



**HAL**  
open science

# Étude des modalités de transposition des séquences d'insertion bactériennes des familles IS91-ISCR

Aurélien Fauconnier

► **To cite this version:**

Aurélien Fauconnier. Étude des modalités de transposition des séquences d'insertion bactériennes des familles IS91-ISCR. Sciences agricoles. Université de Limoges, 2023. Français. NNT : 2023LIMO0101 . tel-04390166

**HAL Id: tel-04390166**

**<https://theses.hal.science/tel-04390166>**

Submitted on 12 Jan 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**Université de Limoges**  
**ED 652 - Biologie, Chimie, Santé (BCS)**  
**Inserm UMR 1092 : Anti-infectieux**

Thèse pour obtenir le grade de  
**Docteur de l'Université de Limoges**  
**Génomique et génétique moléculaire**

Présentée et soutenue par  
**Aurélien FAUCONNIER**

Le 19 décembre 2023

# **Étude des modalités de transposition des séquences d'insertion bactériennes des familles IS91-ISCR**

Thèse dirigée par Mme la Pr Marie-Cécile Ploy et Mme la Dr Cécile Pasternak

## JURY :

Président du jury

Mme la Pr Corinne Arpin, UMR-CNRS 5234, Université de Bordeaux

Rapporteurs

Mme la Pr Corinne Arpin, UMR-CNRS 5234, Université de Bordeaux

Mr le Dr Philippe Glaser, UMR6047, Institut Pasteur

Examineurs

Mme la Dr Virginie Libante, UMR1128 INRAE-UL, Université de Lorraine

Mr le Dr Sébastien Leclercq, UMR 1282 INRAE, Université François Rabelais

Mme la Pr Marie-Cécile Ploy, directrice de thèse, UMR 1092, Université de Limoges

Mme la Dr Sandra Da RE, UMR 1092, Université de Limoges

*À ma famille,  
À mes parents,  
À ma sœur et mon frère,  
À Cécile Pasternak,  
À tous ceux qui m'ont accompagné et aidé pour ce travail.*

*« L'imagination est plus importante que la connaissance car la connaissance est limitée tandis que l'imagination englobe le monde entier [...] Tout est mystère, et la clé d'un mystère est un autre mystère [...] Plus loin on regarde vers le passé, plus loin on voit vers l'avenir [...] Même la nuit la plus sombre prendra fin et le soleil se lèvera »*

***Albert Einstein***

**&**

***Ralph Waldo Emerson***

**&**

***Winston Churchill***

**&**

***Victor Hugo***

## Remerciements

---

*Je remercie vivement la Professeure Corinne Arpin et le Docteur Philippe Glaser, qui ont accepté d'être rapporteurs de cette thèse, ainsi que les autres membres du jury, la Docteure Virginie Libante, le Docteur Sébastien Leclercq, la Docteure Sandra Da RE et ma directrice de thèse la Professeure Marie-Cécile Ploy pour l'intérêt qu'ils ont porté à mon travail et leur présence à ma soutenance. Je tiens également à remercier tous les membres du jury pour l'encouragement dont ils m'ont fait part dans cette fin de thèse.*

*Il me sera très difficile de remercier tout le monde, car c'est grâce à l'aide de nombreuses personnes que j'ai pu mener cette thèse à son terme.*

*À Cécile Pasternak à qui je tiens à adresser des remerciements tout particuliers et chargés d'une profonde gratitude pour ta contribution exceptionnelle à la réalisation de mon projet de thèse, et ce même avant mon arrivée au laboratoire. C'est avec toi et ta disponibilité, ta passion et ton œil avisé que j'ai préparé et réussi le concours de l'ED pour l'obtention d'un financement. Ta force, ta combativité, mais aussi ta présence et ta disponibilité, malgré les défis de santé que tu as dû affronté, ont été une source d'inspiration et de motivation inestimable tout au long de ce projet de recherche. Ta détermination à m'accompagner malgré ces difficultés a été un exemple de résilience et d'engagement qui m'a profondément touché. Ta capacité à maintenir le cap, à continuer à guider mes recherches et à m'apporter un soutien infailible, même dans des moments difficiles que ce soit d'un point de vue personnel ou professionnel pour toi ou pour moi, a été véritablement remarquable. Ta patience, ta bienveillance et ta disponibilité pour discuter de mes idées plus ou moins farfelues dans les mécanismes possibles, de discuter des avancées et d'affiner nos arguments ont été inestimables et excessivement enrichissants. Tes conseils éclairés, ton expertise et ta passion pour la recherche ont été des atouts précieux pour la réussite de ces travaux de thèse. De plus, ton dévouement à l'avancement de la connaissance malgré les obstacles est une leçon précieuse que je n'oublierai jamais. Enfin, je souhaite exprimer ma profonde admiration pour ta personnalité exceptionnelle. Ta volonté de surmonter les difficultés et de continuer à apporter ton expertise aux autres est un exemple à suivre. Personnellement, Cécile, tu étais bien plus qu'une directrice de thèse, tu es et seras un modèle de force, de détermination et d'inspiration.*

*Merci du fond du cœur, Cécile, pour tout ce que tu as fait pour moi. Ton impact sur ma thèse et sur ma vie restera à jamais présent et gravé. Je te remercie avec une gratitude immense.*

*J'espère que nous pourrons avec Marie-Cécile, Sandra et Margaux terminer les dernières expériences afin de publier les derniers travaux auxquels tu participais pour te rendre hommage.*

À Marie-Cécile Ploy à qui je tiens à exprimer ma profonde gratitude pour avoir accepté de m'accueillir au sein du laboratoire UMR1092 pour ma thèse. Je souhaite également te remercier pour ta remarquable implication et ton engagement envers mon projet de recherche, en particulier dans des circonstances difficiles. Les complications avec Thomas et le décès de Cécile ont rendu cette étape de ma thèse complexe et émotionnellement exigeante. Ta décision de prendre en charge la direction de ma thèse dans de telles conditions a été un acte de bienveillance et de soutien qui dépasse largement les attentes et ce même en-dehors du labo quand nous avons des problèmes personnels. Ton expertise, ton leadership et ta compréhension ont été essentiels pour l'avancement de nos recherches, l'écriture du mémoire de thèse et de l'article sur l'élément *orf121* dans un premier temps. Ta capacité à guider, à éclairer et à encourager, même dans des moments difficiles, a été d'une valeur incommensurable. En dépit des défis, ton accueil chaleureux, ton mentorat, et ta confiance en mon travail ont été une source de motivation constante. Ton soutien m'a permis de persévérer et de progresser malgré les obstacles et difficultés de cette 3<sup>e</sup> année de thèse. Je tiens à te remercier du fond du cœur pour ton partage, ta présence bienveillante, ta disponibilité, ta générosité, ton expertise et ton leadership exceptionnel qui ont été des atouts précieux pour cette fin de thèse. Avec une reconnaissance sincère, je te remercie du fond du cœur, pour tout ce que tu as fait pour moi malgré toutes les complications.

À Sandra Da Re à qui je tiens à exprimer ma profonde gratitude pour ton soutien, ton aide précieuse, et ta présence tout au long de ma thèse, ta précieuse contribution et pour ton rôle important dans ma thèse lors de nos innombrables échanges a fait de toi une directrice de thèse officielle. Ton engagement inébranlable et ta disponibilité à répondre à mes innombrables questions ont été des atouts essentiels dans cette réussite. Je tiens également à te remercier pour ta présence constante et ton écoute bienveillante. Ton mentorat informel a contribué de manière significative à mon développement en tant que chercheur. Ton dévouement a été une source de motivation et d'inspiration tout au long de cette aventure.

Je tiens également à remercier l'ensemble de l'unité pour l'accueil sympathique qu'ils m'ont apporté, pour votre soutien, votre aide précieuse et votre bienveillance tout au long de mes travaux de thèse. Votre expertise, partagée avec générosité, a grandement contribué à l'avancement de mes recherches. Les discussions stimulantes, les conseils éclairés et la camaraderie au sein de notre équipe ont créé un environnement propice à l'épanouissement et

la réussite de ce projet. À chacun d'entre vous qui avez pris le temps de répondre à mes questions, de partager vos idées, et de créer un climat de collaboration, je vous adresse un grand merci.

Je tiens à exprimer une profonde gratitude à Margaux Gaschet pour avoir pris en charge la reprise de mes manipulations. Ton engagement, ton professionnalisme et ta précision dans l'exécution de ces tâches ont été d'une valeur inestimable pour le succès de notre projet. Ta volonté de relever ce défi avec enthousiasme et efficacité a eu un impact significatif sur l'avancement de ce projet. Travailler avec toi a été une expérience positive, et je suis reconnaissant d'avoir pu compter sur quelqu'un d'aussi fiable et compétent.

Et bien-sûr, il ne faut pas oublier Anaïs Duprat et Raya Madi à qui je tiens à exprimer ma profonde gratitude pour votre travail exceptionnel et votre dévouement remarquable tout au long de ma thèse. La préparation de centaines et de centaines de litres de LB et pas seulement, ceci a grandement contribué à la réussite de ce projet et de ces résultats. Votre dévouement sans faille pour m'aider pour la préparation du matériel et votre réactivité ont été des atouts indispensables, et ce, avec une énergie et une rigueur qui ont marqué positivement mes travaux. Au-delà du professionnel, je tiens à saluer votre gentillesse, votre soutien et vos encouragements constants. Cela a toujours été un véritable plaisir de venir vous voir, un moment où votre bonne humeur et votre positivité ont apporté un changement d'air et d'envie de continuer. Un merci spécial à Anaïs pour nous avoir fait rire avec ses anecdotes amusantes, ainsi que des éditions spéciales, créant ainsi des moments de détente bienvenus au milieu de mes journées bien remplies. Merci infiniment pour votre aide exceptionnelle et pour avoir rendu chaque interaction agréable et productive.

Je tiens à exprimer le plaisir immense que j'ai eu à travailler avec vous, Florent Di Méo et Veronica Crespi, sur l'analyse *in silico* des doigts de zinc au cours de ce projet passionnant. Cette expérience a été à la fois riche et extrêmement enrichissante, et je tiens à souligner la qualité exceptionnelle de notre collaboration. Les échanges et les discussions d'idées que nous avons eues ont été véritablement superbes. Votre expertise, votre créativité et votre dévouement ont ajouté une dimension unique à notre travail. Chacune de nos interactions a été source d'inspiration, stimulant ainsi notre réflexion et contribuant de manière significative à la qualité globale de notre analyse. Je tiens à vous remercier sincèrement pour votre engagement, votre expertise et votre volonté constante de partager vos connaissances. Merci encore pour cette expérience

collaborative mémorable. Merci encore pour votre dévouement et votre contribution précieuse à ce projet. J'espère que nous aurons l'opportunité de collaborer à nouveau à l'avenir.

Je tiens à ajouter un remerciement tout particulier à Lockman Galal pour ta bienveillance et ta gentillesse exceptionnelles tout au long de notre collaboration sur l'analyse de l'environnement génétique. Ton attitude positive et ton soutien constant ont grandement contribué à faire de cette expérience une aventure non seulement productive, mais aussi agréable. Ta capacité à créer un environnement de travail empreint de respect et de convivialité a été précieuse. Chaque interaction avec toi a été marquée par ta gentillesse, ton écoute attentive et ta volonté sincère de collaborer de manière constructive. La science peut parfois être exigeante, mais ta bienveillance a rendu chaque étape de notre travail plus fluide et agréable. Travailler avec quelqu'un d'aussi attentionné a été un véritable plaisir, et je suis reconnaissant de t'avoir comme collaborateur.

À Mariano Martinez à qui je tiens à exprimer toute ma gratitude pour son soutien et son engagement constants tout au long de mon parcours. Notre rencontre lors de mon stage de master 1 a été le point de départ d'une aventure et d'échanges précieux, et je suis profondément reconnaissant de vous avoir à mes côtés depuis lors. Ton soutien inestimable, tes encouragements sincères et ton expertise dans le domaine des western-blot ont permis de m'aider à débloquent les problèmes que je rencontrais. Je te remercie vivement pour ta disponibilité à répondre à mes questions, à discuter et à m'aider à résoudre des problèmes techniques. Je tiens également à te remercier pour ne jamais avoir cessé de m'inspirer et de m'encourager à poursuivre mes recherches sur le projet, et je n'ai jamais oublié le projet portant sur PknL avec ses fascinants petits domaines répétés « PASTA » pour lequel je te fais toujours part de mes idées. Tes conseils éclairés et ton enthousiasme ont nourri ma passion pour cette discipline de la recherche. Je tiens à t'exprimer ma profonde reconnaissance pour les moments passés à travailler ensemble dans le passé, pour ton soutien indéfectible et pour avoir été un modèle inspirant dans ma vie. Ta présence a été un atout précieux, et je suis honoré de pouvoir toujours échanger avec toi.

À Manon Vastel à qui je tiens à exprimer ma profonde gratitude pour ton soutien indéfectible, tes encouragements constants, ta gentillesse inébranlable, ta bienveillance sincère et ton aide précieuse tout au long de ce projet depuis notre connaissance. Ton impact sur mon parcours a été significatif, et je souhaite te témoigner toute ma reconnaissance. Ton soutien, qu'il soit professionnel ou personnel, a été une source d'inspiration et de force. Tes encouragements ont



illuminé les moments difficiles et ont renforcé ma confiance en moi. Ta gentillesse et ta bienveillance ont créé un environnement propice à l'épanouissement, où le travail s'est transformé en une expérience enrichissante. Ton aide a été plus que précieuse. Merci, Manon, pour ta présence exceptionnelle et bienveillante, cela a été un véritable privilège, et je te suis profondément reconnaissant.

Je tiens à exprimer ma profonde gratitude à ma sœur Fannie Fauconnier et à Alain Cazeneuve pour leurs précieuses contributions à l'écriture de ce mémoire. Votre aide inestimable dans la traque des fautes, vos soutiens constants, vos encouragements motivants, et votre bienveillance ont été des piliers essentiels de la réalisation de ce mémoire. Traquer les fautes peut souvent être un défi, mais avec votre expertise et votre attention minutieuse, nous avons réussi à polir chaque aspect de ce document. Votre rigueur et votre engagement envers l'excellence ont grandement contribué à la clarté et à la qualité de ce travail. Au-delà du travail, nos moments de détente et de défoulement ont été des instants précieux. Ces moments de convivialité ont permis de me vider la tête et d'y voir clair malgré des périodes sombres. Merci pour ces instants de partage, pour votre soutien constant et pour avoir rendu cette expérience non seulement productive, mais aussi agréable.

Pour terminer, je tiens à exprimer ma profonde gratitude à ma famille pour leur soutien inconditionnel tout au long de cette étape importante de ma vie. Leur présence constante malgré la distance, leur gentillesse inébranlable et leurs encouragements ont été des sources de réconfort et de motivation inestimables dans les épreuves difficiles. Les encouragements que vous m'avez prodigués ont été le carburant de ma persévérance. Votre confiance en moi m'a donné la force nécessaire pour aller de l'avant, même dans les moments de doute. Vos paroles bienveillantes ont été un soutien constant, faisant de chaque étape un pas vers la réussite. Merci, du fond du cœur, pour votre dévouement, votre présence, votre gentillesse et vos encouragements. Ma réussite est indissociable de votre soutien constant, et je suis reconnaissant d'avoir une famille aussi exceptionnelle à mes côtés, je vous aime fort ❤️❤️❤️. La famille est le cœur battant de la vie, un lien indéfectible tissé d'amour, de soutien et de moments partagés. Elle est le refuge qui abrite nos joies, nos peines et nos rêves, formant ainsi le socle solide sur lequel repose le récit unique de notre existence.

Merci également à de nombreux artistes qui m'ont permis d'avancer, progresser et ne pas lâcher dans tous les moments de ce projet de thèse en écoutant leurs musiques (Taylor Davis, Alina

*Gingertail, Jada Facer, Chloe Stafler, JFla, Mia Asano, TheSnakeCharmer, First To Eleven, Ally The Piper et bien d'autres encore... et la best Caitlin De Ville). Vive les cover for ever et Original Music !*

## Droits d'auteurs

---

Cette création est mise à disposition selon le Contrat :  
« **Attribution-Pas d'Utilisation Commerciale-Pas de modification 3.0 France** »  
disponible en ligne : <http://creativecommons.org/licenses/by-nc-nd/3.0/fr/>



## Table des matières

---

Introduction .....	20
Chapitre I. La dissémination des gènes de résistances aux antibiotiques par transfert horizontal de gènes .....	26
I.1. La transformation naturelle .....	28
I.2. La conjugaison .....	30
I.3. La transduction .....	30
I.4. La vésiduction .....	33
Chapitre II. Les éléments génétiques mobiles impliqués dans la dissémination des gènes de résistances aux antibiotiques .....	37
II.1. Les plasmides transmissibles .....	38
II.2. Classification des plasmides .....	39
II.3. Les éléments intégratifs conjugatifs (ICE) .....	40
II.4. Les éléments transposables procaryotes .....	44
II.4.1. Les séquences d'insertions (IS) .....	44
II.4.2. Les transposons (Tn) .....	45
II.4.2.1. Les transposons non-composites .....	45
II.4.2.2. Les transposons composites .....	46
II.4.3. Les MITEs, les MICs et les PATEs .....	46
II.5. Intégrons .....	48
Chapitre III. Les séquences d'insertion (IS) procaryotes .....	52
III.1. Classification et diversité .....	52
III.2. Les IS bactériennes « classiques » .....	57
III.2.1. Structure et caractéristiques fonctionnelles .....	57
III.2.2. Mécanismes biochimiques employés par les transposases des IS « classiques » .....	59
III.2.2.1. Les transposases de type DDE .....	59
III.2.2.2. Les transposases de type DDED .....	63
III.2.2.3. Les transposases à sérine .....	64
III.2.2.4. Les transposases à tyrosine .....	66
III.3. Les IS bactériennes atypiques à motif HUH .....	68
III.3.1. Caractéristiques structurales et fonctionnelles .....	68
III.3.2. Mécanismes biochimiques des transposases des IS atypiques .....	68
III.3.3. Les IS à transposase HUH .....	72
III.3.3.1. La famille IS200/IS605 .....	72
III.3.3.1.1. Le sous-groupe IS200 .....	73
III.3.3.1.2. Le sous-groupe IS605 .....	73
III.3.3.1.3. Le sous-groupe IS1341 .....	74
III.3.3.1.4. Les transposases de la famille IS200/IS605 .....	74
III.3.3.1.5. Rôle des extrémités RE et LE .....	75
III.3.3.1.6. Modèle de transposition de la famille IS200/IS605 .....	76
III.3.3.1.7. Lien entre transposition et fourche de réplication .....	77
III.3.3.2. Les familles IS91/ISCR .....	78
III.4. Éléments transposables eucaryotes .....	78
III.4.1. Les Hélitrons avec des transposases HUH Y2 .....	79
III.4.2. Les réplitrons avec des transposases HUH Y1 .....	80
Chapitre IV. Les séquences d'insertions bactériennes atypiques de la famille IS91/ISCR .....	83

IV.1. Les éléments de la famille IS91 .....	83
IV.1.1. Généralités .....	83
IV.1.2. Organisation génétique .....	84
IV.1.2.1. L'extrémité <i>ori</i> S .....	84
IV.1.2.2. L'extrémité <i>ter</i> S .....	86
IV.1.2.3. Contexte génétique .....	89
IV.1.3. Caractéristiques des transposases de la famille IS91 .....	90
IV.1.4. Modèle de transposition de la famille IS91 .....	95
IV.1.4.1. Modèle de transposition dit « concerté » .....	95
IV.1.4.2. Modèle de transposition dit « séquentiel » .....	96
IV.2. Les éléments de la famille ISCR apparentée à IS91 .....	99
IV.2.1. Description des ISCR .....	99
IV.2.2. Organisation génétique .....	101
IV.2.3. Caractéristiques des transposases des ISCR .....	102
Chapitre V. Régulation de la transposition et impact des IS chez les procaryotes .....	107
V.1. La transcription chez les procaryotes .....	107
V.1.1. Le promoteur .....	107
V.1.2. L'ARN polymérase .....	108
V.1.3. La transcription .....	110
V.2. La traduction chez les procaryotes .....	112
V.3. Contrôle transcriptionnel chez les IS .....	114
V.3.1. Promoteurs endogènes .....	114
V.3.2. Promoteurs transitoires .....	114
V.3.3. Termineur de transcription .....	116
V.3.4. Protection contre la transcription « impinging » .....	116
V.3.5. Couplage transcriptionnel .....	117
V.3.6. Stabilité de l'ARNm .....	117
V.4. Contrôle de la traduction des IS .....	117
V.4.1. ARN antisens .....	117
V.4.2. ARNm sans leader .....	118
V.4.3. Séquestration des signaux d'initiation de la traduction .....	118
V.4.4. Couplage traductionnel .....	118
V.5. Les autres niveaux de régulation des IS .....	119
V.5.1. Contrôles post-traductionnels .....	119
V.5.2. Rôle des facteurs de l'hôte .....	119
V.5.3. Rôle des facteurs environnementaux .....	120
Objectifs de thèse .....	122
Partie I. Étude <i>in silico</i> de la diversité des éléments (IS91, IS801, IS1294b) et du rôle d' <i>orf121</i> dans la mobilité de l'élément IS91 .....	123
I.1. L' <i>orf121</i> codé par IS91 régule négativement la transposition d'IS91 .....	123
I.2. Résultats complémentaires de l'analyse <i>in silico</i> .....	185
I.3. Discussion .....	195
Partie II. Étude structurale de la transposase IS91 .....	200
II.1. Matériels et méthodes .....	201
II.1.1. Analyse <i>in vivo</i> .....	201
II.1.2. Prédiction de la structure par AlphaFold2 .....	202
II.1.3. Modélisation de la dynamique de TnpA <sub>IS91</sub> .....	203

II.2. Résultats <i>in vivo</i> .....	204
II.3. Résultats <i>in silico</i> .....	209
II.4. Discussion .....	218
Partie III. Étude de la transposition d'autres éléments de la famille IS91 : ISCR et ISKpn22 .....	224
III.1. Matériels et Méthodes .....	224
III.2. Résultats.....	226
III.2.1. Analyse des transposases.....	226
III.2.2. Efficacité de transposition.....	227
III.3. Discussion .....	232
Conclusion et perspectives.....	235
Références bibliographiques .....	239
Annexes.....	272

## Table des illustrations

---

Figure 1 : Chronologie de la découverte des antibiotiques et des résistances aux antibiotiques	21
Figure 2 : Mécanisme d'action des antibiotiques	22
Figure 3 : Présentation des différents mécanismes moléculaires de la résistance aux antibiotiques	23
Figure 4 : Présentation des principaux mécanismes de transfert de matériel génétique	27
Figure 5: schéma simplifié du processus de transformation chez les bactéries Gram négatives et Gram positives	29
Figure 6 : Cycle lytique et lysogénique des phages	31
Figure 7 : Présentation des 3 mécanismes de transduction	32
Figure 8 : Représentation du transfert horizontal de gènes par l'intermédiaire d'une vésicule membranaire (VM)	34
Figure 9 : Diversité des éléments génétiques mobiles	37
Figure 10 : Schéma de la constitution génétique des plasmides dits transmissibles	39
Figure 11 : Contexte génétique d'un ICE	41
Figure 12 : Cycle de vie d'un élément intégratif conjugatif (ICE)	42
Figure 13 : Représentation schématique lors de l'intégration au site spécifique d'un ICE	43
Figure 14 : Schéma d'une séquence d'insertion	44
Figure 15 : Transposons non-composite et composite	45
Figure 16 : Origine des MITEs, MICs et PATEs	47
Figure 17 : Les différentes régions d'un intégron et son fonctionnement	49
Figure 18 : Expansion et élimination des séquences d'insertions induisant la rationalisation du génome	54
Figure 19 : Répartition des familles d'IS	56
Figure 20 : Principales voies de transposition des transposases DDE	60
Figure 21 : Les deux modes principaux de transposition	62
Figure 22 : Mécanisme des recombinases à sérine	65
Figure 23 : Voies proposées d'excision et d'intégration par les transposases à tyrosine	67
Figure 24 : Organisation de certaines endonucléases de la superfamille HUH	69
Figure 25 : Mécanisme de réaction des protéines HUH	70
Figure 26 : Organisation génétique des membres de la famille IS200/IS605	72
Figure 27 : Structure cristallographique de la transposase TnpA <sub>IS608</sub> en interaction avec l'extrémité RE	75
Figure 28 : Modèle de transposition de l'élément IS608	77

Figure 29 : Orientation de l'insertion et l'excision de IS608 par rapport à la direction de la fourche de réplication.....	78
Figure 30 : Caractéristiques du transposon Helraiser.....	80
Figure 31 : Les endonucléases HUH et les réplitrons.....	81
Figure 32 : Alignement des régions <i>oriS</i> des éléments IS91, IS801, IS1294 et IS1294b.....	84
Figure 33 : Structures secondaires des régions <i>oriS</i> des éléments IS91, IS801, du phage $\phi$ X174 et du plasmide pC194 .....	86
Figure 34 : Alignement des régions <i>terIS</i> des éléments IS91, IS801, IS1294 et IS1294b ....	87
Figure 35 : Principe du mécanisme « <i>one-ended transposition</i> ».....	88
Figure 36 : Organisation génétique des membres actifs connus de la famille IS91 .....	89
Figure 37 : Alignement des domaines conservés des éléments IS91, IS801 et IS1294b avec les protéines Rep et la protéine gpA .....	91
Figure 38 : Modèle de transposition de la famille IS91 dit « concerté » .....	96
Figure 39 : Modèle de transposition de la famille IS91 dit « séquentiel ».....	98
Figure 40 : Alignement des domaines conservés des éléments ISCR1, ISCR2, ISCR3, ISCR4 et ISCR5 avec les éléments de la famille IS91 (IS91, IS801 et IS1294b) et les protéines Rep et la protéine gpA.....	103
Figure 41 : Organisation du promoteur bactérien .....	107
Figure 42 : Core de l'ARN polymérase. ....	109
Figure 43 : Interaction entre l'ARN polymérase et les différentes régions d'un promoteur..	110
Figure 44 : Ancrage initial de l'holoenzyme de l'ARN polymérase à un promoteur. ....	111
Figure 45 : Organisation du site de liaison du ribosome (RBS).....	113
Figure 46 : Augmentation de l'expression du gène <i>tnpA</i> ou activation des gènes par l'insertion d'une IS.....	115
Figure 47 : Distribution des polymorphismes des nucléotides du promoteur $P_{orf121}$ .....	185
Figure 48 : Distribution des polymorphismes des nucléotides du promoteur $P_{tnpA}$ .....	186
Figure 49 : Distribution des polymorphismes des extrémités <i>terIS1294b</i> et <i>oriIS1294b</i> .....	187
Figure 50 : Distribution des polymorphismes des extrémités <i>terIS91</i> et <i>oriIS91</i> .....	188
Figure 51 : Distribution des polymorphismes de la transposase $tnpA_{IS1294b}$ .....	190
Figure 52 : Distribution des polymorphismes de la transposase $tnpA_{IS91}$ .....	192
Figure 53 : Organisation génétique de l'élément IS91.....	192
Figure 54 : Séquence peptidique d'Orf121 et distribution des polymorphismes d'Orf121 ...	193
Figure 55 : Effet <i>in vivo</i> des mutations ponctuelles des transposases du motif CxxC de Orf121 et TnpA en C-terminale.....	198
Figure 56 : Orf121 un dimère ?.....	199
Figure 57 : Effet des mutations ponctuelles des cystéines du domaine I et VI de la transposase $TnpA_{IS91}$ .....	206



Figure 58 : Prédiction de la structure globale de la transposase TnpA <sub>IS91</sub> et la matrice d'erreur d'alignement prédite.....	210
Figure 59 : Position des domaines important de la structure globale de la transposase TnpA <sub>IS91</sub> .....	211
Figure 60 : Impact de la présence de zinc ou non sur la structure globale de la transposase TnpA <sub>IS91</sub> .....	213
Figure 61 : Effet des mutations ponctuelles sur la flexibilité ou la rigidité des domaines de la transposase TnpA <sub>IS91</sub> .....	214
Figure 62 : Effet des mutations ponctuelles du ZF1 de la transposase TnpA <sub>IS91</sub> .....	215
Figure 63 : Effet des mutations ponctuelles du ZF2 de la transposase TnpA <sub>IS91</sub> .....	217
Figure 64 : Localisation et structure des doigts de zinc ZF1 et ZF2 de TnpA <sub>IS91</sub> .....	218
Figure 65 : Un sillon potentiel permettant le passage de l'ADN au niveau de la transposase TnpA <sub>IS91</sub> .....	219
Figure 66 : Superposition des structures AlphaFold2 prédites des transposases IS91, IS801 et IS1294b.....	220
Figure 67 : Représentation hypothétique du doigt de zinc ZF1 des transposases IS91, IS1294b, IS801, ISKpn22, ISCR1 et ISCR2.....	222
Figure 68 : Alignement des transposases TnpA <sub>IS91</sub> , TnpA <sub>IS801</sub> , TnpA <sub>IS1294b</sub> , TnpA <sub>ISKpn22</sub> , TnpA <sub>ISCR1</sub> et TnpA <sub>ISCR2</sub> .....	226
Figure 69 : Effet <i>in vivo</i> des mutations ponctuelles des transposases ISKpn22, ISCR1 et ISCR2 sur une IS mimant IS91, ISCR1 et ISCR2.....	230
Figure 70 : Contexte génétique de quatre gènes en amont et en aval de l'élément IS91 ...	236

## Table des tableaux

---

Tableau 1 : Principales caractéristiques et résumé de chaque famille et sous-groupe d'IS..	55
Tableau 2 : Souches et plasmides utilisés lors de l'étude des domaines I et VI.....	202
Tableau 3 : Analyse des sites cibles d'insertion en fonction des mutations ponctuelles des cystéines du domaine I et VI.....	207
Tableau 4 : Pourcentage de OET observé en fonction des mutations ponctuelles des cystéines des domaines I et VI.....	208
Tableau 5 : Souches et plasmides utilisés lors de l'étude de la transposition des éléments ISCR et ISKpn22.....	225
Tableau 6 : Pourcentage de OET observé en fonction des mutations ponctuelles des transposases ISKpn22, ISCR1 et ISCR2 une IS mimant IS91, ISCR1 et ISCR2.....	231

## Liste des abréviations

---

ADN	Acide DésoxyriboNucléique
ADNdb (ADNds)	Acide DésoxyriboNucléique double brin
ADNsb (ADN ss)	Acide DésoxyriboNucléique simple brin
ARN	Acide RiboNucléique
ARNt	Acide RiboNucléique de transfert
ARNm	Acide RiboNucléique messenger
ARNr	Acide RiboNucléique ribosomique
ACP	Acyl carrier protein
A	Adénine (acide nucléique) ou Alanine (acide aminé)
AF2	AlphaFold2
Å	Angstrom
R	Arginine
N	Asparagine
D	aspartate / Acide aspartique
orf	Cadre ouvert de lecture
CR	Cercle Roulant
C	Cytosine (acide nucléique) ou Cystéine (acide aminé)
ZF	doigt de zinc / Zinc Finger
ZF1	doigt de zinc 1 / Zinc Finger 1
ZF2	doigt de zinc 2 / Zinc Finger 2
ICE	Elément Intégratif Conjugatif
EGM	Elément Génétique Mobile
RE	extrémité droite
LE	extrémité gauche
E	glutamate / Acide glutamique
Q	Glutamine
G	Guanine (acide nucléique) ou Glycine (acide aminé)
H	Histidine
IHF	Integration Host Factor
I	Isoleucine
kb	Kilobase
kDa	kiloDalton
L	Leucine
K	Lysine
Mb	Mégabase
M	Méthionine
MITE	Miniature Inverted Repeat Transposable Element
MIC	Mobile Insertion Cassette
OET	One-Ended Transposition
<i>oriT</i>	origine de Transfert
pb	paire de base

PATE	Palindrome-Associated Transposable Element
pLDDT	per-residue Local Distance Difference Test
F	Phénylalanine
PAE	Predicted Aligned Error
P	Proline
P <sub>OUT</sub>	Promoteur dirigé vers l'extérieur
T4CP	Protéine de Couplage de Type IV
PCR	Réaction en chaîne par polymérase
3'CS	région conservée en 3'
5'CS	région conservée en 5'
DR	Répétitions Directes
U	résidu hydrophobe volumineux
<i>rcr</i>	Rolling Circle Replication
RMSD	Root-Mean-Square Deviation
RMSF	Root-Mean-Square Fluctuations
SOS	Save Our Souls
IS	Séquence d'Insertion
ISCR	Séquence d'Insertion de Région Commune
IR	Séquence Inversée Répétée
IRR	Séquence Inversée Répétée droite
IRL	Séquence Inversée Répétée gauche
S	Sérine
SD	Shine-Dalgarno
RBS	site de liaison du ribosome
T4SS	Système de Sécrétion de Type IV
T	Thréonine
T	Thymine
THG	Transfert Horizontal de Gènes
<i>TnpA/tnpA</i>	transposase
Tn	Transposon
W	Tryptophane
Y	Tyrosine
PTU	Unités Taxonomiques des Plasmides
U	Uracile
V	Valine
VM	Vésicule Membranaire
WT	Wild Type
Zn	Zinc

## Introduction

---

Depuis plusieurs millénaires, l'Homme a connu de nombreuses pandémies d'origine bactérienne dont la plus connue est la peste. Elle s'étale sur plusieurs siècles et est causée par l'agent *Yersinia pestis*. Ce pathogène a notamment été à l'origine de la peste de Justinien au VI<sup>ème</sup> siècle tuant plusieurs millions de personnes (Harbeck et al., 2013), ou encore de la peste noire qui décima au XIV<sup>ème</sup> siècle entre 30% et 50% des populations touchées (Haensch et al., 2010; Halioua, 2004).

En parallèle, durant ces millénaires, les bactéries et les champignons sont utilisés par l'Homme pour leur action antibiotique sous la forme de cataplasme traditionnel de pain moisi. Ce cataplasme fut utilisé pour traiter les plaies ouvertes à travers le monde, en Chine, en Grèce ou encore en Égypte, depuis plus de 2000 ans. A ce jour, le plus ancien document connu et conservé est le papyrus de Eber datant de 1550 av. J.-C. Sur ce dernier, on y trouve une liste de remèdes utilisant du pain moisi et de la terre (Haas, 1999). Ainsi et depuis plus de 1000 ans, les anglo-saxons savaient préparer une recette qui se révélait capable d'anéantir *Staphylococcus aureus* pourtant résistant à la méthicilline (Harrison et al., 2015).

Avant la découverte des antibiotiques, Louis Pasteur décrivait déjà un phénomène appelé antibiose, en mettant en lumière qu'une bactérie pouvait sécréter des substances permettant de tuer une autre bactérie (Brunel, 1951). Durant les années 1890, Emmerich et Löw ont utilisé un composé, appelé pyocyanase, produit par *Pseudomonas aeruginosa*, ayant une activité contre de nombreux pathogènes (Emmerich, Rudolf & Löw, Oscar, 1889). En dépit de leurs découvertes, le développement d'antibiotiques est attribué à Paul Ehrlich qui a conçu le salvarsan en 1910 afin de traiter la syphilis causée par *Treponema pallidum* (Gelpi et al., 2015).

En 1928, le premier agent antimicrobien faisant partie de la famille des  $\beta$ -lactamines a été découvert par Alexander Fleming et est connu sous le nom de pénicilline (Fleming, 1929). Son efficacité a été démontrée par son action inhibitrice sur la croissance bactérienne de *Staphylococcus* (Aminov, 2010).

L'Age d'or des antibiotiques est marqué entre les années 1940 et 1960. Durant cette période, il a été découvert plusieurs classes d'antibiotiques. Bien que l'ensemble de ces antibiotiques soient actuellement utilisés en clinique, leur efficacité a fortement diminué en raison de l'augmentation de la résistance aux antibiotiques. En effet, lors de la découverte d'un antibiotique, les premières résistances peuvent avoir lieu les premières années suivant la mise sur le marché de l'antibiotique (figure 1). Ce phénomène a été amplifié par, entre autres, une utilisation abusive et non adéquate des antibiotiques entraînant *in fine* une augmentation rapide de la résistance (Prescott, 2014).

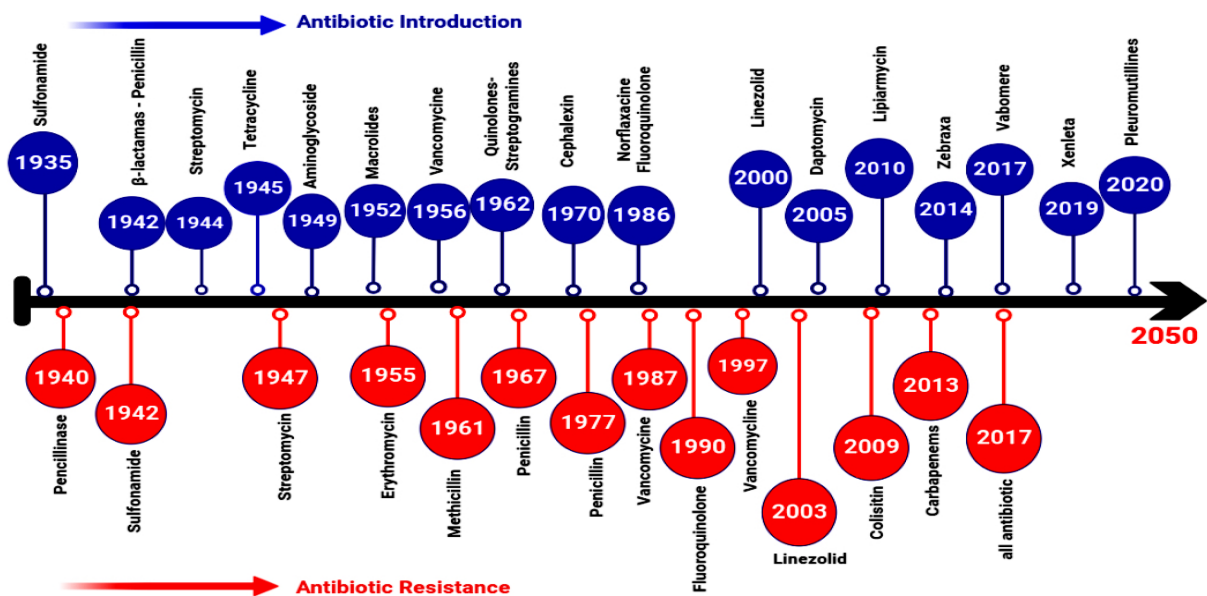


Figure 1 : Chronologie de la découverte des antibiotiques et des résistances aux antibiotiques

Issue de Helmy et al., 2023

*En Haut et en bleu : année de la découverte de l'antibiotique. En bas et en rouge : année à laquelle la résistance à un antibiotique a été rapportée.*

Une étude récente a permis d'estimer que la résistance aux antibiotiques était associée de 4,95 millions de décès dans le monde en 2019, dont 1,27 millions directement attribuables à la résistance (Murray et al., 2022). De plus, cette résistance bactérienne aux antibiotiques associée à des résistances à d'autres antimicrobiens (antiviraux, antiparasitaires) entraînerait le décès d'au moins 10 millions de personnes par an d'ici 2050. Étant précisé que cela représentera évidemment un coût économique important pouvant atteindre jusqu'à 100 000 milliards de dollars (O'Neill, 2016, 2014).

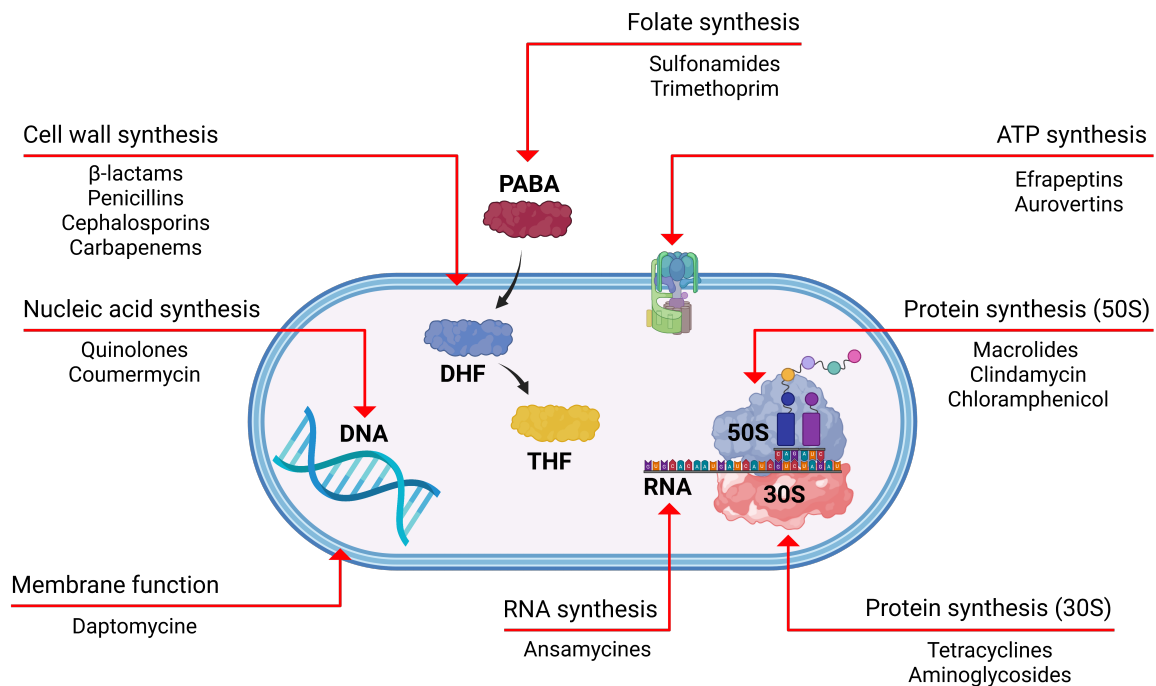


Figure 2 : Mécanisme d'action des antibiotiques

Inspirée de Mammeri, 2013

Aujourd'hui, près de 80 % des antibiotiques proviennent du genre *Streptomyces* vivant dans le sol (Barka et al., 2016).

Les antibiotiques utilisés pour le traitement des infections bactériennes sont souvent classés en fonction de leur mécanisme d'action principal. Il existe six modes d'action principaux (figure 2) :

- Interférence avec la synthèse de la paroi cellulaire, comme les β-lactamines qui ciblent les enzymes de synthèse du peptidoglycane, entraînant la lyse de la bactérie (Džidić et al., 2008; Kahne et al., 2005; Reynolds, 1989).
- Inhibition de la synthèse des protéines. La biosynthèse des protéines est catalysée par les ribosomes ainsi que d'autres facteurs cytoplasmiques. Le ribosome bactérien 70S est composé de deux sous-unités ribonucléoprotéiques, les sous-unités 30S et 50S (Yoneyama and Katsumata, 2006). Certains antibiotiques inhibent la biosynthèse des protéines en ciblant la sous-unité 30S (tétracyclines) ou 50S (chloramphenicol, macrolides) du ribosome (Johnston et al., 2002; Vannuffel and Cocito, 1996; Yoneyama and Katsumata, 2006).
- Interférence avec la synthèse des acides nucléiques, comme les fluoroquinolones qui inhibent l'enzyme ADN gyrase qui est formée de deux sous-unités A et de deux sous-unités B. Le rôle de l'ADN gyrase est d'entailler l'ADN double brin et d'introduire des supercoils négatifs pour refermer ensuite les extrémités entaillées. Les fluoroquinolones se lient à la sous-unité A de la gyrase et interfèrent avec sa fonction de coupure et de recollage des brins. Chez les bactéries

à Gram positif, les fluoroquinolones ciblent la topoisomérase IV qui coupe et sépare le brin d'ADN fille après la réplication de l'ADN (Higgins et al., 2003; Wise, 1999; Yoneyama and Katsumata, 2006).

- Inhibition de la voie métabolique de l'acide folique : les sulfamides et le triméthoprim sont des inhibiteurs distincts dans le métabolisme de l'acide folique (Yoneyama and Katsumata, 2006). Dans cette voie métabolique, les sulfamides inhibent la dihydroptéroate synthase et le triméthoprim agit à un stade ultérieur de la synthèse de l'acide folique et inhibe l'enzyme dihydrofolate réductase.
- Inhibition de la fonction membranaire, comme la daptomycine qui crée une perturbation en se liant à la membrane cellulaire conduisant à une dépolarisation rapide du potentiel membranaire qui aboutit finalement à la mort de la bactérie (Heidary et al., 2018; Taylor and Palmer, 2016).
- Inhibition de l'ATP Synthase, comme les antibiotiques utilisés en médecine vétérinaire, les efrapeptines et aurovertines qui inhibent à la fois la synthèse et l'hydrolyse de l'ATP par l'ATP synthase (Abrahams et al., 1996; van Raaij et al., 1996).

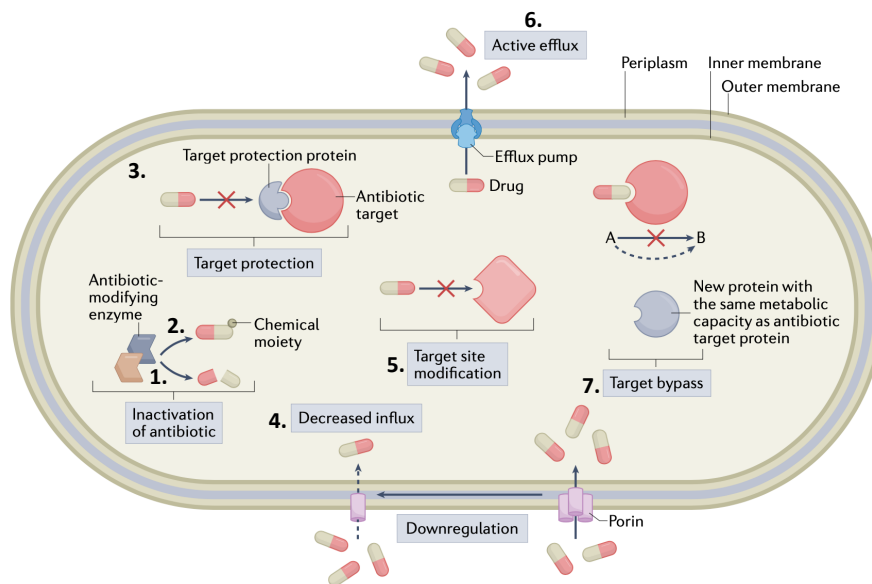


Figure 3 : Présentation des différents mécanismes moléculaires de la résistance aux antibiotiques

Adaptée de Darby et al., 2023

1. L'inactivation de l'antibiotique implique une enzyme capable d'hydrolyser le groupe fonctionnel de l'antibiotique.
2. Modification de l'antibiotique par l'ajout d'un groupement chimique (acétyle, phosphate...).
3. La protection de la cible implique l'association entre le site cible de l'antibiotique et une protéine de protection.
4. La diminution de l'influx est médiée par des modifications de la structure membranaire en diminuant la perméabilité membranaire.
5. La modification de la cible comprend diverses altérations de la cible, telles que la méthylation de l'ARNr 23S ou de l'ARNr 16S, des altérations des précurseurs du peptidoglycane.
6. Les pompes d'efflux éliminent l'antibiotique de la cellule en utilisant l'énergie provenant de l'hydrolyse de l'ATP dans les pompes ABC.
7. Le contournement des cibles implique la création de cibles antibiotiques supplémentaires ou de sous-unités qui ne sont pas susceptibles de se lier à l'antibiotique.



Afin de survivre en présence d'antibiotiques, les bactéries ont mis en place différents mécanismes de résistance par (figure 3) :

- L'inactivation de l'antibiotique, comme la production de  $\beta$ -lactamases qui confère une résistance aux  $\beta$ -lactamines (Ur Rahman et al., 2018).
- La modification de la cible, comme la mutation de l'ADN gyrase ou de la topoisomérase IV dans la résistance aux quinolones et fluoroquinolones (Correia et al., 2017).
- La protection de la cible, comme la protéine de la cassette F (ABC-F) de liaison à l'ATP dans la protection du ribosome par les macrolides (Sharkey et al., 2016).
- Le remplacement de la cible, comme dans la résistance aux sulfamides où une dihydroptéroate synthase moins affine pour ces molécules est synthétisée (Sköld, 2000).
- L'efflux, qui exporte les antibiotiques hors des bactéries comme la tétracycline, les aminoglycosides ou  $\beta$ -lactamines (Markley and Wencewicz, 2018; Pachori et al., 2019; Zapun et al., 2008).
- La diminution de perméabilité membranaire (Pachori et al., 2019; Zapun et al., 2008).

Néanmoins, il ne fait nul doute que les antibiotiques permettent de sauver d'un décès prématuré de nombreuses personnes ayant eu à lutter contre des infections bactériennes ; ils ont également permis de favoriser le développement économique (Cole, 2014). Cependant, le mésusage des antibiotiques engendre une résistance aux antibiotiques des agents pathogènes qui continuent d'émerger, de se transmettre rapidement et de se diversifier par la sélection de mutations spontanées. A titre d'illustration, le Centers for Disease Control and Prevention (CDC) a estimé qu'aux États-Unis, 20 % des infections résistantes aux antibiotiques du pays sont attribuables à l'utilisation de ces agents antimicrobiens dans les activités agricoles (CDC, 2013; Spellberg et al., 2016). Ainsi, il en ressort que l'utilisation non appropriée des antibiotiques génère une pression sélective ayant pour conséquence directe de favoriser la survie des bactéries résistantes aux antibiotiques utilisés, et *in fine* leur dissémination (Laxminarayan et al., 2013).

Il est important de noter que la résistance aux antibiotiques n'est pas récente, il s'agit d'un phénomène naturel car de nombreux micro-organismes du sol produisent des antibiotiques et se préservent eux-mêmes de leur production par des mécanismes de résistance.

L'augmentation de la résistance chez des espèces pathogènes sous l'effet de la pression sélective et des transmissions de résistances entre bactéries mais aussi entre individus (hommes, animaux) est extrêmement préoccupante (Levy and Marshall, 2004; O'Neill, 2014). Des études ont détecté aussi des bactéries résistantes ou gènes de résistance aux antibiotiques dans d'autres environnements tels que l'eau, le sol et l'air (Forsberg et al., 2012; Li et al., 2018; Zhang et al., 2009). Cette résistance aux antibiotiques est ainsi une problématique mondiale qui implique tous les écosystèmes et menace la santé humaine (WHO, 2014).

Les différents mécanismes de résistances aux antibiotiques peuvent être transférés de deux manières :

- Soit la résistance est portée par le chromosome : Dans ce cas, la transmission est verticale. La résistance est transmise à la descendance lorsque la bactérie se divise (Ellabaan et al., 2021). Des mutations spontanées peuvent conférer un avantage à la bactérie dans un environnement en présence d'antibiotiques (Binet and Maurelli, 2005; Wang et al., 2001).
- Soit cette résistance est portée par des éléments génétiques mobiles : Dans ce cas, le transfert de gène est latéral *via* le Transfert Horizontal de Gènes (THG) d'une bactérie à l'autre par différents mécanismes comme la transformation, la transduction ou la conjugaison (Stevenson et al., 2017). Le THG permet d'une part d'acquérir du matériel génétique supplémentaire, comme les gènes de résistance aux antibiotiques et d'autre part, il permet l'échange entre différentes espèces (Paquola et al., 2018). Ce nouveau patrimoine génétique joue un rôle important dans l'acquisition, l'accumulation et la dissémination des gènes de résistance aux antibiotiques (Burrus and Waldor, 2004). La présence d'antibiotique peut avoir une incidence sur l'augmentation de la résistance aux antibiotiques en induisant davantage le THG (Headd and Bradford, 2018; Prudhomme et al., 2006).

## Chapitre I. La dissémination des gènes de résistances aux antibiotiques par transfert horizontal de gènes

---

Le transfert horizontal de gènes (THG) est un processus qui permet à de l'ADN simple brin (ADNsb) ou double brin (ADNdb) d'être transféré d'une bactérie à une autre. L'ensemble des différents THG est présenté en figure 4 (Sun, 2018, 2016). Ces transferts de gènes ont joué un rôle majeur dans la plasticité des génomes bactériens, ainsi que dans leur évolution et leur adaptation, comme le transfert de gènes de virulence et de gènes de résistance aux antibiotiques (Lerminiaux and Cameron, 2019; Sobecky and Hazen, 2009).

Parmi les trois mécanismes canoniques du THG (Arnold et al., 2022; Soucy et al., 2015), les plus couramment étudiés et connus sont :

- La transformation avec de l'ADN libre,
- La transduction par des bactériophages,
- La conjugaison impliquant des plasmides et des éléments conjugatifs intégratifs.

Étant précisé que la conjugaison est le mode le plus important et largement répandu chez les bactéries (Arnold et al., 2022; Soucy et al., 2015). Plus récemment, un nouveau mécanisme de THG appelé « vésiduction » a été identifié, dans lequel le transfert de gène est médié par des vésicules membranaires (VMs) (Cai et al., 2016; Domingues and Nielsen, 2017).

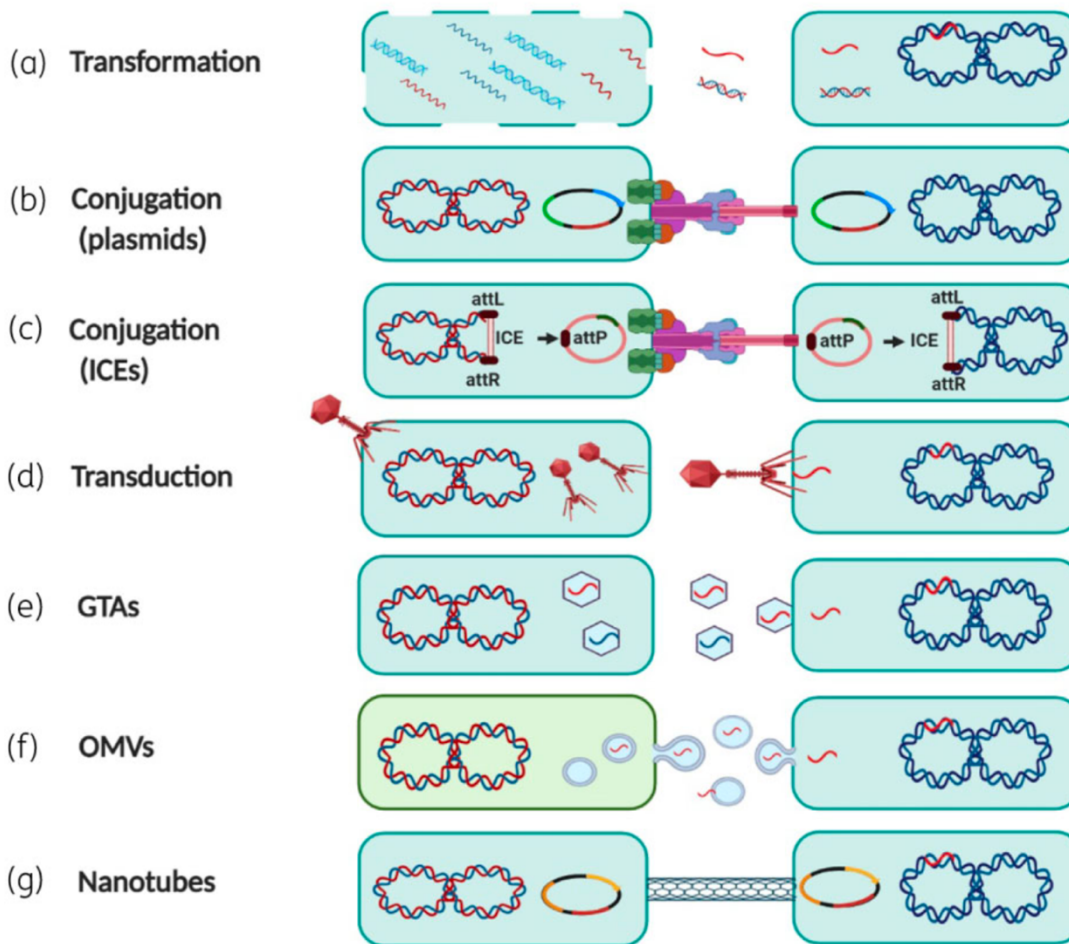


Figure 4 : Présentation des principaux mécanismes de transfert de matériel génétique

Issue de Liu et al., 2022

- (a) La transformation permet l'incorporation d'ADN dit « nu » qui est sous la forme simple ou double brin de l'environnement par une bactérie compétente.
- (b) La conjugaison des plasmides est médiée par le T4SS permettant un contact physique entre deux cellules et le transfert de celui-ci.
- (c) Les ICE sont présents sur le chromosome de la bactérie donneuse. Après son excision au niveau des sites d'attachement, attL et attR. Ils permettent le passage d'une forme linéaire à circulaire. Cette forme circulaire pourra être transférée en utilisant la conjugaison médiée par un T4SS. Dès la transmission à la cellule receveuse, l'ICE est recircularisé, puis il s'intègre dans le génome de ce nouvel hôte.
- (d) La transduction est le transfert de matériel génétique médié par des bactériophages.
- (e) Les agents de transfert de gènes (GTAs) sont des séquences dérivées de phages permettant l'encapsulation de l'ADN. Ainsi la transduction qui s'en suit permet le transfert du matériel génétique dans une autre bactérie.
- (f) La vésiduction permet la formation de vésicule membranaire pouvant encapsuler de l'ADN par une bactérie émettrice. Elle permet aussi le transfert de matériel génétique dans une bactérie réceptrice.
- (g) Les nanotubes sont des structures qui permettent le transfert de plasmides non conjugatifs entre bactéries.

## I.1. La transformation naturelle

En 1928, Griffith identifiait la transformation naturelle lors d'une expérience au cours de laquelle des souris étaient infectées soit avec des souches de *Streptococcus pneumoniae* virulentes dotées d'une capsule, soit avec des souches non-capsulées non virulentes. Dans le premier cas, l'infection était létale pour les souris tandis que dans le second cas les souris survivaient. En revanche, si l'infection était réalisée avec un mix de pneumocoques non capsulés et de pneumocoques capsulés mais préalablement inactivés, alors l'infection était létale pour les souris. Par ces expériences, Griffith a pu démontrer que la souche au préalable avirulente devenait virulente (Griffith, 1928).

Ce n'est que bien plus tard que les chercheurs ont été en mesure de démontrer qu'il s'agissait de l'acquisition de gène codant la capsule par la souche non capsulée. Cela lui permettait de devenir virulente.

La transformation est un mécanisme biologique permettant l'acquisition de l'ADN exogène en 3 étapes (Yu et al., 2022) :

- La capture,
- L'internalisation,
- L'intégration dans le génome bactérien.

La transformation est un phénomène qui permet de transférer des fragments d'ADN allant de 7 à 50 kb et souvent impliqués dans l'acquisition de résistance aux antibiotiques (Blokesch, 2017; Cai et al., 2016; Johnston et al., 2014). Il est estimé que 1 % des espèces bactériennes sont naturellement transformables (espèces dites compétentes) avec des représentants chez les bactéries à Gram positif et à Gram négatif (Nazarian et al., 2018). L'ADN transféré peut s'intégrer ensuite au génome de la bactérie receveuse par recombinaison homologue à condition que les espèces donneuses et receveuses soient phylogénétiquement proches.

Bien que le principe global de la transformation soit similaire, les enveloppes bactériennes des bactéries à Gram positif et à Gram négatif révèlent des différences, ainsi que certaines contraintes. La régulation de la compétence naturelle nécessite plus d'une vingtaine de gènes différents, ils sont présentés dans la figure 5 de manière partielle et simplifiée (Mell and Redfield, 2014). L'ADN extracellulaire est capturé par la machinerie de transformation via un pilus et il est internalisé sous la forme d'ADN simple brin dans la cellule via les protéines Com. Ensuite, l'ADN est pris en charge pour être intégré dans le génome bactérien par recombinaison homologue.

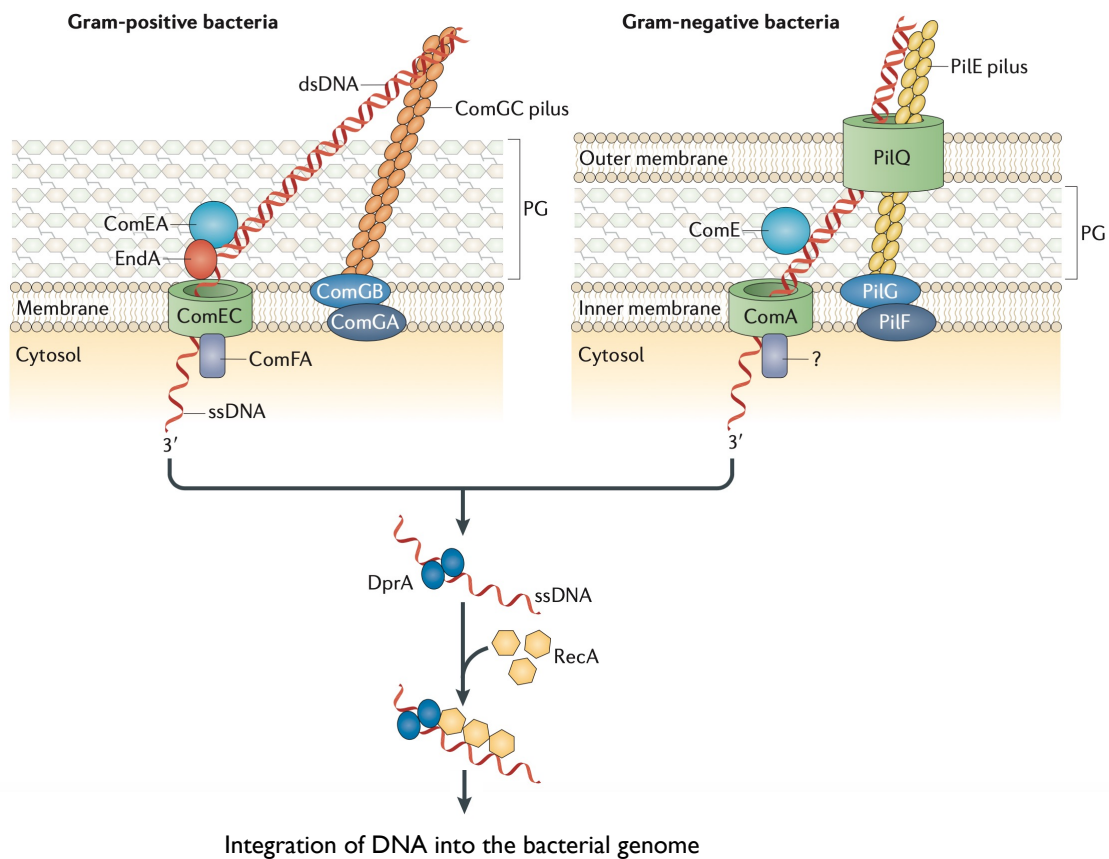


Figure 5: schéma simplifié du processus de transformation chez les bactéries Gram négatives et Gram positives.

Adaptée de Johnston et al., 2014

Chez les bactéries à Gram positif, l'ADNdb exogène est capté par le pilus de transformation qui est composé des sous-unités ComGC. Par ailleurs, chez les Firmicutes, lors de l'internalisation de l'ADN, des protéines supplémentaires sont nécessaires (ComGA et ComGB). Après la réception de l'ADNdb par ComEA, celui-ci est dégradé en ADNsb par la nucléase EndA (chez *Streptococcus pneumoniae*). L'ADNsb est transporté vers le cytoplasme via le pore transmembranaire ComEC. La translocase ATPase-dépendante ComFA a été identifiée uniquement chez les Firmicutes, par laquelle ComFA permet de piloter l'internalisation de l'ADNsb dans le cytoplasme.

Chez les bactéries à Gram négatif, le canal à sécrétine PilQ de la membrane externe permet au pilus PilE de capter l'ADNdb exogène pour le transporter dans le périplasma. Dans le périplasma, l'ADNdb est réceptionné par la protéine ComE et dégradé par une nucléase inconnue à ce jour. Ensuite, l'ADNsb va être internalisé dans le cytoplasme via le pore ComA.

Lorsque l'ADN exogène est dans le cytoplasme, le mécanisme est similaire chez les bactéries à Gram positif et à Gram négatif. L'ADN internalisé est reconnu par des protéines de liaison à l'ADNsb et il est lié à la protéine DprA qui recrute la recombinaise RecA. Ensuite, RecA polymérise le long de l'ADNsb et permet l'intégration de cet ADN dans le génome de la bactérie par recombinaison homologue.

D'un point de vue clinique, ce mécanisme a démontré son intérêt pour l'acquisition de gènes de résistance à différents antibiotiques pour un certain nombre de pathogènes humains (*Acinetobacter baumannii*, *Neisseria meningitidis*, *Streptococcus pneumoniae* ...) (Bowler et al., 1994; Janoir et al., 1999; Lu et al., 2020; Traglia et al., 2019; Winter et al., 2021). Il a été démontré que des désinfectants (chlore, triclosan) augmentaient la compétence (Jin et al., 2020; Lu et al., 2020).

## **I.2. La conjugaison**

Lederberg et Tatum en 1946 ont démontré l'existence d'un nouveau mécanisme impliquant un plasmide qui permet le transfert de gènes chez *Escherichia coli* (Lederberg and Tatum, 1946). La conjugaison, définie comme l'échange d'information génétique d'une bactérie dite donneuse vers une autre bactérie dite receveuse, peut se produire principalement soit entre bactéries à Gram négatif, soit entre à Gram positif. Dans de rares cas, la conjugaison peut toutefois avoir lieu entre les deux types de Gram, tel que *Enterococcus faecalis* à Gram positif et *Escherichia coli* à Gram négatif (Abajy et al., 2007; Arends et al., 2013).

Il convient de préciser que le processus de la conjugaison est initié par la cellule donneuse qui doit établir un contact avec la surface de la cellule receveuse. Ce contact intervient grâce au pilus conjugatif chez les bactéries à Gram négatif. En général, les gènes nécessaires à la conjugaison se trouvent sur un plasmide dit conjugatif, ces derniers seront étudiés lors de la sous-partie II.1.2 du prochain chapitre.

## **I.3. La transduction**

En 1952, Zinder et Lederberg démontraient que l'infection des bactéries par des bactériophages permettait la transmission de gènes entre souches bactériennes (Zinder and Lederberg, 1952). Alors que nous pensions que tous les transferts de gènes médiés par les phages se produisaient soit par la transduction généralisée, soit par la transduction spécialisée, il a récemment été découvert la transduction latérale.

Les phages peuvent être produits soit à partir du cycle lytique, soit à partir du cycle lysogénique présenté en figure 6 (Canchaya et al., 2003). Le choix de l'un de ces cycles dépend des conditions de croissance de l'hôte ainsi que des concentrations de phages dans l'environnement (Pradeep Ram and Sime-Ngando, 2010; St-Pierre and Endy, 2008; Zeng et al., 2010).

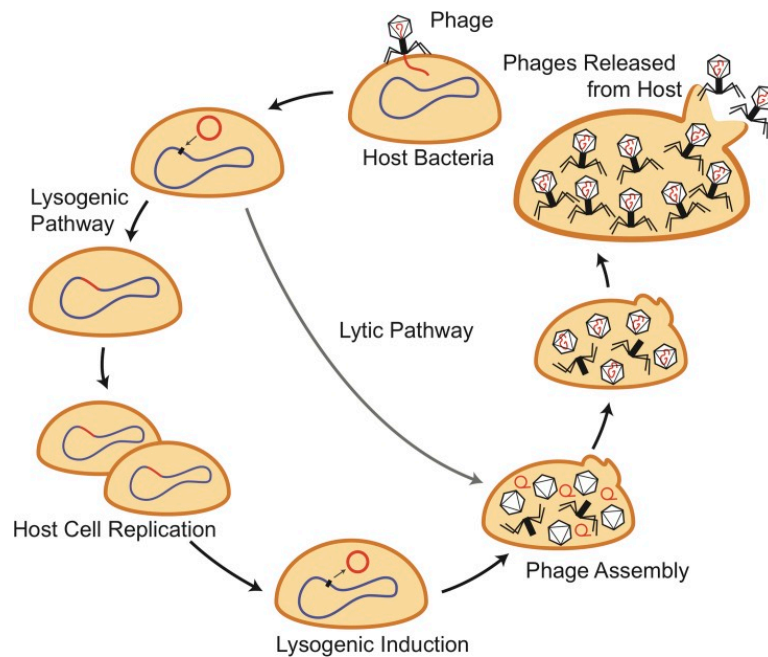


Figure 6 : Cycle lytique et lysogénique des phages

Issue de Chiang et al., 2019

*Le génome viral est représenté en rouge et le génome de l'hôte en bleu. Les phages dits tempérés ont la capacité d'insérer leur génome dans le chromosome de l'hôte. Ainsi, le phage devient un prophage, dès lors nous parlons d'un cycle lysogénique. Cependant, après l'induction lysogénique, les prophages s'excisent et entrent dans le cycle de vie lytique. Ils répliquent leur génome, suivi de l'encapsidation et de la libération des virions dans l'environnement.*

Le cycle lysogénique des phages se caractérise par son intégration temporaire au génome de l'hôte après son infection, nous les appelons alors « prophages » (Brolund et al., 2013; Xue et al., 2015). Pour l'hôte, l'intérêt d'un cycle lysogénique est que des changements phénotypiques sont acquis lors de l'intégration du génome de phage (virulence, mobilité, compétition...) (Bondy-Denomy and Davidson, n.d.; Koskella, 2013; Obeng et al., 2016). Dans ce contexte, le prophage est transmis à la descendance lors de la réplication cellulaire, cette transmission peut avoir des conséquences délétères pour l'hôte (Paul, 2008). En effet, si la viabilité de l'hôte est menacée, le phage peut entrer en phase lytique, se répliquer et entraîner *in fine* le décès de l'hôte. Même dans l'hypothèse où les prophages sont inactivés suite à des mutations, ils demeurent en mesure de lyser la cellule hôte (Asadulghani et al., 2009; Matos et al., 2013).

Le cycle lytique démarre par la réplication du génome du bactériophage qui sera ensuite circularisé en subissant une réplication  $\theta$ , suivie d'une réplication en cercle roulant. Ensuite, l'encapsidation de l'ADN est assurée par l'enzyme terminase du phage qui reconnaît un site spécifique du phage (Rao and Feiss, 2015). Lors de l'encapsidation de l'ADN dans la capsid, lorsque la capsid atteint sa capacité maximale en quantité d'ADN, les terminases de type pac effectuent une seconde coupure non spécifique pour achever cette encapsidation (Streisinger et al., 1967, 1964). Aux termes du clivage terminal, les terminases restent liés à l'ADN et réinitialisent l'encapsidation, c'est ainsi



que de nouvelles procapsides sont remplies au fur et à mesure que les terminases avancent le long de l'ADN viral (Casjens and Hayden, 1988). Enfin, des virions sont lâchés dans l'environnement lors de la lyse de la cellule hôte.

La transduction généralisée a été mise en évidence chez le phage P22 de Salmonella (Zinder and Lederberg, 1952). Lors de cet événement, les phages peuvent encapsider de l'ADN bactérien chromosomique ou plasmidique et le transférer à une autre bactérie (figure 7). Ce phénomène est généré par la reconnaissance de sites dit pseudo-pac, homologues du site pac, par les terminases du phage (Chelala and Margolin, 1976; Schmieger, 1982; Thierauf et al., 2009).

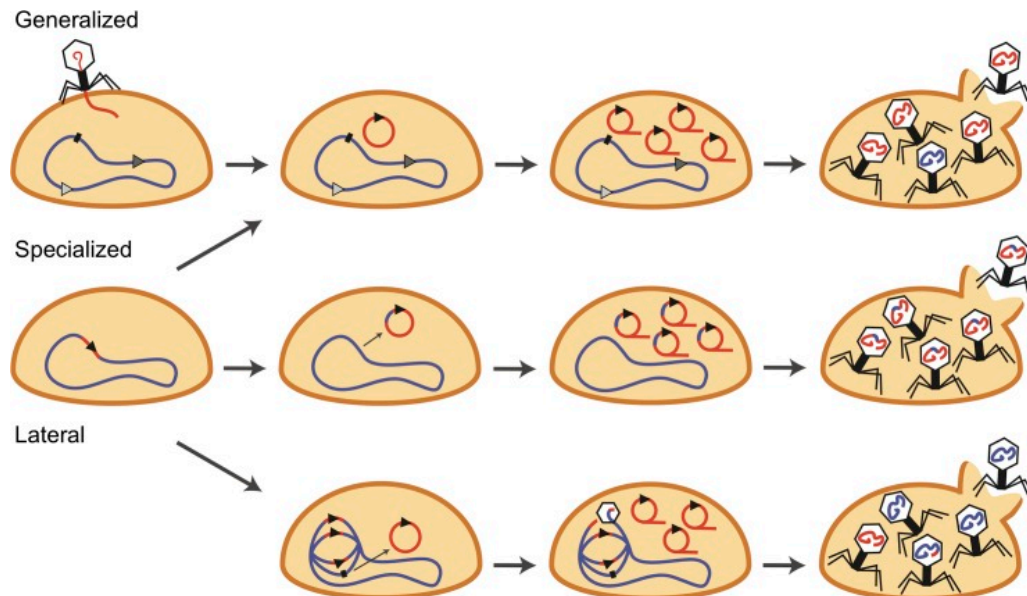


Figure 7 : Présentation des 3 mécanismes de transduction.

Issue de Chiang et al., 2019

*Le génome viral est représenté en rouge, le génome de l'hôte en bleu et les sites pac du phage en noir. Après l'infection par un phage dans la transduction généralisée (en haut) ou après l'induction de prophages dans la transduction spécialisée (au milieu), le génome viral va subir une répllication thêta, suivi d'une répllication en cercle roulant. Ensuite, la terminase du phage initie l'encapsidation de l'ADN à partir des sites pac du phage (triangles noirs) ou des sites pseudo-pac (triangles gris). Dans le cas de la transduction latérale (en bas), la répllication thêta a lieu avant l'excision du prophage, ainsi la terminase exprimée initie l'encapsidation de l'ADN in situ alors que le prophage est encore attaché au chromosome bactérien.*

La transduction spécialisée a été découverte chez le coliphage  $\lambda$  (Morse et al., 1956). Contrairement au mécanisme généralisé, la transduction spécialisée est limitée au transfert d'un ensemble spécifique de gènes. Lors de l'événement d'excision prophagique, une partie de l'ADN viral est joint à un segment adjacent de l'ADN du chromosome bactérien hôte dans le fragment excisé. La molécule d'ADN hybride se réplique, puis est encapsidée et transférée dans de nouvelles cellules hôtes.

Les transductions généralisées et spécialisées sont généralement considérées comme des erreurs commises par les phages entraînant l'encapsulation d'ADN de l'hôte.

Récemment, le mécanisme de transduction latérale a été mis en évidence dans les phages tempérés de *Staphylococcus aureus* (J. Chen et al., 2018). La voie « classique » des prophages est l'excision, la réplication et l'encapsidation. Or, chez les phages tempérés de *Staphylococcus aureus*, l'étape de l'excision du chromosome bactérien est retardée jusqu'à un stade avancé du programme lytique des phages. Ce retard permet à la terminase des phages d'être exprimée et d'initier l'encapsidation de l'ADN *in situ* alors que le prophage est encore sur le chromosome bactérien (Davidson, 2018). Ce mécanisme de retard d'excision semble faire partie intégrante du cycle de vie des phages tempérés de *Staphylococcus aureus*.

Il a été démontré différents liens entre la transduction et la résistance aux antibiotiques, à savoir :

- Le prophage  $\Phi$ 1207.3 de *Streptococcus pyogenes* est porteur des gènes *mef(A)* et *msr(D)* qui confèrent une résistance aux macrolides (Iannelli et al., 2014),
- Le bactériophage de type P1 de *Escherichia coli* est porteur du gène *bla<sub>SHV-2</sub>* qui confère une résistance aux  $\beta$ -lactamines (Billard-Pomares et al., 2014),
- Le phage  $\phi$ C2 de *Clostridioides difficile* assure quant à lui la transduction du Tn6215 codant une résistance à l'érythromycine grâce au gène *erm(B)* (Goh et al., 2013).

#### **1.4. La vésiduction**

Le mécanisme de la vésiduction repose sur la libération de vésicules membranaires (VMs). Les VMs sont sécrétées par les bactéries et participent à divers processus biologiques comme l'exportation de métabolites cellulaires, la communication intercellulaire, la virulence bactérienne, la modulation de la réponse immunitaire de l'hôte et le transfert de matériel génétique (Berleman and Auer, 2013; Bonnington and Kuehn, 2014; Furuyama and Sircili, 2021; Pérez-Cruz et al., 2013).

La présence de matériel génétique dans les VMs a été rapportée chez les bactéries à Gram négatif et à Gram positif, les mycoplasmes et les archées (Brown et al., 2015; Deatherage and Cookson, 2012; Orench-Rivera and Kuehn, 2016). Bien que les fragments présents soient généralement de petite taille, ayant une origine chromosomique, plasmidique et/ou phagique, ils peuvent toutefois atteindre jusqu'à 370 kb (Chiura et al., 2011; Dorward et al., 1989; Klieve et al., 2005; Orench-Rivera and Kuehn, 2016). Le processus de vésiduction est présenté de façon simplifiée en figure 8.

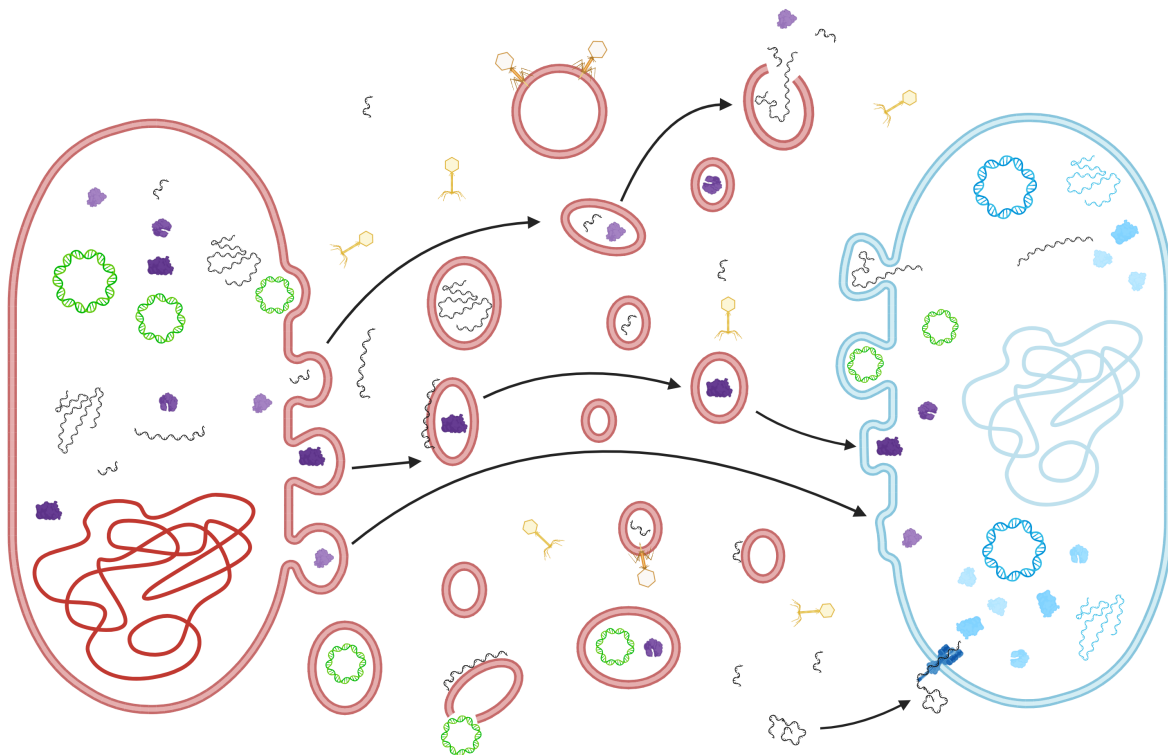


Figure 8 : Représentation du transfert horizontal de gènes par l'intermédiaire d'une vésicule membranaire (VM)

Created with BioRender.com

La cellule donneuse est en rouge (membrane et chromosome), ses protéines en violet, ses plasmides en vert et l'ADNs en noir. Tous les éléments de la cellule receveuse sont en bleu. En jaune sont représentés les bactériophages. Les VM formées à partir de la membrane de la cellule émettrice peuvent contenir du matériel génétique à l'intérieur ou lié à leur surface. Quand ces VM sont lysées, l'ADN présent à l'intérieur est sensible aux nucléases. De plus, des bactériophages peuvent également se lier aux VMs permettant un processus THG de type transduction. Le processus est multidirectionnel. Une fois que la VM atteint une cellule réceptrice, elle fusionne à sa membrane plasmique permettant l'entrée de l'ADN présent à l'intérieur. Ensuite, cet ADN est libéré dans le cytoplasme. Cet ADN sera incorporé au génome de la bactérie réceptrice par recombinaison homologue.

Les VMs peuvent être obtenues par deux voies principales :

- La voie du « blebbing » qui est la libération des VMs par décollement de la membrane.

La voie du « blebbing » se caractérise par le décollement de la membrane qui est le résultat de la déstabilisation de la membrane. Ensuite, il a été démontré que dans la liaison entre le peptidoglycane et la membrane externe chez *Escherichia coli*, un mutant du gène *nlpI* qui participe à la modulation de la dynamique du peptidoglycane induit une augmentation de la formation de VMs sans lyse cellulaire (Schwechheimer et al., 2015). Une autre étude chez *Haemophilus influenzae* a démontré que l'accumulation de phospholipides au niveau du feuillet externe favorisait la libération des VMs (Roier et al., 2016). Les VMs peuvent être également formées à partir soit de protéines incorrectement repliées, soit de l'accumulation de fragments de peptidoglycane dans l'espace périplasmique (McBroom et al., 2006; McBroom and Kuehn, 2007).

- La libération de VMs par l'action de l'endolysine qui permet la lyse cellulaire.

L'action enzymatique des endolysines conduit à la mort cellulaire permettant la formation de VMs. Les endolysines sont utilisées par des phages à ADN double brin afin de lyser l'hôte pour libérer les virions. Il a été mis en évidence chez *Pseudomonas aeruginosa* que, lors d'un stress qui endommage l'ADN, l'expression d'une endolysine conduit à une dégradation du peptidoglycane (Turnbull et al., 2016). Cette dégradation se traduit par l'arrondissement de la cellule, puis son explosion. Certains fragments de la membrane sont capables de s'arrondir et de s'auto-assembler afin de former des VMs. Il est intéressant de noter que les VMs formées par cette lyse cellulaire peuvent contenir des endolysines capables de lyser d'autres cellules (Li et al., 1998). En plus des endolysines issues des phages, les bactéries sont également capables de produire des autolysines bactériennes qui peuvent être impliquées dans la formation de VMs (Hayashi et al., 2002; Wang et al., 2018).

Les VMs peuvent contenir du matériel génétique, soit interne, soit externe à la surface des VMs (Fulsundar et al., 2014; Renelli et al., 2004). Dans l'environnement, les VMs peuvent être la cible de phage et subir une transduction phagique permettant de véhiculer l'ADN de phage (Biller et al., 2014).

Bien que les mécanismes de délivrance des toxines aux cellules de mammifères par les VMs soient bien connus et caractérisés (Kesty et al., 2004; O'Donoghue and Krachler, 2016), la spécificité du mécanisme de délivrance du matériel génétique par les VMs à la surface de la cellule hôte reste relativement obscure. En effet, selon l'espèce receveuse différentes hypothèses sont évoquées. À titre d'illustration : les VMs peuvent être lysées, l'ADN est alors acquis par la cellule receveuse *via* la transformation ou les VMs peuvent s'attacher à la membrane externe puis être internalisées par la cellule réceptrice (Domingues and Nielsen, 2017; Fulsundar et al., 2014).

Il a aussi été montré que des VMs permettaient d'introduire du matériel génétique dans le noyau de cellules eucaryotes (Bitto et al., 2017).

Enfin, de nombreuses études ont mis en avant que les VMs permettaient le transfert de gènes de résistance aux  $\beta$ -lactamines chez des espèces d'intérêt clinique (*Acinetobacter baumannii*, *Pseudomonas aeruginosa* et *Escherichia coli*...), ainsi que des gènes de virulence chez *Pseudomonas aeruginosa*, *Vibrio cholerae*, *Helicobacter pylori* et *Escherichia coli*...(Caruana and Walper, 2020)

Très concrètement, il ressort de ce qui précède que :

- La transformation naturelle se caractérise par un processus régulé d'acquisition d'ADN libre dans l'environnement,
- La conjugaison est un processus nécessitant un contact de cellule à cellule,
- La transduction fait intervenir des bactériophages dont le rôle est de transférer l'ADN entre bactéries par le biais d'une infection,
- La vésiduction implique la formation de VM pour permettre le transfert d'ADN.

Le transfert de gènes implique principalement des gènes présents sur des éléments génétiques mobiles (EGM).

## Chapitre II. Les éléments génétiques mobiles impliqués dans la dissémination des gènes de résistances aux antibiotiques

Afin de mieux comprendre ce chapitre, nous devons d'abord introduire la notion de mobilome. Le mobilome est défini par l'ensemble des éléments génétiques mobiles (EGM) tels que les plasmides, les transposons et les séquences d'insertion (figure 9) (Anani et al., 2020). La circulation de gènes via ces EGM a lieu à deux niveaux, ce qui entraîne un impact important sur l'évolution bactérienne : soit de manière intra-génomique avec les transposons et les séquences d'insertion, soit de manière inter-génomique avec les plasmides, les bactériophages et les transposons/ICE. Les médiateurs de transferts intra-génomiques sont spécifiques (Gibson et al., 2015; Hu et al., 2016), contrairement aux médiateurs de transferts inter-génomiques qui peuvent être impliqués dans une large gamme d'hôtes de transfert (Garcillán-Barcia et al., 2011; Zrimec and Lapanje, 2018). Les EGM permettent l'augmentation de la diversité génétique, l'adaptation par l'acquisition ou la perte de gènes, ainsi que l'adaptation à des environnements spécifiques (Ghaly and Gillings, 2018; Lagier et al., 2015). Les EGM peuvent être à l'origine de la propagation de nombreuses résistances aux antibiotiques chez les espèces bactériennes (Malachowa and DeLeo, 2010).

