signal intensities on microarray (Figure 1). Secondly, the differences between exponential and stationary phase were less pronounced in PAO1 than in B13. Therefore, the difference between ICE*clc* gene expression in stationary and exponential phase was often no longer statistically significant. As far as the catabolic genes are concerned, both the *clc* and the *amn* genes are expressed in PAO1 during exponential and stationary phase (*amn* genes on Figure 1B, 1D, *clc* genes on Figure 1A, 1C), whereas in B13 they are primarily expressed during exponential growth. Also the strong signal present in B13 on

the antisense strand of the *c/c* genes in stationary phase appeared less clearly in PAO1 (Figure 1D).

Interestingly, almost all the genes in the core module (plus the *intB13* gene) are expressed to a much lower level in PAO1 stationary phase cells than in B13 (Figure 1C, 2^{12} compared to 2^{16} in B13). On the other hand, the fold-change difference in expression of *intB13* between exponential and stationary phase in this region remains similar in both PAO1 and B13 (i.e. eightfold induction during stationary phase). Similar as in B13, expression of *orf81655* in PAO1 is one of the highest among all ICE*c/c* genes and also displays the highest fold difference between the two phases.

ICE*c/c* expression analysis of PAO1-ICE*c/c* cells growing on 10 mM 3CBA compared to 15 mM succinate as sole carbon and energy source, showed that cells growing on succinate still express the *c/c* operon, although to a much weaker extent than when exposed to 3CBA (signal of 2^{10} on succinate compared to 2^{14} on 3CBA, Figure 2A). Growth on succinate did not lead to activation of the *intB13* gene and the core module genes in stationary phase, (Figure 2C, 2D).

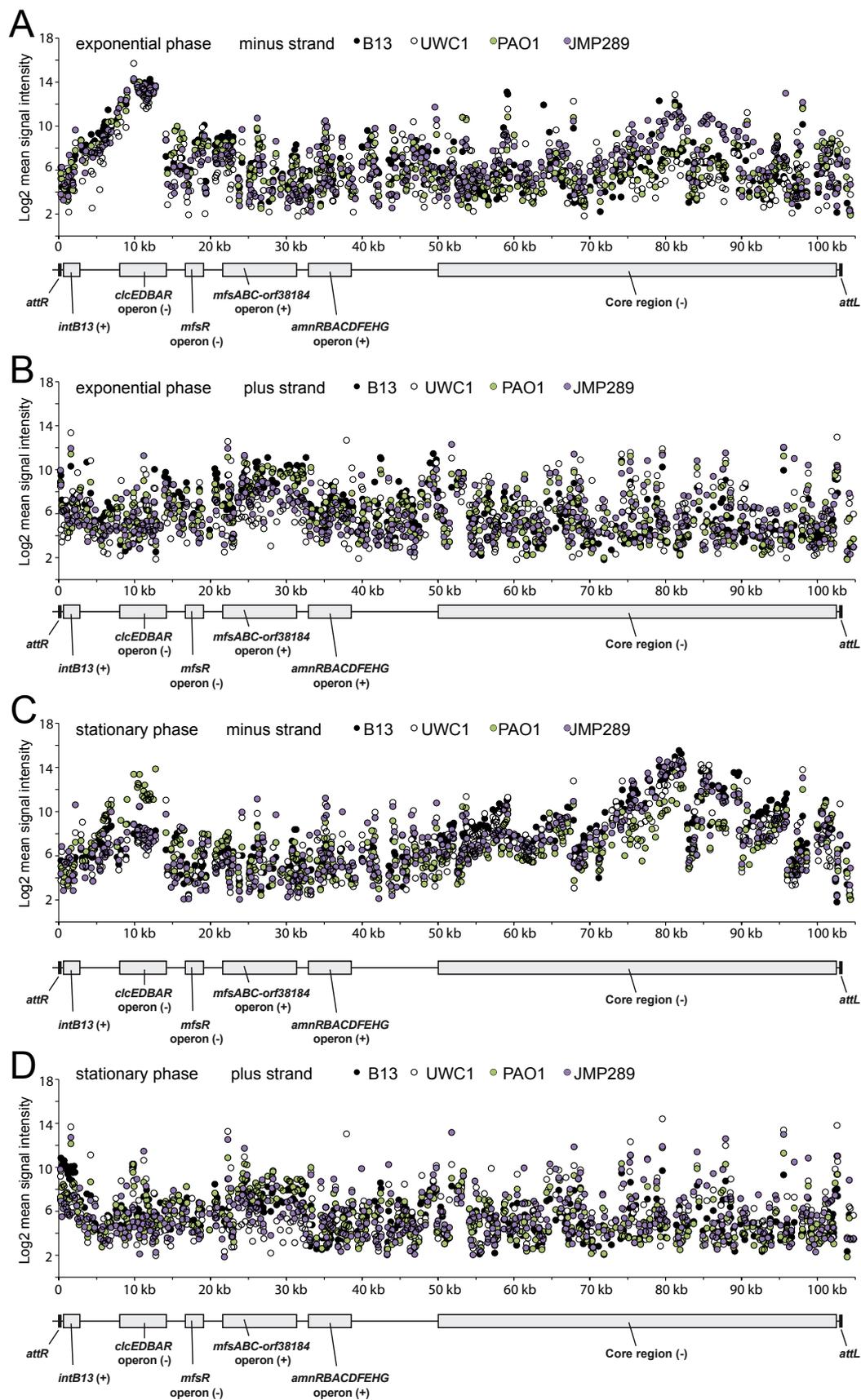


Figure 1. Growth phase-dependent expression of ICE*clc* in all investigated hosts grown on 10 mM 3CBA during exponential and stationary phase. A, transcriptomic profile of the "minus"

strand in exponential phase. B, transcriptomic profile of the "plus" strand in exponential phase. C, transcriptomic profile of the "minus" strand in stationary phase. D, transcriptomic profile of the "plus" strand in stationary phase. Probe signal intensities are in \log_2 scale (y axis) plotted over the length of the *ICEclc* sequence (x axis, in bp). The linear map of the *ICEclc* is aligned below. B13-*ICEclc* (black dots), UWC1-*ICEclc* (white dots), PAO1-*ICEclc* (green dots), JMP289-*ICEclc* (purple dots).

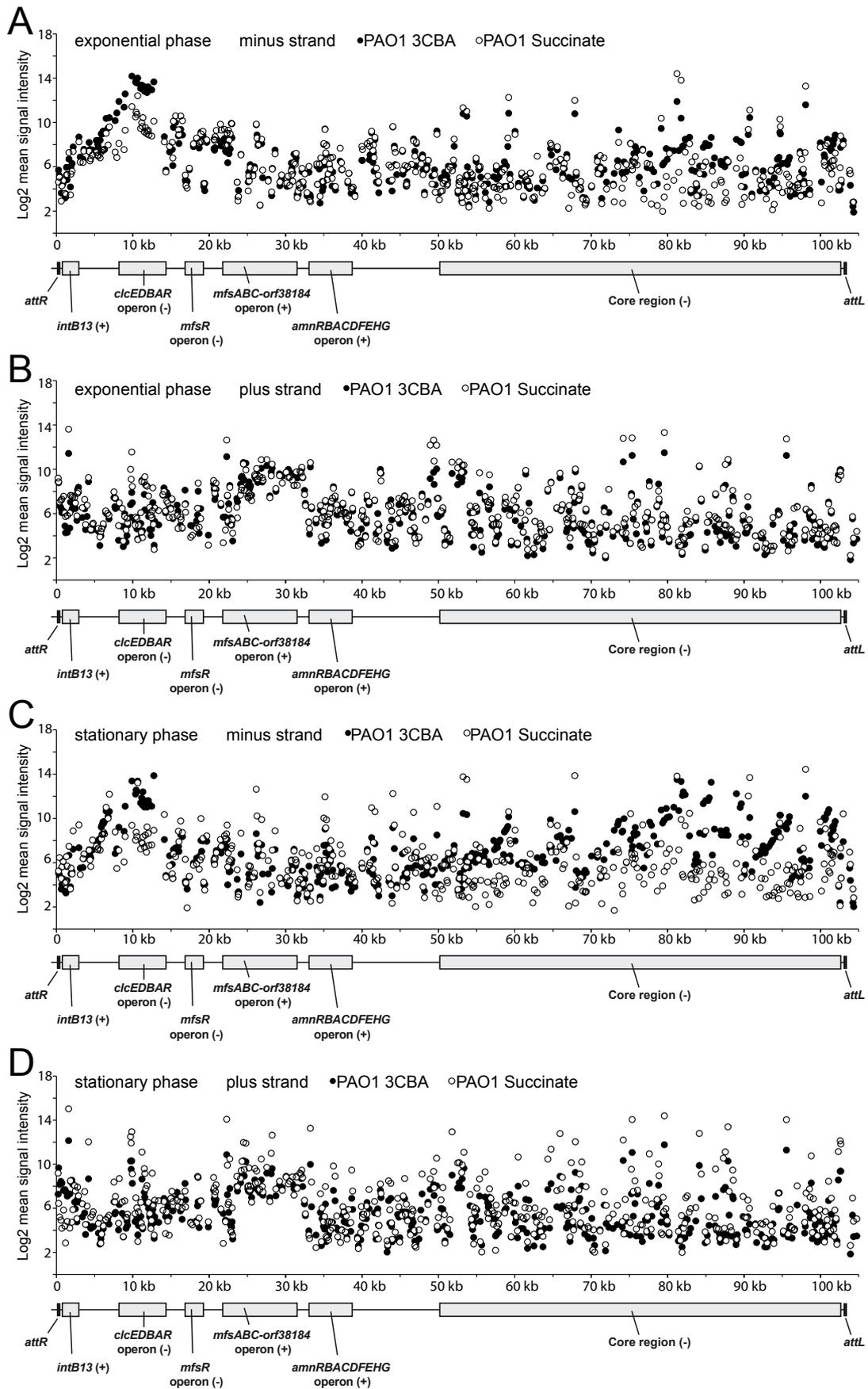


Figure 2. Growth phase-dependent expression of ICE*clc* in PAO1-ICE*clc* grown either on 10 mM 3CBA or on 15 mM succinate. A, transcriptomic profile of the "minus" strand in exponential phase. B, transcriptomic profile of the "plus" strand in exponential phase. C, transcriptomic profile of the "minus" strand in stationary phase. D, transcriptomic profile of the "plus" strand in stationary phase. Probe signal intensities are in \log_2 scale (y axis) plotted over the length of the ICE*clc* sequence (x axis, in bp). The linear map of the ICE*clc* is aligned below. 10 mM 3CBA (black dots), 15 mM succinate (white dots).

ICE*clc* expression in *C. necator* JMP289. ICE*clc* transcriptome analysis of *C. necator* JMP289 (ICE*clc*) first of all shows that the *clc* genes are expressed to similar levels in *C. necator* as in B13 during exponential phase, but that some other genes in the 'catabolic' part of ICE*clc* (e.g., the *amn* genes) are expressed to a lower level (Figure 1B). Interestingly, genes in the core module are expressed to a higher level during exponential phase in *C. necator* than in the three other strains (Figure 1C).

Overall 54 genes of ICE*clc* were statistically significantly differentially expressed between exponential and stationary phase in *C. necator*, which is about half the number as occurred in B13, but more than for PAO1. For example, the *intB13* gene is much less expressed in JMP289 than in B13 in stationary phase (2^8 compared to 2^{11} in B13 stationary phase, Figure 1D). Actually, most of the transcripts in the core module are not statistically differentially expressed in *C. necator* between exponential and stationary phase, but this is due to the surprisingly high expression taking place already in exponential phase (up to 2^{12} compared to 2^8 in the corresponding probes in B13 and PAO1, Figure 1A). One could conclude from this that the expression of ICE*clc* in exponential phase is not as tightly controlled in JMP289 as in B13.

ICE*clc* expression in *P. putida* UWC1. Overall, signal intensities and probe hybridization patterns of ICE*clc* gene expression in stationary phase were very similar between UWC1 and B13. Exceptions include *intB13*, the expression of which in UWC1 was more reminiscent of that of PAO1 and JMP289 (all three strains with intensities of 2^8 compared to the 2^{11} of B13, Figure 1D), or *orf86310* and *orf75419* (putative helicase), which were clearly weaker expressed in UWC1 than in B13. Moreover, the *amn* genes are remarkably weakly expressed in exponential and stationary phase UWC1 cells, even less than in the other strains (Figure 1C, 1D). Finally, it is interesting to mention that expression of ICE*clc* genes in UWC1, PAO1 and JMP289 in several cases (e.g., *clc* genes or *orf81655*) reaches similar levels as in B13, despite them having only one ICE*clc* copy compared to two copies in B13. However, it is also possible that this cannot be measured completely correctly because of probe saturation at the highest expression values (above $\geq 2^{16}$).

DISCUSSION

Microarray hybridizations indicated a number of statistically significant changes of ICE*clc* expression in different hosts. The most striking effect observed was that the clear bipartite behaviour of ICE*clc* along growth phases (i.e. catabolism in exponential phase versus core module in stationary phase) as seen in B13 was present in UWC1, but less so in PAO1 and JMP289. Whether this is the result of failed host control over the ICE is unknown.

Interestingly, microarray data indicated that expression of the catabolic genes on ICE*clc* in PAO1 remains much higher in stationary phase than in the other three strains. This may be the result of the inability of the transcription factors responsible for the control of the catabolic operons (ClcR, AmnR) to exert their action as tightly as in B13 or UWC1. There is precedent for such behavior. For example, the *ant* and *car* operons of plasmid pCAR1 (coding for anthranilate and carbazole degradation, respectively) as well as various putative transporters are differentially expressed depending on the host cell carrying the plasmid [8, 9] It has been suggested that this may have led to differential capacity of the pCAR1 host to metabolize carbazole [10]. Although we see different ICE*clc* *clc* and *amn* expression differences among the hosts we do not currently have information on 3CBA or 2AP degradation rates by those strains. It is interesting that the *amn* operon is differentially expressed in various ICE*clc*-bearing species even when cells are growing on 3CBA. Thus, it could be interesting to measure 2AP degradation rates in 3CBA- or 2AP-growing cultures. The potential importance of cross-activation between *amn*

and *clc* genes for the host could be further tested by producing specific deletions of *amn* or *clc* genes on ICE*clc*.

Expression of the ICE*clc* core genes in stationary phase cells of PAO1-ICE*clc* was clearly weaker than in B13. The consequence of this might be that ICE*clc* transfer rates from PAO1 as donor are lower than from B13, which is in agreement with previous measurements on ICE*clc* transfer frequencies (10^{-4} to 10^{-5} per donor PAO1 compared to 10^{-2} per donor B13 [4]). This behavior is also reminiscent of what has been found for pCAR1 in different host bacteria [8, 9] Indeed, pCAR1 genes for plasmid partition (*parAB*) and conjugation (*trh/tra* genes) are differentially expressed in six different hosts (mainly) among the genus *Pseudomonas*, correlating to pCAR1 transfer frequencies [10]. However, as mentioned by Shintani and coworkers [9], conjugation efficiency is dependent on donor, plasmid and recipient factors, thus making it difficult to conclude that varying transcriptional levels of HGT-associated genes are the sole cause of differing transfer frequencies. For ICE*clc* we know that only a few percent of cells in stationary phase are transfer-competent donors, and are those which actually express ICE*clc* core genes [4]. Therefore, the magnitude of ICE*clc* core gene expression seen in microarray hybridizations of cDNA produced from stationary phase B13 cells (Fig. 1C) is due to only a few percent of transfer-competent donor cells. It could thus very well be that the weaker "global" induction of ICE*clc* core genes in *P. aeruginosa* PAO1-ICE*clc* is due to a lower proportion of P_{int} -expressing cells, and that the lower proportion of cells is the reason for the lower transfer frequencies of ICE*clc* from PAO1 as donor. Unfortunately, this proportion of

P_{int} -expressing cells has not been measured yet for PAO1-ICE clc .

A further interesting observation from the current work was that a number of ICE clc core genes are expressed earlier in exponential phase in JMP289 compared to stationary phase in B13, PAO1 and UWC1. Also here, it would be interesting to use singly copy P_{int} -gfp reporter fusions to test whether JMP289 activates its transfer competence (tc) earlier than B13.

So far the microarray data thus suggested that the host cells may influence the ICE clc genetic program. However, no specific ICE clc or host gene factor has been identified so far, which may be responsible for these gene expression differences. As demonstrated by Miyazaki and coworkers in 2012, RpoS in B13 is instrumental for the expression of the *inrR* gene, a key factor to promote bistability and integrase expression [11]. On the basis of previous work regarding the role of RpoS in strain B13 our primary hypothesis would be that RpoS is different among the four hosts, which then influences ICE clc global expression. Experiments to verify this could include knocking out native *rpoS*-s, which should not be too complicated given that the genome sequences are known, and knocking in *rpoS* from B13 via a mini-transposon insertion. Subsequently, such strains could be re-examined for ICE clc transcription. Additionally, as demonstrated for pCAR1 and various of its bacterial hosts, there may be a complex crosstalk of plasmid and chromosomal factors that influences each other transcriptomes [10, 12, 13]. Thus, it is highly possible that the host-specific difference in ICE clc transcriptome is caused by more than one factor, and that ICE clc itself can

modify the host chromosome transcription, which on its turn interferes with ICE*clc* expression.

MATERIAL AND METHODS

Growth conditions and harvesting. *P. putida* UWC1 (ICE clc), *P. aeruginosa* PAO1 (ICE clc), and *C. necator* JMP289 (ICE clc), were cultivated in minimal medium (MM) based on the type 21C medium [14]. This MM was complemented with 10 mM 3CBA or 15 mM succinate, and the bacteria were grown at 30°C. Bacterial growth was assessed from culture turbidity at 600 nm (OD₆₀₀). Cells were recovered during exponential phase (OD₆₀₀ of 0.4) and 48 hours after the entrance in stationary phase. For RNA isolation, 100 ml of culture was immediately harvested by centrifugation (at 15,000 × *g* for 1 min at 4°C) and the supernatant was decanted. Cell pellets were resuspended in 4 ml RNeasy Protect Bacteria Reagent (QIAGEN GmbH). After 5 min incubation, the suspensions were centrifuged again (at 5,000 × *g* for 5 min at room temperature); the supernatant was discarded and pellets were stored at -80°C.

RNA isolation. Prior to RNA extraction, pellets were slowly thawed, then resuspended in 0.5 ml TES buffer [10 mM Tris-HCl (pH 8.0), 1 mM EDTA, 100 mM NaCl], followed by addition of and mixing with 0.25 ml lysis solution [20 mM sodium acetate (pH 5.5), 1 mM EDTA, 0.5% SDS]. After that, the total RNA was further purified using the hot acid-phenol method as described previously [15]. RNA samples were purified from contaminating DNA by treatment with 50 U of DNase I (RNase free; Roche) during 1 h at 37°C. Finally, the RNA was dissolved in 50 µl diethylpyrocarbonate (DEPC)-treated water and the RNA concentration was quantified by measuring the

absorbances at 260 and 280 nm on a NanoDrop spectrophotometer (Witec AG). The integrity of RNA was determined by agarose gel electrophoresis and the absence of DNA was verified by PCR.

Microarray design. A series of 950 non-overlapping 50-mer probes was designed to cover both coding and non-coding regions of the ICE clc sequence (Acc. No. AJ617740) at approximate distances of 200 bp. Probes were designed using the program OLIGOARRAY version 2.1 [16] with a melting temperature range of 92 to 99°C and a probe GC content range of 52 to 72%. Probes were further designed to not cross-hybridize with gene products from the following potential host strains of the ICE clc element: *Burkholderia xenovorans* LB400 (Acc. No. CP000270-CP000272), *P. putida* F1 (Acc. No. CP000712), *P. putida* KT2440 (Acc. No. AE015451), *P. aeruginosa* PAO1 (Acc. No. AE004091), *Cupriavidus necator* JMP134 (Acc. No. CP000090-CP000093), and *Ralstonia metallidurans* CH34 (Acc. No. CP000352-CP000355). An additional 93 probes were designed to target housekeeping genes from the potential host strains and 8 probes were designed to target positive/negative controls (GFP, luciferase, and mCherry transcripts, [17]). The microarray was manufactured by Agilent Technologies (Santa Clara, CA) in the 8 × 15,000 probe format and each unique probe was synthesized at six randomized spatial locations on the array. The microarray design has been deposited in the NCBI Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo>) under accession number GSE20461.

Microarray hybridization and analysis. Total RNA was isolated and purified from the various ICEc/c hosts cultures during exponential growth on 3CBA and during subsequent stationary phase after 48 h as described above. For microarray hybridizations, cDNA was synthesized from total RNA and directly labeled with cyanine-3-dCTP using a modification of a protocol described elsewhere [18]. Briefly, each 50- μ L reaction contained 10 μ g of total RNA, 1.25 μ g of random hexanucleotide primers (Promega), 100 μ M each of unlabeled dATP, dGTP, and dTTP (Invitrogen), 25 μ M of unlabeled dCTP (Invitrogen), 25 μ M of cyanine-3-labeled dCTP (Perkin-Elmer), 25 U SUPERase•In (Ambion), and 400 U Superscript II reverse transcriptase (Invitrogen). Reactions were performed by heating at 42°C for 2 hours followed by 70°C for 10 min. RNA was then removed by adding 100 mM NaOH, heating to 65°C for 20 min, and neutralizing with 100 mM HCl and 300 mM sodium acetate (pH 5.2). Labeled cDNA products were purified using the MinElute PCR purification kit (Qiagen), and the quantity and incorporation frequency of cyanine-3-labeled dCTP were calculated using the MICROARRAY function on a NanoDrop Spectrophotometer. Sixty ng of labeled cDNA was then loaded onto each microarray, hybridized for 17 hours at 65°C, and washed and scanned as described for labeled cRNA in the One-Color Microarray-Based Gene Expression Analysis Manual (Agilent). The fragmentation step (heating to 60°C for 30 minutes) was omitted.

Hybridization signal intensities were quantified from microarray image scans using AGILENT FEATURE EXTRACTION software version 9.5.3 (Agilent). Microarray data were normalized and globally scaled over the array using GENESPRING GX software with the RMA algorithm and quantile normalization [19, 20]. Mean

probe signals were calculated for each of the three biological replicates and were plotted against their position on the ICEc/c sequence for both strands and for RNAs isolated during exponential and stationary phases. Microarray data of *P. knackmussii* B13 were used from previous experiments [7] and are available in the GEO database (<http://www.ncbi.nlm.nih.gov/geo>) under the accession number GSE20461.

REFERENCES

1. Sentchilo V, Czechowska K, Pradervand N, Minoia M, Miyazaki R, Van Der Meer JR. *Intracellular excision and reintegration dynamics of the ICE_{clc} genomic island of Pseudomonas knackmussii sp. strain B13*. Mol Microbiol 2009, 72:1293–1306.
2. Ravatn R, Zehnder AJB, van der Meer JR. *Low-frequency horizontal transfer of an element containing the chlorocatechol degradation genes from Pseudomonas sp. strain B13 to Pseudomonas putida F1 and to indigenous bacteria in laboratory-scale activated-sludge microcosms*. Appl Environ Microbiol 1998, 64:2126–2132.
3. Springael D, Peys K, Ryngaert A, Roy SV, Hooyberghs L, Ravatn R, Heyndrickx M, Meer J-R van der, Vandecasteele C, Mergeay M. *Community shifts in a seeded 3-chlorobenzoate degrading membrane biofilm reactor: indications for involvement of in situ horizontal transfer of the clc-element from inoculum to contaminant bacteria*. Environ Microbiol 2002, 4:70–80.
4. Gaillard M, Pernet N, Vogne C, Hagenbüchle O, van der Meer JR. *Host and invader impact of transfer of the clc genomic island into Pseudomonas aeruginosa PAO1*. Proc Natl Acad Sci 2008, 105:7058–7063.
5. Klockgether J, Wurdemann D, Wiehlmann L, Tummeler B. *Transcript profiling of the Pseudomonas aeruginosa genomic islands PAGI-2 and pKLC102*. Microbiology 2008, 154:1599–1604.
6. Reinhard F, Miyazaki R, Pradervand N, van der Meer JR. *Cell differentiation to “mating bodies” induced by an integrating and conjugative element in free-living bacteria*. Curr Biol 2013, 23:255–259.
7. Gaillard M, Pradervand N, Minoia M, Sentchilo V, Johnson DR, van der Meer JR. *Transcriptome analysis of the mobile genome ICE_{clc} in Pseudomonas knackmussii B13*. BMC Microbiol 2010, 10:153.
8. Miyakoshi M, Nishida H, Shintani M, Yamane H, Nojiri H. *High-resolution mapping of plasmid transcriptomes in different host bacteria*. BMC Genomics 2009, 10:12.
9. Shintani M, Tokumaru H, Takahashi Y, Miyakoshi M, Yamane H, Nishida H, Nojiri H. *Alterations of RNA maps of IncP-7 plasmid pCAR1 in various Pseudomonas bacteria*. Plasmid 2011, 66:85–92.
10. Nojiri H. *Impact of catabolic plasmids on host cell physiology*. Curr Opin Biotechnol 2013, 24:423–430.