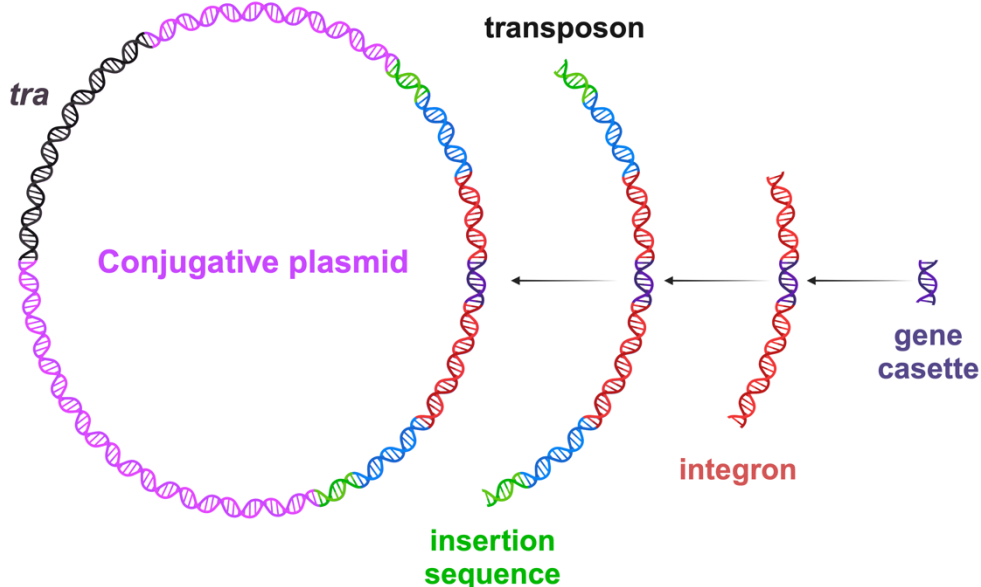


Figure 9 : Diversité des éléments génétiques mobiles

Created with BioRender.com

Les cassettes de gènes (violette) s'insèrent dans les intégrons (rouge) grâce à des sites spécifiques. Ensuite, les intégrons ont la possibilité d'être flanqués par des séquences d'insertions (vert) pour former des transposons. À leur tour, les transposons peuvent être insérés dans un plasmide conjugatif (rose) qui est transmissible. Ainsi, ces différents éléments s'imbriquent les uns dans les autres telles des « poupées russes ». Les gènes tra (noir) permettent une mobilité inter-génomique.

## II.1. Les plasmides transmissibles

Les plasmides sont des séquences extra-chromosomiques que nous trouvons chez de nombreux microorganismes (bactéries, archées et eucaryotes) (Funnell and Phillips, 2004). La taille des plasmides varie de 5 à 500 kb, bien que des plasmides peuvent avoir une taille de 2 kb à plus de 1 Mb (Biet et al., 2002; Harrison et al., 2010; Rozhon et al., 2010). Les plasmides transmissibles sont transmis par conjugaison (Frost et al., 2005; Sota and Top, 2008). En 2010, il a été rapporté qu'environ 14% des plasmides entièrement séquencés sont supposés conjugatifs (Smillie et al., 2010). Récemment, il a été publié que sur 10 000 plasmides, 42% d'entre eux étaient putativement transmissibles (Garcillán-Barcia et al., 2023). Les plasmides peuvent coder pour des gènes de fonctions spécifiques tels que l'autoréplication, le partitionnement et le transfert conjugatif (Bennett, 2008). La conjugaison fait partie des mécanismes les plus efficaces pour la transmission des éléments génétiques entre les bactéries (Aminov, 2011; Guglielmini et al., 2011). Ces plasmides codent principalement des gènes qui permettant une meilleure adaptation de l'hôte à un nouvel environnement, comme l'acquisition de nouvelles sources de nutriments ou une résistance aux antibiotiques (Rankin et al., 2011; Slater et al., 2008). Les plasmides porteurs de gènes de résistance aux antibiotiques les plus fréquemment retrouvés permettent une résistance aux aminosides,  $\beta$ -lactamines et les tétracyclines (Bennett, 2008; Pal et al., 2015). Enfin, les plasmides empêchent la cellule de les tuer après une ségrégation grâce aux systèmes toxine-antitoxine (TA) qui assurent ainsi le maintien du plasmide dans la cellule bactérienne (Hayes, 2003).

Nous distinguons deux types de plasmides transmissibles (figure 10) (Smillie et al., 2010) :

- Les plasmides dit « conjugatifs » codent les quatre composantes nécessaires à la conjugaison, à savoir : une origine de transfert (*oriT*), et un groupe de gènes pour la mobilité (MOB) qui codent une relaxase, une protéine de couplage de type IV (T4CP) et un système de sécrétion de type IV (T4SS).
- Les plasmides dit « mobilisables » contiennent *a minima* l'*oriT*, et la relaxase. Dans certains cas, ils peuvent contenir une protéine de couplage de type IV. Cependant, la transmissibilité du plasmide est assurée par un plasmide conjugatif corésident pour exprimer un système de sécrétion de type IV (T4SS).

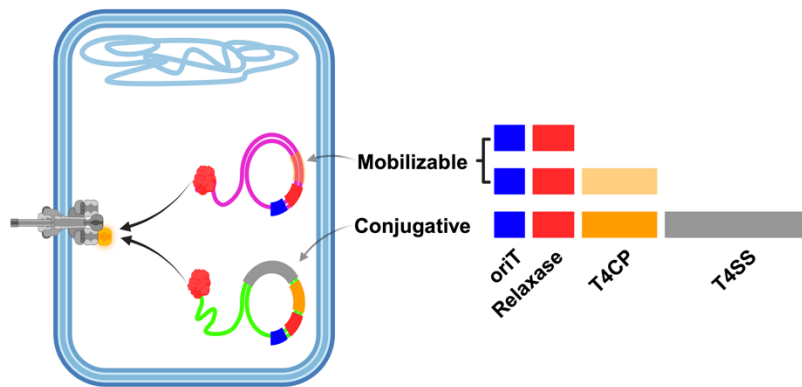


Figure 10 : Schéma de la constitution génétique des plasmides dits transmissibles

Created with BioRender.com

Représentation schématique de la présence ou de l'absence des quatre modules nécessaires à la conjugaison : origine de transfert (*oriT*) (bleu), une relaxase (rouge), une protéine de couplage de type IV (*T4CP*) (orange) et un système de sécrétion de type IV (*T4SS*) (gris). Le plasmide mobilisable et conjugatif sont représentés en rose et vert, respectivement. Le système de *T4SS* assemblé est en gris au niveau de la cellule donneuse.

La relaxase et l'*oriT* sont communes à tous les plasmides transmissibles. La relaxase se lie à l'*oriT* pour former le relaxosome (Fürste et al., 1989). L'*oriT* est une courte séquence d'ADN qui est requise en *cis* pour que le plasmide soit transmissible *via* la conjugaison. La relaxase a pour rôle de catalyser les étapes initiales et finales de la conjugaison (Byrd and Matson, 1997). L'étape initiale commence par le clivage simple brin de l'*oriT*, cette réaction permet de produire le brin d'ADN qui sera transféré chez la bactérie donneuse. Ensuite, l'étape finale est la ligature de l'ADN transporté chez la bactérie receveuse permettant la reconstitution du plasmide circulaire double brin qui a été transféré (de la Cruz et al., 2010).

La *T4CP* est impliquée dans l'interaction entre le relaxosome et le canal du *T4SS* (Llosa et al., 2003; Mihajlovic et al., 2009). De plus, *T4CP* aurait un rôle dynamique dans le transport de l'ADN à travers le *T4SS* (Llosa et al., 2002; Tato et al., 2007). Le canal conjugatif permet principalement la sécrétion de protéines, en l'espèce, la relaxase liée à l'ADN permet son transfert vers la cellule receveuse (Garcillán-Barcia et al., 2007).

## II.2. Classification des plasmides

Un effort important de consensus a été fait afin de permettre le classement des plasmides. A ce titre, c'est en 1976 que Novick introduit les groupes d'incompatibilité comme base de la classification des plasmides, l'incompatibilité plasmidique étant nommée Inc (Novick et al., 1976). Il convient de préciser qu'un groupe d'incompatibilité se définit comme : l'ensemble des plasmides, partageant un ou plusieurs éléments des systèmes de répllication ou de partition des plasmides, qui empêche leurs coexistences stables dans la même cellule en absence de pression de sélection (Novick, 1987).



Par la suite, un autre système de classification fondé sur les relations phylogénétiques entre les relaxases des plasmides transmissibles par conjugaison bactérienne a été mis en place (Garcillán-Barcia et al., 2011, 2009; Smillie et al., 2010). Aujourd'hui, nous comptons neuf classes de gènes MOB qui englobent la diversité phylogénétique des relaxases des plasmides transmissibles. En effet, certaines relaxases ont une préférence pour les plasmides d'une taille donnée, dès lors il apparaît une spécialisation dans les plasmides mobilisables ou dans les plasmides conjugatifs (Coluzzi et al., 2022; Garcillán-Barcia et al., 2009).

Récemment, une nouvelle nomenclature de classification fondée sur la définition d'unités taxonomiques des plasmides (PTU) a été mise en avant. Le PTU se base sur des métriques d'identité nucléotidique moyenne permettant la génération d'un schéma universel de classification des plasmides. Ce système est applicable à tous les taxons bactériens (Garcillán-Barcia et al., 2023). Ainsi, la classification basée sur les PTUs se traduit comme un système de classification naturel pour les plasmides *via* la formation de groupes génomiques cohérents, cela est semblable aux espèces moléculaires (Garcillán-Barcia et al., 2023). *In fine*, cette classification a permis l'introduction d'une échelle de gamme d'hôtes de six grades. Le grade I regroupe des plasmides présents strictement dans des isolats de la même espèce bactérienne alors que le grade VI rassemble des plasmides capables de coloniser différentes classes d'un phylum ou encore de franchir la barrière du phylum (Redondo-Salvo et al., 2020).

Il a été constaté qu'une grande majorité des PTUs des plasmides mobilisables et conjugatifs, soit 81%, étaient répartis dans le grade III. Ce grade se compose des plasmides capables de circuler dans des bactéries de genres différents mais de la même famille. *A contrario*, 84% des PTUs des plasmides non mobilisables sont restreints aux grades I ou II, ces grades se composent essentiellement des plasmides présents dans la même espèce ou le même genre, respectivement (Garcillán-Barcia et al., 2023).

### **II.3. Les éléments intégratifs conjugatifs (ICE)**

Les ICE sont caractérisés par deux facteurs déterminants (Alvarez-Martinez and Christie, 2009; Christie et al., 2014; Trokter et al., 2014) :

- Ils sont intégrés dans le génome de l'hôte,
- Ils codent un système de conjugaison opérationnel et fonctionnel ainsi qu'un système de sécrétion de type 4 assurant le transfert vers une cellule receveuse.

Initialement, les ICE étaient nommés « transposons conjugatifs » car il avait été démontré qu'ils étaient capables à la fois de transposer intra-cellulairement et de conjuguer inter-cellulairement (Clewell et al., 1995). Ainsi ces éléments sont transmis, soit verticalement lorsqu'ils sont intégrés au génome hôte, soit horizontalement après excision et transfert à la cellule receveuse par conjugaison (Delavat et al., 2017). Ce n'est que plus tard qu'ils ont été qualifiés d'éléments conjugatifs intégratifs

(ICE), bien qu'il ait été décidé que les termes « transposon conjugatif » et « ICE » étaient interchangeables (Burrus et al., 2002a; Roberts et al., 2008).

Les ICEs sont porteurs de divers gènes auxiliaires qui confèrent à leurs hôtes un éventail de fonctions supplémentaires (capacités métaboliques, colonisation, résistance aux antibiotiques ou aux métaux) (Johnson and Grossman, 2015). Les ICE ont une taille allant d'environ 20 kb (Tn916) à plus de 500 kb (ICEM/SymR7A) et sont généralement en mosaïques et modulaires (Osborn and Böltner, 2002; Ramsay et al., 2006; Roberts et al., 2001; Roberts and Mullany, 2009).

Le contexte génétique d'un ICE et le cycle de vie d'un ICE sont présentés en figure 11 et en figure 12, respectivement.

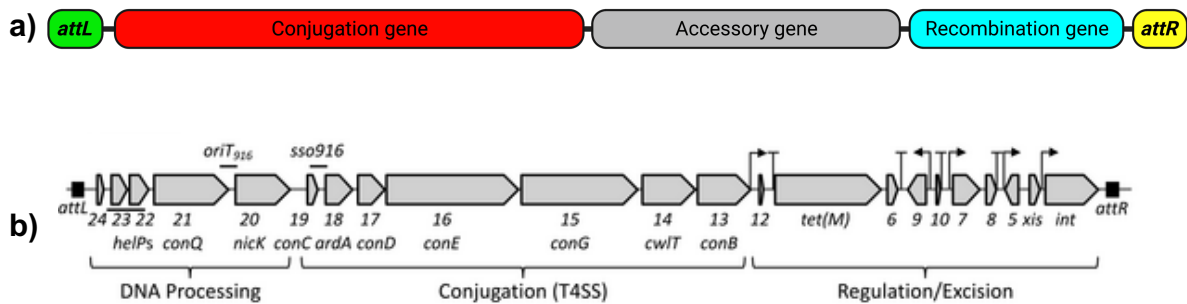


Figure 11 : Contexte génétique d'un ICE

Created with BioRender.com et issue de (Bean et al., 2022)

a) Représentation schématisée simplifiée des différentes régions d'un ICE, les sites *attL* (vert) et *attR* (jaune) permettent l'intégration ou l'excision de l'ICE du chromosome. Les gènes exprimant l'intégrase et les facteurs de régulation de l'excision et de l'intégration sont représentés en bleu. Les gènes exprimant le système de conjugaison sont représentés en rouge et les gènes accessoires codant des gènes de résistance ou de régulation sont représentés en gris.

b) Présentation du contexte génétique du Tn916.

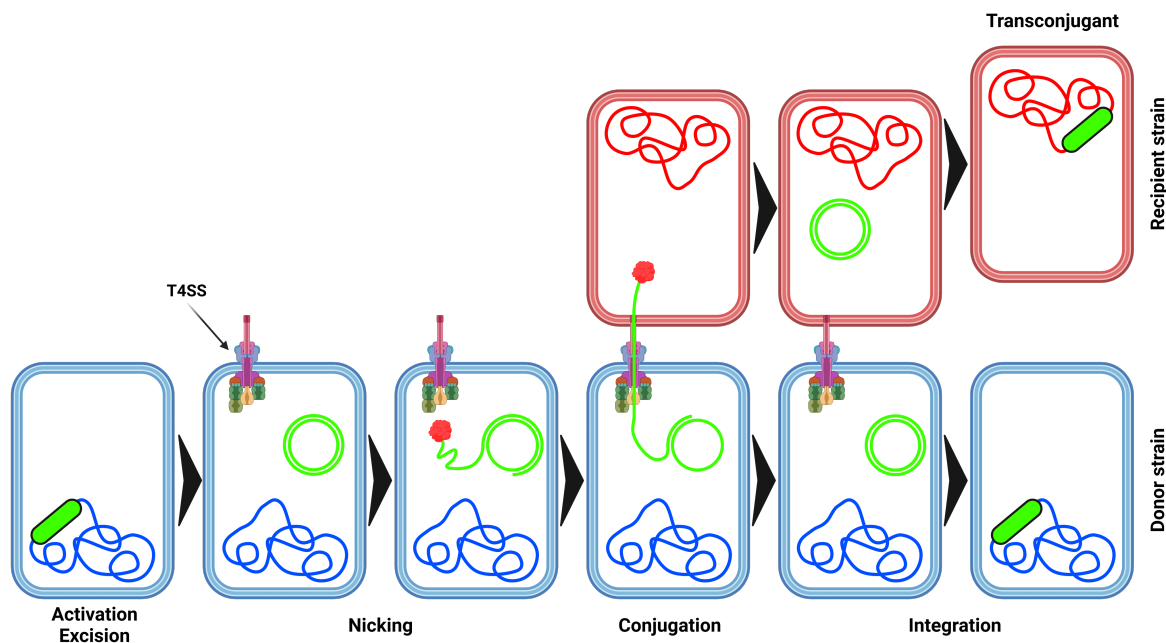


Figure 12 : Cycle de vie d'un élément intégratif conjugatif (ICE)

Created with BioRender.com

La cellule donneuse et son chromosome sont représentés en bleu. La cellule receveuse et son chromosome sont en rouge. La forme intégrée de l'ICE est représentée par un rectangle vert alors que la forme excisée est représentée par des cercles verts. La boule rouge représente la relaxase. Voir le texte pour la description.

Les ICEs s'intègrent et s'excisent du génome bactérien grâce à une intégrase qui assure la recombinaison spécifique (Grindley et al., 2006). En temps normal, les ICEs sont présents sur le chromosome de la bactérie hôte et les gènes de conjugaison ne sont pas exprimés (Wozniak and Waldor, 2010). Ces gènes sont activés soit en réponse à des conditions cellulaires (phase de croissance) ou à des stimuli extra-cellulaires (tétracycline), soit de manière stochastique (Delavat et al., 2017). Parmi ces gènes, le gène *int* code une intégrase qui catalyse l'excision de l'ICE permettant la formation d'un intermédiaire circulaire qui possède les caractéristiques d'un plasmide conjugatif (Johnson and Grossman, 2015). La plupart des ICEs utilisent une intégrase à tyrosine, à l'exception de certains cas dans lesquels ils utilisent une intégrase à sérine ou à DDE (Brochet et al., 2009; Burrus et al., 2002a; Burrus and Waldor, 2004). De plus, le gène *xis* code le facteur excisionase dont la fonction est de réguler l'activité de l'intégrase vers l'excision, en favorisant un appariement entre les sites *attL* et *attR* et en inhibant l'appariement entre les sites *attB* et *attI* (figure 13) (Fogg et al., 2014; Groth and Calos, 2004; Lee et al., 2007).

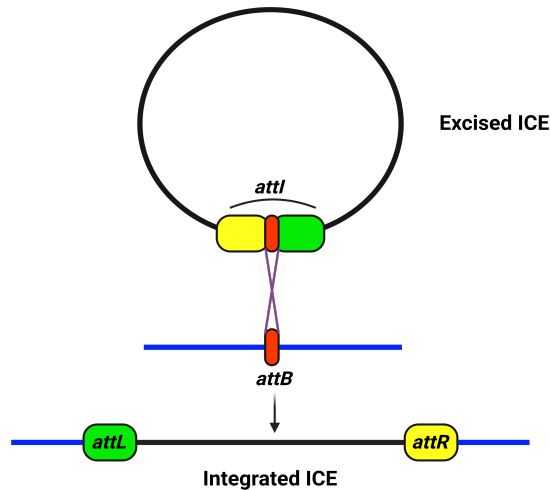


Figure 13 : Représentation schématique lors de l'intégration au site spécifique d'un ICE

Created with BioRender.com

Lors de l'intégration dans le chromosome hôte (trait bleu) au site spécifique *attB* (rouge), l'intégrase de l'ICE catalyse une recombinaison site spécifique entre les régions identiques du site *attI* de l'ICE (trait noir) et du site *attB* (rouge). Cela permet la co-intégration de l'ICE dans le chromosome de l'hôte et la formation des sites *attL* (vert) et *attR* (jaune).

Le transfert de l'ICE vers la cellule receveuse est dépendant de nombreux gènes dont une relaxase, qui est une endonucléase de la superfamille HUH (Burrus et al., 2002b; Chandler et al., 2013). De plus, le T4SS de conjugaison codée par les ICEs est homologue à celui codé par les plasmides conjugatifs (Alvarez-Martinez and Christie, 2009; Kado, 2014). La relaxase se lie à l'*oriT* de l'intermédiaire circulaire pour former le relaxosome. Cela permet le transfert de l'ICE d'une bactérie donneuse à une bactérie receveuse via le principe décrit dans le *chapitre 2 sous partie II.1.2*. Dès lors que l'intermédiaire circulaire est transféré dans la cellule receveuse, la relaxase catalyse la ligature des extrémités 5' et 3' de l'ADN simple brin transféré permettant ainsi la formation d'un ADN simple brin fermé (Johnson and Grossman, 2015). Cet ADN est converti en ADN double brin par la synthèse du second brin. Ensuite, l'intégrase de l'ICE catalyse l'intégration de l'élément dans le chromosome de la receveuse. Chez le donneur, une fois l'ADN double brin de l'ICE reformé, il s'intègre à nouveau dans le chromosome de la cellule donneuse (Wozniak and Waldor, 2010). Bien que de nombreux ICEs ciblent un seul site *attB* dans le chromosome similaire au site *attP* de l'ICE, souvent localisé dans un gène d'ARNt, quelques-uns n'ont pas de spécificité de site comme le Tn916 (Johnson and Grossman, 2015; Roberts and Mullany, 2009).

Parmi les exemples d'ICE, on peut citer des ICE codant des fonctions métaboliques comme le Tn5276 de *Lactococcus lactis* et l'ICE CTnScr94 chez *Salmonella* permettent la fermentation du saccharose, ou des fonctions de virulence avec l'îlot de pathogénicité PAPI-1 de *Pseudomonas aeruginosa* (Carter et al., 2010; Hochhut et al., 1997; Rauch and De Vos, 1992).

De nombreux ICEs contiennent aussi d'autres éléments génétiques tels que des transposons composites (Juhas et al., 2007; Toleman and Walsh, 2011), comme l'ICE*HpaT3T1* chez *Haemophilus parainfluenzae* avec les gènes de résistance à la tétracycline encadrés par deux copies d'IS10 formant le Tn10, ou un intégron avec des gènes de résistance aux antibiotiques comme l'ICE SXT ou l'ICEEc2 (Hochhut et al., 2001; Roche et al., 2010).

## II.4. Les éléments transposables procaryotes

Ce sont des séquences d'ADN capables de se déplacer d'un endroit à un autre, d'un site donneur vers un site cible du même génome ou sur un autre réplicon. Contrairement à la recombinaison homologue, la transposition ne requiert pas d'identité de séquence entre l'élément mobile et l'ADN, elle utilise une enzyme particulière : la transposase.

### II.4.1. Les séquences d'insertions (IS)

Les séquences d'insertion se déterminent comme les éléments transposables les plus simples avec une taille comprise entre 0,7 et 2,5 kb. Jusqu'à récemment, on définissait les séquences d'insertion comme des segments d'ADN contenant uniquement les informations génétiques nécessaires et suffisantes à leur transposition (Berg et al., 1989; Craig et al., 2002).

Le gène *tnpA* code une transposase (TnpA) qui est une enzyme catalytique. La fonction de TnpA est de reconnaître des séquences répétées inversées (IRL, left et IRR, right) et de catalyser la réaction de transposition de l'élément (figure 14). Les séquences IR situées aux extrémités sont indispensables à la mobilité des IS. La taille de ces séquences IR varie d'environ 20-40 pb (Huisman et al., 1989; Johnson and Reznikoff, 1983; Makris et al., 1988; Mugnier et al., 2009). Les IS seront développées dans le *chapitre III*.

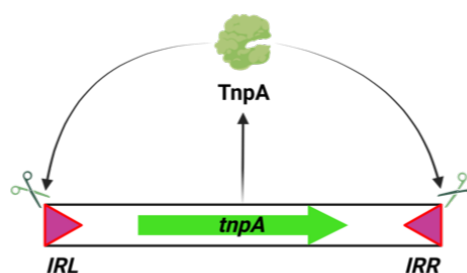


Figure 14 : Schéma d'une séquence d'insertion

Created with BioRender.com

Chez la majorité des IS, le gène *tnpA* (flèche verte indiquant le sens de transcription) est flanqué de régions appelées répétitions inversées terminales (IR -L (left) / IR -R (right), triangle violet). Le gène *tnpA* permet de coder pour une transposase (TnpA). TnpA va reconnaître les extrémités IRL et IRR et les cliver pour permettre la mobilité de l'élément.

## II.4.2. Les transposons (Tn)

Les Tn sont plus grands que les IS et contiennent un ou des gènes supplémentaires dont les fonctions ne sont pas impliquées dans le processus de transposition. Ces gènes sont considérés comme facultatifs et ils sont aussi appelés fonctions « auxiliaires » ou gènes « passagers ». Les Tn sont soit non composites, soit composites (figure 15) (Malachowa and DeLeo, 2010). De plus, certains transposons contiennent des promoteurs qui peuvent entraîner la transcription de gènes en aval, cela peut générer l'activation et l'expression de divers phénotypes auparavant silencieux (Szuplewska and Bartosik, 2009). Les Tn ont largement participé à la propagation des multirésistances bactériennes aux antibiotiques par le fait qu'ils ont la possibilité d'être portés par des éléments transférables (Iyer et al., 2013).

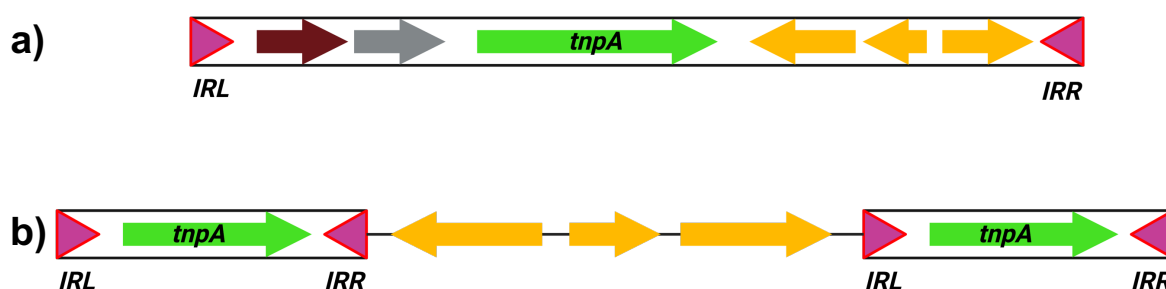


Figure 15 : Transposons non-composite et composite

Created with BioRender.com

Les flèches représentent l'orientation des gènes, en vert le gène *tnpA*, en jaune les gènes de résistance aux antibiotiques, en marron les gènes de résistance aux métaux lourds et en gris les gènes de fonction inconnues. Les extrémités de l'IS sont représentées par des triangles violets indiquant les répétitions inversées terminales (IR -L (left) / IR -R (right)).

a) Représentation schématique d'un transposon non-composite où plusieurs gènes auxiliaires/facultatifs sont représentés entre l'IRL et l'IRR faisant partie intégrante de l'IS.

b) Représentation schématique d'un transposon composite où deux copies d'IS encadrent des gènes de résistance aux antibiotiques.

### II.4.2.1. Les transposons non-composites

Les transposons non composites, appelés aussi transposons unitaires, sont bordés de séquences répétées inversées terminales, comme les IS. Néanmoins, ils possèdent leur propre gène de transposase, permettant leur transposition en intra-cellulaire. Ces transposons hébergent un gène *tnpA* qui code la transposase et un gène *tnpR* qui code la résolvasse.

Ces transposons ont des tailles variables entre 4 kb et 20 kb et ils portent généralement un ou plusieurs gènes facultatifs, comme des gènes de résistance aux antibiotiques ou aux métaux lourds.

Le Tn5563a appartient à la famille du transposon Tn3 qui est le prototype de ce groupe de Tns unitaires. En effet, les transposons de la famille Tn3 sont principalement impliqués dans le transfert des gènes de résistance *bla* induisant une résistance aux  $\beta$ -lactamines (Partridge, 2011; Partridge

and Hall, 2005). Nous pouvons citer l'exemple de Tn1331 qui confère une résistance aux carbapénèmes chez *Klebsiella pneumoniae* (Rinkel et al., 1994).

Différents transposons non-composites portent des gènes de résistance aux antibiotiques (Partridge et al., 2018).

#### **II.4.2.2. Les transposons composites**

Les Tns composites contiennent un gène ou un groupe de gènes flanqués de deux IS identiques, appelées IS-L (left) et IS-R (right) (Schaefer and Kahn, 1998). Ainsi, les deux IS sont suffisamment proches l'une de l'autre pour pouvoir interagir de concert dans la mobilisation du segment d'ADN qu'elles encadrent.

La transposition d'un Tn composite est catalysée par la transposase codée par l'une ou les deux IS. Ces IS peuvent être en orientation directe, c'est-à-dire dans la même orientation, ou inverse. En effet, seule l'une des deux IS du transposon composite code pour une transposase fonctionnelle, l'autre IS code soit une transposase mutée ayant perdu sa fonction, soit éventuellement un régulateur de la transposition.

Le segment d'ADN encadré par les IS n'intervient pas dans le mécanisme de transposition et peut coder pour n'importe quelle fonction, cela est le cas des gènes de résistance aux antibiotiques ou des gènes de dégradation des xénobiotiques (Bennett, 2008; Nojiri et al., 2004; Schaefer and Kahn, 1998). Un même couple d'IS peut border des segments d'ADN d'origine différente formant alors différents transposons composites. L'existence de tels Tn composites indique que de nouveaux Tn continuent perpétuellement de se produire dans les cellules hôtes ayant des IS dans leur génome.

Différents Tn sont porteurs de gènes de résistance aux antibiotiques (Partridge et al., 2018). Par exemple, Tn5 est encadré par deux copies d'IS50, entourant trois gènes codant la résistance à la kanamycine, la bléomycine et la streptomycine, respectivement (Genilloud et al., 1988; Goryshin et al., 1998; Sasakawa et al., 1982).

#### **II.4.3. Les MITEs, les MICs et les PATEs**

Les dérivés des éléments transposables non autonomes sont les MITEs et les MICs. De nombreux génomes procaryotes sont truffés de fragments d'IS et de petits dérivés non-autonomes. En principe, leur transposition peut être catalysée par la transposase de l'élément complet correspondant si elle est exprimée en *trans* dans la même cellule (Delilhas, 2011).

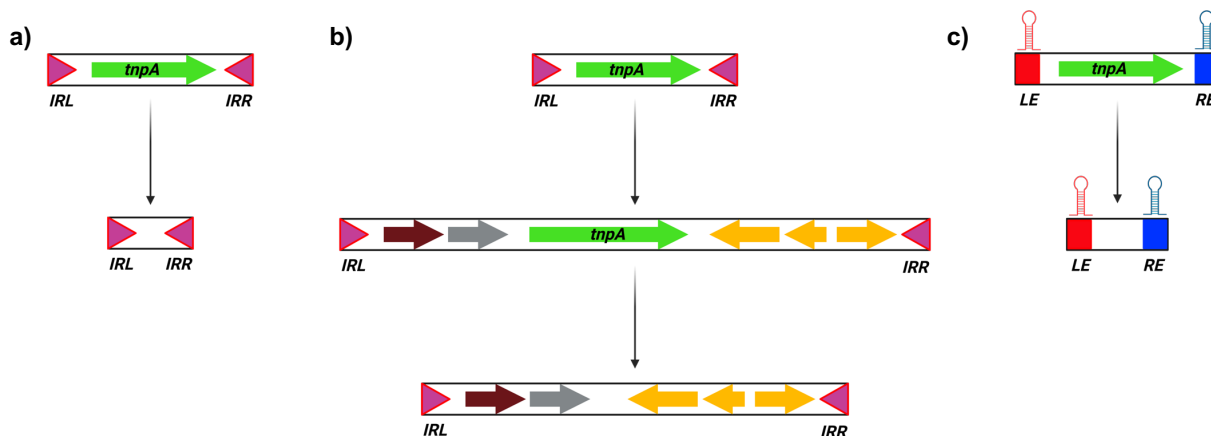


Figure 16 : Origine des MITEs, MICs et PATEs

Created with BioRender.com

Les flèches représentent l'orientation des gènes, en vert le gène *tnpA*, en jaune les gènes de résistance aux antibiotiques, en marron les gènes de résistance aux métaux lourds et en gris les gènes de fonction inconnue. Les extrémités de l'IS sont représentées par des triangles violets indiquant les répétitions inversées terminales (IR -L (left) / IR -R (right)) et les palindromes subterminaux sont représentés par des rectangles rouge (LE – Left End) et bleu (RE – Right End).

a) En haut de la figure se trouve le schéma classique d'une IS et en bas, le schéma d'un MITE qui est probablement dérivé de la délétion interne de la transposase.

b) Après le passage d'une IS vers une forme de transposon non-composite où plusieurs gènes auxiliaires/facultatifs sont représentés entre l'IRL et IRR, le MIC serait probablement dérivé de la délétion interne de la transposase.

c) En effet, comme pour les MITEs, la génération de PATE serait probablement dérivée de la délétion interne de la transposase de l'élément.

Les structures appelées MITEs (Miniature Inverted Repeat Transposable Elements) ont d'abord été identifiées chez *Neisseria spp.* et sont apparentées aux IS avec une transposase DDE (figure 16.a) (Correia et al., 1986, 1988; J. Filée et al., 2007). Elles incluent des représentants des membres des familles IS1, IS4, IS5, IS6 et Tn3 (Brügger et al., 2002; J. Filée et al., 2007). Les MITEs sont de petite taille (~300 pb). Elles incluent des IRs terminales gauche et droite, correctement orientées. Toutefois, à la différence des IS, aucun gène codant une transposase ne se trouve entre les deux IRs. Par ailleurs, ces MITEs génèrent généralement des courtes répétitions directes flanquantes (généralement le dinucléotide TA) (Craig et al., 2002; Delilhas, 2008). Les MITEs sont probablement dérivées de la délétion interne de l'IS ; certaines contiennent de courtes séquences non codantes entre les IRs qui pourraient ou non être dérivées de l'IS (Delilhas, 2011). Les MITEs sont considérées comme des éléments transposables non-autonomes et mobilisables *in trans* par les transposases des copies génomiques parentales complètes (Bardaji et al., 2011). Les MITEs sont généralement co-transcrites avec les gènes situés en amont (De Gregorio et al., 2006). De plus, dans certains cas, elles contiennent des promoteurs fonctionnels dans leur région centrale qui sont reconnus par les facteurs de l'hôte modulant *in fine* l'expression des gènes voisins (Delilhas, 2008).



Les MICs pour « Mobile Insertion Cassette » (figure 16.b) est un autre groupe d'éléments dérivés d'IS et apparentés aux MITEs (Chen et al., 1999; De Palmenaer et al., 2004). Ces éléments, comme les MITEs, sont flanqués d'IR, ne codent pas de transposase et génèrent des répétitions directes (DR) dans le site cible. Cependant, à l'inverse des MITEs, ces éléments transportent des gènes variés et présentent un nombre de copies relativement faible (De Palmenaer et al., 2004). Le sous-groupe IS231 appartenant à la famille IS4 comprend de nombreux exemples de ces différents types de dérivés : MITEs et MICs (De Palmenaer et al., 2008).

Enfin, des structures similaires aux MITEs ont été identifiées et incluent les dérivés de la famille IS200/IS605 (J. Filée et al., 2007). Ces structures s'apparentent à des éléments avec d'autres types de transposase (différentes de DDE). Ces structures ont été nommées PATEs pour « Palindrome-associated Transposable Elements », ce sont des structures qui contiennent aux extrémités uniquement les structures secondaires sub-terminales LE (Left End) et RE (Right End), ainsi que la séquence cible nécessaire au clivage (présenté en figure 16.c) (Dyall-Smith et al., 2011). Ces structures ont été détectées dans les génomes d'Archées, de certaines cyanobactéries et Salmonelles (Siquier et al., 2014; Teo et al., 2022).

## II.5. Intégrons

Ce sont Hatch Stokes et Ruth Hall qui ont décrit pour la première fois les intégrons (Stokes and Hall, 1989). Il s'agit d'un système qui permet de capturer et d'exprimer efficacement des gènes exogènes. Nous distinguons deux grandes familles d'intégrons :

- Les intégrons sédentaires chromosomiques : ils sont strictement retrouvés sur le chromosome bactérien. Ces intégrons jouent un rôle important en tant que réservoir de gènes car ils sont capables de porter des centaines de cassettes de gènes. La plupart des protéines codées par ces cassettes gènes ont une fonction inconnue. Ainsi, ces réseaux de cassettes constituent un point chaud pour la diversité du génome bactérien (Hall, 2012).
- Les intégrons de résistance ou intégrons mobiles.

Les intégrons de résistance permettent de capturer et d'exprimer efficacement des gènes de résistance aux antibiotiques et aux antiseptiques (Stokes and Hall, 1989). Ces intégrons sont naturellement immobiles, toutefois ils sont liés à des éléments génétiques mobiles tels que des ICE ou des plasmides conjugatifs permettant leur mobilité (Stokes et al., 2001). Un intégron de résistance se compose d'une région conservée en 5' (5'CS) et d'une région variable.

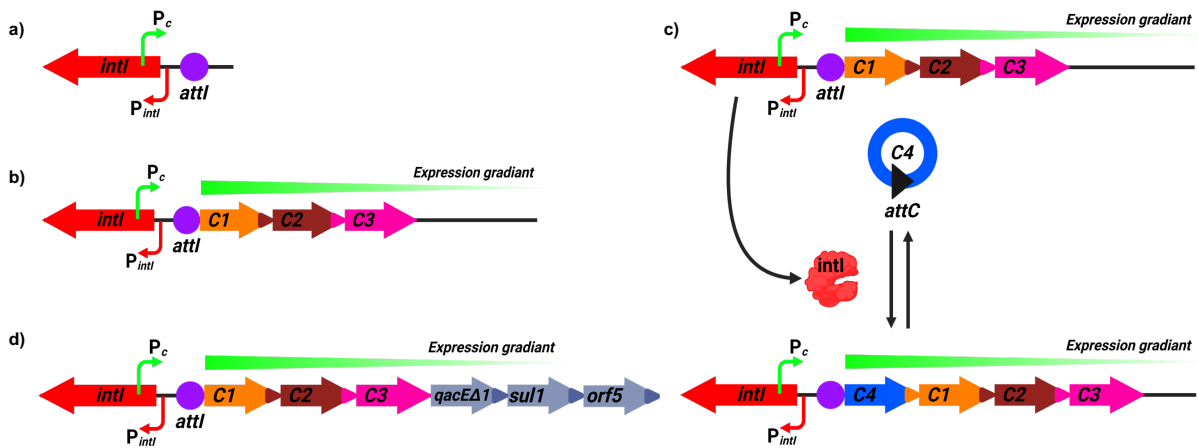


Figure 17 : Les différentes régions d'un intégron et son fonctionnement

Created with BioRender.com

- a) La région 5'CS des intégrons se caractérise par la présence d'un gène *intI* codant pour l'intégrase *IntI* (flèche rouge) qui est exprimé à partir du promoteur *PintI* (petite flèche rouge). Dans la séquence codante de *intI*, un promoteur nommé *Pc* (flèche verte) permet l'expression du réseau de cassette. Le site *attI* (point violet) permet l'intégration des cassettes génétiques dans la région variable.
- b) Après le site *attI* de la région 5'-CS, nous avons un enchainement de cassettes exprimant une diversité de gènes appelée région variable. En effet, le promoteur *Pc* va permettre l'expression des cassettes de gènes. À ce titre, plus la cassette est proche de la région 5'-CS, plus les gènes seront exprimés. A contrario, plus la cassette est éloignée de la région 5'-CS, moins elle a de chance d'être exprimée.
- c) Le gène *intI* permet l'expression de l'intégrase *intI*. Celle-ci va permettre, soit l'intégration d'une cassette circulaire médiée par une recombinaison entre le site *attI* et le site *attC* de la cassette : l'intégration de cette cassette s'effectue en première position ; soit elle peut permettre l'excision d'une ou des cassettes par la recombinaison entre deux sites *attC*. Cela permet la libération d'une cassette sous forme circulaire.
- d) Après la région 5-CS et la région variable se trouvent la région 3'-CS. Dans le cas des intégrons de classe 1, il est commun de retrouver 3 gènes conservés du côté 3' de la région variable.

La région 5'CS, appelée aussi « plateforme fonctionnelle » se compose de trois éléments essentiels (figure 17.a) :

- Un gène *intI* qui, sous le contrôle du promoteur *PintI*, code pour une intégrase qui est une recombinase à tyrosine dont la séquence en acides aminés permet de déterminer plusieurs classes d'intégrons (Mazel, 2006; Messier and Roy, 2001). Il existe cinq classes d'intégrons de résistances : les intégrons de classe 1, 2 et 3 retrouvés fréquemment dans un contexte clinique (Partridge, 2011), l'intégron de classe 4 décrit sur l'ICE SXT de *Vibrio cholerae* et l'intégron de classe 5 sur le plasmide pRSV1 de *Vibrio salmonicida* (Hochhut et al., 2001; Sørum et al., 1992).
- Un site *attI* reconnu par l'intégrase permettant l'intégration des cassettes de gènes dans la région variable de l'intégron (Partridge et al., 2000).
- Un promoteur appelé « *Pc* » dont il peut exister plusieurs variants selon la classe d'intégrons et qui est situé le plus souvent dans la séquence codante de l'intégrase et permettant l'expression des cassettes en aval (Collis and Hall, 1995; Jové et al., 2010; Lévesque et al., 1994).

Les intégrons acquièrent de nouveaux gènes appelés « cassettes de gènes » dans la région variable (figure 17.b) (Hall et al., 1991). Chacune de ces cassettes est délimitée par un site de recombinaison en tige-boucle appelé site *attC* (Hall et al., 1991; Rowe-Magnus et al., 1999). Les cassettes de gènes sont intégrées par recombinaison spécifique entre le site *attI* de la plateforme fonctionnelle et le site *attC* de la cassette par l'intégrase (Stokes et al., 1997). De plus, les cassettes intégrées ont la capacité d'être excisées sous forme d'éléments d'ADN circulaires (figure 17.c) (Collis and Hall, 1992a, 1992b). Généralement, les cassettes de gènes sont dépourvues de promoteur, c'est pourquoi, les cassettes de gènes sont exprimées à partir du promoteur *Pc*, selon un gradient, à savoir que les cassettes les plus proches du promoteur sont les plus exprimées (figure 17.d) (Collis and Hall, 1995).

Certains intégrons de résistances peuvent être flanqués de MITE comme décrit chez *Acinetobacter* et *Enterobacter*. La recombinaison entre ces deux régions MITEs flanquant l'intégron permet son excision du chromosome par transposition (Gillings et al., 2009; Poirel et al., 2009). Ainsi, les MITEs pourraient conférer à l'intégron un mécanisme de mobilité.

Il ressort de ce qui précède que :

- Les plasmides sont des réplicons extra-chromosomiques dont nous trouvons de deux types :
  - Les plasmides dits « mobilisables » dont la transmissibilité du plasmide est assurée par un plasmide conjugatif corésident.
  - Les plasmides dits « conjugatifs » codent l'ensemble de la machinerie nécessaire à son transfert d'une cellule donneuse vers une cellule receveuse grâce à un système de sécrétion de type IV (T4SS).
- Les ICE ou transposons conjugatifs sont localisés sur le chromosome et capables d'excision du génome grâce à une intégrase avec une mobilité inter-génomique médiée par un système de conjugaison identique au plasmide conjugatif.
- Les éléments transposables procaryotes sont des séquences d'ADN capables de se déplacer d'un endroit à un autre, d'un site donneur vers un site cible du même génome ou sur un autre réplicon comme un plasmide. Pour leur mobilité intra-chromosomique, ces éléments transposables utilisent une enzyme particulière : la transposase.
- Les intégrons permettent de capturer et d'exprimer efficacement des gènes exogènes grâce à une intégrase et un promoteur Pc.

## Chapitre III. Les séquences d'insertion (IS) procaryotes

---

De nombreux génomes procaryotes sont mosaïques, composés d'un « génome central » ou core genome, constitué de gènes essentiels et de gènes de ménage. Le génome central comprend aussi des segments d'ADN constituant le « mobilome » se définissant comme une variété de gènes accessoires acquis faisant partie du pan-génome (Medini et al., 2005; Tettelin et al., 2008). Les séquences d'insertion (IS) font partie du « mobilome » ; elles se définissent comme les éléments génétiques mobiles les plus simples contenant uniquement des gènes liés à leur transposition et à la régulation de leur transposition (Craig et al., 2002; Kleckner, 1981; Mahillon and Chandler, 1998). La mobilité de ces IS peut conduire à l'expression, l'acquisition ou la suppression de fonctions chez l'hôte bactérien (Chao and McBroom, 1985; Elena et al., 1998).

### III.1. Classification et diversité

Les premières IS ont été décrites en 1972. Elles ont été identifiées par l'insertion de segment d'ADN dans de nombreux gènes (Hirsch et al., 1972).

Bien que les changements générés par les IS soient souvent des mutations délétères (Elena et al., 1998; Mahillon and Chandler, 1998), ces éléments sont abondants dans les génomes bactériens (Touchon and Rocha, 2007; Wagner, 2009). Il existe différentes hypothèses pour expliquer la persistance des IS chez les populations bactériennes :

- L'hypothèse de l'ADN égoïste qui est définie par des IS capables de persister chez l'hôte bactérien grâce à leur capacité à répliquer bien qu'elles n'apportent aucune contribution à l'hôte (Doolittle and Sapienza, 1980; Orgel and Crick, 1980).
- L'hypothèse adaptative selon laquelle la persistance de l'IS chez l'hôte s'expliquerait par son rôle dans la capacité à induire des mutations (Bichsel et al., 2013; Blot, 1994; Le Rouzic and Deceliere, 2005). Certains travaux ont souligné l'importance des IS comme source de diversité génétique pour l'hôte car la mobilité de ces éléments permet des changements génomiques qui sont dans certains cas bénéfiques (Bickhart et al., 2009; Casacuberta and González, 2013; El Gharniti et al., 2012; McEvoy et al., 2007). De plus, si une mutation bénéfique prévaut grâce à la sélection naturelle chez une population bactérienne, les IS peuvent se fixer en même temps que cette mutation (Chao et al., 1983; Martiel and Blot, 2002). A titre d'illustration, chez les hôtes exposés à des antibiotiques, les IS augmentent la capacité d'adaptation car impliquées dans la mobilisation et/ou l'expression des gènes de résistance aux antibiotiques (Boutoille et al., 2004; Casagrande Proietti et al., 2012; Nagel et al., 2011; Poirel et al., 2003).
- L'hypothèse neutre appelée aussi « effet de la dérive » qui s'explique par la dérive de manière aléatoire des IS vers une fixation dans le génome bactérien (Iranzo et al., 2014; Lynch and Conery, 2003). Si la taille de la population bactérienne est grande, alors la dérive de l'IS devient

plus faible (Le Rouzic et al., 2007; Lynch and Conery, 2003). Ainsi, quand les populations bactériennes sont suffisamment grandes, elles peuvent éliminer efficacement les IS qui ont des effets néfastes (Rankin et al., 2010). Bien que les coûts d'adaptation induits par les mouvements des IS sont relativement faibles, voire proches de la neutralité, il semblerait cependant difficile d'éliminer ces éléments (Dolgin and Charlesworth, 2006).

Ces trois hypothèses ne sont pas mutuellement exclusives pour expliquer l'abondance des IS du fait des interactions complexes entre le transfert horizontal de gènes, la distribution des effets d'adaptation, la dérive et le taux de transposition (Berg et al., 1989; Deceliere et al., 2006; Dolgin and Charlesworth, 2006).

Afin de limiter ou de supprimer l'activité de transposition, des mécanismes ont été mis en place par les IS et/ou leurs hôtes (Blot, 1994; Hall, 1999; Kleckner, 1981; Nagy and Chandler, 2004). Cependant, afin de répondre à des défis environnementaux, il a été observé une augmentation de l'activité de transposition (Le Rouzic and Capy, 2005; Nagel et al., 2011; Pfeifer and Blaseio, 1990). Ce phénomène a été appelé un « burst » de transposition qui se caractérise par une augmentation du taux de transposition. Chez *Escherichia coli* par exemple, le « burst » a favorisé l'adaptation à un environnement à forte osmolarité par une augmentation du taux de mutations bénéfiques (Stoebel and Dorman, 2010). Ainsi, ces changements d'activités de transposition sont un mode de réponse aux stress dans des environnements changeants. Ainsi, le « burst » permet d'augmenter le nombre d'IS et *in fine* de favoriser la persistance des IS (Foster, 2007; Le Rouzic and Capy, 2005; Mennecier et al., 2006).

Enfin, plusieurs études ont montré que certains génomes bactériens, ayant connu une expansion du nombre de copies d'IS, seraient la première étape du processus de réduction du génome et d'optimisation des génomes (Gil et al., 2008; Moran and Plague, 2004; Plague et al., 2008). En effet, cette expansion serait influencée par le mode de vie et plus précisément par le passage d'un mode de vie libre à intracellulaire, comme dans le cadre de l'endosymbiose (Moran and Plague, 2004; Plague et al., 2008). Il en résulterait une accumulation irréversible de mutations dans un environnement intracellulaire confiné du fait d'un milieu riche, ainsi de nombreux gènes de bactéries essentielles à la forme de vie libre deviennent inessentiels (Andersson and Kurland, 1998; Moran, 1996). Cette expansion des IS contribuerait à l'accumulation de pseudogènes par recombinaison intrachromosomique entre les IS et conduirait à une réduction du génome (figure 18) (Andersson and Andersson, 1999; Lawrence et al., 2001; Mira et al., 2001). Ce phénomène d'expansion et de réduction du génome favoriserait la symbiose entre deux organismes (Bordenstein and Reznikoff, 2005). En effet, il a été prouvé que l'endosymbiote *Wolbachia sp.* a été soumis à plusieurs vagues d'expansion et d'élimination des IS (Cerveau et al., 2011). En outre, il a été avancé que ce phénomène de réarrangements et suppressions à grande échelle du génome hôte médié par les IS a amélioré la capacité bactérienne à combattre les défenses de l'hôte, comme chez *Bordetella*, en

modifiant les antigènes de surface et les circuits de régulation (Parkhill et al., 2003; Preston et al., 2004).

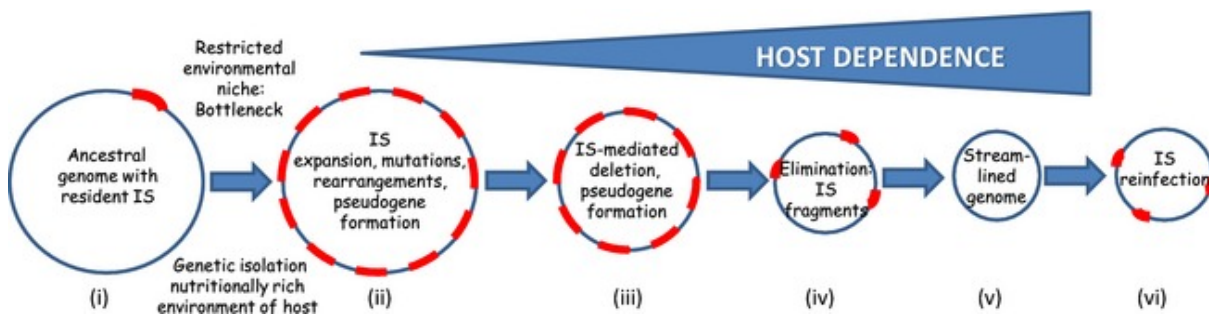


Figure 18 : Expansion et élimination des séquences d'insertions induisant la rationalisation du génome

Issue de Siguier et al., 2014

La figure montre schématiquement, de gauche à droite, les événements qui ont conduit à l'évolution de la dépendance à l'égard de l'hôte chez les bactéries.

(i) Le chromosome ancestral comprenant un faible nombre d'IS résidents (rouges).

(ii) L'expansion des IS résulte de l'isolement et de la formation de goulets d'étranglement au sein d'un organisme hôte. Elle s'accompagne d'une mutation favorisée par l'insertion de nouvelles copies d'IS et par les activités de transposition qui y sont liées, à savoir la délétion et le réarrangement.

(iii) Avec le temps, les IS auront tendance à subir une délétion avec les séquences d'ADN.

(iv) En fin de compte, une délétion importante conduit à la génération de fragments d'IS non autonomes et à leur élimination.

(vi) Il en résulte des génomes rationalisés, « libres » d'IS, qui peuvent être « réinfectés » par l'IS lors de rares contacts avec d'autres souches porteuses d'IS (v).

Actuellement, plus de 5000 IS ont été décrites grâce à l'avènement du séquençage à haut débit des génomes. Il convient de préciser que depuis 1988, le nombre d'IS ne cesse d'augmenter (Ross et al., 2021). La classification de ces éléments est nécessaire pour faire face à leur nombre élevé et à leur diversité.

Plusieurs critères sont utilisés pour classer les IS :

- L'organisation génétique,
- La taille et la structure des séquences inversées répétées, IR,
- La similarité de la séquence primaire des acides aminés de la transposase,
- Les préférences en matière de sites cibles,
- La longueur des répétitions cibles,
- La nature biochimique de la transposase pour le clivage des brins d'ADN.

Actuellement, les IS sont classées sur la biochimie de la transposase qui définit un mécanisme moléculaire de transposition, propre à chaque groupe. Cinq grands types de transposase sont actuellement définis, à savoir les transposases : DDE, DEDD, les sérines, les tyrosines et HUH (tableau 1).

Tableau 1 : Principales caractéristiques et résumé de chaque famille et sous-groupe d'IS

Abréviations : O = Oui ; N= Non ; DR = direct repeat (ou répétition directe) ; IR = inverted repeat (répétition inverse) ; ORF = open reading frame (cadre ouvert de lecture) ; ISNCY = Insertion Sequence Not Classified Yet

Famille	Sous-groupes	Taille	DR (pb)	Extrémités	IRs	Nb des ORF	Frameshift	Résidus catalytiques	Mécanisme
IS1	—	740–1180	8–9	—	O	2	ORFAB	—	Copier-coller et cointégrats
	Un seul ORF	800–1200	0–9	GGnnnTG	N	1	—	DDE	
	ISMhu11	900–4600	0–10	—	O	2	ORFAB	—	
IS1595	ISPha2	1000–1150	8	GGCnnTG	—	1	—	DDNK	Copier-coller ?
	ISPha2+pass	1500–2600	8	—	—	1+pass	—	—	
	ISH4	1000	8	CGCTCTT	—	—	—	DDNK	
	IS1016	700–745	7–9	GGGgctg	—	—	—	DDEK	
	IS1595	900–1100	8	CcTGATT	O	1	—	DDNK+ER4R7	
	ISSod11	1000–1100	8	nnnGcnTATC	—	—	—	DDHK+ER4R7	
	ISNwi1	1080–1200	8	ggnnatTAT	—	—	—	DDEK+ER4	
	ISNwi1+pass	1750–4750	8	—	—	1+pass	—	—	
ISNha5	3450–7900	8	CGGnnTT	—	1	—	DDER/K		
IS3	IS150	1200–1600	3–4	TG	—	—	—	—	Copier-coller
	IS407	1100–1400	4	TG	—	—	—	—	
	IS51	1000–1400	3–4	TG	O	2	ORFAB	DDE	
	IS3	1150–1750	3–4	TGa/g	—	—	—	—	
	IS2	1300–1400	5	TG	—	—	—	—	
IS481	—	950–1300	4–15	TGT	O	1	—	DDE	Copier-coller ?
IS1202	ISAbba32	1450–1870	5–6	TGT	—	—	—	DDE	Copier-coller
	IS7de1	1320–1780	16–17	TAT/TGT	O	1	—	DDE	
	IS1202	1440–1900	27–28	TGT	—	—	—	—	
IS4	IS10	1200–1350	9	CT	—	—	—	—	Couper-coller
	IS50	1350–1550	8–9	C	—	—	—	—	
	ISPepr1	1500–1600	7–8	-T-AA	—	—	—	—	
	IS4	1400–1600	10–13	-AAT	O	1	—	DDE	
	IS4Sa	1150–1750	8–10	CA	—	—	—	—	
	ISH8	1400–1800	10	?	—	—	—	—	
	IS231	1450–5400	10–12	CAT	—	1 ou + *	—	—	
IS701	—	1400–1550	4	—	O	1	—	DDE	—
IS701	ISAbba11	—	—	—	—	—	—	DDE	—
ISH3	—	1225–1500	4–5	C-GT	O	1	—	DDE	—
IS1634	—	1500–2000	5–6	C	O	1	—	DDE	—
IS5	IS903	950–1150	9	GG	—	—	—	—	DDE
	ISL2	850–1200	2–3	—	—	—	—	—	
	ISH1	900–1150	8	-GC	O	1	—	—	
	IS5	1000–1500	4	Ga/g	—	—	—	—	
	IS1031	850–1050	3	GAa/g	—	—	—	—	
	IS427	800–1000	2–4	Ga/g	—	2	ORFAB	—	
IS1182	—	1330–1950	0–60	—	O	1	—	DDE	—
IS6	—	700–900	8	GG	O	1	—	DDE	cointégrats
IS21	—	1750–2600	4–8	TG	O	2*	—	DDE	—
IS30	—	1000–1700	2–3	—	O	1	—	DDE	Copier-coller
IS66	—	2000–3000	8–9	GTAA	Y	3*	—	DDE*	—
	ISBst12	1350–1900	8–9	GTAA	Y	1	—	DDE	
IS256	—	1200–1500	8–9	Ga/g	—	—	—	—	Copier-coller
	IS1249	1300	0–10	GG	Y	1	—	DDE	
	ISC1250	1250	0–9	GG	—	—	—	—	
ISH6	—	1450	8	GGT	O	1	—	DDE	—
ISLre2	—	1500–2000	9	—	O	1	—	DDE	—
ISKra4	ISAzba1	1400–2900	0	—	—	1 or +*	—	—	DDE
	ISMich2	1250–1400	8	GGG	Y	1 or 2	ORFAB	—	
	ISKra4	1400–3700	9	GGG	—	1 or +*	—	—	
IS630	—	1000–1400	2*	—	O	1 or 2	ORFAB	DDE	Couper-coller
IS982	—	1000	3–9	AC	O	1	—	DDE	—
IS1380	—	1550–2000	4–5	CC	O	1	—	DDE	—
ISAs1	—	1200–1500	8–10	CAGGG	O	1	—	—	—
ISL3	—	1300–2300	8	GG	O	1	—	—	—
Tn3	—	>3000	0	GGGG	O	>1	—	DDE	cointégrats
ISAzoz13	—	1250–2200	0–4	Ga/g	O	1	—	—	—
IS110	—	1200–1550	0	—	N	1	—	DEDD	—
	IS1111	—	—	—	O	—	—	—	
IS91	—	1500–2000	0	—	N	1	—	HUH/Y2	Cercle roulant
IS200/IS605	IS200	600–750	0	—	0	1*	—	HUH/Y1	« peel-and-paste »
	IS605	1300–2000	—	—	—	2*	—	HUH/Y1**	
IS607	—	1700–2500	0	—	N	2*	—	Serine**	—
ISNCY*	IS892	1600	0–8	CTAG	O	2	—	—	—
	ISLbi1	1400–1500	5	—	O	1	—	—	
	ISMae2	1400–2400	9	CAG	O	1	—	—	
	ISPlu15	800–1000	0	—	N	1	ORFAB	—	
	ISA1214	1000–1200	8–12	—	O	2	—	—	
	ISC1217	1200	6–8	TAG	O	1	—	—	
	ISM1	1300–1600	8–9	—	O	1	—	—	
ISDol1	1600–1900	6–7	—	O	1	—	DDE		



En plus de la biochimie de la transposase, les IS peuvent être classées en plusieurs types selon les extrémités qui bordent l'élément, ainsi que la génération ou non d'une duplication du site d'insertion de l'élément. Les IS ont donc été classées en groupes ou en familles apparentés qui sont souvent divisées en sous-groupes. En 2014, Siguier et al. ont réalisé une représentation de la distribution des familles d'IS en fonction de la biochimie de la transposase et le nombre de membres par famille (figure 19) (Siguier et al., 2014).

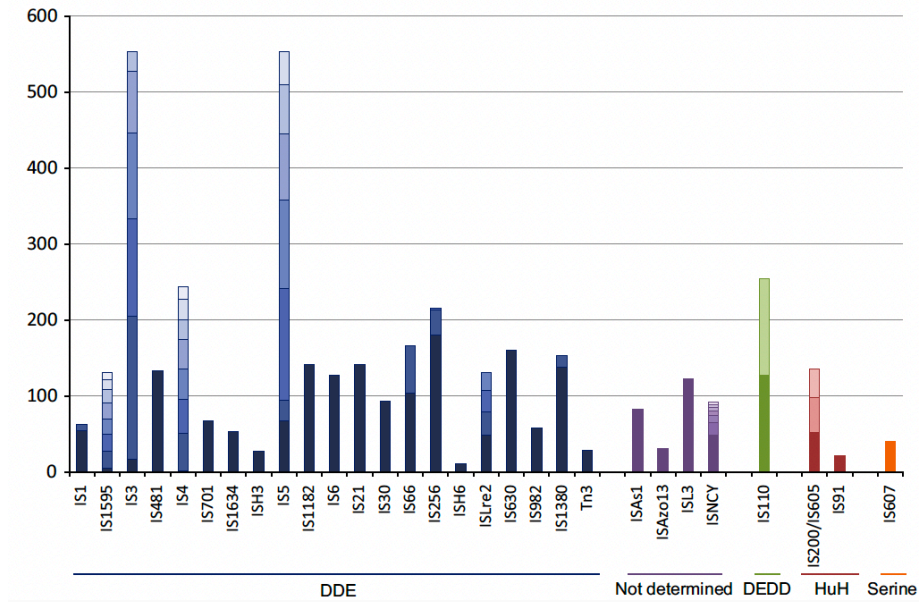


Figure 19 : Répartition des familles d'IS

Issue de Siguier et al., 2014

L'histogramme montre le nombre d'IS d'une famille donnée issue de la base de données ISfinder en juin 2013. Les cases horizontales indiquent le nombre et la taille relative des différents sous-groupes au sein de la famille. Ils sont regroupés par couleur en fonction du type de transposase utilisé : DDE (bleu), DEDD (vert), HUH (rouge), sérine (orange) et indéterminé (violet).

Un simple numéro a été attribué au départ à chaque IS (Campbell et al., 1979). Cependant, à cause de l'augmentation constante d'IS décrites, une nouvelle nomenclature dû être instaurée. La nomenclature actuelle est : la première lettre du genre suivie des deux premières lettres de l'espèce et d'un numéro (Chandler and Mahillon, 2000).

## III.2. Les IS bactériennes « classiques »

### III.2.1. Structure et caractéristiques fonctionnelles

Comme vu précédemment, les IS sont définies comme un segment d'ADN court, généralement cryptique sur le plan phénotypique, codant uniquement pour les enzymes nécessaires à sa transposition et capables d'insertion (Craig et al., 2015, 2002; Siguier et al., 2015). Les IS classiques se caractérisent de la manière suivante (Berg et al., 1989; Craig et al., 2002) :

- Une longueur comprise entre 0,7 et 2,5 kb,
- Sont compactes et comportent un ou deux orfs,
- Se terminent par des séquences répétitives terminales imparfaites (IR) flanquantes.

Les IS codent une transposase qui est une enzyme catalytique catalysant les clivages de l'ADN et les transferts de brins conduisant *in fine* au mouvement de l'IS. Dans certain cas, elles codent également des protéines régulatrices. Compte tenu de leur structure compacte, les IS ont mis en place des stratégies pour optimiser l'expression de leurs gènes, tel que le déplacement programmé des cadres ribosomiques et le réalignement transcriptionnel (Baranov et al., 2001; Chandler and Fayet, 1993; Prère et al., 2011; Sharma et al., 2011; Siguier et al., 2014).

La fonction de la transposase est de reconnaître et cliver les séquences IR. En effet, ces IR sont indispensables à la mobilité des IS. La taille de ces séquences IR varie de 20 à 40 pb (Huisman et al., 1989; Johnson and Reznikoff, 1983; Makris et al., 1988; Mugnier et al., 2009). Les IR sont organisées en deux domaines fonctionnels (Huisman et al., 1989; Johnson and Reznikoff, 1983; Normand et al., 2001; Zerbib et al., 1990) :

- Le domaine appelé « a » : positionné à l'intérieur de l'IR et intervenant dans la liaison de la transposase.
- le domaine appelé « b » : positionné à l'extérieur de l'IR, il est impliqué dans les clivages et les réactions de transfert de brin conduisant à la transposition de l'élément.

Cela étant, les membres de la famille IS21 portent de multiples séquences répétées aux deux extrémités représentant plusieurs sites de liaison de la transposase (Berger et al., 2001; Mahillon and Chandler, 1998).

Un modèle général d'organisation fonctionnelle des transposases semble émerger (Haren et al., 1999; Stalder et al., 1990; Tavakoli et al., 1997; Wiegand and Reznikoff, 1994). En effet, les activités spécifiques de liaison à l'ADN sont majoritairement situées dans la région N-terminale, tandis que le domaine catalytique est souvent localisé vers l'extrémité C-terminale. Outre la présence des domaines fonctionnels de liaison à l'ADN et de catalyse, de nombreuses transposases ont la capacité de générer des formes multimériques essentielles à leur activité (Haren et al., 2000, 1999).

Un type particulier de structure, les leucines zippers, joue un rôle vital dans la multimérisation d'un certain nombre de transposases (Haren et al., 2000, 1998; Lewis et al., 2011).

Enfin, il a été démontré que certaines IS ne présentent que peu de spécificités de séquence dans le choix de leur cible. Il a été mis en exergue que l'élément IS630 exige un dinucléotide <sup>5'</sup>TA dans la cible (Feng and Colloms, 2007). Or, d'autres familles comme IS200/IS605 et IS91, exigent de courtes séquences tétra- ou penta-nucléotidiques (Guynet et al., 2009; Mendiola and de la Cruz, 1989). De plus, il a été mis en évidence que les éléments IS1 et IS186, présentent une certaine spécificité pour les séquences riches en AT et en GC, respectivement (Meyer et al., 1980; Sengstag et al., 1986).

D'autre part, certaines IS ou certains transposons pourraient être attirés par les fourches de réplication, ils présentent un fort biais d'orientation indiquant une préférence de brin à la fourche (Peters and Craig, 2001a, 2001b; Ton-Hoang et al., 2010). Dans d'autres cas, l'IS cible des fourches de réplication bloquées (S. He et al., 2015). De plus, si les IS ne sont pas bien orientées par rapport à la fourche de réplication, ces éléments sont très certainement inactifs et incapables de transposer (Ton-Hoang et al., 2010).

Enfin, les membres des familles IS110, IS3 et IS4 s'insèrent dans des structures secondaires potentielles telles que les palindromes extragéniques répétés (Clément et al., 1999; Tobes and Pareja, 2006; Wilde et al., 2003), les intégrons (Post and Hall, 2009; Tetu and Holmes, 2008) ou encore les extrémités d'autres éléments transposables (Hallet et al., 1991; Partridge and Hall, 2003a). En outre, il a été démontré que les éléments IS21, IS30 et IS911 s'insèrent tous à proximité des séquences qui ressemblent à leurs propres IR (Loot et al., 2004; Olasz et al., 1997; Reimann and Haas, 1987).

Une autre caractéristique des IS classiques : lors de l'insertion de l'élément, elles génèrent de courtes séquences directement répétées (DR) de l'ADN cible. Cette duplication est générée par la réparation de l'espacement entre les sites de clivage. La longueur du DR est généralement comprise entre 2 et 14 pb. Les DR sont caractéristiques d'un élément donné et génèrent une duplication de longueur fixe. C'est déterminé par l'architecture du complexe de transposition qui impose des contraintes sur la distance entre les clivages sur chaque brin de l'ADN cible (Dyda et al., 2012; Montañó and Rice, 2011). Dans certains cas, l'absence de DR peut être observée et traduire le résultat d'une recombinaison homologue inter- ou intra-moléculaire entre deux IS, où chacune a des séquences DR différentes. Dans ce cas, une seule copie de la séquence DR est située sur chacun des produits de délétion réciproque (Susu He et al., 2015).

Pour finir, les éléments IS1549, IS1634 et IS1630 semblent générer de longs DR de longueur très variable (Calcutt et al., 1999; Plikaytis et al., 1998; Vilei et al., 1999). Cependant, le mécanisme impliqué dans la génération de ces longs DR est actuellement inconnu.

## III.2.2. Mécanismes biochimiques employés par les transposases des IS « classiques »

### III.2.2.1. Les transposases de type DDE

Les transposases DDE présentent des domaines catalytiques qui se caractérisent par une triade conservée d'acides aminés essentiels : aspartate (D), aspartate (D), glutamate (E) (Craig et al., 2002). Cette triade permet de coordonner un ion métallique divalent et utilise un OH comme nucléophile dans une réaction de transestérification (Haren et al., 1999; Hickman and Dyda, 2015). Les transposases DDE ont une séquence de taille variable, entre 0,7 à 2,5 kpb de long avec un ou deux cadres de lecture ouverts selon les éléments (Siguier et al., 2014). De plus, les IS se terminent par des IR imparfaits et génèrent de courts DR de longueur variable lors de l'insertion.

Les transposases DDE sont le type le plus abondant chez les organismes procaryotes et eucaryotes. Il existe de nombreuses familles de transposase DDE dont la famille IS3 est la plus répandue chez les espèces bactériennes (Ross et al., 2021).

De façon générale, les transposases DDE présentent deux domaines :

- Le domaine qui assure la liaison à l'ADN spécifique à la séquence, ce domaine est généralement situé vers l'extrémité N-terminale,
- Le domaine catalytique qui est davantage localisé vers l'extrémité C-terminale.

D'un point de vue fonctionnel, cet arrangement n'est pas anodin car il permet la liaison co-translationnelle à l'ADN, c'est-à-dire que la protéine se replie au cours de la traduction permettant à la chaîne polypeptidique naissante de commencer à se lier directement à l'extrémité IS. Ce phénomène a été démontré chez plusieurs éléments comme IS1, IS3, IS30 (Haren et al., 1998; Stalder et al., 1990; Zerbib et al., 1987). De plus, les domaines Hélix-tour-Hélix (HTH) et doigt de zinc ont été impliqués comme domaines de liaison à l'ADN, ce qui rappelle le cas chez l'élément IS1 et IS911 (Ohta et al., 2004; Rousseau et al., 2004).

Toutes les transposases DDE contiennent un motif commun et un repliement similaire à celui de la RNase H (Rice and Baker, 2001). Il a été démontré que ce repliement chez la RNase H permet un positionnement proche des résidus DDE (Nowotny, 2009). La triade DDE est souvent suivie d'un résidu basique, généralement K ou R et elle porte le résidu E conservé pouvant être remplacé par un résidu N ou H (Ross et al., 2021). Les éléments de la famille IS3 ont des similitudes avec les intégrases rétrovirales (IN) que ce soit par l'espacement de la triade DDE ou par l'apparition de résidus conservés supplémentaires (Fayet et al., 1990; Haren et al., 1999).

Les transposases DDE partagent une chimie de transposition de base, mais elles varient dans les étapes menant à la formation d'un intermédiaire d'insertion unique (Hickman et al., 2010a; Hickman and Dyda, 2015). Ces transposases catalysent le clivage d'un seul brin d'ADN ce qui génère un groupement  $^3\text{OH}$  aux extrémités. Ce groupement est utilisé comme nucléophile pour attaquer le

squelette phosphate de l'ADN cible. Cependant, la transposition de ces IS se produit *via* des intermédiaires d'ADN à double brin, ce qui nécessite le traitement du second brin pour libérer l'IS des séquences d'ADN flanquantes dans la molécule donneuse. Ce processus peut se dérouler de différentes manières et il est spécifique à chaque famille d'IS. Ces variations, pour lesquelles le second brin est traité pour la formation d'un intermédiaire, expliquent les différences dans la transposition de chaque famille d'IS (Curcio and Derbyshire, 2003; Hickman and Dyda, 2015; Turlan and Chandler, 2000).

En outre pour certaines IS, le second brin n'est pas clivé mais la réplication qui suit le transfert du premier brin fusionne les molécules du donneur et de la cible, cela permet de générer des cointégrats avec une copie directement répétée à chaque jonction donneur/cible (Craig et al., 2002). Ce processus de réplication est appelé « la transposition répllicative » ou « Target Primed Replicative Transposition », les IS6, le phage Mu et le transposon Tn3 bénéficient de ce processus (figure 20).

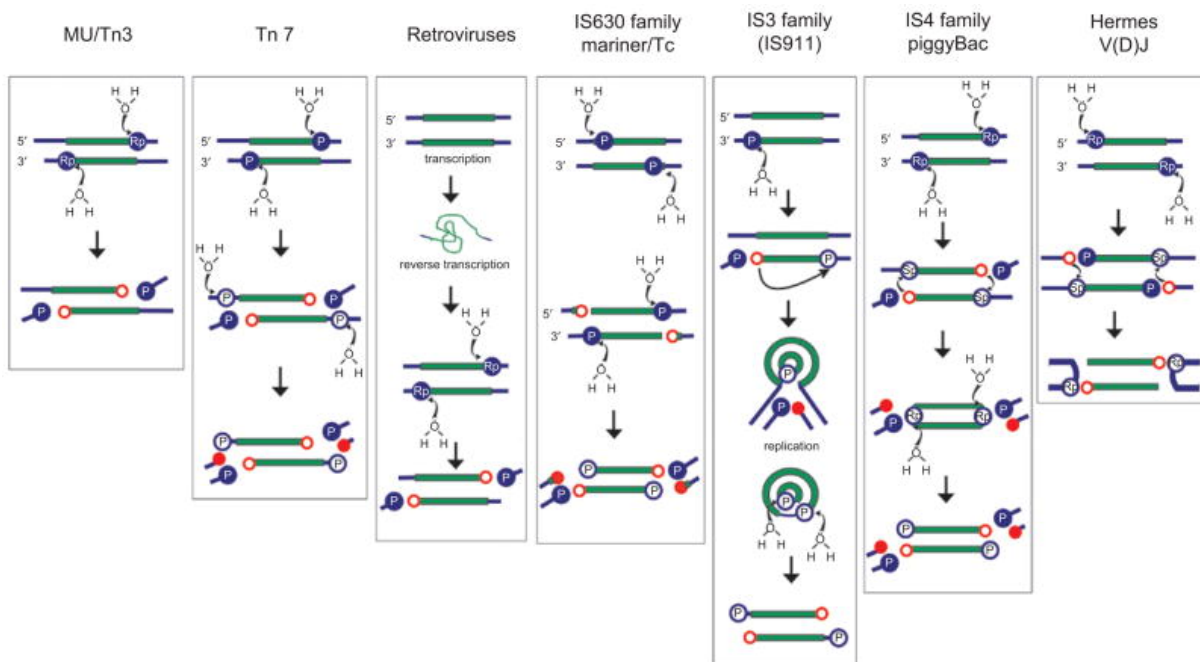


Figure 20 : Principales voies de transposition des transposases DDE

Issue de Hickman et al., 2010a  
 Pour la description, voir texte.

Pour les IS de la famille IS4 et les transposons Tn5 et Tn10, après l'attaque nucléophile initiale, le 3'OH libre est utilisé pour attaquer le brin opposé dans une réaction de transfert trans-brin. Cela permet de générer un intermédiaire en épingle à cheveux transitoire à l'extrémité de l'IS (Haniford and Ellis, 2015). Ce pont est ensuite clivé en utilisant H<sub>2</sub>O comme nucléophile pour libérer l'IS et régénérer le 3'OH sur le brin transféré. Ainsi, l'IS est prête pour le transfert de brin et l'intégration. Ce processus de réplication est appelé « couper-coller ».

Chez les éléments de la famille IS630, l'ADN donneur flanquant est éliminé. En effet, dans ce mécanisme le brin non transféré est clivé de 2 bases avant le clivage du brin transféré (Feng and Colloms, 2007). Ce processus est le même que leurs cousins eucaryotes : les transposons de la famille Tc1/mariner (Plasterk, 1996; Richardson et al., 2009).

Concernant la famille IS3, après l'attaque nucléophile initiale, le  $3^{\circ}\text{OH}$  libre de l'IS sert à attaquer plusieurs nucléotides à l'extérieur de la deuxième extrémité pour générer un pont simple brin laissant un  $3^{\circ}\text{OH}$  libre sur le flanc de l'IS (Chandler et al., 2015). Ensuite, la réplication régénère une copie intacte de l'IS reconstituant le plasmide donneur et elle produirait un ADN intermédiaire à double brin circulaire. La formation de l'intermédiaire circulaire permet l'assemblage d'un promoteur transitoire fort (Ton-Hoang et al., 1997) pour piloter la synthèse de la transposase. Ainsi, une fois généré, l'intermédiaire circulaire est engagé à terminer la transposition. Ce mécanisme de transposition est connu sous le nom de « copier-coller » ou « Donor Primed Replicative Transposition ».

Enfin, dans le dernier modèle démontré uniquement chez les eucaryotes de la famille hAT et la recombinaison V(D)J : la coupure initiée par la transposase permet la libération d'un  $3^{\circ}\text{OH}$  au flanc du brin non transféré qui à son tour peut attaquer le second brin, ce qui permet la formation d'une épingle à cheveux sur l'ADN flanquant et un  $3^{\circ}\text{OH}$  sur le brin transféré (Zhou et al., 2004).

En outre, de nombreuses familles produisent des intermédiaires circulaires double brin. Toutefois, cela ne signifie pas nécessairement qu'elles utilisent toutes le même mécanisme. En effet, un cercle pourrait formellement être généré par une excision impliquant une recombinaison des deux brins (figure 21) (Hickman and Dyda, 2015). Ces différences se reflètent dans les différentes familles d'IS.

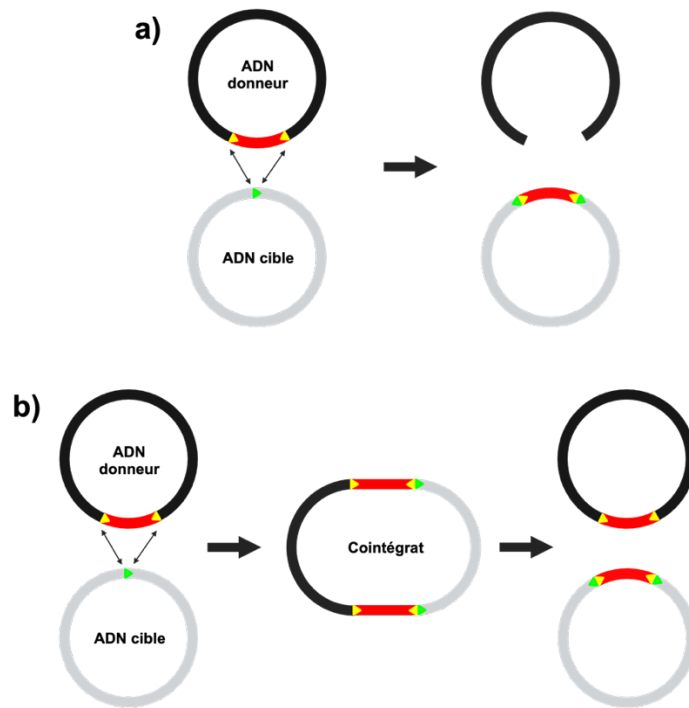


Figure 21 : Les deux modes principaux de transposition

Created with BioRender.com

a) Dans la transposition conservative, l'élément transposable (représenté par un rectangle rouge) est excisé de l'ADN donneur (cercle noir) par des coupures double brin à chaque extrémité (triangles jaunes), puis transféré dans l'ADN cible.  
 b) La transposition répliquative conduit à la formation d'un cointégrat dans lequel l'ADN donneur et cible sont fusionnés, la réplication de l'élément génère deux copies dont l'une est attachée à l'ADN donneur et l'autre est liée à la cible. Le cointégrat est résolu par recombinaison homologue de site spécifique entre les deux copies de transposon. Des répétitions directes (représentées par des triangles vert) sont généralement générées, dans les deux modes de transposition, au niveau du site d'insertion.

Chez les transposases DDE, le site cible d'insertion est choisi par une interaction directe soit entre l'ADN cible et la transposase, soit entre la transposase et d'autres protéines accessoires. Le transposon Tn10 et l'élément IS10 s'insèrent par préférence au niveau de la séquence consensus 5'GCTNAGC (Halling and Kleckner, 1982). De plus, il a été démontré que les séquences adjacentes à ce consensus ont un impact sur la sélectivité de la cible. Le transposon Tn9 et l'élément IS1 privilégient des sites riches en nucléotides AT (Zerbib et al., 1985). Enfin, les transposons eucaryotes de la famille Tc/mariner génèrent une duplication 5'TA et ciblent la séquence consensus 5'CAYATATRTG (Korswagen et al., 1996). *A contrario*, le transposon Tn7 s'insère spécifiquement au site *attTn7* dans le chromosome de *Escherichia coli* et très rarement au niveau d'autres sites non-*attTn7* (Waddell and Craig, 1988).

### III.2.2.2. Les transposases de type DDED

Les transposases à motif catalytique DEDD se caractérisent par l'enchaînement des acides aminés : aspartate (D), glutamate (E), aspartate (D), aspartate (D). A ce jour, elles ne concernent que la famille IS110 (Siguier et al., 2014).

En 1985, l'élément IS110 a été identifié chez *Streptomyces coelicolor* comme étant un élément dérivé du bactériophage phiC31 porteur d'un gène de résistance à la viomycine (Chater et al., 1985). Cette famille IS110 est très répandue et est retrouvée chez environ 130 espèces de bactéries et d'archées différentes (Ross et al., 2021). De plus, certains membres de cette famille sont spécifiques d'une espèce bactérienne comme IS900 et IS901, spécifiques de *Mycobacterium paratuberculosis*, ou encore l'IS1111 de *Coxiella burnetti* et l'IS6110 de *Mycobacterium tuberculosis* qui sont utilisés comme marqueurs d'identification de ces souches (Hoover et al., 1992; Roychowdhury et al., 2015; Rozental et al., 2012; Vary et al., 1990).

Les membres de cette famille ont une taille moyenne d'environ 1450 pb. La famille IS110 est différente de la famille des transposases de type DDE car :

- Elle ne possède pas de séquences IR,
- Elle ne génère généralement pas de duplication directement répétée lors de l'insertion.

Ces différences majeures permettent de constater que les transposases DEDD utilisent un mécanisme de transposition différent de celui des transposases DDE. Cela explique que les membres de la famille IS110 présentent généralement une seule *orf* codant la transposase DEDD.

La transposition des transposases DEDD implique des intermédiaires de jonctions de Holliday qui doivent être résolus en utilisant un mécanisme de type RuvC (Ariyoshi et al., 1994; Buchner et al., 2005; Duckett et al., 1988). Ce type de recombinaison serait cohérent avec la relation étroite entre les éléments de la famille IS110 et les invertases Piv/MooV qui, pendant l'inversion, résolvent les structures de jonctions de Holliday (Tobiason et al., 1999).

Par ailleurs, il est connu que plusieurs membres de la famille IS110 produisent des intermédiaires circulaires à double brin (Lauf et al., 1999; Perkins-Balding et al., 1999). La transposition conduit souvent à des insertions de dimères d'IS en tandem, comportement qui pourrait indiquer un certain type de mécanisme d'insertion par cercle roulant, tel que cela a été observé pour les éléments de la famille IS91 (présenté dans le prochain chapitre IV sous partie 1).

Un fait intéressant est que les intermédiaires circulaires à double brin entraînent l'assemblage d'un promoteur de jonction formé : d'un élément promoteur -35 à l'extrémité droite orienté vers l'extérieur et d'un élément promoteur -10 à l'extrémité gauche orienté vers l'intérieur (Duval-Valentin et al., 2001; Ton-Hoang et al., 1997). A l'origine, cette caractéristique a été identifiée dans les formes circulaires d'IS492 pour la famille IS110 (Perkins-Balding et al., 1999).



En dépit des points abordés, il reste de nombreuses inconnues dans la famille IS110 en ce qui concerne leur mécanisme de transposition et leur spécificité d'insertion et de séquence.

### III.2.2.3. Les transposases à sérine

Les transposases à sérine ne concernent que la famille des IS607, auparavant connues sous le nom IS1535. En 1999, IS1535 fut le premier membre de cette famille à avoir été identifié chez *Mycobacterium tuberculosis* (Gordon et al., 1999). Cette étude de Gordon et al. a permis d'identifier différentes séquences d'insertion apparentées (IS1535, IS1539 et IS1602). Ces éléments ont été regroupés dans une nouvelle famille d'IS appelée famille IS1535. L'élément IS607 a été identifié pour la première fois en 2000 dans une souche de *Helicobacter pylori* (Kersulyte et al., 2000). Et depuis cette IS a été décrite dans environ 20 % des souches de *H. pylori* dans le monde. Des membres de la famille IS607 ont également été isolés chez des cyanobactéries et des archées (J. Filée et al., 2007; Kuno et al., 2010). De plus, des séquences apparentées ont été trouvées dans des génomes eucaryotes et des virus (Jonathan Filée et al., 2007; Gilbert and Cordaux, 2013).

Il convient de souligner que les éléments de la famille IS607 ont une taille moyenne de 2000 pb et portent deux cadres de lecture qui se chevauchent. La configuration entre les deux orfs (TnpA et TnpB) est la suivante : le codon stop de l'orf en amont chevauche le codon de départ de l'orf en aval (Ross et al., 2021). Ce chevauchement suggère que sur le plan de la transcription, l'expression du gène en aval est couplée à celle du gène en amont (Huber et al., 2019; McCarthy and Gualerzi, 1990). Concernant l'élément IS607, il a été rapporté la présence de deux petites orfs en amont du gène de la transposase *tnpA* (Kersulyte et al., 2004, 2000).

L'ensemble des membres de la famille IS607 présente aux extrémités plusieurs courtes séquences directement répétées mais qui ne sont pas apparentées (Boocock and Rice, 2013; W. Chen et al., 2018). De plus, ces régions peuvent contenir une séquence répétée imparfaite plus longue, proche des extrémités, mais à des distances différentes. En outre, Il a été signalé qu'un grand nombre de membres de la famille IS607 porte un trinuéclotide répété à chaque extrémité. Il a aussi été suggéré que ces trinuéclotides seraient impliqués dans le mécanisme de transposition (Boocock and Rice, 2013).

Chez *Escherichia coli*, il a été démontré que l'IS607 s'insérait dans un dinuéclotide <sup>5'</sup>GG (Boocock and Rice, 2013; W. Chen et al., 2018). Dans certains cas, l'insertion de l'élément génère des répétitions cibles de 2 pb, *a contrario* dans d'autre cas aucune répétition cible n'a été observée.

Chez *Helicobacter pylori*, il a été constaté que l'élément IS609 s'insérait à son extrémité gauche à côté d'un trinuéclotide <sup>5'</sup>TAT dans de nombreux sites chromosomiques (Kersulyte et al., 2004).

Concernant le mécanisme de transposition de cette famille, nous ne trouvons que peu de données expérimentales disponibles. Il est supposé que la famille IS607 utilise un mécanisme de

transposition formant des intermédiaires circulaires fermés à double brin. Cela est notamment démontré par la présence d'une jonction LE-RE (Left-End – Right-End) identifiée par PCR dans un système de transposition chez *Escherichia coli* (Boocock and Rice, 2013). En outre, les deux extrémités de l'IS seraient liées par le dinucléotide <sup>5'</sup>GG dans l'intermédiaire circulaire pouvant être clivé par TnpA. La recombinaison entre le site de coupure dans la forme circulaire et un dinucléotide cible permettrait l'insertion de l'élément dans la matrice receveuse (figure 22) (Boocock and Rice, 2013). Selon ce modèle, la recombinaison se déroule au sein d'un complexe tétramérique qui relie l'ADN donneur et receveur (Grindley et al., 2006). La sérine conservée attaquerait l'ADN sur une liaison phosphodiester qui le clive, permettant la formation d'un intermédiaire covalent protéine-ADN <sup>5'</sup>phosphosérine en déplaçant le <sup>3'</sup>OH. On suppose que par un mécanisme de pivotement à 180°, la liaison protéine-ADN <sup>5'</sup>phosphosérine se retrouve face à un groupement <sup>3'</sup>OH libre. La recombinaison se produit alors par l'inversion chimique de la réaction de clivage, c'est-à-dire que les <sup>3'</sup>OH attaquent les liaisons <sup>5'</sup>phosphosérines. Ainsi, la réaction de recombinaison est chimiquement réversible. Ce modèle se résume en 3 étapes importantes : le clivage, l'échange et la religation. Cette transposition ne nécessite pas l'intervention de cofacteurs ou d'ions métalliques (Hickman and Dyda, 2015).

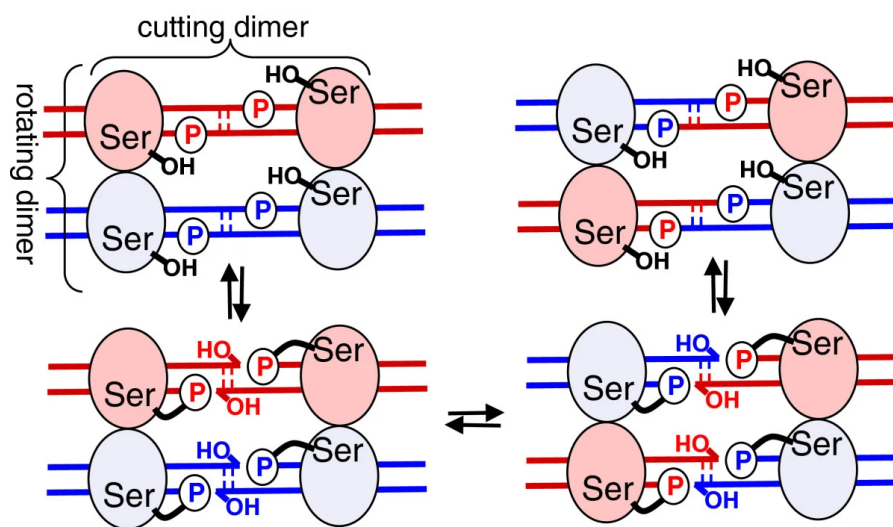


Figure 22 : Mécanisme des recombinases à sérine

Issue de Boocock and Rice, 2013  
 Pour la description, voir le texte.

#### III.2.2.4. Les transposases à tyrosine

La plupart des transposons conjugatifs ou ICE sont mobilisés par des tyrosine transposases, dont il est suggéré qu'elles sont structurellement et mécaniquement liées aux tyrosines recombinases site-spécifiques bien caractérisées comme Cre, ou l'intégrase du phage  $\lambda$  (Grindley et al., 2006; Hickman et al., 1997). Les transposases de la famille des transposons conjugatifs Tn916 et du transposon conjugatif CTnDOT sont les plus étudiées (Roberts and Mullany, 2009; Wood and Gardner, 2015).

C'est en 1981 que pour la première fois l'existence d'une résistance à un antibiotique, transférable, codée par le chromosome, et non par un plasmide, a été supposée. Il a été montré que la résistance à la tétracycline était conférée par un élément intégratif appelé transposon conjugatif et désigné sous le nom de Tn916 (Franke and Clewell, 1981). Il s'agissait du premier transposon conjugatif porteur d'un gène de résistance aux antibiotiques. Ils ont ainsi émis l'hypothèse que Tn916 pourrait expliquer la généralisation de la résistance à la tétracycline parmi les streptocoques. La plupart des éléments de la famille Tn916 possède le gène *tet(M)* conférant une résistance aux tétracyclines (Flanagan et al., 1994; Roberts et al., 2006).

Les éléments CTnDOT sont responsables de la majorité des transferts de résistance aux antibiotiques chez *Bacteroides* (G. Whittle et al., 2002). Cet élément est un transposon conjugatif de 65 kb qui porte des gènes codant pour la résistance aux antibiotiques érythromycine (*ermF*) et tétracycline (*tetQ*). Il est intéressant de remarquer que la tétracycline stimule le transfert conjugatif de CTnDOT (Gabrielle Whittle et al., 2002).

Ces transposases semblent avoir adapté un mécanisme de recombinaison spécifique de site, passant par une jonction de Holliday, pour catalyser les étapes de la transposition. Les transposases à tyrosine clivent l'ADN en utilisant un résidu tyrosine du site actif pour attaquer le phosphate scissile et former une liaison covalente <sup>3</sup>phosphotyrosine. Cette polarité de clivage est opposée à celle des trois autres types de transposases qui forment des liaisons covalentes <sup>5</sup>phosphotyrosine. Par analogie avec les recombinases à tyrosine, un complexe tétramérique s'assemblerait et dans lequel chaque sous-unité cliverait l'un des quatre brins des sites de recombinaison (figure 23). Lors de l'étape d'excision, la recombinaison entre les deux extrémités du transposon (*attL* et *attR*) entraîne la formation d'un intermédiaire circulaire libre dans lequel les extrémités sont accolées. Pour générer cet intermédiaire, les transposases à tyrosine effectuent des coupes décalées aux extrémités du transposon, dans lesquelles des paires de résidus tyrosine du site actif sont liées de manière séquentielle et covalente. L'intermédiaire circulaire est ensuite transféré dans la cellule réceptrice par conjugaison. Une fois dans la cellule réceptrice, un nouvel événement de recombinaison a lieu, c'est alors que l'intermédiaire est intégré au niveau d'un site *attB*.

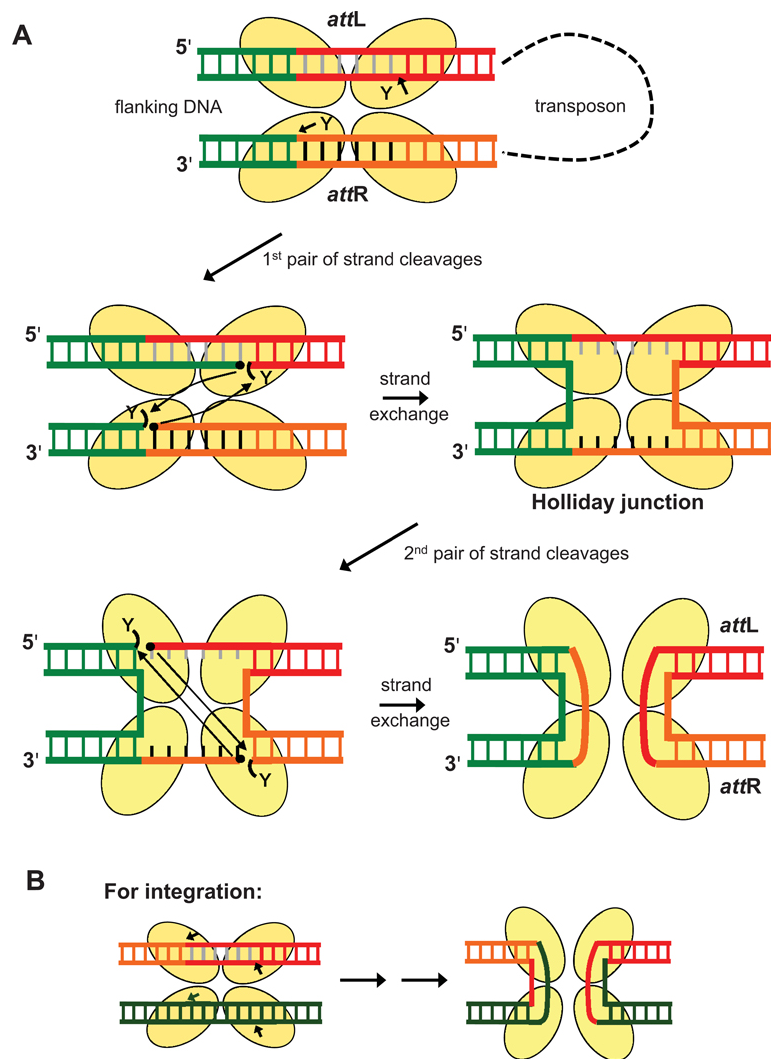


Figure 23 : Voies proposées d'excision et d'intégration par les transposases à tyrosine.

Issue de Hickman and Dyda, 2015

En haut, on voit un assemblage tétramérique réunissant l'extrémité gauche (orange) et l'extrémité droite (rouge) d'un transposon circulaire excisé avec un ADN cible (vert). Pour la description, voir le texte.

Les transposons conjugatifs présentent un spectre de spécificité de ciblage dans leurs exigences d'homologie entre la jonction *attL/attR* et le site *attB* (Rajeev et al., 2009). Certains s'insèrent essentiellement au hasard, tandis qu'au contraire d'autres présentent une spécificité relativement stricte reflétant la machinerie de recombinase spécifique de site. D'autres encore occupent une position intermédiaire avec des exigences de séquence strictes, bien que généralement très courtes, pour l'endroit où ils s'intègrent (Rajeev et al., 2009; Roberts and Mullany, 2009; Waters and Salysers, 2013).

Le transposon Tn916 peut utiliser plusieurs sites cibles et il a été démontré que ceux-ci sont riches en AT. Chez *Butyrivibrio proteoclasticus* B31620, la caractérisation de 123 sites d'insertion dans le

génomique a démontré la présence d'une séquence consensus TTTTTTATATAAAAA qui est utilisée pour l'insertion du transposon Tn916 (Cookson et al., 2011).

CTnDOT est très sélectif dans les sites d'insertion car il s'intègre uniquement dans quelques sites cibles (Bedzyk et al., 1992; Cheng et al., 2000).

### **III.3. Les IS bactériennes atypiques à motif HUH**

Si les transposons « couper-coller » codant pour des transposases DD(E/D) sont les plus courants, une deuxième classe d'éléments code pour des endonucléases HUH. Les endonucléases HUH reconnaissent spécifiquement l'ADN simple brin (ADNsb/ss). Elles sont largement répandues dans tous les domaines du vivant (Chandler et al., 2013).

#### **III.3.1. Caractéristiques structurelles et fonctionnelles**

Les membres des familles appartenant aux IS bactériennes atypiques diffèrent fondamentalement des IS classiques par leur organisation et leur structure. En effet, contrairement aux IS classiques qui portent des répétitions inversées à leurs extrémités, les IS atypiques présentent des structures palindromiques subterminales avec des fonctions distinctes (Garcillán-Barcia et al., 2002; S. He et al., 2015).

D'une part, ces IS atypiques s'insèrent spécifiquement en 3' de tétra- ou penta-nucléotides appelé « site cible d'insertion ». De plus, l'insertion de ces IS n'entraîne pas de duplication du site cible d'insertion, contrairement aux IS classiques (Garcillán-Barcia et al., 2002; S. He et al., 2015).

D'autre part, la transposase des IS atypiques ne partage pas les mêmes caractéristiques que celle des IS classiques. Leurs caractéristiques spécifiques comprennent un motif HUH, un site catalytique comprenant une ou deux tyrosines et parfois un doigt de zinc potentiel (Chandler et al., 2013). Ces IS catalysent le clivage et la jonction de substrats d'ADN sous forme simple brin avec un mécanisme de réplication par cercle roulant.

#### **III.3.2. Mécanismes biochimiques des transposases des IS atypiques**

Le premier membre identifié de la superfamille HUH est la protéine A (gpA) du bactériophage ΦX174 (Kornberg and Baker, 1992). Par la suite, de nombreuses protéines apparentées à gpA ont été identifiées par analyse bio-informatique (Ilyina and Koonin, 1992; Koonin and Ilyina, 1993), comme les protéines impliquées dans la réplication virale et plasmidique par réplication en cercle roulant (CR), la transposition de l'ADN ou encore le transfert conjugatif de plasmides (Garcillán-Barcia et al., 2002, 2009; Kapitonov and Jurka, 2001; Ronning et al., 2005; Toleman et al., 2006a; Ton-Hoang et al., 2005). L'ensemble de ces protéines présente certains motifs protéiques conservés, notamment le motif appelé « HUH ». Ce motif HUH est composé de deux résidus histidine (H) qui

sont séparés par un résidu hydrophobe volumineux (U). De plus, cette protéine présente un motif nommé Y formant un site catalytique et contenant un ou deux résidus tyrosine.

Les motifs HUH et Y sont retrouvés dans les séquences d'insertion de la famille IS91/ISCR ou IS608 de *Helicobacter pylori*, membre de la famille IS200/IS605. Ces deux motifs sont aussi partagés par les protéines Rep de certains plasmides et phages (pC194, pUB110, φX174...) (figure 24) (Chandler et al., 2013; Garcillán-Barcia et al., 2002; S. He et al., 2015; Ross et al., 2021).

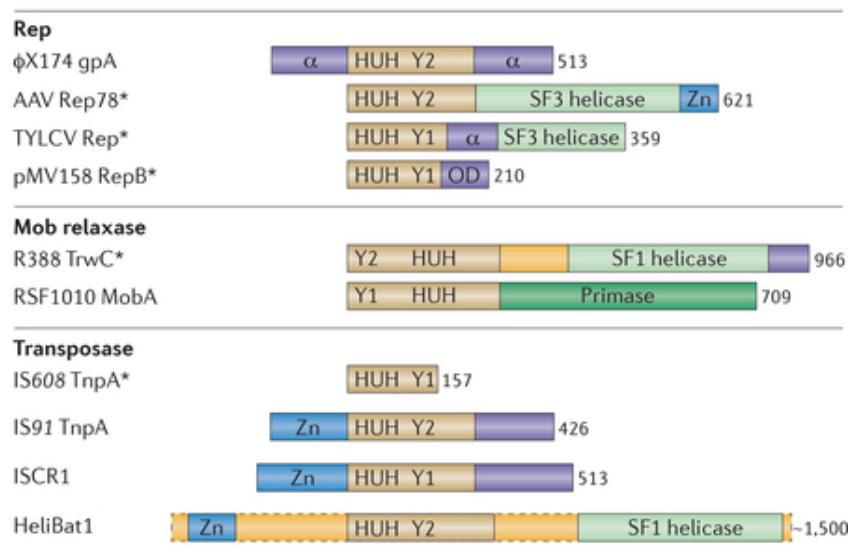


Figure 24 : Organisation de certaines endonucléases de la superfamille HUH  
Issue de Chandler et al., 2013

Nous trouvons deux grands groupes dans la super famille HUH :

- Le groupe HUH Y1 comportant une seule tyrosine : composé d'un grand nombre de protéines de virus eucaryotes (Rosario et al., 2012), des relaxases de plasmides conjugatifs (de la Cruz et al., 2010; Guglielmini et al., 2011), des protéines Rep où l'intermédiaire de réplication est un simple brin (comme pour pUB110) (Gruss and Ehrlich, 1989), les transposases de la famille des IS200/IS605 (Ronning et al., 2005; Ton-Hoang et al., 2005), ainsi que les ISCR qui sont apparentées à la famille IS91 (Toleman et al., 2006a).
- Le groupe HUH Y2 comportant deux tyrosines : décrit chez certaines cyanobactéries, des plasmides archéens, des parvovirus, le virus de l'hépatite C (Ross et al., 2021), et incluant la gpA du bactériophage φX174 et les protéines Rep des phages isométriques à ADN simple brin et double brin (comme le phage P2) (Odegrip and Haggård-Ljungquist, 2001), certaines relaxases de la famille MOB<sub>F</sub> (Chandler et al., 2013), les transposases de la famille IS91 et les héliatrons (Garcillán-Barcia et al., 2002; Kapitonov and Jurka, 2001).

Les protéines HUH utilisent un mécanisme particulier et unique permettant la catalyse d'une rupture et la jonction de l'ADN simple brin (figure 25).

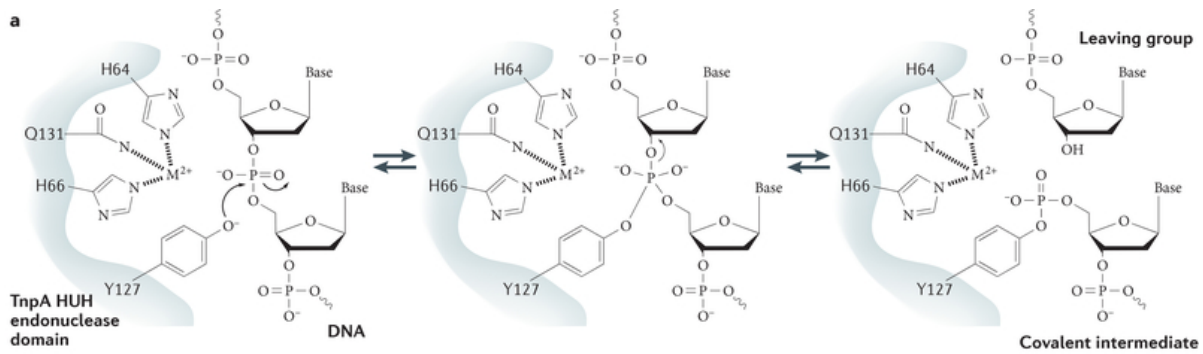


Figure 25 : Mécanisme de réaction des protéines HUH

Issue de Chandler et al., 2013

Le domaine HUH associé au ligand de l'ion métallique divalent ( $M^{2+}$ ), coupe l'ADN pour former un intermédiaire covalent et libérer le groupe partant clivé. Pour plus de description voir le texte.

La ou les tyrosine(s) active(s) du site catalytique donne(nt) lieu à la formation d'un intermédiaire  $5'$ phosphotyrosine et d'un  $3'$ OH au niveau du site de clivage. Bien que le  $3'$ OH libre pourrait être utilisé dans différentes fonctions, la fonction la plus probable serait d'amorcer la réplication comme cela est observé pour les protéines HUH : les relaxases conjuguatives, les plasmides à réplication CR et les protéines Rep de phages simple brin. Par ailleurs, ce  $3'$ OH pourrait avoir une fonction de nucléophile pour le transfert de brin, ce qui permettrait lors de la fin de la réplication CR de résoudre l'intermédiaire phosphotyrosine pour la transposition ou le transfert conjugatif.

De plus l'enzyme HUH présente une polarité de clivage inverse à celle des recombinases à tyrosine. Ceci lui permet de produire des intermédiaires  $3'$ phosphotyrosine générant des groupes  $5'$ OH libres (Grindley et al., 2006). Contrairement au  $3'$ OH, le  $5'$ OH ne peut pas être utilisé comme amorce de réplication. Par ailleurs, l'activité de l'enzyme HUH nécessite un ion métallique divalent qui permet de faciliter le clivage en localisant et en polarisant la liaison phosphodiester scissile, contrairement aux recombinases à tyrosine indépendantes qui ne nécessitent pas de cofacteurs (Ross et al., 2021). Il est supposé que  $Mg^{2+}$  ou  $Mn^{2+}$  sont les cofacteurs divalents physiologiques, bien que, dans des conditions *in vitro*, d'autres ions métalliques divalents peuvent être utilisés (Boer et al., 2009, 2006; Datta et al., 2003; Edwards et al., 2013; Larkin et al., 2005). Les histidines du motif HUH permettent de fournir deux des trois ligands qui sont nécessaires à la coordination des ions métalliques (Ross et al., 2021). Néanmoins, dans la superfamille HUH, le troisième ligand est variable mais il reste de nature polaire, à savoir : un glutamate, une asparagine, une histidine ou une glutamine (Ross et al., 2021).

Au sein de la superfamille HUH, l'ordre entre les motifs HUH et Y varie. En effet, chez les protéines du groupe Rep et les IS, le motif Y est situé en aval du motif HUH ; dans le groupe des relaxases, il est en amont (Ross et al., 2021). Cette permutation entre les deux motifs modifie la topologie du

domaine (Dyda and Hickman, 2003; Guasch et al., 2003; Koonin and Ilyina, 1993). Malgré ces changements, dans toute la superfamille HUH, la conformation tridimensionnelle des résidus du site actif est quasiment identique.

En plus des domaines HUH et Y, d'autres domaines conservés sont souvent présents et ont, la plupart du temps, une fonction inconnue. Parmi ces domaines conservés, certains domaines codent pour une hélicase ATP-dépendante, un motif de liaison au zinc ou une primase (Bruand and Ehrlich, 2000; Chang et al., 2002; Clérot and Bernardi, 2006; Kapitonov and Jurka, 2001; Odegrip et al., 2000). À titre d'illustration, l'activité du domaine hélicase ADN est localisée en C-terminal et peut générer des substrats d'ADN simple brin qui sont nécessaires aux enzymes HUH (Brister and Muzyczka, 1999; Clérot and Bernardi, 2006; Kapitonov and Jurka, 2001). Cependant, cette tâche peut être effectuée par le recrutement d'une hélicase hôte (Bruand and Ehrlich, 2000; Chang et al., 2002; Odegrip et al., 2000). L'action de l'activité de 3' vers 5' de l'hélicase facilite le déroulement de l'ADN à la fourche de réplication du brin matrice qui est nécessaire au processus de réplication CR.

Un autre rôle des nucléases HUH est de reconnaître, puis de se lier, à des structures tige-boucles présentant des sites de clivage qui peuvent être situés soit dans la structure même, soit à côté de celle-ci. L'importance de ces structures en tige-boucle a été démontrée dans de nombreux systèmes comme la réplication des plasmides, la conjugaison, les virus et la transposition (Boer et al., 2006; Brister and Muzyczka, 2000; Messing et al., 2012; Orozco and Hanley-Bowdoin, 1996; Ronning et al., 2005; Ton-Hoang et al., 2012, 2005). La formation de tiges boucles *in vivo* peut avoir lieu dans de nombreuses conditions physiologiques (Bikard et al., 2010).

Des analyses structurales ont démontré que les structures tige-boucle ont la possibilité d'être reconnues de différentes manières (Edwards et al., 2013; Guasch et al., 2003; Hickman et al., 2010b, 2004; Ronning et al., 2005) :

- Soit la séquence nucléotidique de la tige d'ADN double brin est reconnue pour sa séquence,
- Soit elle est reconnue pour les irrégularités de la structure de la tige,
- Soit la séquence nucléotidique de la tige boucle est reconnue de manière spécifique.

Dans plusieurs cas, l'ADN adjacent à la tige-boucle sous forme simple brin était important pour la reconnaissance de cette tige. En effet, il a été rapporté que les relaxases établissent un contact avec les bases situées sur le site de clivage et la structure tige-boucle (Edwards et al., 2013; Guasch et al., 2003; Larkin et al., 2003). Par ailleurs, les protéines TnpA<sub>REP</sub>, apparentées aux transposases IS200/IS605, ont démontré une importance cruciale des nucléotides du côté 5' de la tige-boucle pour la reconnaissance et la liaison spécifique de la séquence (Messing et al., 2012). Cependant, d'autres membres comme les virus adéno-associé (AAV) ou la protéine RepB du plasmide ont un mode de liaison plus complexe en raison de leur reconnaissance de deux régions distinctes (Hickman et al., 2004; Ruiz-Masó et al., 2007)



### III.3.3. Les IS à transposase HUH

#### III.3.3.1. La famille IS200/IS605

Les transposases de la famille IS200/IS605 comportent uniquement le domaine HUH et une seule tyrosine catalytique (Y1). En 1983, IS200 a été le premier membre de cette famille à être identifié chez *Salmonella Typhimurium* insérée dans l'opéron histidine (Edwards et al., 2013). Par la suite, en 1998, l'élément IS605 a été identifié chez *Helicobacter pylori* (Guasch et al., 2003). Certains éléments de la famille IS200 ont été identifiés à la fois chez les bactéries et chez les archées, suggérant un transfert horizontal de gènes entre les bactéries et les archées (J. Filée et al., 2007).

La famille IS200/IS605 est largement distribuée, décrite chez 45 genres bactériens différents et elle compte 153 membres distincts. Cette famille a été divisée en trois sous-groupes (figure 26) :

- Sous-groupe IS200 qui contient seulement le gène *tnpA* codant la transposase,
- Sous-groupe IS605 qui contient *tnpA* et le gène *tnpB* qui code une endonucléase guide),
- Sous-groupe IS1341 qui ne contient que le gène *tnpB*.

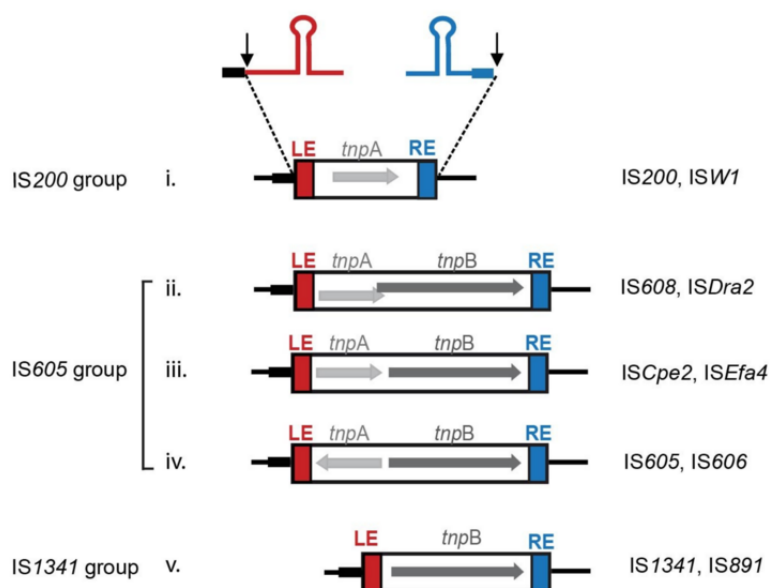


Figure 26 : Organisation génétique des membres de la famille IS200/IS605

Issue de S. He et al., 2015

Les extrémités gauche (LE) et droite (RE) portant le palindrome subterminal sont représentées par des cases rouges et bleues, respectivement. Les sites de clivage gauche et droit sont représentés par des cases noires et bleues respectivement. Les positions de clivage sont indiquées par de petites flèches verticales. Flèches grises : les gènes codant *tnpA* et *tnpB* ; (i) groupe IS200 avec *tnpA* seul, (ii) à (iv) groupe IS605 avec *tnpA* et *tnpB* dans différentes configurations et (v) groupe IS1341 avec *tnpB* seul.

### III.3.3.1.1. Le sous-groupe IS200

Les membres du sous-groupe IS200 sont distribués chez les bactéries à Gram positif et à Gram négatif, ainsi que des archées (Beuzón et al., 2004). Ces membres font partie des plus petites IS autonomes avec une taille comprise entre 600 et 700 pb. Malgré son abondance au sein des génomes du genre *Salmonella*, l'élément IS200 ne contribue que très faiblement à des mutations ; seulement quelques insertions ont été documentées (Beuzón et al., 2004).

Ceci peut être expliqué par le fait que le gène *tnpA<sub>IS200</sub>* est exprimé à partir d'un promoteur  $P_L$  faible qui est localisé dans l'extrémité gauche de l'IS, appelée Left-End (LE) . Les extrémités LE ou RE (Right End) forment des palindromes (Lam and Roth, 1983). Ces deux extrémités sont les sites de liaison de la transposase. Par ailleurs, l'extrémité LE comporte un palindrome subterminal supplémentaire par rapport à l'extrémité RE. Ces deux palindromes subterminaux dans l'extrémité LE jouent un rôle spécifique dans la régulation de l'expression du gène *tnpA<sub>IS200</sub>* (Beuzón et al., 1999). En effet, il a été démontré que le palindrome se situant le plus à l'extérieur chevauche le promoteur  $P_L$  de *tnpA<sub>IS200</sub>*, et agirait comme un terminateur de transcription bidirectionnelle en amont du gène. Cela a pour conséquence d'arrêter jusqu'à 80 % des transcrits. Le second palindrome réprimerait la traduction de l'ARNm en séquestrant le site de liaison au ribosome (RBS). Par ailleurs, il a été mis en évidence la présence d'un petit ARN anti-sens dans la région RE. Celui-ci a été identifié comme substrat de la protéine Hfq impliquée dans la régulation post-transcriptionnelle (Sittka et al., 2008). Hfq se lie directement à la partie 5' du transcrit *tnpA<sub>IS200</sub>* bloquant alors l'accès au RBS (Ellis et al., 2015a).

### III.3.3.1.2. Le sous-groupe IS605

Contrairement au sous-groupe IS200, les éléments du sous-groupe IS605 sont plus longs et ont une taille moyenne de 1700 kpb. Cela s'explique par la présence du gène *tnpB* en sus du gène *tnpA*. Les gènes *tnpA* et *tnpB* peuvent présenter des configurations différentes selon les éléments de la famille :

- Soit les gènes sont divergents,
- Soit les gènes sont orientés dans la même direction, avec *tnpA* en amont de *tnpB*,
- Soit les deux gènes se chevauchent.

Les éléments IS605, IS606 et IS608 ont été identifiés dans de nombreuses souches de *Helicobacter pylori* (Kersulyte et al., 2002, 1998).

Les deux membres les mieux caractérisés de cette famille sont IS608 et ISDra2 de *Deinococcus radiodurans*. Ces derniers possèdent des gènes *tnpA* et *tnpB* qui se chevauchent.

### III.3.3.1.3. Le sous-groupe IS1341

Contrairement aux deux autres sous-groupes, ce groupe présente uniquement le gène *tnpB*. Trois copies de l'élément IS1341 ont été identifiées dans la souche *Thermophilic bacterium PS3* (Murai et al., 1995). L'élément IS891 a été identifié dans une souche de cyanobactérie filamenteuse, en plusieurs copies sur le chromosome (Bancroft and Wolk, 1989).

Malgré l'identification de plusieurs éléments de ce sous-groupe, il n'existe aucune preuve expérimentale nous permettant d'affirmer si ces éléments sont ou pas actifs pour la transposition.

### III.3.3.1.4. Les transposases de la famille IS200/IS605

Les transposases de la famille IS200/IS605 ne comportent qu'un seul résidu tyrosine au niveau du site catalytique. En outre, l'élément IS608 utilise son résidu tyrosine Y127 pour réaliser une attaque nucléophile de la liaison phosphodiester aux sites de clivage.

Il faut rappeler que les transposases des éléments IS608, ISDra2 et ISC1474 ont été cristallisées (Barabas et al., 2008; Hickman et al., 2010b; Lee et al., 2006). Contrairement aux endonucléases HUH Y2 caractérisés qui agissent en monomérique, les endonucléases HUH Y1 forment des dimères obligatoires avec deux sites actifs (figure 27). En outre, la dimérisation entre deux monomères s'effectue en fusionnant leurs feuillets  $\beta$  en un grand feuillet  $\beta$  central, ce dernier est ainsi pris en sandwich entre des hélices  $\alpha$  (Ross et al., 2021). La configuration en *trans* est composée du motif HUH fourni par un monomère et du résidu tyrosine fourni par l'autre. *A contrario*, la configuration en *cis* est composée du motif HUH et du résidu tyrosine d'un même monomère. Il est supposé que la conformation en *trans* est active pendant le clivage, au cours duquel la tyrosine agit comme nucléophile, alors que la conformation en *cis* est supposée fonctionner pendant le transfert de brin, au cours duquel le  $3'OH$  est le nucléophile attaquant. Cependant, bien que des résultats expérimentaux soutiennent l'existence de la configuration en *cis*, seule la configuration en *trans* a été cristallisée (He et al., 2013; Hickman et al., 2010b; Ronning et al., 2005).

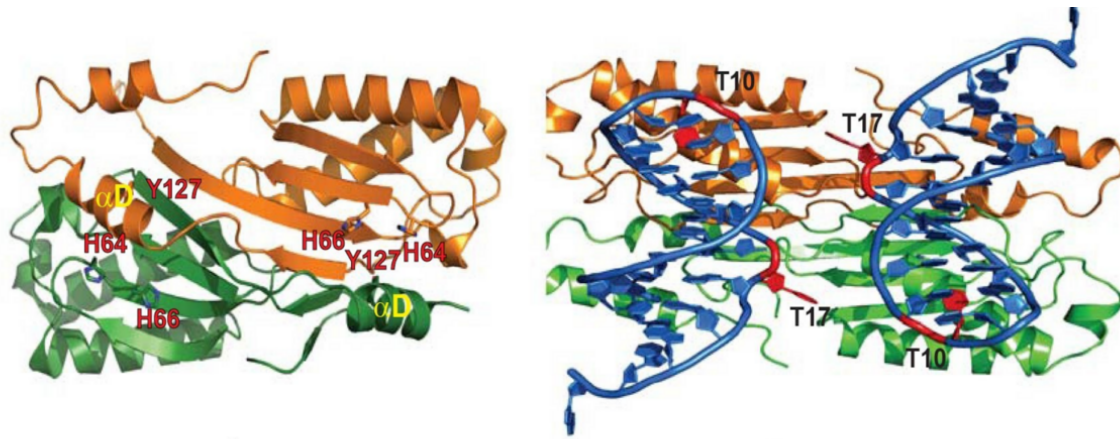


Figure 27 : Structure cristallographique de la transposase TnpA<sub>IS608</sub> en interaction avec l'extrémité RE

Issue de Ross et al., 2021

À gauche, la structure cristallographique de la TnpA<sub>IS608</sub> seule formant un dimère. Les deux monomères sont colorés en vert et en orange, respectivement. Les positions de l'hélice  $\alpha$ D et des résidus catalytiques sont indiquées. À droite, la structure de la TnpA<sub>IS608</sub> qui interagit avec l'extrémité RE représenté en bleu. Le T17 extra-hélique et le T situé dans la boucle en épingle à cheveux sont indiqués en rouge.

De plus, la comparaison des structures cristallisées des différents complexes protéine-ADN a démontré que la transposase pouvait être dans des états de configuration active (le résidu tyrosine est proche du motif HUH) et inactive (le résidu tyrosine est éloigné du motif HUH) (Barabas et al., 2008; Lee et al., 2006; Ronning et al., 2005; Ross et al., 2021).

Le gène *tnpA* codant la transposase réalise les étapes de clivage et de jonction. Le produit TnpB est codée uniquement par les sous-groupes IS1341 et IS605 et n'est pas nécessaire pour la transposition de l'IS608 ou de l'ISDra2 chez *Escherichia coli* et *Deinococcus radiodurans*, respectivement (Pasternak et al., 2010; Ton-Hoang et al., 2005). De plus, des membres de la famille IS200/IS605 portent des gènes *tnpB* qui ont subi ou subissent une altération de ce gène, cela suppose que TnpB ne joue pas un rôle essentiel (Pasternak et al., 2013).

### III.3.3.1.5. Rôle des extrémités RE et LE

Il a été mis en évidence que les extrémités IS608 et ISDra2 portent des palindromes subterminaux imparfaits formant des irrégularités. Ces irrégularités aident la transposase TnpA à distinguer les brins « supérieurs » et « inférieurs » (Hickman et al., 2010b; Ronning et al., 2005). De plus, la structure cristallographique de TnpA<sub>IS608</sub> a démontré que celle-ci se liait spécifiquement aux bases T10 dans la tige-boucle et T17 dans l'extrémité RE (He et al., 2011; Ronning et al., 2005). Des résultats similaires ont été obtenus avec la transposase TnpA<sub>ISDra2</sub> (Hickman et al., 2010b).

Les sites de clivage de l'IS608 conduisent à une insertion de l'IS dans une orientation spécifique avec son extrémité gauche 3' au tétranucléotide 5'-TTAC. *A contrario*, TnpA<sub>ISDra2</sub> reconnaît et s'insère en 3' d'un pentanucléotide 5'-TTGAT combinant des interactions ADN-ADN et protéine-ADN

(Hickman et al., 2010b). En effet, les quatre premiers nucléotides sont impliqués dans les interactions ADN-ADN et le cinquième nucléotide dans les interactions ADN-protéine (TnpA<sub>ISDra2</sub>).

### III.3.3.1.6. Modèle de transposition de la famille IS200/IS605

IS608 et ISDra2 sont très bien décrits d'un point de vue biochimique. Ces connaissances ont été acquises à l'aide d'un système *in vitro* pour l'élément IS608. Un tel système nécessite une transposase purifiée, des substrats d'ADN IS608 simple brin qui seront reconnus par la transposase et des ions métalliques divalents, tels que Mg<sup>2+</sup> ou Mn<sup>2+</sup> (Barabas et al., 2008; Guynet et al., 2008; Ton-Hoang et al., 2005). Des expériences similaires ont été réalisées avec ISDra2 (Hickman et al., 2010b; Pasternak et al., 2010; Ton-Hoang et al., 2010). Il a été montré que seul les substrats d'ADNss du brin « supérieur » sont reconnus alors que ceux du brin « inférieur » sont réfractaires (Guynet et al., 2008; Ton-Hoang et al., 2005).

Le clivage par la transposase TnpA du brin supérieur au niveau des sites de clivage gauche et droit LE et RE conduit à une excision sous la forme d'un intermédiaire circulaire d'ADNss (figure 28). Le clivage aux extrémités permet donc une liaison covalente 5'phosphotyrosine avec l'IS et produit un groupement 3'OH libre de l'ADN donneur à l'extrémité LE. *A contrario*, le clivage du côté RE libère un groupement 5'phosphotyrosine au niveau de l'ADN donneur et un groupement 3'OH libre à l'extrémité RE de l'IS. Ainsi, le groupement 3'OH libre de l'extrémité RE agit comme nucléophile pour attaquer la liaison 5'phosphotyrosine de l'extrémité LE. Cette attaque nucléophile permet de refermer l'IS, ce qui permet *in fine* la formation d'un intermédiaire circulaire ADNss. Dans cet intermédiaire, les deux extrémités (gauche et droite) sont accolées et forment donc une jointure du transposon. Il est important de noter que l'excision de l'IS est accompagnée de la formation d'une jonction donneuse conservant la séquence cible sur le brin d'ADN donneur.

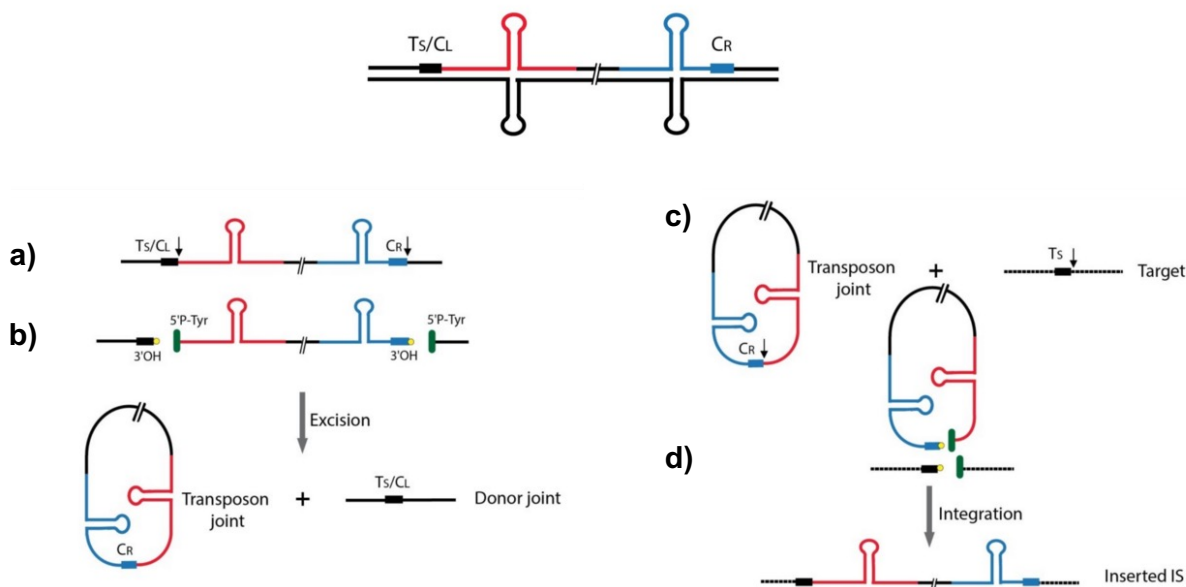


Figure 28 : Modèle de transposition de l'élément IS608

Adaptée de S. He et al., 2015

En haut est présentée l'organisation de IS608 où les extrémités gauche (LE) et droite (RE) avec un palindrome subterminal sont en rouge et en bleu, les sites de clivage gauche ( $T_S / C_L$ ) et droite ( $C_R$ ) sont représentés par des boîtes noires et bleues, respectivement.

a) Les structures du brin supérieur sont reconnues et clivées par la TnpA (flèches verticales).

b) Lors du clivage, une liaison 5' phosphotyrosine (vert) est formée avec LE, et avec le flanc 3' de RE et un 3'-OH (cercle jaune) est formé au niveau du flanc gauche et de RE. L'excision du cercle simple brin avec LE et RE accolés est accompagnée de la formation d'une jonction donneuse conservant la séquence cible.

c) Le cercle simple brin avec la jonction de transposon et ADN cible (noir) avec le site cible. La transposase catalyse le clivage de la jonction du transposon et de la cible simple brin pour l'intégration de l'élément.

Pour l'intégration de l'intermédiaire circulaire, celui-ci est clivé au niveau de la jonction entre les deux extrémités RE/LE. La TnpA catalyse le clivage de la jonction de l'IS et de la cible simple brin. Enfin, l'élément est intégré dans une séquence cible spécifique ( $5^{\circ}$ CATT (IS608)). Il s'insère à l'extrémité gauche 3' du site cible. Il convient de préciser que cette spécificité de la cible est une autre caractéristique inhabituelle de la transposition IS200/IS605.

### III.3.3.1.7. Lien entre transposition et fourche de réplication

Il existe un lien étroit entre la transposition et la fourche de réplication. Cela se traduit par une préférence pour une cible d'ADNss sur la matrice du brin retardé. En effet, il a été démontré que la transposition des éléments IS903, IS10, IS50 était modulée par la fourche de réplication (Dodson and Berg, 1989; Hu and Derbyshire, 1998; Roberts et al., 1985). D'une part, il a été montré que l'excision de l'IS608 est sensible au sens de réplication de l'élément. En effet, elle est plus fréquente lorsque le brin actif se trouve sur le brin retardé (figure 29) (Ton-Hoang et al., 2010). La longueur de l'ADNss du brin retardé dépend de la synthèse des fragments d'Okazaki par la primase DnaG (Wu et al., 1992; Zechner et al., 1992). L'inactivation transitoire de l'activité de la DnaG augmente la

longueur moyenne de l'ADNss entre les fragments d'Okazaki, permettant ainsi d'augmenter la fréquence d'excision de l'IS608. De plus, les fourches de réplication bloquées semblent être des cibles préférentielles pour l'insertion de l'IS608 (Ton-Hoang et al., 2010).

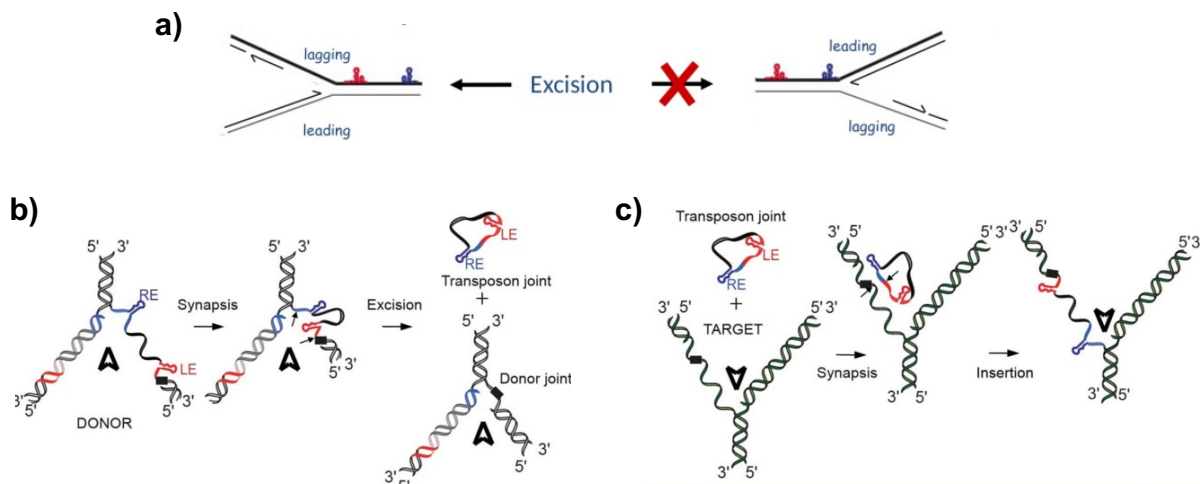


Figure 29 : Orientation de l'insertion et l'excision de IS608 par rapport à la direction de la fourche de réplication

Adaptée de Ross et al., 2021

- a) Disposition du brin actif IS608 (en haut) par rapport à la direction de la fourche de réplication. Les fragments d'Okazaki sur le brin retardé sont indiqués par de courtes lignes. Le sens de la synthèse de l'ADN est indiqué par des demi-flèches.
- b) Excision de l'intermédiaire circulaire simple brin à partir de la matrice de brin retardé d'un plasmide donneur. La pointe de la flèche indique la direction de la réplication.
- c) Intégration de l'extrémité droite (RE) puis de l'extrémité gauche (LE) dans la cible simple brin à la fourche de réplication.

### III.3.3.2. Les familles IS91/ISCR

Les transposases de la famille IS91 ont été les premières à être identifiées avec un domaine HUH Y2 (Garcillán-Barcia et al., 2002). Elles comprennent aussi un motif N-terminal de liaison au zinc, ainsi que des domaines supplémentaires dont la fonction n'a pas encore été identifiée.

Un groupe d'éléments apparentés, les ISCRs, souvent associés à divers gènes de résistance aux antibiotiques, portent une *orf* (nommée CR ou région commune) ressemblant aux transposases de la famille IS91 mais possédant une seule tyrosine (Toleman et al., 2006a).

Les éléments des familles IS91/ISCR seront abordés et détaillés lors du chapitre IV.

### III.4. Éléments transposables eucaryotes

Il existe des similarités évidentes entre les éléments transposables procaryotes et eucaryotes, telles que (Hickman et al., 2010a; Hickman and Dyda, 2015; Montañó and Rice, 2011) :

- La biochimie de base de la transposition est identique.
- Ils ont pour certains des organisations et une taille similaire
- La plupart n'est pas porteur de gènes « passagers ».

Néanmoins, la différence principale entre les éléments transposables procaryotes et eucaryotes réside dans la régulation de l'expression et de l'activité de la transposase (Nagy and Chandler, 2004). De plus, la contrainte majeure des éléments transposables eucaryotes est imposée par le noyau qui sépare physiquement le processus de transposition de celui de l'expression de la transposase. *A contrario*, l'absence de noyau chez les éléments transposables procaryotes permet que la transcription et la traduction soient couplées.

Il est supposé que la plupart des grandes familles d'éléments transposables soit apparue avant ou peu de temps après l'émergence des eucaryotes. En effet, ces éléments sont dans la majorité des cas présents dans les génomes de plusieurs supergroupes eucaryotes (Wells and Feschotte, 2020).

La disponibilité croissante de génomes entiers et assemblés conduit à la découverte de nouveaux groupes d'éléments transposables eucaryotes, dont trois grands groupes et un nouveau groupe émergent ont été décrits :

- Les hélitrons (Kapitonov and Jurka, 2001),
- Les Cryptons codant pour une recombinase de type tyrosine (Goodwin et al., 2003),
- Les Polintons/Mavericks autosynthétisés (Kapitonov and Jurka, 2006; Pritham et al., 2007),
- Les Réplitrans (Craig, 2023).

#### **III.4.1. Les Hélitrons avec des transposases HUH Y2**

Les hélitrons constituent le premier groupe d'éléments transposables eucaryotes à avoir été découverts par analyse computationnelle de séquences de génomes entiers (Kapitonov and Jurka, 2001). Les hélitrons nommés Aie, AthE1, Atrep et Basho ont été décrits pour la première fois à partir du génome de *Arabidopsis thaliana* (Doutriaux et al., 1998; Kapitonov and Jurka, 1999). Des hélitrons ont été trouvés dans au moins une espèce de tous les principaux groupes d'eucaryotes, à l'exception des Rhizaria qui se composent d'organismes unicellulaires ressemblant à des amibes (Brown et al., 2012; Kapitonov and Jurka, 2007). Ces éléments s'observent y compris chez les plantes et le ver *Caenorhabditis elegans*, ainsi que certains mammifères tels que les chauves-souris (Kapitonov and Jurka, 2007, 2001). Les études génomiques ont révélé que l'activité des hélitrons pouvait contribuer à une fraction significative d'environ 2 % des génomes de plantes et d'invertébrés dans lesquels ils ont été découverts (Kapitonov and Jurka, 2001).

Les transposases « RepHel » des hélitrons comportent un domaine HUH Y2 fusionné à une hélicase de la superfamille 1B et elles présentent souvent des domaines accessoires (Thomas and Pritham, 2015). En effet, ces transposases sont proches des séquences d'insertion procaryotes, des éléments de la famille IS91 qui présentent un domaine HUH Y2 ou encore des éléments de la famille ISCR qui présentent un domaine HUH Y1.



Cependant, aucun élément héliatron n'a été démontré comme actif pour la mobilité. C'est pourquoi, sur la base d'une séquence consensus, une copie fonctionnelle a été reconstituée et appelée Helraiser (figure 30) (Grabundzija et al., 2016). Cette transposase présente un domaine N-terminal large, un motif HUH, deux résidus tyrosines formant le site catalytique et en C-terminal, un domaine hélicase. La séquence codante de la transposase est délimitée par des séquences terminales gauche et droite (LTS et RTS) se terminant par des sites  $5^{\prime}\text{TC}/\text{CTAG}^3^{\prime}$  caractéristiques de la famille Helibat1 des héliatrons (Pritham and Feschotte, 2007). Par ailleurs, l'extrémité RTS présente un palindrome subterminal de 19 pb localisé en amont du site de clivage (Grabundzija et al., 2016). Précisons qu'il a été démontré que Helraiser était capable de former des intermédiaires circulaires fermés comme pour les éléments IS91 (del Pilar Garcillán-Barcia et al., 2001). Ces intermédiaires présentaient tous une jonction covalente du dinucléotide  $5^{\prime}\text{TC}$  de l'extrémité LTS et du tétranucléotide  $\text{CTAG}^3^{\prime}$  de l'extrémité RTS. Par ailleurs, il a été démontré que cet intermédiaire était inséré dans le dinucléotide  $5^{\prime}\text{AT}$ , ce qui est une caractéristique des héliatrons.

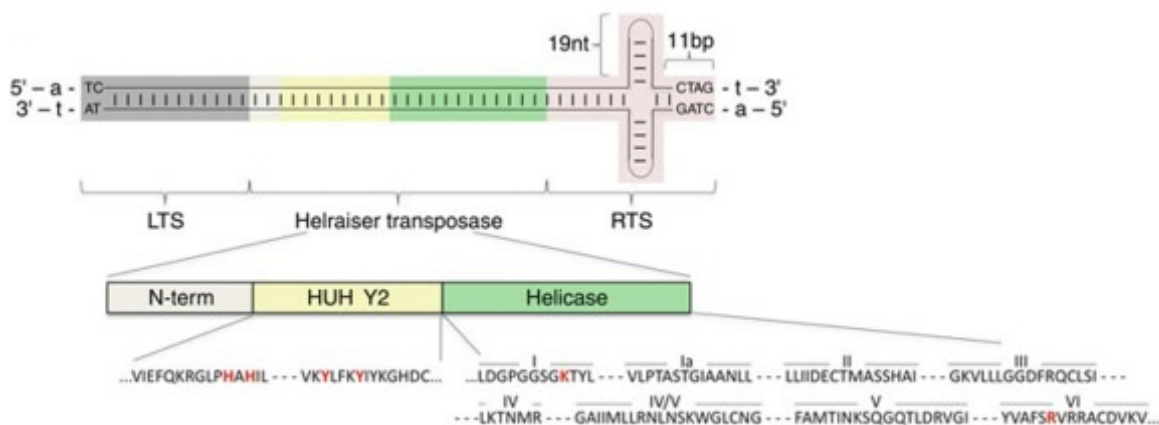


Figure 30 : Caractéristiques du transposon Helraiser

Issue de Grabundzija et al., 2016

*Les séquences terminales LTS et RTS sont en majuscules, et les séquences flanquantes du site cible de l'hôte sont en minuscules. Les rectangles blancs, jaunes et verts représentent respectivement les domaines N-terminal, endonucléase et hélicase de la transposase Helraiser. Les motifs d'acides aminés conservés sont indiqués ci-dessous. Les positions des résidus utilisés pour la mutagenèse sont indiquées en rouge.*

### III.4.2. Les réplitrons avec des transposases HUH Y1

C'est en 2023 que les réplitrons ont été décrits pour la première fois. Ils forment un groupe important de transposases eucaryotes et ils codent une endonucléase HUH Y1. Les réplitrons sont présents dans les génomes d'algues vertes et de plantes, de diverses stramenopiles, y compris les algues brunes. Cette présence suggère une origine ancienne et une large distribution eucaryote (Craig, 2023).

Le Replitrón-1 de l'algue verte unicellulaire *Chlamydomonas reinhardtii* présente une spécificité de site cible, ainsi qu'une augmentation du nombre de copies lors de la transposition. Il convient de souligner qu'il présente tout de même des signatures d'insertion inhabituelles qui pourraient être non canoniques pour les réplitrons en général (Craig, 2023).

Les trois motifs des endonucléases HUH sont clairement définis par leurs résidus hautement conservés chez les réplitrons (figure 31) avec des caractéristiques propres indiquant que les réplitrons sont éloignés de la plupart des groupes de transposase eucaryotes et procaryotes déjà décrits mais plutôt proches des CRESS-DNA virus.

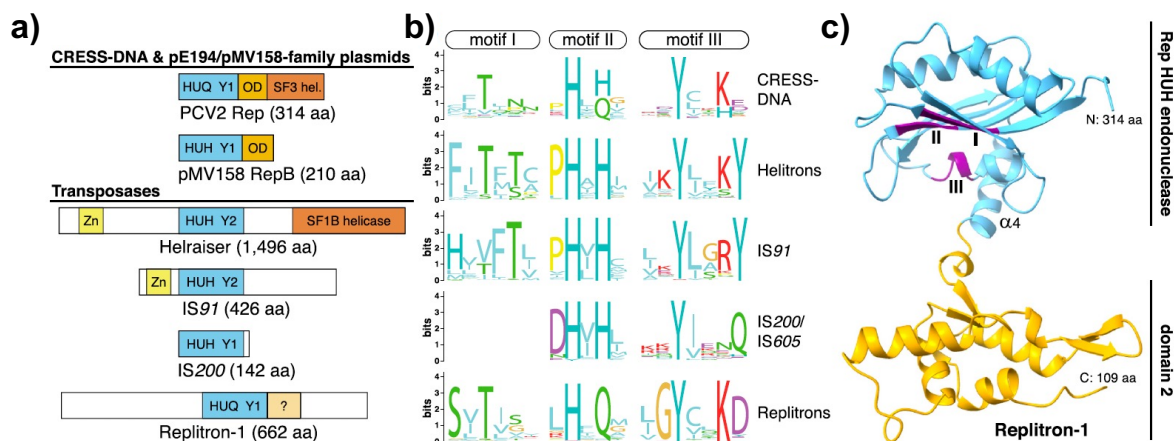


Figure 31 : Les endonucléases HUH et les réplitrons

Issue de Craig, 2023

a) Les domaines HUH Y sont représentés par une boîte bleue. Les domaines hélicases fusionnés sont représentés par des boîtes vermillon, les domaines d'oligomérisation par des boîtes orange, et les domaines putatifs de type doigt de zinc par des boîtes jaunes.

b) Logos de séquences pour les motifs conservés de certaines endonucléases HUH.

c) Structure tertiaire prédictive par AlphaFold2 du domaine Rep et du deuxième domaine de la transposase Replitrón-1. Les motifs I, II et III sont surlignés en violet.

Il a été montré que les hélitrons s'insèrent dans une orientation de 5' vers 3' entre les dinucléotides hôtes AT ou TT, tandis que les éléments procaryotes IS91 s'insèrent dans une orientation de 5' vers un site cible tétranucléotidique (Garcillán-Barcia et al., 2002; Thomas and Pritham, 2015). Les réplitrons suivent probablement un schéma similaire, potentiellement avec une insertion 3' vers un dinucléotide cible, bien que cela ne soit pas encore confirmé (Craig, 2023).

Ainsi, les propriétés structurelles et mécanistiques des réplitrons démontrent des variations sur des thèmes communs présents parmi les endonucléases HUH. Une caractéristique remarquable des réplitrons est leurs courtes répétitions terminales directes, non observées chez d'autres endonucléases HUH (Craig, 2023). Ces points de rupture coïncident spécifiquement avec les extrémités droites du réplitron et il est supposé que la transposition ait pu provoquer des ruptures double brin, contrairement à l'activité attendue d'une endonucléase HUH qui effectue des ruptures

simple brin. L'explication envisagée serait la présence de gènes Fanzor, exprimés chez *Chlamydomonas reinhardtii*, comme cela a été démontré chez les hélitrons (Bao and Jurka, 2013). En effet, les Fanzors sont des homologues eucaryotes des protéines TnpB codées par certains éléments procaryotes de la famille IS200/IS605 et qui ne sont pas essentiels à la transposition (Pasternak et al., 2013). Il a été démontré que les protéines TnpB sont des endonucléases guidées par l'ARN et elles clivent l'ADN (Altae-Tran et al., 2021; Karvelis et al., 2021). Il a été suggéré qu'il s'agit probablement des ancêtres évolutifs des protéines Cas12 des systèmes CRISPR-Cas de classe 2. C'est pourquoi, les Fanzors pourraient avoir une activité de clivage de l'ADN, ce qui suggère qu'ils peuvent fonctionner comme des helppeur facilitant la transposition (Kojima, 2020). Il est possible que les Fanzors de *Chlamydomonas reinhardtii* contribuent à la transposition par un mécanisme encore inconnu et par lequel les deux brins sont clivés par l'action combinée de l'endonucléase HUH et des Fanzors. Cela produirait des duplications non canoniques sur le site cible lors de l'intégration.

## Chapitre IV. Les séquences d'insertions bactériennes atypiques de la famille IS91/ISCR

---

Actuellement, IS91 est le prototype de la famille IS91 et d'une famille apparentée, les ISCR, ces deux familles sont adjacentes à des déterminants de virulence et à des gènes de résistance aux antibiotiques (Burland et al., 1998; Garcillán-Barcia and de la Cruz, 2002; Schlör et al., 2000; Toleman et al., 2006a).

### IV.1. Les éléments de la famille IS91

#### IV.1.1. Généralités

L'élément IS91, paradigme de la famille IS91 et des éléments ISCR apparentés, fut identifié en 1984 (Garcillán-Barcia et al., 2002; Zabala et al., 1982)

Par la suite, IS91 a été caractérisé comme une séquence d'insertion ayant la capacité de transposer dans une variété d'autres plasmides comme pACYC184, R388 et pBR322 (Zabala et al., 1982). Dans la majorité des cas, IS91 était présente en une copie unique, tandis que dans une petite minorité de cas, elle se présente sous la forme de co-intégrations apparentes (Diaz-Aroca et al., 1984). En effet, l'élément IS91 s'est révélé être une IS fonctionnelle qui transpose à des fréquences relativement élevées dans différents plasmides. De plus, plusieurs IS apparentées à l'élément IS91 ont été identifiées et deux d'entre elles sont actives dans la transposition : les éléments IS801 et IS1294 provenant d'un plasmide de *Escherichia coli* multi-résistant aux antibiotiques et d'un plasmide de *Pseudomonas syringae*, respectivement (Romantschuk et al., 1991; Tavakoli et al., 2000).

Une étude des bases de données en utilisant la séquences de la transposase de IS91 a démontré la présence de transposases très similaires dans plusieurs genres bactériens (*Bergeyella*, *Fusobacterium*, *Rhizobium*, *Shewanella*, *Pseudomonas*, *Klebsiella*, *Shigella* et *Escherichia*) (Garcillán-Barcia et al., 2002). De plus, des éléments apparentés ont été détectés chez *Mesorhizobium*, *Pseudomonas*, *Vibrio* et *Salmonella*.

L'élément IS91 a été initialement caractérisé dans des plasmides codant pour l' $\alpha$ -hémolysine (Hly) chez *Escherichia coli* (Zabala et al., 1982). Ainsi, IS91 est généralement associé à divers îlots de pathogénicité plasmidique et chromosomique qui hébergent l'opéron  $\alpha$ -hémolysine (Garcillán-Barcia et al., 2002; Garcillán-Barcia and de la Cruz, 2002; Hacker et al., 1983). De plus, il a été suggéré que cet élément était impliqué dans la dissémination d'autres déterminants de pathogénicité (Knapp et al., 1984). En effet, des éléments de la famille IS91 ont été localisés à côté de plusieurs autres déterminants de virulence dans des souches entéro-pathogènes, entérohémolytiques et entérotoxigènes d'*Escherichia coli* (Burland et al., 1998; Wolf et al., 1997).

Il a aussi été démontré que les éléments IS91-like ont été fréquemment observés à proximité de gènes de dégradation des xénobiotiques et qu'ils sont impliqués dans leur mobilité (Garcillán-Barcia and de la Cruz, 2002). En effet, il a été suggéré que les éléments de la famille IS91 pourraient être des points chauds pour l'insertion d'autres éléments transposables. L'analyse de la majorité des isoformes de la famille IS91 met en évidence que les éléments IS91-like n'apparaissent pas comme des séquences d'insertion classiques, à savoir comme des répétitions directes flanquant des cadres de lecture ouverts adjacents et formant des transposons composites (Garcillán-Barcia and de la Cruz, 2002; Mahillon and Chandler, 1998).

#### IV.1.2. Organisation génétique

Les IS atypiques ne possèdent pas de séquences inversées répétées aux extrémités contrairement aux IS classiques, mais possèdent des séquences spécifiques appelées structures palindromiques subterminales appelées IR<sub>CL</sub> et IR<sub>CR</sub> qui ne possèdent pas la même propriété fonctionnelle (Garcillán-Barcia et al., 2002; Ross et al., 2021).

##### IV.1.2.1. L'extrémité *oriS*

L'*oriS*<sub>91</sub> ou *ori91* est constituée d'une séquence terminale conservée de 23 nucléotides, riche en A-T avec une répétition inversée de 9 pb et d'une seconde répétition inversée, plus longue et riche en A-T appelée IR<sub>T</sub>. La structure palindromique subterminale du côté *oriS* rappelle sans nul doute les origines de répliation des phages (φX174) et des plasmides (pC194) (figure 33). Cette structure est aussi retrouvée chez certains éléments de la famille IS91, comme IS801 ; *a contrario* chez IS1294 ou IS1294*b*, aucune similitude n'a été constatée, ces dernières étant dépourvues de répétitions inversées présentant une région riche en A-T (figure 32) (Garcillán-Barcia et al., 2002; Yassine, 2015; Tavakoli et al., 2000).

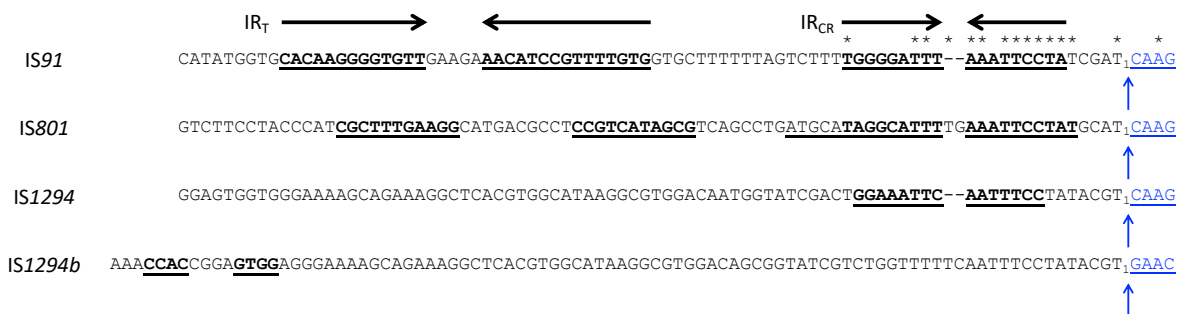


Figure 32 : Aligement des régions *oriS* des éléments IS91, IS801, IS1294 et IS1294*b*

D'après et inspiré de Garcillán-Barcia et al., 2002; Yassine, 2015

Aligement des régions *oriS* où les nucléotides du palindromisme subterminal de IR<sub>T</sub> et IR<sub>CR</sub> sont indiqués en gras et souligné (flèche noire). L'astérisque (\*) indique tous les nucléotides conservés entre les quatre éléments. En bleu, le tétranucléotide représentant le site cible d'insertion et la flèche bleue représentant le site de coupure.

Il a été démontré qu'au moins 82 pb de l'*ori91* sont nécessaires à la transposition de IS91 (Mendiola et al., 1994). Cette extrémité a été nommée *oriS*, car cette séquence serait à l'origine de la réplication par CR lors de la transposition de l'élément (Garcillán-Barcia et al., 2002).

Par ailleurs, les membres de la famille IS91 ne génèrent pas de duplication de site-cible lors de l'insertion, mais ils présentent une spécificité de séquence cible (Diaz-Aroca et al., 1987). En effet, IS91, IS801 et IS1294 s'insèrent à l'extrémité d'un tétranucléotide cible (Mendiola and de la Cruz, 1989; Richter et al., 1998; Tavakoli et al., 2000). Les éléments IS91 et IS801 s'insèrent majoritairement soit en <sup>5'</sup>CTTG, soit en <sup>5'</sup>GTTC, à des taux égaux (Mendiola et al., 1994; Richter et al., 1998), comme observé avec d'autres protéines de la famille HUH, comme celle du phage φX174 (Mendiola et al., 1994). Ces deux sites constituent 70 % des occurrences de la séquence cible. IS1294 peut s'insérer aux deux sites mais les études montrent une insertion préférentielle au site <sup>5'</sup>CTTG (Tavakoli et al., 2000). En revanche, l'élément IS1294*b* s'insère indifféremment aux deux sites <sup>5'</sup>CTTG et <sup>5'</sup>GTTC bien que cet élément ait été caractérisé avec le tétranucléotide cible <sup>5'</sup>GTCC dans la séquence originelle (Yassine, 2015). La présence de l'un de ces sites est essentielle pour la mobilité de l'élément et leur délétion entraîne une abolition de la transposition (Mendiola et al., 1994). Des analyses de mutations dans le site <sup>5'</sup>GTTC ont montré que les mutations des positions 1, 3 et 4 entraînent une diminution de la transposition supérieure à 90% alors qu'une mutation de la position 2 inhibe la mobilité de l'élément. De même, bien que les sites reconnus par IS801 peuvent présenter un degré de dégénérescence plus élevé car il reconnaît également les séquences cible <sup>5'</sup>GTCC ou <sup>5'</sup>CTCG, où le T en position 2 reste très conservé (Richter et al., 1998). Ces résultats suggèrent que le T en seconde position du site cible d'insertion est essentiel à l'insertion des éléments de la famille IS91. De plus, il a été démontré l'importance du nucléotide A après le site cible d'insertion (Yassine, 2015).

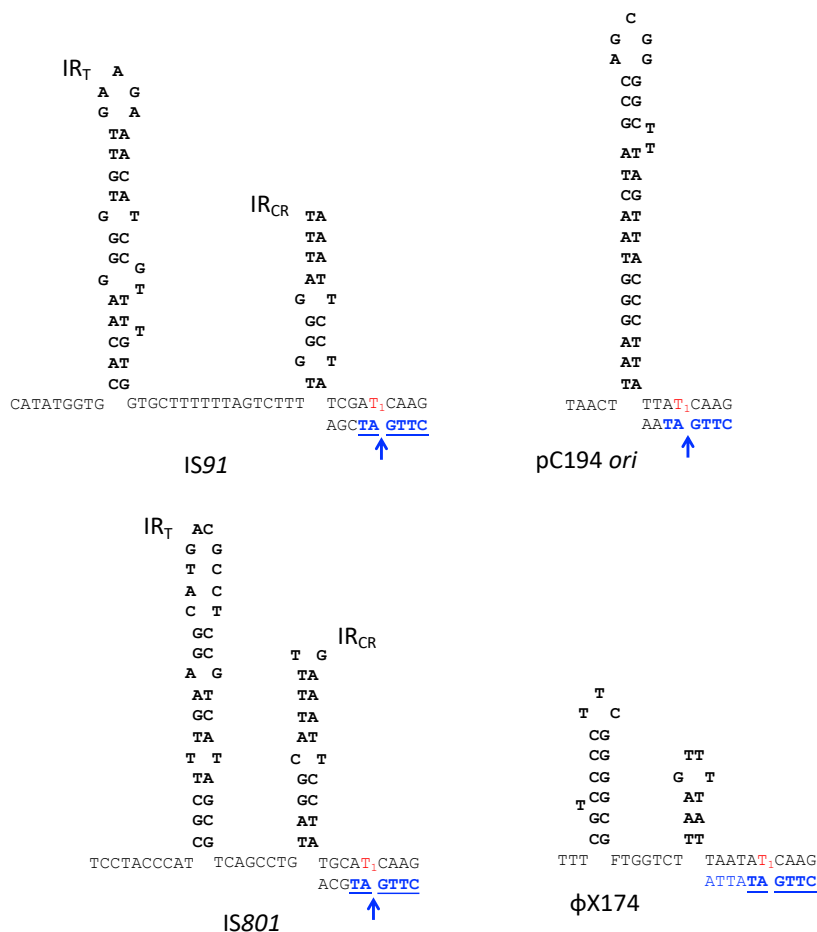


Figure 33 : Structures secondaires des régions *oriS* des éléments IS91, IS801, du phage  $\phi$ X174 et du plasmide pC194

D'après et inspiré de Garcillán-Barcia et al., 2002

*Le site de coupure est indiqué en bleu et l'endroit de coupure est symbolisé par une flèche.*

Il est important de souligner que l'extrémité *ori* de la famille IS91 est proche des *ori* de plasmides et de phages se répliquant par le mécanisme de cercle roulant (rolling circle, CR) (figure 33), suggérant que le mécanisme de transposition des éléments de la famille IS91 s'effectue par ce même mécanisme (Garcillán-Barcia et al., 2002; Mendiola et al., 1994).

#### IV.1.2.2. L'extrémité *terIS*

L'extrémité gauche de IS91 est appelée *ter91* ou *terIS<sub>91</sub>*. Le *terIS<sub>91</sub>* est constitué d'une séquence terminale conservée, IR<sub>CL</sub>, riche en GC de 17 pb avec une répétition inversée de 6 pb ; ce qui est aussi le cas pour les éléments IS801 et IS1294 (figure 34) (Garcillán-Barcia et al., 2002; Mendiola et al., 1994, 1992). Cette extrémité a été nommée *ter*, pour terminaison, la séquence marque donc la terminaison de la réplication par cercle roulant lors de la transposition de l'élément (Garcillán-Barcia et al., 2002).

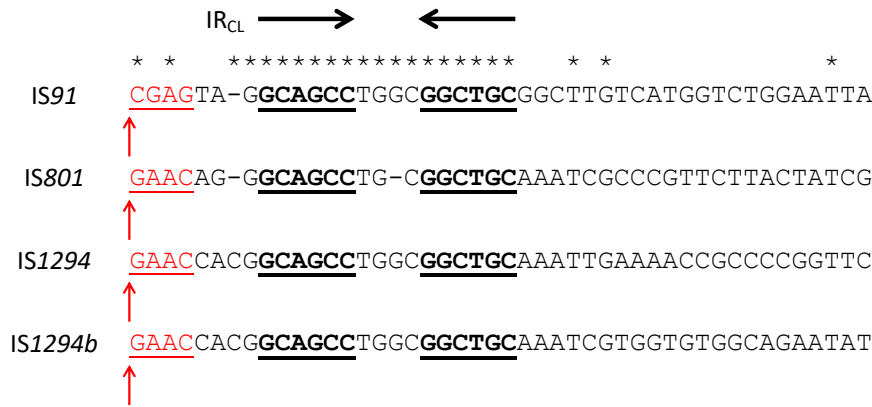


Figure 34 : Alignement des régions *terIS* des éléments IS91, IS801, IS1294 et IS1294b

D'après et inspiré de Garcillán-Barcia et al., 2002; Yassine, 2015

Alignement des régions *terIS* où les nucléotides du palindrome subterminal de IRCL sont indiqués en gras et souligné (flèche noire). L'astérisque (\*) indique tous les nucléotides conservés entre les quatre éléments. En rouge, le site de clivage et la flèche rouge représentent la localisation où la coupure s'effectue.

Cependant, le *terIS* peut ne pas être reconnu. Dans cette situation, cela entraîne la capture de fragments d'ADN adjacents à l'extrémité, il s'agit d'un processus connu sous le nom de One-ended transposition (OET). Pour les éléments IS91 et IS1294, cette terminaison échoue à une fréquence d'environ 1 à 10 % (Garcillán-Barcia et al., 2002; Tavakoli et al., 2000). Ainsi, l'extrémité *terIS* est présumée ne pas être complètement efficace, comme les terminateurs de transcription rho-indépendants, ce qui permettrait de mobiliser des séquences supplémentaires d'ADN adjacent à cette extrémité (figure 35). Ainsi, les éléments IS91-like peuvent jouer un rôle important dans la dissémination des gènes de virulence bactérienne ou de résistance aux antibiotiques (Garcillán-Barcia et al., 2002; Toleman et al., 2006a; Toleman and Walsh, 2011). De plus, il y a également des cas où la région *terIS* est supprimée (Garcillán-Barcia et al., 2002; Mendiola et al., 1994; Tavakoli et al., 2000), comme décrit chez *Pseudomonas vancouverensis* SMET-B et certaines souches de *Acinetobacter* (Castro-Gutierrez et al., 2022).



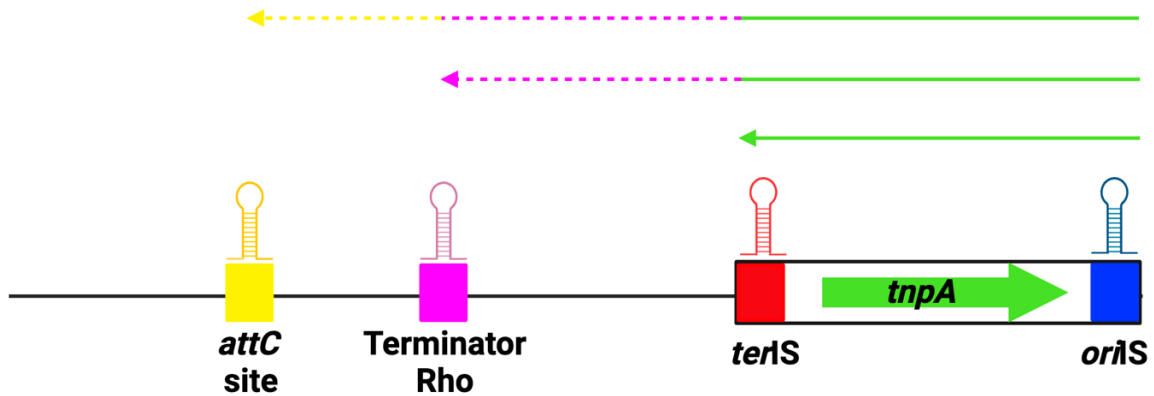


Figure 35 : Principe du mécanisme « one-ended transposition »  
Created with BioRender.com

Les flèches représentent l'orientation des gènes, en vert le gène *tnpA* et les palindromes subterminaux sont représentés par des rectangles rouge et bleu pour les extrémités *terIS* et *oriIS*, respectivement. Les terminateurs alternatifs sont représentés en rose comme le terminateur Rho ou en jaune comme un site *attC*. Quand le *terIS* (LE) est bien reconnu, seule l'IS est mobile (flèche verte), quand celui-ci n'est pas reconnu, la séquence d'ADN adjacente est emportée par la transposition jusqu'à un site *terIS* alternatif : soit un terminateur Rho (trait vert + flèche rose), soit un site *attC* (trait vert et rose + flèche jaune).

Mendiola et al. ont démontré que la séquence inverse répétée à l'extrémité *terIS*<sub>91</sub> n'était pas requise pour la transposition de l'élément. Cependant, ils ont aussi observé la formation de multimères en tandem de 2 à 4 copies du plasmide donneur par le mécanisme de OET (Mendiola et al., 1994).

La séquence *terIS* comprend un tétranucléotide de clivage. Ce site de clivage est différent entre les éléments *IS91* (5'CTCG), *IS801* et *IS1294* (5'GTTC) (Mendiola et al., 1994; Richter et al., 1998; Tavakoli et al., 2000). Il convient de souligner que le tétranucléotide de clivage 5'GTTC est également retrouvé chez l'élément *IS1294b* (Yassine, 2015). Le tétranucléotide 5'CTCG de l'élément *IS91* est essentiel pour sa reconnaissance car la mutation des trois premiers nucléotides a rendu le *terIS* inactif, cela conduit *in fine* à de multiples insertions en tandem (Bernales et al., 1999). Les travaux sur *IS1294b* montrent que les mutations du site de clivage de 5'GTTC vers 5'GTGC, 5'ATTC, 5'GCTC, 5'GCCC n'affecte aucunement l'efficacité de la transposition (Yassine, 2015). Toutefois, l'analyse de dix colonies avec un site de clivage 5'ATTC n'a pas permis de mettre en évidence la mise en œuvre de mécanisme de OET.

Ainsi, les extrémités *oriIS* et *terIS* sont des palindromes subterminaux qui formeraient des structures secondaires et serviraient de signaux de reconnaissance pour la liaison ADN-transposase (Garcillán-Barcia et al., 2002; Ross et al., 2021).

### IV.1.2.3. Contexte génétique

La structure génétique des trois membres de la famille IS91 actifs pour la transposition sont présentés en figure 36, à savoir : l'élément canonique IS91, IS801 et IS1294 / IS1294b (Diaz-Aroca et al., 1984; Yassine, 2015; Romantschuk et al., 1991; Tavakoli et al., 2000). Du côté *ori*S, IS801 et IS91 ont deux palindromes subterminaux, tandis que IS1294b n'a qu'un palindromique subterminal présumé. Du côté *ter*S, les 3 éléments ont un palindromique subterminal unique contenant le site de clivage, différent entre IS91 (5'CTCG) et les deux autres éléments (5'GTTC). Enfin, seul l'élément IS91 possède un gène supplémentaire, appelé *orf121*. Ce gène code probablement une protéine de 121 acides aminés qui présente la particularité suivante : le codon de terminaison de l'*orf121* chevauche le codon d'initiation du gène de la transposase, *tnpA* (5'TGATG). Ainsi ces deux gènes seraient couplés sur le plan de la traduction par le chevauchement d'un seul nucléotide.

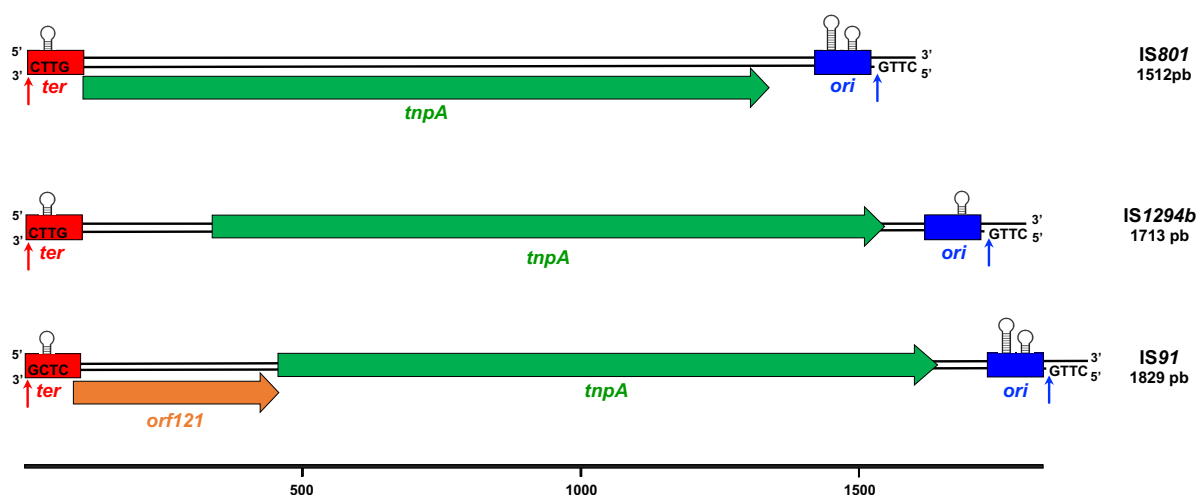


Figure 36 : Organisation génétique des membres actifs connus de la famille IS91

Les extrémités *ter* et *ori* portant les séquences palindromiques sont représentées par des cases rouges et bleues, respectivement. Les positions de clivage aux extrémités *ter* et *ori* sont indiquées par de petites flèches verticales rouges et bleues, respectivement. La séquence d'ADN du site de clivage de l'extrémité *ter* spécifique à l'élément et le site cible tétranucléotidique adjacent à *ori* sont également indiqués. Les séquences codantes *Orf121* et *tnpA* sont représentées par des flèches orange et vertes, respectivement.

Mendiola et al. ont mené des expériences de mutagenèse dans le gène *tnpA*, et ont montré que les éléments mutés étaient non transposables (Mendiola et al., 1992). En revanche, les mutations dans le gène *orf121* de IS91 n'affectaient aucunement la fréquence de transposition. Cependant, ils ont conclu que l'élément *Orf121* pourrait être impliqué dans la sélection de la cible dans la mesure car ils ont montré que certains mutants avaient une préférence marquée à 88 % pour une insertion en 5'CTTG et en 5'CTCG (Bernales et al., 1999; Garcillán-Barcia et al., 2002), alors que l'élément IS91 s'insère à des taux égaux en 5'CTTG ou 5'GTTC (Mendiola et al., 1994). Aucune *orf* similaire à *orf121* n'a pu être trouvée dans IS801 ou IS1294. Néanmoins, l'espacement entre le *ter*S<sub>1294b</sub> et le gène

*tnpA*<sub>IS1294b</sub>, laisse la possibilité d'avoir une *orf* en amont du gène *tnpA*<sub>IS1294b</sub> contrairement à l'élément IS801.

Sur le brin complémentaire, IS91 contient deux *orfs* potentiels, capables de coder pour des protéines de 108 acides aminés (*orf108*) et 106 acides aminés (*orf106*). Cette observation peut poser la question d'ARN anti-sens permettant la régulation de la transposition. Comme pour *orf121*, il n'y a pas d'équivalent dans d'autres membres de la famille IS91. Néanmoins, actuellement il n'existe aucune donnée suggérant que *orf108* et *orf106* sont exprimés, ni qu'elles jouent un rôle dans l'activité de transposition.

Chez d'autres éléments de la famille IS91, plusieurs comprennent un *orf* supplémentaire de fonction inconnue qui se situe en amont (ISCARN110, ISSde12, ISShvi3, ISSod25 et ISWz1) du gène *tnpA* codant la transposase (Ross et al., 2021).

#### **IV.1.3. Caractéristiques des transposases de la famille IS91**

Les séquences protéiques des transposases de la famille IS91, notamment les éléments IS91, IS801 et IS1294b, ont des motifs conservés avec la séquence du premier membre identifié de la superfamille HUH, à savoir la protéine A (gpA) du bactériophage ΦX174 et des protéines Rep des plasmides comme pLAB1000, pC194 ou pL1, mais également des hélitrons (figure 37) (Chandler et al., 2013; Garcillán-Barcia et al., 2002; Ross et al., 2021).