11. Miyazaki R, Minoia M, Pradervand N, Sulser S, Reinhard F, van der Meer JR. *Cellular variability of RpoS expression underlies subpopulation activation of an integrative and conjugative element*. Plos Genet 2012, 8:e1002818.
12. Yun C-S, Suzuki C, Naito K, Takeda T, Takahashi Y, Sai F, Terabayashi T, Miyakoshi M, Shintani M, Nishida H, Yamane H, Nojiri H. *Pmr, a histone-like protein H1 (H-NS) family protein encoded by the IncP-7 plasmid pCAR1, is a key global regulator that alters host function*. J Bacteriol 2010, 192:4720–4731.
13. Miyakoshi M, Shintani M, Inoue K, Terabayashi T, Sai F, Ohkuma M, Nojiri H, Nagata Y, Tsuda M. *Parl, an orphan ParA family protein from Pseudomonas putida KT2440-specific genomic island, interferes with the partition system of IncP-7 plasmids: Plasmid partition inhibitor from P. putida KT2440*. Environ Microbiol 2012, 14:2946–2959.
14. Gerhardt P, Murray R, Costilow R, Nester E, Wood W, Krieg N, Briggs Phillips G. *Manual of methods for general bacteriology*. 1981.
15. Baumann B, Snozzi M, Zehnder AJ, Van Der Meer JR. *Dynamics of denitrification activity of Paracoccus denitrificans in continuous culture during aerobic-anaerobic changes*. J Bacteriol 1996, 178:4367–4374.
16. Rouillard J-M. *OligoArray 2.0: design of oligonucleotide probes for DNA microarrays using a thermodynamic approach*. Nucleic Acids Res 2003, 31:3057–3062.
17. Shaner NC, Campbell RE, Steinbach PA, Giepmans BNG, Palmer AE, Tsien RY. *Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein*. Nat Biotechnol 2004, 22:1567–1572.
18. Charbonnier Y, Gettler B, François P, Bento M, Renzoni A, Vaudaux P, Schlegel W, Schrenzel J. *A generic approach for the design of whole-genome oligoarrays, validated for genotyping, deletion mapping and gene expression analysis on Staphylococcus aureus*. BMC Genomics 2005, 6:95.
19. Bolstad BM, Irizarry RA, Astrand M, Speed TP. *A comparison of normalization methods for high density oligonucleotide array data based on variance and bias*. Bioinformatics 2003, 19:185–193.
20. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP. *Exploration, normalization, and summaries of high density oligonucleotide array probe level data*. Biostatistics 2003, 4:249–264.

CHAPTER 4

CHARACTERIZATION OF A TETR- TYPE REPRESSOR REGULATED MAJOR FACILITATOR SYSTEM ON THE INTEGRATIVE AND CONJUGATIVE ELEMENT ICECLC IN *PSEUDOMONAS KNACKMUSSII* B13.

Nicolas Pradervand, Ryo Miyazaki, Jan Roelof van der Meer

author affiliations:

key words: Efflux system, mfsR, genomic island

running title: TetR-type regulated major facilitator system

ABSTRACT

The integrative and conjugative element ICE*clc* originally found in *Pseudomonas knackmussii* B13, encodes, among others, metabolism of chlorocatechol and 2-aminophenol, two compounds otherwise poisonous to many bacteria. Besides metabolic and horizontal gene transfer functions, ICE*clc* carries a patchwork of various genes, including various transcriptional regulators. In this study, we report of the role of a transcriptional repressor of the Tetracycline repressor family, named MfsR. By using a combination of microarrays, reporter fusions and bandshift assays, we could identify the two DNA targets of MfsR, P_{mfsR} (the promoter of the *mfsR* gene) and P_{mfsA}, the promoter of an operon named *mfsABC-orf38184* and predicted to encode an efflux system.

INTRODUCTION

Homeostasis is one of the defining features of living organisms, and the ability to regulate the concentrations of molecules and ions within the cell is key to survival, especially when cells are exposed to changing environments. Substances that are toxic to the cell even at relatively low concentrations, must be driven out via active transport, using transmembrane transporters and efflux pumps. Efflux pumps can be divided into five main categories: i) the resistance-nodulation-cell division type (RND), ii) the major facilitator family (MFS), iii) small multidrug resistance pumps (SMR), iv) multidrug and toxic compound extrusion systems (MATE), and v) ATP-binding cassette or ABC systems [1]. Activity of efflux pumps in the cell is resource-consuming and, therefore, their synthesis is typically tightly controlled [2]. Control is achieved by dedicated transcriptional regulators that can react to chemical stimuli and trigger expression of genes coding for the efflux pumps, by which the cell can regain its physiological balance [3]. Tetracycline-type regulators (TetR-type) are one of the most prominent families of transcriptional regulators among prokaryotes [4]. TetR-type regulators are often associated with the efflux control of a wide range of antibiotics and organic compounds, although they have also been shown to regulate other cellular processes [5], [6]. Several members of this family are well-studied and interesting, such as TetR itself [7], AcrR [8], QacR [9], TtgR [10], or AdeN [11]. TetR family members are typically transcriptional repressors expressed at a low but constitutive level and need the interaction with a ligand ("inducer") to derepress the promoters they are controlling [12, 13]. Interestingly, in many cases it is the substrate for the

efflux pump itself that functions as the inducer [4]. The specificity of the ligand-repressor interaction is determined by the physico-chemical properties of the repressor's binding pocket. Several TetR-like repressors involved in multidrug drug resistance, such as QacR [14] or TtgR [13], have been reported to recognize with varying affinity a range of ligands, sharing similar properties or structures. Ligand binding results in altered conformation of the regulator (itself in complex with DNA), impedes binding of the regulator to the operator DNA, and subsequently leads to derepression of the genes for the efflux pumps [10, 15]. TetR-like genes are usually (but not exclusively) in a configuration in which the regulator gene lays directly upstream of its target gene, but is divergently oriented [16]. Thus, promoters of both the regulator and the target are often overlapping.

Many of the above-mentioned efflux systems play a role in antibiotic resistance, therefore, their distribution and propagation in clinically relevant strains poses a major health concern [17]. Species-to-species transmission of the genes for efflux systems is typically mediated by horizontal gene transfer processes (HGT), in which mobile genetic elements like plasmids or integrative and conjugative elements (ICEs) play a crucial role [18–20]. Recently, we described a new class of widespread integrative and conjugative elements (ICE) among β - and γ -*proteobacteria*, which, through a region permissive for gene insertions, can carry a wide range of auxilliary gene functions. The currently best-studied model of this family, ICE*clc*, can be found in *Pseudomonas knackmussii* B13 [21], *Ralstonia* sp. strain JS705 [21, 22] and *Burkholderia xenovorans* LB400 [23]. ICE*clc* confers the ability to degrade 3-chlorocatechol and 2-aminophenol [24]. Homologues to ICE*clc*

occur in *Acidovorax* sp. strain JS42, *Bordetella petrii* or *Herminiimonas arsenoxydans*, which carry genes for e.g., heavy metal resistance or other. ICE*clc* is normally integrated in one or more specific sites in the host chromosome (3' 18-bp of the gene for tRNA^{Gly}), but can excise in a small proportion of cells in stationary phase and transfer to a new recipient (species), where it re-integrates [25, 26]. Previously, we noted that ICE*clc*, apart from catabolic genes also carries a gene cluster for a putative MFS efflux system [24]. In this work, we study the regulation and possible roles of this putative MFS efflux system. By microarray and mutant analysis, we identify on ICE*clc* a TetR-type regulator, which controls expression of itself and of the putative MFS efflux system. Promoter analysis and electrophoretic mobility shift experiments with purified N-terminal His6-tagged repressor protein showed that it specifically binds DNA at operator-like sequences presents in both target promoters. The ICE*clc* TetR system is unusual in that the regulator is widely separated from its target operon and, additionally, controls a set of two other transcriptional regulators.

RESULTS

A tetR major facilitator system encoded on ICE*clc*. Close inspection of the ICE*clc* sequence (GenBank accession number AJ617740) indicated the presence of a set of genes for a putative major facilitator system (MFS) efflux pump (*orf32963-34495-36077*) and for a putative *tetR*-type regulator (*orf18502*), which we propose to rename here as *mfsABC* and *mfsR*, respectively, in anticipation of the results described below. The *mfsR*-like gene, however, was distantly located from the *mfs*-genes (Figure 1A). Gene organization and microarray analysis of ICE*clc* gene expression suggested further that the *mfsABC* genes are part of a transcriptional unit that continues with three further open reading frames (Figure 1). The three subsequent open reading frames, *orf37143*, *orf37489* and *orf38184* are predicted to code for an unknown protein, a PagL-like protein and a putative esterase of the alpha-beta hydrolase family, respectively. Mutants were created of *Pseudomonas putida* UWC1 carrying one copy of ICE*clc*, in which the open reading frames *mfsR* or *mfsABC* were precisely deleted. *P. putida* UWC1 (ICE*clc*- Δ *mfsABC*) and *P. putida* UWC1 (ICE*clc*- Δ *mfsR*) were tested in comparison to *P. putida* UWC1 (ICE*clc*) wild-type for their sensitivity to 16 antibiotics and 5 other compounds in a MIC-assay (Minimal Inhibitory Concentration, see *Experimental methods*). Furthermore, BIOLOG phenotypic arrays were used to screen differential chemical sensitivities of *P. putida* UWC1 (ICE*clc*- Δ *mfsABC*) compared to that of *P. putida* UWC1 (ICE*clc*) wild type (Figure S1). No significant differences in MIC-values were obtained for the tested compounds but the BIOLOG tests suggested that chromium (III) chloride

might be toxic for the strain lacking the *mfsABC* genes, whereas the wild type would be more resistant to the presence of this cation (Figure S1). Independent experimental repetitions measuring growth rates, growth yields and ethidium-bromide staining of cells, were not able to confirm the BIOLOG result (data not shown). Therefore, we cannot conclude whether the *mfsABC*-system is responsible for chromium chloride resistance. In order to determine whether gene *mfsR* would be implicated in regulating *mfsABC* expression, we quantified mRNA formation in this region of ICE*clc* using semi-tiling microarray hybridization [27] in cells of *P. putida* UWC1 (ICE*clc*) compared to UWC1 (ICE*clc*- Δ *mfsR*) during exponential growth on 3-chlorobenzoate as sole carbon and energy source, and in subsequent stationary phase. As shown in figure 1B, the operon *mfsABC-orf38184* is expressed on average 12.6-fold more in mutant *P. putida* UWC1 (ICE*clc*- Δ *mfsR*) than in UWC1 (ICE*clc*) wild type. The difference in expression level was independent from the growth phase (Figure 1C, D). This suggested that *mfsR* was directly or indirectly implicated in regulating *mfsABC*.

Regions upstream of *mfsR* and *mfsABC* share similar sequence motifs. Since *mfsR* putatively encodes a TetR-type transcriptional regulator, and most members of this family are known to be autorepressors [4], we hypothesized that the upstream regions of *mfsABC* and *mfsR* would contain similar DNA motifs acting as MfsR binding sites. Indeed, Blastn comparison of the sequences upstream of *mfsR* and of *mfsA* revealed an identical 23 bp box, which we designated OP1 (Figure 2A-B). In addition, each region contained a second imperfect repeat of OP1, named OP2 (for the *mfsR* region) and OP3

(for *mfsA*). Interestingly, OP2 lays 94 bp upstream of OP1 in the *mfsR* promoter region, whereas OP3 is located 89 bp downstream of OP1 in the *mfsA* promoter region (Figure 2B). Both OP1 boxes have an internal palindromic sequence GTACCN₉GGTAC, of which OP2 and OP3 only share the first part GTACCN₉GGT(NN) (Figure 2C). Interestingly, in both promoter regions the two OP boxes are located in the same orientation but on opposite strands (Figure 2C).

MfsR is a regulator that represses both its own promoter and that of the *mfsABC-orf38184* cluster. In order to verify the implication of MfsR in regulating the *mfsR* and *mfsA* promoter regions, reporter gene fusions were constructed. A 656-bp region upstream of *mfsR* ($P_{mfsR-11}$, Figure 2A) and a 486-bp region upstream of *mfsABC* (P_{mfsA-6} , Figure 2B) were each fused to a promoterless *mcherry* gene and introduced in single copy on the chromosome of *P. putida* UWC1 using mini-Tn5 delivery. Consistent with our hypothesis of *mfsR* encoding a repressor, mCherry expression from both *mfsR* and *mfsA* promoters was low in *P. putida* UWC1 (*ICEclc*) but 10-15 times higher in *P. putida* UWC1 without *ICEclc* (Figure 3). mCherry expression was also 7-fold higher in *P. putida* UWC1 (*ICEclc-ΔmfsR*), but was diminished to wild-type *ICEclc* levels when complementing *mfsR* in single copy using mini-Tn7 delivery. Low wild-type mCherry expression also occurred when *P. putida* UWC1 without *ICEclc* was complemented with a single copy mini-Tn7 *mfsR* copy (Figure 3). Taken together, this demonstrated that MfsR alone is sufficient to repress expression of both *mfsR* and *mfsA*-promoters.

None of the aromatic substrates for which ICE*clc* genes encode metabolic enzymes [24], such as 3-chlorobenzoate, 3-chlorocatechol or 2-aminophenol led to a derepression of the P_{mfsA} -*mcherry* fusion in *P. putida* UWC1 (ICE*clc*) when added to the culture media (data not shown).

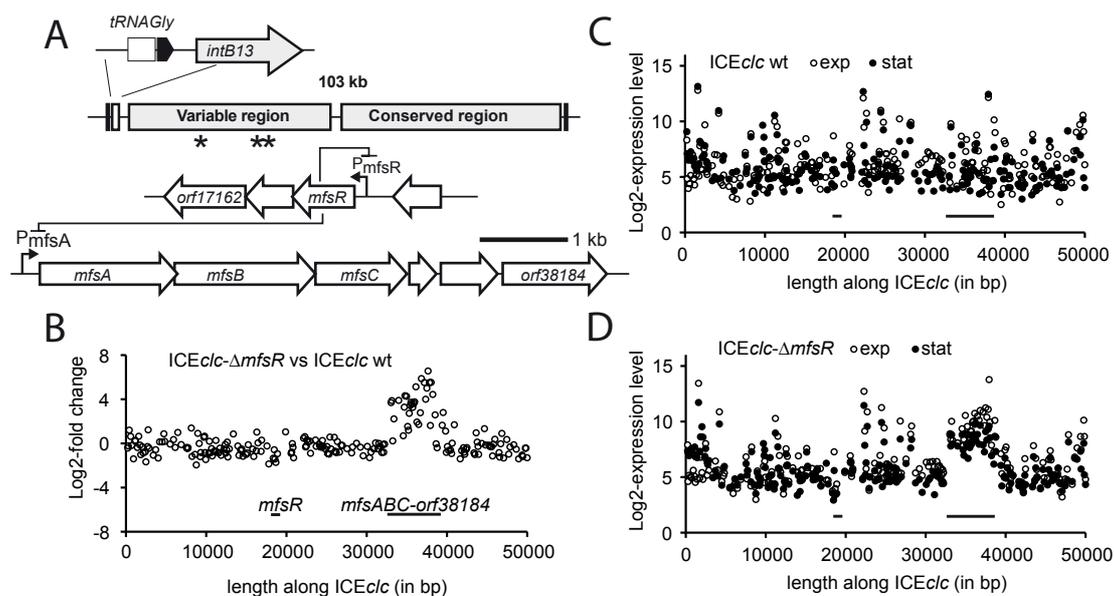


Figure 1. A, Genetic map of ICE*clc* and the relevant regions for microarray analysis. Arrows indicate predicted open reading frames and their orientation. * and **, depict the location of *mfsR* and *mfsABC-orf38184*, respectively. B, Microarray results of the log₂-fold change in mean expression of ICE*clc* in *P. putida* UWC1 (ICE*clc*- Δ *mfsR*) compared to *P. putida* UWC1 (ICE*clc*) wild type, in exponential phase with 10 mM 3CBA as sole carbon and energy source; displayed as a function of length along ICE*clc*. Circles, microarray probes of ICE*clc* "plus strand", i.e. the strand on which operon *mfsABC-orf38184* is encoded (represented by a large black bar). Note that gene expression from *mfsR* is not shown since it lays on the minus strand (its position is indicated by the smaller black bar). C, Log₂-absolute signal intensity of *P. putida* UWC1 ICE*clc* gene expression (position 1-50,000) in exponential (white dots) and stationary (black dots) phase upon growth on 3CBA. D, as C, for *P. putida* UWC1 ICE*clc*- Δ *mfsR* in exponential (white dots) and stationary (black dots) phase.

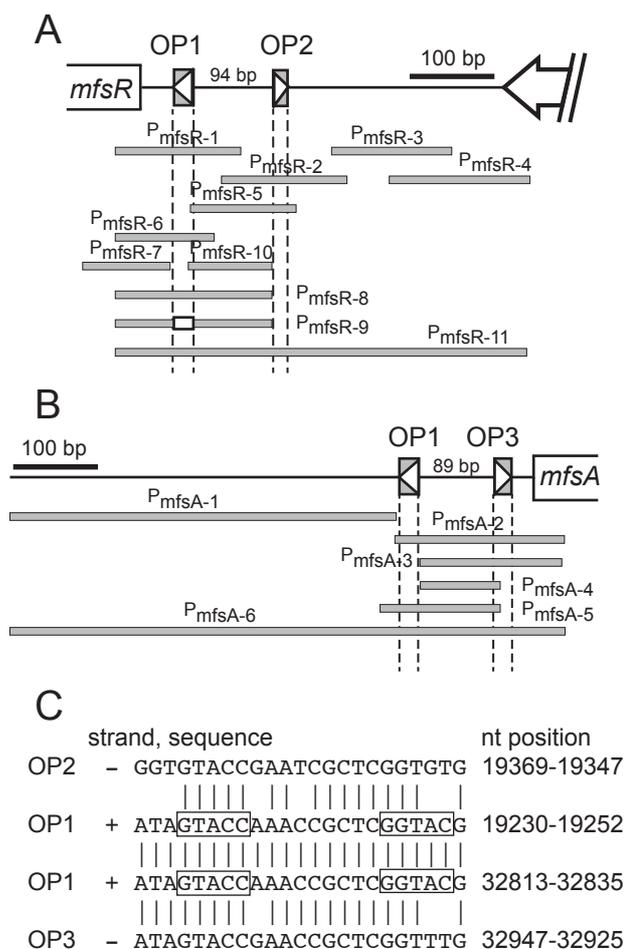


Figure 2. A, overview of the promoter region upstream of *mfsR*. Grey bars, PCR-generated subfragments used in this study. Grey boxes, operator-like sequences OP1 and OP2, the arrows within the OP boxes represent the direction of the homologous OP sequence. Fragment P_{mfsR-9} has its OP1 sequence replaced by a *xhoI* site (white box). B, Promoter region upstream of *mfsA*. C, details of the OP boxes and their nucleotide comparisons. Note the palindromic sequences within the rectangles in both OP1 but incomplete in OP2 and OP3.

MfsR binds to both *mfsR* and *mfsA* promoter regions. In order to verify whether MfsR directly binds to both P_{mfsR} and P_{mfsA} promoter regions and whether the OP1, OP2 and OP3 boxes are involved in binding, electrophoretic mobility shift assays were performed with purified MfsR. The *mfsR* open reading frame was hereto cloned into the *Escherichia coli* expression vector pET22b, which adds a 6His-tag to the C-terminal end of the MfsR polypeptide (MfsR-His6). Nickel-NTA purified MfsR-His6 from *E. coli* has an observed size of 25 kDa (predicted size 25.8 kDa), whereas glutaraldehyde crosslinking resulted in the formation of a 50 kDa protein band (Figure 4). This suggests MfsR to form homodimers in solution, which is in

agreement with several observations that TetR family members occur as homodimers [28, 29].

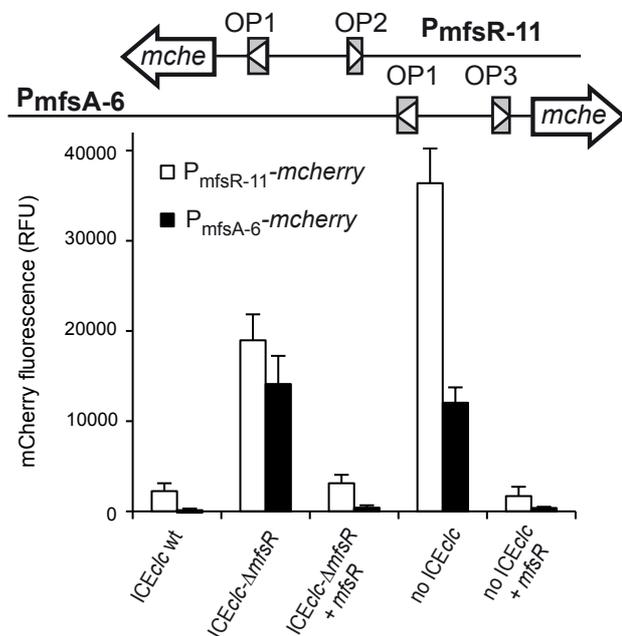


Figure 3. Reporter gene expression from *mfsR* and *mfsA* promoters in presence or absence of *mfsR* in *P. putida* UWC1. $P_{mfsR-11}$ -mcherry (white bars) and P_{mfsA-6} -mcherry (black bars). Strains of *P. putida* UWC1 as indicated below. mCherry fluorescence of the cultures was measured 20 hours after inoculation in MM with 15 mM succinate as sole carbon and energy source and is normalized by the culture density (RFU, relative fluorescence units).

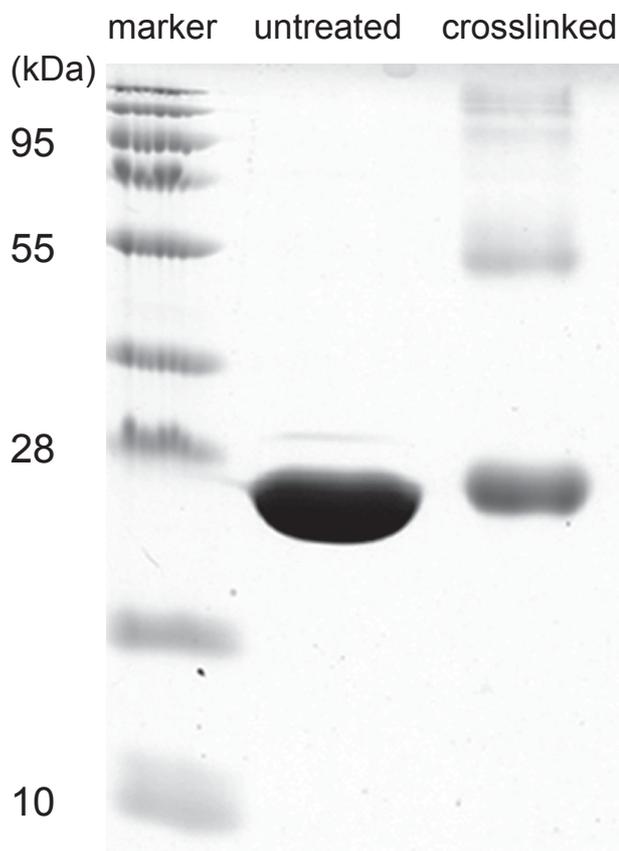


Figure 4. Purification of MfsR-His6 protein. Gel shows untreated Ni-NTA purified MfsR-His6 from *E. coli* extract and glutaraldehyde-crosslinked protein on 12% SDS PAGE. Ladder, protein size marker in kiloDalton (kDa).

Purified MfsR-His6 was incubated with a variety of different length fluorescently end-labelled fragments (depicted in figure 2A-B) from both *mfsR* or *mfsA* promoters, and the resulting DNA-protein complexes were separated on non-denaturing polyacrylamide gels and imaged using fluorometry (Figure 5). As expected, MfsR-His6 binds to the "complete" P_{mfsR} and P_{mfsA} promoter regions ($P_{mfsR-11}$ and P_{mfsA-6} , respectively, Figure 5), but not to a random other *ICEclc* intergenic region IG_{85934} . Strangely enough, both P_{mfsR} and P_{mfsA} - MfsR-His6 complexes yielded two identifiable bands (Figure 5 $P_{mfsR-11}$ +/-, P_{mfsA-6} +/-), suggesting binding of multiple MfsR-His6 protein complexes. In addition, every subfragment containing at least the OP1 box (fragments P_{mfsR-1} , P_{mfsR-6} , P_{mfsR-8} , $P_{mfsR-11}$, P_{mfsA-2} , P_{mfsA-5} , P_{mfsA-6}), the OP2 (fragments P_{mfsR-2} , P_{mfsR-5} , $P_{mfsR-11}$) or OP3 boxes (fragments P_{mfsA-2} , P_{mfsA-3} , P_{mfsA-6}), were bound by purified MfsR-His6 (Figure 5). In contrast, none of the fragments without at least one of the OP boxes (fragments P_{mfsR-3} , P_{mfsR-4} , P_{mfsR-7} , P_{mfsR-9} , $P_{mfsR-10}$, P_{mfsA-1} , P_{mfsA-4}) bound MfsR-His6. Collectively, these results clearly indicated that MfsR directly binds to fragments encompassing at least one, but any of the OP boxes. The exact sequence of MfsR interaction on the operator DNA may lay entirely or partially within the predicted OP box sequence (Figure 2C).

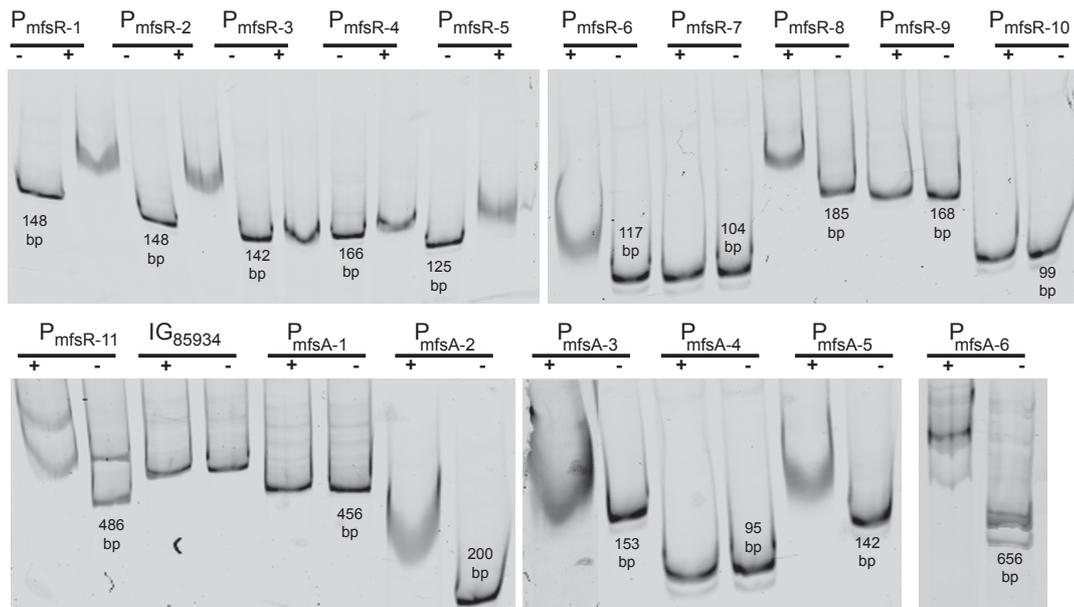


Figure 5. DNA binding of purified MfsR-His6 to *mfsR* and *mfsA* promoters. Images show electrophoretic mobility shift experiments with different fragments of the regions upstream of *mfsR* and *mfsA* (see figure 2A-B for mapping). 0.5 nM of DNA fragments labeled at the 5' end with Dyomics681 infrared dye were added in each reaction. -, no purified MfsR-His6 added. +, MfsR-His6 added (0.5 μ M per reaction). Numbers in the lanes indicate the fragment size.