Domaine I		Domaine II	
Rep-pLAB1000	42-RVKECGEVLRFVA-----DDEGRRLRLYQTFCKSRLLCPLCNWR	99-RKTGRFLFLTLTAENA	
Rep-pL1	41-DVRGCGEVLRFVK-----IG-EHLKLYQTFCHKRLCPLCNWR	97-EPKGRFLFLTLTVKNA	
phageΦX174 (gpA)	137-RKLAERPLKDIR----DEYNYPKKKGKDECSRLLEASTMKS	195-GWFIVFDLTLTADDRL	
TnpA <sub>IS91</sub>	37-KIMACGTTLMGYTQWCCSSPDCCHTKKVCFRCKSRSCPHCGVK	96-PWQH--IVFTLPCQYW	
TnpA <sub>IS801</sub>	50-KMLACGTSILGVKHYTCGNHSCPHVKYLCNTCHCRACPSGK	109-PWQH--IVFTLPDTLW	
TnpA <sub>IS1294b</sub>	42-KMLACGTRILGVKEYNCDKPECPHVRYVTNSCGSRACPSGK	101-DWVH--IVFTLPDTLW	
Domaine III		Domaine IV	
Rep-pLAB1000	148-STEITINK-NGTYH <sup>Q</sup> H <sup>M</sup> H <sup>V</sup> LLLVFKPTYFK	222-SAQETAKY <sup>Q</sup> VKKS	
Rep-pL1	149-STEITVNEQDGSYNQ <sup>H</sup> L <sup>H</sup> VLLLVFKSSYFK	222-SAEETAKY <sup>E</sup> EVKS	
phageΦX174 (gpA)	252--VPEYGTANGRL <sup>H</sup> F <sup>H</sup> AV <sup>H</sup> FMRTLPTGSV	340-VGFY <sup>V</sup> VAKY <sup>V</sup> VNKK	
TnpA <sub>IS91</sub>	146-V-IHTWGR---DQQW <sup>H</sup> P <sup>H</sup> I <sup>H</sup> L-STTAGGV	246-VAVY <sup>F</sup> FGSY <sup>L</sup> LKKP	
TnpA <sub>IS801</sub>	159-A-LHTYGR---RLN <sup>W</sup> H <sup>P</sup> H <sup>V</sup> H <sup>L</sup> -SVTAGGL	259-TVNY <sup>L</sup> LGRY <sup>L</sup> LKKP	
TnpA <sub>IS1294b</sub>	151-A-IHTYGR---RLN <sup>W</sup> H <sup>P</sup> H <sup>V</sup> H <sup>V</sup> -SVTCGGL	251-TARY <sup>L</sup> LGRY <sup>L</sup> LKKP	
Domaine V		Domaine VI	
Rep-pLAB1000	255-LAGSRQISYGGFLFKEIR	309-QNY--FI-----	
Rep-pL1	255-LAGTRQISYGGLLKQIK	312-QNY--FI-----	
phageΦX174 (gpA)	381-LFRIRMSRNEFGMKMLTM	416-TPFNQILKQNAK	
TnpA <sub>IS91</sub>	307-GF--RMVRYGFLSPVK	356-DPLK <sup>C</sup> IL <sup>C</sup> GSQM	
TnpA <sub>IS801</sub>	320-HF--RMLRYFGFLANRV	370-DPFR <sup>C</sup> VL <sup>C</sup> GARM	
TnpA <sub>IS1294b</sub>	312-FF--KMVRYFGFLANRV	362-DPFE <sup>C</sup> VL <sup>C</sup> GGRM	

Figure 37 : Alignement des domaines conservés des éléments IS91, IS801 et IS1294b avec les protéines Rep et la protéine gpA

D'après et inspiré de Garcillán-Barcia et al., 2002; Ross et al., 2021

Entouré en rouge les acides aminés qui sont conservés par l'ensemble des membres alignés et entourés en vert, uniquement les acides aminés qui sont conservés par l'élément de la famille IS91. Dans le domaine III les histidines du motif HUH sont indiquées en rouge et dans le domaine IV les tyrosines du site catalytique sont également indiquées en rouge. Le motif putatif du doigt de zinc formé par les cystéines du domaine VI est indiqué en rouge.

**Le domaine I** comprend une région riche en cystéines avec 6 cystéines et 1 histidine conservées dans une région comprenant environ 40 acides aminés, organisées de la manière suivante : « C.X<sub>11</sub>.**C**.X<sub>4</sub>.**C**.X<sub>1</sub>.**H**. X<sub>7</sub>.C.X<sub>4</sub>.C.X<sub>2</sub>.C ». Les éléments soulignés et en gras sont absents des protéines Rep et de la gpA. Ce motif pourrait être un motif de liaison au zinc, mais aucune donnée n'a démontré cette hypothèse. Les données existantes sur le motif putatif de liaison au zinc se concentrent uniquement sur l'élément IS1294b (Yassine, 2015). Chez cet élément, des mutations des cystéines (C46, C58, C63, C73, C78, C81) et de l'histidine (H65) entraînent une perte de l'activité de transposition suggérant donc un rôle critique de ces acides aminés dans l'activité de la transposase de l'élément IS1294b, avec un rôle probable de cette région dans l'interaction avec l'ADN. D'ailleurs, les mutations des cystéines C22, C100 et C171 ne modifiaient pas l'activité de transposition. Cependant, la structure et l'alignement de la région ne présentent aucune similarité ni avec des motifs de doigt de zinc connus, ni avec d'autres domaines de liaison avec un ion métal

**Le domaine II** se caractérise par la conservation de 8 acides aminés : « W.X<sub>1</sub>.H.X<sub>1</sub>.V.F.T.L.P.X<sub>3</sub>.W ». Dans ce motif, les acides « T.L » sont également chez les protéines Rep et la gpA de ΦX174. Actuellement, il n'existe aucune donnée sur le rôle potentiel de chaque acide aminé de ce motif dans la transposition des éléments de la famille IS91 ou des protéines Rep. Cependant, par homologie avec les hélitrons, il a été suggéré que ce motif aurait une fonction dans la reconnaissance de l'ADN ou serait impliqué dans le positionnement des acides aminés du motif HUH pour la coordination de l'ion métallique (Tarasova and Khayat, 2021; Tompkins et al., 2021).

**Le domaine III** présente une constellation de trois histidines invariantes avec au total 8 acides aminés conservés agencés de la manière suivante : « R.X<sub>3</sub>.W.**H**.P.**H**.X<sub>1</sub>.**H**. X<sub>1</sub>.S.X<sub>1</sub>.T ». Les histidines du motif HUH sont indiquées en rouge . Ce motif est conservé pour la protéine Rep-pLAB1000 et la protéine gpA du phage ΦX174, alors que la protéine Rep-pL1 présente uniquement deux histidines. Ce motif démontre parfaitement que les transposases de famille IS91 appartiennent à la super famille des protéines HUH. En effet, dans cette super famille, le motif HUH sert à coordonner des ions métalliques divalents qui seront nécessaires pour la réaction de catalyse. Les transposases de cette famille catalysent et clivent l'ADN sous forme simple brin en favorisant la formation d'un intermédiaire covalent 5'-phosphotyrosine et des groupements 3'-OH libres au niveau du site de clivage (Chandler et al., 2013). Cependant, les endonucléases HUH peuvent aussi reconnaître et se lier à des structures d'ADN en tige boucle, (Chandler et al., 2013). Ce motif HUH aurait ainsi un rôle dans la reconnaissance des séquences palindromiques subterminales des éléments de la famille IS91. En outre, l'élément IS91 se caractérise par la conservation de trois résidus histidines en position H157, H159 et H161, dont les deux dernières formeraient le motif HUH. Cependant et à ce jour, aucune évaluation de l'impact de la mutation ponctuelle de l'une ou de l'autre de ces trois histidines n'a été effectuée pour l'élément IS91. Néanmoins, chez IS1294b, la mutation de l'histidine H164 en alanine du motif HUH « W.**H**<sub>162</sub>.P.**H**<sub>164</sub>.V.**H**<sub>166</sub>.V » est critique pour la transposition de l'élément IS1294b alors que la mutation de l'histidine H162 en alanine n'affecte pas la fréquence de transposition de l'élément IS1294b (Yassine, 2015). Aucune donnée n'est disponible concernant la mutation de l'histidine H166. De plus, il a été démontré que la mutation ponctuelle de l'histidine H153 en alanine et de l'arginine R157, situées à proximité du motif HUH de l'élément IS1294b, étaient critiques pour la transposition à des taux non détectables. Cela est très intéressant car ces deux résidus histidine et arginine sont également conservés chez les relaxases TrwC du plasmide R388 et Tral des plasmides de la famille F (Guasch et al., 2003; Ilangovan et al., 2017; Larkin et al., 2007). Ces relaxases sont des transférases de brins d'ADN qui catalysent les étapes initiales et finales du traitement de l'ADN pendant le transfert conjuguatif d'ADN de cellule à cellule (Clewell and Helinski, 1969; Lanka and Wilkins, 1995). Il a été montré que la mutation de ces deux histidines entraîne une diminution de la fréquence du transfert conjuguatif. Il a été montré que les deux résidus histidines du motif HUH permettent la coordination de l'ion métallique chez ces relaxases. Cependant, la coordination nécessite l'intervention d'un troisième résidu histidine situé à

proximité du motif HUH, à savoir l'histidine 150 pour TrwC et l'histidine 146 pour Tral. De plus, il a été suggéré que les arginines en position 154 pour TrwC et 150 pour Tral permettaient de stabiliser l'intermédiaire pentavalent coordonné. Ces données permettent de montrer que les trois résidus histidines génèrent la liaison à l'ion métallique en le chargeant positivement, ce qui conduit à une polarisation du phosphate scissile dont la conséquence est de le rendre plus sensible à l'attaque nucléophile médiée par les tyrosines catalytiques. L'ensemble de ces résultats suggère fortement que les résidus HUH de la famille IS91 sont impliqués dans la fixation d'un ion métallique divalent, intervenant en qualité de co-facteur de la réaction de transposition, comme décrit pour les relaxases TrwC et Tral (Guasch et al., 2003; Hickman et al., 2002; Larkin et al., 2007) et pour le Virus Adéno-Associé (AAV).

**Le domaine IV** contient une paire de tyrosines formant le centre catalytique de la transposase, soit « Y.X<sub>1</sub>.G. X<sub>1</sub>.Y.L.K.K.P ». Pour l'élément IS91, ces tyrosines sont en position 249 (Y249) et 253 (Y253). Les éléments soulignés et en gras indiquent les deux tyrosines conservées du site catalytique identique à la protéine gpA. En effet, Y249 et Y253 de la transposase de l'élément IS91 occupent des positions homologues aux résidus Y343 et Y347 du centre actif de la protéine gpA du phage ΦX174 (Mendiola and de la Cruz, 1992; van Mansfeld et al., 1986). Des mutations des tyrosines du site catalytique de la protéine gpA du phage ΦX174 ont eu pour conséquence directe la production de phages non viables (Hanai and Wang, 1993). De plus, il a été montré que les tyrosines Y343 et Y347 de la protéine gpA sont directement impliquées dans les réactions initiales de clivage et de transfert de brin dans la catalyse du substrat d'ADN. Chez IS91, il a été démontré que la mutation de l'une des deux ou des deux tyrosines Y249 et Y253 abolissait l'activité de transposition de l'élément IS91 suggérant que ces deux tyrosines ont un rôle crucial dans l'activité de la transposase de l'élément IS91 (del Pilar Garcillán-Barcia et al., 2001). Des résultats similaires ont été observés pour l'élément IS1294b : la mutation des tyrosines Y254 ou Y258 en phénylalanine avaient pour conséquence une abolition de la transposition de l'élément IS1294b (Yassine, 2015). Alors que la mutation des tyrosines Y236 ou Y240 en phénylalanine n'impacte pas l'activité de la transposase IS1294b. Ainsi, ces données indiquent clairement que les résidus Y249 et Y253 de l'élément IS91, ainsi que les résidus Y254 et Y258 de l'élément IS1294b, font partie du centre catalytique de la transposase par analogie avec la protéine gpA du phage ΦX174. C'est pourquoi, nous pouvons admettre que les sites catalytiques des protéines TnpA de la famille IS91 sont impliquées dans les réactions de clivage et de transfert de brins d'ADN lors de la transposition de l'élément. L'ensemble de ces résultats indique que les transposases de la famille IS91 seraient des analogues fonctionnels de la protéine gpA du phage ΦX174, permettant une répllication par un mécanisme de RC. Bien qu'un certain nombre d'acides aminés autour du site catalytique sont conservés, aucune donnée n'est publiée sur leur importance ou leur rôle dans l'activité de ce site.

**Le domaine V** se caractérise par la conservation de 8 acides aminés organisés de la manière suivante : « F.**R**.M.X<sub>1</sub>.R.**Y**.X<sub>1</sub>.**G**.F.L », les éléments soulignés et en gras sont présents également chez les protéines Rep et la protéine gpA. Actuellement, il n'existe aucune donnée sur le rôle potentiel de chaque acide aminé de ce motif dans la transposition des éléments de la famille IS91.

Pour terminer, **le domaine VI** une région conservée à l'extrémité C-terminale supplémentaire a été mise en évidence pour les éléments de la famille IS91 (figure 37). Celle-ci est absente de la protéine gpA du bactériophage ΦX174 et des protéines Rep des plasmides pLAB1000, pC194 ou pL1. Ce motif conservé se caractérise de la manière suivante : « D.P.X<sub>2</sub>.**C**.X<sub>1</sub>.L.**C**.G.X<sub>2</sub>.M ». Ce motif présente une organisation typique des doigts de zinc à savoir : « C.X<sub>2</sub>.C » qui pourrait être impliqué dans la liaison à l'ADN ou dans une interaction protéine-protéine. L'élément IS91 se caractérise par la conservation de deux résidus cystéines en position 360 et 363, dont aucune évaluation de l'impact de mutations ponctuelles n'a à ce jour été effectuée. L'extrémité C-terminale de la transposase de l'élément IS1294b comporte 4 cystéines. Les mutations de chacune d'entre elles ont montré uniquement l'importance des cystéines 366 et 369 (Yassine, 2015). En effet, ces deux cystéines sont parfaitement alignées sur les deux cystéines conservées de la famille IS91.

En résumé, les gènes *tnpA* des éléments de la famille IS91 codent une transposase de type HUH. Cette dernière est proche de la protéine A (gpA) du bactériophage ΦX174 et des protéines Rep de plasmides comme pLAB1000, pC194 ou pL1 (Chandler et al., 2013; Garcillán-Barcia et al., 2002; Ross et al., 2021). Les transposases de la famille IS91 possèdent une triade de résidus catalytiques HUH impliquée dans la fixation d'un ion métallique indispensable au mécanisme réactionnel (Chandler et al., 2013). L'alignement a démontré que les transposases de la famille IS91 possédaient deux tyrosines conservées formant le centre catalytique de la transposase, de façon similaire à la protéine gpA du phage ΦX174.

#### IV.1.4. Modèle de transposition de la famille IS91

Au fil des années, de nombreuses preuves ont été accumulées en faveur d'un mécanisme de RC pour les éléments de la famille IS91 :

- i) La structure palindromique subterminale du côté *ori*. Cette dernière rappelle indubitablement les origines de réplication des phages ( $\phi$ X174) et des plasmides de la famille pC194 (Garcillán-Barcia et al., 2002),
- ii) La spécificité du site cible d'insertion, <sup>5'</sup>CTTG (éléments IS91, IS1294 et IS801), qui rappelle le site de clivage des plasmides de la famille pC194 (<sup>5'</sup>CTTG) et des phages de type  $\Phi$ X174 (<sup>5'</sup>CTTG) (Garcillán-Barcia et al., 2002; Gruss and Ehrlich, 1989; Novick, 1989),
- iii) La reconnaissance bipolaire, c'est-à-dire l'orientation *ori*IS et *ter*IS, qui est une autre similitude avec les protéines initiatrices de réplication CR (Espinosa et al., 1995; Garcillán-Barcia et al., 2002),
- iv) La similarité entre les transposases de la famille IS91 et les endonucléases de la superfamille HUH (Chandler et al., 2013),
- v) L'observation d'intermédiaires circulaires simple brin et double brin (del Pilar Garcillán-Barcia et al., 2001).

Compte tenu de ces similarités et des données de la littérature, deux modèles de transposition ont été proposés pour la transposition CR des éléments de la famille des IS91.

##### IV.1.4.1. Modèle de transposition dit « concerté »

Le premier modèle, connu sous le nom de modèle de transposition dit « concerté » ou CR avec transfert de brin coordonné, a été développé par Mendiola et al. en 1994 (figure 38) (Mendiola et al., 1994). A partir de ce modèle, nous aurions la formation d'un dimère de la transposase dont le rôle serait de former un complexe initial permettant de regrouper les ADNs donneurs et les ADNs receveurs. Le premier monomère de la transposase permet la catalyse et le clivage d'un monocaténaire en <sup>3'</sup> du tétranucléotide cible <sup>5'</sup>GTTC et en <sup>5'</sup>*ori* de l'IS grâce à la tyrosine catalytique. Ce clivage provoque la libération d'un groupement <sup>3'</sup>OH sur le réplicon donneur et la formation d'un complexe 5'-phosphotyrosine se crée entre le monomère et la région <sup>5'</sup>*ori* de l'IS. Parallèlement, un clivage similaire est effectué par le second monomère dans l'ADN receveur. Ces événements vont créer un complexe transposase-ADN où le premier monomère sera lié de façon covalente au brin donneur, tandis que le second monomère sera complexé au brin receveur. Ensuite, le déplacement du complexe transposase-ADN vers le site cible s'opère. Une fois ce site atteint, le complexe 5'-phosphotyrosine induit une attaque de la cible <sup>3'</sup>OH de l'ADN receveur. Simultanément à ce phénomène, nous constatons une réplication de l'IS de l'ADN donneur qui est amorcée par la machinerie cellulaire. Il s'ensuit un déplacement progressif du brin clivé jusqu'à atteindre le site *ter* de l'IS, puis jusqu'au niveau du site de clivage comme le <sup>5'</sup>CTCG pour IS91. Ce nouveau clivage permet la génération d'un groupement <sup>3'</sup>OH au niveau *ter* de l'IS. Ce groupement <sup>5'</sup>*ori* libre au niveau



du *terIS* va subir une transestérification avec la liaison 5'-phosphotyrosine formée entre la transposase et l'ADN receveur cible. Cette réaction entraîne la formation d'une boucle de l'IS simple brin au niveau du site cible de l'ADN receveur, ce dernier étant ensuite résolu par la machinerie de réplication bactérienne.

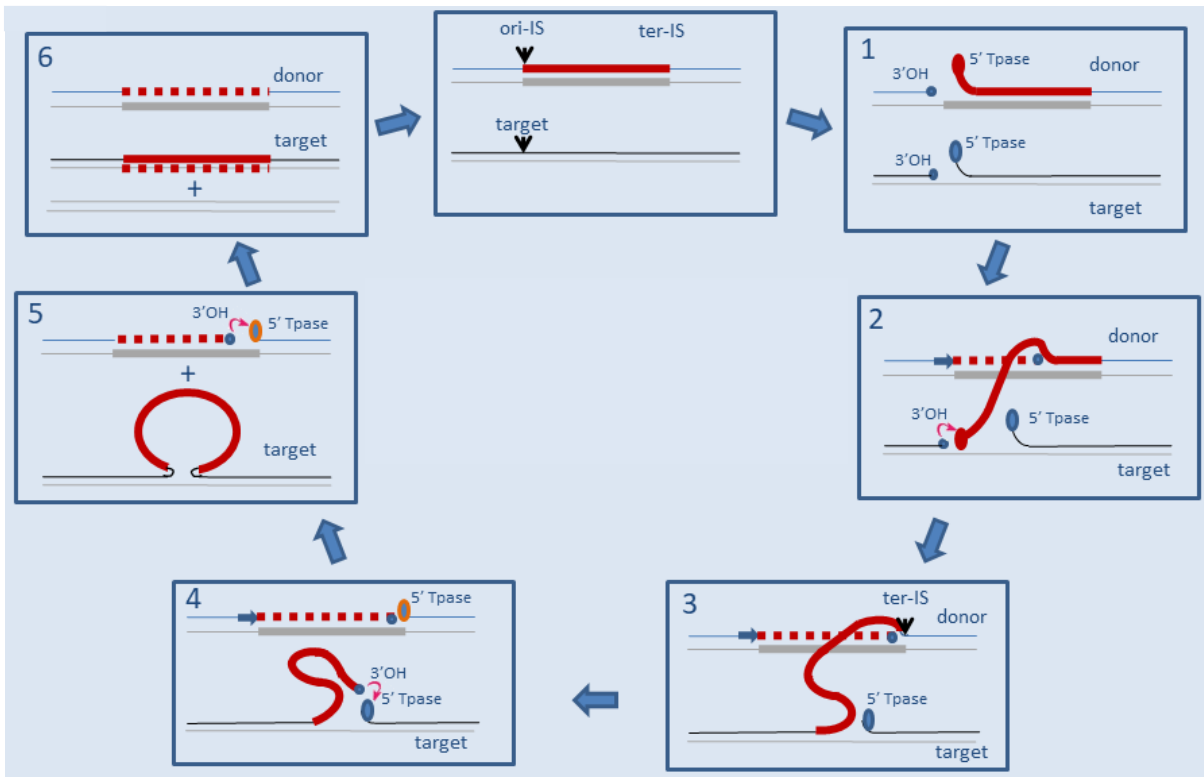


Figure 38 : Modèle de transposition de la famille IS91 dit « concerté »

Issue de Ross et al., 2021

*Le brin déplacé du transposon est en rouge ; le brin non déplacé est en gris ; l'ADN nouvellement répliqué, ligne rouge pointillée. Pour la description de la figure, voir texte.*

#### IV.1.4.2. Modèle de transposition dit « séquentiel »

Le second modèle, connu sous le nom de modèle de transposition dit « séquentiel » ou CR « peel and paste », a été proposé par Tavakoli et al. pour l'élément IS1294. Il a été décrit pour l'élément IS91 par De la Cruz et al. (figure 39) (Garcillán-Barcia et al., 2002; Tavakoli et al., 2000). Comme dans le modèle « concerté », la transposase catalyse le clivage d'un monocaténaire en 3' du tétranucléotide cible  $5^{\text{GTTTC}}$  et en  $5^{\text{ori}}$  de l'IS grâce à la première tyrosine de son site catalytique. Comme précédemment, cette coupure entraîne la libération d'un groupement  $5^{\text{ori}}$  sur le réplicon donneur et la formation d'un complexe 5'-phosphotyrosine entre la transposase et la région  $5^{\text{ori}}$  de l'IS. En parallèle, la libération du groupement  $3^{\text{OH}}$  permet d'initier la réplication du brin « leading » de l'ADN donneur par la machinerie de la réplication bactérienne. Cette réplication entraîne le

déplacement du brin clivé le long de l'ADN donneur jusqu'à atteindre le site *ter* de l'IS. Ensuite, après la reconnaissance du site de clivage <sup>5'</sup>CTCG de l'IS91, la première tyrosine clive de nouveau. Cette seconde coupure permet la libération d'un intermédiaire circulaire qui est refermé par une attaque nucléophile du groupement <sup>3'</sup>OH du brin néosynthétisé sur la liaison 5'-phosphotyrosine au niveau du *ter*IS. À ce stade, des intermédiaires circulaires simple brin et double brin ont été observés sans pour autant savoir si ces intermédiaires sont fonctionnels dans la transposition de l'élément IS91 (del Pilar Garcillán-Barcia et al., 2001). Ensuite, le transfert de cet intermédiaire est médié et initié par la transposase lors de la reconnaissance, dès lors le clivage de l'ADN receveur se produit au niveau de l'extrémité <sup>3'</sup> du site cible d'insertion comme <sup>5'</sup>GTTC. Ce clivage permet la formation de 2 extrémités avec des groupements <sup>3'</sup>OH sur l'intermédiaire circulaire clivé et l'ADN cible. Ce groupement <sup>3'</sup>OH est capable d'attaquer la liaison 5'-phosphotyrosine sur le brin d'ADN opposé permettant l'insertion de l'IS sur le brin d'ADN receveur. Cette intégration de l'IS est finalisée par la machinerie de la réplication bactérienne. En effet, ce modèle « séquentiel » implique l'excision d'un seul brin de l'IS formant un intermédiaire circulaire ce qui est similaire au modèle qui a été proposé pour la réplication du bactériophage  $\phi$ X174.

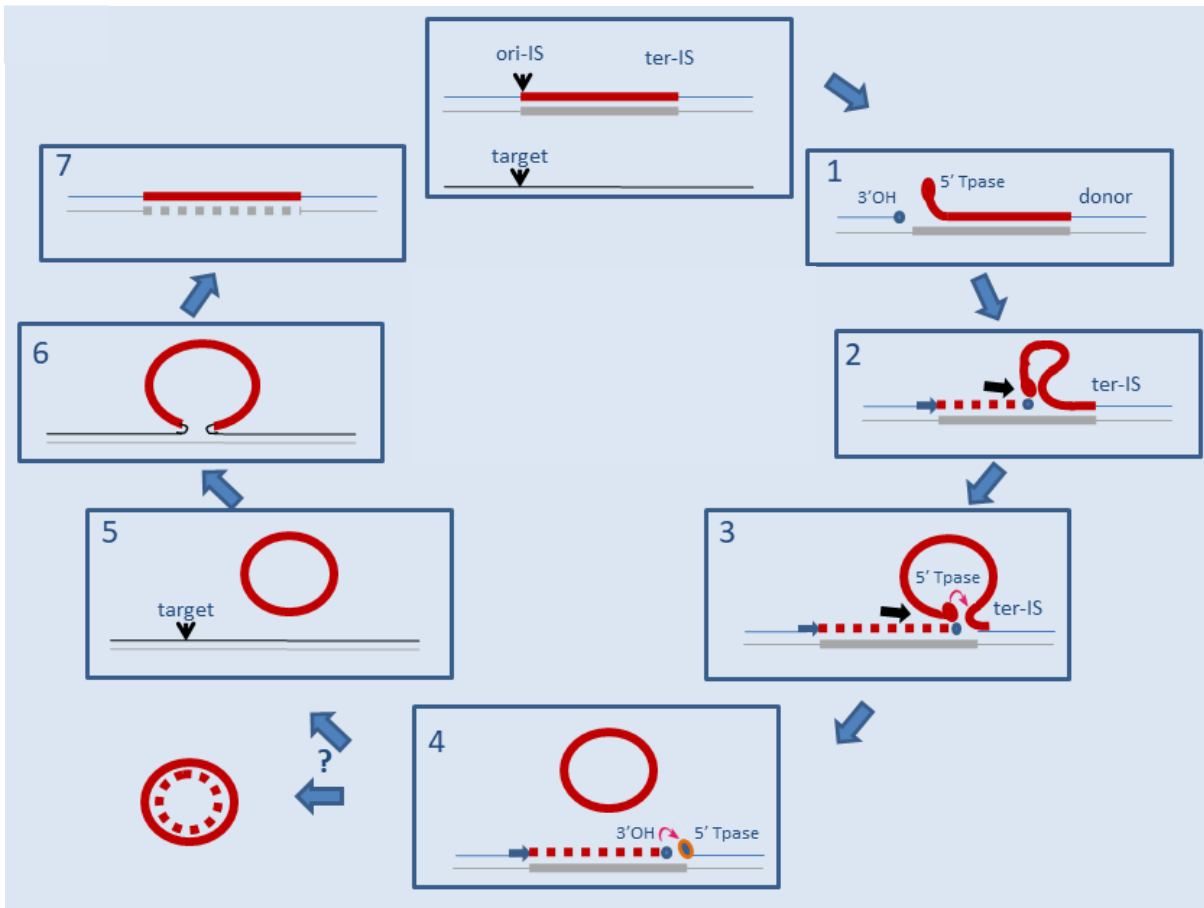


Figure 39 : Modèle de transposition de la famille IS91 dit « séquentiel »

Issue de Ross et al., 2021

Le brin déplacé du transposon est en rouge ; le brin non déplacé est en gris ; l'ADN nouvellement répliqué, ligne rouge pointillée. Pour la description de la figure, voir texte.

Toutefois, il convient de souligner que ces analyses manquent de preuves directes permettant de démontrer que les intermédiaires circulaires d'ADNsb ou d'ADNdb sont des intermédiaires fonctionnels dans la transposition d'IS91. Néanmoins, la possibilité qu'ils soient des intermédiaires de transposition fonctionnels est étayée par l'implication de *orf121* de l'élément IS91 dans la spécificité d'insertion de cet IS (Bernales et al., 1999). En effet, il a été exposé que les mutants *orf121* utilisent <sup>5</sup>CTCG comme site de clivage de seconde préférence. Or, <sup>5</sup>CTCG est précisément la séquence de clivage qui est reconnue du côté *terIS* de l'élément IS91 et est également la séquence adjacente à l'*oriIS* dans les intermédiaires circulaires. Il est donc possible que la transposase puisse cliver les intermédiaires d'ADNsb à cet endroit de jonction, à savoir entre le *terIS* et l'*oriIS*. Bien que cela n'ait pas encore démontré, la conséquence de ce clivage serait de permettre d'insérer l'intermédiaire circulaire d'ADN dans un ADN cible. Si c'est le cas, alors le clivage et la répllication par le mécanisme CR sont dissociés de l'insertion de la cible et se produisent en amont.

L'existence d'une telle preuve permettrait de trancher en faveur d'un modèle de transposition par CR de type séquentiel.

Par conséquent, nous pouvons conclure que chaque modèle implique des réactions de transestérification médiées par les résidus de tyrosine, ainsi qu'une génération des extrémités  $3'$ OH libres permettant la réplication de l'ADN. De plus, ces deux modèles démontrent parfaitement que l'*oriIS* permet d'initier la réplication par CR et que l'extrémité *terIS* marquerait la fin de cette réplication par CR. Néanmoins, en l'absence de données expérimentales en ce sens, il n'est pas possible de trancher entre le modèle concerté et le modèle séquentiel.

Plus récemment, il a été suggéré qu'un autre groupe de séquences, appelé « ISCR » (IS avec une « Common Regions »), apparentées à la famille IS91 auraient aussi une transposition CR (Sohn et al., 2009; Toleman et al., 2006a).

## **IV.2. Les éléments de la famille ISCR apparentée à IS91**

Les éléments de la famille ISCR sont généralement adjacents à des gènes de résistance aux antibiotiques. Ainsi, ces éléments joueraient un rôle prédominant dans la dissémination et l'expression des gènes de résistance aux antibiotiques (Lallement et al., 2018; Toleman et al., 2006b, 2006a; Toleman and Walsh, 2010).

### **IV.2.1. Description des ISCR**

En 1993, les ISCR ont été décrites pour la première fois sous le nom de CR après la région 3'conservée (3'CS) des intégrons de résistance de classe 1, In6 et In7 (Stokes et al., 1993). La région 3'CS est dupliquée et un ou plusieurs gènes de résistance aux antibiotiques sont présents entre les deux copies de la région 3'CS (Valentine et al., 1994). Pour autant ces gènes de résistance ne sont pas des cassettes d'intégrons car dépourvus de site *attC* en aval du gène. Chez ces intégrons In6 et In7, il a été observé une orf de 341 pb nommée CR pour « Common Region ». Par la suite, cette région a été renommée *orf513* présentant un cadre ouvert de lecture de 341 codons (Partridge and Hall, 2003b).

Les éléments ISCR1 n'apparaissent pas comme étant spécifiques d'une niche particulière (Toleman et al., 2006a). En effet, les premiers éléments ISCR1 étaient associés en majorité à des gènes de résistance aux antibiotiques et isolés principalement chez des bactéries environnementales (Arduino et al., 2002; Sabaté et al., 2002). Les éléments ISCR2 ont été isolés du bétail (Boyd et al., 2001; Cloeckert et al., 2000). Par la suite, les ISCR ont été identifiées dans des bactéries à Gram négatif de multirésistance d'intérêt clinique (Arduino et al., 2003).

De par leur structure et par leurs différents domaines composant leur transposase, les éléments ISCR ont été identifiés comme étant apparentés aux membres de la famille IS91 (Toleman et al.,

2006a). En 2007, le gène de la transposase des éléments ISCR a été renommé *rcr* pour « rolling circle replication ». Cette nouvelle appellation résulte de son mécanisme de transposition supposé par CR et par son homologie avec les protéines Rep des plasmides et phages se répliquant par ce mécanisme (Levings et al., 2007).

La famille ISCR est composée de 18 membres. Les éléments ISCR codent une transposase d'environ 500 acides aminés avec une identité entre les ISCR allant de 4 à 88 % (Toleman et al., 2006a). De plus, ces éléments possèdent des domaines fonctionnels d'une transposase de type HUH.

L'élément *ISCR1*, très répandu, proviendrait de l'environnement aquatique car il a été isolé chez des espèces marines comme *Aeromonas salmonicida* et *Shewanella algae* (Rhodes et al., 2000; Toleman et al., 2006a). L'élément *ISCR1* est toujours retrouvé en aval de la région 3'CS et des intégrons de résistance de classe 1 (Collis and Hall, 1995). Une grande diversité de gènes de résistance aux antibiotiques a été décrit en aval de l'élément *ISCR1* codant pour :

- Résistance aux  $\beta$ -lactamines (*bla<sub>CTX-M-1</sub>*, *bla<sub>PER-7/8</sub>*, *bla<sub>NDM-1</sub>*, *bla<sub>IMP-1</sub>*...) (Iguchi et al., 2014; Li et al., 2014; Su et al., 2008; Tada et al., 2017),
- Résistance au triméthoprim (dfrA23, dfrA10...) (Stokes et al., 1993; Villa and Carattoli, 2005),
- Résistance aux quinolones (*qnrA1*, *qnrA3*) (Cheng et al., 2016; Lascols et al., 2008),
- Résistance aux aminosides (*aph(3')*-*Vla*) (Naas et al., 1999),
- Résistance aux phénicolés (*catA*) (Stokes et al., 1993).

En 2000, l'élément *ISCR2* a été identifié en amont du gène *floR* conférant une résistance au florfénicol chez *Escherichia coli*, puis chez *Vibrio cholerae* dans l'ICE SXT (Cloeckaert et al., 2000; Hochhut et al., 2001). Le gène *floR* serait issu du pathogène de certains poissons tel que *Pseudomonas damsalae subsp piscida* (Toleman et al., 2006a). Par ailleurs, l'élément *ISCR2* a aussi été retrouvé au niveau du microbiote intestinal de l'Homme (Pal et al., 2016). Ces éléments sont principalement localisés sur des plasmides au sein de transposons, comme le *Tn6329* qui est formé par deux copies d'IS26, ou encore au sein de l'îlot de pathogénicité *AbaR* chez *Acinetobacter baumannii* (Beyrouthy et al., 2017; Seputiene et al., 2012). Contrairement à l'élément *ISCR1* qui est associé aux intégrons de résistance de classe 1, l'élément *ISCR2* est souvent retrouvé en amont du gène de résistance aux sulfonamides *sul2* sans pourtant démontrer une séquence cible particulière (Rådström and Swedberg, 1988). En aval de l'élément *ISCR2*, une grande diversité de gènes de résistance aux antibiotiques ont été décrits :

- Résistance aux sulfonamides (*sul2*...) (Cloeckaert et al., 2001),
- Résistance au triméthoprim (*dfrA9*, *dfrA20*...) (Jansson and Sköld, 1991; Kehrenberg and Schwarz, 2005),
- Résistance aux macrolides (*erm42*) (Harmer et al., 2015),

- Résistance aux aminosides (*rmtH*) (O'Hara et al., 2013),
- Résistance aux phénicolés (*floR*) (Wang et al., 2016),
- Résistance à la tétracyclines (*tetRtetY*) (Leclercq et al., 2016).

#### IV.2.2. Organisation génétique

Comme nous l'avons précédemment exposé, les éléments ISCR sont apparentés à la famille IS91 et ils possèdent des séquences spécifiques *oriS* et *terS*. Contrairement à certains éléments de la famille IS91, pour lesquels le clivage de l'*oriS* et du *terS* a été démontré, des séquences *oriS* et *terS* ont été proposées pour les éléments ISCR sans que leur fonctionnalité dans la transposition de ces éléments n'ait été étudiée (Garcillán-Barcia et al., 2002; Yassine, 2015; Toleman et al., 2006a).

Une séquence consensus de l'*oriS* a été établie pour les éléments de la famille IS91, à savoir :  $5'GxtTTTxAAaTTCCTATxCaT^3'$  (x : A, C, G ou T) (Garcillán-Barcia et al., 2002; Ross et al., 2021; Tavakoli et al., 2000). L'alignement de l'extrémité *oriS* des éléments ISCR a permis de mettre en évidence une origine potentielle de réplication, à savoir :  $5'GcgTTTgAAcTTCCTATACxx^3'$  (Toleman et al., 2006a). Contrairement aux éléments IS91 et IS801, les éléments ISCR ne possèdent qu'un seul palindrome subterminal à l'extrémité *oriS*, c'est aussi le cas chez les éléments IS1294 et IS1294b qui sont tous deux fonctionnels pour la transposition.

Une séquence consensus du *terS* a été établie pour les éléments IS91, IS801 et IS1294,  $5'GCAGCCxxxxGGCTGC^3'$  (x : A, C, G ou -) (Garcillán-Barcia et al., 2002; Ross et al., 2021). Cependant, l'élément ISCR1 ne possède pas de *terS* défini. L'élément ISCR2 présente un palindrome subterminal potentiel,  $5'AGTGxxxxxCAC^3'$  (x : C, G ou T). L'élément ISCR1 est toujours associé aux intégrons de classe 1. La région où devrait se situer le *terS*, est constitué par les gènes *qacED1* et *sul1* (résistance aux ammoniums quaternaires et aux sulfamides) de la région 3'CS de l'intégron de classe 1 (Toleman et al., 2006a, 2006b). Toleman et al. ont émis l'hypothèse selon laquelle le *terS* de ISCR1 disparaîtrait lorsque celui-ci serait répliqué au-delà du *terS* en utilisant le mécanisme de transposition « *one-ended transposition* » vers un second site au sein de l'intégron de classe 1 (Toleman et al., 2006a). Ceci aurait entraîné la dégénérescence du *terS* originel de l'élément ISCR1. Cela permettrait d'expliquer l'association entre l'élément ISCR1 et l'intégron de résistance de classe 1. Cependant, l'extrémité *terS* devrait présenter une diversité de gènes si ISCR1 était responsable du déplacement des gènes de résistance aux antibiotiques en utilisant le mécanisme de transposition « *one-ended transposition* ». Or, cette diversité de gènes de résistance chez ISCR1 n'est observée qu'au niveau de l'extrémité *oriS* au niveau d'une séquence  $5'GTGGTTTATACTTCATACCC^3'$ . En conséquence, cela ne permet pas de connaître le site cible de l'élément ISCR1. Il serait utile d'effectuer une analyse de la diversité des isoformes de l'élément ISCR1 afin de connaître l'étendue de la diversité de cet élément et de pouvoir éventuellement

émettre l'hypothèse d'un site *terIS* et d'un site cible d'insertion. L'absence d'un site cible d'insertion pourrait laisser supposer que, par homologie aux données connues pour les éléments de la famille IS91, les éléments ISCR1 identifiés seraient inactifs pour la transposition. Par ailleurs, deux promoteurs potentiels P<sub>OUT</sub>, orientés vers l'extérieur de l'ISCR1, ont été mis en évidence dans l'*oriIS* et ont été nommés P<sub>CR1-1</sub> et P<sub>CR1-2</sub>. (Mammeri et al., 2005; Rodriguez-Martinez et al., 2006). Ils participent à l'expression des gènes en aval de l'élément ISCR1 (Lallement et al., 2018). Pour terminer, les travaux de thèse de Claire Lallement au laboratoire ont permis de mettre en lumière que le promoteur P<sub>r<sub>cr1</sub></sub>, permettant l'expression du gène de la transposase *r<sub>cr1</sub>*, est sous le contrôle du répresseur LexA et est donc régulé par la réponse SOS (Lallement, 2018).

L'élément ISCR2 s'est avéré plus facile à définir que l'élément ISCR1 en raison de la diversité des séquences en amont et en aval de la localisation de l'IS. De plus, cet élément possède des extrémités *oriIS* et *terIS* entourant le gène *r<sub>cr2</sub>* codant pour la transposase. En effet, ces caractéristiques ont permis de rapidement définir un site cible d'insertion de l'élément <sup>5'</sup>GTTG, situé en amont, riche en A/T de l'extrémité *oriIS* (Ross et al., 2021). Ainsi et contrairement à l'élément ISCR1, ISCR2 possède toutes les caractéristiques d'une séquence d'insertion complète de la famille IS91. Toutefois, actuellement il n'existe aucune donnée sur l'activité de la transposase ou sur l'expression du gène *r<sub>cr2</sub>* à partir d'un promoteur intrinsèque.

#### **IV.2.3. Caractéristiques des transposases des ISCR**

Les transposases de la famille ISCR sont très proches des éléments de la famille IS91 (figure 40) (Chandler et al., 2013; Garcillán-Barcia et al., 2002; Ross et al., 2021; Toleman et al., 2006a).

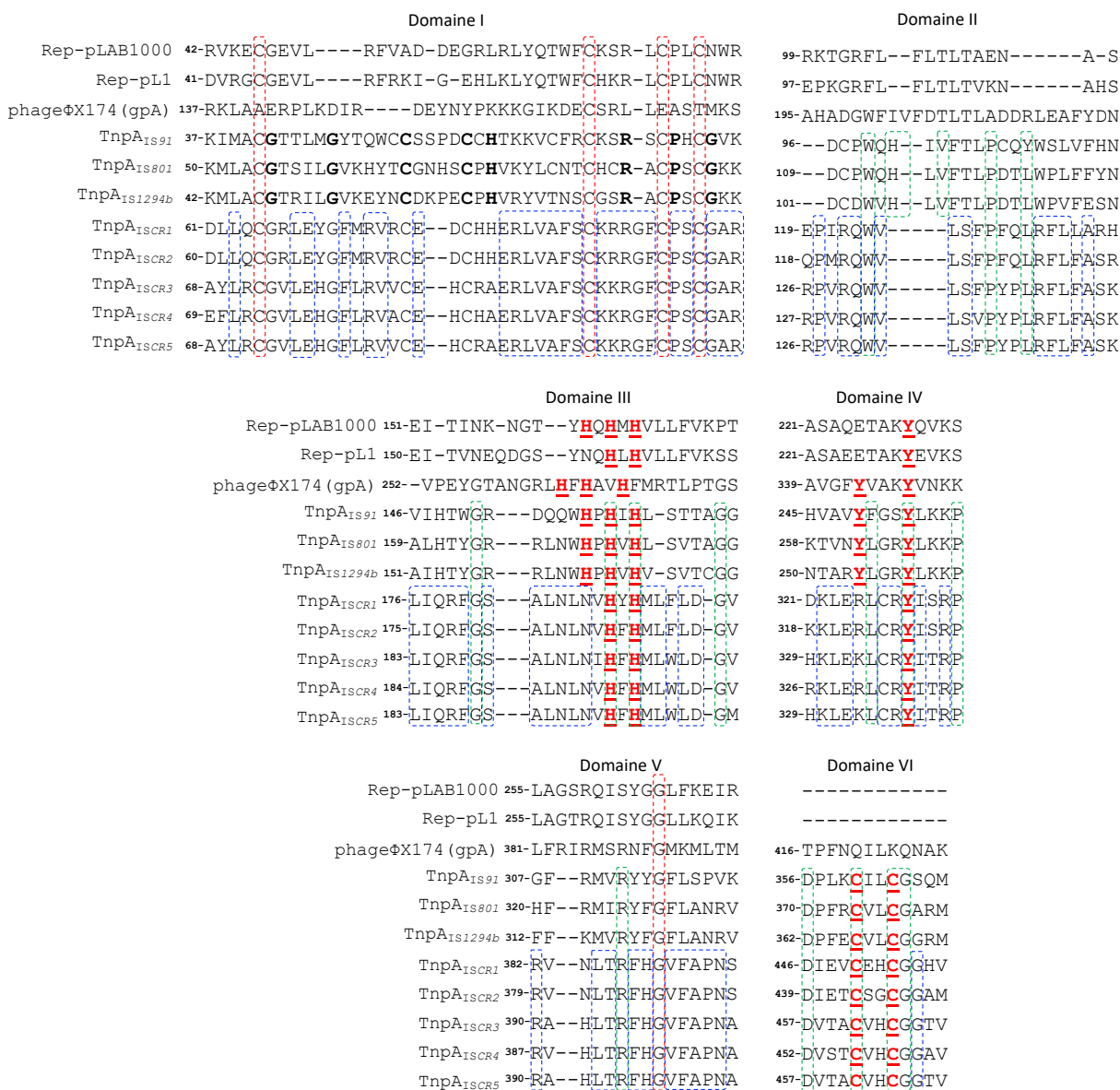


Figure 40 : Alignement des domaines conservés des éléments ISCR1, ISCR2, ISCR3, ISCR4 et ISCR5 avec les éléments de la famille IS91 (IS91, IS801 et IS1294b) et les protéines Rep et la protéine gpA

D'après et inspiré de Garcillán-Barcia et al., 2002; Ross et al., 2021; Toleman et al., 2006a

Entouré en rouge, les acides aminés qui sont conservés par l'ensemble des membres alignés et entouré en vert, uniquement les acides aminés qui sont conservés par les éléments de la famille IS91/ISCR. Entouré en bleu, les acides aminés qui sont conservés uniquement chez les ISCR. Dans le domaine III les histidines du motif HUH sont indiquées en rouge et dans le domaine IV les tyrosines du site catalytique sont également indiquées en rouge. Le motif putatif du doigt de zinc formé par les cystéines du domaine VI est indiqué en rouge.

En effet, les éléments ISCR présentent cinq motifs ainsi que la région conservée à l'extrémité C-terminale supplémentaire absente des protéines Rep et de la protéine gpA.

**Le domaine I** comprend une région riche en cystéines. En effet, nous retrouvons 6 cystéines conservées, toutefois il n'y a pas l'histidine comme dans la famille IS91. Ces cystéines sont



organisées comme suit : « C.X<sub>11</sub>.C.X<sub>2</sub>.C.X<sub>9</sub>.C.X<sub>5</sub>.C.X<sub>2</sub>.C ». De plus, par rapport aux transposases de la famille IS91 les éléments ISCR ne présentent pas deux des acides aminés entre deux cystéines et les acides aminés entre les cystéines sont conservés par l'ensemble des membres ISCR. Les éléments ISCR3, ISCR4 et ISCR5 présentent une alanine à la place d'une histidine. Les travaux de thèse de Haytham Yassine ont montré que la mutation de l'histidine H65 de l'élément IS1294b en alanine inhibait l'activité de la transposase (Yassine, 2015). Ces résultats suggèrent fortement que les transposases ISCR3, ISCR4 et ISCR5 ne seraient pas actives pour la transposition.

Précédemment, nous avons vu que **le domaine II** de la transposases des éléments de la famille IS91 était « W.X<sub>1</sub>.H.X<sub>1</sub>.V.F.T.L.P.X<sub>3</sub>.W ». *A contrario*, les éléments ISCR présentent un autre motif conservé à la même position mais avec toujours la conservation de la leucine (L), « P.X<sub>1</sub>.R.Q.X<sub>1</sub>.V.L.S.X<sub>1</sub>.P.X<sub>2</sub>.L.R.F.L.X<sub>1</sub>.A ». En gras, les acides aminés conservés présents également chez la famille IS91. Aucune donnée n'existe sur le rôle et la fonction de ce motif dans la transposition des éléments ISCR ou IS91.

**Le domaine III** présente seulement deux histidines invariantes pour les éléments ISCR contre trois pour les éléments de famille IS91. La configuration avec seulement deux histidines a déjà été décrite pour la protéine Rep-L1. Les ISCR présentent le motif suivant : « L.I.Q.R.F.G.S.A.L.N.L.N.L.X<sub>1</sub>.H.X<sub>1</sub>.H.M.L.X<sub>1</sub>.L.D.G », en rouge et en gras ce sont les histidines du motif HUH et seulement en gras, ce sont les acides aminés conservés avec les éléments de la famille IS91. Ce motif démontre que les transposases de la famille ISCR appartiennent à la super famille des protéines HUH. Notons qu'aucune évaluation de l'impact de la mutation ponctuelle de l'une ou de l'autre de ces histidines n'a encore été réalisée pour l'élément ISCR. Il convient de préciser que bien qu'un motif HUH ait été identifié pour les ISCR, celui-ci présente entre les deux résidus histidines soit une tyrosine chez ISCR1 c'est-à-dire un acide aminé ayant un indice d'hydrophobicité plutôt neutre (0,29), soit une phénylalanine pour les ISCR2, 3, 4 et 5, c'est-à-dire un acide aminé hydrophobe mais qui contient un cycle aromatique. Il n'existe actuellement aucune donnée relative à l'impact de ces changements dans le motif HUH des transposases *rcr*. Par ailleurs, les résultats décrits chez l'élément IS1294b suggèrent que l'arginine en position R157 qui est proche du motif HUH s'avère critique pour la transposition (Yassine, 2015). Cette arginine est conservée au sein de la famille IS91 mais pas pour les éléments ISCR. En effet, à la place de l'arginine, les éléments ISCR présentent une sérine (S), laissant supposer que les éléments ISCR pourraient ne pas être actifs pour la transposition.

**Le domaine IV** formant le site catalytique ne comporte qu'une seule tyrosine pour les transposases *rcr* contre deux pour les éléments de la famille IS91. Ce motif se caractérise de la manière suivante : « K.L.E.X<sub>1</sub>.L.C.R.Y.I.X<sub>1</sub>.R.P ». En gras et en rouge, il s'agit de la position des tyrosines catalytiques supposées et seulement en gras, ce sont les acides aminés conservés (présents également chez la famille IS91). Dès lors, les membres de la famille ISCR semblent plus étroitement liés à la famille

des protéines Rep qu'à celle de la protéine gpA du phage  $\Phi$ X174. En effet, ces membres ne contiennent qu'un seul résidu tyrosine conservé contrairement à la protéine gpA et la famille IS91 qui en possèdent deux.

**Le domaine V** se caractérise par la conservation de 15 acides aminés : « R.X<sub>2</sub>.L.T.**R**.F.H.**G**.V.F.A.P.N ». En gras, il s'agit de l'acide aminé qui est présent également chez les éléments de la famille IS91 et en gras souligné, il s'agit de l'acide aminé conservé avec les protéines Rep et la protéine gpA en sus des éléments de la famille IS91. Actuellement, il n'existe aucune donnée sur le rôle potentiel de chaque acide aminé de ce motif dans la transposition des éléments de la famille IS91 ou des éléments ISCR.

Pour terminer, **le domaine VI** comprend la région conservée à l'extrémité C-terminale supplémentaire, qui a été mise en évidence pour les éléments de la famille IS91.

L'ensemble de ces données indiquent que les éléments ISCR font partie de la super famille des endonucléases HUH au vu :

- De leur proximité avec les protéines Rep, la protéine gpA et les éléments de famille IS91 de leur transposase,
- De leur organisation génétique avec des régions *oriS* et *terS*, ainsi que la présence des structures tige-boucles à leurs extrémités.

Ces éléments ISCR devraient présenter un mécanisme de transposition semblable à celui des éléments de la famille IS91, selon un mécanisme CR. Néanmoins, les multiples différences qui ont été mises en évidence avec les éléments de la famille IS91 pourraient indiquer des divergences. Soulignons que l'activité de transposition pour les éléments ISCR n'a jamais été démontrée, à l'exception de la transposase ISCR1 pour laquelle il a été montré que la transposase *rcr1* est capable de cliver l'extrémité *oriISCR1* de l'élément, mais dans des conditions de très forte quantité de transposase purifiée (500 ng) et sans ADN compétiteur (Quiroga et al., 2013).

Contrairement aux éléments IS91 qui semblent plutôt associés à l'homme ou aux mammifères, les éléments ISCR sont issus d'autres environnements, aquatiques ou telluriques (Toleman et al., 2006a). Cela pourrait expliquer notre difficulté à étudier la mobilité des éléments ISCR dans la transposition contrairement à celle des éléments de la famille IS91. En effet, la mobilité des éléments ISCR pourrait être induite par des facteurs hôtes ou de l'environnement d'où ils sont originaires, dans des conditions encore inconnues. Une autre possibilité serait que la diversité des éléments ISCR soit plus associée à des échanges de gènes par recombinaison homologue, que par transposition (Boyd et al., 2002; Toleman et al., 2006a). Cette hypothèse est étayée par le fait que les éléments ISCR3, ISCR4, ISCR14 et ISCR16 présentent entre 75% et 97% d'identité en acides aminés de leur transposase et qu'ils sont situés en aval d'un *orf* de type *groE*, ce qui peut suggérer qu'ils pourraient tous avoir un ancêtre commun (Toleman and Walsh, 2008). Soulignons enfin qu'il

n'y a pas encore de démonstration formelle que ces éléments transposent. De plus, comme leur environnement est identique, cela suggère qu'ils sont relativement inactifs dans la transposition et que la diversité des éléments ISCR a été créée par des recombinaisons homologues entre les ISCR.

Pour conclure, des similarités existent entre les différents membres de la super famille HUH. Ces similarités nous amènent à penser que les ISCR transposent de la même façon que les autres transposases HUH. Cependant, bien que les transposases ISCR partagent des similarités avec celles de la famille IS91 et de la famille IS200/IS605, ces deux familles possèdent des mécanismes de transposition différents. Il est probable que les ISCR transposent par l'un de ces deux mécanismes, voir un mécanisme hybride se situant entre celui de la famille IS91 et celui de la famille IS200/IS605 qui n'a qu'une seule tyrosine catalytique. Ainsi, les ISCR pourraient être actives en tant que dimère comme les éléments IS200/IS605, contrairement aux éléments IS91 où il semblerait que la transposition est médiée par un monomère.

Il ressort de ce qui précède que :

- Les éléments de la famille **IS91** :
  - Sont associés à des gènes de virulence et de résistance aux antibiotiques,
  - Ont un contexte génétique et une transposase TnpA (HUH, Y2) qui indiquent que ces éléments font partie de la super famille HUH,
  - Comme les éléments IS801 et IS1294/IS1294b sont actifs pour la transposition,
  - Transposeraient selon un mécanisme de répllication par CR où deux modèles sont proposés, soit le modèle « concerté », soit le modèle « séquentiel ».
- Les éléments de la famille **ISCR** :
  - Sont associés à des gènes de résistance aux antibiotiques,
  - Ont un contexte génétique et une transposase rcr (HUH, Y1) qui indiquent que ces éléments font partie de la super famille HUH,
  - Sont apparentés à la famille IS91, même s'il existe plusieurs différences entre les éléments ISCR et IS91,
  - Ne peuvent pas être considérés comme actifs dans la transposition car aucune donnée expérimentale n'existe et la diversité des éléments ISCR pourrait être le fait de recombinaisons homologues.

## Chapitre V. Régulation de la transposition et impact des IS chez les procaryotes

De nombreux mécanismes classiques de contrôle de l'expression des gènes tels que des répresseurs transcriptionnels chez IS1, IS2 ou des inhibiteurs traductionnels par des ARN anti-sens comme chez IS10, sont connus pour être des acteurs dans la régulation de l'expression de la transposase des IS (Escoubas et al., 1991; Hu et al., 1994; Kleckner et al., 1996). De plus, plusieurs mécanismes coopèrent ensemble dans la régulation de l'expression des gènes et de la transposase. Préalablement à tout développement, il convient à titre liminaire d'aborder les notions de transcription et de traduction chez les procaryotes.

### V.1. La transcription chez les procaryotes

#### V.1.1. Le promoteur

L'étape principale de l'initiation de la transcription est la reconnaissance du promoteur par l'ARN polymérase.

Un promoteur fonctionnel est constitué des éléments -10, -35, -10 étendu et UP. Il s'agit de sites qui sont reconnus par la machinerie transcriptionnelle (Browning and Busby, 2004). Les boîtes -10 et -35 sont définies par rapport au +1 d'initiation de transcription.

Le +1 d'initiation de transcription correspond à la première base transcrite par l'ARN polymérase (figure 41). Ce site d'initiation +1 est généralement un nucléotide A ou G, très rarement un nucléotide C ou T. Le +1 de transcription est localisé environ à 35 pb avant le codon START, c'est-à-dire du gène en aval du promoteur (Shultzaberger et al., 2007). Le codon START le plus courant est ATG. Toutefois, les procaryotes utilisent des codons de départ alternatifs de manière significative, principalement GUG et UUG (Blattner et al., 1997), et plus rarement, le codon AUU (Sacerdot et al., 1982). De plus, deux études ont montré indépendamment qu'au moins 17 codons de départ non AUG peuvent initier la traduction chez *E. coli* (Hecht et al., 2017).

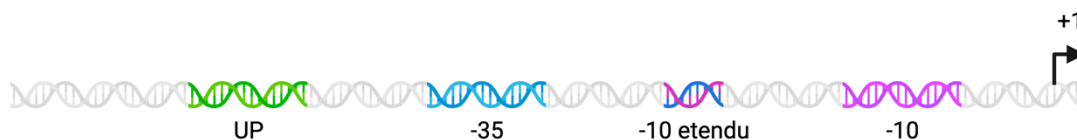


Figure 41 : Organisation du promoteur bactérien  
Created with BioRender.com

Les séquences pour les éléments du promoteur sont représentées en vert pour la séquence UP, en bleu pour la boîte -35, en rose/bleu pour la boîte -10 étendue (rose/bleu) et en rose pour la boîte -10. Le +1 de transcription est indiqué par une flèche noire.

L'hexamère -10 du promoteur est reconnu par le domaine 2 de la sous-unité  $\sigma$  de l'ARN polymérase. De même, l'hexamère -35 du promoteur est reconnu par le domaine 4 de la sous-unité  $\sigma$ . Des séquences hexamères consensuelles ont été établies pour les hexamères -10 et -35 (Campbell et al., 2002). En outre, le facteur  $\sigma^{70}$  est l'ARN polymérase qui reconnaît les séquences consensus hexamériques  $5'$ TATAAT et  $5'$ TTGACA pour les éléments -10 et -35, respectivement. La distance entre les éléments -10 et -35 varie de 15 à 21 pb avec une distance optimum de 17 pb (Hawley and McClure, 1983; Shultzaberger et al., 2007).

Les deux autres éléments importants sont l'élément -10 étendu et l'élément UP. L'élément -10 étendu est un motif de 3-4 pb situé immédiatement en amont de l'hexamère -10 qui est reconnu par le domaine 3 de la sous-unité  $\sigma$  de l'ARN polymérase (Murakami et al., 2002; Sanderson et al., 2003). L'élément UP est une séquence d'environ 20 pb, riche en A et T, située en amont de l'hexamère -35 du promoteur qui est reconnu par les domaines C-terminaux des sous-unités  $\alpha$  de l'ARN polymérase (Ross et al., 2001).

Ainsi, ensemble, les éléments -10, -35, -10 étendu et UP spécifient la liaison initiale de l'ARN polymérase à un promoteur.

Enfin, ils existent d'autres facteurs transcriptionnels que le facteur  $\sigma^{70}$  impliqué dans diverses fonctions (Helmann, 2002; Paget and Helmann, 2003) :

- Stress nutritif ( $\sigma^{38}$  ou  $\sigma^S$ , HrdA, HrdC, HrdD),
- Choc thermique ( $\sigma^{32}$ ),
- Synthèses des flagelles ( $\sigma^{28}$ ),
- Sporulation ( $\sigma^F$  et  $\sigma^G$ ),
- Des fonctions extra-cytoplasmiques (biodisponibilité du fer ( $\sigma^{19}$ ), virulence (Vrel), résistance aux antibiotiques ( $\sigma^W$ )),
- Le facteur  $\sigma^{54}$  impliqué dans la synthèse des flagelles, l'utilisation du nitrate, le chimiotaxisme et le métabolisme énergétique.

### V.1.2. L'ARN polymérase

L'ARN polymérase existe sous deux formes :

- Soit sous la forme d'apoenzyme formant le « core » de l'ARN polymérase qui est constitué de cinq sous-unités.
- Soit sous la forme d'holoenzyme, où le « core » est associé à un facteur  $\sigma$ .

Le « core » de l'ARN polymérase est constitué des cinq sous-unités formant une protéine d'environ 450 kDa (figure 42). :

- Deux sous-unités alpha ( $\alpha$ ),
- Une sous-unité bêta ( $\beta$ ),
- Une sous-unité bêta prime ( $\beta'$ ),
- Une petite sous-unité oméga ( $\omega$ ).

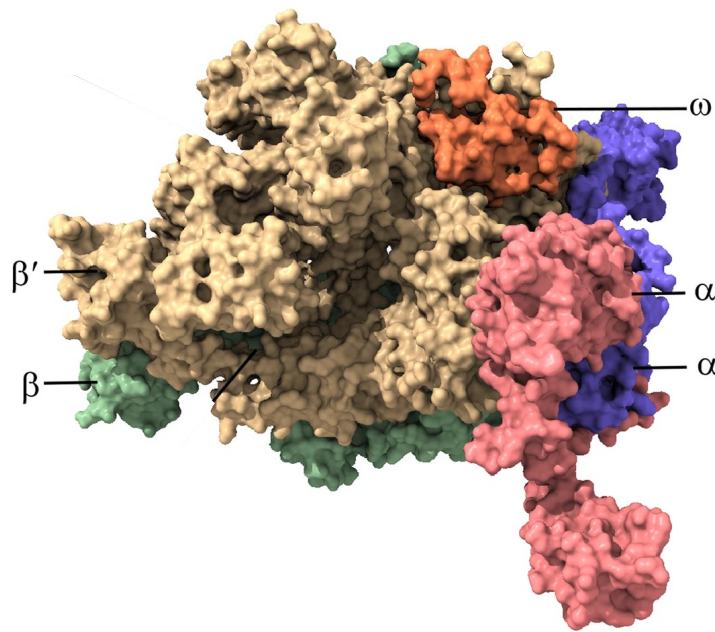


Figure 42 : Core de l'ARN polymérase.

Adaptée de Boyle et al., 2020

*Modèle du noyau de l'ARN polymérase avec les sous-unités  $\beta'$  (jaune),  $\beta$  (vert),  $\alpha$  (rose et violet), et  $\omega$  (orange).*

Seul, le « core » ne peut pas reconnaître les séquences promotrices pour initier la transcription. En outre, pour reconnaître les séquences spécifiques des promoteurs, le « core » nécessite l'intervention d'une dernière sous-unité : le facteur  $\sigma$ .

Le facteur  $\sigma$  a une taille variable de 20 à 70 kDa. Cette sous-unité joue un rôle primordial car elle permet la fixation au promoteur et elle est nécessaire à l'ancrage du « core » à la région promotrice (Burgess et al., 1969; Wösten, 1998).

Si nous prenons l'exemple du facteur  $\sigma 70$ , il possède quatre domaines conservés nommés  $\sigma 1$ ,  $\sigma 2$ ,  $\sigma 3$  et  $\sigma 4$ , ainsi qu'un « linker » reliant les domaines  $\sigma 3$  et  $\sigma 4$  entre eux (Paget and Helmann, 2003). En outre, il a été mis en évidence que les domaines  $\sigma 2$ ,  $\sigma 3$  et  $\sigma 4$  jouent un rôle important dans la reconnaissance de la séquence promotrice (figure 43) (Gross et al., 1998; Murakami et al., 2002) :

- Le domaine  $\sigma 2$  reconnaît et se fixe à l'hexamère -10,
- Le domaine  $\sigma 3$  reconnaît et se fixe à la région -10 étendue,

- Le domaine  $\sigma_4$  reconnaît et se fixe à l'hexamère -35.

Le domaine  $\sigma_1$  se conserve moins bien au sein des facteurs  $\sigma$ . À ce jour, la fonction du domaine  $\sigma_1$  est inconnue.

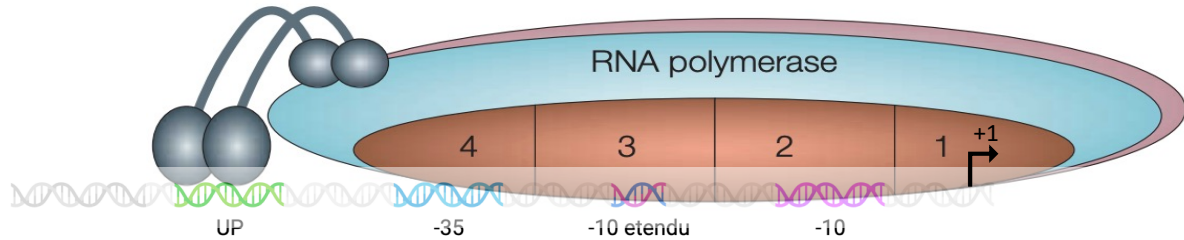


Figure 43 : Interaction entre l'ARN polymérase et les différentes régions d'un promoteur

Created with BioRender.com et adaptée de Browning and Busby, 2004

L'ARN est représenté en cartoon, en bleu les deux sous-unités  $\beta'$  et  $\beta$ , en gris les deux sous-unités  $\alpha$  et les domaines du facteur  $\sigma$  sont en marron numéroté 1 à 4. Les séquences pour les éléments du promoteur sont représentées en vert pour la séquence UP, en bleu pour la boîte -35, en rose/bleu pour la boîte -10 étendue (rose/bleu) et en rose pour la boîte -10. Le +1 de transcription est indiqué par une flèche noire.

### V.1.3. La transcription

L'initiation de la transcription nécessite l'interaction de l'ARN polymérase avec l'ADN promoteur et la formation d'un complexe ouvert. Dans ce complexe, l'ADN duplex est déroulé autour du point de départ de la transcription (deHaseth et al., 1998).

Le rôle des éléments promoteurs est principalement d'amarrer l'ARN polymérase de manière à ce qu'elle soit compétente pour la formation de complexes ouverts, ainsi que les étapes suivantes de la transcription.

La liaison initiale de l'ARN polymérase est initiée par les domaines  $\sigma_2$ ,  $\sigma_3$  et  $\sigma_4$  du facteur sigma sur les séquences conservées du promoteur (figure 44). L'ARN polymérase fixée sur le promoteur est un complexe dit fermé stable qui couvre les 55 premières bases en amont du +1 de transcription.

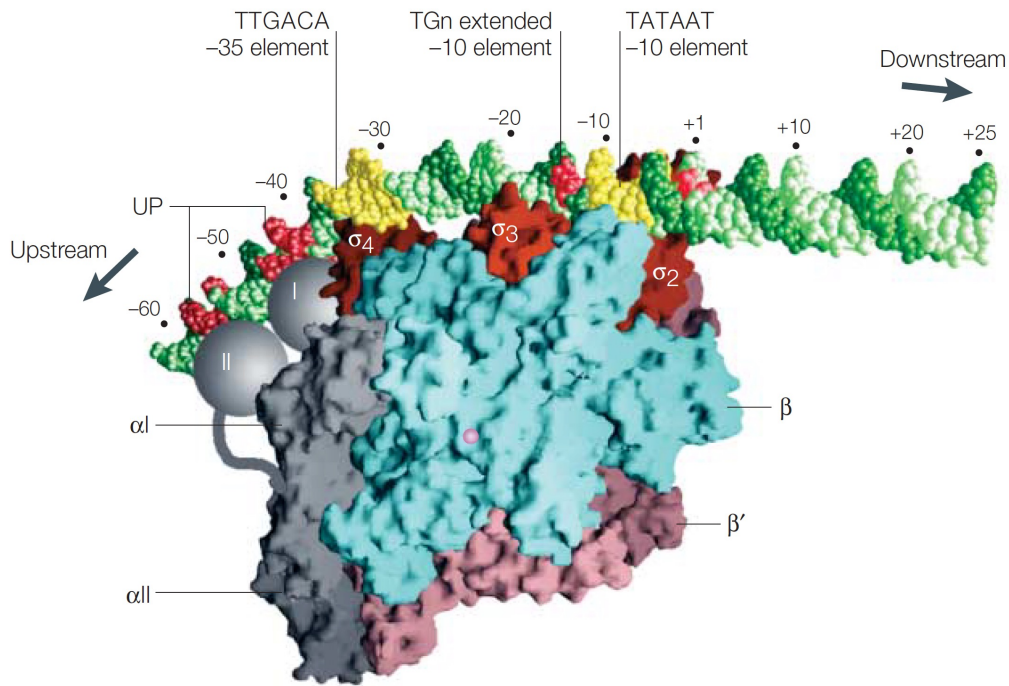


Figure 44 : Ancrage initial de l'holoenzyme de l'ARN polymérase à un promoteur.

Issue de Browning and Busby, 2004

Le brin d'ADNdb est représenté en vert, les éléments -10 et -35 sont en jaune et les éléments étendus -10 et UP sont en rouge sur ce brin d'ADNdb. Les sous-unités de l'ARN polymérase  $\beta$  et  $\beta'$  sont en bleu clair et en rose, respectivement. Les sous-unités  $\alpha$  sont en gris et les différents domaines de  $\sigma$  sont colorés en rouge. Les sphères grises, étiquetées I et II représentent les domaines de l' $\alpha$ CTD qui se lient au promoteur. Le site actif de l'ARN polymérase est indiqué par l'ion  $Mg^{2+}$  (magenta).

Après la liaison initiale de l'ARN polymérase, il s'ensuit un changement de conformation. Cette modification entraîne l'ouverture du brin, juste en aval du point de départ de la transcription, et cela génère le complexe ouvert d'ADN allant approximativement de la position -10 à la position +2 (Tomsic et al., 2001; Tsujikawa et al., 2002). Ceci correspond à l'isomérisation vers le complexe ouvert.

Ensuite, l'ARN polymérase incorpore les deux premiers nucléotides +1 et +2 jusqu'à former un transcrit d'environ 10 nucléotides. Ce petit transcrit a pour conséquence de déstabiliser le complexe et de permettre la dissociation du « core » et du facteur  $\sigma$  (Mooney et al., 2005; Murakami et al., 2002). Ainsi, la transcription se poursuit avec le complexe d'élongation correspondant au « core ».

A la fin de la transcription, il a été mesuré que l'ARN polymérase incorpore environ 25 à 65 nucléotides par seconde *in vivo* chez *Escherichia coli* (Epshtein and Nudler, 2003). La bulle de transcription d'environ 17 pb migre en un complexe d'élongation permettant l'accès au brin complémentaire de façon constante. Au cours de l'élongation, l'ARN polymérase effectue des



pauses ponctuelles réversibles approximativement tous les 100 nucléotides (Kireeva and Kashlev, 2009).

L'élongation de la transcription se termine lorsque le complexe d'élongation est déstabilisé par les terminateurs qui le perturbent et libèrent l'ARN :

- La reconnaissance de la séquence *rut* par le facteur de transcription Rho, qui est une protéine homohexamérique, induit la déstabilisation du complexe d'élongation et provoque la terminaison de la transcription (Richardson, 2002).
- La reconnaissance d'une séquence palindromique, riche en G/C, qui est suivie d'une série de T. En conséquence, lors de l'ouverture de l'ADN, il y aura formation d'une tige-boucle, puis d'un complexe hybride de A et U instable permettant la dissociation du complexe d'élongation (Yarnell and Roberts, 1999).

## **V.2. La traduction chez les procaryotes**

La traduction chez les procaryotes est initiée par le ribosome qui effectue la synthèse biologique des protéines à partir de l'ARNm. Les ribosomes sont constitués de deux sous-unités, 50S et 30S, qui s'emboîtent l'une dans l'autre et travaillent ensemble lors de la synthèse des protéines afin de traduire l'ARNm en une chaîne polypeptidique.

Chez les procaryotes, la sous-unité 30S contient un ARNr (ARN ribosomique 16S) d'environ 1500 nucléotides, tandis que la sous-unité 50S contient deux ARNr 5S et 23S.

L'absence de membrane nucléaire séparant physiquement l'ADN chromosomique de l'environnement cytosolique conduit à reconnaître que la transcription et la traduction sont couplées dans l'espace et dans le temps chez les bactéries et les archées. Le couplage transcription-traduction se produit lorsque les ribosomes se lient aux ARNm naissants et qu'ils commencent à les traduire alors que la transcription est encore en cours, cela forme ainsi un complexe « RNAP-ARNm naissant-ribosome ». (Kohler et al., 2017). Le ribosome se lie au site de liaison du ribosome (RBS) de l'ARNm dès qu'il émerge de l'ARN polymérase (figure 45).

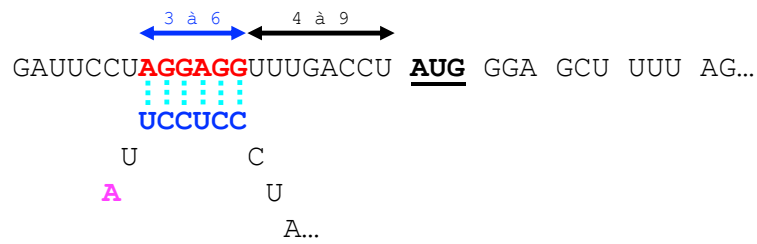


Figure 45 : Organisation du site de liaison du ribosome (RBS)

Le codon d'initiation de traduction qui se lie fMet-tRNA est en gras et souligné. En rouge la séquence consensus Shine-Dalgarno qui est complémentaire de la séquence complémentaire anti-SD de l'ARNr 16S en bleu et l'interaction entre les deux par un trait en pointillé cyan. En rose est indiqué l'extrémité 3'-end de l'ARNr 16S. Les deux flèches indiquent les distances entre les le codon d'initiation et la SD en noir ou la longueur de la SD en bleu.

Les ARNm sont souvent polycistroniques chez les procaryotes, de sorte que les ribosomes doivent trouver d'une part les codons de départ internes et d'autre part, celui qui est le plus proche de l'extrémité 5' (Ma et al., 2002). La complémentarité des nucléotides entre la queue de l'ARN ribosomique 16S et une région située immédiatement en amont du codon de départ est un mécanisme qui aide les ribosomes procaryotes à trouver le bon codon de départ. Cette région, appelée la séquence Shine-Dalgarno (SD), se situe entre quatre à dix bases du codon d'initiation avec un idéal à 7 bases. La séquence SD est une séquence qui se compose de quatre à six nucléotides (Ma et al., 2002; Shine and Dalgarno, 1975, 1974). La séquence SD est reconnue par l'extrémité 3' de l'ARNr 16S qui présente la séquence complémentaire. La séquence SD la plus courante est <sup>5'</sup>AGGAGG et celle de l'extrémité 3' de ARNr 16S contient la séquence CACCUCCUUA<sup>3'</sup> qui peut s'hybrider avec la séquence SD (Amin et al., 2018).

L'élongation de la chaîne polypeptidique implique l'ajout d'acides aminés à l'extrémité carboxyle de la chaîne en croissance. La protéine en croissance quitte le ribosome par le tunnel de sortie du polypeptide dans la grande sous-unité (Mitra et al., 2005).

Dans des conditions de croissance optimales, les ribosomes de *Escherichia coli* traduisent 14 à 17 acides aminés par seconde (Proshkin et al., 2010; Young and Bremer, 1976; Zhu et al., 2016). Cela signifie qu'ils transloquent environ 42 à 51 nucléotides par seconde le long de l'ARNm en cours de traduction. Ainsi, les taux de transcription et de traduction de l'ARNm sont bien adaptés. Les taux de traduction et de transcription varient en fonction des conditions de croissance, néanmoins les taux des deux processus restent coordonnés (Iyer et al., 2018; Proshkin et al., 2010; Vogel and Jensen, 1994).

Le ribosome traduit les codons de l'ARNm jusqu'à ce que le ribosome atteigne un codon d'arrêt sur l'ARNm (UAA, UGA ou UAG).

### V.3. Contrôle transcriptionnel chez les IS

#### V.3.1. Promoteurs endogènes

De nombreux promoteurs endogènes décrits qui dirigent l'expression de la transposase d'IS sont en fait souvent inefficaces ou peu efficaces. De plus, les promoteurs endogènes sont généralement localisés partiellement dans les séquences IR de l'élément, ce qui peut provoquer une régulation de l'expression lors de la fixation de la transposase à ces IR.

Lors de la formation de l'intermédiaire circulaire de l'élément IS911 au cours de la transposition, les deux séquences IR terminales sont liées (Duval-Valentin et al., 2001). Il a été mis en évidence que la synthèse de la transposase est généralement pilotée par un promoteur faible appelé  $P_{IRL}$ , ce dernier se localise à l'extrémité gauche (IRL). Lors de la formation d'un intermédiaire circulaire de transposition, la jonction circulaire crée un promoteur fort appelé  $P_{junc}$ . Ce dernier est 30 à 50 fois plus puissant que le promoteur endogène  $P_{IRL}$ . La boîte -35 du promoteur est située à l'extrémité droite (IRR) et la boîte -10 à l'extrémité IRL pour le promoteur  $P_{junc}$  (Ton-Hoang et al., 1997). En outre, l'inactivation du promoteur  $P_{junc}$  par mutagenèse a fortement réduit la transposition de IS911 *in vivo* (Duval-Valentin et al., 2001). Cette caractéristique a également été identifiée dans les formes circulaires d'IS492 de la famille IS110 (Perkins-Balding et al., 1999).

Enfin, il a été démontré que les éléments IS10, IS50 et IS903 portent des méthylations DAM proches des promoteurs endogènes ou les chevauchant, entraînant une régulation de l'expression de la transposase (Derbyshire and Grindley, 1992; Roberts et al., 1985; Tomcsanyi and Berg, 1989). Lorsque les sites sont entièrement méthylés, le promoteur est inefficace. *A contrario*, lorsque le site est hémiméthylé, le promoteur est plus actif.

#### V.3.2. Promoteurs transitoires

Les IS peuvent aussi contenir des séquences promotrices permettant l'expression des gènes en aval (figure 46) :

- Soit les IS forment un promoteur hybride lors de l'insertion avec la boîte -35 dans l'IS et la boîte -10 dans la séquence d'ADN adjacente.
- Soit les IS ont un promoteur complet comprenant les boîtes -35 et -10 dans une des extrémités de l'IS, ce promoteur étant appelé  $P_{OUT}$  pour « *outward-oriented promoters* ».

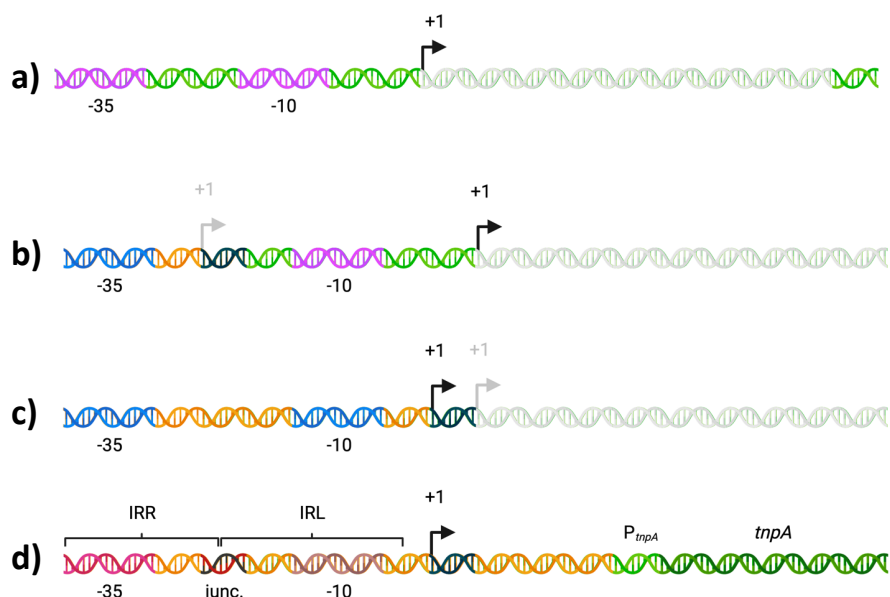


Figure 46 : Augmentation de l'expression du gène *tnpA* ou activation des gènes par l'insertion d'une IS

Created with BioRender.com

a) Promoteur classique d'un gène bactérien avec le gène en gris, en vert la séquence du chromosome ou plasmide et en rose les éléments -35 et -10.

b) Promoteur hybride formé par l'apport de l'élément -35 en bleu de l'IS en orange et sont +1 en gris.

c) Promoteur  $P_{OUT}$  est l'apport d'un promoteur complet par l'insertion d'une IS.

d) Promoteur transitoire fort créé dans l'intermédiaire circulaire de transposition, où l'élément -35 (rouge) et -10 (marron) sont fournis par l'extrémité IRR et IRL lors de la jonction (rouge/noir), respectivement. Ce qui permet de sur-exprimer transitoirement la transposase (vert) située en aval.

IS1490, IS30, IS257, IS2, IS911, IS982, ISVa1 et ISCR1) (Brookfield, 1989; Craig et al., 2002; Lallement, 2018; Mahillon and Chandler, 1998).

Outre des fonctions du métabolisme ou de virulence (Bongers et al., 2003; Coleman et al., 2014; Han et al., 2011; López de Felipe et al., 1996; Wood et al., 1991), il a été montré que la formation de ce type de promoteurs permettait l'expression de nombreux gènes de résistance aux antibiotiques :

- L'élément IS18 est capable d'activer l'expression du gène silencieux *aac(6')-Ij* de *Acinetobacter* sp., ce qui confère une résistance aux aminosides (Rudant et al., 1998),
- L'insertion de l'élément IS2 chez *Escherichia coli* permet l'expression du gène *acrEF* qui confère une résistance aux fluoroquinolones (Jellen-Ritter and Kern, 2001),
- L'insertion de l'élément IS256 en amont du gène *mecA* a augmenté la transcription de ce gène, ce qui confère une résistance à la méticilline chez *Staphylococcus sciuri* (Couto et al., 2003),
- L'élément IS257 confère une résistance au triméthoprime chez *Staphylococcus aureus* (Leelaporn et al., 1994),

- L'élément *ISEcp1B*, associé au gène *bla<sub>CTX-M</sub>*, permet une résistance aux  $\beta$ -lactamines (Lartigue et al., 2006),
- Les éléments *IS1* et *IS10* chez *Salmonella enterica* permettent une surexpression de l'opéron *acrEF* qui confère une résistance aux fluoroquinolones (Olliver et al., 2005),
- Les promoteurs *P<sub>CR1-1</sub>* et *P<sub>CR1-2</sub>* de l'élément *ISCR1* sont impliqués dans l'expression des gènes en aval *bla<sub>CTX-M-9</sub>* et *dfrA19* qui confèrent une résistance aux  $\beta$ -lactamines et au triméthoprim (Lallement et al., 2018), respectivement.

D'autre part, la présence d'une boîte -10 a également été détectée dans l'extrémité gauche de plusieurs éléments. Lors de l'insertion de l'IS, la combinaison de la boîte -10 avec la boîte -35, résidant dans l'extrémité droite ou dans l'ADN adjacent, peut générer un promoteur efficace. Il a été démontré que ce type de promoteurs pouvait aussi être formé par les intermédiaires circulaires, comme cela a été démontré avec *IS911* (Duval-Valentin et al., 2001; Ton-Hoang et al., 1997). L'intérêt de la formation d'un promoteur transitoire au sein de l'intermédiaire circulaire d'une IS est de permettre une forte expression de la transposase et de générer une augmentation conséquente de l'activité de transposition et donc son insertion (Duval-Valentin et al., 2001; Lewis et al., 2004; Reimann et al., 1989). De nombreuses IS ont adopté ce mécanisme comme *IS2*, *IS21*, *IS30*, *IS492* et *IS911* (Craig et al., 2002).

### V.3.3. Termineur de transcription

Chez de nombreuses IS, il n'est pas rare de trouver des séquences semblables aux signaux de terminaison de la transcription dans de nombreux gènes *tnpA* comme chez *IS1*, *IS2*, *IS3*, *IS5*, *IS30* (Berg et al., 1989)... En outre, il a été démontré des fréquences de transpositions réduites pour les transposons *Tn5* et *Tn9*, ainsi que pour les séquences d'insertion *IS1* et *IS5* chez *Escherichia coli K-12* qui portent une mutation dans *Rho* (Datta and Rosner, 1987). De surcroît, de nombreuses IS portent le domaine catalytique dans la région C-terminale de la protéine. En conséquence, lors de la troncature de l'ARNm par un termineur « prématuré », la transcription peut conduire à la production d'un fragment N-terminal de la transposase. Ce fragment peut alors agir comme un composant régulateur de la transposition de l'IS.

Certaines IS ne possèdent pas de codon de terminaison et la transcription continue alors au-delà de l'extrémité de l'IS dans l'ADN flanquant en aval comme décrit chez les éléments *IS240C*, *IS427* et *ISMk1*.

### V.3.4. Protection contre la transcription « impinging »

De nombreuses IS ont développé des mécanismes qui atténuent leur activation en entravant la transcription après leur insertion dans des gènes hôtes actifs. Le « impinging » de la transcription est un mécanisme d'interférence transcriptionnelles qui pourrait résulter de la perturbation des

complexes entre l'extrémité de l'IS et la transposase (Berg et al., 1989). Ce phénomène a été observé dans le cas de l'IS1, puis chez le phage Mu et plus tard chez l'élément l'IS50 (Berg et al., 1989; Chandler and Galas, 1983a; Goosen and van de Putte, 1986; Machida et al., 1982).

### V.3.5. Couplage transcriptionnel

Chez les IS bactériennes, un décalage de cadre de lecture peut se produire au niveau de la transcription. Nous parlons d'un mécanisme de glissement transcriptionnel qui pourrait être relativement fréquent chez les IS (Baranov et al., 2006, 2005).

Au cours de la transcription, il peut y avoir des séquences d'ADN homopolymériques comme AAAAAAAAA et l'ARN polymérase peut alors glisser. Lors de ce glissement, l'ARN polymérase peut insérer des nucléotides non prévus dans la chaîne d'ARNm en cours de transcription. Il en résulte une population d'ARNm dont la séquence se trouve en aval du site de glissement dans les trois phases de lecture possibles par rapport à la région en amont du site de glissement.

### V.3.6. Stabilité de l'ARNm

L'élément IS5 présente deux orfs, la première de 978 pb et la seconde de 324 pb, transcrites à partir des promoteurs  $P_{51}$  et  $P_{5S}$  (Rak et al., 1982), respectivement. Ces deux gènes sont transcrits dans des directions opposées comme dans le cas d'ARN antisens. RAK et al. ont démontré que ces deux orfs étaient transcrits et traduits. Il a été supposé que le petit orf pouvait avoir un contrôle négatif sur les modalités de transposition de l'élément. De plus, leurs résultats indiquent que l'ARNm de la transposase de l'IS5 est instable, ce qui pourrait avoir un impact sur l'efficacité de la transposase.

## V.4. Contrôle de la traduction des IS

### V.4.1. ARN antisens

La synthèse d'ARN anti-sens a été démontrée chez l'élément IS10. En effet, chez cet élément trois promoteurs sont présents (Simons et al., 1983) :

- $P_{IN}$  : promoteur du gène de la transposase d'IS10 qui est intrinsèquement faible ce qui contribue à la faible fréquence de la transposition d'IS10 *in vivo*,
- $P_{III}$  : promoteur faible
- $P_{OUT}$  : promoteur fort situé juste à l'opposé de  $P_{IN}$ , qui pourrait inhiber la transcription de  $P_{IN}$  en *cis*.

Il convient de préciser que le promoteur  $P_{OUT}$  produit un ARN court d'environ 70 nucléotides appelé « ARN-OUT » (Case et al., 1989). Cet ARN est très stable. L'ARN-OUT se replie en un « domaine tige » duplexé et surmonté d'un « domaine boucle » faiblement apparié. L'appariement est initié par l'interaction de la boucle ARN-OUT avec l'extrémité 5' de l'ARN de la transposase, ce qui permet de

séquestrer les signaux d'initiation de la traduction (Case et al., 1989) et bloquer la liaison avec le ribosome (Case et al., 1989).