A single OP1 box is sufficient for MfsR repression. To determine whether both OP boxes in the upstream regions were necessary for repression, we constructed a shorter *mfsR* promoter fragment (P_{mfsR-8}, Figure 2A), fused this to a promoterless *mCherry* and inserted it in single copy onto the chromosome of *P. putida* UWC1. As before for the longer *mfsR* promoter fragment (P_{mfsR-11}, Figure 6), mCherry expression from P_{mfsR-8} was high in *P. putida* UWC1 without *ICEclc* and in *P. putida* UWC1 (*ICEclc-ΔmfsR*) (Figure 6, lanes 4 and 2). In contrast, mCherry expression from P_{mfsR-8} is low in *P. putida* UWC1 (*ICEclc*), *P. putida* UWC1 (*ICEclc-ΔmfsR*) complemented with a single copy mini-Tn7 *mfsR* insertion, or *P. putida* UWC1 mini-Tn7 *mfsR* (Figure 6, lanes 1, 3 and 5, respectively). Deleting the OP1 box in the *mfsR*-promoter

(replacing it with an *XhoI* site) resulted in loss of mCherry expression in *P. putida* UWC1 (P_{mfsR-9} , Figure 6, lane 6). This indicated that the interaction of MfsR with a single OP1 box is sufficient for an efficient repression of P_{mfsR} . Results with the P_{mfsR-9} fragment indicated that the OP1 sequence in P_{mfsR} is also essential for expression. It must be noted that the promoter activity of P_{mfsR-8} in *P. putida* UWC1 is threefold lower than that of $P_{mfsR-11}$ (Figure 6, lane 8). This suggests that the promoter in P_{mfsR-8} is somehow incomplete (yet functional) or that additional *cis* elements, such as enhancers, are missing.

A similar shorter promoter fragment of *mfsA* was not functional in *P. putida* UWC1 (P_{mfsA-3} , Figure 6, lane 9). Because of the promoter geometry this fragment carries the OP3 box closer to the gene start, whereas in the *mfsR* region this is OP1. This finding, however, suggests that fragment P_{mfsA-3} does not encompass the *mfsA* promoter, which must therefore lay further upstream. Bioinformatic promoter searches using SoftBerry's BPRM suggests that a -35 and a -10 box exist upstream and partially overlapping with OP1. However, it is also possible that other essential *cis* elements are missing in that fragment or the promoter is truncated.

Because *mfsR* is distantly located from *mfsABC-orf38184* we finally tested whether its upstream region had retained the classical promoter configuration of TetR/TetA (i.e., two overlapping but divergent promoters). Hereto, the fragment P_{mfsA-6} (Figure 2B) was cloned in the opposite orientation with respect to the *mcherry* gene. *P. putida* UWC1 and *P. putida* UWC1 miniTn7 *mfsR* with a single copy of this promoter fusion, however, did not express any mCherry fluorescence (Figure 6, lanes 12 and 13, $P_{mfsA-6R}$). Therefore, we

conclude that there was no promoter in the reverse direction that would help explain the presence of the secondary OP box.

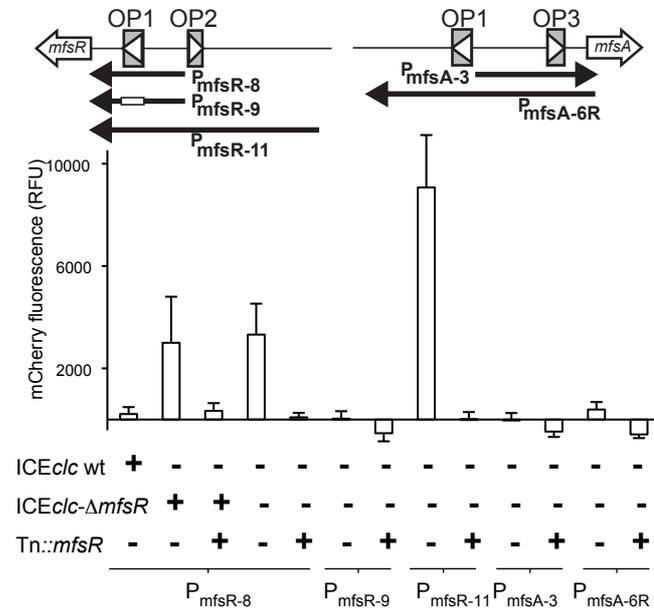


Figure 6. Reporter gene expression of *mfsR* and *mfsA* promoter subfragments fused to a promoterless *mCherry* gene in single copy in *P. putida* UWC1 with (+) or without (-) *ICEclc* or *ICEclc- Δ mfsR*, in presence (+) or absence (-) of complemented *mfsR*. Fluorescence was measured 25.5 hours after inoculation in MM plus succinate, and is normalized by the culture density.

DISCUSSION

TetR-type transcriptional regulators typically repress the expression of efflux pumps genes as well as their own, in a wide range of prokaryotic species [4]. In the present study, we characterized the TetR-like repressor MfsR encoded on the mobile element ICE*clc*. By using microarrays, we could identify the potential targets of MfsR by revealing differential gene expression between wild type ICE*clc*-bearing *P. putida* and a mutant strain lacking *mfsR*. We showed that deletion of *mfsR* leads to the overexpression of an operon-like set of 6 genes termed *mfsABC-38184*, also carried on ICE*clc*. The first three open reading frames of this gene cluster are predicted to encode a tripartite efflux pump of the MFS superfamily of transporters. Promoter reporter fusions confirmed microarray results by demonstrating that the *mfsR* gene product is responsible for the repression of *mfsABC-38184* expression, as well as of itself. The repressor protein, like many of its TetR-like relatives, forms homodimer complexes and even possibly multimers. MfsR binds specifically to DNA boxes that are essential for protein-DNA interaction and which are present in pairs in both promoters P_{mfsR} and P_{mfsA} . These DNA boxes, operators or operator-containing, encompass more or less conserved pseudo-palindromic sequences, that seem to be key features for sequence recognition. Interestingly, a single of these OP boxes (OP1) is already enough for MfsR to correctly prevent transcription of P_{mfsR} . The presence of the OP1 sequence was also required for correct expression (Figure 6, results of P_{mfsR} -9), suggesting that it carries (part of) the promoter sequence itself. Indeed, a possible -35 box (TTGAAG) is located just upstream of OP1 and a putative -

10 (TACTAT, 17 bp spacing) is part of OP1. Hence, the role of the second and more upstream OP box, OP2, is not immediately clear. It is reasonable to assume that a second repressor-binding box in the immediate vicinity (137 bp) of the acting repressor (OP1) can significantly increase the concentration of repressor protein in the region and thus enhance the efficiency of repression [30]. Alternatively, OP2 is a now-defunct relic inherited from a past configuration in which *mfsR* was directly upstream and divergently oriented from *mfsABC* (see further below for why such a scenario may be likely). A past duplication and rearrangement may then have created both OP1-OP2 and the OP1-OP3 regions. Interestingly, the longer *mfsR* promoter fragment in absence of MfsR leads to stronger expression of the reporter gene (Figure 6, compare fragments $P_{mfsR-11}$ with P_{mfsR-8}), which may be the result of enhancer sequences or other hidden promoters located upstream of OP1.

In the *mfsA* promoter region the order of the OP1 and OP3 box is reversed compared to the *mfsR* promoter (Figure 2). Interestingly, however, the shorter version of P_{mfsA} (including OP3 but not OP1) does not invoke expression of the fused reporter gene in absence of MfsR (Figure 6, P_{mfsA-3}). This suggests that similar as for P_{mfsR} , the promoter lies most probably within (at least partly) OP1. Indeed, a putative -10 box overlaps with OP1 in the *mfsA* promoter (TATAGT), with a putative -35 box 17 bp upstream (TTGACA). Both *mfsA* and *mfsR* promoters may have diverged sufficiently in that the reverse orientation of the *mfsA* promoter fragment does not invoke expression of a reporter gene (Figure 6, $P_{mfsA-6R}$). The role of OP3 may thus be to act as an auxiliary repressor-binding site, which may further reduce transcription from the *mfsA* promoter. Even though MfsR of ICE*c/c* belongs to the TetR family, it is a

distant relative of most common TetR members (data not shown). Its closest neighbors consist of an MfsR in *Acidovorax* sp. JS42 (termed ajs_2925, 93% amino acid identity; accession number: CP000539.1) and in *Aeromonas hydrophila* SSU (termed HMPREF1171_02726, 93% amino acid identity; accession number: AGWR0100022.1). In those strains, however, the *mfsR* gene is directly upstream and divergently oriented of an *mfsABC* cluster, with proper conservation and order of similar OP1 and OP2/3 boxes (Figure 7).

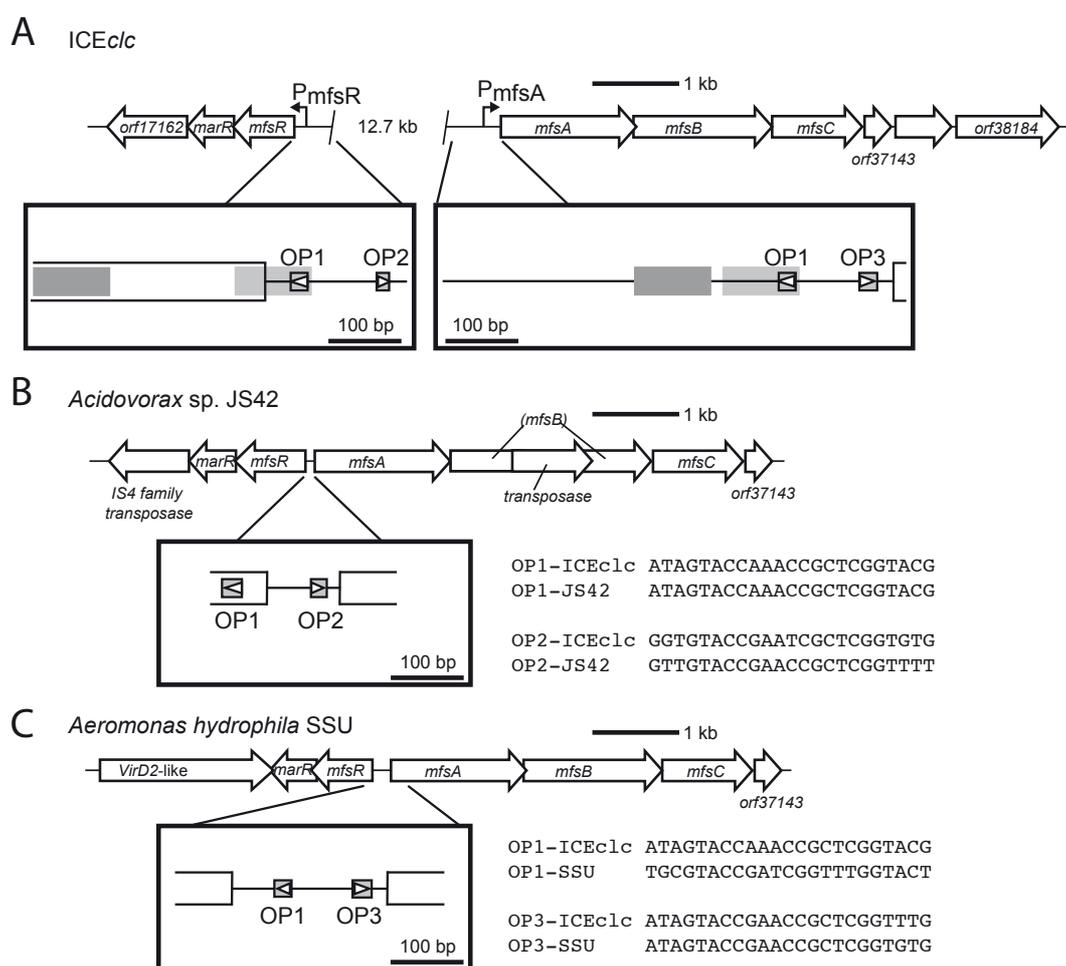


Figure 7. Strain comparisons of the *mfsR/mfsABC* synteny. **A**, *mfsR* and *mfsABC* on ICE_{clc}, where they are separated by 12.7 kb. Details of the promoter regions are illustrated within boxes. Grey shaded regions point to homologies between *mfsR* and *mfsA* promoters. **B**, *mfsR* and *mfsABC* homologs in *Acidovorax* sp. JS42. Note that the annotated *mfsR* in JS42 is longer than its ICE_{clc} counterpart and overlaps with the predicted OP1 box. A transposase of the IS116/IS10/IS902 family has disrupted the sequence of *mfsB* and another transposase (of the IS4 family) is located downstream of *mfsR*. Sequences highlight the similarity between the

OP2 boxes, whereas the OP1 box is 100% conserved (not shown). C, *mfsR* and *mfsABC* in *A. hydrophila* SSU. The *mfsR* and *mfsA* intergenic region is twice the size of that in JS42 (212 bp), and *mfsR* has the same annotated size as for ICE*clc*. Sequence homology between the OP1-boxes and between the OP3-boxes is indicated.

This finding would be consistent with the hypothesis that *mfsR-mfsABC* on a predecessor ICE*clc* were separated by the insertion of other genes (those for 2-aminophenol degradation), during which the promoter regions became duplicated and diverged. Indeed, small regions of overall 70% conserved nucleotide identity occur between *mfsR* and *mfsA* (Figure 7A, shaded regions). It is further interesting to note that the *mfsR-mfsABC* genes in *Acidovorax* sp. strain JS42 are not present on a mobile element closely similar to ICE*clc*, but rather on another ICE termed ICE(Tn4371)6039 from the Tn4371 family [31, 32]. The exact nature of the location of the *mfsR-mfsABC* genes in *A. hydrophila* SSU (part of an ICE or not) cannot be determined yet, since this genome has not been finished (Figure 7). ICE*clc*-LB400, an almost identical ICE*clc* in *B. xenovorans* LB400 [23] carries the same "separated" *mfsR-mfsABC* configuration as ICE*clc* itself. In contrast, other elements related to ICE*clc*, such as ICE-GI6 of *B. petrii* DSM 12804 [33] only have incomplete fragments of the *mfsR/mfsABC* genes (data not shown).

Despite the clear homology of the *mfsR-mfsABC* system on ICE*clc* to known TetR and major facilitator efflux systems, and its distribution among other bacterial genomes, it is not clear yet what the identity of the ligand/inducer for the pumps and for the derepression is. In analogy to other TetR-like regulons, we had assumed that such ligand might be the toxicant effluxed by the MfsABC complex. Unfortunately, all our efforts to identify such toxicants were

so far fruitless. Neither 3-chlorocatechol, 3-chlorobenzoate or 2-aminophenol (whose metabolisms are encoded on *ICEclc*) nor general cell stressors (such as antibiotics and organic solvents) showed statistically significant differences in MIC assays between wild-type and *mfsR*- or *mfsABC*-mutants of UWC1. Neither was the P_{mfsR} -promoter inducible with those substances (data not shown, see material and methods). The only hint came from a BIOLOG Phenotypic Array-screening, which suggested that the *mfsABC*-deletion mutant of *P. putida* UWC1 (*ICEclc*) reacted more sensitively than the wild-type to chromium chloride (Supplementary Figure S1). However, both growth rates and yields of *P. putida* UWC1 (*ICEclc*) and the *mfsABC*-deletion mutant in the absence or presence of 100 mg chromium per L were indistinguishable (F. Delavat, not shown). Absence of clear phenotypic differences upon deletion of the genes for an efflux pump is not uncommon, since many bacterial genomes encode multiple copies of such efflux systems, probably the most extreme being *Pseudomonas aeruginosa* [34–36]. Also *P. putida* encodes seven paralogs of RND-type multidrug efflux systems [37].

Nonetheless, the presence of a TetR-type regulon controlling a set of genes with clear homology to a major facilitator superfamily efflux system on *ICEclc* emphasizes the role of mobile genetic elements in the spread of (potential) resistance genes in the bacterial world. This is further supported by the apparent DNA rearrangements between and within ICEs, as demonstrated by the differences between the *mfs* loci of *P. knackmussii* *ICEclc* and the equivalent of *Acidovorax* sp. strain JS42 ICE(Tn4371)6039. It is for certain, that the more ICEs are discovered and studied, the more their role in the adaptation and shaping of bacterial genomes will be understood.

MATERIALS AND METHODS

Bacterial strains and plasmids. *Escherichia coli* DH5 α (Gibco Life Technologies, Gaithersburg, Md.), DH5 α λ pir and *E. coli* S17-1 λ pir were routinely used for plasmid propagation, clonings. *E.coli* BL21 (DE3) was used for protein overexpression. Strains used during this study are listed in table 1.

Media and growth conditions. *E.coli* and *Pseudomonas putida* were grown at 37°C and 30°C, respectively. *E.coli* was cultivated on Luria-Bertani medium (LB) whereas *P. putida* was grown on LB and 21C mineral medium-based minimal medium [38, 39] complemented with either 10 mM 3CBA or 15 mM succinate as sole carbon and energy sources. When necessary, the following antibiotics were added to culture media at the indicated concentrations : kanamycin (Km) 25 μ g/ml, gentamicin (Gm) 20 μ g/ml, ampicillin (Ap) 100 or 500 μ g/ml.

Strains constructions. Isolation of both chromosomal and plasmid DNA, PCR, restriction enzyme digestion, ligation and electroporation were performed as described by standard procedures [39] Gene deletion mutants were obtained by double recombinations described elsewhere [40]. Complementations of mutations were done by introducing genes as single copies via the Tn7 delivery vector system pUC18-miniTn7_Gm [41]. Nucleotide positions are given according to GenBank accession number AJ617740 (ICE*clc*).

Minimal inhibitory concentration assays (MICs). The minimal inhibitory concentrations of the following compounds were determined for the three strains *P. putida* UWC1 (ICE*clc*) wild type, *P. putida* UWC1 (ICE*clc*- Δ *mfsABC*) and *P. putida* UWC1 (ICE*clc*- Δ *mfsR*): 3-chlorobenzoate (3CBA), 2-aminophenol (2AP), 4-chlorophenol, toluene, ethidium bromide, ciprofloxacin, tetracycline, gentamicin, chloramphenicol, piperacillin/tazobactam, meropenem, erythromycin, aztreonam, imipenem, ofloxacin, trimethoprim, ceftazidime, sparfloxacin, trovafloxacin, moxifloxacin, and levofloxacin. Strains were grown at 30°C with orbital shaking using either MM with 10 mM 3CBA or MMB as growth media. Assays were performed as described elsewhere [42].

Reporter gene fusions. Promoter fragments were produced by PCR, cloned in front of a promoterless *mcherry* gene and delivered in single copy into various strains (listed in Table 1) using the mini-Tn5 delivery system pCK218 {Reference}. Transformants were selected on Km selective medium and verified by specific PCR. For each reporter fusion, at least three independent clones were used as replicates in fluorometry measurements. Reporter strains were grown overnight in minimal medium (MM) with 15 mM succinate Km25 in 96-well black microtiter plates with flat transparent bottoms (Greiner bio-one). 5 μ l of each culture was then used as an inoculum for the wells filled with fresh 195 μ l of MM 15 mM succinate Km25. The plates were subsequently incubated at 30°C with orbital shaking (300 rpm), and their absorbance (OD600) and fluorescence (590, 620) were monitored at regular time intervals during 2 days. The data displayed in figures 2 and 5 show the ratio of fluorescence intensity divided by the absorbance at OD600 minus the

same ratio of a control non-fluorescent strain (carrying an empty reporter fusion).

Induction assay. *P. putida* UWC1 (ICE*clc*) carrying promoter fusion P_{mfsA-6} -*mcherry* was cultured on MM with 15 mM succinate and 25 µg/ml Km in Erlenmeyer flasks. At mid exponential phase, one of the following compounds was added to the cultures: 10 mM 3CBA, 0.5 mM 3-chlorocatechol, 5 mM 2-aminophenol, 4% or 0.4% ethanol (shock induction). The absorbance (OD600) and the mCherry fluorescence were monitored for the next two and half hours in 96 wells microtitre plates and compared to a blank uninduced control. Simultaneously, the same reporter strain was cultured on MM with 10 mM 3CBA and 25 µg/ml Km (smooth induction). The absorbance and the mCherry fluorescence were measured for 35 hours and compared to that of the same strain growing on 15 mM succinate.

BIOLOG Phenotypic MicroArray. Chemical sensitivity differences between *P. putida* UWC1 (ICE*clc*) and *P. putida* UWC1 (ICE*clc*Δ*mfsABC*) were determined by BIOLOG Phenotypic MicroArrays PM09-20 (Biolog Inc., Hayward, California, <http://www.biolog.com>).

Microarray analysis. Transcriptome profiling of *P. putida* UWC1 (ICE*clc*) wild type and *P. putida* UWC1 (ICE*clc*-Δ*mfsR*) were performed as described previously [27]. Strains were grown on minimal medium (MM) with 10 mM 3CBA as sole carbon and energy sources. Cells were harvested at mid-exponential phase and 48 hours after the entrance in stationary phase.

Total RNA was extracted, cy3-labeled cDNA was synthesized by reverse transcription, purified and hybridized to Agilent 8x15k custom ICE*clc* microarray slides. After washing, slides were scanned, data extracted and analyzed using GeneSpring GX. Microarray data have been deposited on GEO database with the accession number : GSE51391

Protein purification. *mfsR* was amplified by PCR and cloned (NdeI-XhoI) into pET-22b (+) (Novagen, Cat. No. 69744-3), thus adding a polyhistidine tag at the C-terminal end. The vector was transformed into *E.coli* BL21 (DE3), which was subsequently grown on LB Ap100 at 25°C, and 220 rpm rotary shaking, and induced with 1 mM IPTG. At mid exponential phase, cells were collected, pelleted down, and lyzed using a French press. His-tagged protein was purified according to procedures outlined by the supplier (Qiagen), and dialysed overnight in *N*-cyclohexyl-3-aminopropanesulfonic acid (CAPS)-based DNA binding buffer (10 mM CAPS pH10.5, 250 mM NaCl, 10 mM Mg Acetate, 10 mM KCl, 0.1 mM EDTA, 1 mM DTT, 5% (v/v) glycerol) using a Slide-A-Lyzer cassette (Thermo Scientific).

Dimerization assay. 62 µg of freshly purified protein were incubated with 0.4% glutaraldehyde for 30 minutes. The reaction mix was then boiled and loaded on 12% SDS AA gel and run for 2 hours at 250 V [43].

Fluorescence-based EMSA. DNA fragments encompassing the P_{mfsR} and P_{mfsA} promoters (listed in figure 2A-B) were produced by PCR and purified from gel. Then, they were ligated into pGEM-T-easy (Promega) and

transformed into *E. coli* DH5 α for propagation. The plasmids were extracted by using a Genomed's Jetstar Midiprep (<http://www.genomed-dna.com>) and linearized by *Sca*I digestion. Linear plasmids were used as templates in the PCR using M13 primers with attached Dyomics681 dye at their 5'-end (Microsynth AG). The resulting Dyomics681-labeled promoter fragments were purified on gel and used in subsequent EMSAs (0.5 nM final concentration). For each EMSA reaction (a total of 20 μ l), 0.2 μ g of Salmon sperm DNA and 1 μ g BSA were added to the CAPS-based DNA-binding buffer. 0.5 μ M of freshly purified MfsR-His6 was added or not in the assays. Reaction mixes were incubated during 30 minutes at room temperature and protected from direct sunlight. Then, samples were loaded on a CAPS-based native 4% acrylamide gel and run for 4-6 hours at 50 V in a CAPS buffer (pH 10.5). Gels were then incubated for 1 hour in gel drying solution (containing 20% ethanol, 10% methanol, 10% isopropanol, 2% glycerol, 58 % H₂O) and directly scanned on a fluorescence bed scanner (680, 700 nm, Li-Cor Odyssey scanner, Li-Cor Biosciences).

Homology searches and synteny comparisons. Blastn and blastp analyses were carried out using individual ICE*c/c* genes sequences as query, and ICEberg [44] and ncbi's nucleotide collection (nr/nt) and non-redundant protein sequences (nr) as databases. Synteny analysis between ICE*c/c* and related genomes was performed with the Artemis Comparison Tool via the webACT site [45, 46].