Une configuration similaire a été identifiée chez l'élément IS30 de *Escherichia coli* (Arini et al., 1997). En effet, cet élément présente deux promoteurs. Le promoteur P<sub>30A</sub> qui contrôle l'expression du gène de la transposase et le promoteur P<sub>30C</sub>, localisé sur le brin opposé, qui code un transcrit de 150 nucléotides. Bien que le mécanisme ne soit pas entièrement décrypté comme pour l'élément IS10, le transcrit à partir du promoteur P<sub>30C</sub> est complémentaire à l'ARN de la transposase chez l'élément IS30 (Arini et al., 1997). De plus, l'augmentation de l'activité du promoteur P<sub>30C</sub> entraîne une diminution de la fréquence de transposition de l'élément IS30.

#### **V.4.2. ARNm sans leader**

Il a été montré que le promoteur de l'élément IS186 était un promoteur relativement fort. Cependant, le niveau de production de la transposase de cet élément est faible malgré ce promoteur fort (Keller, 1996). Ils ont montré que la traduction devait être initiée à l'extrémité 5' de l'ARNm bien que aucune séquence SD évidente ne soit présente. Ainsi, le taux de traduction de ce transcrit sans leader n'était que de 1,6 % comparativement à celui du gène *lac*, ce qui expliquerait la faible capacité de IS186 à transposer.

#### **V.4.3. Séquestration des signaux d'initiation de la traduction**

La séquestration des signaux d'initiation de la traduction dans une structure secondaire de l'ARNm a été observée chez les éléments IS10 et IS50 dont les séquences IR contiennent la SD et le codon START de traduction du gène de la transposase. Ainsi, il a été démontré que lors de la transcription par la présence d'un promoteur externe en amont de l'extrémité IR cela entraîne une séquestration des signaux d'initiation de la traduction issue de la formation de structure secondaire (Davis et al., 1985; Krebs and Reznikoff, 1986). Ceci a été suggéré chez l'élément IS10.

#### **V.4.4. Couplage traductionnel**

Le couplage traductionnel agit lors de l'élongation de la traduction par le ribosome. Ce mécanisme implique un décalage programmé des cadres de traduction entre deux orfs. Concrètement, il s'agit d'un décalage de -1 du cadre de lecture dans lequel le ribosome en cours de traduction glisse d'une base vers l'amont et reprend en phase alternative. En outre, pour effectuer ce glissement, il existe généralement la position des codons dits glissants dans une séquence heptanucléotidique de la sorte : X XXX XXY (X et Y présentent 2 bases nucléiques différentes). Lors du glissement du ribosome, celui-ci se lit comme suit : XXX XXY N (N présente un des quatre acides nucléiques).

Lors d'une analyse approfondie, il a été mis en évidence que la séquence heptanucléotidique <sup>5'</sup>A AAA AAG est couramment retrouvée chez les IS et démontrée comme signal de glissement chez

IS1 (Chandler and Fayet, 1993; Sekine and Ohtsubo, 1989). Néanmoins, il existe d'autres séquences de glissement, comme <sup>5</sup>C CCA AAG ou <sup>5</sup>G GGA AAG, qui sont fréquemment retrouvées chez les IS1 et IS3 (Sharma et al., 2014).

Dans la littérature, il existe de nombreux exemples dans lesquels le couplage traductionnel est suspecté (Siguier et al., 2015) :

- Chez de nombreux membres de la famille IS21, le codon de terminaison de *istA* et le codon d'initiation de *istB* se chevauchent.
- Chez l'élément IS66, le codon de terminaison de *tnpA* et le codon d'initiation de *tnpB* se chevauchent (**ATGA**).
- Chez l'élément IS1167, le cadre de lecture pourrait être réparti entre deux orfs consécutifs par le chevauchement des codons d'initiation et de terminaison (**ATGA**).
- Chez de nombreux membres de la famille IS605, les gènes *tnpA* et *tnpB* se chevauchent légèrement suggérant un couplage traductionnel.

## V.5. Les autres niveaux de régulation des IS

### V.5.1. Contrôles post-traductionnels

Dans la littérature, il existe très peu de données par rapport à des mécanismes de régulation post-traductionnels telle que la protéolyse. Chez l'élément IS903, il a été suggéré que la protéolyse pourrait être impliquée dans la génération de fragment régulateur de la transposase (Derbyshire et al., 1990). Enfin, l'activité de la transposase du bactériophage Mu est instable. Celui-ci nécessite la synthèse de la protéine tout au long du cycle lytique du bactériophage (Pato and Reich, 1984), la transposase étant également instable au cours de ce cycle (Gama et al., 1990). Gama et al. ont émis l'hypothèse selon laquelle l'inactivation fonctionnelle de la protéine n'est pas simplement le résultat de sa protéolyse, la possibilité d'une autre régulation post-traductionnelle est ouverte, telle que l'intervention d'un compétiteur (Gama et al., 1990).

### V.5.2. Rôle des facteurs de l'hôte

L'activité de transposition est fréquemment modulée par divers facteurs de l'hôte : soit par des protéines chaperons, soit par la réponse SOS, soit par des facteurs de l'ADN, soit par d'autres facteurs.

Différentes **protéines chaperons** ont été impliquées dans la régulation d'IS :

- La **protéine Hfq** est impliquée dans la régulation de la transposition chez le transposon Tn10 et l'élément IS10 en se fixant sur le site RBS de l'ARNm de la transposase (Ellis et al., 2015b; Ross et al., 2013).



- Les **IHF** (integration host factor) sont impliquées dans la régulation de la transposition de l'élément *IS1* et du transposon *Tn10* (*IS10*) en se fixant sur les sites de liaison IHF (Chalmers et al., 1998; Gamas et al., 1985; Signon and Kleckner, 1995).
- Les protéines de structuration du nucléoïde, de type **histone H-NS**, jouent également un rôle dans la mobilité de certaines IS. En effet, il a été démontré chez le transposon *Tn10* et *Tn5* que H-NS induit le dépliage du transpososome pour favoriser les échanges intermoléculaires (Ward et al., 2007; Wardle et al., 2009, 2005; Whitfield et al., 2009). Un tel mécanisme a également été suggéré pour l'assemblage du transpososome *IS1* (Shiga et al., 1999).

**La réponse SOS** pour « *Save Our Souls* » est une réponse cellulaire aux dommages à l'ADN. Le système SOS inhiberait la transposition du transposon *Tn5* et de l'élément *IS50* (Kuan and Tessman, 1991; Weinreich et al., 1991). Le promoteur  $P_{rcr1}$ , codant la transposase *rcr1* de l'élément *ISCR1*, est aussi sous la dépendance de la réponse SOS (Lallement, 2018).

**La protéine DnaA** de *Escherichia coli* permet l'initiation de la réplication chromosomique (Fuller and Kornberg, 1983; Funnell et al., 1986; Hirota et al., 1970). Chez le transposon *Tn5*, il a été démontré que la protéine DnaA était capable de reconnaître les extrémités extérieures du transposon (Fuller et al., 1984; Reznikoff, 1993) :

**L'activité polymérase, l'ADN gyrase** et l'activité exonucléase 5' à 3' de **l'ADN polymérase I** sont nécessaires à la transposition du transposon *Tn5* (Sasakawa et al., 1981; Isberg and Syvanen, 1982).

L'« **Acyl carrier protein** » (ACP) stimule le clivage de l'extrémité 3' du transposon chez le transposon *Tn3* (Maekawa et al., 1996).

**Le facteur GTP** (Guanosine TriPhosphate) était nécessaire pour la transposition de l'élément *IS903* et du transposon *Tn552* chez *Escherichia coli* (Coros et al., 2005; Twiss et al., 2005).

**Le facteur  $\sigma S$**  (RpoS) en phase stationnaire régule positivement la transposition du transposon *Tn4652* en phase stationnaire chez *Pseudomonas putida* (Ilves et al., 2001).

### V.5.3. Rôle des facteurs environnementaux

Il a été démontré chez de nombreuses IS que la température semble être un facteur clé de la transposition :

- Une température de 40°C favorise la transposition de l'élément *ISEcp1B* qui est associé au gène de résistance *bla<sub>CTX-M-2</sub>*, (Lartigue et al., 2006)
- Une température de 42°C réduit la transposition de l'élément *IS1* (Reif and Saedler, 1975),
- La transposition de l'élément *IS30* et le transposon *Tn3* sont optimaux à environ 30°C (Kretschmer and Cohen, 1979).

- La température à laquelle la fréquence de transposition la plus élevée a été observée varie entre la température ambiante et 30°C pour les transposons Tn21, Tn501, Tn1721, Tn2501 et Tn3926 (Turner et al., 1990).

Chez d'autres IS, la température joue sur les quantités de protéines. En effet, chez IS911, Haren et al. ont démontré que la production des protéines est fortement réduite à 42°C par rapport à 37°C (Haren et al., 1997). En revanche, la réduction de la température à 30°C entraîne une augmentation significative de la production.

De plus, il a été mis en évidence que ces changements de température impactaient l'état de méthylation du transposon Tam3 (Hashida et al., 2003). En effet, une température plus élevée entraînait une hyperméthylation, tandis qu'une température plus basse entraînait une méthylation réduite.

Enfin, il a été démontré chez deux éléments (ISDra2 et Tn1) que les irradiations augmentaient significativement la fréquence de transposition de ces éléments (Eichenbaum and Livneh, 1998; Menecier et al., 2006). Ceci peut être expliqué par les dommages à l'ADN causés par les irradiations.

Il ressort de ce qui précède que :

- L'activité de transposition est modulée par différents facteurs au niveau transcriptionnel ou traductionnel de la transposase.
- La transposition est régulée par des ARN antisens ou des protéines.
- L'activité de transposition est fréquemment modulée par divers facteurs de l'hôte.
- Les facteurs environnementaux impactent l'activité de transposition, telles que la température ou l'irradiation...

La régulation peut donc intervenir soit par l'élément lui-même, soit par l'hôte, soit par l'environnement.

## Objectifs de thèse

---

Bien qu'ils soient très fréquemment retrouvés à proximité de gènes de virulence et/ou de gènes de résistances aux antibiotiques, les membres de la famille IS91/ISCR demeurent très peu caractérisés notamment ne ce qui concerne les modalités de leur transposition, l'expression et la fonctionnalité de leur transposase. Les derniers travaux datent du début des années 2000 pour l'élément IS91, paradigme de la famille IS91/ISCR. De plus, contrairement aux autres éléments de la famille, IS91 présente une *orf* supplémentaire, appelée *orf121*, dont le rôle n'était pas connu au début de ce travail.

C'est pourquoi, afin de mieux appréhender les fonctionnalités des éléments de la famille IS91/ISCR, nous avons défini les objectifs suivants :

- Étude *in silico* de la diversité des éléments (IS91, IS801, IS1294b) et du rôle d'*orf121* dans la mobilité de l'élément IS91,
- Étude structurale de la transposase IS91,
- Étude de la transposition d'autres éléments de la famille IS91 : ISCR et ISKpn22.

## **Partie I. Étude *in silico* de la diversité des éléments (IS91, IS801, IS1294b) et du rôle d'*orf121* dans la mobilité de l'élément IS91**

---

Jusqu'à présent, aucune analyse des bases de données n'avait été réalisée pour identifier les différentes isoformes et les sites cibles d'insertion des éléments de la famille IS91, actifs en transposition. Ainsi, nous avons répertorié et analysé la diversité des éléments IS91, IS801 et IS1294b à partir de données de la GenBank®.

Concernant l'ensemble des éléments de la famille IS91, seul l'élément IS91 possède un gène supplémentaire, appelé *orf121* dont le codon de terminaison de l'*orf121* chevauche le codon d'initiation du gène de la transposase, *tnpA* (<sup>5'</sup>**TGATG**). Ainsi, ces deux gènes auraient la possibilité d'être couplés sur le plan transcriptionnel et traductionnel par le chevauchement d'un seul nucléotide.

Mendiola et al. ont mené des expériences de mutagenèse dans le gène *tnpA*. Ils ont montré que les éléments mutés étaient non transposables (Mendiola et al., 1992) et qu'il existait une différence significative dans la spécificité d'insertion en l'absence de *orf121*. En revanche, lorsque *orf121* était muté, la fréquence de transposition n'était pas significativement différente. Ils ont donc conclu que le gène *orf121* n'affectait pas la fréquence de transposition mais qu'il modifiait la spécificité de la cible (Garcillán-Barcia et al., 2002).

De ces résultats, il ressort que *orf121* est dispensable pour la transposition. Il convient toutefois de souligner qu'aucune fonction n'a été attribuée à la protéine Orf121 dans le processus de transposition de IS91.

Nous avons donc réalisé un travail expérimental afin de caractériser les interactions possibles entre Orf121 et TnpA.

Ce travail a fait l'objet d'un manuscrit qui sera soumis prochainement.

### **I.1. L'*orf121* codé par IS91 régule négativement la transposition d'IS91**

## IS91-encoded *orf121* regulates negatively IS91 transposition

Aurélien Fauconnier<sup>1</sup>, Sandra Da Re<sup>1</sup>, Margaux Gaschet<sup>1</sup>, Thomas Jové<sup>1a</sup>, Marie-Cécile Ploy<sup>1¶\*</sup>,  
Cécile Pasternak<sup>1¶†</sup>

<sup>1</sup>Univ. Limoges, INSERM, CHU Limoges, RESINFIT, U 1092, F-87000 Limoges, France

a Present address: no longer employed by INSERM.

\*Corresponding author

E-mail: [marie-cecile.ploy@unilim.fr](mailto:marie-cecile.ploy@unilim.fr)

¶ These authors participated equally in the work

**Keywords:** DNA transposition, bacterial insertion sequences, IS91 family, rolling circle replication, translational coupling

**Short title:** *orf121*-mediated inhibition of IS91 transposition

## INTRODUCTION

Mobile genetic elements (MGEs) are major contributors to the spread of bacterial multidrug resistance bacterial strains, posing a global threat to human, animal and environmental health (for recent reviews, see [1,2]). They include genetic elements able to transfer between bacterial cells (such as transmissible plasmids and integrative conjugative elements), and transposable elements (TE) involved in intracellular mobility (such as insertion sequences (IS), transposons and gene cassettes from integrons). TE act by both promoting dissemination of and activating neighboring antibiotic resistance genes (ARG). The *IS91* family represents an atypical IS family for which only four members have been shown active in transposition to date: the canonical *IS91* element [3], *IS801* [4], *IS1294* [5] and *IS1294b*, which is very close to *IS1294* (95% identity) [6]. These IS are usually associated with virulence and antibiotic resistance genes. They encode a transposase belonging to the HUH superfamily of single-strand nucleases [7], and are presumed to transpose by a rolling-circle (RC) mechanism [8], but their contribution to genes dissemination is unclear.

In place of terminal inverted repeats (IR) classically found at IS boundaries and framing the transposase gene *tnpA*, the *IS91* family displays two functionally distinct ends: the *oriIS* sequence where transposition initiates and the *terIS* end where transposition terminates [9]. Nevertheless, for *IS91* and *IS1294*, termination fails at a frequency of about 1 to 10% [5,9] and from 1.7% to 14% for *IS1294b* [6], leading to the capture of DNA fragments adjacent to the *terIS* end, a process known as one-ended transposition (OET). Therefore, *IS91*-like elements may play a significant role in the dissemination of bacterial virulence or antibiotic resistance genes adjacent to *IS91* elements [9–11]. *IS91* family members do not generate target site duplications (TSD) upon insertion, but display target sequence specificity. The recognized target for *IS91*, *IS801*, *IS1294* and *IS1294b* insert is a tetranucleotide, 5'-CTTG or -5'-GTTC [5,6,12,13].

The *oriIS* side of *IS801* and *IS91* contain two imperfect subterminal palindromes, whereas *IS1294* and *IS1294b* have only one presumed palindrome (figure 1). The *terIS* side contains a single

imperfect sub-terminal palindrome flanked by the cleavage site, which differs between IS91 (5'-CTCG) and the other three elements (5'-GTTC).

Previous studies showed the *in vivo* formation of single-stranded (ss) and double-stranded (ds) IS91 circular intermediates during the IS91 transposition. Both ssDNA and dsDNA carried a junction in which *ori91* and *ter91* were abutted [14]. However, it is not yet clear which of these intermediates are functional in the IS91 transposition pathway.

Unlike all other IS91 family members, IS91 carries, upstream the *tnpA* transposase gene, a second *orf* potentially encoding a 121 amino acids polypeptide, named *orf121*. The termination codon of *orf121* overlaps the initiation codon of the *tnpA* gene, suggesting that the expression of both proteins may be coupled. Mutations in *orf121* did not affect transposition frequency but appeared to only alter the choice of the insertion target [15]. To date, no function has been attributed to the *orf121* protein in the transposition process of IS91.

Herein, we studied whether *orf121* product plays a role in the transposition of IS91 into the *E. coli* host. Using a genetic system based on mating-out procedure, we showed that the induction of the expression of *orf121* has an inhibitory effect on the *in vivo* transposition of IS91. We demonstrated that the expression of the two genes *orf121* and *tnpA* were coupled, and that the overlap between the two genes is necessary for the transcription of *tnpA*. We also showed that only the ssDNA circular intermediates can be inserted into a new target sequence and that *orf121* has no impact on the choice of the target site. In addition, we showed that induction of *orf121* expression strongly decreases the OET rate, suggesting that the ORF121 protein may be required for accurate recognition and cleaving of the *terIS91* end.

## RESULTS

### ***orf121* is present and highly conserved in the IS91 element**

When we initiated this work, no database analysis was available to assess the diversity of the different isoforms and the target insertion sites of the active members for transposition of the IS91 family. For this analysis, we have chosen to study the elements IS91, IS801 and IS1294b because the elements IS1294 and IS1294b are very close and IS1294b has been better characterized. We performed an *in silico* analysis and identified 134 isoforms, 225 isoforms and 35 different isoforms for IS91, IS1294b and IS801, respectively.

These different isoforms were mainly found in animals (48.8%) and Homo sapiens (23.4%) for IS91, although any origin was mentioned for 25.5% of the isoforms (table S1). IS1294b was found in animals (14.5%) and Homo sapiens (26.2%) and the origin was unknown in 52.4% of cases. For IS801, when the origin was mentioned, it was found exclusively in plants.

IS91 and IS1294b were found predominantly in Enterobacterales (99.6 % and 100 %, for IS91 and IS1294b, respectively) with a majority in *Escherichia coli* (75% and 52.8%, for IS91 and IS1294b, respectively) followed by *Salmonella enterica* for IS91 and *Shigella spp* for IS1294b (table S2). IS801 was found exclusively in *Pseudomonas* species.

IS91 and IS1294b were mainly found on plasmids and mostly in a single copy (72.62% and 65.77%, respectively) (tables S3 and S4). More than half of the plasmids identified were of the *IncFII* type (table S5).

Target insertion sites for the different isoforms showed similar profiles, with a preference for the two sites 5'-GTTC or 5'-CTTG, followed by 5'-CTCG for IS91 and IS1294b, or 5'-GTCC for IS801 (table S6). All the target sites always have a T in the second position. The cleavage site tetranucleotides are in the majority 5'-CTCG for IS91 and 5'-GTTC for IS1294b and IS801, as previously described.

Analysis of the different isoforms of IS91 showed that the overlap between the *orf121* and *tnpA* genes was well conserved, being present in 79% of cases. For the remaining isoforms, the loss of the overlap



was due to either insertions or deletions in the homo-polymeric regions of the *orf121* sequence (figure S1).

### **Both *orf121* and *tnpA* genes are expressed from the $P_{orf121}$ promoter**

We first investigated the transcriptional expression level of *orf121* and *tnpA* in their natural configuration within the IS91 element (table 1). Using BPRM (<http://www.softberry.com>; [16]) we identified putative  $\sigma^{70}$  promoters and Shine-Dalgarno (SD) sequences for *orf121* and *tnpA* (see figure 2). The activity of the two predicted promoters  $P_{tnpA}$  and  $P_{orf121}$  was determined by using transcriptional fusions to *lacZ* with the *lacZ* ribosome binding site (RBS) (plasmids pCP65 and pCP66, Table 2). The resulting  $\beta$ -galactosidase activities indicated that  $P_{orf121}$  is functional and stronger than  $P_{tnpA}$  (62-fold increase) (figure 3a).

To better understand the expression of the two genes *orf121* and *tnpA* in a natural configuration and the importance of the overlap between the stop codon of *orf121* and the start codon of *tnpA*, we measured the  $\beta$ -galactosidase activity of four constructs with either one ( $P_{tnpA}$ ) or two of the predicted promoters  $P_{orf121}::orf121_{(P_{tnpA})}$  in the presence or absence of the overlap (Table 2). In these conditions, the promoters  $P_{orf121}$  and  $P_{tnpA}$  are associated with their own native RBS sequence. The resulting  $\beta$ -galactosidase activities were lower compared to the ones with the RBS of *lacZ* (figure 3b). The activity of  $P_{tnpA}$  was very low ( $< 1$  MU) and it increased in the presence of *orf121* ( $P_{orf121}::orf121_{(P_{tnpA})}$ ) in both conditions with or without the overlap *orf121-tnpA*. Nevertheless, in the absence of the overlap, the activity of  $P_{orf121}::orf121_{(P_{tnpA})\Delta-1}$  decreased by a 2.28-fold factor.

Finally, to better understand the role of each promoter, we constructed mutants of  $P_{orf121}$  or  $P_{tnpA}$  in the presence of the overlap. We introduced mutations in *orf121* without impacting the amino acid sequence. The mutations were in the -10 box for  $P_{orf121}$  (5'-TATAAA to 5'-CGCGAA) and in the -35 box of  $P_{tnpA}$  (5'-TTGCCG to 5'-TGGCGG). The activity drastically decreased 14.76-fold when the  $P_{orf121}$  was mutated while the decreased, although significant, was less important when  $P_{tnpA}$  was mutated (2.46-fold) (figure 3c).

Altogether, these results suggest that the transcription of the *tnpA* gene arises mainly from the  $P_{orf121}$  promoter. Furthermore, the overlap seems necessary for the coupled transcription of *orf121* and *tnpA* genes.

### **The *orf121* product down-regulates IS91 transposition into *E. coli*.**

We then investigated whether *orf121* might play a regulatory role in the *in vivo* transposition of IS91. For this purpose, we used a genetic system based on a mating-out assay in *E. coli* to measure the mobility of IS91 derivatives into the pOX38Km conjugative plasmid (see figure S2 and Materials and Methods). We compared the transposition frequency of IS91 derivatives cloned into the high-copy number pUC18 plasmid in *E. coli* strains expressing *in trans* *tnpA* alone (plasmid pCP4) or *orf121-tnpA* (plasmid pCP40), under the control of the inducible  $P_{Lac}$  promoter on a compatible plasmid. We used two different IS91 derivatives: IS91-*ter91-cm<sup>R</sup>* (pCP7 plasmid), which contains a  $Cm^R$  cassette flanked by *ori91* and *ter91* ends, and IS91-*ter91::orf121-cm<sup>R</sup>* (pCP39 plasmid) which additionally carries *orf121* under the control of its own promoter (see figure S2, figure S3 and table 1). For both IS91 derivatives, the IS91 sequence includes the last 82 bp of the IS91 right end, including the *ori91*, that have been previously shown to be sufficient for a functional transposition [8] and the *ter91* sequence corresponds to the first 89 bp including the *ter91* cleavage tetranucleotide (5'-CTCG). Adjacent to the *ori91* end, the IS91 donor plasmids contain the 5'-CTTG tetranucleotide necessary for cleaving and transposition. We observed a high decrease (5402-fold) in transposition frequency with IS91-*ter91-cm<sup>R</sup>* when cells expressed both *orf121* and *tnpA*, compared to *tnpA* alone ( $2.76 \times 10^{-3}$  and  $5.12 \times 10^{-7}$ , respectively) (figure 4.a). With cells expressing *tnpA* alone with IS91-*ter91::orf121-cm<sup>R</sup>* where *orf121* is expressed *in cis*, the transposition frequency was also reduced (1.2-fold) compared to IS91-*ter91-cm<sup>R</sup>*. In contrast, with basal *in cis* expression of *orf121*, we observed a significant decrease (11.1-fold) in transposition frequency compared with an IS91-*ter91-cm<sup>R</sup>*. These results indicate that co-expressing *orf121* (*in cis* or *in trans*), in addition to the transposase, significantly decreases the IS91 transposition frequency into the *E. coli* host.

We then compared the frequency of transposition of the *IS91-ter91-cm<sup>R</sup>* derivative in *E. coli* test strains expressing *in trans* *tnpA* and *orf121* under the control of two independent and inducible promoters,  $P_{Lac}$  and  $P_{LtetO-1}$  promoters, respectively (Figure 4b). In cells expressing *orf121::P<sub>LtetO-1</sub>-P<sub>Lac</sub>::tnpA* (from pAF26 plasmid) with *IS91-ter91-cm<sup>R</sup>*, the basal transposition frequency was  $2.38 \times 10^{-5}$  in the absence of IPTG and aTc, where both  $P_{Lac}$  and  $P_{LtetO-1}$  promoters were not expressed. Depressing  $P_{LtetO-1}$  by adding anhydrotetracycline (aTc) significantly decreased the transposition frequency by 2.3-fold compared with the condition without aTc, with a frequency of  $1.05 \times 10^{-5}$ . In the presence of IPTG and absence of aTc, *tnpA* is expressed and the transposition frequency was  $1.07 \times 10^{-3}$ , representing a 45-fold increase compared with the condition without inductor. When both *orf121* and *tnpA* were expressed (presence of IPTG and aTc), the transposition frequency was  $1.25 \times 10^{-4}$ , representing an 8.6-fold reduction compared with the condition when only *tnpA* is expressed (IPTG-only condition). These results confirm that the negative effect observed on *IS91* transposition is due to the expression of *orf121* expression itself and not to the specific configuration *orf121-tnpA*.

In parallel, we quantified *orf121* and *tnpA* transcripts (Figure 5). We showed that, when *orf121* is expressed either *in trans* (*IS91-ter91-cm<sup>R</sup>*) or *in cis* (*IS91-ter91::orf121-cm<sup>R</sup>*) conditions, the mRNA copy number of *tnpA* significantly decreased 6.93-fold (from  $5.84 \times 10^2$  to  $8.46 \times 10^1$ ) and 2-fold (from  $5.84 \times 10^2$  to  $2.92 \times 10^2$ ), respectively. In both conditions, *in trans* or *in cis*, the mRNA copy number of *orf121* were  $2.21 \times 10^5$  and  $2.50 \times 10^2$ , respectively.

These results suggest that *orf121* negatively influences *tnpA* transcript levels. We observed the same trends when both genes are expressed from independent inducible promoters (*orf121::P<sub>LtetO-1</sub>-P<sub>Lac</sub>::tnpA*) with *IS91-ter91-cm<sup>R</sup>*, the *tnpA* mRNA copy number was significantly reduced by 1.78-fold when *orf121* was expressed ( $1.20 \times 10^3$  vs  $6.75 \times 10^2$ ). In the same construction, the mRNA copy number of *orf121* was significantly increased by 1.26-fold in the presence of both inducers than in the presence of aTc-only condition, the numbers copies improve from  $2.65 \times 10^1$  to  $3.33 \times 10^1$  (p-value = 0.008).

We then quantified TnpA and Orf121 products. We used a His-tagged TnpA or ORF121 proteins (either a His6-tag in the N-ter of Orf121 and/or in the C-ter of TnpA) in mating-out conditions and performed Western-immunoblotting. Orf121 was detected only in the soluble part and TnpA was found in the inclusion bodies. The presence of Orf121 and TnpA proteins was confirmed by mass spectrometry (see figure S4).

We first checked that the addition of the His-tags did not affect the transposition frequencies (see figure S5). We observed that when *orf121* and *tnpA* are under the control of a same promoter  $P_{Lac}::orf121tnpA^*$  (plasmid pAF2) the quantity of TnpA decreased by a 4.1-fold factor (4.88 AU) (figure 6). Interestingly, the quantity of TnpA was similar, at around 20 AU (Arbitrary Units), with  $orf121^*::P_{LietO-1}-P_{Lac}::tnpA^*$  (plasmid pAF34,\* indicate localization of His-tag) in the presence of IPTG and whether or not aTc, and  $P_{Lac}::tnpA^*$  (plasmid pAF4) in the presence of IPTG (figure 6). The production of Orf121 was similar whatever the conditions of *orf121* expression. These results showed the lower production of the TnpA protein in the presence of the Orf121 protein, suggesting that the regulation of the coupled expression of *orf121* and *tnpA* is at translational level.

### **Effect of Orf121 on *in vivo* insertion of IS91 circular intermediate into *E. coli*.**

Previously, it was demonstrated that two new DNA species were formed *in vivo* after induction of transposase expression in an independent manner; i.e circular intermediates of single-stranded (ss) and double-stranded (ds) IS91 when *ori91* and *ter91* were fused [17]. However, we do not know which intermediates are functional for IS91 insertion. We used a suicide mating assay (see figure S3; figure S6 and Materials and Methods). The *ter91-ori91* junction present in the substrates (from pCP54 or pCP55 plasmids) used for the insertion assays is shown in Figure 7a.

To test the double-stranded or single-stranded circular intermediates, we performed either a or filter conjugations as described in materials and methods. The positive transformation control pUC18 had a transformation frequency of  $2.96 \times 10^6$ . However, whatever the double-stranded circular

intermediate, the Cm<sup>R</sup> transformant frequency was undetectable ( $< 2.63 \times 10^{-9}$ ) (Figure 7b) suggesting that the dsDNA is not recognized. For experiments with ssDNA, we first observed that, as expected, when we used donor cells carrying the empty plasmid pSW23T (no *ori91-ter91* junction) and recipient cells carrying P<sub>Lac</sub>::*tnpA* or P<sub>Lac</sub>::*orf121tnpA* (from plasmid pCP4, pCP40), the frequency of Cm<sup>R</sup> exconjugants was undetectable ( $< 2,87 \times 10^{-9}$ ) (Figure 7c). This frequency was also undetectable when the donors carried the DNA junction in the orientation *pter91::ori91*. However, when donor cells carried the *pori91::ter91* junction and recipient cells expressed *tnpA* only (P<sub>Lac</sub>::*tnpA*), the frequency of transposition was  $3.33 \times 10^{-5}$ . When recipient cells expressed both *orf121* and *tnpA* (P<sub>Lac</sub>::*orf121tnpA*), the transposition frequency decreased by 11592-fold ( $< 2,87 \times 10^{-9}$ ). We performed the same experiment but with recipient cells expressing *tnpA* and *orf121* under the control of two inducible promoters (*orf121::P<sub>LtetO-1</sub>-P<sub>Lac</sub>::tnpA*). We confirmed that the expression of *orf121* decreased the transposition frequency with the *pori91::ter91* junction with a 3.6-fold reduction (Figure 7d). Molecular analysis by arbitrarily-primed PCR (AP-PCR), of dozens of exconjugants clones, as described in the materials and methods, enabled us to prove that all Cm<sup>R</sup> exconjugants result from insertion of the *pori91::ter91* junction into the chromosome of the recipient cell (see table S7).

These results suggest that only the circular intermediate is activated for insertion and that the strand used is the bottom strand, and that Orf121 negatively impacts transposition through the ability of TnpA to insert the circular intermediate.

### **Orf121 is not involved in the target selection and insertion specificity of IS91.**

It has been previously shown by analysing 18 transposition products that mutations in *orf121* did not affect transposition, but appeared to alter the choice of target [15]. To prove that all the Cm<sup>R</sup> exconjugants we obtained resulted from the transposition of the IS91 derivative into the pOX38Km plasmid, hundreds of transposition products were molecularly analyzed by AP-PCR. The target insertion sites were also examined for every condition. The results showed that expression of *orf121*

*in cis* (pCP39) or *in trans* (pCP40) has a target site specificity similar to that of *tnpA* alone (pCP4) (table 2a) (p value = 0.45). We did not obtain any significant difference in specificity when *orf121* and *tnpA* were under the control of two different promoters table 2b) (p value = 0.29). These results are in agreement with previous results suggesting that the Orf121 element is not involved in the target selection and insertion specificity of IS91.

In addition, we analysed by AP-PCR the tetranucleotides of the cleavage sites on the *ter91* side of a few dozen transposition events. Cleavage sites were examined only in the condition with or without *orf121 in trans*. We did not observe any change in the cleavage site (5'-CTCG) in the presence or absence of Orf121 (table 2c).

### **The *orf121* product is important for accurate recognition and cleaving of the *terIS91* end.**

Previous studies have shown in IS91 that termination at *terIS* fails at a frequency of around 1 to 10% [9], allowing the capture of DNA fragments adjacent to the *terIS* end through OET. An analysis of the percentage of OETs in the presence or absence of Orf121 was carried out using the different mating-out assays described above.

When cells expressing  $P_{Lac}::tnpA$  with IS91-*ter91-cm<sup>R</sup>*, the percentage of OET was 53% (table 3a). In contrast, when *orf121* was expressed *in cis*, the percentage of OET was 28%, representing a 1.89-fold reduction compared to the condition without *orf121*. However, when *orf121* was expressed *in trans*, the percentage of OET was 18%, representing a 2.94-fold reduction compared with the condition without *orf121* and a 1.56-fold reduction than when *orf121* was expressed *in cis*. We then tested cells expressing IS91-*ter91-cm<sup>R</sup>* and *orf121::P<sub>LtetO-1</sub>-P<sub>Lac</sub>::tnpA* allowing control of *orf121* and *tnpA* expression (table 3b). Depression of  $P_{LtetO-1}$  by the addition of aTc significantly decreased the percentage of OET by 2.64-fold compared to the condition without aTc, with a percentage of 11%. When both *orf121* and *tnpA* were expressed (presence of IPTG and aTc), the percentage of OET was 32%, a 1.5-fold reduction compared to the condition where only *tnpA* was expressed (IPTG-only). These results suggest that Orf121 may be required to accurately recognize and cleave the *terIS* end.

## DISCUSSION

Currently, *IS91* is the prototype for the *IS91* family and a related family, the ISCRs, both of which are adjacent to virulence determinants and antibiotic resistance genes [2,9,10,18–20]. These elements therefore play a predominant role in the dissemination and expression of antibiotic resistance genes [10,21–23]. Despite these factors, the *IS91* family remains poorly characterized and there is a need to better understand the regulation of their transposition.

We first performed an *in silico* analysis that revealed a large number of different isoforms for *IS91* and *IS1294b*. In contrast, we detected few isoforms for the *IS801* element, and mainly in plants in *Pseudomonas* spp isolates. When the genetic support was available, we found that *IS91* and *IS1294b* were mainly on IncFII-type plasmids. IncFII-type plasmids have a narrow host spectrum restricted to *Enterobacteriales* and generally associated with the propagation of antibiotic resistance [24–28]. This is in agreement with our data as 99.8 % of the *IS91* and *IS1294b* isoforms were found in *Enterobacteriales*. Furthermore, *IS91* and *IS1294b* have been often described associated with or close to antibiotic resistance genes [6,29–39] suggesting that these IS might play a role in antibiotic resistance mobilization. We found that *IS91* and *IS1294b* are generally found in a single copy, suggesting that the mobilization cannot occur via the formation of composite transposons. Dissemination of antibiotic resistance might then be linked to the IncFII-type plasmids dissemination among *Enterobacteriales*. This suggests that the diversity of antibiotic-resistance genes associated with these IS would be limited. We did not analyze the genetic environment of the detected isoforms. In the future, it will be useful to analyze this environment to understand better the diversity of genes that could be mobilized. The one-ended transposition (OET) is the other possible mechanism of antibiotic resistance gene mobilization with a transposition failing to terminate at *terIS* and mobilizing genes adjacent to the *terIS* sequence.

We thus analysed the OET capacity of *IS91*. We showed that termination at *terIS* fails at a frequency of approximately 50%, suggesting that members of the *IS91* family could indeed be involved in gene

propagation, potentially including antibiotic-resistance genes and virulence genes. It has been shown that IS91-type elements can mobilize up to 40 kb, in the case of IS801 [40]. Previous studies with elements of the IS91 family have shown that *terIS* termination fails at a frequency of around 1-14% [5,6,9], allowing capture of DNA fragments adjacent to the *terIS*. However, it should be noted that our method of analysis is different. Indeed, in previous studies involving IS91 or IS1294, all the genes were on the same plasmid. Indeed, the analysis of IS1294*b* was based on the sucrose counter-selection methodology using a strategy based on the inactivation of the *sacB* gene presenting only few insertion sites. In this study, we used two plasmids, the pOX38Km plasmid as a transposition reporter plasmid with more than 1000 possible insertion sites, and IS91-derived plasmids inserted into a high copy number plasmid (pUC18); therefore, the IS91-derivative is not limiting unlike the transposase which is expressed from a p15A-derived plasmid. enabling more transposition events to be captured. Furthermore, under our *in vivo* analysis conditions, the OET mechanism could therefore be more frequent than previously thought.

Finally, the *in silico* analysis showed most IS91 isoforms had the overlap between the STOP codon of *orf121* and the START codon of *tnpA*, suggesting an important role of this configuration in the regulation of the genes expression. We first showed that both promoters were active but the P<sub>*orf121*</sub> promoter was significantly stronger than the P<sub>*tnpA*</sub> promoter (62-fold). In the native configuration, with the termination codon of *orf121* overlapping the initiation codon of *tnpA*, we showed that both *orf121* and *tnpA* are mainly transcribed from the P<sub>*orf121*</sub> intrinsic promoter suggesting a transcriptional coupling for the IS91 transposase expression. This has been previously shown for IS through transcriptional slippage [41,42]. Indeed, the *orf121* sequence shows three potential RNA polymerase slippage regions (5'-T<sub>90</sub>TTTTTT<sub>96</sub>; 5'-A<sub>192</sub>AACAAAAA<sub>200</sub> and 5'-C<sub>351</sub>CCCCGCCC<sub>360</sub>). Such a transcriptional coupling mechanism generates a translation coupling mechanism known as “termination-reinitiation”. This consists of a -1 shift in the reading frame in which the ribosome undergoing translation slips one base upstream and resumes in the alternative phase so that the downstream gene can be translated, as shown previously for transposition in bacteria and other IS



[43–46]. The *in silico* analysis is also in favor of such a regulation. Indeed, a number of the isoforms (21%) described have deletions or insertions in the homopolymeric zones of *orf121*, resulting in a truncated *orf121* gene or a loss of the *orf121-tnpA* overlap, that could modify the level of *tnpA* expression.

Although we do not know the underlying mechanism of this regulation of *orf121*, we can assume a direct interaction between the two proteins (Orf121-TnpA) in this regulation. However, according to the effect of *orf121* expression on *tnpA* mRNA copy number, we cannot exclude mRNA-mRNA or mRNA-protein interactions or a role of *orf121* in stabilizing the *tnpA* mRNA.

Western-blot analysis showed a significantly lower amount of TnpA in a natural configuration than when *tnpA* gene alone is expressed. These results may suggest a sequestration of translation initiation signals resulting in a decrease in protein quantity. Such a hypothesis has been suggested in IS10 and IS50 [50,51], for which it has been assumed that they generate secondary structures in the mRNA that sequester translation initiation signals of the *tnpA* gene.

Contrary to the results of Bernales et al, 1999 [15], our results showed that the Orf121 product does not contribute to the target selection and insertion specificity of IS91, neither alters the cleavage site 5'-CTCG. These data are consistent with the results obtained from the *in silico* analysis.

Previous studies on the transposition of the IS91 family concerned only the IS91 and IS1294 elements, for which two models of the CR mechanism have been proposed. The first, known as the “concerted transposition model” or CR with coordinated strand transfer, was developed by Mendiola et al in 1994 [8]. In this model, the transposase is dimeric and forms an initial complex that brings together the donor and recipient DNAs, whose single-strand breaks occur at the same time. The second model, known as the “sequential transposition model” or “peel and paste”, has been described for IS91, IS1294 [5,17]. This second model appears to involve a single strand circular intermediate, as proposed for the replication of bacteriophage  $\phi$ X174 (for recent reviews, see [52]). Previous studies showed the *in vivo* formation of single-stranded (ss) and double-stranded (ds) IS91 circular intermediates during the IS91 transposition [17]. Our results provide strong evidence that only the

bottom DNA single- strand can be inserted into a target site in the orientation *oriIS-terIS*, suggesting that IS91 performs transposition based on the “peel and paste” model, in which only circular intermediates are functional in IS91 transposition as shown figure 8. Transfer of such an intermediate is mediated and initiated by the transposase, through the recognition of the target insertion site. As in the IS200/IS605 family, we can assume that the target site is not recognized directly by TnpA but by 4 bases at the foot of the hairpin in the transposition circle which undergo unconventional base pairing with the target sequence of the ssDNA [53]. Cleavage of the recipient DNA then occurs at the 3' end of the insertion target site, and at the tetranucleotide 5'-CTCG at the junction of the circular intermediate *oriIS91::terIS91* (figure 8.f). This cleavage allows the formation of two ends with 3'OH groups on the cleaved circular intermediate and the target DNA. This 3'OH group is capable of attacking the 5'phosphotyrosine bond on the opposite DNA strand, allowing the IS to be inserted into the recipient DNA strand (figure 8.g). This IS integration is completed by the bacterial replication machinery (figure 8.h).

In this work, we showed that the IS91 family has conserved features and that the IS91-specific *orf121* has a pleiotropic role: i) *orf121* negatively influences *tnpA* transcription levels; ii) co-expression of *orf121* (*in cis* or *in trans*), in addition to the transposase, strongly decreased the frequency of IS91 transposition; iii) the Orf121 protein has a negative effect on the insertion step; iv) *orf121* improves recognition and precise cleavage of the *terIS* as we observed less OET when *orf121* is expressed. Furthermore our results gained insights into the mechanism of transposition of IS91. Further experiments on Orf121-TnpA interactions and the “peel and paste” transposition mechanism will allow to continue to decipher the biological relevance+ of this IS91 family. IS91 elements were often described associated to antibiotic-resistance genes. The peculiarities of the regulation and mechanisms of IS91-family transposition would better explain their role in resistance dissemination.

## MATERIALS AND METHODS

### Bacterial strains, media and growth conditions.

Bacterial strains and plasmids are listed in table 1. Details of the plasmid constructs are available on request; see figure S3, table S8 and the extended experimental procedures for schematics of the plasmids used and further information. *E. coli* strains were grown in Lysogeny Broth (LB) at 37°C. Media were supplemented with the appropriate antibiotics used at the following concentrations: spectinomycin 50 µg/mL, kanamycin 25 µg/mL, chloramphenicol 25 µg/mL, nalidixic acid 25 µg/mL, and ampicillin 100 µg/mL. Expression of IS91 *tnpA* alone, or co-expression with *orf121* from the  $P_{lac}$  promoter was induced by adding 0.5 mM IPTG in the media. Expression of *orf121* from the  $P_{LTet-O1}$  promoter was induced by adding 50 nM anhydrotetracycline in the media. Transformation of *E. coli* with plasmid DNA was performed as previously described [54]. All plasmid constructions were verified by DNA sequencing.

### DNA manipulations.

PCR reactions were carried out using Phusion DNA Polymerase (ThermoScientific) to amplify the fragments subsequently used for cloning, or Quick-Load® Taq 2X Master Mix (NEB) for all other applications. All PCR reactions were purified using MACHEREY-NAGEL's NucleoSpin Gel and PCR Clean-up according to the manufacturer's recommendations. Insertion sites of the mini-IS91 into pOX38Km reporter plasmid or *E. coli* *DH5a* genome were mapped by arbitrary priming (AP) PCR. The first PCR round was performed in a final volume of 50 µl with 15 µl from single Cm<sup>R</sup> lysed colonies as a template. The arbitrary primer (ARB1) at a final concentration of 0.8 µM was paired with an Cm<sup>R</sup>-specific primer (Cm-1005) at a final concentration of 0.8 µM and PCR was performed as follows: 2 min 95°C, 6 cycles of 30 s 95°C, 30 s 30°C, 1 min 30 s 72°C; 30 cycles of 30 s 95°C, 30 s 45°C, 2 min 72°C; and finally 72°C for 5 min. The second cycle was performed in a final volume of 50 µl with 5 µl of the purified PCR product from the first cycle as a template. A second arbitrary

primer (ARB2) was paired with primer Cm-1049, each at a final concentration of 0.8  $\mu$ M and PCR was performed as follows: 2 min 95°C, 30 cycles (30 s 95°C, 30 s 57°C, 2 min 72°C); 72°C for 5 min. The PCR products were directly sequenced using the Ori91-1104 primer. To analyse the *terIS* cleavage site, the same procedure was carried out with primer pairs ARB1/Cm-473 for the first cycle of PCR and ARB2/Ter91-337 for the second cycle. The products of this PCR were directly sequenced with the Ter91-276 primer.

Oligonucleotides used for PCR amplification of DNA fragments required for plasmid construction, diagnostic PCR or sequencing are described in table S8.

### ***In silico* analysis of IS91 DNA sequences from the public database.**

The amino acid sequence of the TnpA transposase encoded by IS91 (accession number CAD30457.1), IS801 (accession number CAA40540.1) and IS1294b (accession number AIY68271.1) was blasted with BLASTp (NCBI). For the methodology, see Figure S7. The matching sequences were filtered to retain only those sequences for which the level of amino acid identity (equal or) was greater than 95% for IS91 and IS801 (including truncated TnpA). The threshold for IS1294b was set at (equal or) greater than 98%. Corresponding nucleotide sequences in which the IS and TnpA were partial or truncated were discarded only for IS91 and IS1294b. For each accession number recovered, we analysed the target insertion and clivage sites as well as data relating to the bacterial host, the host of the IS and the genetic support. Genetic support was analysed using the PLSDB site [55,56]. When the PLSDB site gave no results, we searched for plasmid-specific genes (such as *PsiA-B*, *StbA-B*, *RepA*...) or chromosome-specific genes (such as *DNAa*, *SeqA*, *DNAI*) to identify the genetic support. To define each isoform, we used a threshold of 95% nucleotide identity over the entire IS with at least one nucleotide that differs. This data extraction was carried out on 2020-10-15 for IS91 and on 2023-01-11 for IS801 and IS1294b.

### **Measurement of *in vivo* transposition frequencies of IS91 derivatives into *E. coli* by mating-out assay.**

Mating-out assays were performed as previously described [57]. The donor strain was JS219 [58] carrying three plasmids: pCP7 or pCP39 (used as IS91-Cm<sup>R</sup> derivatives donor plasmids; Amp<sup>R</sup>), pCP4, pCP40 or pAF26 (used as transposase +/- *orf121* expression plasmids; Sp<sup>R</sup>) and pOX38Km, the conjugative target plasmid (see figure S2 and S3). The recipient strain was MC1041 (Rif<sup>R</sup> Nal<sup>R</sup>). Following mating with the recipient strain, cells were plated on LB plates supplemented with rifampicin and kanamycin (exconjugants) or nalidixic acid, kanamycin and chloramphenicol (transposants). Transposition frequency was determined by dividing transposants by exconjugants. The frequencies of the transposition event were carried out on more than 5 independent experiments. These experiments were used to calculate mean values and standard deviations.

### **Measurement of the percentage of transposants from one-ended transposition in mating-out assay.**

Several hundred clones resulting from transposition were subcultured into LB plates (Km<sup>R</sup>, Cm<sup>R</sup> and Nal<sup>R</sup>) and LB plates (Km<sup>R</sup>, Amp<sup>R</sup> and Nal<sup>R</sup>). The percentage of one-ended transposition was determined by dividing the number of Km<sup>R</sup>, Amp<sup>R</sup> and Nal<sup>R</sup> clones by the number of Km<sup>R</sup>, Cm<sup>R</sup> and Nal<sup>R</sup> clones and multiplying by 100.

### **Measurement of *in vivo* insertion of IS91 circular intermediates into *E. coli*.**

To test the double-stranded circular intermediate, *E. coli* DH5 $\alpha$  strains containing plasmid pCP4 or pCP40 were transformed with 500 ng of plasmid pCP54 or pCP55 containing the *ori91::ter91* or *ter91::ori91* junction(see figure S3), respectively. 0.5 mM IPTG was added to the medium both during growth of the bacteria prior to preparation of the competent cells and during the expression phase of the transformation. Determination of the total number of viable cells was performed on LB plates supplemented with spectinomycin (Sp<sup>R</sup>) and the insertion of one of the junctions in the *E. coli*

*DH5α* genome was selected on LB plates supplemented with spectinomycin and chloramphenicol (Sp<sup>R</sup>, Cm<sup>R</sup>).

To test the single-stranded circular intermediate, overnight cultures with Sp<sup>R</sup> for donors and Cm<sup>R</sup> + DAP (0.3 mM) for recipients were diluted 1:100, without antibiotics, and further grown to OD<sub>600</sub> = 0.3. The expression of *tnpA* was induced by the addition of 0.5 mM IPTG and/or *orf121* by the addition of aTc at 50 ng/ml in the recipient. The strains were left for a further 1 h at 37°C with agitation. Plasmids were transferred from the *E. coli β2163 (pir-)* donor strain carrying the suicide plasmid (pSW23T, pCP54 or pCP55) into the *E. coli DH5α* target strain containing different *tnpA* and/or *orf121* expression plasmids (pCP4, pCP40 or pAF26). Filters containing the mixture of donor and recipient strains in a 1:8 ratio were incubated for 3 h at 37°C on LB plates (0.3 mM DAP, 0.5 mM IPTG or 0.3 mM DAP, 0.5 mM IPTG, 50 ng/ml aTc). Cells were then suspended in 2 mL LB, and appropriate dilutions were plated on selective LB plates (Sp<sup>R</sup>) for recipients and LB plates (Sp<sup>R</sup>, Cm<sup>R</sup>) for transconjugants (see figure S6). The frequency of Cm<sup>R</sup> transconjugants was calculated as the number of Cm<sup>R</sup>-tagged recipient cells out of the total number of recipient cells. The frequencies of the insertion event per viable cell from 4 independent experiments were used to calculate mean values and standard deviations.

### **Quantification of *orf121* and *tnpA* transcripts.**

For mating-out transcripts, overnight cultures were diluted to OD<sub>600</sub> = 0.2 and grown to the mean log phase (OD<sub>600</sub> = 0.5). For β-galactosidase assay transcripts, overnight cultures were diluted 1:100 and grown to the mean log phase (OD<sub>600</sub> = 0.6). Total RNA was extracted with the NucleoSpin® RNA Extraction Kit (Macherey-Nagel Inc.). Contaminating DNA was removed from RNA samples by using the Turbo DNA-free Kit (Ambion). cDNAs were synthesized from 1 μg of DNase-treated total RNA by using PrimeScript™ RT Reagent kit (TaKaRa Clontech). cDNA was quantified by PerfeCTa® SYBR® Green FastMix® Kit (Quanta BioSciences™) with adequate oligonucleotides

(see table S8). Relative expressions of *orf121* (Primers ORF121 F and ORF121 R) and *tnpA* (Primers TnpA F and TnpA R) genes were estimated by normalizing transcript copy number to those of the housekeeping gene *dxs* (Primer *dxs*-LC3 and *dxs*-LC4). The *tnpA/dxs* or *orf121/dxs* ratio was obtained from 3 independent experiments, which were used to calculate mean values and standard deviations.

### **β-galactosidases assays**

β-galactosidase assays were performed in the *E. coli* MG1656 strain (see table 1; [59]). Overnight cultures were diluted 1:100, and further grown to mid-log phase ( $OD_{600} = 0.4-0.6$ ). The Miller assay was performed as previously described [60]. β-galactosidase activity was obtained from 3 independent experiments, which were used to calculate mean values and standard deviations.

### **Analysis of His-tagged ORF121 or TnpA<sub>IS91</sub> protein expression by western-immunoblotting.**

Overnight cultures were diluted to  $OD_{600} = 0.2$  and grown to mid-log phase ( $OD_{600} = 0.5$ ). Gene expression was induced by the addition of 0.5 mM IPTG and/or 50 nM anhydrotetracycline in the media. Cultures were incubated for 3 h at 37°C, then 1 mL of each culture was centrifuged and stored at -20°C. Cell extracts were prepared using the B-PER bacterial protein extraction kit according to the manufacturer's recommendations at a rate of 100 μL of reagent per OD unit (Thermo Scientific, Ref. 90078) and inclusion body proteins using the solubilization reagent used according to the manufacturer's instructions (Thermo Scientific, Ref. 78115). Cell extract samples were subjected to 15% SDS-PAGE electrophoresis, followed by transfer to a PVDF membrane according to the manufacturer's instructions (Trans-Blot Turbo Transfer System, BioRad). The membrane was treated with a 1:3000 dilution of mouse anti-His-tag antibody (Invitrogen, Ref. MA121315) (1 night at 4°C), then washed 3 times for 10 minutes. Next, the membrane was treated with a 1:3000 dilution of anti-Mouse IgG, HRP antibody (Invitrogen, Ref. A28177), then washed 6 times for 5 minutes. The membrane was revealed with SuperSignal West Pico PLUS (Thermo Scientific, Ref. 34577). The

amount of protein was determined in arbitrary units (AU) by calculating the value of the air under the curve using ImageLab software.

### **Analysis of His-tagged ORF121 or TnpA<sub>IS91</sub> proteins by mass spectrometry.**

Peptides were analyzed by micro-LC-MS/MS using a nanoLC 425 system in micro-flow mode (Eksigent, Dublin, CA, USA) coupled with time-of-flight (TOF) (TripleTOF 5600+ Sciex, Framingham, MA, USA) operating in high-sensitivity mode. Reverse-phase LC was performed via a trap and elute configuration using a trap column (C18 Pepmap100 cartridge, 5 µm pore size; Thermo Fisher Scientific) and an analytical column (ChromXP C18 column, 12 nm, 3 µm pore size, Sciex) with the following mobile phases: loading solvent (water/ACN/TFA 98/2/0.05% (v/v)), solvent A (0.1% (v/v) TFA in water) and solvent B (water/ACN/TFA 5/95/0.1% (v/v)). All samples were loaded, trapped, and desalted using a flow rate of 10 µL/min with loading solvent for 5 minutes. The chromatographic separation was performed at a flow rate of 2 µL/min as follows: initial, 5% solvent B, increased to 25% for 90 minutes, then increased to 95% B for 10 minutes, maintained at 95% for 5 minutes and, finally, decreased to 5% B for re-equilibration. The peptides were then reprocessed via ProteinPilot with the Mascot module using *Escherichia coli* Uniprot 2022\_03. The following parameters were entered: trypsin 1 maximum missed cleavage, cysteine carbamidomethylation (fixed), methionine oxidation (variable), Precursor Tolerance 0.01Da; MS/MS fragment tolerance: 0.05 Da, a  $p < 0.05$  - peptide score  $> 25$  - bold red required and a protein score  $> 100$ .



## **ACKNOWLEDGEMENTS**

A. Fauconnier acknowledges the Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation (MESRI) for his doctoral training grant. This work was supported by fundings from the French Research institute Inserm.

The authors thank Emilie Pinault from BISCEm unit (Univ. Limoges, UAR 2015 CNRS, US 42 Inserm, CHU Limoges) for technical support regarding mass spectrometry.

## REFERENCES

1. von Wintersdorff CJH, Penders J, van Niekerk JM, Mills ND, Majumder S, van Alphen LB, et al. Dissemination of Antimicrobial Resistance in Microbial Ecosystems through Horizontal Gene Transfer. *Front Microbiol.* 2016;7. Available: <https://www.frontiersin.org/articles/10.3389/fmicb.2016.00173>
2. Partridge SR, Kwong SM, Firth N, Jensen SO. Mobile Genetic Elements Associated with Antimicrobial Resistance. *Clin Microbiol Rev.* 2018;31: e00088-17. doi:10.1128/CMR.00088-17
3. Diaz-Aroca E, de la Cruz F, Zabala JC, Ortiz JM. Characterization of the new insertion sequence IS91 from an alpha-hemolysin plasmid of *Escherichia coli*. *Mol Gen Genet MGG.* 1984;193: 493–499. doi:10.1007/BF00382089
4. Romantschuk M, Richter GY, Mukhopadhyay P, Mills D. IS801, an insertion sequence element isolated from *Pseudomonas syringae* pathovar phaseolicola. *Mol Microbiol.* 1991;5: 617–622. doi:10.1111/j.1365-2958.1991.tb00732.x
5. Tavakoli N, Comanducci A, Dodd HM, Lett MC, Albiger B, Bennett P. IS1294, a DNA element that transposes by RC transposition. *Plasmid.* 2000;44: 66–84. doi:10.1006/plas.1999.1460
6. Yassine H. Etude de la séquence d'insertion IS1294b et de son implication dans la dissémination des gènes de résistance aux antibiotiques chez les entérobactéries. These de doctorat, Bordeaux. 2015. Available: <https://www.theses.fr/2015BORD0405>
7. Chandler M, de la Cruz F, Dyda F, Hickman AB, Moncalian G, Ton-Hoang B. Breaking and joining single-stranded DNA: the HUH endonuclease superfamily. *Nat Rev Microbiol.* 2013;11: 525–538. doi:10.1038/nrmicro3067
8. Mendiola MV, Bernales I, de la Cruz F. Differential roles of the transposon termini in IS91 transposition. *Proc Natl Acad Sci U S A.* 1994;91: 1922–1926. doi:10.1073/pnas.91.5.1922
9. Garcillán-Barcia MP, Bernales I, Mendiola MV, de la Cruz F. IS 91 Rolling-Circle Transposition. *Mobile DNA II.* John Wiley & Sons, Ltd; 2007. pp. 889–904.

doi:10.1128/9781555817954.ch37

10. Toleman MA, Bennett PM, Walsh TR. ISCR elements: novel gene-capturing systems of the 21st century? *Microbiol Mol Biol Rev MMBR*. 2006;70: 296–316. doi:10.1128/MMBR.00048-05
11. Toleman MA, Walsh TR. Combinatorial events of insertion sequences and ICE in Gram-negative bacteria. *FEMS Microbiol Rev*. 2011;35: 912–935. doi:10.1111/j.1574-6976.2011.00294.x
12. Mendiola MV, de la Cruz F. Specificity of insertion of IS91, an insertion sequence present in alpha-haemolysin plasmids of *Escherichia coli*. *Mol Microbiol*. 1989;3: 979–984. doi:10.1111/j.1365-2958.1989.tb00247.x
13. Richter GY, Björklöf K, Romantschuk M, Mills D. Insertion specificity and trans-activation of IS801. *Mol Gen Genet MGG*. 1998;260: 381–387. doi:10.1007/s004380050907
14. del Pilar Garcillán-Barcia M, Bernales I, Mendiola MV, de la Cruz F. Single-stranded DNA intermediates in IS91 rolling-circle transposition. *Mol Microbiol*. 2001;39: 494–501. doi:10.1046/j.1365-2958.2001.02261.x
15. Bernales I, Mendiola MV, de la Cruz F. Intramolecular transposition of insertion sequence IS91 results in second-site simple insertions. *Mol Microbiol*. 1999;33: 223–234. doi:10.1046/j.1365-2958.1999.01432.x
16. Salamov VSA, Solovyevand A. Automatic annotation of microbial genomes and metagenomic sequences. *Metagenomics Its Appl Agric Biomed Environ Stud*. 2011; 61–78.
17. del Pilar Garcillán-Barcia M, Bernales I, Mendiola MV, de la Cruz F. Single-stranded DNA intermediates in IS91 rolling-circle transposition. *Mol Microbiol*. 2001;39: 494–501. doi:10.1046/j.1365-2958.2001.02261.x
18. Burland V, Shao Y, Perna NT, Plunkett G, Sofia HJ, Blattner FR. The complete DNA sequence and analysis of the large virulence plasmid of *Escherichia coli* O157:H7. *Nucleic Acids Res*. 1998;26: 4196–4204. doi:10.1093/nar/26.18.4196
19. Garcillán-Barcia MP, de la Cruz F. Distribution of IS91 family insertion sequences in bacterial genomes: evolutionary implications. *FEMS Microbiol Ecol*. 2002;42: 303–313. doi:10.1111/j.1574-

6941.2002.tb01020.x

20. Schlör S, Riedl S, Blass J, Reidl J. Genetic rearrangements of the regions adjacent to genes encoding heat-labile enterotoxins (eltAB) of enterotoxigenic *Escherichia coli* strains. *Appl Environ Microbiol.* 2000;66: 352–358. doi:10.1128/AEM.66.1.352-358.2000
21. Lallement C, Pasternak C, Ploy M-C, Jové T. The Role of ISCR1-Borne POUT Promoters in the Expression of Antibiotic Resistance Genes. *Front Microbiol.* 2018;9: 2579. doi:10.3389/fmicb.2018.02579
22. Toleman MA, Bennett PM, Walsh TR. Common regions e.g. orf513 and antibiotic resistance: IS91-like elements evolving complex class 1 integrons. *J Antimicrob Chemother.* 2006;58: 1–6. doi:10.1093/jac/dkl204
23. Toleman MA, Walsh TR. ISCR elements are key players in IncA/C plasmid evolution. *Antimicrob Agents Chemother.* 2010;54: 3534; author reply 3534. doi:10.1128/AAC.00383-10
24. Toukdarian A. Plasmid Strategies for Broad-Host-Range Replication in Gram-Negative Bacteria. *Plasmid Biology.* John Wiley & Sons, Ltd; 2004. pp. 257–270. doi:10.1128/9781555817732.ch11
25. Carattoli A. Resistance plasmid families in Enterobacteriaceae. *Antimicrob Agents Chemother.* 2009;53: 2227–2238. doi:10.1128/AAC.01707-08
26. Naseer U, Sundsfjord A. The CTX-M conundrum: dissemination of plasmids and *Escherichia coli* clones. *Microb Drug Resist Larchmt N.* 2011;17: 83–97. doi:10.1089/mdr.2010.0132
27. Nicolas-Chanoine M-H, Bertrand X, Madec J-Y. *Escherichia coli* ST131, an intriguing clonal group. *Clin Microbiol Rev.* 2014;27: 543–574. doi:10.1128/CMR.00125-13
28. Takayama Y, Sekizuka T, Matsui H, Adachi Y, Eda R, Nihonyanagi S, et al. Characterization of the IncFII-IncFIB(pB171) Plasmid Carrying bla NDM-5 in *Escherichia coli* ST405 Clinical Isolate in Japan. *Infect Drug Resist.* 2020;13: 561–566. doi:10.2147/IDR.S232943
29. Fortunato G, Vaz-Moreira I, Gajic I, Manaia CM. Insight into phylogenomic bias of blaVIM-2 or blaNDM-1 dissemination amongst carbapenem-resistant *Pseudomonas aeruginosa*. *Int J*

Antimicrob Agents. 2023;61: 106788. doi:10.1016/j.ijantimicag.2023.106788

30. Phuadraksa T, Wichit S, Songtawee N, Tantimavanich S, Isarankura-Na-Ayudhya C, Yainoy S. Emergence of plasmid-mediated colistin resistance *mcr-3.5* gene in *Citrobacter amalonaticus* and *Citrobacter sedlakii* isolated from healthy individual in Thailand. *Front Cell Infect Microbiol.* 2022;12: 1067572. doi:10.3389/fcimb.2022.1067572
31. Lu M, Zhai WS, Du PC, Wang Y, Zhan ZF, Chen S, et al. [Molecular characteristics of ciprofloxacin-cefotaxime-azithromycin co-resistant *Salmonella enterica* Serovar Thompson in foodborne diseases in Hunan Province]. *Zhonghua Yu Fang Yi Xue Za Zhi.* 2022;56: 1745–1750. doi:10.3760/cma.j.cn112150-20220112-00046
32. Li L, Olsen RH, Wang C, Song A, Xiao J, Meng H, et al. First report of two foodborne *Salmonella enterica* subsp. *enterica* serovar *Bovismorbificans* isolates carrying a novel mega-plasmid harboring *blaDHA-1* and *qnrB4* genes. *Int J Food Microbiol.* 2021;360: 109439. doi:10.1016/j.ijfoodmicro.2021.109439
33. Jiang T, Li G, Huang L, Ding D, Ruan Z, Yan J. Genomic and Phylogenetic Analysis of a Multidrug-Resistant *blaNDM*-carrying *Klebsiella michiganensis* in China. *Infect Drug Resist.* 2023;16: 3109–3116. doi:10.2147/IDR.S409544
34. Berglund F, Ebmeyer S, Kristiansson E, Larsson DGJ. Evidence for wastewaters as environments where mobile antibiotic resistance genes emerge. *Commun Biol.* 2023;6: 321. doi:10.1038/s42003-023-04676-7
35. Liu Y, Wu F, Chen Q, Ying Y, Jiang Y, Lu J, et al. Comparative genomics analysis of *Raoultella planticola* S25 isolated from duck in China, with florfenicol resistance. *Comp Immunol Microbiol Infect Dis.* 2020;68: 101398. doi:10.1016/j.cimid.2019.101398
36. Fan Q, Zhang J, Shi H, Chang S, Hou F. Metagenomic Profiles of Yak and Cattle Manure Resistomes in Different Feeding Patterns before and after Composting. *Appl Environ Microbiol.* 2023;89: e0064523. doi:10.1128/aem.00645-23
37. Zhou K, Luo Q, Wang Q, Huang C, Lu H, Rossen JWA, et al. Silent transmission of an

- IS1294b-deactivated *mcr-1* gene with inducible colistin resistance. *Int J Antimicrob Agents*. 2018;51: 822–828. doi:10.1016/j.ijantimicag.2018.01.004
38. Wei B, Cha S-Y, Shang K, Zhang J-F, Jang H-K, Kang M. Genetic diversity of extended-spectrum cephalosporin resistance in *Salmonella enterica* and *E. coli* isolates in a single broiler chicken. *Vet Microbiol*. 2021;254: 109010. doi:10.1016/j.vetmic.2021.109010
39. Yassine H, Bientz L, Cros J, Goret J, Bébéar C, Quentin C, et al. Experimental evidence for IS1294b-mediated transposition of the *bla*CMY-2 cephalosporinase gene in Enterobacteriaceae. *J Antimicrob Chemother*. 2015;70: 697–700. doi:10.1093/jac/dku472
40. Bardaji L, Añorga M, Echeverría M, Ramos C, Murillo J. The toxic guardians - multiple toxin-antitoxin systems provide stability, avoid deletions and maintain virulence genes of *Pseudomonas syringae* virulence plasmids. *Mob DNA*. 2019;10: 7. doi:10.1186/s13100-019-0149-4
41. Baranov PV, Fayet O, Hendrix RW, Atkins JF. Recoding in bacteriophages and bacterial IS elements. *Trends Genet*. 2006;22: 174–181. doi:10.1016/j.tig.2006.01.005
42. Baranov PV, Hammer AW, Zhou J, Gesteland RF, Atkins JF. Transcriptional slippage in bacteria: distribution in sequenced genomes and utilization in IS element gene expression. *Genome Biol*. 2005;6: R25. doi:10.1186/gb-2005-6-3-r25
43. Sharma V, Prère M-F, Canal I, Firth AE, Atkins JF, Baranov PV, et al. Analysis of tetra- and hepta-nucleotides motifs promoting -1 ribosomal frameshifting in *Escherichia coli*. *Nucleic Acids Res*. 2014;42: 7210–7225. doi:10.1093/nar/gku386
44. Chandler M, Fayet O. Translational frameshifting in the control of transposition in bacteria. *Mol Microbiol*. 1993;7: 497–503. doi:10.1111/j.1365-2958.1993.tb01140.x
45. Sekine Y, Ohtsubo E. Frameshifting is required for production of the transposase encoded by insertion sequence 1. *Proc Natl Acad Sci U S A*. 1989;86: 4609–4613. doi:10.1073/pnas.86.12.4609
46. Siguier P, Gourbeyre E, Varani A, Ton-Hoang B, Chandler M. Everyman's Guide to Bacterial Insertion Sequences. *Microbiol Spectr*. 2015;3: MDNA3-0030–2014. doi:10.1128/microbiolspec.MDNA3-0030-2014

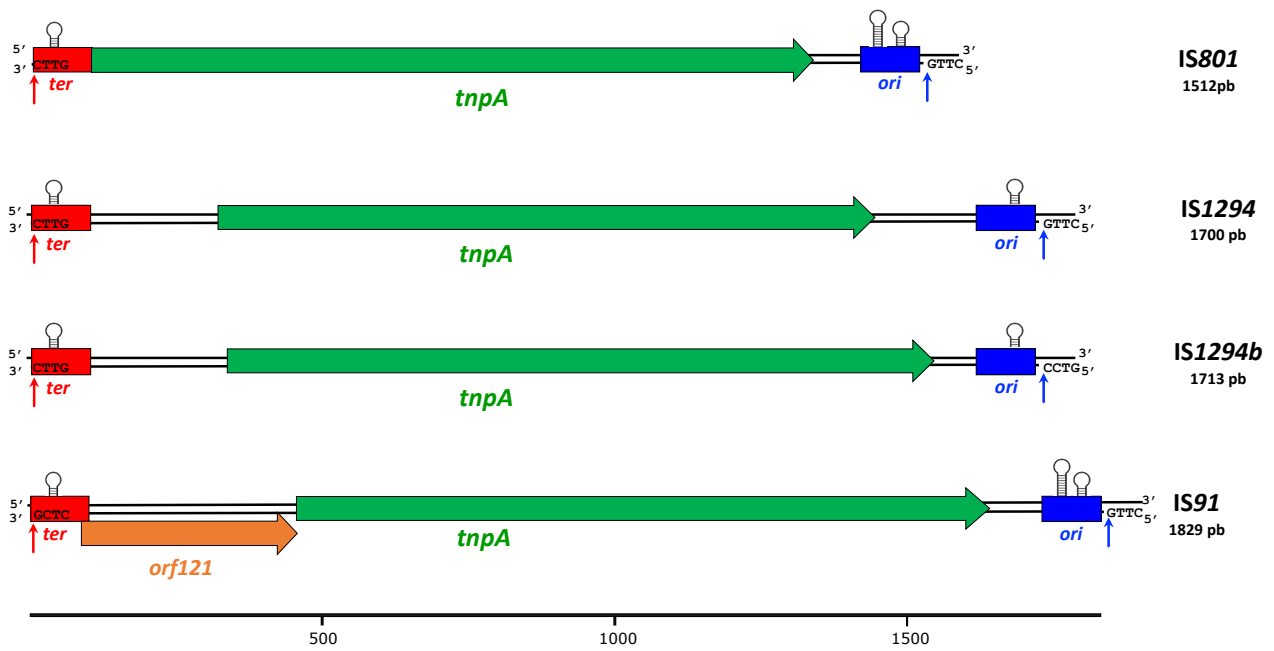
47. Meers C, Le HC, Pesari SR, Hoffmann FT, Walker MWG, Gezelle J, et al. Transposon-encoded nucleases use guide RNAs to promote their selfish spread. *Nature*. 2023;622: 863–871. doi:10.1038/s41586-023-06597-1
48. Altae-Tran H, Kannan S, Demircioglu FE, Oshiro R, Nety SP, McKay LJ, et al. The widespread IS200/IS605 transposon family encodes diverse programmable RNA-guided endonucleases. *Science*. 2021;374: 57–65. doi:10.1126/science.abj6856
49. Karvelis T, Druteika G, Bigelyte G, Budre K, Zedaveinyte R, Silanskas A, et al. Transposon-associated TnpB is a programmable RNA-guided DNA endonuclease. *Nature*. 2021;599: 692–696. doi:10.1038/s41586-021-04058-1
50. Davis MA, Simons RW, Kleckner N. Tn10 protects itself at two levels from fortuitous activation by external promoters. *Cell*. 1985;43: 379–387. doi:10.1016/0092-8674(85)90043-1
51. Rezsöhazy R, Hallet B, Delcour J, Mahillon J. The IS4 family of insertion sequences: evidence for a conserved transposase motif. *Mol Microbiol*. 1993;9: 1283–1295. doi:10.1111/j.1365-2958.1993.tb01258.x
52. Wawrzyniak P, Płucienniczak G, Bartosik D. The Different Faces of Rolling-Circle Replication and Its Multifunctional Initiator Proteins. *Front Microbiol*. 2017;8. Available: <https://www.frontiersin.org/articles/10.3389/fmicb.2017.02353>
53. Guynet C, Achard A, Hoang BT, Barabas O, Hickman AB, Dyda F, et al. Resetting the site: redirecting integration of an insertion sequence in a predictable way. *Mol Cell*. 2009;34: 612–619. doi:10.1016/j.molcel.2009.05.017
54. Hanahan D. Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol*. 1983;166: 557–580. doi:10.1016/s0022-2836(83)80284-8
55. Galata V, Fehlmann T, Backes C, Keller A. PLSDB: a resource of complete bacterial plasmids. *Nucleic Acids Res*. 2019;47: D195–D202. doi:10.1093/nar/gky1050
56. Schmartz GP, Hartung A, Hirsch P, Kern F, Fehlmann T, Müller R, et al. PLSDB: advancing a comprehensive database of bacterial plasmids. *Nucleic Acids Res*. 2022;50: D273–D278.

doi:10.1093/nar/gkab1111

57. Polard P, Prère MF, Fayet O, Chandler M. Transposase-induced excision and circularization of the bacterial insertion sequence IS911. *EMBO J.* 1992;11: 5079–5090. doi:10.1002/j.1460-2075.1992.tb05615.x
58. Cam K, Béjar S, Gil D, Bouché JP. Identification and sequence of gene *dicB*: translation of the division inhibitor from an in-phase internal start. *Nucleic Acids Res.* 1988;16: 6327–6338. doi:10.1093/nar/16.14.6327
59. Espéli O, Moulin L, Boccard F. Transcription attenuation associated with bacterial repetitive extragenic BIME elements. *J Mol Biol.* 2001;314: 375–386. doi:10.1006/jmbi.2001.5150
60. Miller JH. *A Short Course in Bacterial Genetics: A Laboratory Manual and Handbook for Escherichia Coli and Related Bacteria.* CSHL Press; 1992.
61. Demarre G, Guérout A-M, Matsumoto-Mashimo C, Rowe-Magnus DA, Marlière P, Mazel D. A new family of mobilizable suicide plasmids based on broad host range R388 plasmid (IncW) and RP4 plasmid (IncPalpha) conjugative machineries and their cognate *Escherichia coli* host strains. *Res Microbiol.* 2005;156: 245–255. doi:10.1016/j.resmic.2004.09.007
62. Yanisch-Perron C, Vieira J, Messing J. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene.* 1985;33: 103–119. doi:10.1016/0378-1119(85)90120-9
63. Pasternak C, Dulermo R, Ton-Hoang B, Debuchy R, Siguier P, Coste G, et al. ISDra2 transposition in *Deinococcus radiodurans* is downregulated by TnpB. *Mol Microbiol.* 2013;88: 443–455. doi:10.1111/mmi.12194
64. Jové T, Da Re S, Denis F, Mazel D, Ploy M-C. Inverse correlation between promoter strength and excision activity in class 1 integrons. *PLoS Genet.* 2010;6: e1000793. doi:10.1371/journal.pgen.1000793
65. Chandler M, Galas DJ. IS1-mediated tandem duplication of plasmid pBR322. Dependence on *recA* and on DNA polymerase I. *J Mol Biol.* 1983;165: 183–190. doi:10.1016/s0022-2836(83)80249-

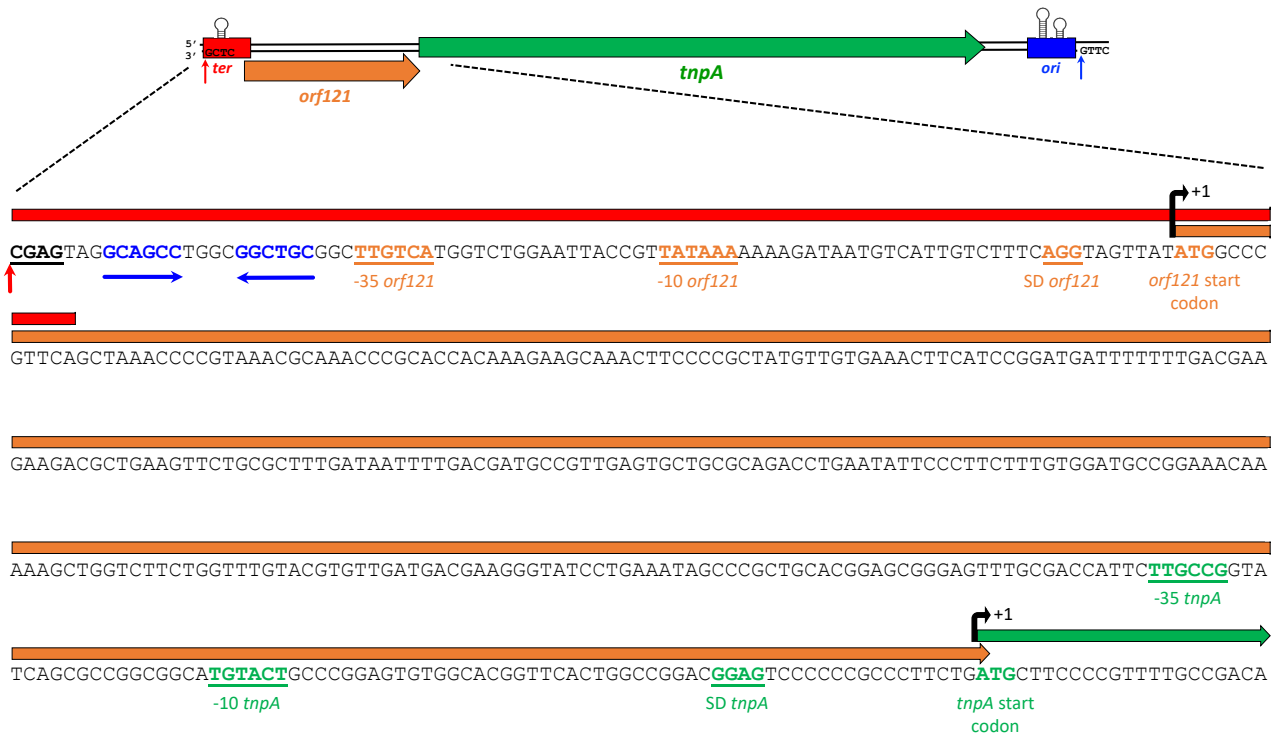


## FIGURES AND LEGENDS



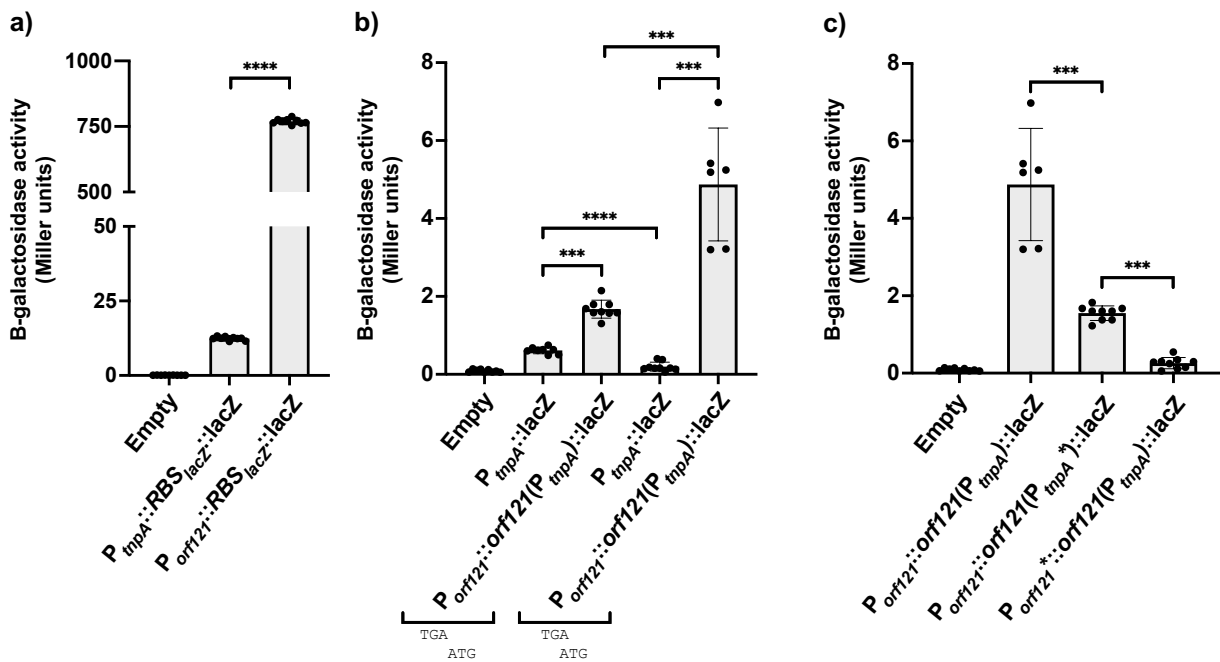
**Figure 1. Genetic organization of active members of the IS91 family (IS801, IS1294, IS1294b and IS91).**

*Ter* and *ori* ends carrying the palindromic sequences are depicted as red and blue boxes, respectively. The cleavage positions at *ter* and *ori* ends are shown by red and blue vertical arrows, respectively. The cleavage site at the *ter* end and the tetranucleotide target site adjacent to *ori* are indicated. *orf121* and *tnpA* encoding sequences are shown as orange and green arrows, respectively.



**Figure 2. DNA sequence of the IS91 5' end from *ter91* to the start of the *tnpA* gene.**

Blue arrows represent the terminal palindrome. The *terIS* region, the *orf121* and *tnpA* genes are indicated by a red rectangle, an orange arrow and a green arrow, respectively. The nucleotides are shown below. The cleavage tetranucleotide is in bold and underlined with a red arrow to indicate the cleavage site. The -35, -10, SD and start codon regions are shown in bold and underlined.



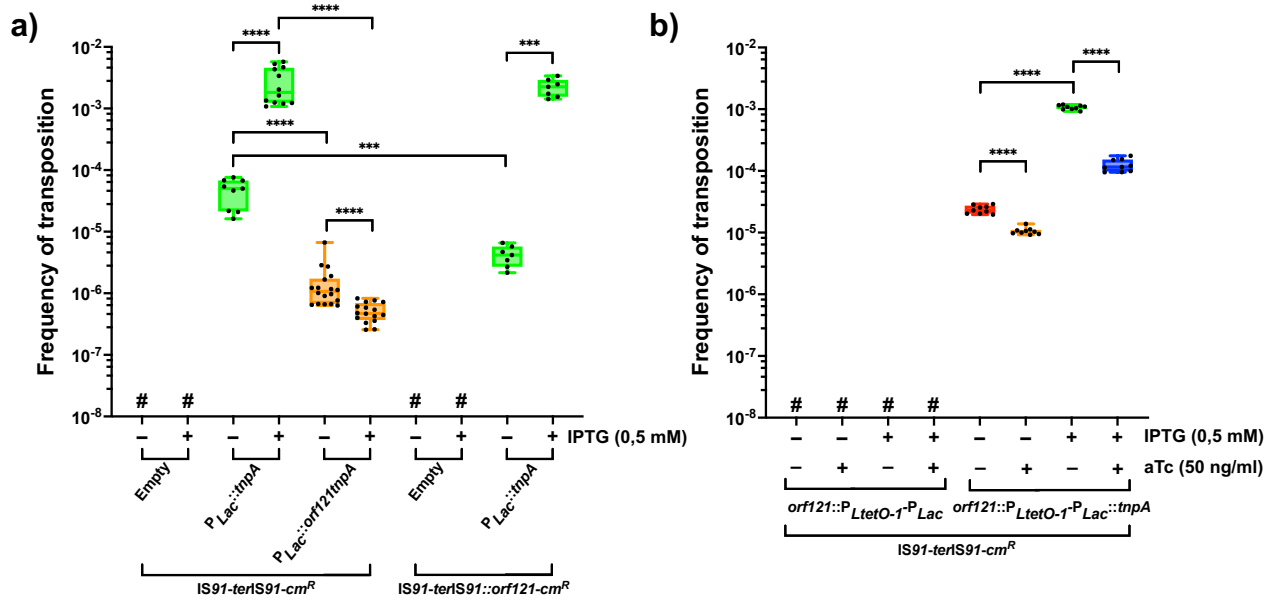
**Figure 3. Activities of the  $P_{orf121}$  and  $P_{tnpA}$  promoters.**

$\beta$ -galactosidase activities were measured from  $lacZ$ -transcriptional fusions with the  $P_{orf121}$  and/or  $P_{tnpA}$  sequences. The transcriptional fusions are described in Table 1. The results are the average of at least 3 independent experiments.  $p = 0.0002$  (\*\*\*) and  $p < 0.0001$  (\*\*\*\*). Empty, empty transposase expression vector (pCP2).

a) activity of the two predicted promoters:  $P_{orf121}$  and  $P_{tnpA}$ .

b) activity of  $tnpA$  with or without the  $orf121$  termination codon overlapping the  $tnpA$  initiation codon.

c) activity of  $tnpA$  with the termination codon of  $orf121$  overlapping its initiation codon in the presence of a mutation in the promoter  $P_{orf121}$  or  $P_{tnpA}$ .

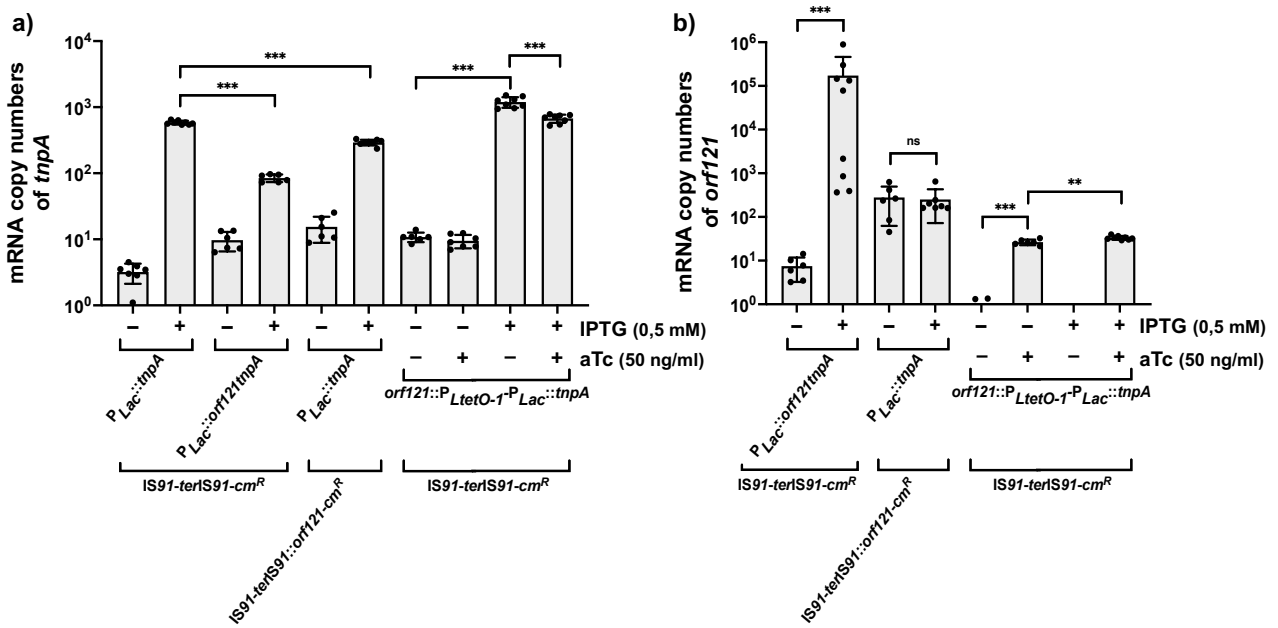


**Figure 4. Expression of *orf121* decreases IS91 transposition frequency.**

Mating-out assays as described in Materials and Methods. Measured transposition frequencies (Error bars indicate the SDs;  $n \geq 5$ ).  $p = 0.0002$  (\*\*\*) and  $p < 0.0001$  (\*\*\*\*). #, undetectable transposition. The - and + indicate the presence or absence of inducer at 0.5 mM IPTG or 50 ng/ml aTc. Empty, empty transposase expression vector;  $P_{lac}::tnpA$ , *tnpA* gene alone is expressed *in trans*;  $P_{lac}::orf121tnpA$ , both ORF121 and *tnpA* gene are expressed *in trans*.

a) *orf121* gene is expressed *in cis* or *in trans*.

b) *orf121* and *tnpA* are under the control of two different promoters.