Table 1. Bacterial strains used in this study.

| | Characteristics | Reference or source | Lab collection strain number | Primers used | |
|-----------------------|--|--|------------------------------|--|--|
| <i>E. coli</i> | | | | | |
| | BL21 (DE3) / pET22b(+)- <i>mfsRHIS₆</i> | this study | 3787 | 110705 + 110706 | |
| | DH5 α / pGEMt-easy-P _{mfsR-1} | this study | 4453 | 120824 + 120825 | |
| | DH5 α / pGEMt-easy-P _{mfsR-2} | this study | 4454 | 120826 + 120827 | |
| | DH5 α / pGEMt-easy-P _{mfsR-3} | this study | 4455 | 120828 + 120829 | |
| | DH5 α / pGEMt-easy-P _{mfsR-4} | this study | 4456 | 120830 + 120831 | |
| | DH5 α / pGEMt-easy-P _{mfsR-5} | this study | 4457 | 120832 + 120833 | |
| | DH5 α / pGEMt-easy-P _{mfsR-6} | this study | 4458 | 120824 + 100210 | |
| | DH5 α / pGEMt-easy-P _{mfsR-7} | this study | 4459 | 121106 + 121107 | |
| | DH5 α / pGEMt-easy-P _{mfsR-8} | this study | 4460 | 121107 + 121109 | |
| | | | | for the fusion : 121211 + 121212 | |
| | DH5 α / pGEMt-easy-P _{mfsR-9} | this study | 4461 | 121106+121107 <i>xhoI</i> -ligated with 121108+121109 for the fusion : 121211 + 121212 | |
| | DH5 α / pGEMt-easy-P _{mfsR-10} | this study | 4462 | 121108 + 121109 | |
| | DH5 α / pGEMt-easy-P _{mfsR-11} | this study | 3859 | 121211 + 121213 | |
| | DH5 α / pGEMt-easy-P _{mfsA-1} | this study | 4463 | 120513 + 121105 | |
| | DH5 α / pGEMt-easy-P _{mfsA-2} | this study | 4464 | | |
| | | | | 121104 + 120512 | |
| | DH5 α / pGEMt-easy-P _{mfsA-3} | this study | 4465 | 121102 + 120512 for the fusion : 121214 + 120512 | |
| | DH5 α / pGEMt-easy-P _{mfsA-4} | this study | 4466 | 121102 + 121103 | |
| | DH5 α / pGEMt-easy-P _{mfsA-5} | this study | 4467 | 121104 + 121103 | |
| | DH5 α / pGEMt-easy-P _{mfsA-6} | this study | 3918 | 120513 + 120512 | |
| | DH5 α / pGEMt-easy-P ₈₅₉₃₄₋₈₈₄₀₀ | Contains promoter region of ICE <i>clc</i> P ₈₅₉₃₄₋₈₈₄₀₀ | this study | 4000 | 120203 + 120204 |
| <i>P. putida</i> UWC1 | | | | | |
| | no ICE <i>clc</i> | Plasmid-less derivative of <i>P. putida</i> mt-2 (TOL), used as recipient for ICE <i>clc</i> , Rif ^R | [25] | 1291 | |
| | no ICE <i>clc</i> + Tn7- <i>mfsR</i> | 1291-derivative with integrated single copy Tn7- <i>mfsR</i> , Gm ^R | this study | 4301 | 100208 + 120222 |
| | ICE <i>clc</i> wt | One copy of ICE <i>clc</i> integrated in <i>tRNAGly</i> -gene #5, obtained via conjugation with <i>P. knackmussii</i> sp B13 and strain 1291 | [20] | 2737 | |
| | ICE <i>clc</i> - Δ <i>mfsR</i> | 2737-derivative with a deletion in <i>mfsR</i> . Subsequent putative hidden gene <i>orf17984</i> also compromised | this study | 3543 | 101105 +101106, 101107 + 101108 |
| | ICE <i>clc</i> - Δ <i>mfsR</i> + Tn7- <i>mfsR</i> | 3543-derivative complemented in trans with Tn7- <i>mfsR</i> with its native promoter | this study | 4161 | |
| | ICE <i>clc</i> - Δ <i>mfsABC</i> | 2737-derivative with deletion in <i>mfsABC</i> . | this study | 4165 | 120716 + 120717, 120718 + 129719 |
| | no ICE <i>clc</i> + P _{mfsR-11} - <i>mcherry</i> | | this study | 3482-3487 | |
| | no ICE <i>clc</i> + Tn7- <i>mfsR</i> + P _{mfsR-11} - <i>mcherry</i> | | this study | 4302-4307 | |
| | ICE <i>clc</i> wt + P _{mfsR-11} - <i>mcherry</i> | | this study | 3497-3502 | |
| | ICE <i>clc</i> - Δ <i>mfsR</i> + P _{mfsR-11} - <i>mcherry</i> | | this study | 3606-3611 | |
| | ICE <i>clc</i> - Δ <i>mfsR</i> + Tn7- <i>mfsR</i> + P _{mfsR-11} - <i>mcherry</i> | | this study | 4282-4287 | |
| | no ICE <i>clc</i> + P _{mfsA-6} - <i>mcherry</i> | | this study | 4272-4277 | |
| | no ICE <i>clc</i> + Tn7- <i>mfsR</i> + P _{mfsA-6} - <i>mcherry</i> | | | 4308-4313 | |
| | | | this study | | |
| | ICE <i>clc</i> wt + P _{mfsA-6} - <i>mcherry</i> | | this study | 4254-4259 | |
| | ICE <i>clc</i> - Δ <i>mfsR</i> + P _{mfsA-6} - <i>mcherry</i> | | this study | 4260-4265 | |
| | ICE <i>clc</i> - Δ <i>mfsR</i> + Tn7- <i>mfsR</i> + P _{mfsA-6} - <i>mcherry</i> | | this study | 4266-4271 | |
| | ICE <i>clc</i> wt + P _{mfsR-8} - <i>mcherry</i> | | this study | 4410-4413 | |
| | ICE <i>clc</i> - Δ <i>mfsR</i> + P _{mfsR-8} - <i>mcherry</i> | | this study | 4414-4417 | |
| | ICE <i>clc</i> - Δ <i>mfsR</i> + Tn7- <i>mfsR</i> + P _{mfsR-8} - <i>mcherry</i> | | this study | 4418-4421 | |
| | no ICE <i>clc</i> + P _{mfsR-8} - <i>mcherry</i> | | this study | 4402-4405 | |
| | no ICE <i>clc</i> + Tn7- <i>mfsR</i> + P _{mfsR-8} - <i>mcherry</i> | | this study | 4406-4409 | |

Chapter 4

| | | |
|---|------------|-----------|
| no ICE clc + P $_{mfsR-9}$ - <i>mcherry</i> | this study | 4392-4395 |
| no ICE clc + Tn7- <i>mfsR</i> + P $_{mfsR-9}$ - <i>mcherry</i> | this study | 4396-4399 |
| no ICE clc + P $_{mfsA-3}$ - <i>mcherry</i> | this study | 4424-4427 |
| no ICE clc + Tn7- <i>mfsR</i> + P $_{mfsA-3}$ - <i>mcherry</i> | this study | 4428-4431 |
| no ICE clc + P $_{mfsA-6rev}$ - <i>mcherry</i> | this study | 4434-4437 |
| | this study | 4438-4441 |
| no ICE clc + Tn7- <i>mfsR</i> + P $_{mfsA-6rev}$ - <i>mcherry</i> | | |

Table 2. Primers list

| primer name | sequence |
|-------------|-------------------------------------|
| 100208 | ATCATGATCGGGAGCGTTTACT |
| 100210 | GCTAGGGAATGTGCTATTCAAAG |
| 101105 | GAATTCTGCTCCAGGACGGTGAACAA |
| 101106 | CTCGAGGGCTGCTCCATTTGGTTTGACT |
| 101107 | CTCGAGGACTTGGTTAGTGC GTTGCATCT |
| 101108 | GGATCCGCAGTGCAGAGTTCCTTTTAGAG |
| 110705 | TTTTCATATGCAGATGCAACGCACTA |
| 110706 | TTTTCTCGAGTGATCGGGAGCGTTTAC |
| 120204 | TTTTTTCTAGAGACCCTCATTACATCGACATGAC |
| 120222 | TTTTGGATCCCCACCTTCGTGGTCAATTC |
| 120512 | TTTTGGATCCAGCAGCGGTTGTGACGTCAT |
| 120513 | TTTTTCTAGAGGAGCTGATGGATGGATGA |
| 120203 | TTTTTGGATCCGGACGGGCTCCTTGAAAAG |
| 120716 | TTTTTCTAGACTGCCTTTGCCGATGC |
| 120717 | GAAAGTTCACCCAATATACGAATAG |
| 120718 | ACTAGTCGCCATTTCGTCCAGCATG |
| 120719 | TTTTGGATCCAATGGCCACCCCATAG |
| 120824 | ATGACTTG GTTAGTGC GTTGCAT |
| 120825 | TCACAAGTCGATATAAGCAATGCG |
| 120826 | CGCATTGCTTATATCGACTTGTGA |
| 120827 | GGCAGGAATGAACTGCGCC |
| 120828 | GGCGCAGTTCATTCCTGCC |
| 120829 | TGCCTGCCAATCTCGGCATCTC |
| 120830 | AAGAGCTGCTGCGACAGCAG |
| 120831 | CCACCTTCGTGGTCAATTC |
| 120832 | CGCTTCAACTTTGAATAGCAC |
| 120833 | CACGCTGGTGTACCGAAT |
| 121102 | GTCAAACCTCAACCACGCCGC |
| 121103 | GGTTTGACATAAAAGAAAAG |
| 121104 | TTGACACAACGCCCGAAATCGGT |
| 121105 | AACACCCCAAGCAGGTGCGA |
| 121106 | TTTTTCTCGAGGCTTCATCTAGGCCCTTGT |
| 121107 | TTGGCGTTCAGCTCGCCGCTGCT |
| 121108 | TTTTCTCGAGCGCTTCAACTTTGAATAGCAC |
| 121109 | GTGCCATAGCGCAATTAAGA |
| 121211 | TTTTTGGATCCATGACTTGGTTAGTGC GTTGCAT |
| 121212 | TTTTTCTAGAGTGCCATAGCGCAATTAAGA |
| 121213 | TTTTTCTAGACCACCTTCGTGGTCAATTC |
| 121214 | TTTTTCTAGAGTCAAACCTCAACCACGCCGC |

ACKNOWLEDGMENTS

The authors gratefully thank JM Enteza and M Giddey for their help with the MIC assays. F Delavat and S Beggah are thanked for their help to test the possible sensitivity of wild-type and mutant strains to chromium chloride. N Pradervand was supported by a fellowship of the Faculty of Biology and Medicine from the University of Lausanne, Switzerland.

SUPPLEMENTARY INFORMATION

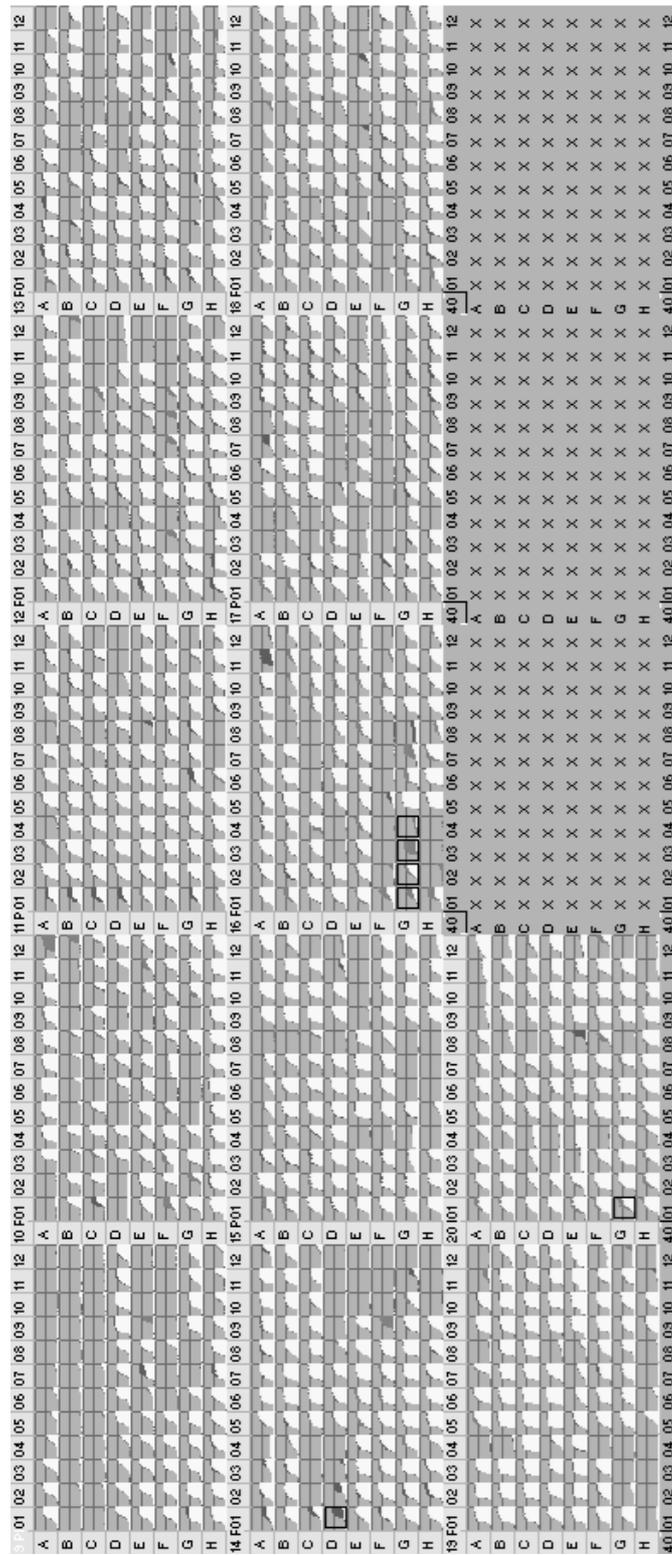


Figure S1. BIOLOG chemical sensitivity analysis of *P. putida* UWC1 (ICEcIc) and *P. putida* UWC1 (ICEcIcΔmfsABC). Each panel of the 96-well plates represents the consensus respiration kinetics (reported by monitoring the reduced tetrazolium dye over time) of *P. putida* UWC1 (ICEcIc) in red, *P. putida* UWC1 (ICEcIcΔmfsABC) in green and the overlay of the two strains in yellow. Boxed panels indicate conditions to which the two strains responded differentially in a reproducible and significant manner. Boxed panels include D01 on plate 14 (cadmium chloride), G01-G04 on plate 16 (chromium III chloride), G01 on plate 20 (Captan).

REFERENCES

1. Misra R, Bavro VN. *Assembly & transport mechanism of tripartite drug efflux systems*. *Biochim Biophys Acta* 2009, 1794:817.
2. Grkovic S, Brown MH, Skurray RA. *Regulation of bacterial drug export systems*. *Microbiol Mol Biol Rev* 2002, 66:671–701.
3. Cases I, de Lorenzo V, Ouzounis CA. *Transcription regulation and environmental adaptation in bacteria*. *Trends Microbiol* 2003, 11:248–253.
4. Ramos JL, Martínez-Bueno M, Molina-Henares AJ, Terán W, Watanabe K, Zhang X, Gallegos MT, Brennan R, Tobes R. *The TetR family of transcriptional repressors*. *Microbiol Mol Biol Rev* 2005, 69:326–356.
5. Shimada T, Hirao K, Kori A, Yamamoto K, Ishihama A. *RutR is the uracil/thymine-sensing master regulator of a set of genes for synthesis and degradation of pyrimidines*. *Mol Microbiol* 2007, 66:744–757.
6. Quiñones B, Pujol CJ, Lindow SE. *Regulation of AHL production and its contribution to epiphytic fitness in Pseudomonas syringae*. *Mol Plant Microbe Interact* 2004, 17:521–531.
7. Hillen W, Berens C. *Mechanisms underlying expression of Tn10 encoded tetracycline resistance*. *Annu Rev Microbiol* 1994, 48:345–369.
8. Ma D, Alberti M, Lynch C, Nikaido H, Hearst JE. *The local repressor AcrR plays a modulating role in the regulation of acrAB genes of Escherichia coli by global stress signals*. *Mol Microbiol* 1996, 19:101–112.
9. Grkovic S, Brown MH, Roberts NJ, Paulsen IT, Skurray RA. *QacR is a repressor protein that regulates expression of the Staphylococcus aureus multidrug efflux pump QacA*. *J Biol Chem* 1998, 273:18665–18673.
10. Terán W, Felipe A, Segura A, Rojas A, Ramos J-L, Gallegos M-T. *Antibiotic-dependent induction of Pseudomonas putida DOT-T1E TtgABC efflux pump is mediated by the drug binding repressor TtgR*. *Antimicrob Agents Chemother* 2003, 47:3067–3072.
11. Rosenfeld N, Bouchier C, Courvalin P, Périchon B. *Expression of the resistance-nodulation-cell division pump adeIJK in Acinetobacter baumannii is regulated by AdeN, a TetR-type regulator*. *Antimicrob Agents Chemother* 2012, 56:2504–2510.
12. Duque E, Segura A, Mosqueda G, Ramos JL. *Global and cognate regulators control the expression of the organic solvent efflux pumps*

- TtgABC and TtgDEF of Pseudomonas putida*. Mol Microbiol 2001, 39:1100–1106.
13. Terán W, Krell T, Ramos JL, Gallegos M-T. *Effector-repressor interactions, binding of a single effector molecule to the operator-bound TtgR homodimer mediates derepression*. J Biol Chem 2006, 281:7102–7109.
 14. Murray DS. *Crystal structures of QacR-diamidine complexes reveal additional multidrug-binding modes and a novel mechanism of drug charge neutralization*. J Biol Chem 2004, 279:14365–14371.
 15. Orth P, Schnappinger D, Hillen W, Saenger W, Hinrichs W. *Structural basis of gene regulation by the tetracycline inducible Tet repressor-operator system*. Nat Struct Biol 2000, 7:215–219.
 16. Ahn SK, Cuthbertson L, Nodwell JR. *Genome context as a predictive tool for identifying regulatory targets of the TetR family transcriptional regulators*. Plos One 2012, 7:e50562.
 17. Rodríguez-Rojas A, Rodríguez-Beltrán J, Couce A, Blázquez J. *Antibiotics and antibiotic resistance: a bitter fight against evolution*. Int J Med Microbiol 2013.
 18. Van Houdt R, Monchy S, Leys N, Mergeay M. *New mobile genetic elements in Cupriavidus metallidurans CH34, their possible roles and occurrence in other bacteria*. Antonie Van Leeuwenhoek 2009, 96:205–226.
 19. Norberg P, Bergström M, Jethava V, Dubhashi D, Hermansson M. *The IncP-1 plasmid backbone adapts to different host bacterial species and evolves through homologous recombination*. Nat Commun 2011, 2:268.
 20. Sentchilo V, Czechowska K, Pradervand N, Minoia M, Miyazaki R, Van Der Meer JR. *Intracellular excision and reintegration dynamics of the ICEclc genomic island of Pseudomonas knackmussii sp. strain B13*. Mol Microbiol 2009, 72:1293–1306.
 21. Ravatn R, Zehnder AJB, van der Meer JR. *Low-frequency horizontal transfer of an element containing the chlorocatechol degradation genes from Pseudomonas sp. strain B13 to Pseudomonas putida F1 and to indigenous bacteria in laboratory-scale activated-sludge microcosms*. Appl Environ Microbiol 1998, 64:2126–2132.
 22. Müller TA, Werlen C, Spain J, Van Der Meer JR. *Evolution of a chlorobenzene degradative pathway among bacteria in a contaminated groundwater mediated by a genomic island in Ralstonia*. Environ Microbiol 2003, 5:163–173.
 23. Chain PS, Deneff VJ, Konstantinidis KT, Vergez LM, Agulló L, Reyes VL, Hauser L, Córdova M, Gómez L, González M. *Burkholderia*

- xenovorans* LB400 harbors a multi-replicon, 9.73-Mbp genome shaped for versatility. Proc Natl Acad Sci 2006, 103:15280–15287.
24. Gaillard M, Vallaeys T, Vorhölter FJ, Minoia M, Werlen C, Sentchilo V, Pühler A, van der Meer JR. *The clc element of Pseudomonas sp. strain B13, a genomic island with various catabolic properties.* J Bacteriol 2006, 188:1999–2013.
 25. Reinhard F, Miyazaki R, Pradervand N, van der Meer JR. *Cell Differentiation to “mating bodies” induced by an integrating and conjugative element in free-living bacteria.* Curr Biol 2013, 23:255–259.
 26. Sentchilo V, Ravatn R, Werlen C, Zehnder AJB, Van Der Meer JR. *Unusual integrase gene expression on the clc genomic island in Pseudomonas sp. strain B13.* J Bacteriol 2003, 185:4530–4538.
 27. Gaillard M, Pradervand N, Minoia M, Sentchilo V, Johnson DR, van der Meer JR. *Transcriptome analysis of the mobile genome ICE_{clc} in Pseudomonas knackmussii B13.* BMC Microbiol 2010, 10:153.
 28. Krell T, Terán W, Mayorga OL, Rivas G, Jiménez M, Daniels C, Molina-Henares A-J, Martínez-Bueno M, Gallegos M-T, Ramos J-L. *Optimization of the palindromic order of the TtgR operator enhances binding cooperativity.* J Mol Biol 2007, 369:1188–1199.
 29. Wenzel M, Lang K, Günther T, Bhandari A, Weiss A, Lulchev P, Szentgyörgyi E, Kranzusch B, Göttfert M. *Characterization of the flavonoid-responsive regulator FrrA and its binding sites.* J Bacteriol 2012, 194:2363–2370.
 30. Hammar P, Leroy P, Mahmutovic A, Marklund EG, Berg OG, Elf J. *The lac repressor displays facilitated diffusion in living cells.* Science 2012, 336:1595–1598.
 31. Haigler BE, Wallace WH, Spain JC. *Biodegradation of 2-nitrotoluene by Pseudomonas sp. strain JS42.* Appl Environ Microbiol 1994, 60:3466–3469.
 32. Ryan MP, Pembroke JT, Adley CC. *Novel Tn4371-ICE like element in Ralstonia pickettii and genome mining for comparative elements.* BMC Microbiol 2009, 9:242.
 33. Lechner M, Schmitt K, Bauer S, Hot D, Hubans C, Levillain E, Loch C, Lemoine Y, Gross R. *Genomic island excisions in Bordetella petrii.* BMC Microbiol 2009, 9:141.
 34. Chuanchuen R, Karkhoff-Schweizer RR, Schweizer HP. *High-level triclosan resistance in Pseudomonas aeruginosa is solely a result of efflux.* Am J Infect Control 2003, 31:124–127.
 35. Chuanchuen R, Beinlich K, Hoang TT, Becher A, Karkhoff-Schweizer RR, Schweizer HP. *Cross-resistance between triclosan and antibiotics*

- in Pseudomonas aeruginosa is mediated by multidrug efflux pumps: Exposure of a susceptible mutant strain to triclosan selects nfxB mutants overexpressing MexCD-OprJ.* Antimicrob Agents Chemother 2001, 45:428–432.
36. Mima T, Joshi S, Gomez-Escalada M, Schweizer HP. *Identification and Characterization of TriABC-OpmH, a triclosan efflux pump of Pseudomonas aeruginosa requiring two membrane fusion proteins.* J Bacteriol 2007, 189:7600–7609.
 37. Nelson K, Weinel C, Paulsen I, Dodson R, Hilbert H, Martin do Santos V, Fouts D, Gill S, Pop M. et al. *Complete genome sequence and comparative analysis of the metabolically versatile Pseudomonas putida KT2440.* Environ Microbiol 2002, 4:799–808.
 38. Gerhardt P, Murray R, Costilow R, Nester E, Wood W, Krieg N, Briggs Phillips G. *Manual of methods for general bacteriology.* 1981.
 39. Sambrook JR. *Molecular cloning: a laboratory manual*, third edn. 2001.
 40. Miyazaki R, van der Meer JR. *A dual functional origin of transfer in the ICEclc genomic island of Pseudomonas knackmussii B13.* Mol Microbiol 2011, 79:743–758.
 41. Choi K-H, Gaynor JB, White KG, Lopez C, Bosio CM, Karkhoff-Schweizer RR, Schweizer HP. *A Tn7-based broad-range bacterial cloning and expression system.* Nat Meth 2005, 2:443–448.
 42. Wayne P. *Clinical and Laboratory Standards Institute. Performance standards for antimicrobial susceptibility testing; approved standard.* Sixteenth informational supplement. Document M100-S16. 2006.
 43. Dubey AK, Baker CS, Suzuki K, Jones AD, Pandit P, Romeo T, Babitzke P. *CsrA regulates translation of the Escherichia coli carbon starvation gene, cstA, by blocking ribosome access to the cstA transcript.* J Bacteriol 2003, 185:4450–4460.
 44. Bi D, Xu Z, Harrison EM, Tai C, Wei Y, He X, Jia S, Deng Z, Rajakumar K, Ou H-Y. *ICEberg: a web-based resource for integrative and conjugative elements found in Bacteria.* Nucleic Acids Res 2011, 40:D621–D626.
 45. Abbott JC, Aanensen DM, Rutherford K, Butcher S, Spratt BG. *WebACT--an online companion for the Artemis comparison tool.* Bioinformatics 2005, 21:3665–3666.
 46. Carver TJ, Rutherford KM, Berriman M, Rajandream M-A, Barrell BG, Parkhill J. *ACT: the Artemis comparison tool.* Bioinformatics 2005, 21:3422–3423.

CHAPTER 5

A UNIQUE OPERON OF THREE TRANSCRIPTIONAL REGULATORS CONTROLS ICE_{CLC} HORIZONTAL GENE TRANSFER IN *PSEUDOMONAS KNACKMUSSII* B13

Nicolas Pradervand, Sandra Sulser, Ryo Miyazaki, Veronica Bellini, Jan
Roelof van der Meer

Key words : tetR, MarR, LysR, HGT, LGT, integrative and conjugative
elements, MGE, GEI, genomic islands

ABSTRACT

ICE*clc* is a mobile genetic element occurring in *Pseudomonas knackmussii* B13, which is capable to transfer itself by conjugation to a variety of proteobacterial genera. Transfer is initiated from specialized "transfer competent" donor cells, which form at low proportions (3-5%) in populations in stationary phase, particularly when having been cultured on 3-chlorobenzoate. The regulatory mechanisms which control the proportion of transfer competent donor cells in culture are not well understood. In this study, we report the discovery and characterization of a locus on ICE*clc* consisting of three consecutive regulatory genes, which play a major role in controlling the level of transfer competence. The cluster is formed by a TetR-type regulator (named MfsR), a MarR and a LysR-type transcriptional activator, named TciR, and its expression is under control of MfsR. Deletion of *mfsR* on ICE*clc* led to a drastic increase of the population of transfer competent cells almost up to 100%, and ICE*clc* transfer rates approaching 1 per donor. In contrast, deletion of *tciR* and of the *marR*-type gene reduced ICE*clc* transfer by three to five orders of magnitude, suggesting that they are an important factor for activating the initiation of ICE*clc* transfer competence.

INTRODUCTION

Comparisons between ever-increasing numbers of sequenced genomes reveal the large extent to which prokaryotic genomes have undergone horizontal gene transfer (HGT). In the past decade, integrative and conjugative elements (ICEs) have been recognized as a new class of mobile genetic elements that contribute to HGT within the microbial world [1–3]. Despite these observations, we know still relatively little on the molecular mechanisms of the diverse ICE life styles. Functional and mobilizable ICEs share the common feature of existing either as an integrated form or as an excised circular form in which both ends are connected [4]. The integrated form occurs at specific sites in the host's chromosome (often in genes for tRNA), and is vertically transmitted to daughter cells by chromosome replication and segregation. The circular form is a transient plasmid-like molecule resulting from the excision from the chromosome, and allows the conjugation machinery to transfer the ICE to new recipient cells. Autonomous plasmid-like replication of the circular form may occur [5], but is not required for the transfer itself. Consequently, the transition from integrated to excised form is the critical decisive step for initiation of ICE self-transfer.

The regulatory mechanisms for control of the switch from integrated to excised state vary widely among different ICE. In several ICE, this switch is the consequence of a cascade of a variety of regulatory factors, such as PhrI/RapI in *ICEBs1*, SetR/SetCD in *SXT*, KorSA/Pra in *pSAM2* and QseM/TraR in *ICEMISym^{R7A}* [6–13].

In pSAM2, the repressor KorSA not only represses itself but also the activator *pra*, which in turn, upregulates the expression of the integrase, the excisionase and the replicase, thus leading to the excision of pSAM2 [9–11]. To this date, the inhibitor of KorSA is not known and therefore the induction conditions of pSAM2 HGT remains mysterious.

Control of excision and HGT in ICEMISym^{R7A} is achieved by the regulator TraR, an acylhomoserine lactone dependent autoactivator, which stimulates indirectly the excision and transfer of ICEMISym^{R7A} via the *msi171* and *msi172* genes [12, 14]. However, the allosteric repressor QseM inhibits TraR and thus prevents HGT [13]. In ICEMISym^{R7A}, QseM itself is regulated at the transcriptional level by increasing concentrations of QseC, which supposedly reaches highest levels only in a fraction of the total population, leading to the repression of QseM and ultimate induction of HGT [13]. RapI is the main activator of HGT of ICEBs1, which stimulates ImmA to proteolyse ImmR, the repressor of the excisionase gene [6]. RapI is repressed in exponential phase by the transition phase regulator AbrB, thus preventing HGT to occur during cell division period and low density of cells. In stationary phase, if many cells already contain ICEBs1, RapI is inhibited by the pentapeptide PhrI, whose biosynthesis is encoded on ICEBs1. In SXT, HGT is triggered by the SOS response, especially after exposition to DNA-damaging antibiotics like ciprofloxacin [7]. RecA stimulates SetR to self-proteolyse, freeing the genes *setC* and *setD* from the transcriptional repression by SetR [15]. Subsequently, SetC and SetD activates several genes encoding essential factors in the HGT machinery, like the integrase, the excisionase and the putative T4SS [8, 16].

Interestingly, many ICE only transfer at very low frequencies, suggesting that the regulatory cascades operate in a bistable manner, forcing a small subset of cells in a subpopulation to follow the path of ICE excision and transfer. This bistability is most pronounced and well-studied for a model ICE named *ICEclc* in *Pseudomonas* [17, 18]. *ICEclc* is originally found in two copies in *Pseudomonas knackmussii* B13 and is member of a wider family of *ICEclc*-like entities distributed among many proteobacterial species. *ICEclc* is integrated in the 3'-part of genes for tRNA^{Gly} but can excise as a result of action of the *IntB13* integrase, the gene for which lays close to one of its ends (Figure 1A). Expression of *intB13* expression in the integrated form is under control of the promoter P_{int} , which by single cell reporter gene analysis was shown to become active only in 3-5% of a bacterial population during stationary phase [17]. Direct single cell visualization further confirmed that only cells which express P_{int} above a threshold are capable of transferring *ICEclc* to new recipients, a bistable state which we recently named "transfer competence" (tc) [19]. Expression of *intB13* is dependent on a variety of factors, most notably a gene named *inrR*, which is present near the other end of *ICEclc* (Figure 1A), and which itself is also bistably expressed in the same cell [17]. Both *inrR* and *intB13* expression are dependent on the stationary phase sigma factor RpoS, the cellular abundance of which we could recently show co-determines the proportion of cells activating P_{int} and P_{inR} [20]. Although RpoS and InrR have been implicated in activating *ICEclc* excision and transfer, it has been recognized that they are not sufficient, and one or more additional factors or regulatory cascades may be necessary for the tc state to develop.

In this study, we report of a new and unexpected level of regulatory control on ICE*clc* transfer. This regulation is encoded by a small operon, containing three consecutive transcriptional regulators. We used random transposon mutagenesis to screen for ICE*clc* transfer-deficient phenotypes. These studies were then further refined by creation of sequence-specific deletion mutants and using a combination of transfer assays, microarrays, RT-PCR and reporter gene-based single cell fluorescence microscopy.

RESULTS

Transposon mutants in *mfsR* are impaired in ICE*clc* transfer. In order to discover ICE*clc*-located factors involved in its self-transfer, a library of *P. knackmussii* B13 mutants was generated by using random Tn5 mutagenesis [21]. The pool of B13 mutants, which likely contains Km^R-gene insertions in both copies of ICE*clc*, was next used as mass donor for conjugation of ICE*clc* to *Pseudomonas putida* UWC1 (Fig. S1). We hypothesized that mutant ICE*clc* with insertions in genes implicated in self-transfer could still be transferred to UWC1 if the second copy of ICE*clc* in the same B13 donor cell would be intact and complement the mutant copy. By selecting for Km-resistant *P. putida* transconjugants, we thus recovered a new library of *P. putida* with likely only single copy mutant ICE*clc*. A total of 1920 mutants of this library was mated clone-by-clone in a 96-well cell suspension-based conjugation assay with a second *P. putida* recipient (Fig. S1). For those donors for which no detectable mutant ICE*clc* transfer occurred, we mapped the location of the Km^R-gene insertion on ICE*clc*. Interestingly, we found 18 clones which had insertions in an ICE*clc* gene previously named *mfsR* (Figure 1B). Apart from a mutant with an insertion in *intB13*, no other mutants with impaired ICE*clc* transfer were found in this screening. At least four different Km^R-gene insertions had occurred in *mfsR*, at ICE*clc* nucleotide positions 19033, 18758, 18730 and 18618 (Figure 1, accession number: AJ617440.2).

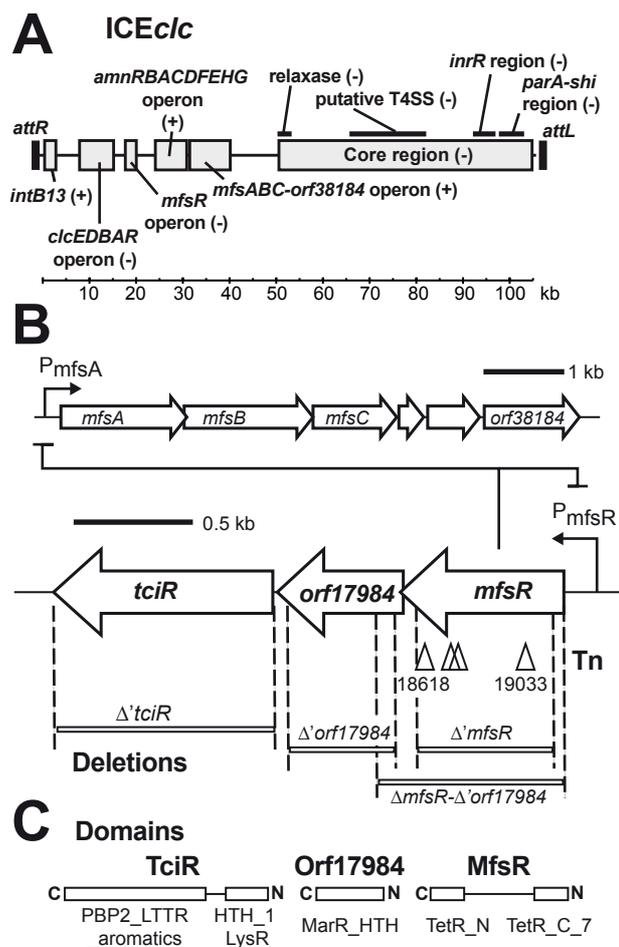


Figure 1. Schematic overview of ICE clc and the location of the genes relevant to this study. A) ICE clc integrated stage with the two flanking 18-bp repeats (as black rectangles). Previously determined gene regions indicated at their approximate location. Important functional regions are depicted as grey boxes accompanied by legends. + or -, indicate the orientation of the coding strand for the respective gene(s) (the + strand being the one of *intB13*). *clcEDBAR* operon, chlorocatechol degradation genes. *amnRBACDFEHG* operon, 2-aminophenol degradation genes. *mfsABC* genes, efflux pump. Scale bar in kilobasepairs (kb). B) Detail of the *mfsR* and *mfsABC* operons. Arrows, predicted open reading frames (the right-to-left orientation indicates the minus strand). Triangles, positions of the Tn5-mediated kanamycin gene insertions (nucleotide positions indicated below, according to the AJ617740.2 numbering). Regions deleted in this study are displayed as white bars with the names of the mutations noted below. Note that the *mfsR* gene product represses the expression of the P_{mfsR} and P_{mfsA} promoters. C) BlastP-predicted domains are represented for each of the three regulatory genes in the *mfsR* operon. C and N, carboxy and amino terminus, respectively.

Frequencies of ICE*clc* transfer from donor strains *P. putida* UWC1-2961 (Km^R-gene insertion at 19033) and UWC1-2962 (insertion at 18618) in a filter-based conjugation assay were 10⁴-fold and 10³-fold lower than of a *P. putida* with one integrated wild-type ICE*clc* copy, respectively (Figure 2A). From this result we concluded that insertional inactivation of *mfsR* or polar effects on downstream-located genes were responsible for impaired ICE*clc* transfer.

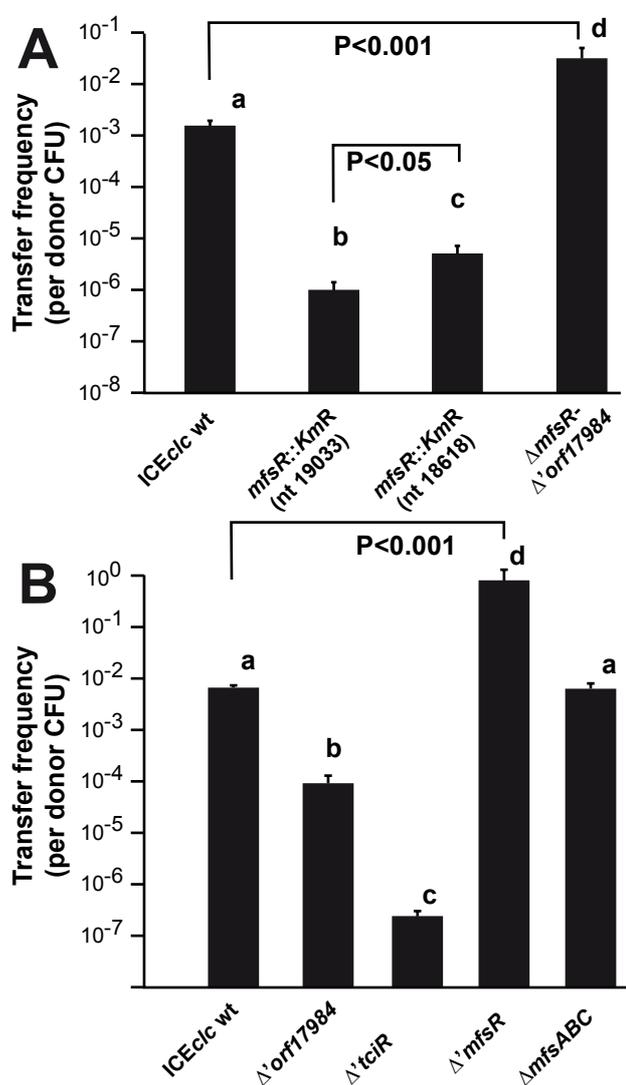


Figure 2. ICE*clc* transfer frequencies of *P. putida* UWC1 with different ICE*clc* genotypes. A) Effects of transposon insertion mutants in *mfsR* and a *mfsR-orf17984* deletion compared to wild-type. B) Effects of *tciR*, *orf17984*, *mfsR* or *mfsABC* deletions compared to wild-type. Bars show mean transfer frequencies as transconjugant (growing on 3CBA, Km- and Gm-resistant) per donor CFU (colony forming unit) from triplicates and the corresponding standard

deviations. Small cap letters indicate statistical significant groups in an Anova with post hoc Tukey test (at the indicated P-values).

The ICE*clc*-encoded *mfsR* operon contains three consecutive transcriptional regulators. The *mfsR* gene is located between the genes for chlorocatechol degradation (*clc*) and those for 2-aminophenol degradation (*amn*) (Figure 1A). *mfsR* is the first open reading frame of a putative operon of three transcriptional regulators, (*mfsR*), *orf17984* and *orf17162*, which we renamed *tc*i*R* (transfer competence inducer regulator) in anticipation of the results described below (Figure 1B). *mfsR* encodes a TetR-like regulator harboring helix-turn-helix motifs TetR_N and TetR_C_7 (pfam0040 and pfam14246, respectively, see Figure 1C). Previously, we demonstrated that *mfsR* is responsible for the repression of its own promoter (P_{mfsR}) and that of a putative efflux system, *mfsABC*, also encoded on ICE*clc* (Pradervand et al., unpublished; Figure 1B). Since MfsR is an autorepressor, this would imply that the two additional downstream putative regulatory genes are under MfsR control. The *orf17984* gene overlaps with the end of the *mfsR* open reading frame by 4 bp and encodes a putative regulator of the MarR family (smart00347 HTH_MARR motif). The last gene of this cluster, *tc*i*R*, starts 24 bp downstream of the stop codon of *orf17984* and is predicted to code for a LysR-type transcriptional regulator harboring a N-terminal HTH_1 motif (pfam00126) and a C-terminal substrate-binding domain (PBP2_LTTR_aromatics_like; cd08414). Reverse transcription of UWC1 (ICE*clc*) RNA isolated from exponential phase-grown cells, followed by specific PCR amplification confirmed that the three genes are transcribed on the same mRNA which ends downstream of *tc*i*R* (data not shown). This would

imply that *mfsR-orf17984-tciR* indeed form a single polycistronic unit, the expression of which is under the control of MfsR.

Effects of *tciR*, *orf17984* and *mfsR* deletions on ICE*clc* transfer frequencies. In order to investigate the role of the three regulators on ICE*clc* transfer, their open reading frames were each individually and partially deleted in separate strains, namely *P. putida* UWC1 (ICE*clc*- Δ '*mfsR*), UWC1 (ICE*clc*- Δ '*orf17984*), UWC1 (ICE*clc*- Δ '*mfsR*- Δ '*orf17984*) and *P. putida* UWC1 (ICE*clc*- Δ '*tciR*) (Figure 1B). ICE*clc* transfer frequencies in plate-mating assays with a gentamicin-resistant *P. putida* UWC1 as recipient were 10²- and 10⁴-fold lower for UWC1 donors with ICE*clc* having internal deletions in *orf17984* and *tciR*, respectively, compared to intact ICE*clc* (Figure 2B). No statistically significant difference was found for ICE*clc* transfer frequencies between *P. putida* UWC1 (ICE*clc*- Δ '*mfsABC*) and UWC1 (ICE*clc* wild-type, Figure 2B). In contrast, mutant ICE*clc* elements from *P. putida* UWC1 (ICE*clc*- Δ '*mfsR*) and UWC1 (ICE*clc*- Δ '*mfsR*- Δ '*orf17984*) transferred with 100- and 10-fold higher frequencies than wild type ICE*clc*, respectively (Figure 2A, B). These results suggested that *tciR* encodes an activator for ICE*clc* transfer. In absence of MfsR, expression of the downstream genes *orf17984* and *tciR* is higher, which results in increased ICE*clc* transfer. The effect of the transposon insertions in *mfsR* (i.e., lower ICE*clc* transfer rates, Figure 2A) may therefore originate from a polar effect on *orf17984-17162* expression.

***tciR* encodes a major activator of the ICE*clc* core region.** A previous study unveiled the global bipartite behavior of the ICE*clc*

transcriptome according to growth phase [22]. Indeed, when *P. putida* UWC1 (ICE*clc*) cells are growing exponentially on 3CBA, expression from the genes in the ICE*clc* core region (roughly the second half of ICE*clc*) plus the integrase *intB13* is silent, whereas they are highly transcribed when cells are in stationary phase (Figure 3A). *P. putida* with mutant ICE*clc* lacking either *orf17984* or *tcIR* strongly diminished expression in the core region and the integrase genes in stationary phase when compared to wild type (Figure 3C, Fig. S2). This suggested that these mutants are not capable of activating ICE*clc* core region expression, which would corroborate the results from ICE*clc* transfer experiments (Figure 2B). Both mutants behaved very similarly (Fig. S2), and their ICE*clc* gene expression is indistinguishable from that of wild type ICE*clc* in exponential phase (Fig. S2A, B). In addition, expression of the *mfsR-orf17984-tcIR* cluster is not statistically significantly different in the *tcIR* or *orf17984* deletion mutants (Figure 4A-C). In contrast, the remarkable effect of deleting or interrupting *mfsR* is that expression from the ICE*clc* core genes already starts in exponentially growing cells [*P. putida* UWC1 (ICE*clc*- Δ '*mfsR*) and (ICE- Δ '*mfsR*- Δ '*orf17984*)], and remains even slightly higher in stationary phase than in wild-type ICE*clc* (Figure 3B, Fig. S3). In addition, *mfsR* mutants derepress the *mfsABC-orf38184* operon (Figure 3B, Fig. S3), as noted previously (Pradervand et al, unpublished), and the *mfsR-orf17984-tcIR* cluster itself, both in exponential and stationary phase cells (Figure 4A, D). Expression analysis of the *P. putida* UWC1 strain with the Km^R-gene insertion in *mfsR* (ICE*clc*-*mfsR*::Km^R, position 19033) showed a combination of effects (Figure 3D, Fig. S3). Similar to *mfsR* mutants, expression from *mfsABC* is higher than wild-type, whereas that of the core region genes and

intB13 is dramatically lower than wild-type in stationary phase cells (Figure 3D). Expression of the *mfsR-orf17984-tciR* region itself is mixed; the first 160 bp of *mfsR*, upstream of the Km^R-gene insertion are higher expressed than in wild-type cells (as seen by 1 probe at position 19191 on Figure 3D), but the downstream genes *orf17984* and *tciR* are lower expressed compared to wild-type (Figure 3D). This indicates, therefore, that insertion of the Km^R-gene caused a polar effect on expression of *orf17984* and *tciR*, which explains the strongly diminished expression of the ICE*clc* core genes in stationary phase. Since gene expression from ICE*clc* is similar in mutants lacking *mfsR* without or with a small part of *orf17984* (Fig. S3), we conclude that it is the LysR-type regulator encoded by *tciR* which is the main activator for ICE*clc* core gene expression. The similar effect of the *orf17984* and *tciR* deletions may thus be caused by again a polar influence of deleting *orf17984* on expression of *tciR* itself (although this is not evident from transcript analysis; see, e.g., Figure 4A). Interestingly, the magnitude difference of ICE*clc* expression in the non-coding strand was much more pronounced for stationary phase *P. putida* UWC1 (ICE*clc*- Δ '*mfsR*) compared to wild-type, than for UWC1 (ICE*clc*- Δ '*mfsR*- Δ '*orf17984*) or UWC1 (ICE*clc*-*mfsR*::Km^R, Fig. S3 right panels, + strand). This suggests that the *orf17984* gene product may have some role in amplifying expression from the ICE*clc* genes.

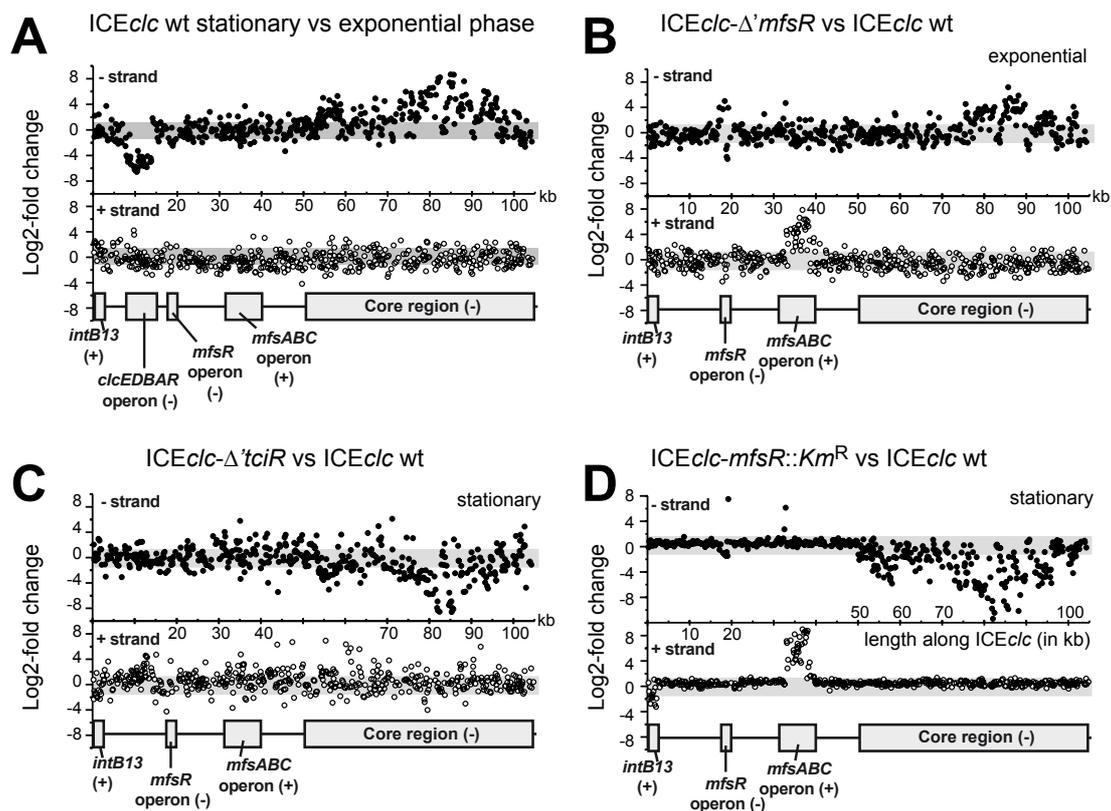


Figure 3. Differential expression of the ICEclc gene region from micro-array data in selected mutant ICEclc versus wild type in *P. putida* UWC1. A) Differential expression of the ICEclc region between stationary and exponential phase cells of wild type *P. putida* UWC1 (ICEclc). B) Differential expression of the ICEclc region between the *mfsR* deletion mutant and wild-type, in exponentially growing cells. C) Comparison of the *tciR* deletion mutant and wild-type, in stationary phase cells. D) Comparison of the *mfsR* transposon insertion mutant versus wild-type, in stationary phase cells. Dots indicate the log₂-fold change of hybridization signal per microarray probe in the comparison, plotted at their distance along the ICEclc sequence (X-axis; in kb). A scheme of ICEclc is redrawn at the bottom of each section, with regions of interest as grey boxes (+ or - indicate the DNA strand on which the region is encoded). Separate displays indicate expression differences on the plus- or the minus-strand. Grey bars in the background indicate the two-fold cut-off level. For a complete set of microarray results, see Supplementary Figures S2 and S3.

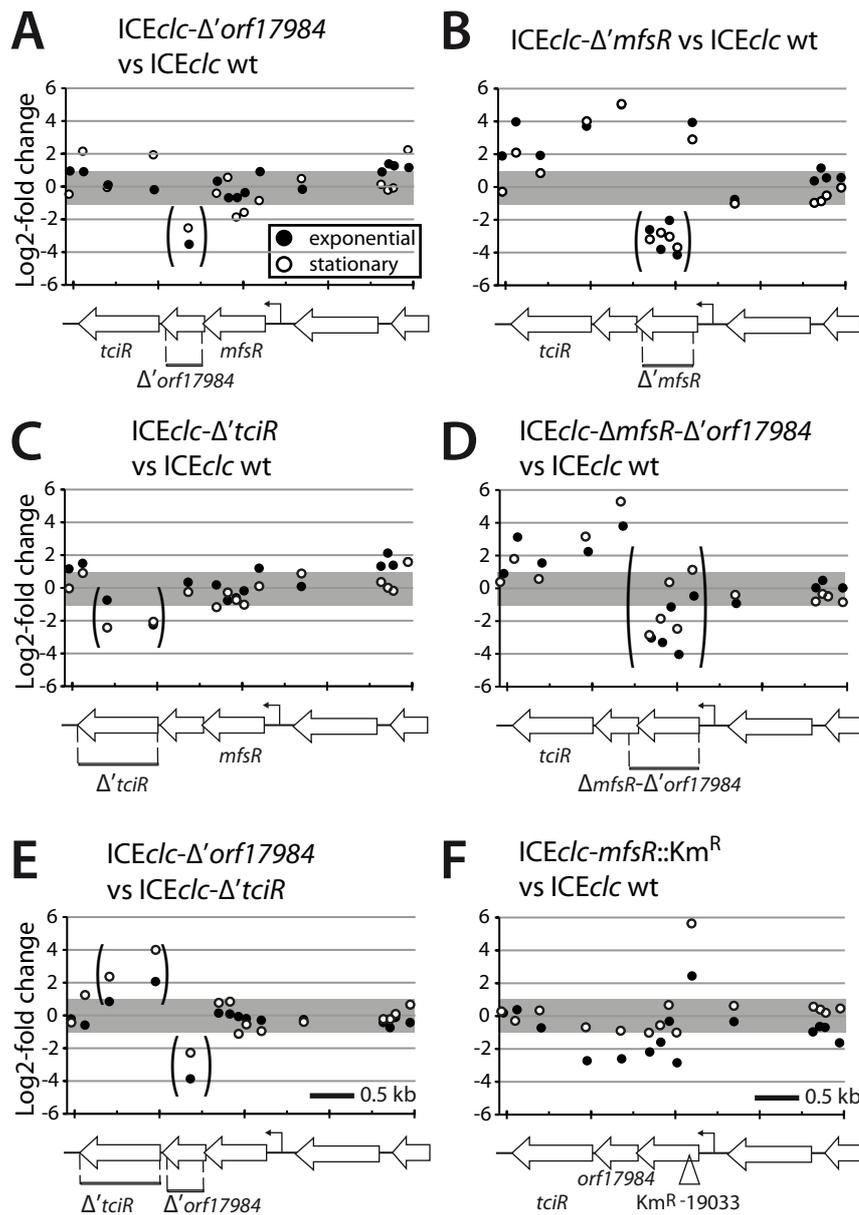


Figure 4. Detailed view on the differential expression of the *mfsR* operon in *P. putida* ICEclc wild-type or mutants. A) *Orf17984* deletion mutant versus wild-type. B) *mfsR* deletion mutant versus wild type. C) *orf17984* deletion mutant versus wild-type. D) *mfsR-orf17984* deletion mutant versus wild-type. E) *orf17984* versus *tciR* deletions. F) *mfsR*-transposon insertion mutant versus wild-type. Panels show log₂-fold change of expression level per microarray probe in this region of ICEclc for exponential (dark dots) and stationary phase cells (white dots). Genetic map of the region drawn at the bottom of each section for clarity. Arrows represent genes, deleted regions are indicated by stippled bars and corresponding probes are within brackets.