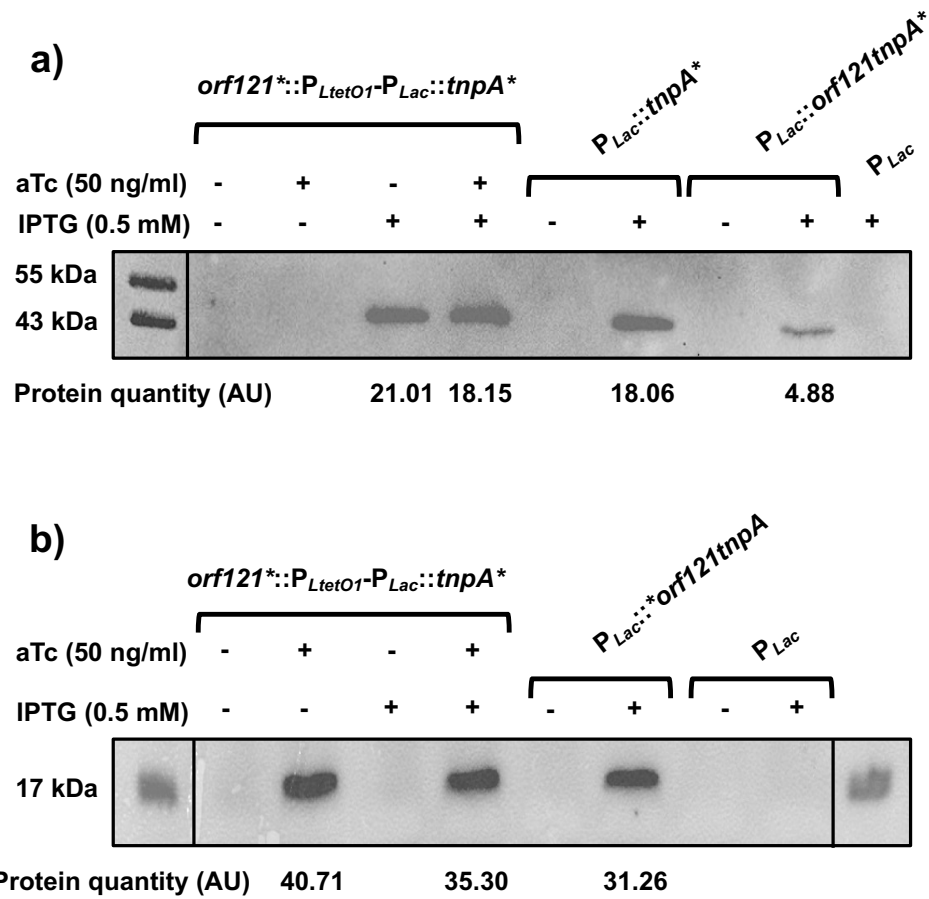


**Figure 5. mRNA copy numbers of *orf121* and *tnpA* genes**

Quantification of *orf121* and *tnpA* transcripts by qRT-PCR as described in Materials and Methods. (Error bars indicate the SDs;  $n \geq 5$ ).  $p = 0.002$  (\*\*) and  $p = 0.0002$  (\*\*\*). The - and + indicate the presence or absence of inducer at 0.5 mM IPTG or 50 ng/ml aTc. When the *orf121* gene is expressed *in cis* or *in trans*, or when *orf121* and *tnpA* are under the control of two different promoters.

a) numbers transcripts of *tnpA* per ratio of the *tnpA* gene to the *dxs* housekeeping gene.

b) numbers transcripts of *orf121* per ratio of the *orf121* gene to the *dxs* housekeeping gene.

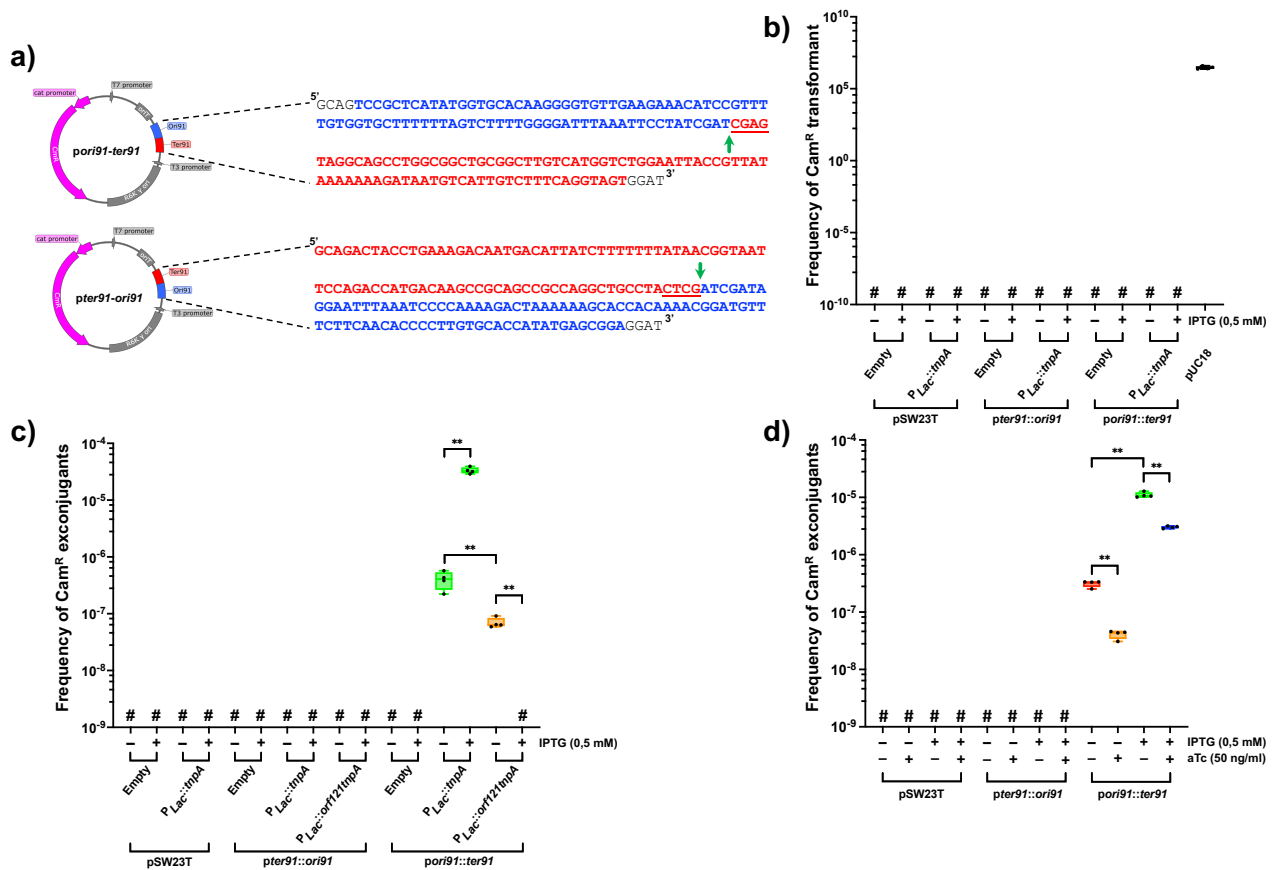


**Figure 6. Detection of Orf121 and TnpA products by western immunoblot**

Analysis of His-tagged ORF121 or TnpAIS91 protein expression by western-immunoblotting as described in Materials and Methods. The - and + indicate the presence or absence of inducer at 0.5 mM IPTG or 50 ng/ml aTc; \*, His-tag position; AU for an Arbitrary Unit quantity of protein.

a) detection of the TnpA product in the insoluble bodies, expected size 47 kDa.

b) detection of Orf121 in the soluble part, expected size 17 kDa.



**Figure 7. Only ssDNA circular intermediates are successfully inserted into a new target sequence.**

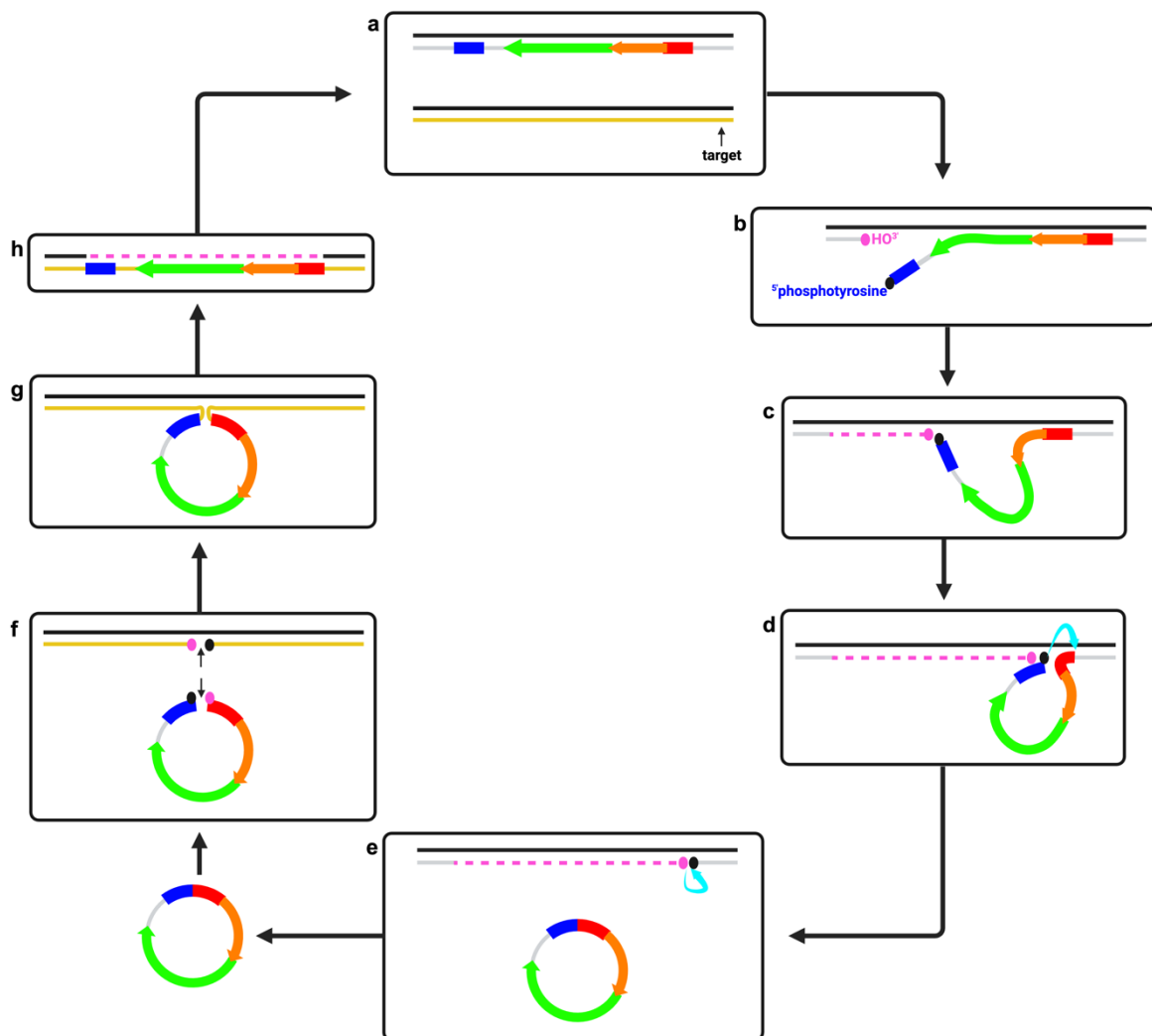
*E. coli* transformation with dsDNA intermediates and suicide conjugation assays to assess ssDNA circles insertion. Values represent the mean and the standard deviation of 4 biological replicates.  $p = 0.002$  (\*\*). The - and + indicate the presence or absence of inducer at 0.5 mM IPTG or 50 ng/ml aTc; #, undetectable exconjugants or transformants.

a) *pter91::ori91* or *pori91::ter91* junction present in the substrates used for insertion analyses.

b) *In vivo* insertion frequencies of the IS91 dsDNA circles. The vertical axis shows the frequency of  $\text{Cm}^R$  transformant resulting from transposition insertion of *pter91::ori91* or *pori91::ter91* into the *E. coli* recipient strain expressing *P<sub>lac</sub>::tnpA* or empty, empty transposase expression vector in the presence or absence of IPTG. pCU18 (black) was used as a positive control for transformation efficiency. #, undetectable  $\text{Cm}^R$  transformant ( $< 2.63 \times 10^{-9}$ ).

c) *In vivo* insertion frequencies of the IS91 ssDNA circles. The vertical axis shows the frequency of Cm<sup>R</sup> exconjugants resulting from transposition insertion of *pter91::ori91* or *pori91::ter91* into the *E. coli* recipient strain expressing empty, empty transposase expression vector, *P<sub>lac</sub>::tnpA* (green) or *P<sub>lac</sub>::orf121tnpA* (orange) in the presence or absence of IPTG. #, undetectable Cm<sup>R</sup> exconjugants (<2.87 x 10<sup>-9</sup>).

d) Suicide conjugation assays to assess ssDNA intermediates insertion. Values represent the mean and the standard deviation of 4 biological replicates. *In vivo* insertion frequencies of the IS91 ssDNA circles. The vertical axis shows the frequency of Cm<sup>R</sup> exconjugants resulting from transposition insertion of *pter91::ori91* or *pori91::ter91* into the *E. coli* recipient strain expressing *orf121::P<sub>LtetO-1</sub>-P<sub>lac</sub>::tnpA* without inducer (red), with only aTc (orange) or only IPTG (green), with both inducers (blue). #, undetectable Cm<sup>R</sup> exconjugants (<3.04 x 10<sup>-9</sup>).





### **Figure 8. Update of the “peel and paste” transposition model of IS91**

The top strand is represented by a black line in the donor and recipient DNA, the bottom strand is grey in the donor DNA and blue in the recipient DNA. The *terIS*, *orf121*, *tnpA* and *oriIS* ends are indicated by a red rectangle, an orange arrow, a green arrow and a blue rectangle, respectively. The free 3'OH groups are represented by a pink sphere and the 5'phosphotyrosine complex by a black sphere. The cyan arrow indicates nucleophilic attacks. Strand replication by the bacterial replication machinery is shown in dotted pink. Created with BioRender.com

## TABLE AND LEGENDS

**Table 1. Bacterial strains and plasmids**

Δ indicates a deletion ; Ω indicates an insertion ; :: indicates a novel joint; \* indicates a mutation.

Strain or plasmid	Surname	Genotype or other relevant characteristics	Source or reference
<b><i>E. coli</i> strains</b>			
DH5α		<i>supE44 ΔlacU(φ80lacZΔM15)</i> <i>hsdR17recA1endA1gyrA96 thi-1 relA1</i>	Laboratory stock
JS219		MC1061, <i>recA1</i> , <i>lacI<sup>f</sup></i>	[58]
MC240		XA103, F <sup>-</sup> , <i>ara</i> , <i>del(lac pro)</i> , <i>gyrA</i> (nalR), <i>metB</i> , <i>argE<sup>am</sup></i> , <i>rpoB</i> , <i>thi</i> , <i>supF</i>	Gift of B. Ton-Hoang
β2163		(F <sup>-</sup> ) RP4-2-Tc::Mu <i>dapA</i>	[61]
<b>Plasmids</b>			
pSU2600		Source of <i>IS91</i> for the <i>orf121</i> gene, <i>tnpA</i> gene and the promoters P <sub><i>orf121</i></sub> and P <sub><i>tnpA</i></sub>	[8]
pUC18		Amp <sup>R</sup> ; MCS from M13mp19 into <i>lacZ'</i>	[62]
pCP5		pUC18; P <sub><i>cat</i></sub> ::Cm <sup>R</sup>	This work
pCP6		pUC18; <i>ter91</i> P <sub><i>cat</i></sub> ::Cm <sup>R</sup>	This work
pAF1		pUC18; P <sub><i>cat</i></sub> ::Cm <sup>R</sup> <i>ori91</i>	This work
pGY13556		pAPT110ΔKan <sup>r</sup> ; P <sub><i>lac</i></sub> :: <i>tnpA</i> <sub><i>ISDra2</i></sub> - <i>his6</i>	[63]
pSW23T		Cm <sup>R</sup>	[61]
pSU38ΔtotlacZ		vector carrying promoterless <i>lacZ</i> ; p15A <i>ori</i> ; Km <sup>R</sup>	[64]
pOX38Km		Km <sup>R</sup>	[65]
<b><i>IS91</i> substrate plasmids</b>			
pCP7	<i>IS91-ter91-cm<sup>R</sup></i>	pUC18; <i>ter91</i> P <sub><i>cat</i></sub> ::Cm <sup>R</sup> <i>ori91</i>	This work
pCP39	<i>IS91-ter91::orf121-cm<sup>R</sup></i>	pUC18; <i>ter91::orf121</i> P <sub><i>cat</i></sub> ::Cm <sup>R</sup> <i>ori91</i>	This work
<b><i>TnpA</i> expression plasmids</b>			
pCP2	<u>Empty</u>	pGY13556ΔtnpA <sub><i>ISDra2</i></sub> Ωlinker79; SpR	This work
pCP4	P <sub><i>lac</i></sub> :: <i>tnpA</i>	pCP2; P <sub><i>lac</i></sub> :: <i>tnpA</i>	This work
pAF4	P <sub><i>Lac</i></sub> :: <i>tnpA</i> *	pCP4; P <sub><i>lac</i></sub> :: <i>tnpA</i> His <sub>6</sub>	This work
pCP40	P <sub><i>lac</i></sub> :: <i>orf121tnpA</i>	pCP2; P <sub><i>lac</i></sub> :: <i>orf121tnpA</i>	This work
pAF2	P <sub><i>Lac</i></sub> :: <i>orf121tnpA</i> *	pCP40; P <sub><i>lac</i></sub> :: <i>orf121tnpA</i> His <sub>6</sub>	This work
pCP57	P <sub><i>Lac</i></sub> ::* <i>orf121tnpA</i>	pCP40; P <sub><i>lac</i></sub> ::His <sub>6</sub> <i>orf121tnpA</i>	This work
pAF25		pCP2; <i>orf121</i> ::P <sub><i>LtetO-1</i></sub> -P <sub><i>lac</i></sub>	This work
pAF26		pCP2; <i>orf121</i> ::P <sub><i>LtetO-1</i></sub> -P <sub><i>lac</i></sub> :: <i>tnpA</i>	This work
pAF34	<i>orf121</i> *::P <sub><i>LtetO-1</i></sub> -P <sub><i>Lac</i></sub> :: <i>tnpA</i> *	pCP2; <i>orf121</i> His <sub>6</sub> ::P <sub><i>LtetO-1</i></sub> -P <sub><i>lac</i></sub> :: <i>tnpA</i> His <sub>6</sub>	This work
<b>Circular intermediates plasmids</b>			
pCP54	<i>pori91::ter91</i>	pSW23T; <i>oriIS91::terIS91</i>	This work
pCP55	<i>pter91::ori91</i>	pSW23T; <i>terIS91::oriIS91</i>	This work
<b>Transcriptional fusion plasmids with <i>lacZ</i></b>			
pCP65	P <sub><i>tnpA</i></sub> ::RBS <sub><i>lacZ</i></sub> :: <i>lacZ</i>	pSU38ΔtotlacZ; P <sub><i>tnpA</i></sub> ::RBS <sub><i>lacZ</i></sub> :: <i>lacZ</i>	This work
pCP66	P <sub><i>orf121</i></sub> ::RBS <sub><i>lacZ</i></sub> :: <i>lacZ</i>	pSU38ΔtotlacZ; P <sub><i>orf121</i></sub> ::RBS <sub><i>lacZ</i></sub> :: <i>lacZ</i>	This work
pCP73	P <sub><i>tnpA</i></sub>	pSU38ΔtotlacZ; <i>orf121</i> ::P <sub><i>tnpA</i></sub> ::RBS <sub><i>tnpA</i></sub> :: <i>lacZ</i> with overlap (-1)	This work
pCP74	P <sub><i>orf121</i></sub> :: <i>orf121</i> (P <sub><i>tnpA</i></sub> )	pSU38ΔtotlacZ; P <sub><i>orf121</i></sub> ::RBS <sub><i>orf121</i></sub> :: <i>orf121</i> ::P <sub><i>tnpA</i></sub> ::RBS <sub><i>tnpA</i></sub> :: <i>lacZ</i> with overlap (-1)	This work
pCP75	P <sub><i>tnpA</i></sub> Δ-1	pSU38ΔtotlacZ; <i>orf121</i> ::P <sub><i>tnpA</i></sub> ::RBS <sub><i>tnpA</i></sub> :: <i>lacZ</i> <b>without</b> overlap (-1)	This work
pCP76	P <sub><i>orf121</i></sub> :: <i>orf121</i> (P <sub><i>tnpA</i></sub> )Δ-1	pSU38ΔtotlacZ; P <sub><i>orf121</i></sub> ::RBS <sub><i>orf121</i></sub> :: <i>orf121</i> ::P <sub><i>tnpA</i></sub> ::RBS <sub><i>tnpA</i></sub> :: <i>lacZ</i> <b>without</b> overlap (-1)	This work
pAF36		pSU38ΔtotlacZ; <i>orf121</i> ::P <sub><i>tnpA</i></sub> *::RBS <sub><i>tnpA</i></sub> :: <i>lacZ</i> with overlap (-1)	This work
pAF37		pSU38ΔtotlacZ; P <sub><i>orf121</i></sub> ::RBS <sub><i>orf121</i></sub> :: <i>orf121</i> ::P <sub><i>tnpA</i></sub> *::RBS <sub><i>tnpA</i></sub> :: <i>lacZ</i> with overlap (-1)	This work

pAF38		<p style="text-align: center;">pSU38<math>\Delta</math>totlacZ;  <math>P_{orf121}^*::RBS_{orf121}::orf121::P_{mpA}::RBS_{mpA}::lacZ</math> with  overlap (-1)</p>	This work
-------	--	---	-----------

**Table 2. Orf121 is not involved in the target selection and insertion specificity of IS91 and tetranucleotides of the cleavage sites**

The n value indicates the number of clones analysed by AP-PCR and the - and + indicate the presence or absence of inducer at 0.5 mM IPTG or 50 ng/ml aTc.

a) Analysis of the target insertion sites in relation to the expression of *orf121* *in cis* or *in trans*

b) Analysis of the target insertion sites when *tnpA* *orf121* are expressed *in trans* from two different promoters.

c) Analysis of *terIS* cleavage sites in the presence or absence of *orf121* *in trans*.

		IS91- <i>Ter91-cm<sup>R</sup></i>		IS91- <i>Ter91::orf121-cm<sup>R</sup></i>
		<i>P<sub>lac</sub>::tnpA</i>	<i>P<sub>lac</sub>::orf121tnpA</i>	<i>P<sub>lac</sub>::tnpA</i>
Insertion target site	n	277	292	240
	5'-CTTG	142	136	109
	5'-GTTC	113	134	113
	5'-GTCC	14	13	12
	5'-GTAC	3	1	0
	5'-CTCG	3	5	2
	5'-GTCCG	1	0	0
	5'-CTTC	1	0	1
	5'-CTGT	0	1	0
	5'-CTGG	0	1	0
	5'-CTGA	0	1	0
	5'-GTGC	0	0	2
	5'-TTTG	0	0	1

		<i>orf121::P<sub>LtetO-1</sub>-P<sub>lac</sub>::tnpA</i>			
		IPTG (0.5 mM)		aTc (50 ng/ml)	
Insertion target site	n	-	-	+	+
		-	+	-	+
		50	49	50	50
	5'-CTTG	31	34	29	36
5'-GTTC	19	15	21	12	
5'-CTCG	0	0	0	2	

		IS91- <i>Ter91-cm<sup>R</sup></i>	
		<i>P<sub>lac</sub>::tnpA</i>	<i>P<sub>lac</sub>::orf121tnpA</i>
cleavage site	n	43	68
	5'-CTCG	43	68

**Table 3. Orf121 protein is required for accurate recognition and cleavage of the *terIS* end**

The n value indicates the number of clones analysed.

a) Impact of *orf121* expression *in cis* or *in trans* on recognition and cleavage of the *terIS* site. The exponents \* indicate that the difference between the percentages is significant compared to the condition without *orf121* expression, i.e. that the observed difference cannot be due to random sampling.

b) Impact of Orf121 on *terIS* site recognition and cleavage. The exponents \* indicate that the difference between the percentages is significant compared to the condition without inducer, i.e. that the observed difference cannot be due to random sampling. The exponents \$ indicate that the difference between the percentages is significant compared to IPTG-only.

**a)**

	<i>P<sub>lac</sub>::tnpA</i>	<i>P<sub>lac</sub>::orf121tnpA</i>
<b>IS91-Ter91-cm<sup>R</sup></b>	53 % n : 1235	18 %* n : 674
<b>IS91-Ter91::orf121-cm<sup>R</sup></b>	29 %* n : 1104	-

**b)**

	<i>orf121::P<sub>LtetO-1</sub>-P<sub>lac</sub>::tnpA</i>			
<b>IPTG (0.5 mM)</b>	-	-	+	+
<b>aTc (50 ng/ml)</b>	-	+	-	+
<b>IS91-ter91-cm<sup>R</sup></b>	29 % n : 1020	11 %* n : 1020	48 %* n : 1020	32 %\$ n : 1020

## SUPPLEMENTAL FIGURES AND LEGENDS

**a)** IS91-V1 85- GATGATTTTTTTGAC  
 IS91-V22 85- GATGA-TTTTTTGAC

IS91-V1 85- GATGATTTTTTTGAC ... 190- GGAAACAAAAAGC  
 IS91-V6 85- GATGA-TTTTTTGAC ... 190- GGAAAC-AAAAGC

IS91-V1 345- GGAGTCCCCCGCCCTTCTGA  
 IS91-V3 345- GGAGTCCCCCGCC-TTCTGA  
 IS91-V13 345- GGAGTCCCCC-CCCTTCTGA

**b)** IS91-V1 56- TTCCCCGC-TATGTTG  
 IS91-V66 56- TTCCCCGC**G**TATGTTG

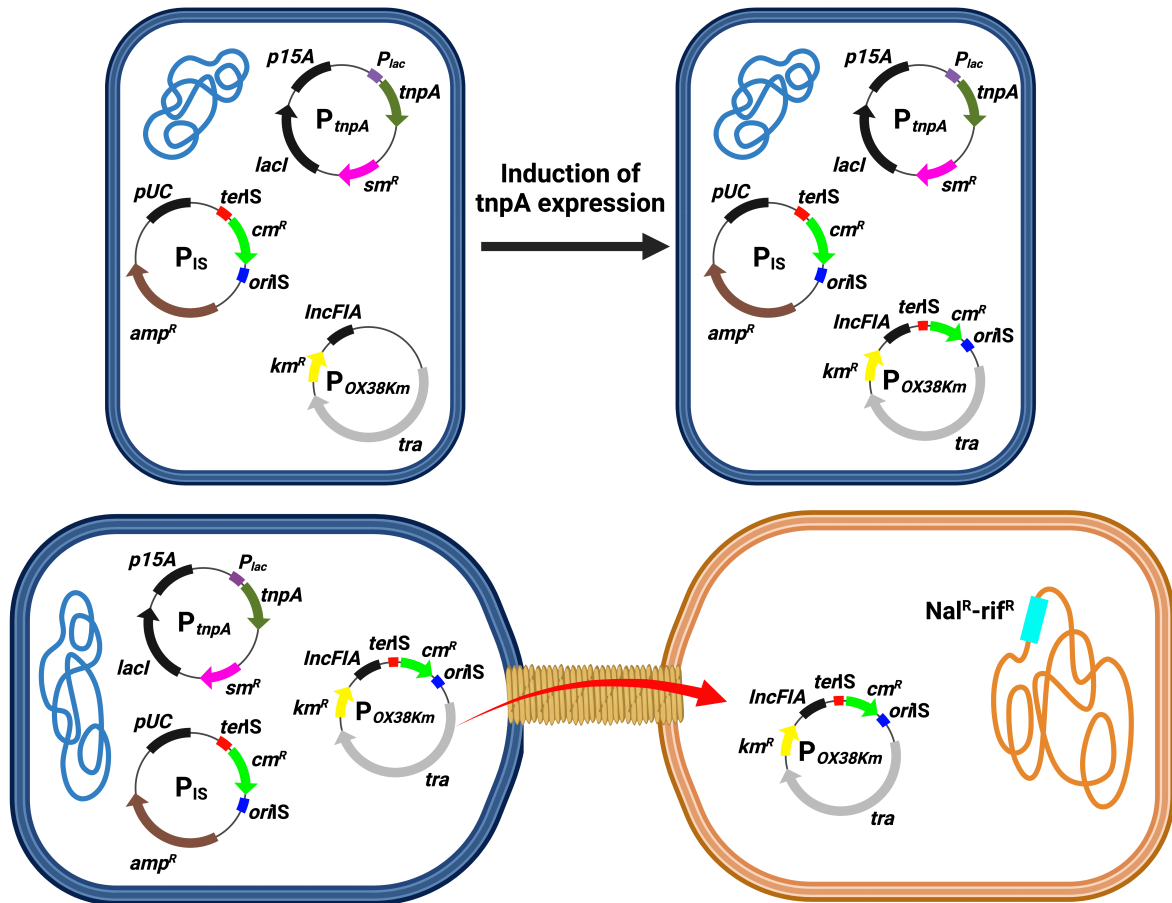
IS91-V1 345- GGAGT--CCCCC-G  
 IS91-V61 345- GGAGT--CCCCC**C**G  
 IS91-V18 345- GGC**GT****TA**CCCCC-G

### Figure S1. Alignment of isoforms with inserted or deleted bases relative to the first sequence described

Alignment of described isoforms with insertions or deletions in the coding sequence of *orf121* with respect to the first described *orf121* sequence of the IS91 element (GenBank: X17114.5.1), as a result of the loss of the -1 overlap between the *tnpA* gene and *orf121*. Insertions are represented in bold and deletions by a -.

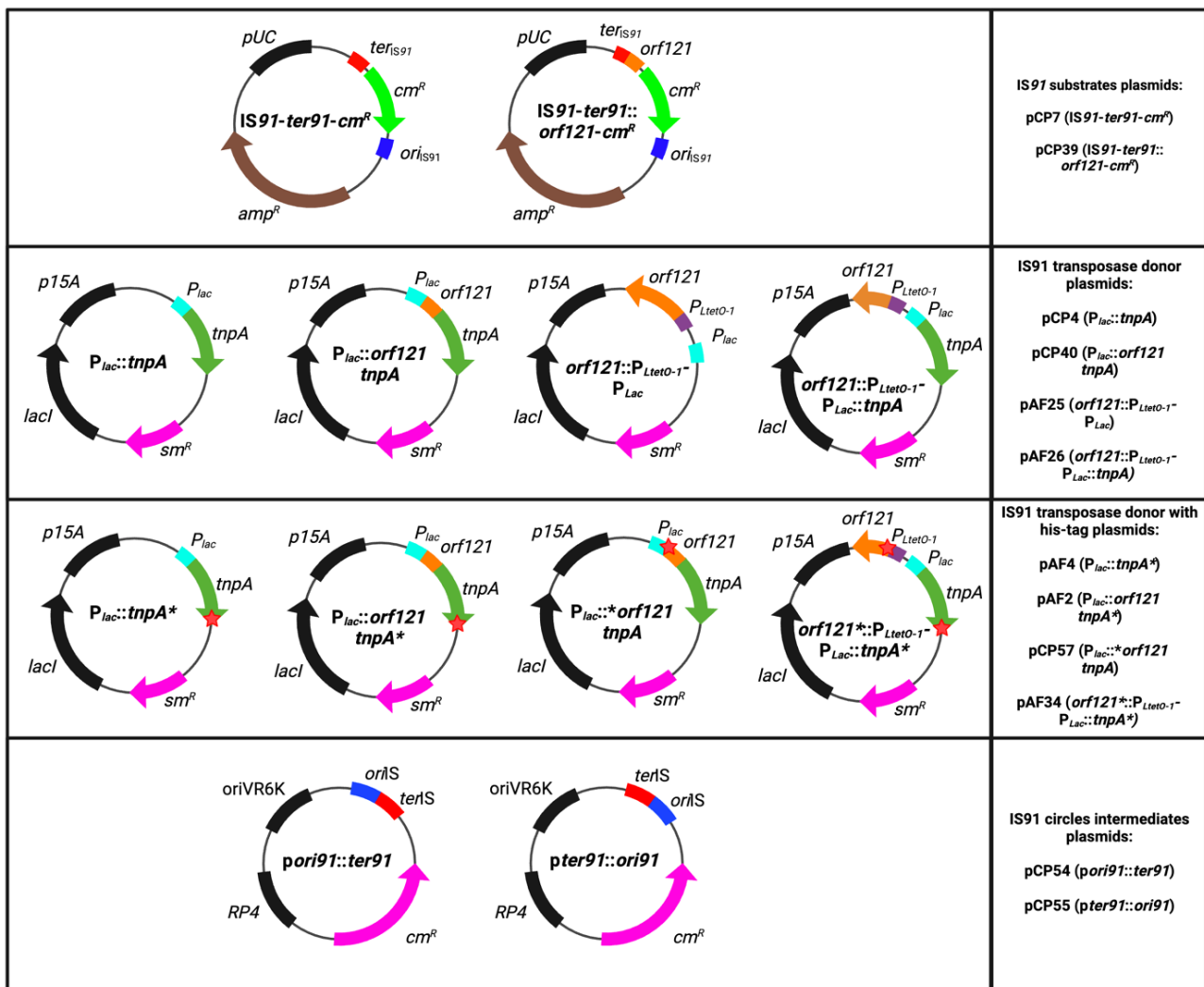
a) Alignment showing deletions in the coding sequence of *orf121*, the deletion of T at position 90 (1 isoform including IS91-V22); the deletion of T at position 90 and of A at position 196 (1 isoform including IS91-V6); the deletion of C at position (2 isoforms including IS91-V3) or the deletion of G at position (12 isoforms including IS91-V13).

b) Alignment showing insertions in the coding sequence of *orf121*, insertion of a G after position 63 (4 isoforms including IS91-V66); insertion of a TA after position 360 (1 isoform including IS91-V18) or insertion of a C after position 356 (1 isoform including IS91-V91).



**Figure S2. Schematic of the mating-out assays**

The *E. coli JS219* carries three plasmids: the IS donor plasmid (a derivative of pUC18; Amp<sup>R</sup> (brown arrow); the mini-IS (*oriIS*, Cm<sup>R</sup> and *terIS* in blue, green and red respectively)), a second plasmid (with an origin of replication p15A; Sm<sup>R</sup> (pink arrow); expressing the IS91 transposase under the control of the IPTG-inducible P<sub>Lac</sub> promoter (khaki)), and the conjugative pOX38Km (a derivative of plasmid F) used as the target for transposition. The donor chromosome and cell are in blue and those of the recipient in orange. The conjugation bridge is stranded and the red arrow indicates the transfer of plasmid pOX38Km from the donor to the recipient cell. Induction of the transposase allows mobilization of the mini-IS from the IS donor plasmid onto the conjugative plasmid. We then use conjugation to isolate the transposition events on the pOX38Km plasmid in the recipient cell. Created with BioRender.com



**Figure S3. Schematic of plasmids used for mating-out assays and suicide conjugation assays**

IS91 substrate plasmids present the *oriIS* (blue) and *terIS* (red) flanking the Cm<sup>R</sup> mobility reporter gene (green), a second marker gene for the OET mechanism which is the Amp<sup>R</sup> gene (brown).

IS91 transposase donor with or without His-tag plasmids feature the P<sub>lac</sub> promoter (cyan) which controls the *tnpA* gene (green) with the upstream presence of *orf121* (orange). When the two genes are expressed under the control of two promoters, the *orf121* gene is under the control of the P<sub>LtetO-1</sub> promoter (purple). These plasmids are selected by the sm<sup>R</sup> resistance gene. The red star represents the position of the His-tag.

IS91 circles intermediates plasmids in which the two junctions *oriIS91* (blue) and *terIS91* (red) are abutted. This plasmid is selected by the Cm<sup>R</sup> gene (pink).



**a)** [A0ABF91SV1|A0ABF91SV1\\_ECOLX](#) Mass: 47331 Score: 520 Queries matched: 19 emPAI: 1.25  
 IS91-like element IS91 family transposase O5=Escherichia coli O141:H4 OX=2861806 GN=K2H89\_26900 PE=4 SV=1  
 Match to: Q8K0X5|Q8K0X5\_ECOLX Score: 520  
 IS91 family transposase O5=Escherichia coli OX=562 GN=tnpA PE=4 SV=1  
 Found in search of 2022\_09\_16\_Bande\_AF4-2.wiff

Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Peptide
1282	355.1965	708.3784	708.3741	6.12	1	26	0.63	1	K.KVCFR.C
2364	460.7411	919.4676	919.4664	1.34	0	32	0.46	1	K.NLHFYAR.K 2365
2504	473.2500	944.4854	944.4828	2.80	0	25	2.9	2	R.GNNVNSR.V
2744	495.2469	988.4792	988.4767	2.61	0	45	0.019	1	R.FSHWVADK.G 2743 2745
3113	537.2867	1072.5588	1072.5593	-0.45	0	51	0.0077	1	R.YYGFSLPVK.R 3114 3116
3515	598.2942	1194.5738	1194.5782	-3.62	0	61	0.00038	1	R.FADIFQQGNR.W
3627	613.8344	1225.6542	1225.6489	4.39	0	71	3.4e-005	1	R.LAELVLJHER.L 3626
4110	500.9176	1499.7310	1499.7368	-3.91	0	74	1.8e-005	1	R.LEHYAGQDEIGLR.Y
4131	509.5533	1525.6381	1525.6409	-1.82	0	46	0.001	1	R.DWNCFLDTHYR.R
4244	571.6437	1711.9093	1711.9105	-0.70	0	70	5.5e-005	1	K.QPEGSVRPVVTVESVTK.I
4369	644.3143	1929.9211	1929.9295	-4.36	0	77	8.1e-006	1	R.VMDNATHAVVYFVGSYK.K
4399	667.6883	2000.0431	2000.0466	-1.76	1	73	2e-005	1	R.KYPELVIPDELAVEGNSK.R 4398

Nominal mass ( $M_r$ ): 47331; Calculated pI value: 9.41

```

1 MLPRFADIFQ QGNRNLNMLE KQPEGSVRPV VTESVTKIMA CGTTLMGYTQ
51 WCCSSPDCCH TKKVCFRCKS RSCPHCGVKA GAQIQYLLS LVPDCPWQHI
101 VFTLPCQWYS LVFHNRLWLA EHSRIAADVI LEICHTDVE PGIFTVIHTW
151 GRDQQWPHI HLSTTAGVVT SGHTWKNLHF YARKVMSMR YRITRLLSRK
201 YPELVIPDEL AVEGNSKRDW NCFLDTHYR GNNVNSRVM DNATHAVVYF
251 GSYLKKPPVP MSRLHYAGQ DEIGLRYNSH RTKREYLLM SGDEFMERFS
301 HWVADKGRM VRYYGFLSPV KRRLLEEVY VITETVRKTA MQIRWRGMYQ
351 RLLKVDPLKC ILCGSRMFT GLKRGYRLAE LVLJHERLAR QQVCG

```

Proteins matching the same set of peptides:  
[Q8K0X5|Q8K0X5\\_ECOLX](#) Mass: 47331 Score: 520 Queries matched: 19  
 IS91 family transposase O5=Escherichia coli OX=562 GN=tnpA PE=4 SV=1

**b)** [A0A2S8J3Z6|A0A2S8J3Z6\\_ECOLX](#) Mass: 47327 Score: 560 Queries matched: 21 emPAI: 0.72  
 IS91 family transposase O5=Escherichia coli OX=562 GN=D5732\_00175 PE=4 SV=1  
 Match to: Q8K0X5|Q8K0X5\_ECOLX Score: 560  
 IS91 family transposase O5=Escherichia coli OX=562 GN=tnpA PE=4 SV=1  
 Found in search of 2022\_09\_16\_Bande\_AF2-2.wiff

Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Peptide
1354	355.1958	708.3770	708.3741	4.14	1	27	0.57	1	K.KVCFR.C
2461	460.7418	919.4690	919.4664	2.86	0	32	0.57	1	K.NLHFYAR.K 2460
2620	473.2495	944.4844	944.4828	1.74	0	33	0.5	1	R.GNNVNSR.V
2839	495.2477	988.4808	988.4767	4.23	0	41	0.048	1	R.FSHWVADK.G 2838
3227	537.2885	1072.5624	1072.5593	2.91	0	55	0.0031	1	R.YYGFSLPVK.R 3226 3228 3229 3231
3666	598.2968	1194.5790	1194.5782	0.73	0	56	0.0014	1	R.FADIFQQGNR.W
3759	409.5579	1225.6519	1225.6489	2.45	0	64	0.00017	1	R.LAELVLJHER.L 3758 3760
1556	750.8760	1499.7374	1499.7368	0.41	0	87	9e-007	1	R.LEHYAGQDEIGLR.Y
4370	571.6445	1711.9117	1711.9105	0.70	0	70	5.4e-005	1	K.QPEGSVRPVVTVESVTK.I 4371
4547	667.6895	2000.0467	2000.0466	0.04	1	85	1.3e-006	1	R.KYPELVIPDELAVEGNSK.R 4545 4546

Nominal mass ( $M_r$ ): 47331; Calculated pI value: 9.41

```

1 MLPRFADIFQ QGNRNLNMLE KQPEGSVRPV VTESVTKIMA CGTTLMGYTQ
51 WCCSSPDCCH TKKVCFRCKS RSCPHCGVKA GAQIQYLLS LVPDCPWQHI
101 VFTLPCQWYS LVFHNRLWLA EHSRIAADVI LEICHTDVE PGIFTVIHTW
151 GRDQQWPHI HLSTTAGVVT SGHTWKNLHF YARKVMSMR YRITRLLSRK
201 YPELVIPDEL AVEGNSKRDW NCFLDTHYR GNNVNSRVM DNATHAVVYF
251 GSYLKKPPVP MSRLHYAGQ DEIGLRYNSH RTKREYLLM SGDEFMERFS
301 HWVADKGRM VRYYGFLSPV KRRLLEEVY VITETVRKTA MQIRWRGMYQ
351 RLLKVDPLKC ILCGSRMFT GLKRGYRLAE LVLJHERLAR QQVCG

```

Proteins matching the same set of peptides:  
[Q8K0X5|Q8K0X5\\_ECOLX](#) Mass: 47331 Score: 560 Queries matched: 21  
 IS91 family transposase O5=Escherichia coli OX=562 GN=tnpA PE=4 SV=1

**c)** [A0AR4ENI7|A0AR4ENI7\\_ECOLX](#) Mass: 14085 Score: 451 Queries matched: 11 emPAI: 0.92  
 IS91, orf2 O5=Escherichia coli OX=562 GN=AH47\_18675 PE=4 SV=1  
 Match to: Q9F440|Q9F440\_ECOLX Score: 451  
 IS91, orf2 O5=Escherichia coli OX=562 GN=NCT10082\_05818 PE=4 SV=1  
 Found in search of 2022\_09\_16\_Bande\_CP57-2.wiff

Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Peptide
862	300.6966	599.3786	599.3755	5.30	1	37	0.098	1	R.SKLPY.R 863 864
4908	632.2963	1262.5780	1262.5779	0.13	0	83	2.1e-006	1	R.VDDEGYPETAR.C 4909 4902
6259	649.6301	1945.8685	1945.8694	-0.47	0	88	3.4e-007	1	K.LHPDFFDEEAEVLR.F 6255 6256 6257 6258

Nominal mass ( $M_r$ ): 14031; Calculated pI value: 5.00

```

1 NARSAKPKR KPAPQSKLP RYVVKLHPDD FFDEEAEVLR RFDNFDDAVE
51 CCADLNIPFF VDAWKKLVF WFRVDDDEGY PEIARCTERE FATILAGISA
101 GGHYCEPGT VHWPDGPPPP F

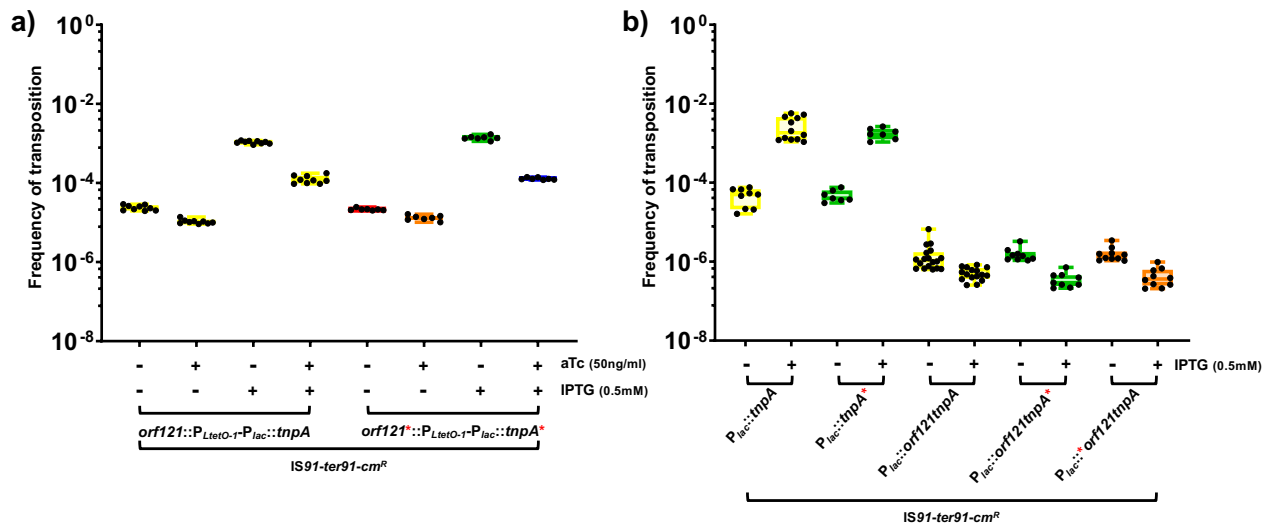
```

Proteins matching the same set of peptides:  
 Uncharacterized protein O5=Escherichia coli TW10509 OX=656449 GN=ERFG\_04669 PE=4 SV=1  
[Q9F440|Q9F440\\_ECOLX](#) Mass: 14031 Score: 451 Queries matched: 11

## Figure S4. Peptides sequences determined by Mass Spectrometry

Analysis of His-tagged ORF121 or TnpA<sub>IS91</sub> proteins by mass spectrometry as described in Materials and Methods.

- a) Detection of TnpA protein when cells expressed *tnpA*-only,  $P_{Lac}::tnpA^*$  (from pAF4 plasmid), identification of the TnpA transposase with 21 peptides (10 distinct) and 26% coverage.
- b) Detection of TnpA protein when cells expressed both *orf121* and *tnpA*,  $P_{Lac}::orf121tnpA^*$  (from pAF2 plasmid), identification of the TnpA transposase with 19 peptides (12 distinct) and 33% coverage.
- c) Detection of Orf121 protein when cells expressed both *orf121* and *tnpA*,  $P_{Lac}::orf121tnpA$  (from pCP57 plasmid), identification of the TnpA transposase with 11 peptides (3 distinct) and 26% coverage.

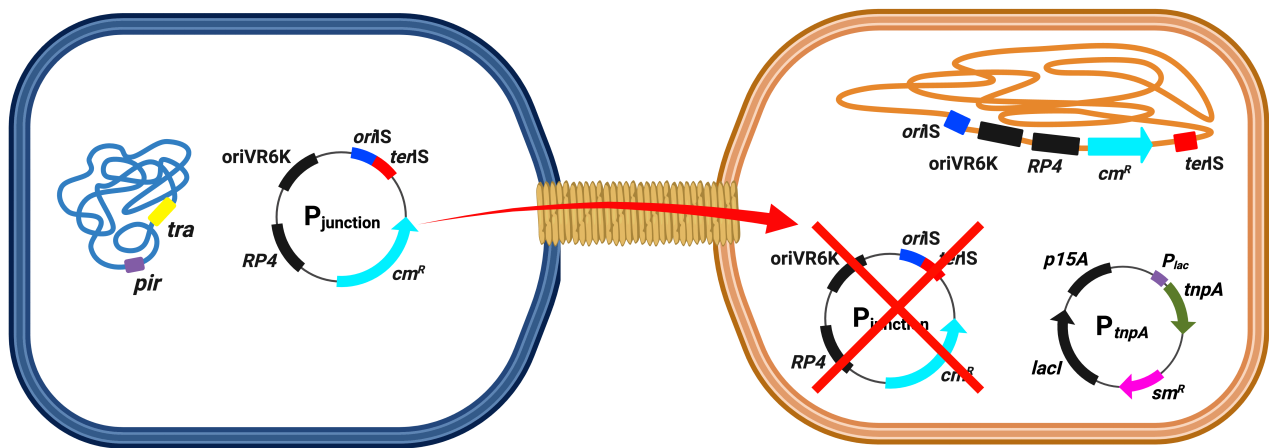


**Figure S5. The addition of His-tag does not modify the activity of transposase or Orf121**

Mating-out assays as described in Materials and Methods. Measured transposition frequencies (Error bars indicate the SDs;  $n \geq 5$ ). #, undetectable transposition. The - and + indicate the presence or absence of inducer at 0.5 mM IPTG or 50 ng/ml aTc. Empty, empty transposase expression vector;  $P_{lac}::tnpA$ , *tnpA* gene alone is expressed *in trans*;  $P_{lac}::orf121tnpA$ , both ORF121 and *tnpA* gene are expressed *in trans*; *orf121::P<sub>LtetO-1</sub>-P<sub>Lac</sub>::tnpA* are expressed independently

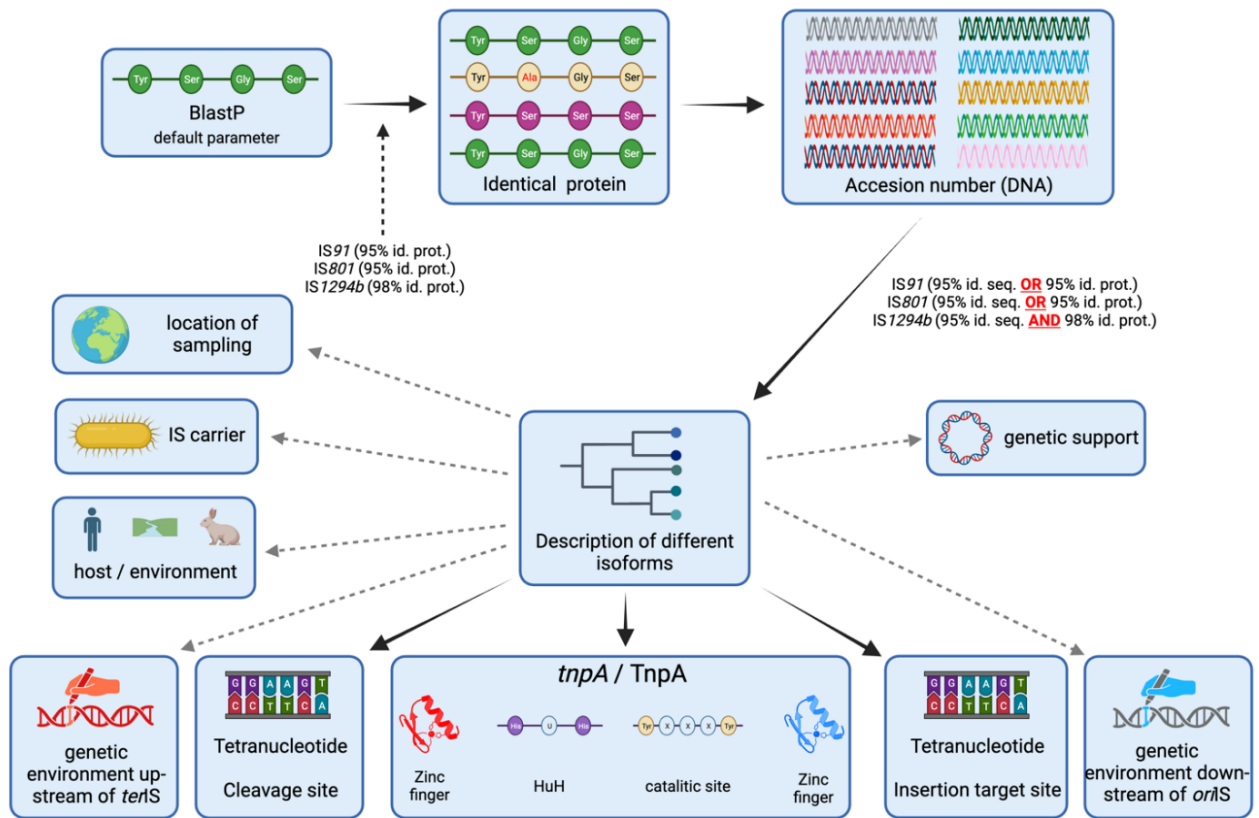
a) *orf121* and *tnpA* are under the control of two different promoters.

b) the *orf121* gene is expressed *in cis* or *in trans*.



**Figure S6. Schematic of the suicide conjugation assays to measure insertions of the IS91 ssDNA circles into the *E. coli* chromosome**

The *E. coli*  $\beta 2163$  donor strain (blue cell) carries the junction plasmid carrying a chloramphenicol resistance gene (*cm<sup>R</sup>*, cyan arrow), the *oriIS* and *terIS* region represented by blue and red rectangles, respectively. The chromosome of the donor strain carries the genes for transfer of the suicide plasmid to the recipient cell. The red arrow indicates conjugation and transfer of the suicide plasmid from the donor cell to the recipient (orange cell). The suicide plasmid cannot be replicated in the recipient cell. This *E. coli* *DH5a* recipient cell carries the plasmid expressing the IS91 transposase under the control of the IPTG-inducible *P<sub>Lac</sub>* promoter (khaki) and the *sm<sup>R</sup>* resistance gene (pink arrow). Expression of the transposase will induce cleavage of the suicide plasmid and allow insertion into the chromosome of the recipient cell.



**Figure S7. Principle of GenBank® in silico analysis of elements IS91, IS801 and IS1294b**

For the description of the figure refer to the materials and methods: “*In silico* analysis of IS91 DNA sequences from the public database”. Created with BioRender.com

## SUPPLEMENTAL TABLES AND LEGENDS

**Table S1. Host of bacteria carrying IS91, IS801, or IS1294b**

	<b>IS91</b>	<b>IS1294b</b>	<b>IS801</b>
<b>n</b>	924	859	74
<b>Animals</b>	451	125	0
<b>Homo Sapiens</b>	216	225	0
<b>Plants</b>	1	1	29
<b>Environment</b>	19	37	0
<b>Food</b>	1	21	0
<b>Unknown</b>	236	450	45

**Table S2. Bacteria carrying IS91, IS801 and IS1294b elements**

	<b>IS91</b>	<b>IS1294b</b>	<b>IS801</b>
n	924	859	74
<i>Escherichia coli</i>	693	454	0
<i>Salmonella enterica</i>	222	77	0
<i>Shigella sonnei</i>	6	172	0
<i>Shigella flexneri</i>	0	82	0
<i>Shigella dysenteriae</i>	0	38	0
<i>Pseudomonas savastanoi</i>	0	0	53
<i>Klebsiella pneumoniae</i>	1	23	0
<i>Pseudomonas syringae</i>	0	0	15
<i>Shigella boydii</i>	0	10	0
<i>Pseudomonas amygdali</i>	0	0	3
<i>Pseudomonas avellanae</i>	0	0	2
<i>Pseudomonas ficuserectae</i>	0	0	1
<i>Providencia sp.</i>	1	0	0
<i>Zoogloea sp.</i>	1	0	0
<i>Escherichia marmotae</i>	0	1	0
<i>Citrobacter farmeri</i>	0	1	0
<i>Enterobacter cloacae</i>	0	1	0

**Table S3. Genetic support of IS**

	<b>IS91</b>	<b>IS1294b</b>	<b>IS801</b>
n	924	859	74
<b>Plasmid</b>	671	565	27
<b>Chromosome</b>	26	27	39
<b>Unknown</b>	227	267	9

**Table S4. Number of copies of IS91, IS801 and IS1294b per replicon**

<b>n</b>	<b>IS91</b>	<b>IS1294b</b>	<b>IS801</b>
	671	565	74
<b>1</b>	790	645	14
<b>2</b>	12	51	4
<b>3</b>	11	14	1
<b>4</b>	11	9	4
<b>5</b>	2	3	0
<b>6</b>	0	3	0
<b>7</b>	2	0	1
<b>8</b>	0	0	1
<b>9</b>	1	0	2



**Table S5. Origin of replication of plasmids carrying IS**

	<b>IS91</b>	<b>IS1294b</b>	<b>IS801</b>
<b>n</b>	671	565	27
<b>IncFII</b>	366	321	0
<b>R64 / IncI1 family</b>	38	45	0
<b>ColE1</b>	16	11	0
<b>IncI1-I(alpha)</b>	10	12	1
<b>IncFIB</b>	3	5	0
<b>IncHI</b>	2	4	0
<b>IncC</b>	1	4	0
<b>IncN</b>	2	1	0
<b>IncI2</b>	1	1	0
<b>IncY</b>	9	0	0
<b>Inc</b>	4	0	0
<b>RK2</b>	4	0	0
<b>IncP</b>	2	0	0
<b>IncHI2</b>	1	0	0
<b>Col</b>	0	2	0
<b>RepHI2</b>	0	1	0
<b>unknown</b>	212	158	26

**Table S6. Insertion target and cleavage sites**

The n value indicates the accession number analysed. X indicate the nucleotide is not known.

<b>INSERTION TARGET SITE</b>			
	IS91	IS1294b	IS801
<b>N</b>	924	859	74
<b>5'-GTTC</b>	590	495	33
<b>5'-CTTG</b>	208	137	24
<b>5'-CTCG</b>	64	92	3
<b>5'-GTCC</b>	20	18	9
<b>5'-CTGG</b>	4	74	2
<b>5'-GTAC</b>	5	2	3
<b>5'-GTGC</b>	3	2	0
<b>5'-GTTG</b>	2	0	0
<b>5'-TTGC</b>	1	0	0
<b>5'-TTTA</b>	1	0	0
<b>5'-TTAT</b>	0	1	0
<b>5'-CTAX</b>	0	1	0
<b>5'-XXXX</b>	26	37	0

<b>CLEAVAGE SITE</b>			
	IS91	IS1294b	IS801
<b>N</b>	924	859	74
<b>5'-CTCG</b>	923	0	0
<b>5'-GTTC</b>	0	858	74
<b>5'-CTCT</b>	2	0	0
<b>5'-GTAC</b>	0	1	0

**Table S7. Analysis of target insertion sites to insertions of the IS91 ssDNA circles into the *E. coli* chromosome**

The n value indicates the number of clones analysed by AP-PCR and the - and + indicate the presence or absence of inducer at 0.5 mM IPTG or 50 ng/ml aTc.

a) Analysis of the target insertion sites of the IS mimicking IS91 in relation to the expression of *orf121* in cis or in trans of *orf121*.

b) Analysis of the target insertion sites of the IS mimicking when *orf121* and *tnpA* are under the control of two different promoters.

a)

		<b><i>P<sub>lac</sub>::tnpA</i>   <i>P<sub>lac</sub>::orf121tnpA</i></b>		
		n	50	20
Insertion target site	5'-CTTG		24	15
	5'-GTTC		13	2
	5'-CTCG		8	1
	5'-GTCC		4	2
	5'-GTAC		1	0

b)

		<b><i>P<sub>LtetQ-1</sub>::orf121-P<sub>lac</sub>::tnpA</i></b>			
		IPTG (0.5 mM)		aTc (50 ng/ml)	
		-	-	+	+
		-	+	-	+
n		57	26	58	58
Insertion target site	5'-GTTC	29	18	27	27
	5'-CTTG	24	8	19	21
	5'-CTCG	3	0	8	10
	5'-GTCC	1	0	2	0
	5'-GTAC	0	0	2	0

**Table S8. Oligonucleotides used for linkers formation, plasmids construction and arbitrary PCR experiments**

<sup>a</sup> tags with restriction site are in bold.

Oligonucleotide	Oligonucleotides sequences (5' → 3') <sup>a</sup>	Use
<b>linkers formation</b>		
CP17 CP18	TATGGGTACCGTCGACGGATCCGCATG CGGATCCGTCGACGGTACCCA	Formation of linker79 (construction of pCP2)
CP21 CP22	AATTCCGAGTAGGCAGCCTGGCGGCTGCGGCT TGTCATGGTCTGGAATTACCGTTATAAAAAA GATAATGTCATTGTCTTTCAGGTAGTG GATCCACTACCTGAAAGACAATGACATTATCT TTTTTATAACGGTAATTCCAGACCATGACAAG CCGCAGCCGCCAGGCTGCCTACTCGG	Formation of linkerTer91 (construction of pCP6)
CP19 CP20	GTCCGCTCATATGGTGCACAAGGGGTGTTGAA GAAACATCCGTTTTGTGGTGCTTTTTTAGTCTT TTGGGATTTAAATTCCTATCGATCAAGA AGCTTCTTGATCGATAGGAATTTAAATCCCCA AAAGACTAAAAAGCACCACAAAACGGATGT TTCTTCAACACCCCTTGTGCACCATATGAGCGG ACTGCA	Formation of linkerOri91 (construction of pCP7)
<b>Plasmids construction</b>		
CP23 CP24	TACTGGATCCTGAGACGTTGATCGGCAC TGAGCTGCAGTGCCACTCATCGCAGTAC	Amplification of the P <sub>cat</sub> :Cam <sup>R</sup> cassette (construction of pCP5)
CP56 CP75	CCGGAATTCGAGTAGGCAGCCTGGCGGC CGCGGATCCTCAGAAGGGCGGGGGACTCC	Amplification of <i>terIS91-orf121</i> -P <sub>tmpA</sub> (construction of pCP39)
CP124 CP125	CCGCTGCAGTTGCTTTCAGGAAAATTTTTCTG CCGCTGCAGATTCTCACCAATAAAAAACG	Addition of the tetR gene to repress the P <sub>LietO-1</sub> promoter in plasmids pAF25, pAF26 and pAF34
CP56 CP75	CCGGAATTCGAGTAGGCAGCCTGGCGGC CGCGGATCCTCAGAAGGGCGGGGGACTCC	Amplification of <i>terIS91-orf121</i> (pCP39 construct)
CP15 CP16	TATACATATGCTTCCCCGTTTTGCCGAC GCGTGGATCCTATTAACAGGATGACTG	Amplification of the CDS of <i>tnpA<sub>IS91</sub></i> (pCP4 construct)
AF4 CP58	ATATGGATCCTCAATGATGATGATGATGATGG CCGCACACCTGCTGTGC TATACATATGGCCCGTTCAGCTAAACC	Amplification of the <i>tnpA<sub>IS91</sub></i> CDS with a C-terminal His6 tag (pAF4 construct)

CP58 CP16	TATACATATGGCCCGTTCAGCTAAACC GCGTGGATCCTATTAACAGGATGACTG	Amplification of the CDS of <i>orf121-tnpA<sub>IS91</sub></i> (pCP40 construct)
AF4 CP58	ATATGGATCCTCAATGATGATGATGATGATGG CCGCACACCTGCTGTCG TATACATATGGCCCGTTCAGCTAAACC	Amplification of the CDS of <i>orf121-tnpA<sub>IS91</sub></i> with a His6 tag at the C-terminus of <i>tnpA<sub>IS91</sub></i> (pAF2 construct)
CP16 CP86	GCGTGGATCCTATTAACAGGATGACTG TATACATATGCATCATCATCATCATGCCC GTTTCAGCTAAACC	Amplification of the CDS of <i>orf121-tnpA<sub>IS91</sub></i> with an N-terminal His6 tag of <i>orf121</i> (pCP57 construct)
CP88 CP89	CTGCCCGCTTTCTCGAGTCCCTATCAGTGAT AGAGATTG CCATATAACTACCTCGGTCAGTGCCTCCTG	Amplification of the P <sub>LtetO-1</sub> portion (170 bp)
CP90 CP91	TCTAGAATTCTCAGAAGGGCGGGGGGACTC GGACGCACTGACCGAGGTAGTTATATGGC	Amplification of the <i>orf121</i> portion (386 pb)
CP88 CP90	CTGCCCGCTTTCTCGAGTCCCTATCAGTGAT AGAGATTG TCTAGAATTCTCAGAAGGGCGGGGGGACTC	Amplification of the <i>orf121::P<sub>LtetO-1</sub></i> (pAF25 construct)
AF68 CP16	CGGGGTACCATGCTTCCCCGTTTTGCCGAC GCGTGGATCCTATTAACAGGATGACTG	Amplification of the CDS of <i>tnpA<sub>IS91</sub></i> (pAF26 construct)
AF69 AF71 AF70 AF73 AF72 AF74	ACACTCCGGGCAGTACATGCCGCGGGCGCTG ATACCGCAAGAATG GCACTGACCGAGGTAGTTATATGCATCATCAT CATCATCATGC ATGATGATGATGATGATGCATATAACTACCTC GGTCAGTGC GTCGGCAAACGGGGAAGCATATGTATATCTC CTTCTTAAA TTTAAGAAGGAGATATACATATGCTTCCCCGTT TTGCCGAC TCTGTATGGAACGGGCATGCGGATCCTCAATG ATGATGATGATGATG	Amplification of the CDS of <i>orf121::P<sub>LtetO-1</sub>-P<sub>Lac</sub>::tnpA<sub>IS91</sub></i> with a His6 tag at the N-terminus of <i>orf121</i> and at the C-terminus of <i>tnpA<sub>IS91</sub></i> (pAF34 construct; <i>in fusion</i> )
CP79 Cm-1049	GCTGCCTACTCGATCGATAGGAATTTAAATCC GCTTCCATGTCCGCAGATGC	Amplification of the <i>oriIS91</i> portion (159 bp)
CP80 Ter91-337	ATTCTATCGATCGAGTAGGCAGCCTGGCGGC TG CAACGTCTCAGGATCCACTACCTG	Amplification of the <i>terIS91</i> portion (113 bp)
CP79	GCTGCCTACTCGATCGATAGGAATTTAAATCC	

CP81	ATATGGATCCTCCGCTCATATGGTGCACAAG	Amplification of the <i>oriIS91</i> -BamHI portion (110 bp)
CP80	ATTCTATCGATCGAGTAGGCAGCCTGGCGGC	Amplification of the <i>terIS91</i> -PstI portion (107 bp)
CP82	GTGGCTGCAGACTACCTGAAAGACAATGAC	
CP81	ATATGGATCCTCCGCTCATATGGTGCACAAG	Amplification of the Junc <i>terIS91</i> - <i>oriIS91</i> -INV fusion fragment by junction PCR (pCP54/55 construct)
CP82	GTGGCTGCAGACTACCTGAAAGACAATGAC	
CP110	CGCGAATTCATGGCCCGTTCAGCTAAAC	Amplification of the fragment containing the $P_{mpA}$ promoter (boxes -35 and -10) with RBS <sub><i>lacZ</i></sub> (pCP65 construct)
CP116	GCGGGATCCATGGCTGTTTCCTGTCCGGCCAG TGAACCGTG	
CP95	CGCGAATTCGAGTAGGCAGCCTGGCGGC	Amplification of the fragment containing the $P_{orf121}$ promoter (boxes -35 and -10) with RBS <sub><i>lacZ</i></sub> (pCP66 construct)
CP96	GCGGGATCCATGGCTGTTTCCTGAAAGACAAT GACATTATC	
CP110	CGCGAATTCATGGCCCGTTCAGCTAAAC	Amplification of the fragment containing the $P_{mpA}$ promoter (box -35 and -10) set with RBS <sub><i>mpA</i></sub> and an overlap between the stop codon of <i>orf121</i> and start of <i>lacZ</i> (pCP73 construct)
CP121	GCGGGATCCATCAGAAGGGCGGGGGGA	
CP95	CGCGAATTCGAGTAGGCAGCCTGGCGGC	Amplification of the fragment containing the $P_{orf121}$ promoter (boxes -35 and -10) set with RBS <sub><i>orf121</i></sub> and an overlap between the stop codon of <i>orf121</i> and start of <i>lacZ</i> (pCP74 construct)
CP121	GCGGGATCCATCAGAAGGGCGGGGGGA	
CP110	CGCGAATTCATGGCCCGTTCAGCTAAAC	Amplification of the fragment containing the $P_{mpA}$ promoter (box -35 and -10) set with RBS <sub><i>mpA</i></sub> and WITHOUT an overlap between the stop codon of <i>orf121</i> and start of <i>lacZ</i> (pCP75 construct)
CP128	GCGGGATCCATTCAGAAGGGCGGGGGGAC	

CP95	CGCGAATTCGAGTAGGCAGCCTGGCGGC	Amplification of the fragment containing the $P_{orf121}$ promoter (box -35 and -10) set with $RBS_{orf121}$ and WITHOUT an overlap between the stop codon of <i>orf121</i> and start of <i>lacZ</i> (pCP76 construct)
CP128	GCGGGATCCATTCAGAAGGGCGGGGGGAC	
CP110 AF80	CGCGAATTCATGGCCCGTTCAGCTAAAC CTGATACCCGCCAGAATGGTCGCAAAC	Insertion of a mutation in the -35 box of the $P_{mpA}$ promoter left portion (305 bp)
CP121 AF79	GCGGGATCCATCAGAAGGGCGGGGGGA CGACCATTCTGGCGGGTATCAGCGCC	Insertion of a mutation in the -35 box of the $P_{mpA}$ promoter on the right-hand side (102 bp)
CP110 CP121	CGCGAATTCATGGCCCGTTCAGCTAAAC GCGGGATCCATCAGAAGGGCGGGGGGA	Amplification of the Junc fusion fragment mutated in box -35 of the $P_{mpA}$ TTGCCG promoter into TGGCGG (pAF36 construct)
CP95 AF80	CGCGAATTCGAGTAGGCAGCCTGGCGGC CTGATACCCGCCAGAATGGTCGCAAAC	Insertion of a mutation in the -35 box of the $P_{mpA}$ promoter, left portion (392 bp)
AF79 CP121	CGACCATTCTGGCGGGTATCAGCGCC GCGGGATCCATCAGAAGGGCGGGGGGA	Insertion of a mutation in the -35 box of the $P_{mpA}$ promoter on the right-hand side (102 bp)
CP95 CP121	CGCGAATTCGAGTAGGCAGCCTGGCGGC GCGGGATCCATCAGAAGGGCGGGGGGA	Amplification of the Junc fusion fragment mutated in box -35 of the $P_{mpA}$ TTGCCG promoter to TGGCGG (pAF37 construct)
CP95 AF78	CGCGAATTCGAGTAGGCAGCCTGGCGGC TTATCTTTTTTCGCGACGGTAATTCCAG	Insertion of a mutation in the -10 box of the $P_{orf121}$ promoter, left portion (72 bp)
AF77 CP121	AATTACCGTCGCGAAAAAAGATAATGTC GCGGGATCCATCAGAAGGGCGGGGGGA	Insertion of a mutation in the -10 box of the $P_{orf121}$ promoter right portion (424 bp)
CP95 CP121	CGCGAATTCGAGTAGGCAGCCTGGCGGC GCGGGATCCATCAGAAGGGCGGGGGGA	Amplification of the Junc fusion fragment mutated in box -10 of the $P_{orf121}$ TATAAA promoter into CGCGAA (pAF38 construct)

Arbitrary-primed (AP)-PCR		
ARB1 Cm-1005	GGCCACGCGTCGACTAGTACNNNNNNNNNNAT CGG GCAGATTACGCGCAG	First round of arbitrary PCR to map IS91 insertion sites into pOX38Km
ARB2 Cm-1049	GGCCTCGCGTCGACTACTTC GCTTCCATGTGCGCAGATGC	Second round of arbitrary PCR to map IS91 insertion sites into pOX38Km
ARB1bis Cm1005-Bis54	GGCCTCGCGTCGACTACTTCNNNNNNNNNNGA CTG CCCTTATTCGCACCTGGCGG	First round of arbitrary PCR to map IS91 insertion sites into <i>E. coli</i> chromosome
ARB2 Cm1049-Bis54	GGCCTCGCGTCGACTACTTC GGCGTCGACGGTATCGATAAGC	Second round of arbitrary PCR to map IS91 insertion sites into <i>E. coli</i> chromosome
Ori91-1104	GCAGTCCGCTCATATGGTGC	Sequencing of arbitrary PCR products
ARB1 Cm-473	GGCCACGCGTCGACTAGTACNNNNNNNNNNAT CGG ATCAACGGTGGTATATCCAGTG	First set of arbitrary PCRs on the terIS91 side in pOX38Km with ARB1
ARB2 Ter91-337	GGCCTCGCGTCGACTACTTC CAACGTCTCAGGATCCACTACCTG	Second set of arbitrary PCRs on the terIS91 side in pOX38Km with ARB2 (pCP7)
ARB2 Orf121-484	GGCCTCGCGTCGACTACTTC GTCTGCGCAGCACTCAACG	Second set of arbitrary PCRs on the terIS91 side in pOX38Km with ARB2 (pCP39)
Ter91-276	TCCAGACCATGACAAGCCGC	For Sanger sequencing of the <i>terIS91</i> side (pCP7)
Ter91-318	CTACCTGAAAGACAATGAC	For Sanger sequencing of the <i>terIS91</i> side (pCP39)
qRT-PCR primer		
ORF121 F	GTGCTGCGCAGACCTGAATAT	<i>orf121</i>
ORF121 R	ACTCCGGGCAGTACATGCC	
TnpA F	TCCCTGGTGTTCACAACCGGT	<i>tnpA</i>
TnpA R	CACACCACCGGCAGTTGTCG	
dxs-LC3	ATGACGTGGCGATTCAAAA	housekeeping gene <i>dxs</i> (1)
dxs-LC4	AGCCGGTATAGAGCATCTGG	

(1) Baltazar M, Bourgeois-Nicolaos N, Larroude M, Couet W, Uwajenezza S, Doucet-Populaire F, Ploy MC, Da Re S. Activation of class 1 integron integrase is promoted in the intestinal



environment. PLoS Genet. 2022 Apr 28;18(4):e1010177. doi: 10.1371/journal.pgen.1010177.  
PMID: 35482826; PMCID: PMC9090394.

## I.2. Résultats complémentaires de l'analyse *in silico*

Premièrement, nous avons réalisé une analyse du polymorphisme des promoteur  $P_{orf121}$  et  $P_{tnpA}$  des isoformes IS91 décrites.

L'analyse du **promoteur  $P_{orf121}$**  (figure 47) montre que la région promotrice -35 ne présente aucune mutation dans la séquence nucléotidique. En revanche, un grand nombre d'isoformes présente un nucléotide supplémentaire T ou A entre le triplet TTA et TAA dans la boîte -10 avec des délétions d'un ou deux nucléotides en amont ou en aval de ces nucléotides supplémentaires. Concernant la séquence SD, une mutation unique est retrouvée de AGG vers TGG.

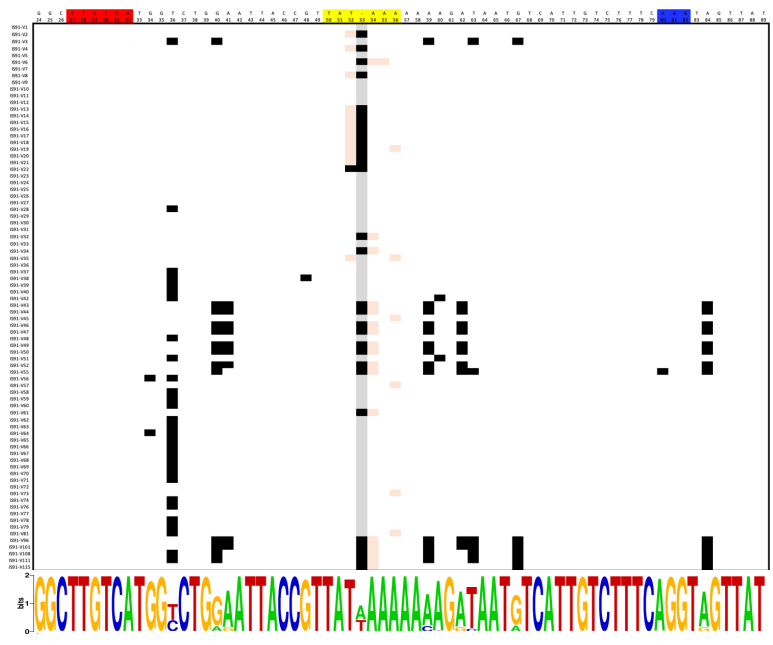


Figure 47 : Distribution des polymorphismes des nucléotides du promoteur  $P_{orf121}$

Représentation de la distribution des polymorphismes en fonction de l'isoforme décrite et de la position de la mutation dans le promoteur  $P_{orf121}$ , par rapport à la première séquence IS91 décrite (GenBank : X17114.5.1). En haut de la matrice, il est indiqué en rouge la boîte -35, en jaune la boîte -10 et en bleu la séquence SD. Dans la matrice thermique, les mutations sont représentées en noir. En saumon est indiqué l'absence des nucléotides. Les colonnes en fond gris indiquent les nucléotides présents uniquement chez certaines isoforme décrites (noir) et absents de la plupart des isoformes dont la première séquence.

L'analyse du **promoteur  $P_{tnpA}$**  est présentée en figure 48. Les boîtes -10 et -35 ne présentent aucune mutation malgré une grande variabilité autour. *A contrario*, des mutations de la séquence SD sont observées, telles que GGAG vers GGCG, pour la plupart des isoformes.

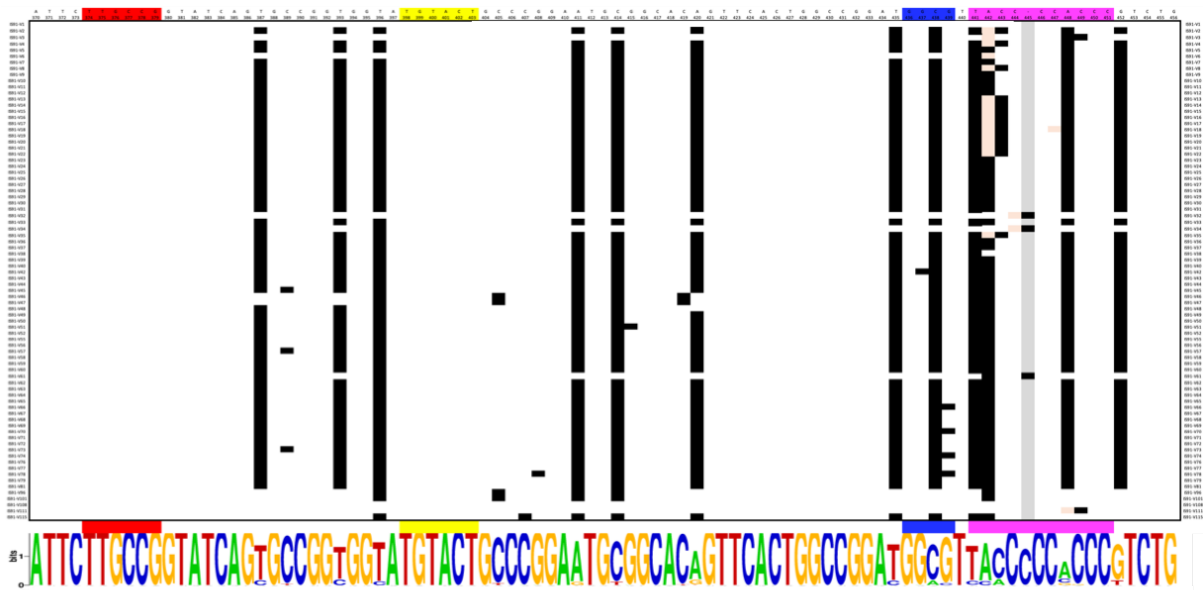


Figure 48 : Distribution des polymorphismes des nucléotides du promoteur  $P_{tnpA}$

Représentation de la distribution des polymorphismes en fonction de l'isoforme décrite et de la position de la mutation dans le promoteur  $P_{tnpA}$ , par rapport à la première séquence IS91 décrite (GenBank : X17114.5.1). En haut de la matrice, il est indiqué en rouge la boîte -35, en jaune la boîte -10, en bleu la séquence SD et en rose une région homo-polymérique. Dans la matrice thermique, les mutations sont représentées en noir et l'absence de mutation en blanc. En saumon est indiqué l'absence des nucléotides. Les colonnes en fond gris indiquent les nucléotides présents uniquement chez certaines isoforme décrites (noir) et absents de la plupart des isoformes dont la première séquence.

**Deuxièmement**, nous avons réalisé une analyse des extrémités *oriS* et *terS* des isoformes IS1294b et IS91.

Une analyse des extrémités *terS*1294b et *oriS*1294b a été réalisée à l'aide d'une carte de la distribution des polymorphismes en fonction de l'isoforme avec les palindromes de chaque extrémité (figure 49). L'extrémité *terS*1294b montre une forte conservation et aucune mutation n'est présente dans le palindrome, alors qu'une grande variabilité est observée pour l'extrémité *oriS*1294b au niveau du palindrome (Yassine, 2015). En revanche, l'extrémité proche du tétranucléotide est très conservée. Il a été montré que les nucléotides T<sub>1</sub>, C<sub>3</sub>, A<sub>6</sub>, T<sub>7</sub>, C<sub>8</sub>, T<sub>10</sub> et A<sub>14</sub> ont un effet délétère sur l'activité de transposition lors de leur mutation (Yassine, 2015). Ainsi, il a été proposé que ces nucléotides soient les sites de liaison de la transposase *via* des interactions spécifiques. En outre, l'alignement de l'ensemble des variants décrits démontre que les 20 premiers nucléotides ( $3^{\text{T}}_{20}$ TTTTCAATTCCTATACGT $1^{\text{G}}$  $5^{\text{T}}$ ) de l'extrémité sont tous conservés chez les 225 isoformes décrites. Ceci corrobore les données expérimentales de

Haytham Yassine et nous permet de supposer que les 20 premiers nucléotides sont nécessaires pour la liaison de la transposase de IS1294b à cette extrémité.

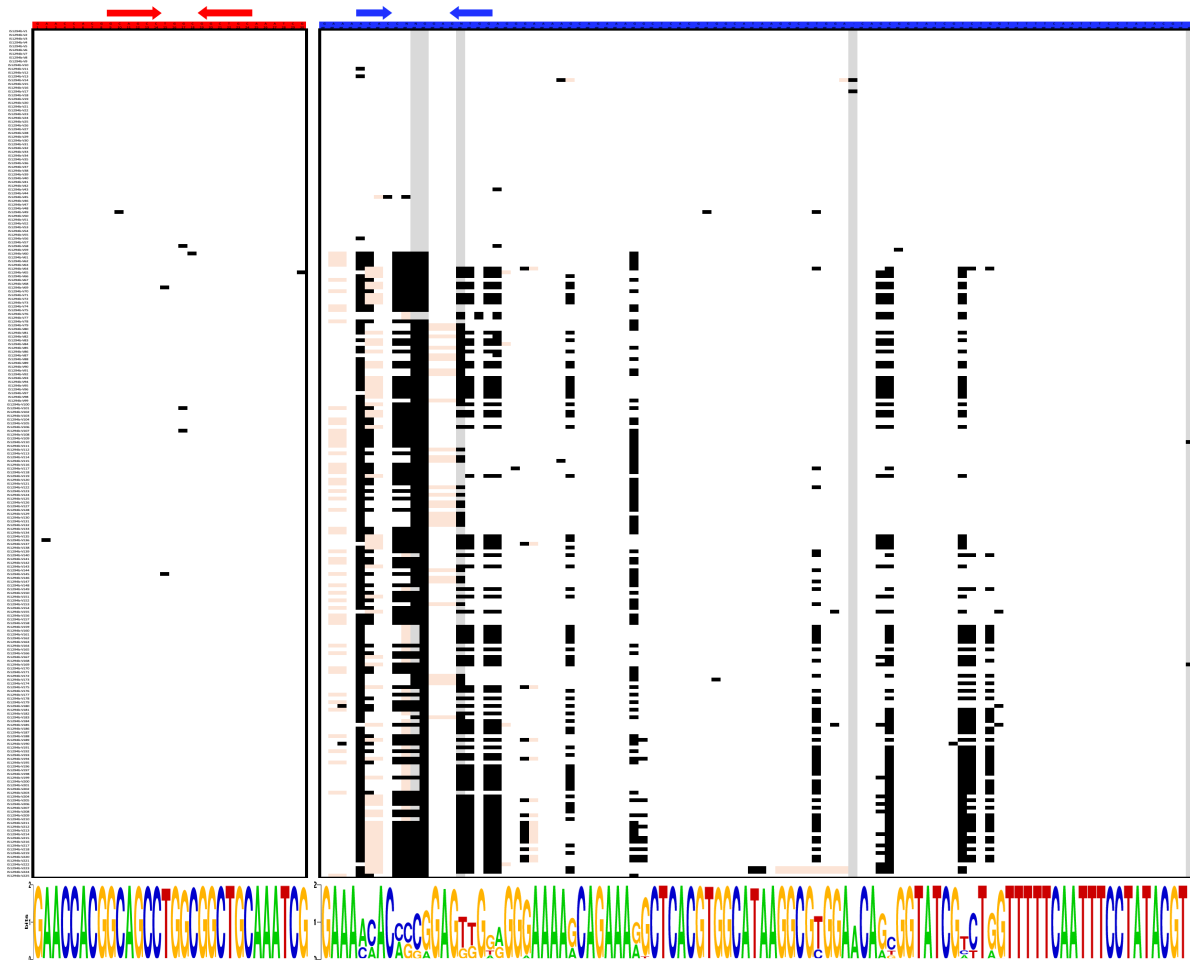


Figure 49 : Distribution des polymorphismes des extrémités *terIS1294b* et *oriIS1294b*

Représentation de la distribution des polymorphismes en fonction de l'isoforme décrite et de la position de la mutation dans la séquence *terIS* (à gauche, haut rouge) et la séquence *oriIS* (à droite, haut bleu) par rapport à la première séquence IS1294b décrite (GenBank : KP017243.1). Chaque palindrome est indiqué par de grandes flèches horizontales rouge pour l'extrémité *terIS* et bleu pour l'extrémité *oriIS*. Dans la matrice thermique, les mutations sont représentées en noir. En saumon est indiqué l'absence de ces nucléotides. Les colonnes en fond gris indiquent les nucléotides présents uniquement chez certaines isoforme décrites (noir) et absents de la plupart des isoformes dont la première séquence.