***tciR* positively regulates *inrR* and *intB13* expressions.** In order to verify results from transfer assays and microarray data, a double promoter-

reporter construct, carrying P_{int} -*gfp* and P_{inR} -*echerry* was inserted in *P. putida* UWC1 (ICE*clc*), *P. putida* UWC1 (ICE*clc*- Δ' *mfsR*), *P. putida* UWC1 (ICE*clc*- Δ' *orf17984*) and *P. putida* UWC1 (ICE*clc*- Δ' *tcIR*). P_{int} and P_{inR} are the respective promoters for the integrase gene *intB13* (integrated form) and the integrase activator gene *inrR*. Previous studies showed that both promoters are active only in a small subpopulation of transfer competent cells [17, 20]. Consistent with previous data, the subpopulation of *P. putida* UWC1 (ICE*clc*) wild-type cells expressing P_{int} and P_{inR} in stationary phase represented a few percent (Figure 5). In contrast, deletion in *mfsR* resulted in almost all cells expressing P_{int} and P_{inR} -promoters (Figure 5). Most interestingly, expression of both promoters in *P. putida* UWC1 (ICE*clc*- Δ' *mfsR*) occurred much earlier than in wild-type cells (Figure 5B). Conversely, *P. putida* UWC1 (ICE*clc*- Δ' *tcIR*) and *P. putida* UWC1 (ICE*clc*- Δ' *orf17984*) did not produce any detectable P_{int} or P_{inR} -expressing cells, neither in exponential nor in stationary phase (Figure 5), in accordance with absence of ICE*clc* core gene activation on microarrays and lower transfer frequencies.

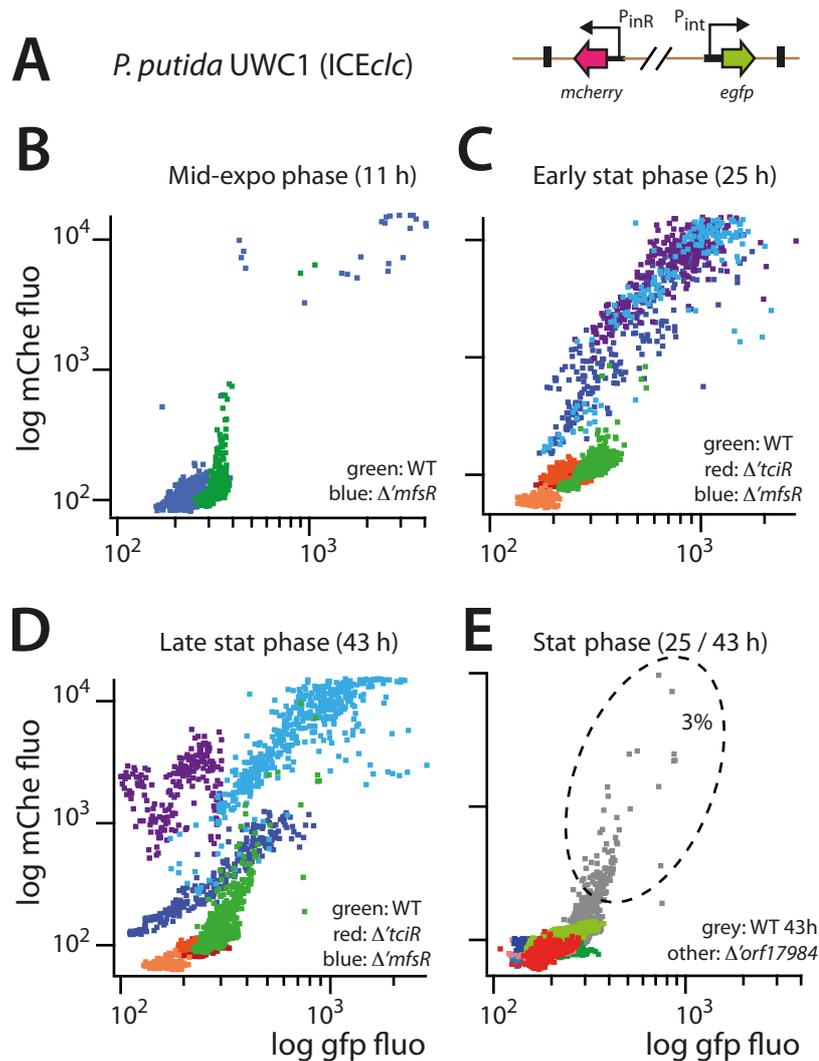


Figure 5. Effect of mutations in the *mfsR* region on the expression of the P_{int} - and P_{inR} -promoters of ICEclc in *P. putida* UWC1. A) Schematic representation of the single copy mini-transposon containing the P_{int} -*egfp* and P_{inR} -*echerry* fusions introduced in the *P. putida* strains. B) Scatter plot of GFP and echerry fluorescence in single cells of *P. putida* UWC1 (ICEclc) wild-type (green dots) and the *mfsR* deletion mutant (blue dots), exponentially growing in microcolonies 11 h after inoculation. C) Expression of GFP and echerry in early stationary phase cells in microcolonies of wild-type (green dots), *mfsR* deletion mutant (blue tints; three clones with different mini-transposon insertion positions) and *tciR* deletion mutant (red tints; three clones). D) as C, in late stationary phase. E) Expression in the *orf17984* deletion mutant in exponential and stationary phase cells in microcolonies (three clones each; red, green or blue tints) versus wild-type (grey dots). Note as example the subpopulation of wild-type cells (dotted ellips in panel E) expressing both reporters, compared to the majority of cells in the *mfsR* deletion mutant but a complete absence of such subpopulation in the *tciR* and *orf17984* deletion mutants.

DISCUSSION

ICE*clc* has two distinctive modes of existence, the integrated form, which is transmitted vertically, and the circular form, which is horizontally transferred. Previous work demonstrated that the switch between these two states does not occur in every cell, but is rather confined to a small subpopulation of *tc* cells, that are defined by higher level of RpoS and by an additional yet unknown factor [17, 19, 20]. In the present study, we report the discovery of a cluster of three regulatory genes, which control activation of the ICE*clc* core genes and consequently, ICE*clc* transfer rates.

The cluster of regulatory genes was accidentally recovered as the main occurring transposon insertion mutant from a library of mutant ICE*clc* stored in *P. putida* UWC1 and tested for self transfer deficiency. Insertions in the first gene of this cluster (previously named *mfsR*), caused 3-4 orders of magnitude decrease in ICE*clc* transfer in filter mating assays (Figure 2). Subsequent analysis of the gene region surrounding these transposon insertions then indicated that *mfsR* is the first of a three-gene operon, followed by two others named tentatively *orf17984* and *tc*i*R* (Figure 1). Interestingly, all three genes in the cluster are predicted to encode transcription factors; *mfsR* coding for a TetR-type repressor, *orf17984* for a MarR-type and *tc*i*R* for a LysR-type regulators. We previously demonstrated that MfsR is a transcriptional repressor of its own promoter, which is located directly upstream of *mfsR*, and of a promoter for a set of genes on ICE*clc* coding for an efflux pump (*mfsABC*, Figure 1). Consequently, MfsR would be expected to exert repression on *mfsR*, *orf17984* and *tc*i*R*, and, indeed, transcript analysis showed that clones

with a deletion in *mfsR* have higher expression of the two downstream genes (Figure 4).

Our first hypothesis to explain the behaviour of the transposon insertion mutants in *mfsR* stipulated a direct implication of MfsR in regulating ICE*clc* transfer genes, since the ICE*clc* core expression was silenced in these mutants (Fig. 3). However, as complementation by an added single copy *mfsR* gene did not restore wild-type phenotype (not shown), and ICE*clc* gene expression in an *mfsR* deletion mutant turned out very different than in the transposon insertion mutant (Figure 3), we suspected implication of the downstream located genes (Figure 1). Indeed, a precise deletion of *tcjR* also caused a 10⁴-fold lower frequency of ICE*clc* transfer compared to wild type (Figure 2B). In addition, deletion of *tcjR* prevented the transcriptional activation of the ICE*clc* core region in a similar but not identical way as the transposon insertion mutants in *mfsR* (Figure 3). The similarity of phenotypes between *P. putida* UWC1 (ICE*clc*-Δ*tcjR*) and *P. putida* UWC1 (ICE*clc*-*mfsR*::Km^R₁₉₀₃₃) therefore likely results from a polar effect of the insertion of the Km-resistance gene within *mfsR* on transcription of *orf17984* and *tcjR*. This can be appreciated in detail in Figure 4F, where the region downstream of the insertion site is not transcribed, but upstream is overexpressed because of the absence of MfsR as autorepressor. We therefore conclude at this point that the gene product of *tcjR* is an activator needed for expression of the ICE*clc* core genes; either directly as a master regulator on a variety of individual ICE*clc* core promoters, or on a first "node" in an activation cascade. The most surprising effect of mutations in the three regulatory genes, however, was that deletions in *mfsR* led to a complete deregulation of the ICE*clc* core

gene expression of the wild-type situation. This became obvious, first, from frequencies of ICE*clc* transfer 10-100 fold higher than wild type, and almost approximating 1 per donor cell CFU (Figure 2). Secondly, the deregulation was obvious in microarray data showing the ICE*clc* core region in the *mfsR* deletion mutants being already transcribed in exponential phase, and even slightly higher than in the wild type in stationary phase (Figure 3, Fig. S2). Finally, single cell epifluorescence microscopy analysis of strains labeled with an $P_{int}\text{-}gfp/P_{inR}\text{-}echerry$ gene cassette, showed massive reporter gene expression in the *mfsR* deletion strain compared to wild-type, whereas *tcIR* and *orf17984* deletions caused further reduction of the proportion of cells expressing the reporters (Figure 5). These results, therefore, suggest that MfsR in wild-type ICE*clc* represses *tcIR*, but when deleted, leads to constant expression of *tcIR*, which helps to initiate the cascade of ICE*clc* core gene expression.

The role of *orf17984* is less clear and yet unsolved. Deletion in *orf17984* produced the same ICE*clc* transcriptome profile as deletion in *tcIR* (Figure 3D). It also resulted in a lower transfer frequency than wild-type but not as low as the deletion in *tcIR* (Figure 2B), and produced no detectable reporter gene expression from P_{int} or P_{inR} (Figure 5). In contrast, mutants with deletions in *mfsR* or *mfsR* plus the first 116 bp of *orf17984* (Figure 1), behaved quite similar in transfer frequency (Figure 2) and showed similar ICE*clc* transcriptomes (Fig. S3), except for the magnitude of expression on the non-coding strands (Fig. S3). Possibly, therefore, the effects of *orf17984* deletion itself are also due to a polar effect on *tcIR* expression and Orf17984 somehow enhances transcriptional control of TciR.

From these results we propose the following model: MfsR prevents the induction of transfer-competence by repressing (mainly) *tcir*. TciR on its turn is needed for activation of the ICE*clc* core genes and subsequent transfer initiation. Deletion of *mfsR* basically results in all cells expressing P_{int} and P_{inR}-promoters, and all cells becoming ICE*clc* donors (Figure 2). Since in wild-type ICE*clc* only a small subset of cells activates the transfer competence program, this must mean that either repression by MfsR is inherently "sloppy" (causing a few cells to escape its control and transcribing *tcir*), or that there is a chemical ligand which specifically derepresses MfsR in a small subset of cells. The resulting TciR would then be the necessary activator to trigger ICE*clc* core expression in cells which on average have highest RpoS levels [20].

Interestingly, quasi global formation of transfer competence across all cells in the *mfsR* deletion mutant did not result in abundant cell death, as we previously described as being a side consequence of becoming transfer competent in wild-type cells [22]. This cell inhibition was caused by the *parA-shi* gene products on ICE*clc* [19]. Previous observations with the *P. putida* *mfsR* transposon insertion mutant (ICE*clc*-*mfsR*::Km^R₁₉₀₃₃) had also shown it yielded a lower proportion of both transfer competent cells and cell growth inhibition, suggesting such linkage [19]. In contrast, the phenotype of the ICE*clc* *mfsR* deletion mutant would suggest that ICE*clc* transfer and cell growth inhibition are two processes that not necessarily need to go together, although in wild-type ICE*clc* they are linked. The question is thus justified as to why control by MfsR on ICE*clc* transfer competence initiation exists, if not to prevent excessive cell growth inhibition through the ICE*clc* encoded factors Shi/ParA?

Interestingly, components of the *mfsR* operon were found to have homologues in other bacteria but not in the same configuration. Blastn and BlastP analysis, complemented by screening of the ICEBerg database [23], showed that the combination of both *mfsR* and *orf17984* as in ICE*clc* only occurs in a similar ICE*clc* of *Burkholderia xenovorans* LB400 (100% amino acid identity), in a Tn4371-related element of *Acidovorax* sp. strain JS42 (93% amino acid identity) and in *Aeromonas hydrophila* SSU (93% amino acid identity, not shown, accession number AGWR01000022.1). More distant *mfsR* homologues (58-57% amino acid identity) are also present in megaplasmids of various *Ralsonia solanacearum* species (CMR15, GMI1000, CFBP2957, PSI07, Figure S4).

On the other hand, *tcjR* proved to be much more widespread than *mfsR*, and homologues can be found in the ICE*clc*-like GI-1 and GI-6 of *Bordetella petrii* DSM12804 (Figure S5B-C), in PAGI-2 of *P. aeruginosa* strain C (Figure S5D), in ICEXcaVe85-1 of *X. campestris* pv. *vesicatoria* str. 85-106 (terminology derived from the ICEberg database), in ICEHaeULPAs1-1 of *Herminiimonas arsenicoxydans*, in ICECmeCH34-1 of *Cupriavidus metallidurans* CH34, in ICETauDSM9187-1 of *Tolumonas auensis* DSM 9187, ICEAxyA8-1 of *Achromobacter xylosoxidans* A8 and many more. Curiously, in *Acidovorax* sp. strain JS42, a *tcjR* homologue appears on an ICE*clc*-like element (ICEAciJS42), whereas - as mentioned above, *mfsR* and *orf17984* co-occur on a Tn4371-type ICE in the same strain. As depicted on Figure S5B-C, the ICE*clc*-like elements in *B. petrii* DSM 12804 harbor three close homologues to TciR, one on GI-1 (81% amino acid identity in 238/294 aa overlap) and two on GI-6. However, the configurations of those *tcjR*-like genes in *B. petrii* differ

much from that of the *mfsR* operon in ICE*clc*, thus suggesting very different expression control. In the more distant ICE*clc* relative PAGI-2, a putative BphR protein shares 79% of its aminoacid content (231/293 aa overlap) with *tcIR*, but it occurs in a very different operonic architecture.

Taken together these results suggest that TciR is a common regulator found in many ICEs, possibly regulating the transfer of all of them. However, its expression regulation in those different mobile elements is unknown but probably does not resemble that in ICE*clc*. Interestingly, insofar as known, transfer rates of the ICE*clc*-like GIs in *B. petrii* were much lower ($\sim 10^{-7}$ per donor) than that of ICE*clc* [24], and PAGI-2 transfer is undetectable [25]. This suggests that the *tcIR* homologues on those elements are (i) very poorly expressed or (ii) cannot be induced any longer, or (iii) that the promoters in the core region have diverged and do not respond anymore to the TciR homologue. Perhaps, therefore, the control of *tcIR* expression by the MfsR autorepressor is a mechanism that evolved specifically in ICE*clc* of B13 and LB400, and was selected for by the resulting efficient conjugative transfer rates. The clear genealogy of TciR to the LysR-type regulator family further suggests that it requires a cofactor or ligand to exert its regulatory control, similar to many regulators of this type [26, 27].

MATERIAL AND METHODS

Strains and culture conditions. Table 1 lists the strains used in this study. *Escherichia coli* DH5 α (Gibco Life Technologies, Gaithersburg, Md.), *E. coli* DH5 α λ pir, *E. coli* BW20767/pRL27 were cultured at 37°C on Luria-Bertani (LB) medium. *Pseudomonas* species were cultured at 30°C on LB or 21C minimal medium (MM) complemented with one of the following carbon sources: 0.5, 5, or 10 mM 3-chlorobenzoate (3CBA), 15 mM succinate, 10 mM fructose [28, 29]. Antibiotics were supplemented to the growth medium to select for maintenance of genetic constructions at the following concentrations: kanamycin (Km) 25 μ g/ml, chloramphenicol (Cm) 20 μ g/ml, rifampicin (Rif) 50 μ g/ml, nalidixic acid (Nal) 50 μ g/ml, gentamicin (Gm) 20 μ g/ml, and ampicillin (Ap) 100 μ g/ml.

Table 1. Strains used in this study and their specifications.

| Strain name | Strain collection number | Description | Reference |
|---|--------------------------|--|-----------|
| <i>E. coli</i> DH5 α | | | |
| <i>E. coli</i> DH5 α λ pir | | | |
| <i>E. coli</i> BW20767/ pRL27 | 1853 | tra+, pRL27 containing hyperactive mini-Tn5 element (<i>oriV</i> , Km ^R). | [21] |
| <i>P. knackmussii</i> B13 | 78 | Original host of ICE <i>clc</i> (2 copies). | [32] |
| <i>P. putida</i> UWC1 | 1291 | plasmid-free derivative of <i>P. putida</i> KT2440, Rif ^R , Provided by Carole Newberry, Cardiff University, Wales, | [33] |

| | | | |
|--|----------------|---|---------------|
| | | UK. | |
| <i>P. putida</i> UWC1 (Nal) | | Spontaneous Nal ^R -mutant of 1291. | This study |
| <i>P. putida</i> UWCGC | 2744 | Tn7-Ptac-echerry-Gm ^R | [30] |
| <i>P. putida</i> UWC1 (ICE <i>clc</i>) | 2737 | 1291-derivative with one ICE <i>clc</i> copy integrated into tRNA-gly 5. | [31] |
| <i>P. putida</i> UWC1 (ICE <i>clc</i>) | 2738 | 1291-derivative with one ICE <i>clc</i> copy integrated into tRNA-gly 6. | [31] |
| <i>P. putida</i> UWC1 (ICE <i>clc</i> -Km ^R ₁₉₀₃₃) | 2961 | Transposon mutant of strain 2737 with a Km ^R -gene inserted at nucleotide position 19033 in ICE <i>clc</i> . | [19] |
| <i>P. putida</i> UWC1 (ICE <i>clc</i> -Km ^R ₁₈₆₁₈) | 2962 | Transposon mutant of strain 2737 with a Km ^R -gene inserted at nucleotide position 18618 in ICE <i>clc</i> . | This study |
| <i>P. putida</i> UWC1 (ICE <i>clc</i> - Δ <i>mfsR</i> - Δ ' <i>orf17984</i>) | 3453 | Derivative of strain 2737 with <i>mfsR</i> and part of <i>orf17984</i> deleted (from nucleotide position 18395 to 19166). | (unpublished) |
| <i>P. putida</i> UWC1 (ICE <i>clc</i> - Δ ' <i>mfsR</i>) | 4322 | Derivative of strain 2737 with an internal deletion in <i>mfsR</i> (from nucleotide position 18581 to 19143). | This study |
| <i>P. putida</i> UWC1 (ICE <i>clc</i> - Δ ' <i>orf17984</i>) | 4372 | Derivative of strain 2737 with an internal deletion in <i>orf17984</i> (from nucleotide position 18032 to 18468). | This study |
| <i>P. putida</i> UWC1 (ICE <i>clc</i> - Δ ' <i>tciR</i>) | 4321 | Derivative of strain 2737 with an internal deletion in <i>tciR</i> (from nucleotide position 17164 to 17985). | This study |
| <i>P. putida</i> UWC1 (ICE <i>clc</i> - Δ <i>mfsABC</i>) | 4165 | Derivative of strain 2737 with a deletion of <i>mfsABC</i> (from nucleotide position 32969 to 37134). | (unpublished) |
| <i>P. putida</i> UWC1 (ICE <i>clc</i> - Δ ' <i>mfsR</i>) + P _{int} - <i>gfp</i> /P _{inR} - | 4469, 4470, | Derivatives of strain 4322 with single copy random insertion of a mini-Tn- | This study |

| | | | |
|--|------------------------|---|------------|
| <i>echerry</i> | 4471 | $P_{int-gfp}/P_{inR-echerry}, Km^R$ | |
| <i>P. putida</i> UWC1 (ICE <i>clc</i> - Δ ' <i>orf17984</i>) + $P_{int-gfp}/P_{inR-echerry}$ | 4475, 4476, 4477 | Derivatives of strain 4372 with single copy random insertion of a mini-Tn- $P_{int-gfp}/P_{inR-echerry}, Km^R$ | This study |
| <i>P. putida</i> UWC1 (ICE <i>clc</i> - Δ ' <i>tcIR</i>) + $P_{int-gfp}/P_{inR-echerry}$ | 4479, 4480, 4481 | Derivatives of strain 4321 with single copy random insertion of a mini-Tn- $P_{int-gfp}/P_{inR-echerry}, Km^R$ | This study |
| <i>P. putida</i> UWC1 (ICE <i>clc</i>) + $P_{int-gfp}/P_{inR-echerry}$ | 3531, 3532, 3533 | Derivatives of strain 2737 with single copy random insertion of a mini-Tn- $P_{int-gfp}/P_{inR-echerry}, Km^R$ | This study |

Table 2. Oligonucleotides used for amplification of ICE*clc* fragments.

| Primer number | Sequence 5'-3' | Purpose | Position on ICE <i>clc</i> |
|---------------|---------------------------------------|---|----------------------------|
| 101105 | gaattcTGCTCCAGGACG GTGAACAA | ICE <i>clc</i> - Δ <i>mfsR</i> - Δ ' <i>orf17984</i> down fragment, EcoRI | 17735- 17754 |
| 101106 | ctcgagGGCTGCTCCATT TGGTTTGACT | ICE <i>clc</i> - Δ <i>mfsR</i> - Δ ' <i>orf17984</i> down fragment, XhoI | 18373- 18394 |
| 101107 | ctcgagGACTTGGTTAGT GCGTTGCATCT | ICE <i>clc</i> - Δ <i>mfsR</i> - Δ ' <i>orf17984</i> up fragment, XhoI | 19167- 19189 |
| 101108 | ggatccGCAGTGCGAGA GTTCCTTTTAGAG | ICE <i>clc</i> - Δ <i>mfsR</i> - Δ ' <i>orf17984</i> up fragment, BamHI | 19714- 19737 |
| 101109 | tttgaattcAGCCCATATGA TAAGCAAGAGTGA | ICE <i>clc</i> - Δ <i>tcIR</i> down fragment, EcoRI | 16644- 16667 |
| 101110 | tttctcgagACGACCTATCT GCTCCGAC | ICE <i>clc</i> - Δ <i>tcIR</i> down fragment, XhoI | 17145- 17163 |
| 101111 | tttctcgagTCACGGCCGT GGTTCTGTGA | ICE <i>clc</i> - Δ <i>tcIR</i> up fragment, XhoI | 17986- 18005 |

| | | | |
|--------|---------------------------------------|---|--------------------------|
| 101112 | ttttgatccGTGGTGACAT TCATGCGTGCCTAT | ICE <i>clc</i> - Δ <i>tc</i> <i>R</i> up fragment, BamHI | 18549- 18572 |
| 120816 | ttttgaattcGCCGAGTTCAT GGAGCG | ICE <i>clc</i> - Δ ' <i>orf17984</i> down fragment, EcoRI | 17404- 17420 |
| 120817 | ttttctcgagTGACCTCGATA GCAAAC | ICE <i>clc</i> - Δ ' <i>orf17984</i> down fragment, XhoI | 18015- 18031 |
| 120818 | ttttctcgagCGCATCAAATT GCTGTG | ICE <i>clc</i> - Δ ' <i>orf17984</i> up fragment, XhoI | 18469- 18485 |
| 120819 | ttttgatccCAACTACCGA CATGATCCAGCGCG | ICE <i>clc</i> - Δ ' <i>orf17984</i> up fragment, BamHI | 19043- 19066 |
| 120820 | ttttgatccGCAGTGCGAG AGTTCCTTTTAGAG | ICE <i>clc</i> - Δ ' <i>mfsR</i> up fragment, BamHI | 19714- 19737 |
| 120821 | ttttctcgagCTGCTCGGTG GCAAGGT | ICE <i>clc</i> - Δ ' <i>mfsR</i> up fragment, XhoI | 19144- 19160 |
| 120822 | ttttgaattcTCACGGCCGT GGTTCTGTGA | ICE <i>clc</i> - Δ ' <i>mfsR</i> down fragment, EcoRI | 17986- 18005 |
| 120823 | ttttctcgagACGCGGCGGT GGTGACATT | ICE <i>clc</i> - Δ ' <i>mfsR</i> down fragment, XhoI | 18562- 18580 |
| 120716 | ttttctagaCTGCCTTTGCC GATGC | ICE <i>clc</i> - Δ <i>mfsABC</i> up fragment, XbaI | 32206- 32221 |
| 120717 | GAAAGTTCACCCAATAT ACGAATAG | ICE <i>clc</i> - Δ <i>mfsABC</i> up fragment | 32944- 32968 |
| 120718 | actagTCGCCATTCGTCC AGCATG | ICE <i>clc</i> - Δ ' <i>mfsABC</i> down fragment, SpeI | 37135- 37152 |
| 120719 | ttttgatccAATGGCCACC CCATAG | ICE <i>clc</i> - Δ ' <i>mfsABC</i> down fragment, BamHI | 37812- 37827 |
| 070934 | AACAAGCCAGGGATGT AACG (tpnRL17-1) | Map Km ^R insertion flanking region in transposon mutants. | not on ICE <i>clc</i> |
| 070935 | CAGCAACACCTTCTTCA CGA (tpnRL13-2) | Map Km ^R insertion flanking region in transposon mutants. | not on ICE <i>clc</i> |

a) small caps, auxiliary sequence absent from ICE*clc*.

b) capital letters, ICE*clc* sequences.

Strain constructions and DNA techniques. DNA purification, PCR, restriction enzyme digestions, DNA ligations and electro-transformations were performed according to standard procedures [29]. Deletions in ICE*clc* genes were created by double recombination techniques as described elsewhere [30]. Nucleotide positions are given according to AJ617740.2 (ICE*clc*). Primers used for strain constructions are listed in table 2.

Random mutagenesis and screening. Random mini-transposon insertions in *P. knackmussii* B13 were generated by mobilization of the pRL27 suicide plasmid from *E. coli* BW20767 in a biparental mating. Hereto both strains were each cultured overnight in 3 ml LB, pelleted down, resuspended in 50 µl sterile saline solution (0.9% NaCl), mixed in a 1:1 (v/v) ratio and incubated on the surface of an LB agar plate for 24 hours at 30°C. The mixture was then resuspended with 1 ml saline solution, which was inoculated in 100 ml MM with 0.5 mM 3CBA plus Km to select for the mini-transposon insertion and Cm to counterselect against *E. coli*, and incubated at 30°C for 16 h with orbital shaking (180 rpm). An aliquot of 3 ml of this pool of enriched Km^R B13 mutants was as used *en masse* as donor in a subsequent mating procedure. Hereto, the 3 ml suspension was pelleted by centrifugation, washed with 3 ml sterile saline and mixed with 3 ml of suspension of *P. putida* UWC1 recipient, that had been grown for 16 h on LB, was pelleted by

centrifugation and resuspended in sterile saline. The mating mixture was again centrifuged, the cell pellet was resuspended in 50 μ l sterile saline solution and spotted on the surface of a MM agar plate containing 0.5 mM 3CBA. The mixture was incubated for 72 hours at 30°C, after which the cells were washed from the plate with 1 ml sterile saline, which was further serially diluted and plated on MM agar plates with 5 mM 3CBA plus Km and Rif to select for transconjugants carrying mutant ICE*c*. Individual colonies were purified, recultured in organized 96-well format and stored at -80°C after addition of and mixing with 15% glycerol. Libraries were replicated and regrown in 100 μ l LB plus Rif for 16 h in 96-well microtiter plates, mixed with 100 μ l *P. putida* UWC1 NaI^R recipient suspension, and incubated at 30°C for 48 h. Then 50 μ l of each well was reinoculated into 170 μ l of MM containing 5 mM 3CBA plus Km, Rif and NaI, and growth was measured by continuous OD-measurements in a multiplate reader (FluoStar Omega, BMG labtech). Absence of growth was taken as indication for absence of ICE*c* transfer, in which case the donor culture was recovered for mapping of the transposon insertion.

Insertion mappings. DIG-labeled primers 070934 or 070935, annealing to one of the ends of the Km^R insert but facing outward, were used (separately) in single-primer PCR with mutant UWC1 donors' DNA as templates. The reactions produced oligonucleotide probes encompassing 5'-DIG, sequence of one of the Km^R insert ends and sequence of ICE*c* downstream the insertion. Such products were used for rough localization of the insertion position by hybridizing to macroblot membranes (Eurogentec,

UK), whose set of 55-mer oligonucleotides covers most of the ICE c genes. Each macroblot membrane was incubated in 4 ml DIG Easy Hyb buffer (DIG Easy Hyb and MB Grade from fish sperm 100 μ g/ml, Roche Diagnostics GmbH, Mannheim, Germany) for 2 hours at 55°C until it was incubated with 4 ml of DIG Easy Hyb buffer supplemented with 12 μ l of the individual DIG-labeled probes for an additional 18 hours. Membranes washing consisted of 4 times 5 minutes with 25 ml 2X SSC 0.1% SDS, followed by 4 times 30 minutes with 25 ml 0.2X SSC 0.1% SDS and 4 times 5 minutes with 25 ml 1X SSC. Membranes were then incubated 30 minutes with 10 ml of 5% blocking agents, followed by 30 minutes with 10 ml anti-DIG antibodies (Roche). Again, membranes were washed 3 times with 25 ml washing buffer (0.5X SSPE 0.2% SDS) for 15 minutes. Then, membranes were incubated with 1 ml detection buffer 1% CSPD (Roche) for 5 minutes, and exposed to Super RX Fuji Medical X-ray film. Once the insertion was roughly mapped, PCR-based sequencing was used to determine exact position on mutant ICE c donor DNA.

Transfer assay. Frequency of ICE c transfer was determined in experimental conditions described previously [31]. *P. putida* UWC1 (ICE c), *P. putida* UWC1 (ICE c -Km^R₁₉₀₃₃), *P. putida* UWC1 (ICE c -Km^R₁₈₆₁₈), *P. putida* UWC1 (ICE c - Δ 'orf17984), *P. putida* UWC1 (ICE c - Δ 'tciR), *P. putida* UWC1 (ICE c - Δ mfsR- Δ 'orf17984), *P. putida* UWC1 (ICE c - Δ 'mfsR), *P. putida* UWC1 (ICE c - Δ mfsABC) were each used as donors, whereas *P. putida* UWCGC (constitutively fluorescent, Gm^R) was used as the recipient. Donors and recipient were each cultivated on 5 mM 3CBA MM and 10 mM fructose MM, respectively and combined on 0.5 mM 3CBA agar plates as a

single concentrated pellet. After 48 hours incubation at 30°C, mating mixes were resuspended, diluted and plated on 5 mM 3CBA MM agar (counting of donors) or 5 mM 3CBA Gm agar Petri dishes (counting of transconjugants). Transconjugants were verified by PCR for carrying ICE*clc*.

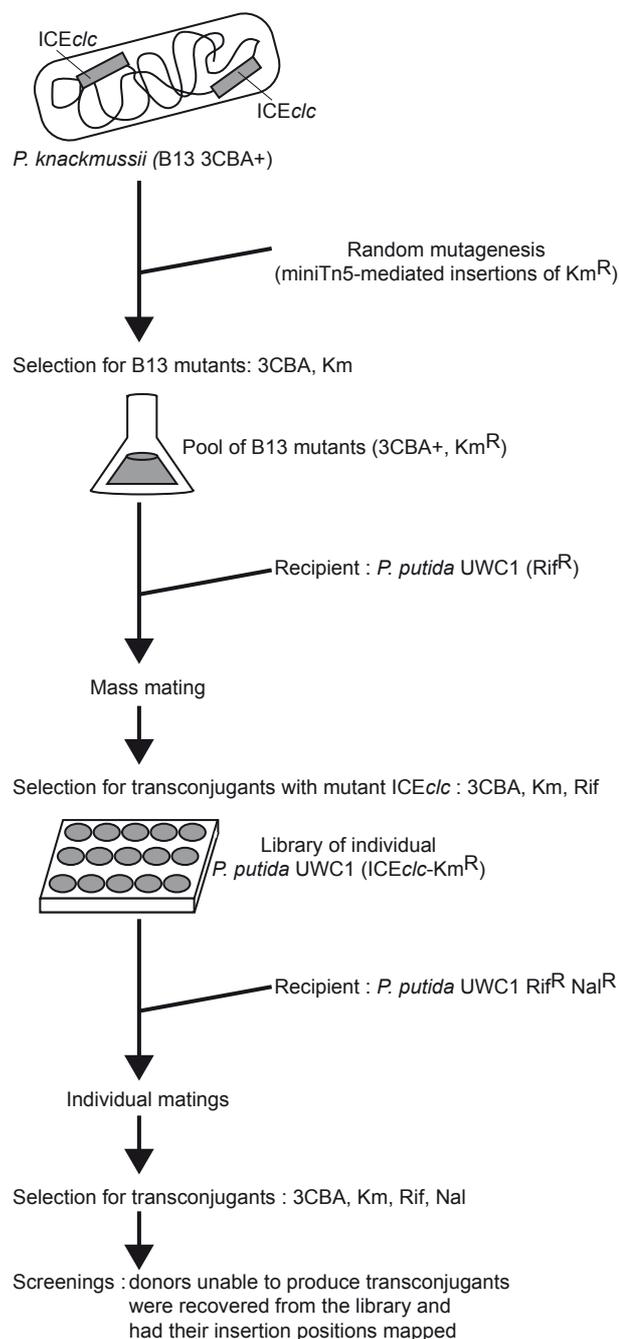
Microarrays. Transcriptomes of *P. putida* UWC1 (ICE*clc*), *P. putida* UWC1 (ICE*clc*), *P. putida* UWC1 (ICE*clc*-Km^R₁₉₀₃₃), *P. putida* UWC1 (ICE*clc*- Δ 'orf17984), *P. putida* UWC1 (ICE*clc*- Δ 'tciR), *P. putida* UWC1 (ICE*clc*- Δ 'mfsR- Δ 'orf17984), *P. putida* UWC1 (ICE*clc*- Δ 'mfsR) were determined by microarray analysis, as described previously [22]. Total RNA was extracted from cells grown on 10 mM 3CBA MM and harvested at mid exponential phase (OD₆₀₀ = 0.6) and 48 hours after the entrance in stationary phase. Reverse transcription using cyanine-labeled dCTP among the dNTPs, produced cy3-labeled cDNA that was further purified and hybridized on 8x15 microarray slides (Agilent, Santa Clara, CA, USA). Slides were washed and scanned. Data were recovered and analyzed using GeneSpring GX. Microarray data can be accessed from the GEO database (accession number : GSE51391).

Time-lapse microscopy. *P. putida* UWC1 strains were precultured o/n at 30°C in LB medium and 100 μ l of the o/n culture were subsequently grown in 20 ml MM 4 mM 3CBA and the appropriate antibiotics. After incubation for 96 hours at 30°C, the cells were 100x diluted in 1xMM. For the preparation of the gel patches, 1 ml of MM containing 1% (w/v) agarose (kept molten at 50°C), 20 μ l of Hutner's, 10 μ l of 10 mM 3CBA and 2 μ l of vitamin solution were mixed together. Four gel patches were prepared by placing 130 μ l of the

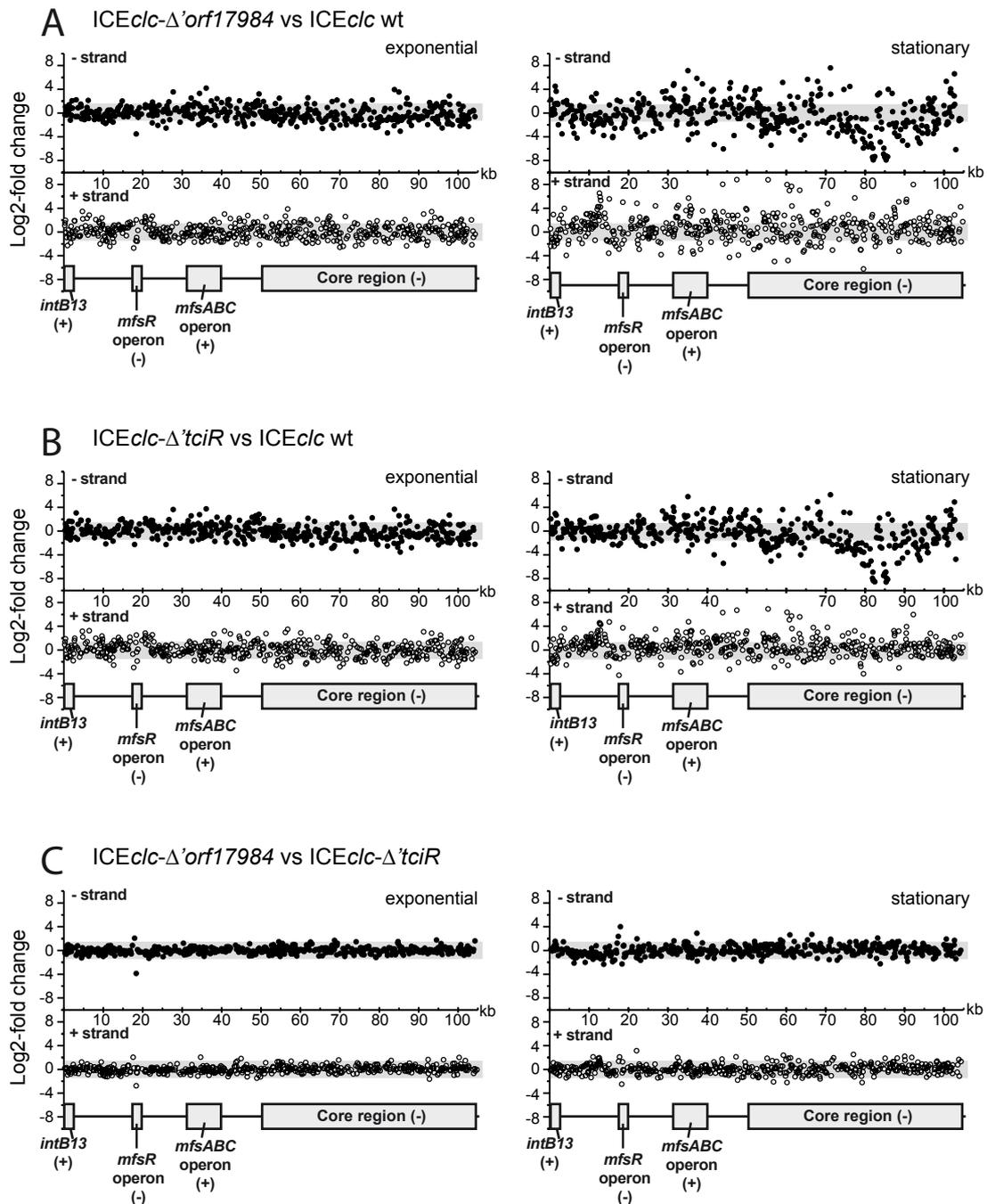
MM mixture per patch on a circular cover slip (42 mm in diameter and 0.17-mm thick) and covering them with a second slip. When the agarose was solid, 6 μ l of the cell dilution were placed onto the agarose gel surface. As soon as the drops were dried, the patches were turned upside down and placed on a new round cover slip. A silicon spacer ring was added and a second circular cover slip was put on top. The whole system was placed into a rigid metal cast POC chamber and fixed with a metal ring.

Images were taken during 48 hours with intervals of 1 hour in 10 regions per agarose patch using a Nikon Inverted Microscope Eclipse Ti-E, equipped with a Perfect Focus System (PFS), pE-100 CoolLED and a Plan Apo λ 100x 1.45 Oil objective. For the analysis of the images, the image acquisition software Micro-Manager 1.4 and MetaMorph (Series 7.5, MDS, Analytical Technologies) were used.

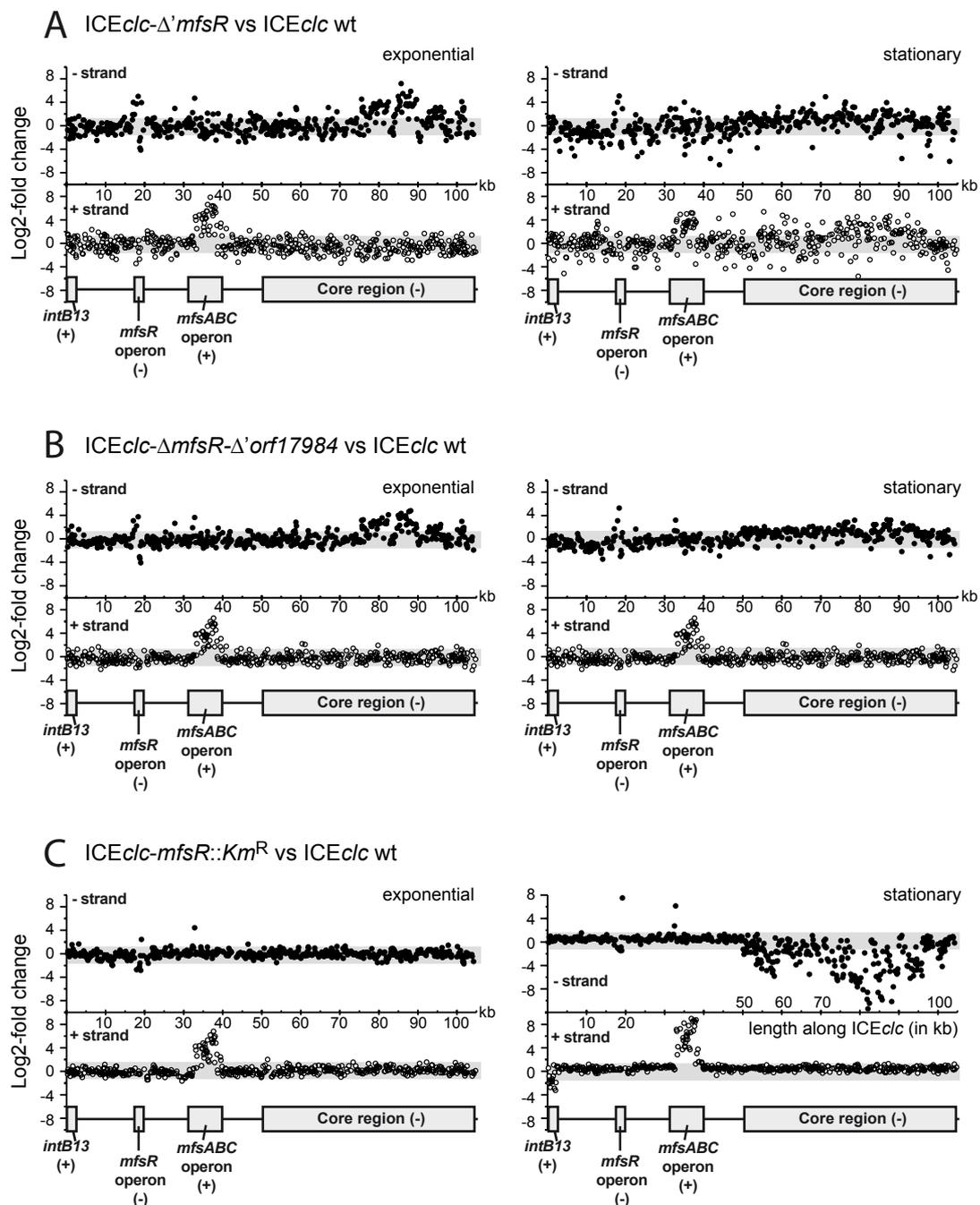
SUPPLEMENTARY DATA



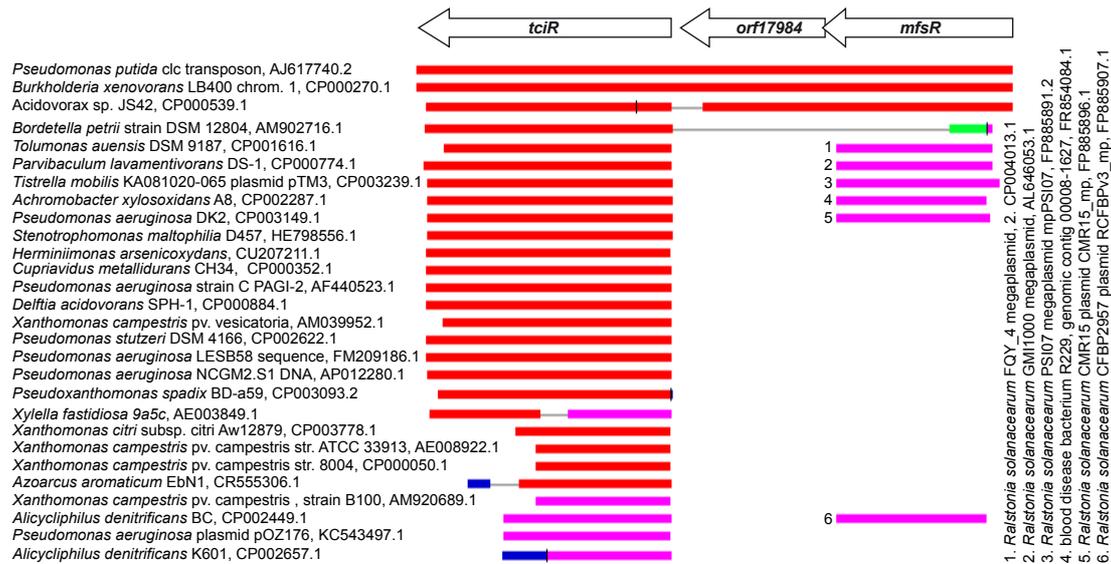
Supplementary Figure S1. Outline of the random mutagenesis and subsequent selection procedure. Original ICE_{clc}-host *P. knackmussii* B13 is randomly mutagenized by miniTn5-mediated insertions of kanamycin resistance inserts (Km^R). Mutant B13 are selected by culturing cells on minimal medium (MM) with Km and 3-chlorocatechol (3CBA) as sole carbon and energy source. The pool of B13 mutants is cultured in batch and mixed with recipient strain *P. putida* UWC1 (resistant to rifampicin, Rif^R). The mixture is incubated in mating conditions for 72 hours and plated on MM with 3CBA, Km and Rif, to select for transconjugants. Individual colonies of transconjugants were restriking and organized into a mutant library in 96-well plates. Each mutant is used as donor in a new 96-well mating with recipient *P. putida* UWC1 resistant to Rif and nalidixic acid (Nal^R). Individual mating mixtures are grown on MM agar with 3CBA, Km, Rif, Nal in order to select for transconjugants. In absence of transconjugant growth, the donor of that particular mating was recovered from the library and had its insertion position mapped.



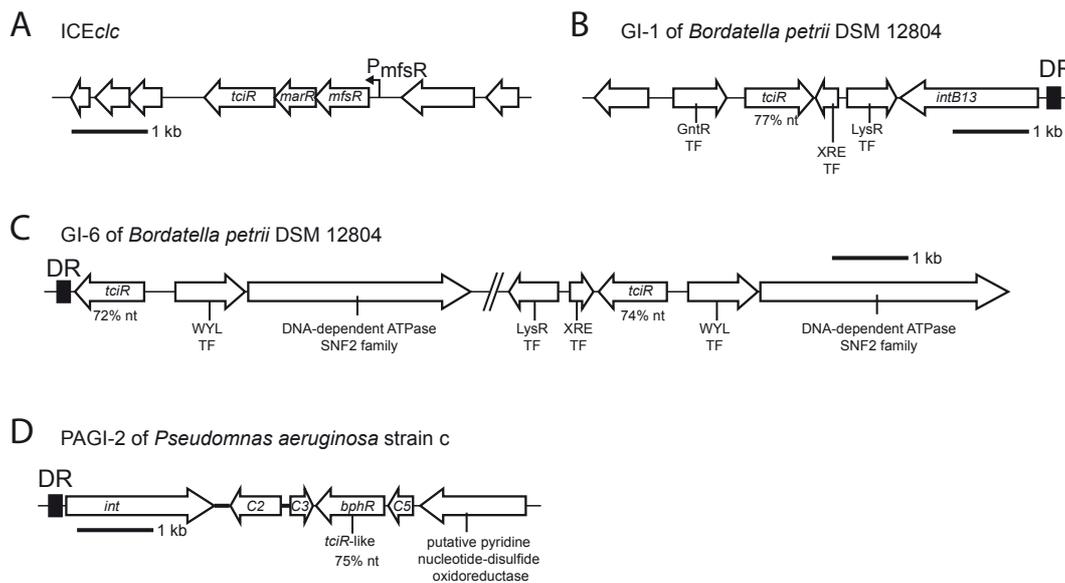
Supplementary Figure S2. Pair-wise comparisons of expression in the ICE*clc* area between *tcIR* or *orf17984* deletions and ICE*clc* wild-type, by using microarray analysis. Panels indicate comparisons of exponentially growing or stationary phase cells, with hybridization signals on the plus- or minus-strand of ICE*clc*. Dots indicate the log₂-fold change of hybridization signal per microarray probe in the comparison, plotted at their distance along the ICE*clc* sequence (X-axis; in kb). A scheme of ICE*clc* is redrawn at the bottom of each section, with regions of interest as grey boxes (+ or - indicate the DNA strand on which the region is encoded). Grey bars in the background indicate the two-fold cut-off level.



Supplementary Figure S3. Pair-wise comparisons of expression in the ICE*clc* area between *mfsR* deletion mutants, *mfsR* transposon insertion mutant and ICE*clc* wild-type, by using microarray analysis. Panels indicate comparisons of exponentially growing or stationary phase cells, with hybridization signals on the plus- or minus-strand of ICE*clc*. Dots indicate the log₂-fold change of hybridization signal per microarray probe in the comparison, plotted at their distance along the ICE*clc* sequence (X-axis; in kb). A scheme of ICE*clc* is redrawn at the bottom of each section, with regions of interest as grey boxes (+ or - indicate the DNA strand on which the region is encoded). Grey bars in the background indicate the two-fold cut-off level.



Supplementary Figure S4. Homologs of the *mfsR* operon. Blastn of the nucleotide sequence of the *mfsR* operon (position 17053 to 19188 from ICE*clc* AJ617740) versus the non redundant nt collection of NCBI (<http://blast.ncbi.nlm.nih.gov/>). Alignment scores, red \geq 200, pink, 80-200, green 50-80, blue 40-50. Black lines, insertions.



Supplementary Figure S5. Genetic configurations surrounding the *tciR* gene in various strains. A, the *tciR*-containing *mfsR* operon found in ICE*clc* of *Pseudomonas knackmussii* B13 and *Burkholderia xenovorans* LB400 (accession numbers, AJ617740.2 and CP000270.1, respectively). B, *tciR* homologue in GI-1 of *Bordetella petrii* DSM 12804 (accession number AM902716.1). C, two *tciR* homologues in GI-6 of *Bordetella petrii* DSM 12804 (accession number AM902716.1). D, a *tciR* homologue (*bphR*) in PAGI-2 of *Pseudomonas aeruginosa* strain C (accession number AF440523.1).