Une analyse des extrémités *terIS91* et *oriIS91* a été réalisée à l'aide d'une carte de la distribution des polymorphismes en fonction de l'isoforme avec les palindromes de chaque

extrémité (figure 50). L'extrémité *terIS91* montre qu'elle ne possède aucun polymorphisme dans sa séquence. Celle-ci est conservée dans la diversité des isoformes décrites.

*A contrario*, l'extrémité *oriIS91* présente une variabilité importante. En effet, le polymorphisme est observé dans les palindromes subterminaux ainsi que dans l'espace inter-palindromique. Mendiola et al. ont montré que les 82 pb de l'extrémité *oriIS91* sont suffisantes pour la transposition (Mendiola et al, 1994). Cependant, ils n'ont pas étudié l'effet de mutations ponctuelles de cette extrémité comme chez *IS1294b*. Ils ont toutefois démontré que la mutation de certaines positions de l'extrémité *oriIS* induisait soit une abolition, soit une diminution drastique de la transposition (Yassine, 2015).

Ainsi, une étude des différentes mutations répertoriées dans l'extrémité *oriIS91* mériterait d'être menée.

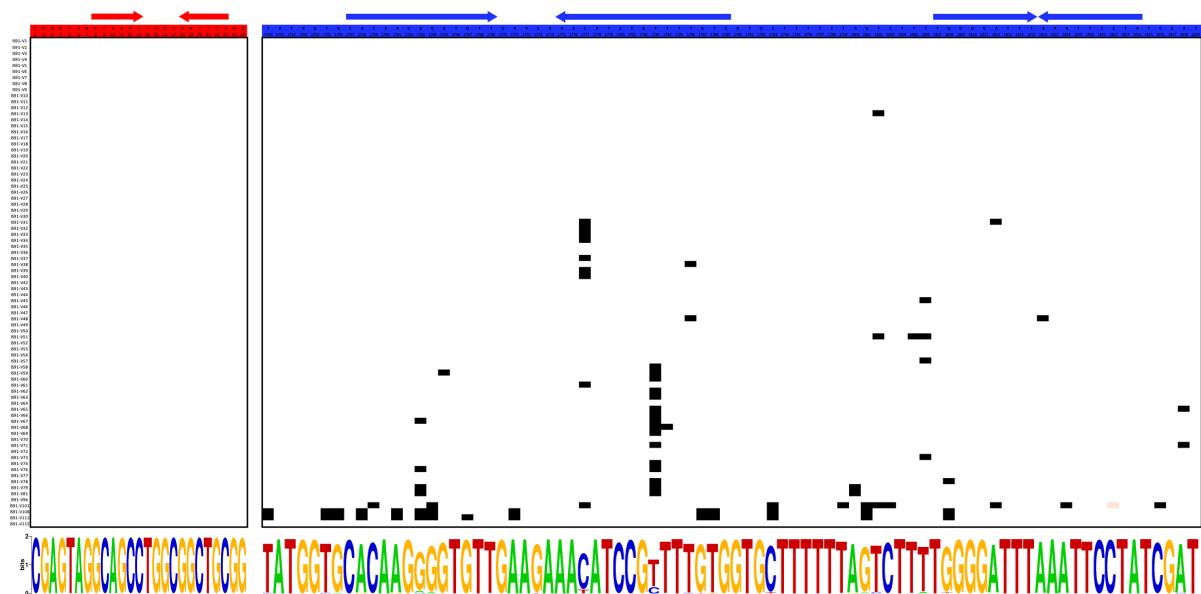


Figure 50 : Distribution des polymorphismes des extrémités *terIS91* et *oriIS91*

Représentation de la distribution des polymorphismes en fonction de l'isoforme décrite et de la position de la mutation dans la séquence *terIS* (à gauche, haut rouge) et la séquence *oriIS* (à droite, haut bleu) par rapport à la première séquence *IS91* décrite (GenBank : X17114.5.1). Chaque palindrome est indiqué par les flèches horizontales rouge pour l'extrémité *terIS* et bleu pour l'extrémité *oriIS*. Les mutations sont représentées en noir. En saumon est indiqué l'absence de ces nucléotides.

**Troisièmement**, nous avons réalisé une analyse du polymorphisme des transposases des isoformes IS1294b et IS91. Nous avons décrit les 225 isoformes à l'aide d'une carte thermique binaire de la distribution des polymorphismes en fonction de l'isoforme décrite et de la position de la mutation dans la transposase TnpA<sub>IS1294b</sub>, par rapport à la première séquence TnpA<sub>IS1294b</sub> décrite (GenBank : AIY68271.1) en figure 51.

L'analyse de la répartition des polymorphismes de la **transposase IS1294b** a révélé de très nombreuses mutations entre chaque domaine conservé. L'étude des 225 isoformes démontrent que les mutations affectent très peu les domaines importants pour l'activité de la transposase, à l'exception de deux acides aminés du domaine I. De plus, la région entre le domaine I et II est très conservée et présente peu de mutations. Concernant les domaines conservés, nous pouvons remarquer que :

- Sept mutations sont retrouvées, une ou deux fois, dans le domaine I aux positions G52C, E55K, C58R, C63Y, P64L et C73Y. Enfin, dans ce même domaine, la mutation L51P est retrouvée quatre fois. L'effet de certaines de ces mutations a été décrit dans les travaux de thèse de Haytham Yassine pour les C58, C63 et C73 (Yassine, 2015). Etant précisé que concernant les mutations L51, G52, E55 et P64, il n'existe aucune donnée.
- Quatre mutations sont retrouvées, une ou deux fois, dans le domaine II aux positions D101N, W102R, V106A et T108I, ainsi que la mutation L113Q qui est retrouvée cinq fois.
- Dans le domaine III, des mutations sont retrouvées une ou deux fois aux positions H164Y, H166Y, V167A, H168P, S170P. Enfin, la mutation W163C est retrouvée douze fois. D'après les données de la littérature, les mutations H166 et H168 seraient critiques pour l'activité de la transposase lors d'une mutation. *A contrario*, la mutation H164 n'altère pas son activité (Yassine, 2015).
- Deux mutations sont retrouvées une fois dans le domaine IV aux positions A254X et K262T à proximité des deux tyrosines catalytiques. Enfin, deux mutations sont retrouvées cinq et quatre fois aux positions L257P/Q et G258S, respectivement. Aucune donnée n'est disponible dans la littérature concernant ces mutations.
- Dans le domaine V, des mutations sont retrouvées une fois aux positions F314Y et M317L. Il n'existe aucune donnée expérimentale pour ce domaine. Le faible nombre de mutations dans ce domaine suggère l'existence d'un rôle potentiel dans la mobilité de l'élément.
- Concernant le domaine VI, deux mutations et une mutation sont observées pour V377E et C379S, respectivement. Dans la littérature, il a été démontré que la mutation de C379 était

critique pour l'activité de la transposase (Yassine, 2015), toutefois il n'existe aucune donnée pour V377.

Force est de constater que les mutations décrites dans les domaines I à VI sont peu fréquentes. Ainsi, nous pouvons suggérer que ces acides aminés joueraient un rôle potentiel dans l'activité de la transposase pour la catalyse de la transposition.

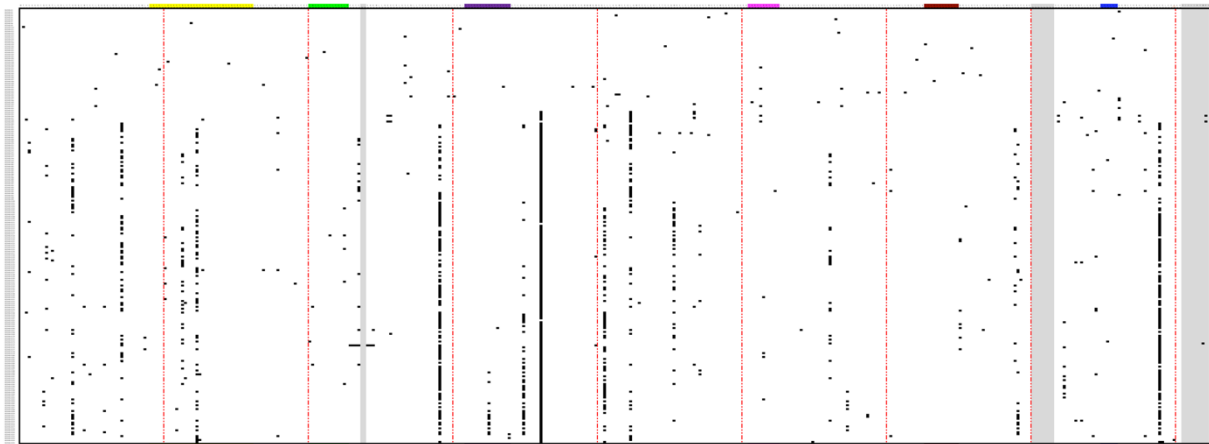


Figure 51 : Distribution des polymorphismes de la transposase  $tnpA_{IS1294b}$

*Représentation de la distribution des polymorphismes en fonction de l'isoforme décrite et de la position de la mutation dans la transposase par rapport à la première séquence  $tnpA_{IS1294b}$  décrite (GenBank: AIY68271.1). Chaque domaine important de la transposase est représenté de couleur différente sur le bord en haut et en bas de la matrice thermique : domaine I (jaune), domaine II (vert), domaine III (violet), domaine IV (rose), domaine V (marron), domaine VI (bleu). Dans la matrice thermique, les mutations sont représentées en noir. Les colonnes en fond gris indiquent les acides aminés présents uniquement chez certaines isoformes décrites (noir) et absents de la plupart des isoformes dont la première séquence. Tous les 50 acides aminés, une ligne rouge en pointillé a été faite.*

Nous avons décrit les 81 isoformes à l'aide d'une carte de la distribution des polymorphismes, description faite en fonction de l'isoforme décrite et de la position de la mutation dans la transposase  $TnpA_{IS91}$ , par rapport à la première séquence  $TnpA_{IS91}$  décrite (GenBank : CAD30457.1) en figure 52.

L'analyse de la **transposase IS91** de la répartition des polymorphismes a révélé qu'entre chaque domaine conservé très peu de mutations ont été observées entre les domaines I et II, ainsi qu'entre les domaines V et VI. Cela pourrait signifier que ces régions inter-domaines joueraient un rôle dans la structure de la protéine ou l'activité de la transposase. Il serait alors

intéressant de cibler certaines mutations afin d'évaluer leurs effets comme V30E, A82T ou R346K, etc... Entre les domaines II et III, III et IV, et IV et V, nous avons observé plus de mutations que pour les autres inter-domaines avec 6 positions présentant une densité importante de mutations. Concernant les domaines conservés, nous pouvons remarquer que :

- Trois mutations sont retrouvées : une, trois et quatre fois dans le domaine I aux positions L45R, V64I et H75N/P, respectivement. Il n'existe aucune donnée dans la littérature de l'effet de ces mutations sur l'activité de la transposase.
- De façon surprenante, le domaine II ne présente aucune mutation signifiant une importance de ce motif dans l'activité de la transposase. En outre, nous pouvons supposer que la mutation d'un seul de ces acides aminés pourrait altérer l'activité de la transposase. Il serait intéressant d'investiguer le motif « VFTLP » d'une part et l'acide aminé T d'autre part, par homologie avec les réplitrons et les héliçons qui conservent également ces deux acides aminés.
- Dans le domaine III, une seule mutation est retrouvée à la position I147N. Il n'existe actuellement aucune donnée sur cette mutation.
- A proximité du site catalytique (domaine IV), une seule mutation est retrouvée à la position A247V et dont l'effet est inconnu.
- Dans le domaine V, des mutations sont retrouvées une fois aux positions L317P, S318E, P319S, V320G et K321E. La mutation R309L est présente quatre fois.
- Concernant le domaine VI, une mutation est observée de C363Y. Dans la littérature, par homologie avec la transposase TnpA<sub>IS1294b</sub>, il a été démontré que sa mutation était critique pour l'activité de la transposase (Yassine, 2015).



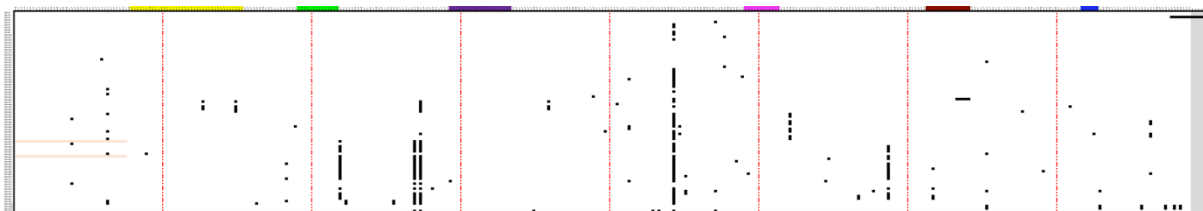


Figure 52 : Distribution des polymorphismes de la transposase tnpA<sub>IS91</sub>

Représentation de la distribution des polymorphismes en fonction de l'isoforme décrite et de la position de la mutation dans la transposase, par rapport à la séquence de la transposase tnpA<sub>IS91</sub> (GenBank : CAD30457.1). Chaque domaine important de la transposase est représenté par une couleur différente sur le bord en haut et en bas de la matrice thermique, domaine I (jaune), domaine II (vert), domaine III (violet), domaine IV (rose), domaine V (marron), l'absence de mutation en blanc. Les colonnes en fond gris indiquent les acides aminés sont présents uniquement chez certaines isoformes décrites (noir) et absents de la plupart des isoformes (dont la première séquence). Les lignes en fond saumon indiquent l'absence de ces acides aminés car la séquence est tronquée. Tous les 50 acides aminés, une ligne rouge en pointillé a été faite.

**Quatrièmement**, nous avons réalisé une analyse du polymorphisme de Orf121 des isoformes IS91. Les travaux précédents et les nôtres ont montré le rôle de *orf121* dans la régulation de la transposition de IS91, en lien avec le chevauchement des deux gènes *orf121* et *tnpA* (figure 53).

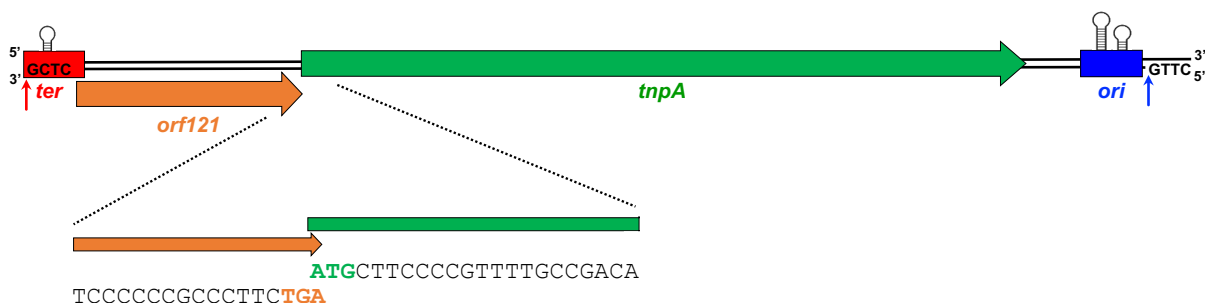


Figure 53 : Organisation génétique de l'élément IS91

Les extrémités *ter* et *ori* portant les séquences palindromiques sont représentées par des cases rouges et bleues, respectivement. Les positions de clivage aux extrémités *ter* et *ori* sont indiquées par de petites flèches verticales rouges et bleues, respectivement. La séquence d'ADN du site de clivage de l'extrémité *ter* spécifique à l'élément et le site cible tétranucléotidique adjacent à *ori* sont également indiqués. Les séquences codantes *Orf121* et *tnpA* sont représentées par des flèches orange et vertes, respectivement. La fin de la séquence codante de *orf121* et *tnpA* sont indiqués et les codons stop et start de *orf121* et *tnpA* sont en orange et vert, respectivement.

Il n'existe aucune donnée ou information concernant le produit Orf121. La séquence peptidique est présentée en figure 54a. Un seul motif semble évident : le CxxC formé de deux

cystéines, ce qui est typique d'un motif doigt de zinc. Il est présent en une seule copie, cela pourrait suggérer une interaction avec une autre protéine ou la formation d'un dimère de Orf121. Un second motif moins évident serait la présence d'une leucine zipper potentiel aux positions **V.X<sub>5</sub>.L.X<sub>5</sub>.V** ou **V.X<sub>6</sub>.L.X<sub>6</sub>.V**.

De façon générale, les acides aminés de Orf121 semblent extrêmement conservés à l'exception de 6 acides aminés en position 41, 56, 61, 90, 118 et 121 avec des mutations identiques en position 56, 61, 118 et 121 (figure 54b). Il se peut que l'ancêtre, qui a acquis ces mutations, ait ensuite muté pour les positions 41 et 90. La mutation 61 se localise précisément à la jonction entre les deux motifs leucine zipper. Enfin, le motif doigt de zinc potentiel est extrêmement conservé et ne présente qu'une seule mutation de P106Q. Ainsi, il apparaît nécessaire d'étudier en profondeur ces mutations afin d'évaluer leur impact sur l'activité de la protéine.

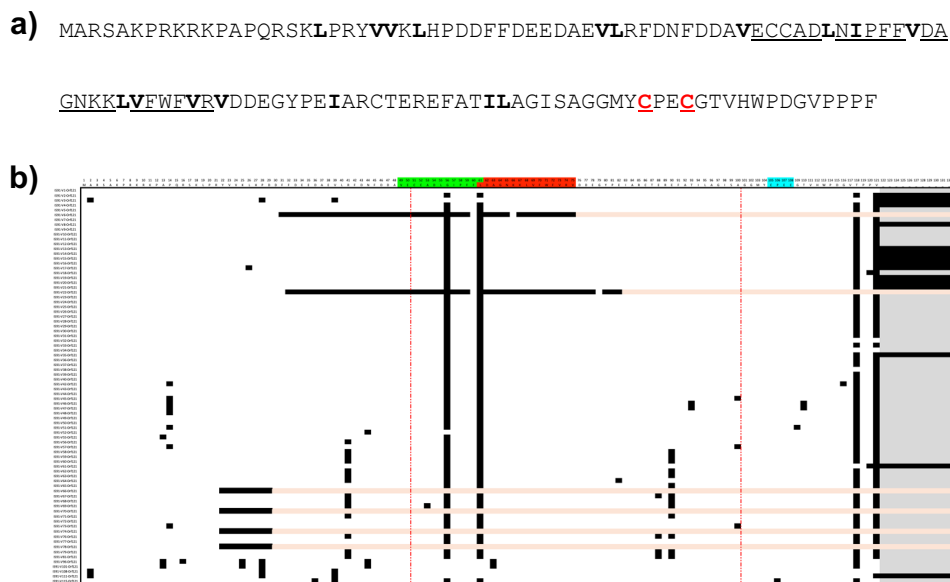


Figure 54 : Séquence peptidique d'Orf121 et distribution des polymorphismes d'Orf121

a) Séquence peptidique de la première séquence Orf121 décrite (GenBank : CAC14584.1).

b) Représentation de la distribution des polymorphismes en fonction de l'isoforme décrite et de la position de la mutation dans la séquence de Orf121, par rapport à la première séquence de Orf121 qui a été décrite (GenBank : CAC14584.1). Les deux domaines leucines zipper supposés sont représentés en vert et en rouge. Un domaine en cyan est indiqué pour la localisation du motif CxxC. Dans la matrice thermique, les mutations sont représentées en noir. Les colonnes en fond gris indiquent les acides aminés absents de la plupart des isoformes dont la première séquence. Les lignes en fond saumon indiquent l'absence de ces acides aminés car la séquence est tronquée. Tous les 50 acides aminés, une ligne rouge en pointillé a été faite.

Il ressort de ce qui précède que :

- L'analyse des isoformes IS91 démontre que :
  - L'extrémité *terIS* est conservée chez toutes les isoformes,
  - L'extrémité *oriIS* présente un nombre conséquent de mutation dans les deux palindromes subterminaux,
  - La transposase TnpA<sub>IS91</sub> présente peu de mutations,
  - L'Orf121 présente un motif doigt de zinc et leucine zipper potentiel.
  
- L'analyse des isoformes IS1294b démontre que :
  - L'extrémité *terIS* est conservée chez toutes les isoformes,
  - L'extrémité *oriIS* présente une conservation des 20 premiers nucléotides qui sont les sites supposés de liaison à la transposase,
  - Le palindrome proposé n'est pas conservé.
  - La transposase TnpA<sub>IS1294b</sub> présente peu de mutation dans les domaines importants.

### I.3. Discussion

Notre analyse *in silico* a démontré que les extrémités *oriS* et *terS* sont extrêmement conservées. L'extrémité *oriS* présente un polymorphisme plus important que l'extrémité *terS*. En outre, la mutation de certaines de ces bases pourrait impacter la mobilisation de l'élément s'il s'agit de sites de liaison potentiels de la transposase. Chez *IS1294b*, il a été démontré l'importance de la base T (-3) du tétranucléotide adjacent au site de coupure de l'*oriS* (Yassine, 2015). De plus, ces auteurs ont montré que 7 bases (A1, G3, T6, A7, G8, A10 et T14) à l'extrémité *oriS* sont nécessaires pour l'activité de transposition *in vivo* car leurs mutations étaient très délétères. Cela correspond à notre analyse de la base de données où les vingt premiers nucléotides sont tous conservés chez les isoformes décrites. De plus, ces bases peuvent être des sites de liaison spécifiques au niveau de la région riche en cystéines identifiée à l'extrémité N-terminale de la protéine TnpA<sub>IS1294b</sub>.

Enfin l'extrémité *terS* des deux éléments ne présente presque aucun polymorphisme.

Ainsi, au niveau des régions *terS* et *oriS*, ces deux séquences restent à explorer chez *IS91* par des expériences de mutagenèse ciblée. Concernant l'extrémité *oriS*, il a été démontré précédemment que les 82 pb de l'extrémité étaient suffisantes pour fonctionner dans la transposition (Mendiola et al., 1994). Nous avons observé un certain nombre de mutations qui sont dans les palindromes terminaux de l'extrémité *oriS*. Il serait intéressant d'investiguer ces mutations, car il pourrait s'agir de nucléotides de fixation de la transposase. Ces palindromes présentent des bases extra-hélicases dont l'étude serait nécessaire. Il a été montré l'importance de telles bases extra-hélicases dans les sites *attC* de cassettes d'intégrons quant à leur rôle dans la sélection des brins d'ADN. On peut donc se poser la question du rôle potentiel de ces bases extra-hélicases dans la sélection du brin bottom lors de la transposition de *IS91*.

La transposase *IS1294b* démontre un polymorphisme plus important que la transposase *IS91*. Pour la plupart des polymorphismes observés, les mutations sont localisées dans les régions inter-domaines. De plus, notre analyse a montré que les domaines II et V sont extrêmement conservés. L'étude de leur rôle dans la transposition serait pertinent.

Enfin, des études ont démontré que la méthylation à proximité de promoteurs de gènes varie considérablement, une plus grande méthylation des promoteurs étant corrélée à une transcription faible ou inexistante (Suzuki and Bird, 2008). Cela est pour *IS91* car un site de

méthylation est localisé dans la boîte -10 du promoteur *tnpA* (<sup>5'</sup>**TGTACT**). En outre, selon l'état de méthylation de ce site, la transcription du gène *tnpA* pourrait être activée ou non comme cela a été démontré chez de nombreux gènes (Reisenauer and Shapiro, 2002; Suzuki and Bird, 2008; Zweiger et al., 1994). Les sites de méthylation chez *E. coli* sont largement répandus le long du génome et se rencontrent environ tous les 200-300 pb. Ces sites peuvent affecter la liaison et l'activité des facteurs de transcription et de l'ARN polymérase, en particulier lorsque le nombre de sites de méthylation est supérieur à la distribution moyenne (Mohapatra and Biondi, 2017). Ceci peut être le cas du gène *tnpA* car sa séquence codante présente cinq sites de méthylation, dont quatre sont regroupés dans une région de 250 pb.

De plus, il n'est pas exclu que l'expression des gènes pourrait être régulé par d'autres facteurs comme les protéines CRP, IHF et OmpR. Il est supposé pour l'élément *ISCR1* que ces protéines pourraient intervenir au sein de la région promotrice du gène *rcr1* (Lallement, 2018).

La présence de *orf121* n'explique pas à elle seule la régulation négative de l'élément *IS91*, ce que démontre l'expression des gènes *orf121* et *tnpA* en tête-bêche. En effet, lorsque les deux gènes sont sous le contrôle de deux promoteurs différents, la diminution de la transposition est moindre par rapport à la configuration native entre les deux gènes. Ceci suggère la présence éventuelle d'un autre facteur impactant la mobilité de l'élément. Ainsi, il a été décrit sur le brin complémentaire deux *orfs* potentielles, *orf108* et *orf106* qui pourraient jouer un rôle dans la régulation de la transposition en tant qu'ARN-antisens (Garcillán-Barcia et al., 2002), tel que démontré pour les éléments *IS10* et *IS30* (Arini et al, 1997 ; Case et al, 1989).

Bien que nous ayons montré l'effet négatif de *orf121*, nous ne connaissons pas précisément la manière dont *orf121* agit négativement sur la transposition. Ainsi nous avons deux hypothèses :

- La première est une interaction directe protéine-protéine. En effet, nous avons supposé que la protéine Orf121 interagissait avec la transposase par le motif CxxC en C-terminal des deux protéines.
- La seconde hypothèse est que le motif CxxC permet la formation d'un dimère de Orf121. Ainsi, avec le leucine zipper potentiel de Orf121, une forme dimérique pourrait interagir avec les ADN *oriIS91* et *terIS91*. Cette hypothèse expliquerait l'effet de Orf121 sur la capacité réduite de transposition quand elle se fixe sur *oriIS91* agissant comme un

compétiteur de la transposase et quand elle se fixe sur *terIS91*, elle améliore sa reconnaissance.

Afin de tester ces hypothèses, deux mutants ponctuels ont été construits lorsque *orf121* et *tnpA* ont une configuration naturelle sous le contrôle du  $P_{lac}$  :

- Un mutant de la cystéine 105 en sérine de la protéine Orf121 ( $P_{lac}::orf121_{C105S}tnpA$ ),
- Et un mutant de la cystéine 360 en sérine de la transposase ( $P_{lac}::orf121tnpA_{C360S}$ ).

Au préalable, nous avons vérifié que la conjugaison était stable (entre  $5 \times 10^{-2}$  et  $9 \times 10^{-2}$ ) entre les réplicats biologiques pour ne pas impacter et induire une variabilité dans les résultats de transposition induite par la fluctuation de la conjugaison. La figure 55 présente les fréquences de transposition associées aux différentes mutations générées dans la construction  $P_{lac}::orf121tnpA$ .

Nous observons que la fréquence de transposition dans la construction sauvage avait une activité basale de  $1,47 \times 10^{-6}$ . Cette fréquence diminuait à  $5,12 \times 10^{-7}$  en présence d'IPTG exprimant *orf121*, soit une diminution de 2,88 fois. En présence des mutations des cystéines en position 105 ou 360, nous avons observé une augmentation significative de 34,7 fois de la fréquence de transposition par rapport à la construction sauvage quand la cystéine 105 de l'élément Orf121 est mutée en sérine et l'activité de transposition chutait de 11,98 fois quand la cystéine 360 de la transposase était mutée en sérine.

Ce résultat préliminaire confirme que le motif CxxC en C terminale de la protéine Orf121 est important pour son activité, soit dans la formation du dimère, soit dans l'interaction avec la transposase.

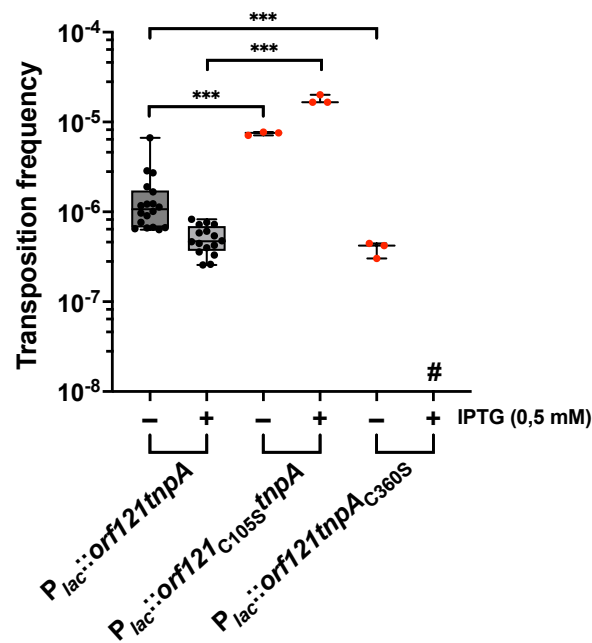


Figure 55 : Effet *in vivo* des mutations ponctuelles des transposases du motif CxxC de Orf121 et TnpA en C-terminale

Les barres d'erreur indiquent les écarts types ;  $n \geq 5$  (point noir) et  $n = 3$  (point rouge). #, transposition indétectable. Le - et + indiquent la présence d'inducteur ou non à 0,5 mM.  $P_{lac}::orf121tnpA$  indique la construction témoin sans mutation qui est exprimée en trans ;  $P_{lac}::orf121_{C105S}tnpA$  et  $P_{lac}::orf121tnpA_{C360S}$  indique les constructions portant une mutation ponctuelle qui est exprimée en trans. Le dérivé IS91 porté par le plasmide donneur de transposon est  $_{GTC-5'}terIS91-cm^R-orfIS91_{GTC-5'}$ .

Afin de vérifier l'hypothèse de la formation d'un dimère, nous avons réalisé une détection par Western-immunoblot en condition non dénaturante de Orf121 (figure 56). Précisons que dans une condition dénaturante, la protéine Orf121 est détectée à 17 kDa. Or, dans la condition non dénaturante, deux espèces sont détectées : une bande à 17 kDa et une légèrement en-dessous 26 kDa. Cela suggère fortement la formation d'un dimère de la protéine Orf121.

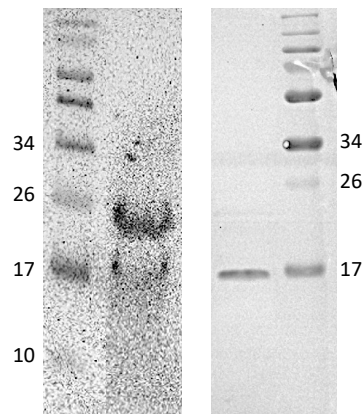


Figure 56 : Orf121 un dimère ?

*Analyse par western-immunoblot de Orf121 à partir d'un gel de polyacrylamide sans agent dénaturant (SDS) à gauche et en présence de SDS dans des conditions normales d'analyse par western-immunoblot à droite.*

Enfin, la séquence peptidique d'Orf121 a permis de mettre en évidence un motif de liaison au zinc potentiel en C-terminal et deux leucines zipper potentiels. L'analyse du polymorphisme de Orf121 a révélé que plusieurs acides aminés mériteraient d'être étudiés afin de mieux comprendre l'activité et le rôle de la protéine Orf121.



## Partie II. Étude structurale de la transposase IS91

---

Au cours de l'introduction dans le *chapitre 4 sous partie IV.2.3*, nous avons vu que les domaines I et VI de la transposase des membres de la famille IS91 et ISCR étaient conservés mais leur fonction est inconnue. Le premier domaine en N-terminal (domaine I) de l'élément IS91 comprenant une région conservée riche en cystéine, il pourrait constituer un motif putatif de liaison au zinc (figure 37). De plus, il existe une région conservée en C-terminal (domaine VI) supplémentaire uniquement chez les éléments de la famille IS91.

Parmi les domaines de liaison à l'ADN, le motif du doigt de zinc (ZF) est le plus répandu. Il est principalement constitué de résidus d'acides aminés, en particulier de cystéines, qui coordonnent des ions zinc (Berg et al., 2015). Ces cystéines forment des liaisons covalentes avec le zinc, stabilisant ainsi la structure.

En général, un seul domaine à ZF n'est pas suffisant pour assurer une liaison de haute affinité avec une séquence cible d'ADN spécifique (Berg and Godwin, 1997; Klug and Schwabe, 1995). Des études structurales des protéines complexes à doigt de zinc avec l'ADN ont démontré que la reconnaissance moléculaire spécifique à une séquence est réalisée par l'interaction des bases de l'ADN avec les chaînes latérales de la surface de la partie N-terminale de l'hélice  $\alpha$  (Wolfe et al., 2000). La reconnaissance moléculaire de l'ADN par le ZF procaryote semble être particulière et très différente de celle observée chez les eucaryotes (Russo et al., 2015). De plus, il est admis que les ZF possèdent un rôle dans la stabilisation des protéines et facilitent même le repliement des sous-unités protéiques (Kluska et al., 2018). Enfin, il existe un nombre croissant de preuves qui suggèrent que la perte d'un ZF entraîne immédiatement une perturbation de la structure et de la fonction unique de l'ensemble du domaine (Padjasek et al., 2020).

Les transposases des éléments IS91, IS801, IS1294 et IS1294b ont un pourcentage en cystéines de 4,1%, 3,4%, 4,6% et 3,7%, respectivement, ce qui est supérieur à la valeur moyenne (1,03%-1,19%) de cystéine d'une protéine (Miseta and Csutora, 2000). Ces différences pourraient suggérer qu'il existerait probablement un rôle de ces cystéines, notamment dans les domaines I et VI, dans l'activité de ces transposases en formant un ou plusieurs ZF.

Nous avons donc voulu appréhender le rôle des domaines I et VI chez la transposase TnpA<sub>IS91</sub> par une approche *in vivo* et *in silico*. Ainsi, nous avons décidé d'étudier l'impact d'une série de mutations ponctuelles de cystéine en sérine en position 41, 52, 53, 58, 59, 65, 68, 73 et 76 du

domaine I, puis en position 360 et 363 du domaine VI, ainsi que l'histidine 60 en alanine du domaine I.

## **II.1. Matériels et méthodes**

### **II.1.1. Analyse *in vivo***

Nous avons utilisé un système génétique basé sur le test de « *mating-out* » chez *Escherichia coli* pour mesurer la mobilité des dérivés des IS91 dans le plasmide conjugatif pOX38Km. Tous les résultats obtenus en « *mating-out* » ont été analysés à l'aide d'un test de Mann-Whitney à deux groupes non-appariés.

La stratégie expérimentale de mating-out est la même que celle décrite en *Partie I sous partie 1.1* dans l'article « *IS91-encoded orf121 regulates negatively IS91 transposition* ». La liste des souches et des plasmides utilisés est présentée dans le tableau 2. La liste des amorces utilisés pour la construction des plasmides est disponible en annexe 6.

Tableau 2 : Souches et plasmides utilisés lors de l'étude des domaines I et VI

$\Delta$  indique une suppression ;  $\Omega$  indique une insertion ; :: indique une nouvelle articulation ; \* indique une mutation

Nom	Surnom	Caractéristiques	Source / référence
<b>Souches utilis</b>			
JS219		MC1061, <i>recA1</i> , <i>lacI<sup>q</sup></i>	<i>Cam et al., 1988</i>
MC240		XA103, F <sup>-</sup> , <i>ara</i> , <i>del(lac pro)</i> , <i>gyrA</i> ( <i>nalR</i> ), <i>metB</i> , <i>argE<sup>am</sup></i> , <i>rpoB</i> , <i>thi</i> , <i>supF</i>	Gift of B. Ton-Hoang
<b>Plasmide de référence et mère</b>			
pSU2600		Source of IS91 of TnpA <sub>IS91</sub>	<i>Mendolia et al, 1994</i>
pUC18		Amp <sup>r</sup> ; MCS from M13mp19 into <i>lacZ'</i>	<i>Yanisch-Perron et al., 1985</i>
pGY13556		pAPT110 $\Delta$ Kan <sup>r</sup> ; P <sub>lac</sub> :: <i>tnpA</i> <sub>ISDra2-</sub> <i>his6</i> ; <i>spc<sup>R</sup>/sm<sup>R</sup></i>	<i>Pasternak et al, 2013</i>
pOX38Km		Km <sup>R</sup>	<i>Chandler and Galas, 1983b</i>
<b>Plasmide mimant une IS</b>			
pCP7	IS91- <i>ter91-cm<sup>R</sup></i>	pUC18; <i>ter91</i> P <sub>cat</sub> ::Cam <sup>R</sup> <i>ori91</i>	This work
<b>Dérivé de TnpA<sub>IS91</sub> avec les mutations ponctuelles dans les domaines I et VI</b>			
pCP2	P <sub>Lac</sub>	pGY13556 $\Delta$ tnpA <sub>ISDra2</sub> $\Omega$ linker79	This work
pCP4	P <sub>lac</sub> :: <i>tnpA</i>	pCP2; P <sub>lac</sub> :: <i>tnpA</i>	This work
pAF8	pTnpA <sub>C68S</sub>	pCP2; P <sub>lac</sub> :: <i>tnpA</i> <sup>*C68S</sup>	This work
pAF9	pTnpA <sub>C73S</sub>	pCP2; P <sub>lac</sub> :: <i>tnpA</i> <sup>*C73S</sup>	This work
pAF11	pTnpA <sub>C53S</sub>	pCP2; P <sub>lac</sub> :: <i>tnpA</i> <sup>*C53S</sup>	This work
pAF12	pTnpA <sub>C58S</sub>	pCP2; P <sub>lac</sub> :: <i>tnpA</i> <sup>*C58S</sup>	This work
pAF13	pTnpA <sub>C41S</sub>	pCP2; P <sub>lac</sub> :: <i>tnpA</i> <sup>*C41S</sup>	This work
pAF14	pTnpA <sub>C59S</sub>	pCP2; P <sub>lac</sub> :: <i>tnpA</i> <sup>*C59S</sup>	This work
pAF15	pTnpA <sub>C58/59S</sub>	pCP2; P <sub>lac</sub> :: <i>tnpA</i> <sup>*C58/59S</sup>	This work
pAF16	pTnpA <sub>C65S</sub>	pCP2; P <sub>lac</sub> :: <i>tnpA</i> <sup>*C65S</sup>	This work
pAF17	pTnpA <sub>C76S</sub>	pCP2; P <sub>lac</sub> :: <i>tnpA</i> <sup>*C76S</sup>	This work
pAF18	pTnpA <sub>C52S</sub>	pCP2; P <sub>lac</sub> :: <i>tnpA</i> <sup>*C52S</sup>	This work
pAF19	pTnpA <sub>C52/53S</sub>	pCP2; P <sub>lac</sub> :: <i>tnpA</i> <sup>*C52/53S</sup>	This work
pAF27	pTnpA <sub>H60A</sub>	pCP2; P <sub>lac</sub> :: <i>tnpA</i> <sup>*H60A</sup>	This work
pAF28	pTnpA <sub>C360S</sub>	pCP2; P <sub>lac</sub> :: <i>tnpA</i> <sup>*C360S</sup>	This work
pAF29	pTnpA <sub>C363S</sub>	pCP2; P <sub>lac</sub> :: <i>tnpA</i> <sup>*C363S</sup>	This work
pAF40	pTnpA <sub>C360/363S</sub>	pCP2; P <sub>lac</sub> :: <i>tnpA</i> <sup>*C360/363S</sup>	This work

### II.1.2. Prédiction de la structure par AlphaFold2

La modélisation de la structure de la transposase TnpA<sub>IS91</sub> a été effectuée à partir du modèle fiable et robuste issu de l'outil de prédiction de structure AlphaFold2 (AF2) (Jumper et al., 2021; Varadi et al., 2022).

AlphaFold propose estime la qualité de sa prédiction, en autres, en mesurant par résidu la déviation positionnelle de ce dernier par rapport aux structures résolues similaires. La mesure de confiance par résidu, appelée pLDDT (per-residue local distance difference test) est

comprise entre 0 et 100. Il convient de noter que la confiance du modèle peut considérablement varier le long de la séquence. Les valeurs de confiance les plus basses peuvent être soit associées à une mauvaise prédiction mais également à un désordre intrinsèque. La mesure de confiance par résidu de pLDDT est représenté par un code couleur :

- Les résidus dont le pLDDT est supérieur à 90% sont prédits avec une haute confiance (en bleu).
- Les résidus dont le pLDDT est compris entre 70% et 90% sont prédits avec une bonne confiance (en cyan).
- Les résidus dont le pLDDT est compris entre 50 et 70% sont prédits avec une confiance moyenne et doivent être traités avec prudence (en jaune).
- Les résidus dont le pLDDT est inférieur sont prédits avec une faible confiance et doivent être considérés avec une attention particulière (en orange).

Outre la structure secondaire, AF2 permet également d'évaluer la structure tertiaire, i.e., l'agencement des différents domaines entre eux. La confiance dans la prédiction peut être mesurée par « Predicted Aligned Error » (PAE). Ce résultat indique l'erreur dans la prédiction de la position en Å d'un résidu  $i$  par rapport à un résidu  $j$ . Une PAE d'inférieur à 3.5 Å propose une haute confiance de la structure tertiaire, dans la mesure où 3.5 Å représente la résolution moyenne des structures de Cryo-ElectroMicroscopie.

### II.1.3. Modélisation de la dynamique de TnpA<sub>IS91</sub>

Aucune donnée sur la structure de la transposase TnpA<sub>IS91</sub> par microscopie électronique cryogénique (cryo-EM) n'étant à ce jour disponible, nous avons utilisé la modélisation par dynamique moléculaire pour élucider et comprendre le comportement des molécules d'intérêt biologique *in situ*.

Ces simulations reposent sur la Seconde Loi de Physique Classique, dite Newtonienne, en considérant les atomes comme des boules avec une masse et les liaisons comme des ressorts. En assumant les caractères physico-chimiques des différents types d'atomes (charge atomique partielle, planarité, dispersion, etc..), il est possible de calculer l'énergie de ces structures (à l'aide d'une équation définissant le champ de force) à partir de laquelle il est possible de dériver les forces exercées sur chaque atome. En d'autres termes, en connaissant les vitesses des particules (définies notamment par la température), les positions des atomes mais également les forces exercées sur les particules, il est possible de calculer la position des atomes à un intervalle de temps faible (i.e., 2 femtosecondes) et de répéter l'opération  $N$

fois (ou N itérations). Aujourd'hui, il est commun de pouvoir échantillonner les conformations de systèmes tels que TnpA<sub>IS91</sub> sur 10 ms (en considérant 5 réplicats indépendant de 2 ms chacun, soit 10<sup>9</sup> interactions) en une vingtaine de jours de calcul.

En analysant le comportement des macromolécules sur ce temps de simulations, il est également possible d'en déduire son comportement physico-chimique structural permettant d'obtenir non plus une image de la molécule à un instant t, mais une « vidéo » dynamique de la molécule en solution. Cette approche est d'autant plus intéressante en comparant des structures sauvages avec des mutants et en décryptant les différences structurales et dynamiques pour mieux comprendre la fonction de molécules d'intérêt biologique.

Nous avons évalué l'écart quadratique moyen (RMSD). Le RMSD est défini comme la mesure de la distance moyenne entre les atomes de deux molécules superposées. Lorsqu'un système dynamique fluctue autour d'une position moyenne bien définie, la RMSD par rapport à la moyenne dans le temps est appelée RMSF ou fluctuation quadratique moyenne.

## II.2. Résultats *in vivo*

Nous nous sommes intéressés au rôle éventuel des cystéines dans l'activité de transposition *in vivo* de IS91. À cette fin, nous avons étudié une série de mutations ponctuelles de cystéine en sérine dans la transposase TnpA<sub>IS91</sub> en position 41, 52, 53, 58, 59, 65, 68, 73 et 76 du domaine I, et en position 360 et 363 du domaine VI, ainsi que l'histidine 60 en alanine du domaine I.

Au préalable, nous avons vérifié que la conjugaison était stable (entre 5 x 10<sup>-2</sup> et 9 x 10<sup>-2</sup>) entre les réplicats biologiques et techniques pour ne pas impacter et induire une variabilité dans les résultats de transposition induite par la fluctuation de la conjugaison (figure 57a).

La figure 57b présente les fréquences de transposition associées aux différentes mutations générées dans la transposase TnpA<sub>IS91</sub>. Nos analyses de mutagénèses nous ont permis d'identifier quatre résidus cystéines critiques pour l'activité de la transposase : C41, C68, C73 et C76. En effet, quand ces cystéines sont mutées, l'activité de transposition est indétectable.

Les cystéines 360 et 363 ont présenté un effet drastique sur la transposition, cela suggère qu'elles jouent aussi un rôle majeur dans l'activité de transposition.

Les mutation simple des cystéines 53 et 58 en sérine, ainsi que les doubles mutations des couples de cystéine 52/53 et 58/59, ont démontré un effet intermédiaire sur l'activité de transposition. Cela nous laisse supposer qu'elles auraient un rôle important dans la

transposition. En effet, la mutation d'un résidu pourrait être compensée par celui qui est adjacent ou proche. Cela reste à confirmer par l'étude *in silico* de l'effet des mutations ponctuelles du motif du doigt de zinc dans la transposase TnpA<sub>IS91</sub>.

En revanche, la mutation des cystéines 52, 59 et 65 ainsi que la mutation de l'histidine 60 en alanine n'ont aucun impact sur l'activité de transposition.

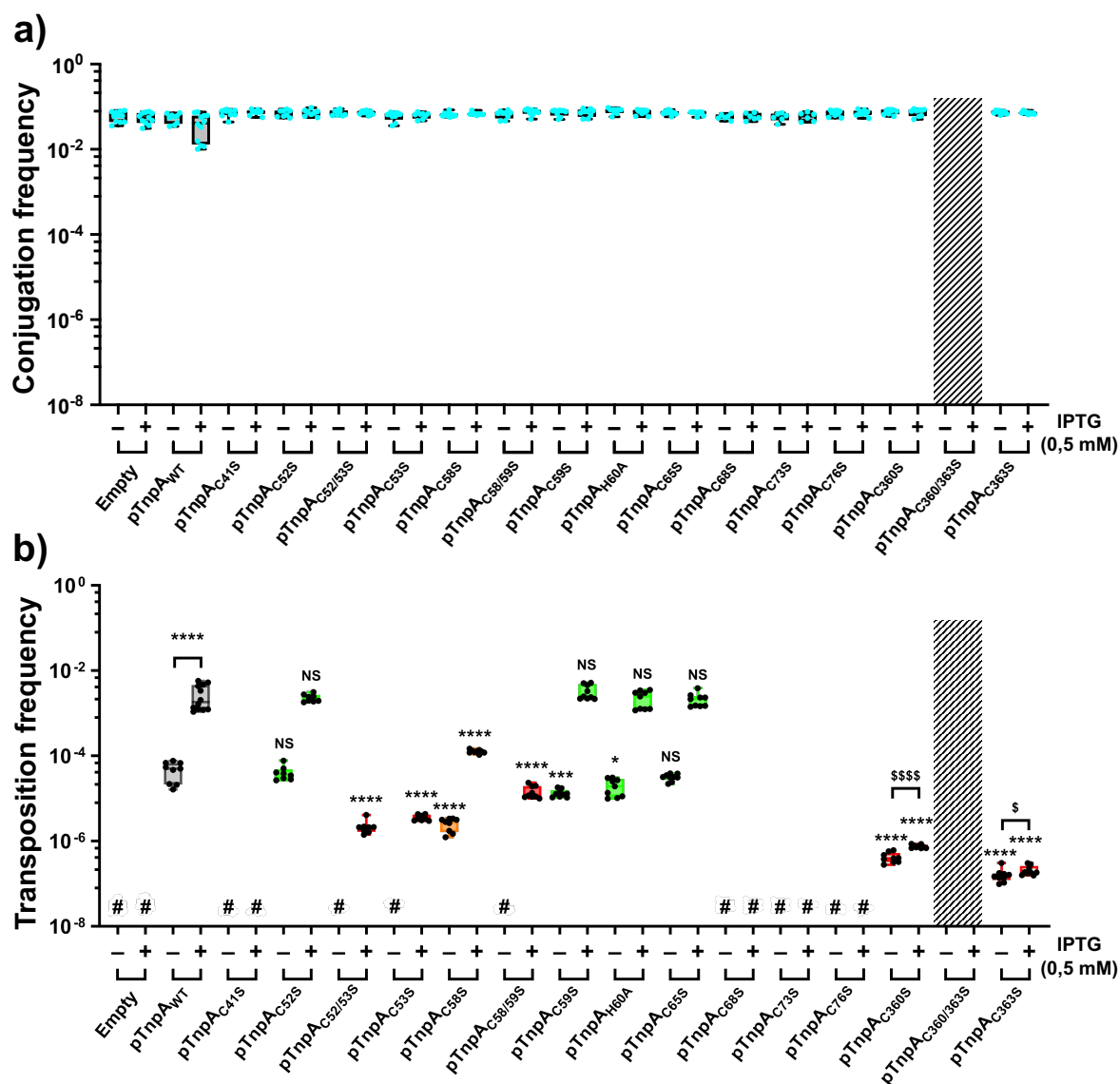


Figure 57 : Effet des mutations ponctuelles des cystéines du domaine I et VI de la transposase

TnpA<sub>IS91</sub>

Les barres d'erreur indiquent les écarts types ;  $n \geq 5$ . #, transposition indétectable. La zone hachurée en noir indique que les données ne sont pas encore disponibles. Le – et + indiquent la présence d'inducteur ou non à 0,5 mM. Les astérisques indiquent les différences significatives (0,1234 (ns) ; 0,0332 (\*) ; 0,0021 (\*\*) ; 0,0002 (\*\*\*) et <0,0001 (\*\*\*\*)). Empty, vecteur d'expression de la transposase vide ; pTnpA<sub>WT</sub>, le gène *tnpA* seul et sans mutation est exprimé en trans ; pTnpA<sub>C41S</sub>, le gène *tnpA* seul portant la mutation de la cystéine 41 en sérine est exprimé en trans. Le dérivé IS91 porté par le plasmide donneur de transposon est  $G_{CTC-5}terIS91-cm^R-orilS91_{GTTTC-5}$ . a) Les fréquences de conjugaisons observées. b) Les fréquences de transposition observées.

Nous avons aussi analysé les sites cibles d'insertion en présence des différents cystéines mutées. Concernant les transposants obtenus en mating-out, nous avons réalisé une analyse moléculaire des sites cibles d'insertion présentée en tableau 3. Quand nous regardons les résultats pour la mutation C59S ou C65S présente un équilibre entre les sites cibles 5'GTTC et 5'CTTG. *A contrario*, les résultats pour la mutation C53S ou C58S présentent un déséquilibre entre les sites cibles 5'GTTC et 5'CTTG avec une préférence marquée pour le site 5'GTTC. Nous avons décidé de réaliser un test du  $\chi^2$  pour vérifier si les résultats obtenus avec les sites cibles 5'GTTC et 5'CTTG, en présence des mutations des cystéines. Une valeur de  $\chi^2$  de 20.53 a été obtenue, elle correspond à une valeur P = 0.004. Cela signifie que la présence de certaines mutations des cystéines a modifié la sélection du site cible et la spécificité d'insertion dans ces constructions. Ces résultats suggèrent que ces cystéines pourraient être impliquées dans la sélection de la cible et de la spécificité d'insertion de IS91. Toutefois, il est plus envisageable que ces cystéines soient impliquées dans la stabilité de l'interaction avec l'ADN.

Tableau 3 : Analyse des sites cibles d'insertion en fonction des mutations ponctuelles des cystéines du domaine I et VI

Target tetranucleotide	pTnpA expression plasmid							
	pTnpA <sub>IS91</sub>	pTnpA <sub>IS91</sub> C52S	pTnpA <sub>IS91</sub> C52/53S	pTnpA <sub>IS91</sub> C53S	pTnpA <sub>IS91</sub> C58S	pTnpA <sub>IS91</sub> C58/59S	pTnpA <sub>IS91</sub> C59S	pTnpA <sub>IS91</sub> C65S
5'GTTC	113	42	38	56	55	39	42	42
5'CTTG	142	52	52	37	27	47	41	44
5'CTCG	3	2	1	1	8	2	3	1
5'GTAC	3	2	1	2	0	3	2	0
5'GTCC	14	1	0	3	2	4	3	7
5'CTTC	1	0	0	0	1	0	1	2
5'GTCG	1	0	0	0	0	0	0	0
5'GTGC	0	0	0	0	3	0	3	0
5'CTGG	0	0	0	0	0	0	1	0
5'GTTG	0	0	0	0	0	0	1	0
5'GTTA	0	0	0	0	0	0	0	1
TOTAL	277	99	92	99	96	95	97	97



Nous avons aussi analysé le pourcentage de OET pour les différentes transposases mutées (tableau 4). L'analyse statistique a démontré que les pourcentages d'OET par rapport à la transposase WT :

- La mutation de la cystéine 363 avait un pourcentage d'OET similaire à la transposase TnpA<sub>IS91</sub> WT.
- La mutation simple des cystéines 52, 59, 65, 360 et de l'histidine 60 présentait un pourcentage d'OET significativement plus faible. Des résultats similaires ont été obtenus pour les doubles mutants 52/53 et 58/59.
- Les mutants 53 et 58 présentait une augmentation significative du pourcentage d'OET.

Tableau 4 : Pourcentage de OET observé en fonction des mutations ponctuelles des cystéines des domaines I et VI

L'indice <sup>a</sup> et <sup>b</sup> indique que pourcentage d'OET significativement plus faible ou plus élevé par rapport à la transposase sauvage, respectivement.

<i>TnpA expression plasmid</i>	<i>% one-ended transposition</i>
<i>pTnpA<sub>IS91</sub></i>	53,04% n = 1235
<i>pTnpA<sub>IS91</sub> C52S</i>	41,35% <sup>a</sup> n = 1260
<i>pTnpA<sub>IS91</sub> C52/53S</i>	30,00% <sup>a</sup> n = 660
<i>pTnpA<sub>IS91</sub> C53S</i>	68,22% <sup>b</sup> n = 900
<i>pTnpA<sub>IS91</sub> C58S</i>	69,26% <sup>b</sup> n = 950
<i>pTnpA<sub>IS91</sub> C58/59S</i>	44,44% <sup>a</sup> n = 900
<i>pTnpA<sub>IS91</sub> C59S</i>	43,33% <sup>a</sup> n = 900
<i>pTnpA<sub>IS91</sub> H60A</i>	46,47% <sup>a</sup> n = 1020
<i>pTnpA<sub>IS91</sub> C65S</i>	43,54% <sup>a</sup> n = 1300
<i>pTnpA<sub>IS91</sub> C360S</i>	41,35% <sup>a</sup> n = 208
<i>pTnpA<sub>IS91</sub> C363S</i>	46,67% n = 60

Ces résultats suggèrent un rôle important de ces cystéines dans la reconnaissance et le clivage du site *terIS91*. Les différences relevées peuvent être expliquées par l'effet des mutations qui pourraient soit :

- Impacter la structure,
- Impacter la stabilité de la transposase,
- Impacter la capacité de reconnaissance de l'ADN.

### **II.3. Résultats *in silico***

L'étude *in silico* a été réalisée grâce à une collaboration avec le Dr Florent DI MEO et le Dr Veronica CRESPI de l'unité INSERM 1248 P&T (Pharmacology & Transplantation) de l'Université de Limoges.

La génération de la structure 3D à l'aide de AF2 est présentée en figure 58a. La structure prédite présente une confiance globalement élevée. Il est en revanche important de noter que les domaines de 241 à 271 présentent un pLDDT inférieure à 50. Bien qu'il soit indispensable de traiter avec une attention particulière les données obtenues à partir des simulations, nous avons observé, à partir de ce même modèle prédit, que les cystéines d'intérêt (représentées par des sphères) forment spontanément des doigts de Zinc (Zinc Finger, ZF) tel que suggéré initialement.

Ainsi, nous avons défini deux doigts de zinc :

- Les cystéines C41, C68, C73 et C76 forment le zinc finger 1 (ZF1).
- Les cystéines C53, C58, C360 et C363 forment le zinc finger 2 (ZF2).

Les cystéines qui composent le ZF1 et ZF2 sont présentes dans des régions prédites avec un indice de confiance très élevé (valeur pLDDT, bleu et cyan). Ainsi, nous pouvons considérer cette prédiction de structure 3D pour l'analyse dynamique des doigts de zinc afin de les comparer aux données expérimentales obtenues.

Par ailleurs, la matrice PAE (figure 58b) prédit également l'agencement prédit des domaines avec une résolution de bonne confiance, notamment pour ZF1 et ZF2.

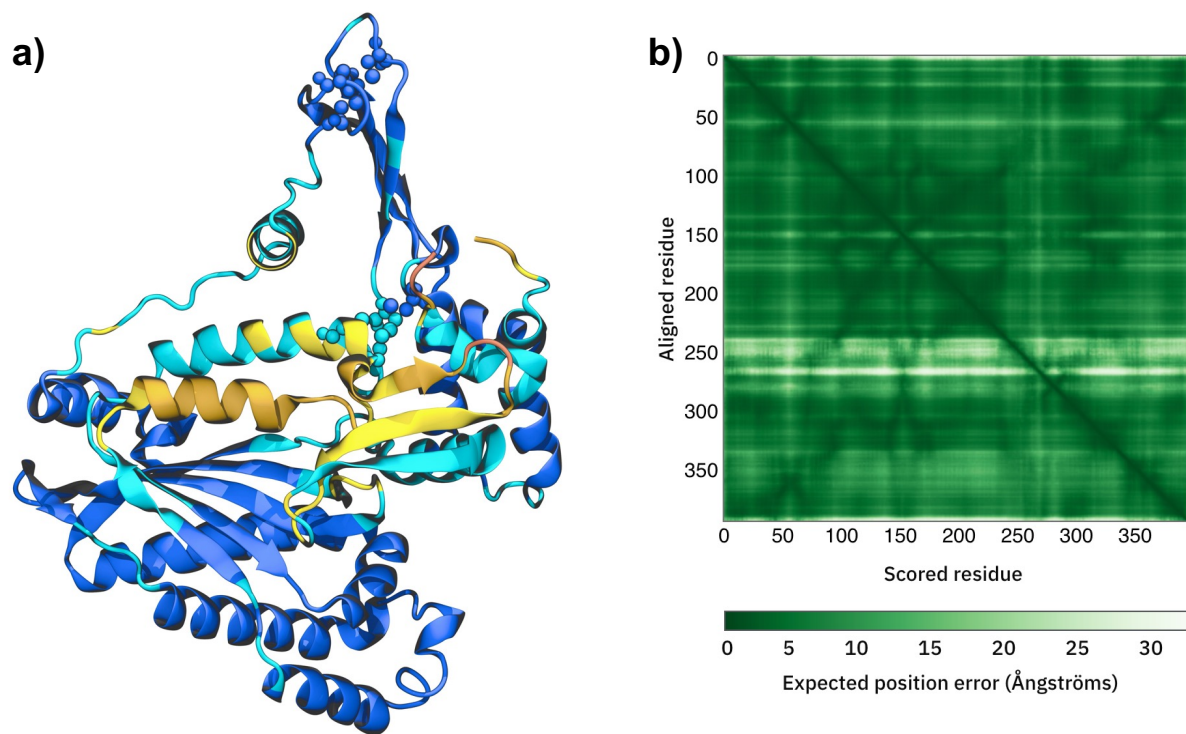


Figure 58 : Prédiction de la structure globale de la transposase TnpA<sub>IS91</sub> et la matrice d'erreur d'alignement prédite

a) Structure de la transposase TnpA<sub>IS91</sub> à l'aide d'AlphaFold2. b) Matrice de la prédiction d'erreur d'alignement de la transposase TnpA où il est indiqué l'erreur dans la prédiction de la position en Å d'un résidu  $i$  par rapport à un résidu  $j$  par un gradient de couleur vert foncé vers blanc.

La figure 59 représente la structure de la transposase de IS91 avec les différents domaines.

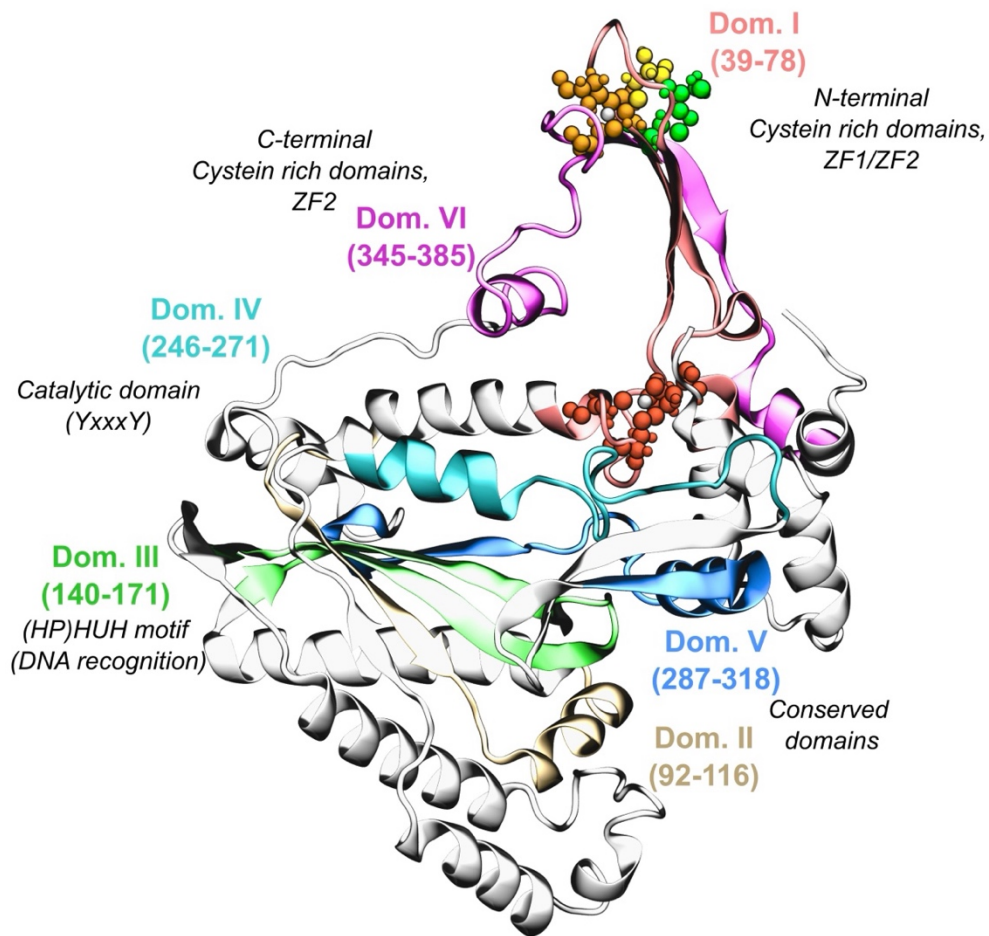


Figure 59 : Position des domaines important de la structure globale de la transposase TnpA<sub>IS91</sub>  
 Le domaine I, II, III, IV, V et VI est représenté en rouge, brun, vert, cyan, bleu et rose, respectivement. La localisation dans la séquence en acide aminé est indiquée entre parenthèse sous le nom du domaine. Les cystéines sont représentées par des sphères rouge, orange, jaune et vertes respectivement.

Le domaine 1 est indiqué en rouge. Il comprend les acides aminés de la position 39-78 et est impliqué dans le ZF1 et ZF2.

Le domaine 2 est en brun et comprend les acides aminés de la position 92-116. La fonction de ce domaine est inconnue.

Le domaine 3 est indiqué en vert et comprend les acides aminés de la position 140-171. Ce domaine comporte le motif HUH impliqué dans la reconnaissance des palindromes terminaux.

Le domaine 4 est en cyan et comprend les acides aminés de la position 246-271. Ce domaine comporte le site catalytique qui catalyse les coupures de l'ADN.

Le domaine 5 est en bleu et comprend les acides aminés de la position 287-318. La fonction de ce domaine est inconnue.

Le domaine 6 est en rose. Il comprend les acides aminés de la position 345-385 et est impliqué uniquement dans le ZF2.

Sur cette figure, nous avons aussi représenté les différentes cystéines qui ont été mutées en représentation sphérique et par couleur. Les 4 cystéines (C41, C68, C73 et C76) dont l'activité *in vivo* est abolie sont représentées en rouge. En jaune et en orange : les cystéines (C53, C52/53, C58, C58/59, C360 et C363) qui ont présenté une activité réduite après mutation en sérine. Enfin, les cystéines (C52 et C59), représentées en vert, sont celles dont la mutation n'a pas impacté l'activité de la transposase *in vivo*.

Dans un premier temps, nous avons évalué l'impact de la présence de zinc ou non sur la dynamique de la structure globale de la transposase sauvage TnpA<sub>IS91</sub> à partir des modèles de la transposase non mutée (figure 60). L'analyse en dynamique a démontré que la structure globale est conservée en présence de zinc avec liaison du zinc aux cystéines pour ZF1 et ZF2. *A contrario*, en l'absence de zinc, l'analyse a démontré que, même si les structures secondaire et tertiaire globale sont conservées, le domaine I est fortement perturbé structuralement. Ces résultats suggèrent fortement que les cystéines du ZF1 seraient impliquées dans la structuration de la protéine par la fixation d'un atome de Zinc. Cette explication est étayée par la proximité du « core » de la protéine avec la présence des domaines III et IV qui sont importants pour la mobilité et l'activité de la transposase.

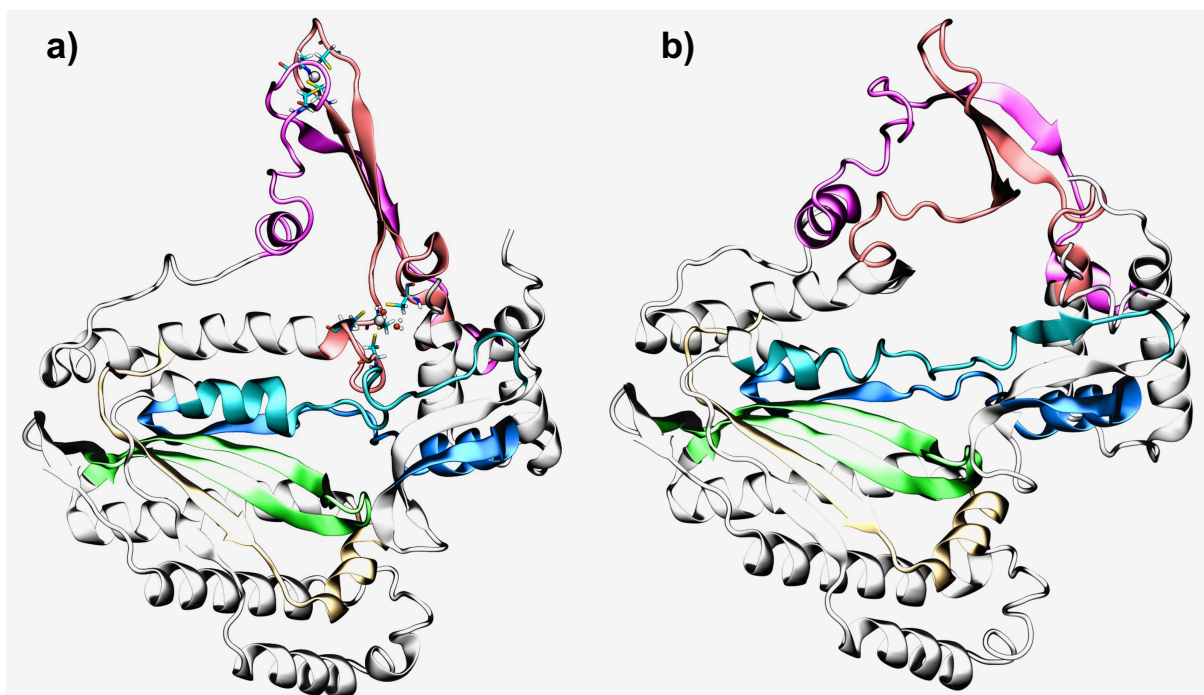


Figure 60 : Impact de la présence de zinc ou non sur la structure globale de la transposase TnpA<sub>IS91</sub>. Les domaines I, II, III, IV, V et VI sont représentée en rouge, brun, vert, cyan, bleu et rose, respectivement. a) En présence de zinc : les cystéines du ZF1 et ZF2 sont représentées par des bâtonnets et l'atome de zinc par une sphère blanche. b) Conséquence de l'absence de zinc sur la transposase TnpA<sub>IS91</sub>.

Nous avons évalué la flexibilité des domaines de la transposase TnpA<sub>IS91</sub> sauvage en présence ou non de zinc, mais également les mutants au niveau des ZF1 et ZF2 avec la méthode RMSF (figure 61). L'analyse a montré qu'en l'absence de zinc les domaines I et VI sont plus flexibles et dynamiques (figure 61a). La région en amont du domaine I et celle entre les domaines IV-V présentent des fluctuations similaires aux domaines I et VI. De façon globale, les résultats montrent aussi que la transposase est plus flexible et dynamique en présence des mutations ponctuelles des cystéines (figure 61b, 61c et 61d).

En conséquence, ces résultats confortent l'idée que le ZF1 est impliqué dans la stabilité de la protéine et que ZF2 serait davantage impliqué dans une interaction avec l'ADN car les cystéines C53, C58, C360 et C363 du ZF2 montrent une forte flexibilité en présence de zinc.

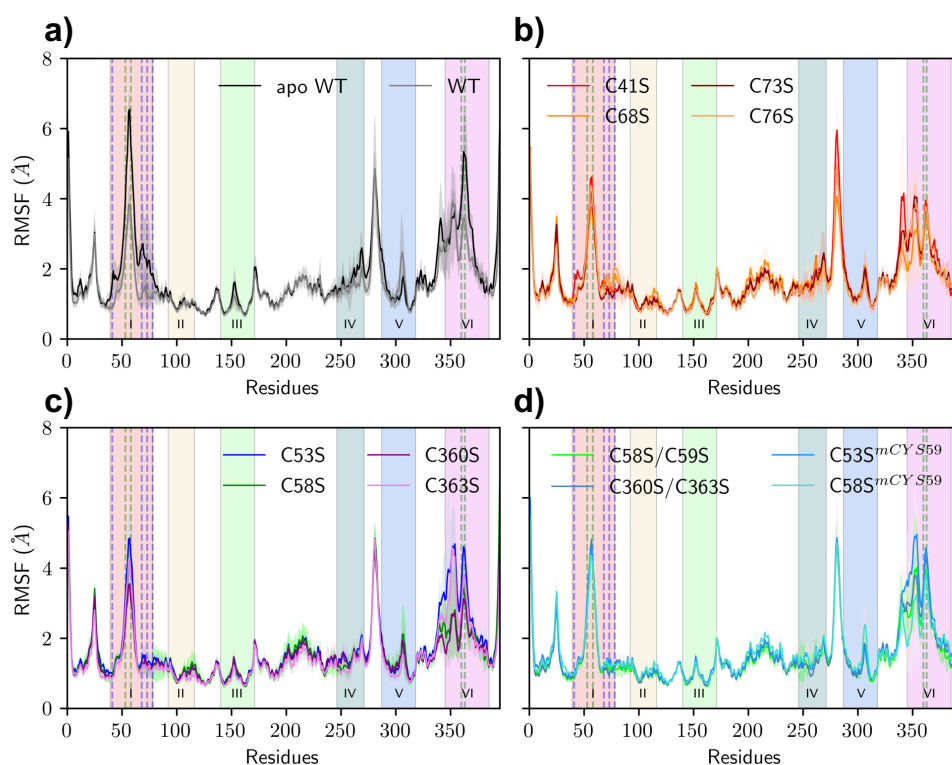


Figure 61 : Effet des mutations ponctuelles sur la flexibilité ou la rigidité des domaines de la transposase TnpA<sub>1S91</sub>

Les domaines I, II, III, IV, V et VI sont représentés en rouge, brun, vert, cyan, bleu et rose, respectivement. La position des résidus est représentée en axe X et la distance RMSF (Å) en axe Y.

- a) Effet de la présence ou non de zinc sur la flexibilité et la dynamique de la transposase sauvage.
- b) Effet de la mutation des cystéines C41, C68, C73 et C76 sur la flexibilité et la dynamique en présence de zinc.
- c) Effet de la mutation des cystéines C53, C58, C360 et C363 sur la flexibilité et la dynamique en présence de zinc.
- d) Effet de la mutation des doubles mutations sur la flexibilité flexibles et la dynamique en présence de zinc.

Nous avons aussi regardé l'effet des mutations ponctuelles des cystéines du doigt de zinc putatif en présence de zinc. L'analyse du ZF1 en présence des quatre cystéines (C41, C68, C73 et C76) montre une coordination N=4 du Zinc par les 4 groupement thiolates dirigés vers le Zinc, adoptant une géométrie tétraédrique. Il est important de noter que le nombre de coordination N=4.9 (figure 62a) doit être considéré avec attention. En effet, l'analyse dynamique montre une proximité d'une molécule d'eau en moyenne, mais ne restant pas tout au long de la simulation et s'échangeant au cours du temps. De plus, l'orientation de la molécule d'eau ne dénote pas d'une liaison dirigée entre l'oxygène et le Zn et ne modifie pas la géométrie tétraédrique caractéristique du Zn coordonné à quatre ligands (figure 62b). Les

résultats de l'analyse dynamique montrent que l'eau devient un ligand du zinc dès lors que l'une des quatre cystéines est mutée.

En présence d'une mutation ponctuelle, le centre métallique peut se lier à six ligands ( $N = 6$ ), l'ion zinc est hexacoordonné.

L'effet de la mutation de la cystéine 68 du ZF1 est présenté en figure 62c. Contrairement à la condition sauvage, l'analyse en dynamique du mutant présente plusieurs molécules d'eau capables de se lier au zinc.

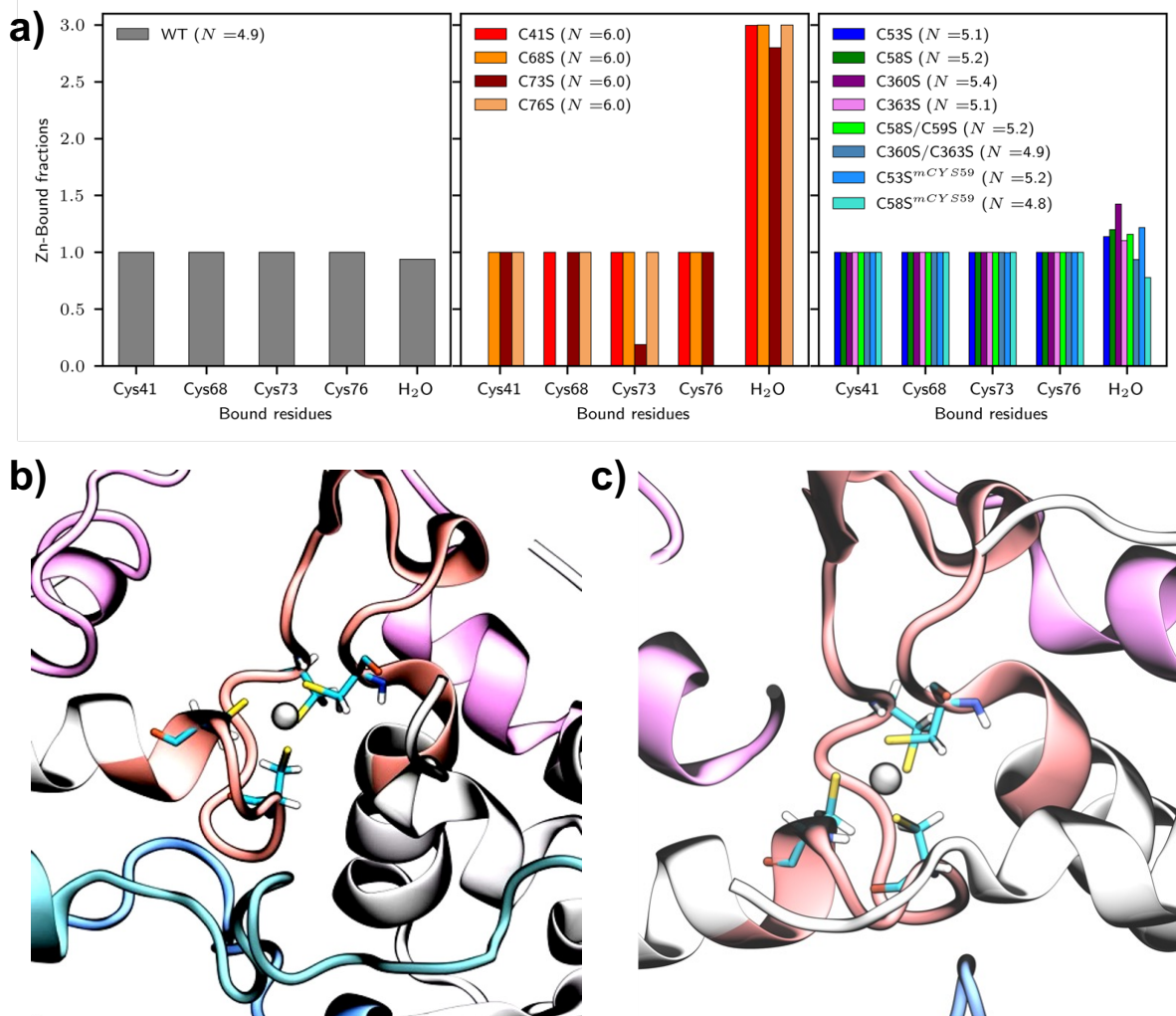


Figure 62 : Effet des mutations ponctuelles du ZF1 de la transposase TnpA<sub>IS91</sub>

a) Capacité de liaison du zinc aux cystéines 41, 68, 73, 73 et à H<sub>2</sub>O. A gauche est présenté le graphique dans la condition WT, au milieu, l'effet des mutations des cystéines du ZF1 et à droite l'effet des mutations des cystéines ZF2 sur la capacité de liaison du zinc ZF1.

Les domaines I et VI sont représentés en rouge et rose, respectivement. Les cystéines du ZF1 sont représentées en bâtonnet et l'atome zinc par une sphère blanc. b) Condition sans mutation. c) Condition avec la mutation de la cystéine 68 en sérine.



Pour ZF2, l'analyse dynamique montre que l'eau devient un ligand du zinc quand l'une des quatre cystéines est mutée. En effet, en l'absence de mutation, le centre métallique présente la capacité à se lier à quatre atomes ( $N = 4$ ), l'ion zinc est tétraédrique (figure 63a). *A contrario*, en présence d'une mutation ponctuelle, le centre métallique devient lié à six atomes ( $N = 6$ ), l'ion zinc est hexacoordonné et les molécules d'eau deviennent des ligands du zinc. De plus, les cystéines adjacentes aux cystéines 53 et 58 ne remplacent pas une des cystéines mutées dans la liaison au zinc.

A l'instar du ZF1, l'analyse du ZF2 montre qu'en présence des quatre cystéines (C53, C59, C360 et C363), l'eau n'agit pas comme un ligand de zinc car le numéro de coordination  $N$  est strictement égal à 4 (figure 63b). Ces résultats montrent que les mutations du ZF2 n'influencent pas la conformation de coordination du ZF1 qui reste égale à 4.

L'effet de la mutation de la cystéine 360 du ZF2 est présenté en figure 63c. Contrairement à la condition sauvage, l'analyse en dynamique du mutant présente plusieurs molécules d'eau capables de se lier au zinc.

Les activités de transposition plus faibles du système muté ZF2 ne sont pas liées à la substitution de la cystéine dans la configuration ZF2. Ceci soulève l'interrogation de l'implication de la structure locale du ZF2 dans la liaison à l'ADN.

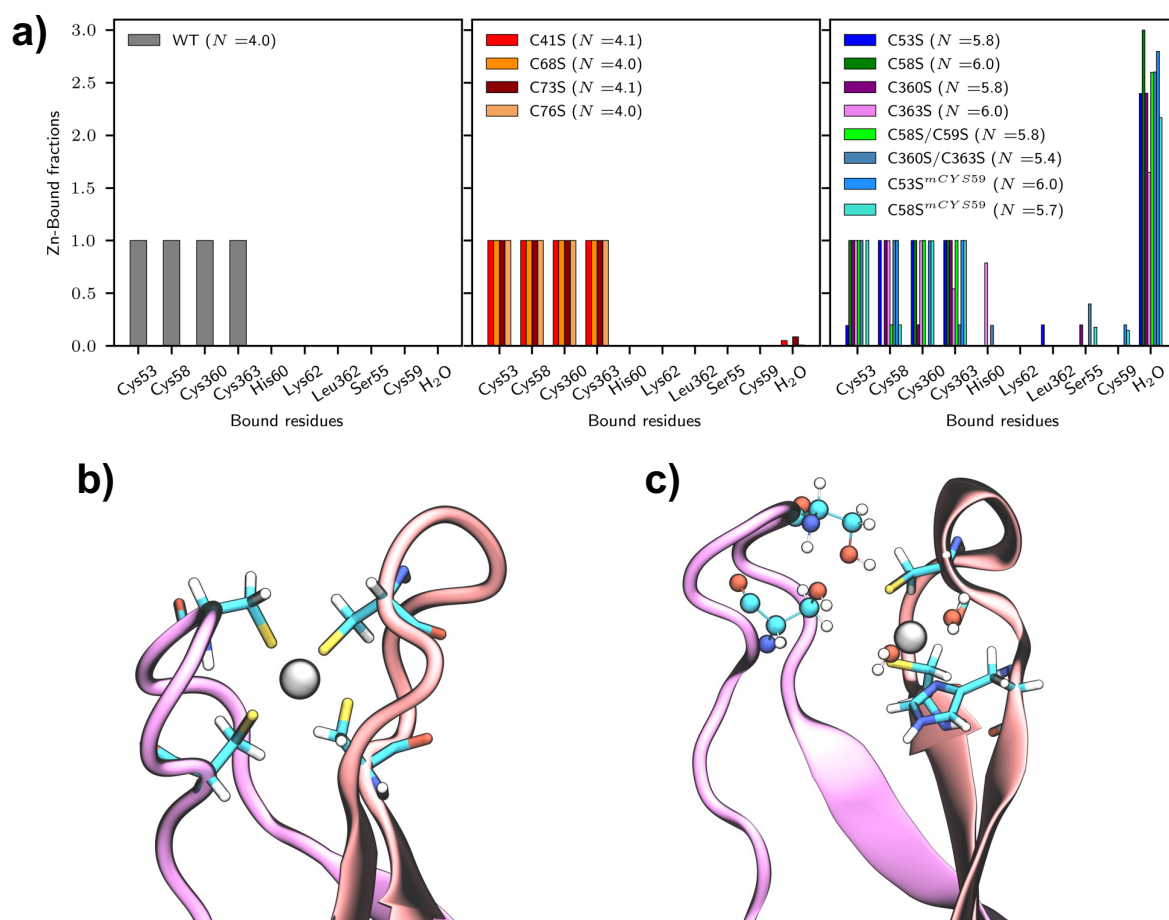


Figure 63 : Effet des mutations ponctuelles du ZF2 de la transposase TnpA<sub>1S97</sub>

a) Capacité de liaison du zinc par rapport aux cystéines 53, 58, 360, 363, 59 ainsi que H<sub>2</sub>O et l'histidine 60, la leucine 362 et la sérine 55. A gauche est présenté le graphique dans la condition WT, au milieu, l'effet des mutations des cystéines ZF1 sur la capacité de liaison du zinc ZF2 et à droite l'effet des mutations des cystéines ZF2. Le domaine I et VI est représenté en rouge et rose, respectivement. Les cystéines du ZF2 sont représentées en bâtonnet et l'atome zinc par une sphère blanc. b) Condition WT et sans mutation. c) Condition avec la mutation de la cystéine 360 en sérine.

## II.4. Discussion

Les investigations menées *in vivo* et *in silico* ont permis de mieux appréhender l'impact et le rôle des domaines I et VI, ces derniers étant conservés chez les éléments de la famille IS91. Nous avons constaté l'importance de huit cystéines pour l'activité de la transposase TnpA<sub>IS91</sub> (figure 64). En effet, nous avons montré que six cystéines sont localisées dans le domaine I et deux cystéines dans le domaine VI. L'ensemble des résultats exposés nous a permis de définir deux doigts de zinc que nous avons nommés ZF1 et ZF2 :

- ZF1 serait impliqué dans la stabilité et la structuration de la transposase. Il ferait intervenir les cystéines 41, 68, 73 et 76.
- ZF2 serait impliqué dans une interaction avec l'ADN. Il ferait intervenir les cystéines 53, 58, 360 et 363.