REFERENCES

1. Burrus V, Waldor MK. *Shaping bacterial genomes with integrative and conjugative elements*. Research in Microbiology 2004, 155:376–386.
2. Burrus V, Pavlovic G, Decaris B, Guédon G. *Conjugative transposons: the tip of the iceberg*. Molecular microbiology 2002, 46:601–610.
3. Dobrindt U, Hochhut B, Hentschel U, Hacker J. *Genomic islands in pathogenic and environmental microorganisms*. Nature Reviews Microbiology 2004, 2:414–424.
4. Wozniak RAF, Waldor MK. *Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow*. Nature Reviews Microbiology 2010, 8:552–563.
5. Lee CA, Babic A, Grossman AD. *Autonomous plasmid-like replication of a conjugative transposon*. Molecular microbiology 2010, 75:268–279.
6. Auchtung JM, Lee CA, Monson RE, Lehman AP, Grossman AD. *Regulation of a Bacillus subtilis mobile genetic element by intercellular signaling and the global DNA damage response*. Proceedings of the National Academy of Sciences of the United States of America 2005, 102:12554–12559.
7. Beaber JW, Hochhut B, Waldor MK. *SOS response promotes horizontal dissemination of antibiotic resistance genes*. Nature 2004, 427:72–74.
8. Burrus V, Waldor MK. *Control of SXT integration and excision*. Journal of Bacteriology 2003, 185:5045–5054.
9. Sezonov G, Hagège J, Friedmann A, Guerineau M. *Characterization of pra, a gene for replication control in pSAM2, the integrating element of Streptomyces ambofaciens*. Molecular Microbiology 1995, 17:533–44.
10. Sezonov G, Duchêne A-M, Friedmann A, Guérineau M, Pernodet J-L. *Replicase, excisionase, and integrase genes of the Streptomyces element pSAM2 constitute an operon positively regulated by the pra gene*. Journal of bacteriology 1998, 180:3056–3061.
11. Sezonov G, Possoz C, Friedmann A, Pernodet J-L, Guerineau M. *KorSA from the Streptomyces integrative element pSAM2 is a central transcriptional repressor: target genes and binding sites*. Journal of Bacteriology 2000, 182:1243–1250.
12. Ramsay JP, Sullivan JT, Jambari N, Ortori CA, Heeb S, Williams P, Barrett DA, Lamont IL, Ronson CW. *A LuxRI-family regulatory system controls excision and transfer of the Mesorhizobium loti strain R7A symbiosis island by activating expression of two conserved hypothetical*

- genes*. *Molecular Microbiology* 2009, 73:1141–1155.
13. Ramsay JP, Major AS, Komarovskiy VM, Sullivan JT, Dy RL, Hynes MF, Salmond GPC, Ronson CW. *A widely conserved molecular switch controls quorum sensing and symbiosis island transfer in Mesorhizobium loti through expression of a novel antiactivator*. *Molecular Microbiology* 2013, 87:1–13.
 14. Ramsay JP, Sullivan JT, Stuart GS, Lamont IL, Ronson CW. *Excision and transfer of the Mesorhizobium loti R7A symbiosis island requires an integrase IntS, a novel recombination directionality factor RdfS, and a putative relaxase RlxS*. *Molecular Microbiology* 2006, 62:723–734.
 15. Beaber JW, Waldor MK. *Identification of operators and promoters that control SXT conjugative transfer*. *Journal of Bacteriology* 2004, 186:5945–5949.
 16. Beaber JW, Hochhut B, Waldor MK. *Genomic and functional analyses of SXT, an integrating antibiotic resistance gene transfer element derived from Vibrio cholerae*. *Journal of Bacteriology* 2002, 184:4259–4269.
 17. Minoia M, Gaillard M, Reinhard F, Stojanov M, Sentchilo V, van der Meer JR. *Stochasticity and bistability in horizontal transfer control of a genomic island in Pseudomonas*. *Proceedings of the National Academy of Sciences* 2008, 105:20792–20797.
 18. Gaillard M, Vallaeyts T, Vorhölter FJ, Minoia M, Werlen C, Sentchilo V, Pühler A, van der Meer JR. *The clc element of Pseudomonas sp. strain B13, a genomic island with various catabolic properties*. *Journal of bacteriology* 2006, 188:1999–2013.
 19. Reinhard F, Miyazaki R, Pradervand N, van der Meer JR. *Cell Differentiation to “mating bodies” induced by an integrating and conjugative element in free-living bacteria*. *Current Biology* 2013, 23:255–259.
 20. Miyazaki R, Minoia M, Pradervand N, Sulser S, Reinhard F, van der Meer JR. *Cellular variability of RpoS expression underlies subpopulation activation of an integrative and conjugative element*. *PLoS Genetics* 2012, 8:e1002818.
 21. Larsen R, Wilson M, Guss A, Metcalf W. *Genetic analysis of pigment biosynthesis in Xanthobacter autotrophicus Py2 using a new, highly efficient transposon mutagenesis system that is functional in a wide variety of bacteria*. *Archives of Microbiology* 2002, 178:193–201.
 22. Gaillard M, Pradervand N, Minoia M, Sentchilo V, Johnson DR, van der Meer JR. *Transcriptome analysis of the mobile genome ICE_{clc} in Pseudomonas knackmussii B13*. *BMC microbiology* 2010, 10:153.
 23. Bi D, Xu Z, Harrison EM, Tai C, Wei Y, He X, Jia S, Deng Z, Rajakumar

- K, Ou H-Y. *ICEberg: a web-based resource for integrative and conjugative elements found in Bacteria*. *Nucleic Acids Research* 2011, 40:D621–D626.
24. Lechner M, Schmitt K, Bauer S, Hot D, Hubans C, Levillain E, Loch C, Lemoine Y, Gross R. *Genomic island excisions in Bordetella petrii*. *BMC Microbiology* 2009, 9:141.
 25. Klockgether J, Wurdemann D, Wiehlmann L, Tummler B. *Transcript profiling of the Pseudomonas aeruginosa genomic islands PAGI-2 and pKLC102*. *Microbiology* 2008, 154:1599–1604.
 26. Tropel D, Van Der Meer JR. *Bacterial transcriptional regulators for degradation pathways of aromatic compounds*. *Microbiology and Molecular Biology Reviews* 2004, 68:474–500.
 27. Maddocks SE, Oyston PCF. *Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins*. *Microbiology* 2008, 154:3609–3623.
 28. Gerhardt P, Murray R, Costilow R, Nester E, Wood W, Krieg N, Briggs Phillips G. *Manual of methods for general bacteriology*. 1981.
 29. Sambrook JR. *Molecular cloning: a laboratory manual*, third edn. 2001.
 30. Miyazaki R, van der Meer JR. *A dual functional origin of transfer in the ICE_{clc} genomic island of Pseudomonas knackmussii B13*. *Molecular Microbiology* 2011, 79:743–758.
 31. Sentchilo V, Czechowska K, Pradervand N, Minoia M, Miyazaki R, Van Der Meer JR. *Intracellular excision and reintegration dynamics of the ICE_{clc} genomic island of Pseudomonas knackmussii sp. strain B13*. *Molecular microbiology* 2009, 72:1293–1306.
 32. Dorn E, Hellwig M, Reineke W, Knackmuss H. *Isolation and characterization of a 3-chlorobenzoate degrading pseudomonad*. *Arch Microbiol* 1974, 99:61–70.
 33. McClure NC, Weightman AJ, Fry JC. *Survival of Pseudomonas putida UWC1 containing cloned catabolic genes in a model activated-sludge unit*. *Applied and Environmental microbiology* 1989, 55:2627–2634.

CHAPTER 6

GENERAL DISCUSSION

The purpose of this thesis was to explore the regulatory network underlying the control of horizontal transfer of ICE*clc*. By using a combination of microarray and complementary techniques, we could propose a map of 15 mRNAs synthesized in the core module of ICE*clc*, a region predicted to encode most HGT functions. In the process, we discovered that ICE*clc* transcriptome had a strong bipartite behavior depending on the growth phase of the host cells. In stationary phase after growth on 3CBA, the core module is globally transcribed, whereas the cargo module becomes silent and vice-versa in exponential phase. Then, by generating a library of random ICE*clc* mutants in *P. putida* UWC1 and screening for HGT-deficient phenotypes, it was possible to identify the *mfsR* operon as a major actor in the regulation of transfer. The *mfsR* operon, located outside the previously defined core module, is composed of three consecutive transcriptional regulator genes, *mfsR*, *orf17984*, *tcIR*, belonging to the families of TetR, MarR and LysR regulators, respectively. We could demonstrate that it is most likely TciR, which is responsible for the global activation of the core module in stationary phase and for the emergence of the so-called tc cells (transfer competent cells, eventually becoming donors). We could show that TciR is under the transcriptional control of MfsR, a repressor acting both on its own promoter P_{mfsR} and on the promoter of the *mfsABC-orf38184* operon P_{mfsA} . The *mfsABC-orf38184* operon is supposed to encode an efflux pump system of the major facilitator superfamily, as well as three further enzymes. Deletion of the pumps did not affect ICE*clc* transfer frequency, therefore, it is unlikely to have any role in HGT. Deletion of the pump genes did not have a clear phenotype in antimicrobial compound resistance, either. In many described

efflux systems controlled by a TetR-type repressor, the substrate of the pumps is a toxic molecule that inhibits the repressor by allosterically interacting with it (thus becoming a ligand) [1]. The resulting ligand-bound repressor loses affinity for its DNA operator and therefore alleviates repression on the pumps genes as well as on the repressor gene. The synthesized efflux pumps drive the ligand molecule outside the cell, thus returning the system to equilibrium. If such ligand-induced mechanism would exist between MfsR and the MfsABC efflux system, as depicted in Figure 1, the identification of the ligand molecule would be a priority as it may be the key to ultimately explain the activation of the core module via TciR. During this study, we made two educated guesses about the mysterious ligand molecule, if there is any: (i) the molecule may be a substrate or an intermediate of the pathways provided by ICE*clc* cargo module, thus linking HGT to the specific ICE*clc* fitness, as it is the case for SXT, CTnDOT or Tn916 [2–4]. (ii) The molecule may be a general cell toxicant, thus linking stress situations for the host cell with ICE escape, as is the case for ICE*Bs1* or ICE*St1* [5,6]. In order to verify the first hypothesis, i.e., the ligand is a substrate/intermediate of the ICE*clc* pathways, we tested whether the P_{mfsA} -*echerry* reporter fusion would be induced by exposition to 3CBA, 3CLC or 2AP. Simultaneously, we tested in MIC assays whether a host carrying an ICE*clc* with deleted *mfsABC* genes or deleted *mfsR* would be more sensitive to those compounds than a wild type ICE*clc* host. None of these two approaches were conclusive in identifying a responding molecule.

In order to challenge the second hypothesis, i.e., the ligand is a general toxic molecule, we tested a wide range of halogenated compounds, organic

solvents and antibiotics both with the inducer and toxicant approaches. Some of those compounds, like ethanol or chloramphenicol were selected because of homologous predicted domains between MfsR and AcrR or AefR [7,8]. As none of the tested molecules gave clear phenotypes, we expanded the toxicant search by comparing *P. putida* UWC1 (ICE*clc*) and *P. putida* UWC1 (ICE*clc*- Δ *mfsABC*) in a BIOLOG chemical sensitivity Phenotype Array (PM09-PM20). The only convincing result from this screening was that CrCl₃ in concentrations between 100 and 500 mg/L affected the growth of the Δ *mfsABC* mutant consistently to a higher degree than that of wild type. This result, however, could not be confirmed by growth rate and growth yield measurements of wild-type and *mfsABC* deletion mutant in two media, challenged or not with CrCl₃.

Another possibility could be that the MfsR repressor does not need (anymore) a ligand molecule to derepress the *mfsR* operon and trigger the HGT program. Indeed, one could imagine that MfsR could be naturally erring on the P_{mfsR} promoter as suggested by (i) the higher constitutive expression of the P_{mfsR}-*echerry* reporter fusion than P_{mfsA}-*echerry*. As for the proposed model of HGT regulation in ICEMISym^{R7A}, a complex system of autoregulatory loops fluctuating above a threshold in certain cells can explain the triggering of HGT program [9–11]. Indeed, it is possible that in rare occasions certain cells display a stronger leaky expression of the *mfsR* operon. In those cells, TciR will then be higher expressed as well, and can activate the tc pathway, when simultaneously higher levels of the stationary phase sigma factor RpoS are present [12–14]. Subsequently, tc cells arise and globally activate the core module and the integrase gene, ultimately leading to the transfer of ICE*clc*.

Differential repression by MfsR may be a consequence of varying stability in exponential or stationary phase, but this was never tested.

Finally, more work is necessary to understand the targets of the *mfsR-orf17984-tciR* transcription factors. How is TciR activating the core region genes? Does it bind each of the 15 identified promoters in the core region, or does it start a cascade at one particular promoter? Does it need co-activator(s), like other LysR-type transcriptional regulators? What is the influence of RpoS on TciR target(s) ?

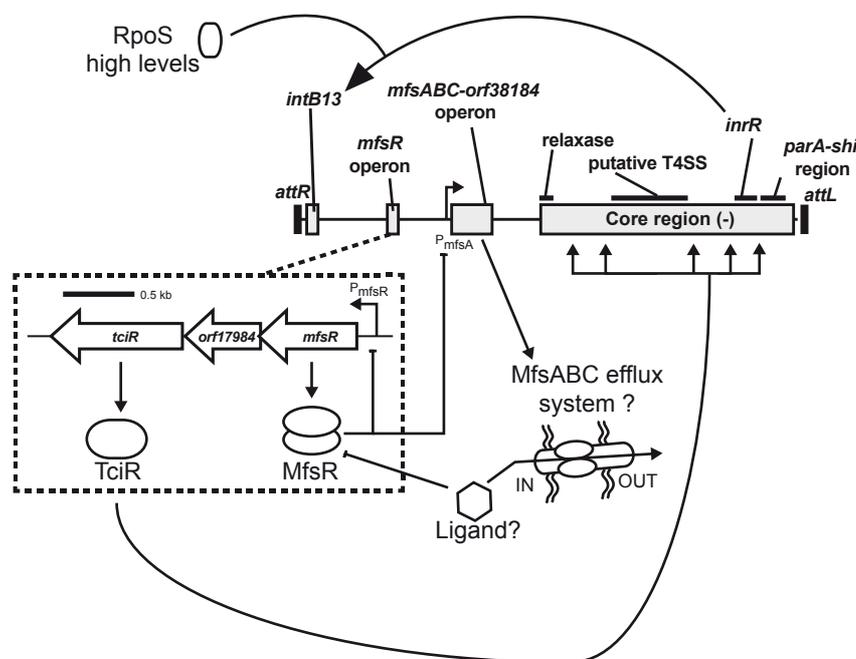


Figure 1. Working model of ICE clc regulation. The *mfsR* operon is under the transcriptional control of the MfsR repressor, as is the *mfsABC-orf38184* operon. The MfsR repressor dimerizes and specifically binds to a pair of operator boxes (not shown) in the P_{mfsR} and P_{mfsA} promoters, inhibiting transcription. The *mfsABC-orf38184* operon putatively encodes a three-partite efflux pumps (MfsABC) and three putative enzymes with unknown functions (not shown). A ligand molecule may inhibit allosterically the MfsR repressor. By analogy to other TetR-efflux systems, this ligand molecule would be hypothetically driven outside the cell by the MfsABC pumps. Besides the *mfsR* and *orf17984* genes, derepression of the *mfsR* operon leads to the synthesis of TciR, a regulator that activates the entire core module (possibly via

inrR) during stationary phase after growth on 3CBA, in cells that happen to have the highest level of RpoS.

The results of this work present two major perspectives, (i) better understanding of ICE*clc*-type transfer control and (ii) the possible stimulation of (beneficial) HGT if it is possible to increase transfer competence formation in donor cells.

Assuming that transfer of ICE*clc* type elements is not solely stochastic, but requires some ligand, which is not a toxic metal (like CrCl₃), I propose that induction of ICE-transfer in bacterial communities in waste water treatment plants or in sites contaminated with organohalides might help to accelerate spontaneous adaptation and natural attenuation. One could easily imagine that higher horizontal transfer rates of catabolic MGEs *in situ* could convert a certain fraction of native bacteria in a system into effective bioremediation actors [15–21]. If this would be feasible, engineers could benefit from the help of bacteria simultaneously (i) adapted to the local environment, (ii) already present on site in great number (no need for culture and dispersal) and (iii) "upgraded" with the biochemical power to clean up many chlorinated aromatics via the CLC -"hub"- catabolism.

On the other hand, I conclude that the discovery of the potential role of TciR as a global activator of ICE*clc* HGT will provide a non-negligible help for future studies of ICE*clc* lifestyle and HGT machinery. Indeed, monitoring ICE*clc* molecular events is rendered difficult by the fact that it is active only in tc cells,

which occur under natural conditions in 3-5% of the population. As confessed by members of the Grossman laboratory, the rapid and deep understanding of ICEBs1 was enabled by the fact that they could induce ICEBs1 transfer in >95% of the cells via *rapI* overexpression [5,22]. Mutants such as the deletion of *mfsR* leading to overexpression of *tcIR* will ease the detection and investigation of the lifestyle-related phenomena in ICEc1c. Finally, interesting experiments can be done to assess the ecological importance of bistability of HGT compared to majority-based HGT, as it exists in certain systems, like pCF100 [23].

REFERENCES

1. Ramos JL, Martínez-Bueno M, Molina-Henares AJ, Terán W, Watanabe K, et al. *The TetR family of transcriptional repressors*. Microbiol Mol Biol Rev 2005, 69: 326–356.
2. Beaber JW, Hochhut B, Waldor MK. *SOS response promotes horizontal dissemination of antibiotic resistance genes*. Nature 2004, 427: 72–74.
3. Shoemaker NB, Salyers AA. *Tetracycline-dependent appearance of plasmidlike forms in Bacteroides uniformis 0061 mediated by conjugal Bacteroides tetracycline resistance elements*. J Bacteriol 1988, 170: 1651–1657.
4. Celli J, Trieu-Cuot P. *Circularization of Tn916 is required for expression of the transposon-encoded transfer functions: characterization of long tetracycline-inducible transcripts reading through the attachment site*. Mol Microbiol 1998, 28: 103–117.
5. Auchtung JM, Lee CA, Monson RE, Lehman AP, Grossman AD. *Regulation of a Bacillus subtilis mobile genetic element by intercellular signaling and the global DNA damage response*. Proc Natl Acad Sci USA 2005, 102: 12554–12559.
6. Bellanger X, Roberts AP, Morel C, Choulet F, Pavlovic G, et al. *Conjugative transfer of the integrative conjugative elements ICESt1 and ICESt3 from Streptococcus thermophilus*. J Bacteriol 2009, 191: 2764–2775.
7. Su CC, Rutherford DJ, Yu EW. *Characterization of the multidrug efflux regulator AcrR from Escherichia coli*. Biochem Biophys Res Commun 2007, 361: 85–90.
8. Kawakita Y, Taguchi F, Inagaki Y, Toyoda K, Shiraishi T, et al. *Characterization of each aefR and mexT mutant in Pseudomonas syringae pv. tabaci 6605*. Mol Genet Genomics 2012: 1–12.
9. Ramsay JP, Sullivan JT, Stuart GS, Lamont IL, Ronson CW. *Excision and transfer of the Mesorhizobium loti R7A symbiosis island requires an integrase IntS, a novel recombination directionality factor RdfS, and a putative relaxase RlxS*. Mol Microbiol 2006, 62: 723–734.
10. Ramsay JP, Sullivan JT, Jambari N, Ortori CA, Heeb S, et al. *A LuxRI-family regulatory system controls excision and transfer of the Mesorhizobium loti strain R7A symbiosis island by activating expression of two conserved hypothetical genes*. Mol Microbiol 2009, 73: 1141–1155.
11. Ramsay JP, Major AS, Komarovskiy VM, Sullivan JT, Dy RL, et al. *A widely conserved molecular switch controls quorum sensing and symbiosis island transfer in Mesorhizobium loti through expression of a*

- novel antiactivator*. Mol Microbiol 2013, 87: 1–13.
12. Minoia M, Gaillard M, Reinhard F, Stojanov M, Sentchilo V, et al. *Stochasticity and bistability in horizontal transfer control of a genomic island in Pseudomonas*. Proc Natl Acad Sci 2008, 105: 20792–20797.
 13. Miyazaki R, Minoia M, Pradervand N, Sulser S, Reinhard F, et al. *Cellular variability of RpoS expression underlies subpopulation activation of an integrative and conjugative element*. Plos Genet 2012, 8: e1002818.
 14. Reinhard F, Miyazaki R, Pradervand N, van der Meer JR. *Cell differentiation to “mating bodies” induced by an integrating and conjugative element in free-living bacteria*. Curr Biol 2013, 23: 255–259.
 15. Bathe S, Schwarzenbeck N, Hausner M. *Bioaugmentation of activated sludge towards 3-chloroaniline removal with a mixed bacterial population carrying a degradative plasmid*. Bioresour Technol 2009, 100: 2902–2909.
 16. Bathe S, Schwarzenbeck N, Hausner M. *Plasmid-mediated bioaugmentation of activated sludge bacteria in a sequencing batch moving bed reactor using pNB2*. Lett Appl Microbiol 2005, 41: 242–247.
 17. Dejonghe W, Goris J, El Fantroussi S, Hofte M, De Vos P, et al. *Effect of dissemination of 2,4-dichlorophenoxyacetic acid (2,4-D) degradation plasmids on 2,4-D degradation and on bacterial community structure in two different soil horizons*. Appl Environ Microbiol 2000, 66: 3297–3304.
 18. Springael D, Top EM, Seeger H, Aguzzi A. *Horizontal gene transfer and microbial adaptation to xenobiotics: new types of mobile genetic elements and lessons from ecological studies*. Trends Microbiol 2004, 12: 53–58.
 19. Top EM, Springael D. *The role of mobile genetic elements in bacterial adaptation to xenobiotic organic compounds*. Curr Opin Biotechnol 2003, 14: 262–269.
 20. Top EM, Springael D, Boon N. *Catabolic mobile genetic elements and their potential use in bioaugmentation of polluted soils and waters*. Fems Microbiol Ecol 2002, 42: 199–208.
 21. Van der Meer JR, Sentchilo V. *Genomic islands and the evolution of catabolic pathways in bacteria*. Curr Opin Biotechnol 2003, 14: 248–254.
 22. Lee CA, Babic A, Grossman AD. *Autonomous plasmid-like replication of a conjugative transposon*. Mol Microbiol 2010, 75: 268–279.
 23. Dunny GM. *The peptide pheromone-inducible conjugation system of Enterococcus faecalis plasmid pCF10: cell–cell signalling, gene transfer, complexity and evolution*. Philos Trans R Soc B Biol Sci 2007, 362: 1185.

CURRICULUM VITAE

Nicolas Pradervand

Born in Fribourg, Switzerland, November 8th 1983

Nationality : Swiss

- 2008 - **PhD in Life Sciences**
2013 Faculty of Biology and Medicine, University of Lausanne, Switzerland
University of Lausanne's PhD fellowship 2008
- PhD thesis *Regulation of ICElc horizontal transfer in Pseudomonas knackmussii B13.*
Lab of Jan Roelof van der Meer
Department of Fundamental Microbiology
- 2006 - **Master of Sciences in Genomics and Experimental Biology**
2008 Faculty of Biology and Medicine, University of Lausanne, Switzerland
- Long step MSc thesis *Identification of transfer functions of the CLC element in Pseudomonas sp. strain B13.*
Lab of Jan Roelof van der Meer
Department of Fundamental Microbiology
- Short step MSc thesis *Molecular insights into the regulation of antifungal activity in the biocontrol bacterium Pseudomonas fluorescens CHAO.*
Lab of Christoph Keel
Department of Fundamental Microbiology
- 2003 - **Bachelor of Sciences in Biology** (specialization in Biochemistry)
2006 Faculty of Sciences, University of Fribourg, Switzerland
- BSc thesis *Does the aminoacylated state of tRNA influence its stability in Trypanosoma brucei ?*
Lab of André Schneider
Department of Cell and Developmental Biology
- 1999 - **Swiss Federal Maturity**
2003 Collège de Gambach, Fribourg, Switzerland

LIST OF PUBLICATIONS

Miyazaki R, Minoia M, Pradervand N, Sentchilo V, Sulser S, Reinhard F and van der Meer JR. *The clc element and related genomic islands in Proteobacteria.* Published in Bacterial Integrative Mobile Genetic Elements, edited by Adam P. Roberts and Peter Mullany. 2013 Landes Bioscience. ISBN: 978-1-58706-660-3.

Reinhard F, Miyazaki R, Pradervand N, van der Meer JR. *Host cell differentiation to primitive "mating bodies" induced by an integrating and conjugative element in free living bacteria.* Current Biology. 2013 Feb 4;23(3):255-9.

Miyazaki R, Minoia M, Pradervand N, Sulser S, Reinhard F, van der Meer JR. *Cellular variability of RpoS expression underlies subpopulation activation of an integrative and conjugative element.* PLoS Genet. 2012; 8(7):e1002818.

Gaillard M*, Pradervand N*, Minoia M, Sentchilo V, Johnson DR, van der Meer JR. *Transcriptome analysis of the mobile genome ICE_{clc} in Pseudomonas knackmussii B13.* BMC Microbiol. 2010 May 26;10:153.
(*Shared first authors)

Sentchilo V, Czechowska K, Pradervand N, Minoia M, Miyazaki R, van der Meer JR. *Intracellular excision and reintegration dynamics of the ICE_{clc} genomic island of Pseudomonas knackmussii sp. strain B13.* Molecular Microbiology. 2009 Jun;72(5):1293-306 .

In preparation :

Nicolas Pradervand, Ryo Miyazaki, Jan Roelof van der Meer. *Characterization of a TetR-type repressor regulated major facilitator system on the integrative and conjugative element ICE_{clc} in Pseudomonas knackmussii B13.*

Nicolas Pradervand, Sandra Sulser, Ryo Miyazaki, Veronica Bellini, Jan Roelof van der Meer. *A unique operon of three transcriptional regulators controls ICE_{clc} horizontal gene transfer in Pseudomonas knackmussii B13.*