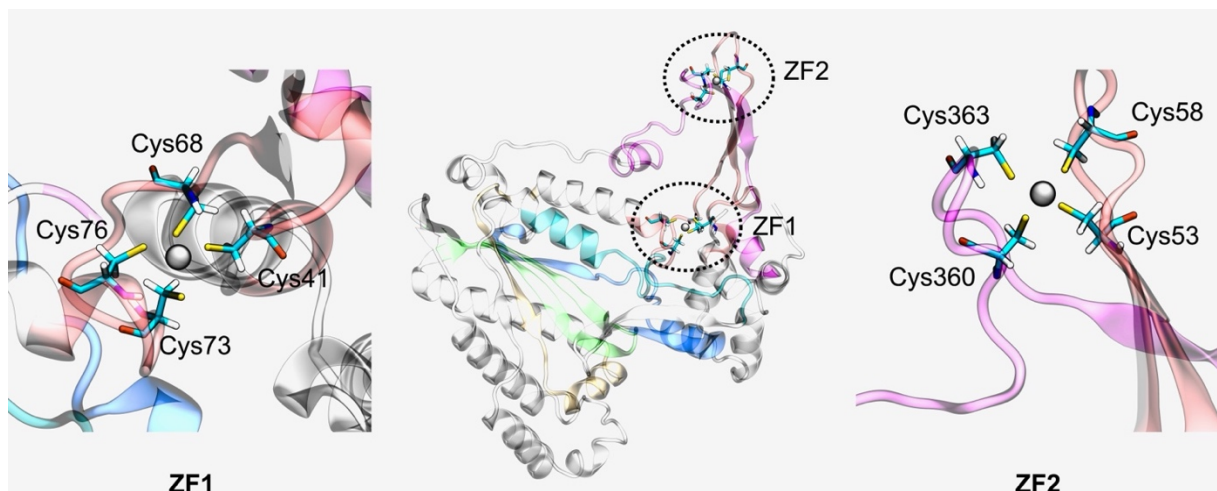


Figure 64 : Localisation et structure des doigts de zinc ZF1 et ZF2 de TnpA<sub>IS91</sub>

Les domaines I, II, III, IV, V et VI sont représentés en rouge, brun, vert, cyan, bleu et rose, respectivement. Les cystéines du ZF1 et ZF2 sont représentées en bâtonnet. A gauche est représenté un zoom sur le ZF1 et à droite sur le ZF2 avec l'atome zinc représenté par une sphère blanche. Enfin, au milieu, la localisation des ZF1 et ZF2 chez la transposase TnpA<sub>IS91</sub>.

L'implication de ZF2 dans une interaction avec l'ADN est soutenue par sa localisation à l'extrémité du doigt. En effet, cette extrémité est excentrée par rapport au « core » de la transposase. De plus, lorsque nous mettons les résidus de deux histidines H159 et H161 du motif HUH et les deux tyrosines Y249 et Y253 du site catalytique sur la transposase (figure 65), nous constatons que le haut du doigt contenant le ZF2 pointe vers l'extérieur de la protéine

et que le ZF1 est à proximité du motif HUH et du site catalytique. Cette constatation renforce l'idée que ZF1 et ZF2 ont des rôles distincts dans la fonction de la transposase.

En effectuant une rotation de la transposase, nous apercevons la présence d'un sillon qui pourrait permettre le passage de l'ADN. Dans le prolongement de ce sillon, nous remarquons le haut du doigt avec ZF2. Enfin, dans ce sillon, nous avons observé la présence du domaine V conservé dans la famille IS91 dont la fonction est inconnue. Au vu de cette structure, le domaine V pourrait intervenir dans l'interaction avec l'ADN en permettant l'ouverture ou non de ce sillon.

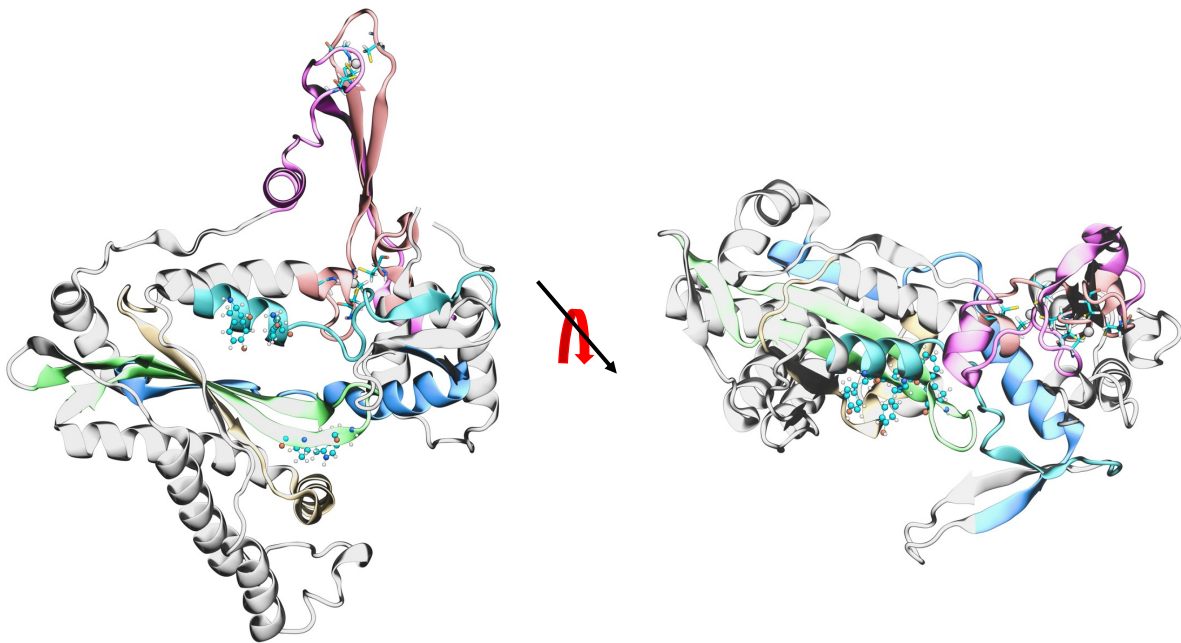


Figure 65 : Un sillon potentiel permettant le passage de l'ADN au niveau de la transposase TnpA<sub>IS91</sub>  
*Les domaines I, II, III, IV, V et VI sont représentés en rouge, orange, vert, cyan, bleu et rose, respectivement. Les cystéines du ZF1 et ZF2 sont représentées en bâtonnet, et les tyrosines du motif catalytique et les histidines du motif HUH en sphère. La flèche noire indique l'axe de rotation et la flèche rouge la rotation effectuée.*

Pour terminer, nous avons réalisé un alignement des structures 3D issues d'AlphaFold des transposases TnpA<sub>IS91</sub>, TnpA<sub>IS1294b</sub> et TnpA<sub>IS801</sub>. Cet alignement permet de mettre en évidence que ces trois transposases ont une structure similaire (figure 66). Ainsi, les résultats obtenus avec la transposase TnpA<sub>IS91</sub> pourraient être appliqués à ces autres transposases de la famille IS91.

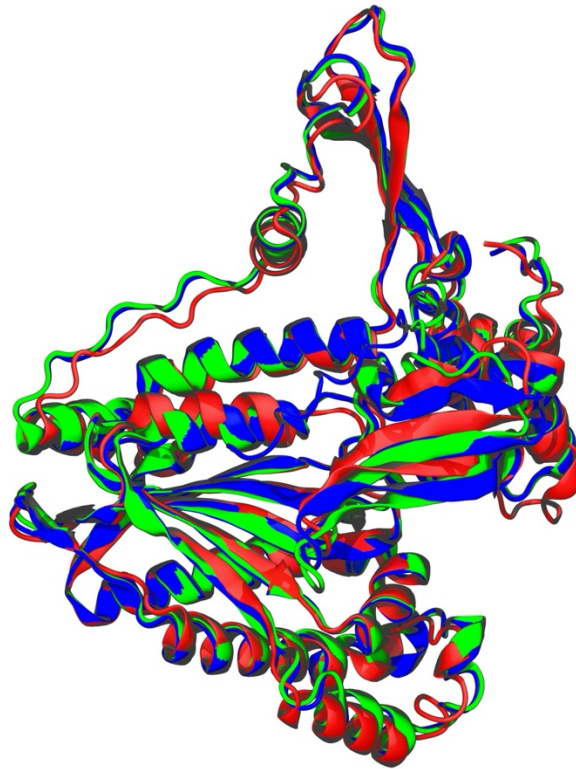


Figure 66 : Superposition des structures AlphaFold2 prédites des transposases IS91, IS801 et IS1294b

*Alignement des structures prédites sur AF2 pour les transposases IS91 (id. prot. Q8KMX5, rouge), IS801 (id. prot. P24607, bleu) et IS1294b (id. prot. S0F1K3, vert).*

Néanmoins, il est important de rappeler que l'utilisation de ces outils de modélisation doivent être considérés avec attention en attendant la résolution expérimentale de la structure d'intérêt. Dans le cas de TnpA<sub>IS91</sub>, le modèle d'AF2 prédit nativement les motifs de fixation au Zinc de type Zinc-Finger (ZF).

Néanmoins, la modélisation moléculaire reste un outil de choix pour prédire l'impact structural issue d'une mutation.

Au vu de nos résultats, nous pouvons supposer une structure similaire entre les transposases de IS91, IS1294b et IS801, connues pour être transposables . La question de l'existence des doigts de zinc et leur rôle se pose aussi pour les autres éléments de la famille IS91. Une représentation du motif du ZF1 du côté N-terminal pour les TnpA<sub>IS91</sub>, TnpA<sub>IS1294b</sub> et TnpA<sub>IS801</sub>, mais aussi TnpA<sub>ISCR1</sub> (id protéine : CAJ84008.1) , TnpA<sub>ISCR2</sub> (id protéine : AUN95242.1) et

TnpA<sub>ISKpn22</sub>, qui a été isolée dans une cassette de gène d'un intégron (id protéine : H1ACB9) (figure 67).

On peut voir que les trois éléments actifs pour la transposition (TnpA<sub>IS91</sub> TnpA<sub>IS1294b</sub> et TnpA<sub>IS801</sub>) présentent le même profil. En effet, les cystéines importantes du doigt de zinc putatif de la transposase de IS91 sont positionnées identiquement.

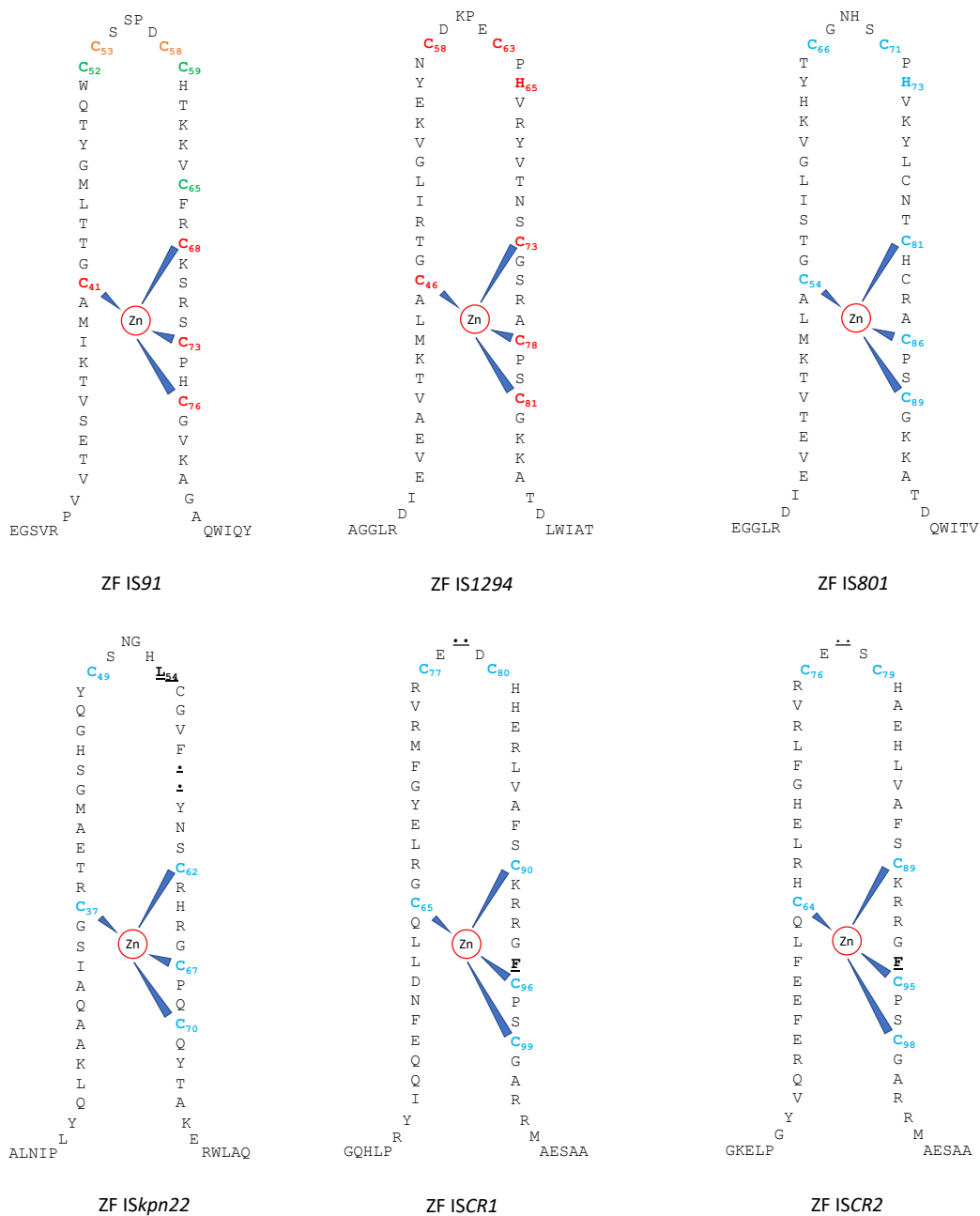


Figure 67 : Représentation hypothétique du doigt de zinc ZF1 des transposases IS91, IS1294b, IS801, ISkpn22, ISCR1 et ISCR2

Les cystéines ou l'histidine sont représentées en rouge pour un effet délétère sur l'activité de la transposase, en orange pour un effet intermédiaire et en vert pour aucun effet. En bleu sont représentées les cystéines conservées par rapport aux éléments IS91, IS1294b, IS801. En gras et souligné sont indiquées les mutations ou les insertions, puis les délétions par des points soulignés.

Chez les transposases TnpA<sub>ISKpn22</sub>, TnpA<sub>ISCR1</sub> et TnpA<sub>ISCR2</sub>, nous constatons l'absence de deux résidus en haut du doigt (correspondant à la localisation entre les cystéines C53 et C58 de l'élément IS91), ce qui peut avoir un effet important sur l'activité de la transposase comme l'ont montré les mutations C53 et C58 chez IS91. Ainsi, par homologie avec les résultats obtenus pour la transposase TnpA<sub>IS91</sub>, ces trois transposases pourraient être inactives pour la transposition. De plus, les transposases TnpA<sub>ISCR1</sub> et TnpA<sub>ISCR2</sub> présentent une phénylalanine entre deux cystéines de localisation correspondant aux cystéines C68 et C73 de l'élément IS91 et pour lesquelles aucune activité de transposition n'a été détectée en présence de mutations. Enfin, la transposase TnpA<sub>ISKpn22</sub> possède une leucine en position 54 au lieu d'une cystéine comme pour les éléments étant actifs dans la transposition. La mutation ponctuelle en cette position chez IS91 avait permis de constater une réduction faible mais significative de l'activité de transposition par rapport à la condition sans inducteur.

Ainsi, ces représentations nous laissent supposer que la configuration spatiale au niveau du doigt de zinc putatif est similaire pour les éléments IS91, IS1294b et IS801. De plus, par homologie avec celles de l'élément IS91, cela suggère que les transposases ISKpn22, ISCR1 et ISCR2 seraient inactives pour la transposition. Néanmoins des investigations plus poussées par une approche *in vivo* seraient nécessaires pour mieux caractériser ces observations.



## Partie III. Étude de la transposition d'autres éléments de la famille IS91 : ISCR et ISKpn22

---

Les éléments de la famille ISCR sont généralement adjacents à des gènes de résistance aux antibiotiques. Ainsi, ces éléments joueraient un rôle prédominant dans la dissémination et l'expression des gènes de résistance aux antibiotiques (Lallement et al., 2018; Toleman et al., 2006b, 2006a; Toleman and Walsh, 2010). Cependant, aucune preuve de leur activité de transposition n'a jamais été démontrée à ce jour.

La transposase TnpA<sub>ISKpn22</sub>, a été isolée dans une cassette de gène d'un intégron de classe 1 In722 (Genbank : AB616660, id protéine : H1ACB9). Ses deux extrémités *oriIS* et *terIS* ne sont pas connues, ce qui rend son étude compliquée. Actuellement, aucune donnée n'existe sur la capacité de ISKpn22 à être actives dans la transposition. Nous savons uniquement que la séquence protéique de la transposase TnpA<sub>ISKpn22</sub>, est très proche de celle de la famille IS91 avec un score E-value de  $10^{-93}$  avec la transposase ISVsA10 de la famille IS91. De plus, la transposase ISKpn22 possède deux tyrosines catalytiques contrairement aux ISCR qui possèdent une seule tyrosine.

C'est pourquoi, nous avons décidé d'étudier l'activité des transposases ISKpn22, ISCR1 et ISCR2 et leur capacité propre à reconnaître les éléments IS91, ISCR1 et ISCR2.

### III.1. Matériels et Méthodes

Nous avons utilisé des expériences de « *mating-out* » chez *Escherichia coli* comme précédemment décrit (*Partie I sous partie 1.1* dans l'article « *IS91-encoded orf121 regulates negatively IS91 transposition* »). Tous les résultats obtenus en « *mating-out* » ont été analysés à l'aide d'un test de Mann-Whitney à deux groupes non-appariés.

La liste des souches et des plasmides utilisés est présentée dans le tableau 5. La liste des amorces utilisées pour la construction des plasmides est disponible en annexe 6.

Tableau 5 : Souches et plasmides utilisés lors de l'étude de la transposition des éléments ISCR et ISKpn22

$\Omega$  indique une insertion ; :: indique une nouvelle articulation ; \* indique une mutation

Nom souche / plasmide	Surnom	Génotypes / Caractéristiques	Source / référence
<b>Souche utilisées</b>			
JS219		MC1061, <i>recA1</i> , <i>lacI<sup>q</sup></i>	<i>Cam et al., 1988</i>
MC240		XA103, F <sup>-</sup> , <i>ara</i> , <i>del(lac pro)</i> , <i>gyrA</i> ( <i>nalR</i> ), <i>metB</i> , <i>argE<sup>am</sup></i> , <i>rpoB</i> , <i>thi</i> , <i>supF</i>	Gift of B. Ton-Hoang
Salmonella enterica Serovar Keurmassar		Source of ISCR1 for <i>rcr1</i>	<i>Garnier et al., 2006</i>
Photobacterium damsela subsp. Piscicida		Source of ISCR2 for <i>rcr2</i>	Juiz-Río et al., 2005
<b>Plasmide de référence et mère</b>			
pSU2600		Source of IS91 for TnpA <sub>IS91</sub>	<i>Mendolia et al, 1994</i>
pKPI-6		Source of ISKpn22 contenant l'intégron de classe 1 In722 ; GenBank : AB616660	<i>Shigemoto et al., 2012</i>
pUC18		Amp <sup>r</sup> ; MCS from M13mp19 into <i>lacZ'</i>	<i>Yanisch-Perron et al., 1985</i>
pGY13556		pAPT110Δkan <sup>r</sup> ; P <sub>lac</sub> :: <i>tnpA</i> <sub>ISDra2-his6</sub>	<i>Pasternak et al, 2013</i>
pOX38Km		Km <sup>R</sup>	<i>Chandler and Galas, 1983b</i>
<b>Plasmide mimant une IS</b>			
pCP7	mini-IS91	pUC18 ; <i>ter91</i> P <sub>cat</sub> ::Cam <sup>R</sup> <i>ori91</i>	This work
pAF24	mini-ISCR1	pUC18; <i>terIS91</i> P <sub>cat</sub> ::Cam <sup>R</sup> <i>oriISCR1</i>	This work
pCP67	mini-ISCR2	pUC18; <i>terISCR2</i> P <sub>cat</sub> ::Cam <sup>R</sup> <i>oriISCR2</i>	This work
<b>Plasmide exprimant les transposases</b>			
pCP2	P <sub>Lac</sub>	pGY13556ΔtnpA <sub>ISDra2</sub> Ωlinker79	This work
pCP4	P <sub>lac</sub> :: <i>tnpA</i>	pCP2; P <sub>lac</sub> :: <i>tnpA</i>	This work
pNV1	TnpA <sub>ISCR1</sub>	pCP2; P <sub>lac</sub> :: <i>rcr1</i>	This work
pAF5	TnpA <sub>ISCR1</sub> Y2	pCP2; P <sub>lac</sub> :: <i>rcr1</i> *R325Y	This work
pAF20	TnpA <sub>ISCR1</sub> HIH Y2	pCP2 ; P <sub>lac</sub> :: <i>rcr1</i> *Y190I ; R325Y	This work
pAF21	TnpA <sub>ISCR1</sub> HIH	pCP2; P <sub>lac</sub> :: <i>rcr1</i> *Y190I	This work
pNV2	TnpA <sub>ISCR2</sub>	pCP2; P <sub>lac</sub> :: <i>rcr2</i>	This work
pAF6	TnpA <sub>ISCR2</sub> Y2	pCP2; P <sub>lac</sub> :: <i>rcr2</i> *R322Y	This work
pAF22	TnpA <sub>ISCR2</sub> HIH Y2	pCP2; P <sub>lac</sub> :: <i>rcr2</i> *F189I ; R322Y	This work
pAF23	TnpA <sub>ISCR2</sub> HIH	pCP2; P <sub>lac</sub> :: <i>rcr2</i> *F189I	This work
pCP44	TnpA <sub>ISKpn22</sub>	pCP2; P <sub>lac</sub> :: <i>tnpA</i> <sub>ISKpn22</sub>	This work
pAF7	TnpA <sub>ISKpn22</sub> L54C	pCP2; P <sub>lac</sub> :: <i>tnpA</i> <sub>ISKpn22</sub> *L54C	This work

## III.2. Résultats

### III.2.1. Analyse des transposases

Nous avons tout d'abord comparé la séquence protéique des différentes transposases (figure 68). Les éléments *ISCR1* et *ISCR2* présentent une tyrosine (Y) et une phénylalanine (F) au niveau du motif HUH, respectivement. De plus, ce motif HUH est composé de deux résidus histidine (H) qui sont séparés par un résidu hydrophobe volumineux (U). En outre, le motif HUH est impliqué dans la reconnaissance et la liaison à des structures d'ADN en épingle à cheveux, ce qui permet la reconnaissance des sites de clivage pouvant être localisés soit dans l'épingle à cheveux, soit dans l'ADNsb adjacent à la structure. Afin de démontrer s'il existe ou non un effet des résidus Y et F sur la capacité de mobilité des éléments *ISCR*, nous avons décidé d'effectuer une mutation ponctuelle de ces résidus en isoleucine (I) comme chez la transposase de *IS91*.

	Domaine I	Domaine III	Domaine IV
TnpA <sub>IS91</sub>	MA <b>C</b> GGTTLMGYTQW <b>C</b> SSPD <b>C</b> CHTKKVCFR <b>C</b> KSR-S <b>C</b> PH <b>C</b> GV	DQQW <b>H</b> PH <b>I</b> HL-S <b>T</b> T	VAV <b>Y</b> FGS <b>Y</b> LKKP
TnpA <sub>IS801</sub>	LA <b>C</b> GTSILGVKHY <b>T</b> CGNHS <b>C</b> PHVKYLCNT <b>C</b> HCR-A <b>C</b> PS <b>C</b> GK	RLNW <b>H</b> PH <b>V</b> HL-S <b>V</b> T	TVN <b>Y</b> LG <b>R</b> Y <b>L</b> KKP
TnpA <sub>IS1294b</sub>	LA <b>C</b> GTRILGVKEYN <b>C</b> DKPE <b>C</b> PHVRYVTNS <b>C</b> GSR-A <b>C</b> PS <b>C</b> GK	RLNW <b>H</b> PH <b>V</b> HV-S <b>V</b> T	TAR <b>Y</b> LG <b>R</b> Y <b>L</b> KKP
TnpA <sub>ISKpn22</sub>	SG <b>C</b> RTEAMGSHGQ <b>Y</b> CSNGH <b>L</b> CGVF--YNS <b>C</b> RHR-G <b>C</b> PQ <b>C</b> YT	NLSEHP <b>H</b> I <b>H</b> CL <b>I</b> TH	VAK <b>Y</b> LAR <b>Y</b> MRGG
TnpA <sub>ISCR1</sub>	LQ <b>C</b> GRLEYGFMRV <b>R</b> C <b>E</b> --D <b>C</b> HHERLVAFS <b>C</b> KRRGF <b>C</b> PS <b>C</b> GA	ALN <b>L</b> NV <b>H</b> Y <b>H</b> M <b>L</b> F <b>L</b> D	KLE <b>R</b> LC <b>R</b> Y <b>I</b> SRP
TnpA <sub>ISCR2</sub>	LQ <b>C</b> GRLEYGFMRV <b>R</b> C <b>E</b> --D <b>C</b> HHERLVAFS <b>C</b> KRRGF <b>C</b> PS <b>C</b> GA	ALN <b>L</b> NV <b>H</b> F <b>H</b> M <b>L</b> F <b>L</b> D	KLE <b>R</b> LC <b>R</b> Y <b>I</b> SRP

Figure 68 : Alignement des transposases TnpA<sub>IS91</sub>, TnpA<sub>IS801</sub>, TnpA<sub>IS1294b</sub>, TnpA<sub>ISKpn22</sub>, TnpA<sub>ISCR1</sub> et TnpA<sub>ISCR2</sub>

Au vu de ces données, par homologie avec les 3 transposases connues pour être actives dans la transposition, nous avons décidé de tester différentes combinaisons :

- La transposase TnpA<sub>IS91</sub> a été utilisée comme contrôle positif de la transposition.
- Les transposases sauvages TnpA<sub>ISKpn22</sub>, TnpA<sub>ISCR1</sub> et TnpA<sub>ISCR2</sub>
- Les transposases *ISCR1* et *ISCR2* avec une mutation ponctuelle du motif HUH de HYH et HFH vers HIH dans les deux cas : TnpA<sub>ISCR1</sub> HIH et TnpA<sub>ISCR2</sub> HIH.
- Les deux transposases *ISCR1* et *ISCR2* ont été testées avec une mutation ponctuelle au niveau du site catalytique pour passer d'une tyrosine à deux tyrosines aux sites catalytiques (comme pour *IS91*), RLCRY vers YLCRY : TnpA<sub>ISCR1</sub> Y2 et TnpA<sub>ISCR2</sub> Y2.
- Enfin, les deux mutations ponctuelles ont été testées pour les transposases *ISCR*. Ainsi, les deux transposases *ISCR1* et *ISCR2* présentent une mutation ponctuelle de HYH

RLCRY et HFH RLCRY vers HIH YLCRY dans les deux cas : TnpA<sub>ISCR1</sub> HIH Y2 et TnpA<sub>ISCR2</sub> HIH Y2.

- La transposase ISKpn22 avec une mutation ponctuelle de la leucine 54 en cystéine : TnpA<sub>ISKpn22</sub> L54C.

L'activité de ces transposases a été testée sur des mini-IS de IS91, ISCR1 et ISCR2, avec une cassette Cm<sup>R</sup> encadrée par les extrémités *oriIS91* et *terIS91*, *oriISCR1* et *terIS91*, *oriISCR2* et *terISCR2*, respectivement. Nous avons décidé d'utiliser le *terIS91* avec *oriISCR1* car le *terISCR1* n'est pas caractérisé.

Les tétranucléotide de clivage sont *terIS91* (5'CTCG) et *terISCR2* (5'TCCC). Adjacents à l'extrémité *oriIS*, les plasmides donneurs contiennent le tétranucléotide 5'CTTG nécessaire au clivage et à la transposition.

### III.2.2. Efficacité de transposition

Au préalable, nous avons vérifié que la conjugaison était stable (entre  $5 \times 10^{-2}$  et  $9 \times 10^{-2}$ ) entre les réplicats biologiques et techniques pour ne pas impacter et induire une variabilité dans les résultats de transposition induite par la fluctuation de la conjugaison.

La figure 69 présente les fréquences de transposition associées aux différentes substitutions générées dans les transposases TnpA<sub>ISKpn22</sub>, TnpA<sub>ISCR1</sub> ou TnpA<sub>ISCR2</sub> avec une mini-IS91, mini-ISCR1 ou mini-ISCR2.

**Premièrement**, la transposase TnpA<sub>IS91</sub> la fréquence de transposition était comparable aux études précédentes avec la mini-IS91 (*partie I.1 et partie II.2* des résultats), avec une activité basale de  $3,14 \times 10^{-5}$  et de  $2,43 \times 10^{-3}$  en présence d'IPTG, soit une augmentation significative de 77,26 fois. Les résultats montrent :

- Une transposition indétectable de la transposase de IS91, TnpA<sub>IS91</sub>, à l'état basal avec les mini-ISCR1 et mini-ISCR2. Cette activité était augmentée en présence d'IPTG induisant l'expression de la transposase mais restait faible ( $3,00 \times 10^{-7}$  et  $3,17 \times 10^{-7}$ , respectivement) en comparaison de l'activité de transposition avec la mini-IS91 ( $3,14 \times 10^{-5}$  à l'état basal et  $2,43 \times 10^{-3}$  après induction par IPTG).
- Les transposases sauvages TnpA<sub>ISKpn22</sub>, TnpA<sub>ISCR1</sub> et TnpA<sub>ISCR2</sub> ne présentent aucune activité de transposition, quelle que soit la mini-IS utilisée dans nos conditions de laboratoire (expression de la transposase basale ou induite). Des résultats similaires sont

observés avec les mutations ponctuelles pour les transposases  $TnpA_{ISCR1 Y2}$  et  $TnpA_{ISCR2 HIH}$ , mutées dans le site catalytique ou le site HIH, respectivement.

- La mutation ponctuelle de la leucine 54 en cystéine de la transposase  $TnpA_{ISKpn22}$  a permis de détecter une activité de transposition, relativement faible, mais significative, par rapport à la condition basale sans mutation. Une fréquence de transposition de  $2,42 \times 10^{-7}$ ,  $1,62 \times 10^{-7}$  et  $1,60 \times 10^{-7}$  a été observée en présence d'IPTG avec respectivement la mini-IS91, la mini-ISCR1 et la mini-ISCR2.
- La double mutation de ISCR1,  $TnpA_{ISCR1 HIH Y2}$ , nous avons détecté une activité de transposition en présence d'IPTG de  $2,70 \times 10^{-7}$ ,  $2,65 \times 10^{-7}$  et  $3,14 \times 10^{-7}$  soit une augmentation significative de 7,45 fois, 6,36 fois et 8,64 fois avec la mini-IS91, mini-ISCR1 et mini-ISCR2 par rapport à la condition basale où l'activité de transposition n'était pas détectable.

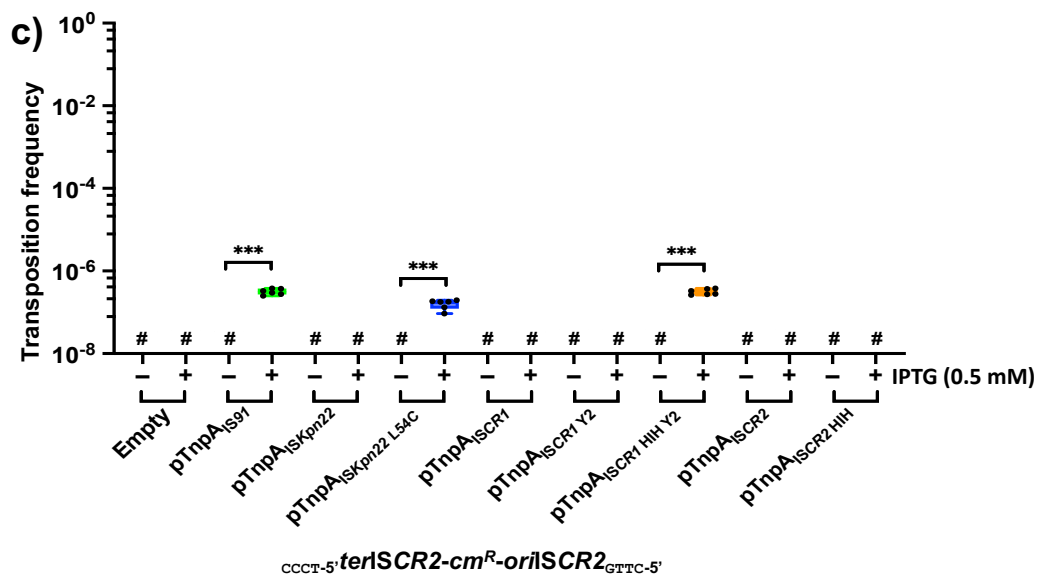
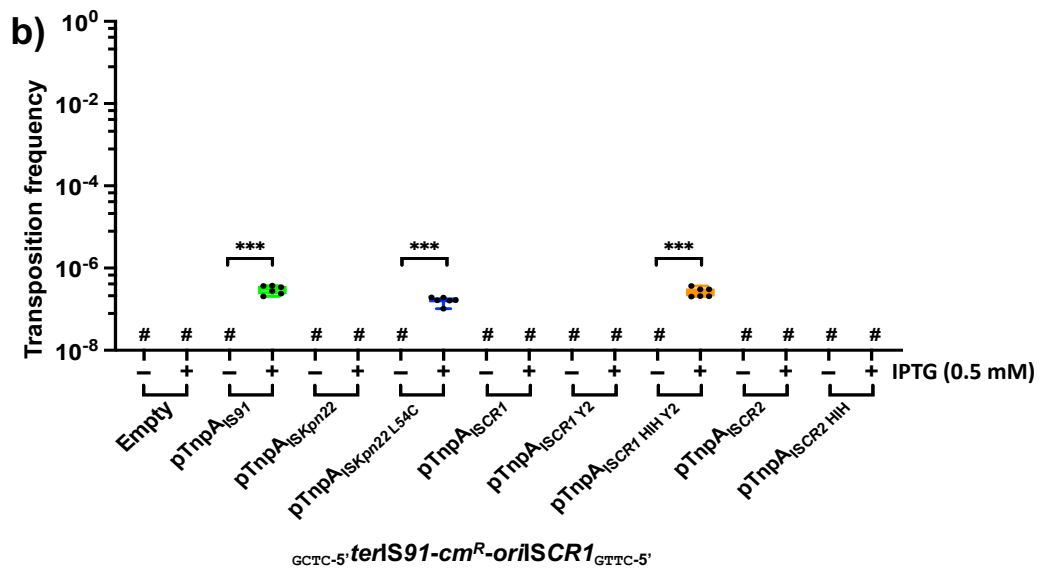
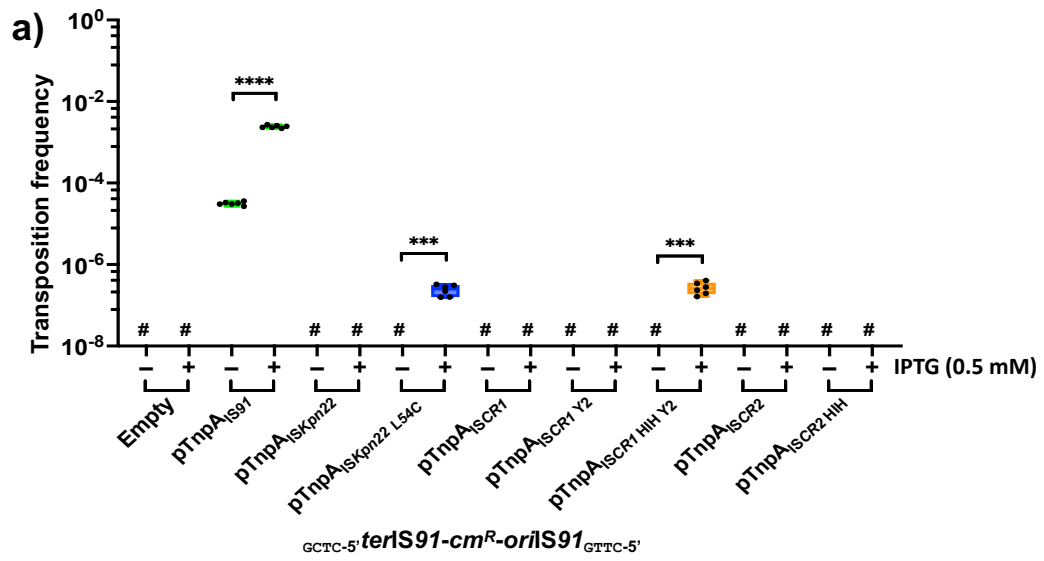


Figure 69 : Effet *in vivo* des mutations ponctuelles des transposases ISKpn22, ISCR1 et ISCR2 sur une IS mimant IS91, ISCR1 et ISCR2

Les barres d'erreur indiquent les écarts types ;  $n \geq 5$ . #, transposition indétectable. Le – et + indiquent la présence d'inducteur ou non à 0,5 mM. Empty, vecteur d'expression de la transposase vide ; pTnpA<sub>IS91</sub>, pTnpA<sub>ISCR1</sub>, pTnpA<sub>ISCR2</sub> et pTnpA<sub>ISKpn22</sub> le gène *tnpA* seul et sans mutation est exprimé en trans ; pTnpA<sub>ISCR1</sub> Y2 or HIH Y2, pTnpA<sub>ISCR2</sub> HIH et pTnpA<sub>ISKpn22</sub> L54C, le gène *tnpA* seul portant une ou deux mutations ponctuelles est exprimé en trans.

\*\*\*  $p < 0,0002$ , \*\*\*\*  $p < 0,0001$

a) Les fréquences de transposition observées avec le dérivé mini-IS91 <sub>GCTC-5</sub>terIS91-cm<sup>R</sup>-oriIS91<sub>GTTTC-5</sub>.

b) Les fréquences de transposition observées avec le dérivé mini-ISCR1 <sub>GCTC-5</sub>terIS91-cm<sup>R</sup>-oriISCR1<sub>GTTTC-5</sub>.

c) Les fréquences de transposition observées avec le dérivé mini-ISCR2 <sub>CCCT-5</sub>terISCR2-cm<sup>R</sup>-oriISCR2<sub>GTTTC-5</sub>.

Nous avons aussi analysé les taux de OET pour les différentes transposases (tableau 6). L'analyse statistique a démontré que la différence des pourcentages de OET n'était pas significativement différente entre les trois transposases pour lesquelles une transposition a été observée ( $TnpA_{IS91}$ ,  $TnpA_{ISKpn22 L54C}$  et  $TnpA_{ISCR1 HIH Y2}$ ). En moyenne, il a été observé environ 50% de OET pour ces transposases quelle que soit la mini-IS testée.

Tableau 6 : Pourcentage de OET observé en fonction des mutations ponctuelles des transposases  $ISKpn22$ ,  $ISCR1$  et  $ISCR2$  une IS mimant  $IS91$ ,  $ISCR1$  et  $ISCR2$

La valeur n indique le nombre de clones qui ont été analysés.

a) Les pourcentages de OET observées avec le dérivé mini- $IS91_{GCTC-5}terIS91-cm^R-orlIS91_{GTTTC-5}$ .

b) Les pourcentages de OET observées avec le dérivé mini- $ISCR1_{GCTC-5}terIS91-cm^R-orlISCR1_{GTTTC-5}$ .

c) Les pourcentages de OET observées avec le dérivé mini- $ISCR2_{CCCT-5}terISCR2-cm^R-orlISCR2_{GTTTC-5}$ .

**a)** *TnpA expression plasmid*    % one-ended transposition

$pTnpA_{IS91}$	51,67% n = 960
$pTnpA_{ISKpn22 L54C}$	48,28% n = 29
$pTnpA_{ISCR1 HIH Y2}$	56,00% n = 25

**b)** *TnpA expression plasmid*    % one-ended transposition

$pTnpA_{IS91}$	55,17% n = 29
$pTnpA_{ISKpn22 L54C}$	50,00% n = 12
$pTnpA_{ISCR1 HIH Y2}$	50,00% n = 20

**c)** *TnpA expression plasmid*    % one-ended transposition

$pTnpA_{IS91}$	54,55% n = 33
$pTnpA_{ISKpn22 L54C}$	60,00% n = 10
$pTnpA_{ISCR1 HIH Y2}$	46,15% n = 26



Dans nos conditions de laboratoire, nous avons montré que les éléments *ISKpn22*, *ISCR1* et *ISCR2* ne sont pas actifs pour la transposition. Cependant, nous avons démontré que la double mutation ponctuelle (HIH, YLCRY) chez *ISCR1* permet d'observer une activité de transposition. La mutation ponctuelle chez *ISKpn22* en position 54 d'une leucine vers une cystéine permet également d'observer une activité de transposition. De plus, nous avons démontré que les transposases TnpA<sub>IS91</sub>, TnpA<sub>ISCR1 HIH Y2</sub> et TnpA<sub>ISKpn22 L54C</sub> étaient capables de reconnaître l'extrémité *oriS* de *IS91*, *ISCR1* et de *ISCR2* avec un tétranucléotide cible <sup>5</sup>CTTG et l'extrémité *terIS* de l'élément *IS91* pour les mini-IS de *IS91*, *ISCR1* et le *terIS* de *ISCR2* pour la mini-*ISCR2*.

### III.3. Discussion

Nos résultats montrent que les éléments *ISKpn22*, *ISCR1* et *ISCR2* ne seraient pas actifs pour la transposition. Toutefois il est intéressant de noter que la transposase de l'élément *IS91* est capable de reconnaître l'extrémité *oriS* de *ISCR1* et de *ISCR2*. Ainsi, dans un contexte naturel, la diversité des éléments *ISCR1* et *ISCR2* aurait pu être créée par l'élément *IS91*. Au vu de la similarité de structure, il est possible que les transposases des éléments *IS1294*, *IS1294b* et *IS801* soient aussi capables de reconnaître ces extrémités *oriS*. Si c'était le cas ces IS auraient pu aussi participer à la diversité des éléments *ISCR1* et *ISCR2*. Par une approche de « mating-out », il serait intéressant de détecter la capacité des transposases *IS1294*, *IS1294b* et *IS801* à mobiliser la mini-*IS91*, la mini-*ISCR1* et la mini-*ISCR2*. De plus, il serait intéressant d'étudier aussi si la transposase TnpA<sub>IS91</sub> est capable de reconnaître l'*oriS* des autres éléments *ISCR*.

La mobilisation entre les éléments de la famille *IS91* et *ISCR* pourrait avoir plusieurs conséquences majeures d'un point de vue épidémiologique, en particulier en ce qui concerne la transmission de gènes de résistance aux antibiotiques et de virulence chez l'homme et les animaux. De plus, les éléments *IS91*, *IS1294* et *IS1294b* pourraient participer activement à la propagation de la résistance aux antibiotiques par une mobilisation inter-familiale (*IS91*-*ISCR*). *A contrario*, il est très peu envisageable que l'élément *IS801* participe à cette

mobilisation car cette IS est exclusivement retrouvée chez des plantes et uniquement le genre *Pseudomonas*.

Ainsi, la capacité de ces éléments à se mobiliser entre eux pourrait contribuer à la diversification des profils de résistance aux antibiotiques. Si des gènes de virulence sont transportés conjointement avec les gènes de résistance, cela pourrait augmenter la virulence des souches bactériennes.

C'est pourquoi une analyse approfondie par une approche *in silico* de l'environnement génétique des éléments de la famille IS91 et ISCR retrouvés régulièrement dans des isolats cliniques seraient intéressante en cherchant spécifiquement les gènes de résistance aux antibiotiques, de virulence.

La mobilisation des éléments IS91 et ISCR pourrait influencer la dynamique des infections bactériennes, en particulier dans les contextes cliniques et vétérinaires.

Une étude précédente a suggéré que chaque transposase est hautement spécialisée dans la reconnaissance de ses propres terminaisons, cela a été étayé par le fait que la transposase IS91 n'a pas réussi à reconnaître efficacement *terIS801* comme terminateur de transposition (Bernales, 1997). Or, nous démontrons ici que la transposase TnpA<sub>IS91</sub> est capable de reconnaître efficacement le *terISCR2*. Étant donné que le motif de l'extrémité *terIS* (figure 34) des éléments IS1294, IS1294b et IS801 est similaire à celle de IS91, il est fort probable que les éléments IS1294, IS1294b et IS801 n'aient en fait pas une telle spécificité de terminaison.

De plus, nous avons mis en évidence que les transposases TnpA<sub>ISCR1</sub><sup>HIH Y2</sup> et TnpA<sub>ISKpn22</sub><sup>L54C</sup> étaient capables de reconnaître le *terIS91* et le *terISCR2*. Cependant, l'activité détectée de transposition pour TnpA<sub>ISCR1</sub><sup>HIH Y2</sup> et TnpA<sub>ISKpn22</sub><sup>L54C</sup> reste très faible. Ainsi, pour obtenir une transposition plus active, l'ajout de deux cystéines en haut du doigt ZF1 et/ou l'enlèvement de la phénylalanine entre deux cystéines du ZF1 pourrait être nécessaire.

Pour les transposases ISCR1 et ISCR2, soulignons que dans des conditions *in vitro*, la transposase ISCR1 était capable de cliver l'extrémité *oriISCR1* de l'élément dans des conditions de très fortes quantités de transposase purifiée (500 ng) et sans ADN compétiteur (Quiroga et al., 2013). Ainsi, comme l'élément ISCR1 a une seule tyrosine catalytique il est envisageable que les ISCR agissent en tant que dimère, comme les transposases HUH de la famille IS200/IS605 (HUH Y1) qui ne présentent qu'une seule tyrosine catalytique. Chaque dimère apporterait ainsi chacun une tyrosine afin de pouvoir effectuer le clivage avec deux tyrosines catalytiques, tel qu'attendu.

Plusieurs hypothèses sur l'origine des éléments ISCR ont été émises, ISCR1 proviendrait d'environnements aquatiques (*Aeromonas salmonicida* et *Shewanella algae*) (Rhodes et al., 2000; Toleman et al., 2006a) et ISCR2 serait originaire d'un pathogène de certains poissons (*Pseudomonas damsalae* subsp. *piscida*) (Toleman et al., 2006a). Ainsi, les transposases ISCR1 et ISCR2 sont probablement originaires de l'environnement aquatique, ce contrairement aux éléments IS91 et IS1294b qui seraient d'origine humaine. Il est donc possible que les ISCR nécessitent des facteurs propres à l'environnement aquatique ou au poisson hôte de la bactérie porteuse.

## Conclusion et perspectives

---

Les éléments de la famille *IS91* et de la famille *ISCR* sont des séquences d'insertion qui, bien que fréquemment décrites en contexte clinique, sont peu caractérisées. Nos objectifs de travail étaient donc de mieux comprendre le paradigme de la famille, *IS91*, ainsi que les modalités de transposition de la famille *IS91* et de la famille *ISCR*.

Dans **un premier temps**, nous avons répertorié et analysé les caractéristiques des éléments *IS91*, *IS1294b* et *IS801* disponibles dans la Genbank® grâce à une étude *in silico*. Ce travail nous a permis de démontrer que :

- Ces éléments s'insèrent préférentiellement en 5'GTTC et 5'CTTG avec toujours un T en seconde position.
- Leurs tétranucléotides de clivage sont hautement conservés.
- Les éléments *IS91* et *IS1294b* sont principalement retrouvés chez les entérobactéries isolées, chez l'Homme ou les animaux. Ces deux éléments sont principalement présents en une seule copie sur des plasmides ayant une origine de répllication IncFII. *IS801* est isolée exclusivement chez *Pseudomonas* sp. et chez les plantes.
- L'analyse des isoformes *IS91* et *IS1294b* a montré que ces IS sont très conservées (*terIS*, *TnpA* et *oriIS*) et que la configuration *orf121-tnpA<sub>IS91</sub>*, ainsi que la protéine Orf121 sont également très conservées.

Bien que des publications font état de la présence de l'élément *IS91* et *IS1294b* à proximité de gènes de résistance, aucune étude de l'environnement génétique global n'a encore été réalisée. J'ai commencé à réaliser une analyse de ces environnements grâce au logiciel GCsnap (<https://github.com/JoanaMPereira/GCsnap>; (Pereira, 2021)) et en collaboration avec le docteur Lokman GALA de l'équipe Inserm -U1094 EpiMaCT. Les résultats préliminaires ont permis de mettre en évidence en amont et en aval de l'élément *IS91* :

- La présence de nombreux systèmes toxine-antitoxine de type II (PemK/MazF family, RelE/ParE family, toxin CcdA/CcdB, Phd/YefM family, antitoxin YacA, PemK/PemI et death-on-curing family toxin),
- La présence d'autres systèmes toxine-antitoxine (heat-stable enterotoxin ST-I group a, heat-stable enterotoxin ST-II, Hok family toxin, HlyD),

- La présence de gènes de résistance aux antibiotiques (*aph(3'')-Ib*, TEM-1, *tetR*, *tetB*, la famille de transposons de type EamA),
- La présence de gènes de résistance aux métaux lourds (*tehB* et *merP*)
- La présence de nombreux éléments génétiques appartenant à la famille IS1, IS66, IS21, IS3, IS6/26 et le transposon Tn3.
- La présence de nombreux gènes impliqués dans la conjugaison : *TrbJ*, *TraM*, *TraL*, *TraN*, *TraC*, *TraX*, *TraW*, *TraV*...

Ainsi, nous avons pu définir un résultat préliminaire global de l'environnement génétique présenté en figure 70.

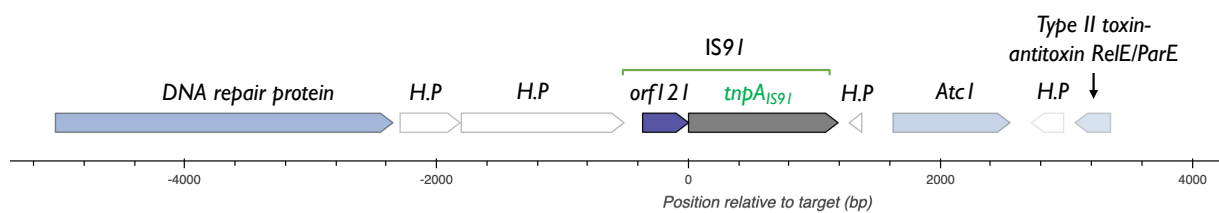


Figure 70 : Contexte génétique de quatre gènes en amont et en aval de l'élément IS91

Il convient toutefois de préciser que des mises au point sont encore nécessaires afin de tirer tout le potentiel des jeux de données grâce au logiciel GCsnap. Au terme de cette mise au point, une analyse à plus grande échelle sera réalisée sur les séquences de l'élément IS91, puis de l'élément IS1294b. De plus, une analyse séparant les IS plasmidiques et chromosomiques sera réalisée.

Ainsi l'analyse de l'environnement génétique permettrait de mieux comprendre l'association entre les plasmides de type IncFII et les éléments IS91 et IS1294b, ainsi que les gènes associés à ces éléments. De plus, ces deux IS sont généralement retrouvées en une seule copie par support génétique. Cela suggère que la formation de transposons composites pour ces deux éléments est plutôt rare. Ainsi, lors de l'analyse de l'environnement génétique, il serait intéressant dans le contexte naturel de détecter des insertions résultant d'une transposition ou d'une transposition OET.

Nous avons constaté que très peu d'isoformes ont été décrites pour l'élément IS801. Elles ont été trouvées exclusivement dans des plantes porteuses de *Pseudomonas spp*, sans préférence pour le support génétique entre chromosome et plasmide, sachant qu'il n'a pas été possible de déterminer le type de plasmide dans lequel l'élément IS801 a été trouvé.

Cependant, l'analyse des bases de données peut être biaisée. En effet, pour un grand nombre de numéros d'accès (représentant en moyenne un quart des cas), l'hôte ou la bactérie a été isolé, mais le support génétique et le type d'origine de réplication du plasmide ne sont pas disponibles.

Dans **un second temps**, nous avons mis en évidence la contribution de Orf121 dans la régulation de la transposition de l'élément IS91 :

- Nos résultats suggèrent que *orf121* et le *tnpA* sont principalement transcrits en un ARNm bicistronique à partir du promoteur du  $P_{orf121}$ . La transcription des deux gènes est couplée et le chevauchement est nécessaire à cette transcription.
- La co-expression de Orf121 (*en cis* ou *en trans*), en plus de la transposase, diminue fortement et significativement la fréquence de transposition de IS91 dans l'hôte *E. coli*.
- La protéine Orf121 a un effet négatif sur l'efficacité de la transposition, son effet étant indépendant de la configuration naturelle entre *orf121* et  $tnpA_{IS91}$ .
- L'intermédiaire circulaire activé pour l'insertion de l'élément IS91 est monocaténaire et il concerne uniquement le brin inférieur. De plus, l'expression séparée des gènes *orf121* et *tnpA* met en évidence l'impact négatif du produit Orf121 sur la capacité de la transposase TnpA à insérer l'intermédiaire circulaire.
- Orf121 n'est pas impliquée dans la sélection de la cible et la spécificité d'insertion de IS91.
- Le produit *orf121* est important pour la reconnaissance précise et le clivage de l'extrémité *terIS91* permettant de limiter la mobilisation des gènes adjacents.

Ainsi, pour mieux comprendre la fonctionnalité de Orf121, il conviendrait d'étudier le motif CxxC de Orf121 et le leucine zipper potentiel par des mutations ponctuelles. Il est également important de réitérer des réplicats des données préliminaires et d'améliorer la détection par Western-immunoblot en condition non dénaturante de Orf121. De plus, des expériences de retard sur gel permettraient de confirmer la fixation ou non de la transposase TnpA sur le brin bottom ou top. On pourrait aussi envisager d'étudier ainsi la possibilité de Orf121 à se fixer aux extrémités *oriIS91* et *terIS91*.

Dans **un troisième temps**, nous avons identifié et caractérisé deux doigts de zinc chez l'élément IS91. Le premier doigt de zinc que nous avons appelé ZF1 est impliqué dans la structure de la protéine et a fait intervenir les cystéines 41, 68, 73 et 76. Le second doigt de

zinc que nous avons appelé ZF2 serait impliqué dans l'interaction avec l'ADN et ferait intervenir les cystéines 53, 58, 360 et 363.

Actuellement, le double mutant des cystéines 360 et 363 en sérine est en cours de construction et sera analysé dès que possible. Cependant et au regard des résultats déjà obtenus relatifs aux mutations simples de chacune de ces cystéines, nous pouvons légitimement supposer qu'aucune transposition n'interviendra. In fine, cela nous permet de suggérer que ces deux résidus ont un rôle crucial dans la transposition de l'élément IS91.

Dans **un quatrième temps**, nous avons mis en évidence que les éléments IS*Kpn22*, IS*CR1* et IS*CR2* ne sont pas actifs pour la transposition. Toutefois, nous avons démontré que la double mutation ponctuelle (H1H, YLCRY) chez IS*CR1* permet de restaurer une activité de transposition. Chez IS*Kpn22* en position 54 d'une leucine vers une cystéine, la mutation ponctuelle permet également de restaurer une activité de transposition. De plus, nous avons démontré que les transposases TnpA<sub>IS91</sub>, TnpA<sub>ISCR1 H1H Y2</sub> et TnpA<sub>ISKpn22 L54C</sub> étaient capables de reconnaître les *oriS* de IS91, IS*CR1* et IS*CR2* avec un tétranucléotide cible 5'CTTG et l'extrémité *terS* de IS91 et IS*CR2*. Bien qu'au cours de cette étude nous ayons mis en évidence de nombreux éléments importants dans la mobilité des éléments IS91, IS*CR1*, IS*CR2* et IS*Kpn22*, de nouvelles interrogations ont été soulevées. Actuellement, les plasmides exprimant TnpA<sub>ISCR1 H1H</sub>, TnpA<sub>ISCR2 Y2</sub>, et TnpA<sub>ISCR2 H1H Y2</sub> sont en cours de construction et les analyses *in vivo* seront réalisées prochainement. De plus, pour les plasmides dont l'activité a déjà été testée, les analyses moléculaires des insertions sont en cours, à savoir : TnpA<sub>IS91</sub>, TnpA<sub>ISKpn22 L54C</sub> et TnpA<sub>ISCR1 H1H Y2</sub>. Il serait également intéressant d'étendre cette étude aux trois autres éléments connus pour être actifs dans la transposition (IS801, IS1294 et IS1294b).

L'analyse approfondie de cette mobilisation inter-familiale met en lumière l'importance cruciale de la surveillance continue de ces éléments dans les contextes cliniques, vétérinaires et environnementaux, afin de mieux comprendre et contrôler la propagation de la résistance aux antibiotiques.

## Références bibliographiques

---

- Abajy, M.Y., Kopeć, J., Schiwon, K., Burzynski, M., Döring, M., Bohn, C., Grohmann, E., 2007. A type IV-secretion-like system is required for conjugative DNA transport of broad-host-range plasmid pIP501 in gram-positive bacteria. *J. Bacteriol.* 189, 2487–2496. <https://doi.org/10.1128/JB.01491-06>
- Abrahams, J.P., Buchanan, S.K., Van Raaij, M.J., Fearnley, I.M., Leslie, A.G., Walker, J.E., 1996. The structure of bovine F1-ATPase complexed with the peptide antibiotic efrapeptin. *Proc. Natl. Acad. Sci. U. S. A.* 93, 9420–9424.
- Altae-Tran, H., Kannan, S., Demircioglu, F.E., Oshiro, R., Nety, S.P., McKay, L.J., Dlakić, M., Inskeep, W.P., Makarova, K.S., Macrae, R.K., Koonin, E.V., Zhang, F., 2021. The widespread IS200/IS605 transposon family encodes diverse programmable RNA-guided endonucleases. *Science* 374, 57–65. <https://doi.org/10.1126/science.abj6856>
- Alvarez-Martinez, C.E., Christie, P.J., 2009. Biological diversity of prokaryotic type IV secretion systems. *Microbiol. Mol. Biol. Rev.* MMBR 73, 775–808. <https://doi.org/10.1128/MMBR.00023-09>
- Amin, M.R., Yurovsky, A., Chen, Y., Skiena, S., Fitcher, B., 2018. Re-annotation of 12,495 prokaryotic 16S rRNA 3' ends and analysis of Shine-Dalgarno and anti-Shine-Dalgarno sequences. *PLoS One* 13, e0202767. <https://doi.org/10.1371/journal.pone.0202767>
- Aminov, R.I., 2011. Horizontal gene exchange in environmental microbiota. *Front. Microbiol.* 2, 158. <https://doi.org/10.3389/fmicb.2011.00158>
- Aminov, R.I., 2010. A brief history of the antibiotic era: lessons learned and challenges for the future. *Front. Microbiol.* 1, 134. <https://doi.org/10.3389/fmicb.2010.00134>
- Anani, H., Zgheib, R., Hasni, I., Raouf, D., Fournier, P.-E., 2020. Interest of bacterial pangenome analyses in clinical microbiology. *Microb. Pathog.* 149, 104275. <https://doi.org/10.1016/j.micpath.2020.104275>
- Andersson, J.O., Andersson, S.G., 1999. Insights into the evolutionary process of genome degradation. *Curr. Opin. Genet. Dev.* 9, 664–671. [https://doi.org/10.1016/s0959-437x\(99\)00024-6](https://doi.org/10.1016/s0959-437x(99)00024-6)
- Andersson, S.G., Kurland, C.G., 1998. Reductive evolution of resident genomes. *Trends Microbiol.* 6, 263–268. [https://doi.org/10.1016/s0966-842x\(98\)01312-2](https://doi.org/10.1016/s0966-842x(98)01312-2)
- Arduino, S.M., Catalano, M., Orman, B.E., Roy, P.H., Centrón, D., 2003. Molecular epidemiology of orf513-bearing class 1 integrons in multiresistant clinical isolates from Argentinean hospitals. *Antimicrob. Agents Chemother.* 47, 3945–3949. <https://doi.org/10.1128/AAC.47.12.3945-3949.2003>
- Arduino, S.M., Roy, P.H., Jacoby, G.A., Orman, B.E., Pineiro, S.A., Centron, D., 2002. blaCTX-M-2 is located in an unusual class 1 integron (In35) which includes Orf513. *Antimicrob. Agents Chemother.* 46, 2303–2306. <https://doi.org/10.1128/AAC.46.7.2303-2306.2002>
- Arends, K., Celik, E.-K., Probst, I., Goessweiner-Mohr, N., Fercher, C., Grumet, L., Soellue, C., Abajy, M.Y., Sakinc, T., Broszat, M., Schiwon, K., Koraimann, G., Keller, W., Grohmann, E., 2013. TraG encoded by the pIP501 type IV secretion system is a two-domain peptidoglycan-degrading enzyme essential for conjugative transfer. *J. Bacteriol.* 195, 4436–4444. <https://doi.org/10.1128/JB.02263-12>
- Arini, A., Keller, M.P., Arber, W., 1997. An antisense RNA in IS30 regulates the translational expression of the transposase. *Biol. Chem.* 378, 1421–1431. <https://doi.org/10.1515/bchm.1997.378.12.1421>
- Ariyoshi, M., Vassilyev, D.G., Iwasaki, H., Nakamura, H., Shinagawa, H., Morikawa, K., 1994. Atomic structure of the RuvC resolvase: a holliday junction-specific endonuclease from *E. coli*. *Cell* 78, 1063–1072. [https://doi.org/10.1016/0092-8674\(94\)90280-1](https://doi.org/10.1016/0092-8674(94)90280-1)
- Arnold, B.J., Huang, I.-T., Hanage, W.P., 2022. Horizontal gene transfer and adaptive evolution in bacteria. *Nat. Rev. Microbiol.* 20, 206–218. <https://doi.org/10.1038/s41579-021-00650-4>
- Asadulghani, M., Ogura, Y., Ooka, T., Itoh, T., Sawaguchi, A., Iguchi, A., Nakayama, K., Hayashi, T., 2009. The defective prophage pool of *Escherichia coli* O157: prophage-prophage interactions potentiate horizontal transfer of virulence determinants. *PLoS Pathog.* 5, e1000408. <https://doi.org/10.1371/journal.ppat.1000408>
- Bancroft, I., Wolk, C.P., 1989. Characterization of an insertion sequence (IS891) of novel structure from the cyanobacterium *Anabaena* sp. strain M-131. *J. Bacteriol.* 171, 5949–5954. <https://doi.org/10.1128/jb.171.11.5949-5954.1989>
- Bao, W., Jurka, J., 2013. Homologues of bacterial TnpB\_IS605 are widespread in diverse eukaryotic transposable



elements. *Mob. DNA* 4, 12. <https://doi.org/10.1186/1759-8753-4-12>

Barabas, O., Ronning, D.R., Guynet, C., Hickman, A.B., Ton-Hoang, B., Chandler, M., Dyda, F., 2008. Mechanism of IS200/IS605 family DNA transposases: activation and transposon-directed target site selection. *Cell* 132, 208–220. <https://doi.org/10.1016/j.cell.2007.12.029>

Baranov, P.V., Fayet, O., Hendrix, R.W., Atkins, J.F., 2006. Recoding in bacteriophages and bacterial IS elements. *Trends Genet.* 22, 174–181. <https://doi.org/10.1016/j.tig.2006.01.005>

Baranov, P.V., Gurvich, O.L., Fayet, O., Prère, M.F., Miller, W.A., Gesteland, R.F., Atkins, J.F., Giddings, M.C., 2001. RECODE: a database of frameshifting, bypassing and codon redefinition utilized for gene expression. *Nucleic Acids Res.* 29, 264–267. <https://doi.org/10.1093/nar/29.1.264>

Baranov, P.V., Hammer, A.W., Zhou, J., Gesteland, R.F., Atkins, J.F., 2005. Transcriptional slippage in bacteria: distribution in sequenced genomes and utilization in IS element gene expression. *Genome Biol.* 6, R25. <https://doi.org/10.1186/gb-2005-6-3-r25>

Bardaji, L., Añorga, M., Jackson, R.W., Martínez-Bilbao, A., Yanguas-Casás, N., Murillo, J., 2011. Miniature transposable sequences are frequently mobilized in the bacterial plant pathogen *Pseudomonas syringae* pv. *phaseolicola*. *PLoS One* 6, e25773. <https://doi.org/10.1371/journal.pone.0025773>

Barka, E.A., Vatsa, P., Sanchez, L., Gaveau-Vaillant, N., Jacquard, C., Meier-Kolthoff, J.P., Klenk, H.-P., Clément, C., Ouhdouch, Y., van Wezel, G.P., 2016. Taxonomy, Physiology, and Natural Products of Actinobacteria. *Microbiol. Mol. Biol. Rev.* MMBR 80, 1–43. <https://doi.org/10.1128/MMBR.00019-15>

Bean, E.L., McLellan, L.K., Grossman, A.D., 2022. Activation of the integrative and conjugative element Tn916 causes growth arrest and death of host bacteria. *PLoS Genet.* 18, e1010467. <https://doi.org/10.1371/journal.pgen.1010467>

Bedzyk, L.A., Shoemaker, N.B., Young, K.E., Salyers, A.A., 1992. Insertion and excision of *Bacteroides* conjugative chromosomal elements. *J. Bacteriol.* 174, 166–172. <https://doi.org/10.1128/jb.174.1.166-172.1992>

Bennett, P.M., 2008. Plasmid encoded antibiotic resistance: acquisition and transfer of antibiotic resistance genes in bacteria. *Br. J. Pharmacol.* 153 Suppl 1, S347-357. <https://doi.org/10.1038/sj.bjp.0707607>

Berg, D.E., Howe, M.M., Ajioka, J.W., 1989. *Mobile DNA: American Society for Microbiology Washington.*

Berg, J.M., Godwin, H.A., 1997. Lessons from zinc-binding peptides. *Annu. Rev. Biophys. Biomol. Struct.* 26, 357–371. <https://doi.org/10.1146/annurev.biophys.26.1.357>

Berg, J.M., Tymoczko, J.L., Gatto Jr., G.J., Stryer, L., 2015. *Biochemistry 8th edition, New York: W H Freeman.*

Berger, N., Heller, A.E., Störmann, K.D., Pfaff, E., 2001. Characterization of chimeric enzymes between caprine arthritis–encephalitis virus, maedi–visna virus and human immunodeficiency virus type 1 integrases expressed in *Escherichia coli*. *J. Gen. Virol.* 82, 139–148. <https://doi.org/10.1099/0022-1317-82-1-139>

Berleman, J., Auer, M., 2013. The role of bacterial outer membrane vesicles for intra- and interspecies delivery. *Environ. Microbiol.* 15, 347–354. <https://doi.org/10.1111/1462-2920.12048>

Bernales, I., 1997. *TransposicioAn por cõArculo rodante de IS91 (PhD Thesis).* PhD Thesis, Universidad de Cantabria, Spain.

Bernales, I., Mendiola, M.V., de la Cruz, F., 1999. Intramolecular transposition of insertion sequence IS91 results in second-site simple insertions. *Mol. Microbiol.* 33, 223–234. <https://doi.org/10.1046/j.1365-2958.1999.01432.x>

Beuzón, C.R., Casadesús, J., 1997. Conserved structure of IS200 elements in *Salmonella*. *Nucleic Acids Res.* 25, 1355–1361. <https://doi.org/10.1093/nar/25.7.1355>

Beuzón, C.R., Chessa, D., Casadesús, J., 2004. IS200: an old and still bacterial transposon. *Int. Microbiol. Off. J. Span. Soc. Microbiol.* 7, 3–12.

Beuzón, C.R., Marqués, S., Casadesús, J., 1999. Repression of IS200 transposase synthesis by RNA secondary structures. *Nucleic Acids Res.* 27, 3690–3695. <https://doi.org/10.1093/nar/27.18.3690>

Beyrouthy, R., Robin, F., Hamze, M., Bonnet, R., 2017. IncFIIk plasmid harbouring an amplification of 16S rRNA methyltransferase-encoding gene *rmtH* associated with mobile element ISCR2. *J. Antimicrob. Chemother.* 72, 402–406. <https://doi.org/10.1093/jac/dkw435>

Bichsel, M., Barbour, A.D., Wagner, A., 2013. Estimating the fitness effect of an insertion sequence. *J. Math. Biol.* 66, 95–114. <https://doi.org/10.1007/s00285-012-0504-2>

- Bickhart, D.M., Gogarten, J.P., Lapierre, P., Tisa, L.S., Normand, P., Benson, D.R., 2009. Insertion sequence content reflects genome plasticity in strains of the root nodule actinobacterium *Frankia*. *BMC Genomics* 10, 468. <https://doi.org/10.1186/1471-2164-10-468>
- Biet, F., Cenatiempo, Y., Fremaux, C., 2002. Identification of a replicon from pTXL1, a small cryptic plasmid from *Leuconostoc mesenteroides* subsp. *mesenteroides* Y110, and development of a food-grade vector. *Appl. Environ. Microbiol.* 68, 6451–6456. <https://doi.org/10.1128/AEM.68.12.6451-6456.2002>
- Bikard, D., Loot, C., Baharoglu, Z., Mazel, D., 2010. Folded DNA in action: hairpin formation and biological functions in prokaryotes. *Microbiol. Mol. Biol. Rev. MMBR* 74, 570–588. <https://doi.org/10.1128/MMBR.00026-10>
- Billard-Pomares, T., Fouteau, S., Jacquet, M.E., Roche, D., Barbe, V., Castellanos, M., Bouet, J.Y., Cruveiller, S., Médigue, C., Blanco, J., Clermont, O., Denamur, E., Branger, C., 2014. Characterization of a P1-like bacteriophage carrying an SHV-2 extended-spectrum  $\beta$ -lactamase from an *Escherichia coli* strain. *Antimicrob. Agents Chemother.* 58, 6550–6557. <https://doi.org/10.1128/AAC.03183-14>
- Biller, S.J., Schubotz, F., Roggensack, S.E., Thompson, A.W., Summons, R.E., Chisholm, S.W., 2014. Bacterial vesicles in marine ecosystems. *Science* 343, 183–186. <https://doi.org/10.1126/science.1243457>
- Binet, R., Maurelli, A.T., 2005. Frequency of spontaneous mutations that confer antibiotic resistance in *Chlamydia* spp. *Antimicrob. Agents Chemother.* 49, 2865–2873. <https://doi.org/10.1128/AAC.49.7.2865-2873.2005>
- Bitto, N.J., Chapman, R., Pidot, S., Costin, A., Lo, C., Choi, J., D’Cruze, T., Reynolds, E.C., Dashper, S.G., Turnbull, L., Whitchurch, C.B., Stinear, T.P., Stacey, K.J., Ferrero, R.L., 2017. Bacterial membrane vesicles transport their DNA cargo into host cells. *Sci. Rep.* 7, 7072. <https://doi.org/10.1038/s41598-017-07288-4>
- Blattner, F.R., Plunkett, G., Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B., Shao, Y., 1997. The complete genome sequence of *Escherichia coli* K-12. *Science* 277, 1453–1462. <https://doi.org/10.1126/science.277.5331.1453>
- Blokesch, M., 2017. In and out-contribution of natural transformation to the shuffling of large genomic regions. *Curr. Opin. Microbiol.* 38, 22–29. <https://doi.org/10.1016/j.mib.2017.04.001>
- Blot, M., 1994. Transposable elements and adaptation of host bacteria. *Genetica* 93, 5–12. <https://doi.org/10.1007/BF01435235>
- Boer, D.R., Ruíz-Masó, J.A., López-Blanco, J.R., Blanco, A.G., Vives-Llàcer, M., Chacón, P., Usón, I., Gomis-Rüth, F.X., Espinosa, M., Llorca, O., del Solar, G., Coll, M., 2009. Plasmid replication initiator RepB forms a hexamer reminiscent of ring helicases and has mobile nuclease domains. *EMBO J.* 28, 1666–1678. <https://doi.org/10.1038/emboj.2009.125>
- Boer, R., Russi, S., Guasch, A., Lucas, M., Blanco, A.G., Pérez-Luque, R., Coll, M., de la Cruz, F., 2006. Unveiling the molecular mechanism of a conjugative relaxase: The structure of TrwC complexed with a 27-mer DNA comprising the recognition hairpin and the cleavage site. *J. Mol. Biol.* 358, 857–869. <https://doi.org/10.1016/j.jmb.2006.02.018>
- Bondy-Denomy, J., Davidson, A.R., n.d. When a virus is not a parasite: the beneficial effects of prophages on bacterial fitness - PubMed [WWW Document]. URL <https://pubmed.ncbi.nlm.nih.gov/24585054/> (accessed 10.11.23).
- Bongers, R.S., Hoefnagel, M.H.N., Starrenburg, M.J.C., Siemerink, M.A.J., Arends, J.G.A., Hugenholtz, J., Kleerebezem, M., 2003. IS981-mediated adaptive evolution recovers lactate production by *ldhB* transcription activation in a lactate dehydrogenase-deficient strain of *Lactococcus lactis*. *J. Bacteriol.* 185, 4499–4507. <https://doi.org/10.1128/JB.185.15.4499-4507.2003>
- Bonnington, K.E., Kuehn, M.J., 2014. Protein selection and export via outer membrane vesicles. *Biochim. Biophys. Acta* 1843, 1612–1619. <https://doi.org/10.1016/j.bbamcr.2013.12.011>
- Boocock, M.R., Rice, P.A., 2013. A proposed mechanism for IS607-family serine transposases. *Mob. DNA* 4, 24. <https://doi.org/10.1186/1759-8753-4-24>
- Bordenstein, S.R., Reznikoff, W.S., 2005. Mobile DNA in obligate intracellular bacteria. *Nat. Rev. Microbiol.* 3, 688–699. <https://doi.org/10.1038/nrmicro1233>
- Boutoille, D., Corvec, S., Caroff, N., Giraudeau, C., Espaze, E., Caillon, J., Plésiat, P., Reynaud, A., 2004. Detection of an IS21 insertion sequence in the *mexR* gene of *Pseudomonas aeruginosa* increasing beta-lactam resistance. *FEMS Microbiol. Lett.* 230, 143–146. [https://doi.org/10.1016/S0378-1097\(03\)00882-6](https://doi.org/10.1016/S0378-1097(03)00882-6)

- Bowler, L.D., Zhang, Q.Y., Riou, J.Y., Spratt, B.G., 1994. Interspecies recombination between the penA genes of *Neisseria meningitidis* and commensal *Neisseria* species during the emergence of penicillin resistance in *N. meningitidis*: natural events and laboratory simulation. *J. Bacteriol.* 176, 333–337. <https://doi.org/10.1128/jb.176.2.333-337.1994>
- Boyd, D., Cloeckaert, A., Chaslus-Dancla, E., Mulvey, M.R., 2002. Characterization of variant *Salmonella* genomic island 1 multidrug resistance regions from serovars Typhimurium DT104 and Agona. *Antimicrob. Agents Chemother.* 46, 1714–1722. <https://doi.org/10.1128/AAC.46.6.1714-1722.2002>
- Boyd, D., Peters, G.A., Cloeckaert, A., Boumedine, K.S., Chaslus-Dancla, E., Imberechts, H., Mulvey, M.R., 2001. Complete nucleotide sequence of a 43-kilobase genomic island associated with the multidrug resistance region of *Salmonella enterica* serovar Typhimurium DT104 and its identification in phage type DT120 and serovar Agona. *J. Bacteriol.* 183, 5725–5732. <https://doi.org/10.1128/JB.183.19.5725-5732.2001>
- Boyle, W.K., Hall, L.S., Armstrong, A.A., Dulebohn, D.P., Samuels, D.S., Gherardini, F.C., Bourret, T.J., 2020. Establishment of an in vitro RNA polymerase transcription system: a new tool to study transcriptional activation in *Borrelia burgdorferi*. *Sci. Rep.* 10, 8246. <https://doi.org/10.1038/s41598-020-65104-y>
- Brister, J.R., Muzyczka, N., 2000. Mechanism of Rep-mediated adeno-associated virus origin nicking. *J. Virol.* 74, 7762–7771. <https://doi.org/10.1128/jvi.74.17.7762-7771.2000>
- Brister, J.R., Muzyczka, N., 1999. Rep-mediated nicking of the adeno-associated virus origin requires two biochemical activities, DNA helicase activity and transesterification. *J. Virol.* 73, 9325–9336. <https://doi.org/10.1128/JVI.73.11.9325-9336.1999>
- Brochet, M., Da Cunha, V., Couvé, E., Rusniok, C., Trieu-Cuot, P., Glaser, P., 2009. Atypical association of DDE transposition with conjugation specifies a new family of mobile elements. *Mol. Microbiol.* 71, 948–959. <https://doi.org/10.1111/j.1365-2958.2008.06579.x>
- Brolund, A., Franzén, O., Melefors, O., Tegmark-Wisell, K., Sandegren, L., 2013. Plasmidome-analysis of ESBL-producing *Escherichia coli* using conventional typing and high-throughput sequencing. *PloS One* 8, e65793. <https://doi.org/10.1371/journal.pone.0065793>
- Brookfield, J.F.Y., 1989. Mobile DNA: edited by Douglas E. Berg and Martha M. Howe, American Society for Microbiology, 1989. \$75.00 (member)/\$95.00 (non member) (i + 972 pages) ISBN 1 55581 005 5. *Trends Biochem. Sci.* 14, 428–429. [https://doi.org/10.1016/0968-0004\(89\)90306-X](https://doi.org/10.1016/0968-0004(89)90306-X)
- Brown, L., Wolf, J.M., Prados-Rosales, R., Casadevall, A., 2015. Through the wall: extracellular vesicles in Gram-positive bacteria, mycobacteria and fungi. *Nat. Rev. Microbiol.* 13, 620–630. <https://doi.org/10.1038/nrmicro3480>
- Brown, M.W., Kolisko, M., Silberman, J.D., Roger, A.J., 2012. Aggregative multicellularity evolved independently in the eukaryotic supergroup Rhizaria. *Curr. Biol. CB* 22, 1123–1127. <https://doi.org/10.1016/j.cub.2012.04.021>
- Browning, D.F., Busby, S.J., 2004. The regulation of bacterial transcription initiation. *Nat. Rev. Microbiol.* 2, 57–65. <https://doi.org/10.1038/nrmicro787>
- Bruand, C., Ehrlich, S.D., 2000. UvrD-dependent replication of rolling-circle plasmids in *Escherichia coli*. *Mol. Microbiol.* 35, 204–210. <https://doi.org/10.1046/j.1365-2958.2000.01700.x>
- Brügger, K., Redder, P., She, Q., Confalonieri, F., Zivanovic, Y., Garrett, R.A., 2002. Mobile elements in archaeal genomes. *FEMS Microbiol. Lett.* 206, 131–141. <https://doi.org/10.1111/j.1574-6968.2002.tb10999.x>
- Brunel, J., 1951. Antibiosis from Pasteur to Fleming. *J. Hist. Med. Allied Sci.* 6, 287–301. <https://doi.org/10.1093/jhmas/vi.summer.287>
- Buchner, J.M., Robertson, A.E., Poynter, D.J., Denniston, S.S., Karls, A.C., 2005. Piv site-specific invertase requires a DEDD motif analogous to the catalytic center of the RuvC Holliday junction resolvases. *J. Bacteriol.* 187, 3431–3437. <https://doi.org/10.1128/JB.187.10.3431-3437.2005>
- Burgess, R.R., Travers, A.A., Dunn, J.J., Bautz, E.K., 1969. Factor stimulating transcription by RNA polymerase. *Nature* 221, 43–46. <https://doi.org/10.1038/221043a0>
- Burland, V., Shao, Y., Perna, N.T., Plunkett, G., Sofia, H.J., Blattner, F.R., 1998. The complete DNA sequence and analysis of the large virulence plasmid of *Escherichia coli* O157:H7. *Nucleic Acids Res.* 26, 4196–4204. <https://doi.org/10.1093/nar/26.18.4196>
- Burrus, V., Pavlovic, G., Decaris, B., Guédon, G., 2002a. Conjugative transposons: the tip of the iceberg. *Mol. Microbiol.* 46, 601–610. <https://doi.org/10.1046/j.1365-2958.2002.03191.x>
- Burrus, V., Pavlovic, G., Decaris, B., Guédon, G., 2002b. The ICES<sub>t1</sub> 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. [https://doi.org/10.1016/s0147-619x\(02\)00102-6](https://doi.org/10.1016/s0147-619x(02)00102-6)

Burrus, V., Waldor, M.K., 2004. Shaping bacterial genomes with integrative and conjugative elements. *Res. Microbiol.* 155, 376–386. <https://doi.org/10.1016/j.resmic.2004.01.012>

Byrd, D.R., Matson, S.W., 1997. Nicking by transesterification: the reaction catalysed by a relaxase. *Mol. Microbiol.* 25, 1011–1022. <https://doi.org/10.1046/j.1365-2958.1997.5241885.x>

Cai, J., Wu, G., Jose, P.A., Zeng, C., 2016. Functional transferred DNA within extracellular vesicles. *Exp. Cell Res.* 349, 179–183. <https://doi.org/10.1016/j.yexcr.2016.10.012>

Calcutt, M.J., Lavrrar, J.L., Wise, K.S., 1999. IS1630 of *Mycoplasma fermentans*, a novel IS30-type insertion element that targets and duplicates inverted repeats of variable length and sequence during insertion. *J. Bacteriol.* 181, 7597–7607. <https://doi.org/10.1128/JB.181.24.7597-7607.1999>

Cam, K., Béjar, S., Gil, D., Bouché, J.P., 1988. Identification and sequence of gene *dicB*: translation of the division inhibitor from an in-phase internal start. *Nucleic Acids Res.* 16, 6327–6338. <https://doi.org/10.1093/nar/16.14.6327>

Campbell, A., Berg, D.E., Botstein, D., Lederberg, E.M., Novick, R.P., Starlinger, P., Szybalski, W., 1979. Nomenclature of transposable elements in prokaryotes. *Gene* 5, 197–206. [https://doi.org/10.1016/0378-1119\(79\)90078-7](https://doi.org/10.1016/0378-1119(79)90078-7)

Campbell, E.A., Muzzin, O., Chlenov, M., Sun, J.L., Olson, C.A., Weinman, O., Trester-Zedlitz, M.L., Darst, S.A., 2002. Structure of the bacterial RNA polymerase promoter specificity sigma subunit. *Mol. Cell* 9, 527–539. [https://doi.org/10.1016/s1097-2765\(02\)00470-7](https://doi.org/10.1016/s1097-2765(02)00470-7)

Canchaya, C., Proux, C., Fournous, G., Bruttin, A., Brüßow, H., 2003. Prophage genomics. *Microbiol. Mol. Biol. Rev.* MMBR 67, 238–276, table of contents. <https://doi.org/10.1128/MMBR.67.2.238-276.2003>

Carter, M.Q., Chen, J., Lory, S., 2010. The *Pseudomonas aeruginosa* pathogenicity island PAPI-1 is transferred via a novel type IV pilus. *J. Bacteriol.* 192, 3249–3258. <https://doi.org/10.1128/JB.00041-10>

Caruana, J.C., Walper, S.A., 2020. Bacterial Membrane Vesicles as Mediators of Microbe - Microbe and Microbe - Host Community Interactions. *Front. Microbiol.* 11, 432. <https://doi.org/10.3389/fmicb.2020.00432>

Casacuberta, E., González, J., 2013. The impact of transposable elements in environmental adaptation. *Mol. Ecol.* 22, 1503–1517. <https://doi.org/10.1111/mec.12170>

Casagrande Proietti, P., Bietta, A., Coletti, M., Marenzoni, M.L., Scorza, A.V., Passamonti, F., 2012. Insertion sequence IS256 in canine pyoderma isolates of *Staphylococcus pseudintermedius* associated with antibiotic resistance. *Vet. Microbiol.* 157, 376–382. <https://doi.org/10.1016/j.vetmic.2011.12.028>

Case, C.C., Roels, S.M., Jensen, P.D., Lee, J., Kleckner, N., Simons, R.W., 1989. The unusual stability of the IS10 anti-sense RNA is critical for its function and is determined by the structure of its stem-domain. *EMBO J.* 8, 4297–4305. <https://doi.org/10.1002/j.1460-2075.1989.tb08616.x>

Casjens, S., Hayden, M., 1988. Analysis in vivo of the bacteriophage P22 headful nuclease. *J. Mol. Biol.* 199, 467–474. [https://doi.org/10.1016/0022-2836\(88\)90618-3](https://doi.org/10.1016/0022-2836(88)90618-3)

Castro-Gutierrez, V., Fuller, E., Garcillán-Barcia, M.P., Helgason, T., Hassard, F., Moir, J., 2022. Dissemination of metaldehyde catabolic pathways is driven by mobile genetic elements in Proteobacteria. *Microb. Genomics* 8, mgen000881. <https://doi.org/10.1099/mgen.0.000881>

CDC, 2013. Antibiotic resistance threats in the United States, 2013 [WWW Document]. URL <https://stacks.cdc.gov/view/cdc/20705> (accessed 10.11.23).

Cerveau, N., Leclercq, S., Leroy, E., Bouchon, D., Cordaux, R., 2011. Short- and long-term evolutionary dynamics of bacterial insertion sequences: insights from *Wolbachia* endosymbionts. *Genome Biol. Evol.* 3, 1175–1186. <https://doi.org/10.1093/gbe/evr096>

Chalmers, R., Guhathakurta, A., Benjamin, H., Kleckner, N., 1998. IHF modulation of Tn10 transposition: sensory transduction of supercoiling status via a proposed protein/DNA molecular spring. *Cell* 93, 897–908. [https://doi.org/10.1016/s0092-8674\(00\)81449-x](https://doi.org/10.1016/s0092-8674(00)81449-x)

Chandler, M., de la Cruz, F., Dyda, F., Hickman, A.B., Moncalian, G., Ton-Hoang, B., 2013. Breaking and joining single-stranded DNA: the HUH endonuclease superfamily. *Nat. Rev. Microbiol.* 11, 525–538. <https://doi.org/10.1038/nrmicro3067>

Chandler, M., Fayet, O., 1993. Translational frameshifting in the control of transposition in bacteria. *Mol. Microbiol.*

7, 497–503. <https://doi.org/10.1111/j.1365-2958.1993.tb01140.x>

Chandler, M., Fayet, O., Rousseau, P., Ton Hoang, B., Duval-Valentin, G., 2015. Copy-out-Paste-in Transposition of IS911: A Major Transposition Pathway. *Microbiol. Spectr.* 3. <https://doi.org/10.1128/microbiolspec.MDNA3-0031-2014>

Chandler, M., Galas, D.J., 1983a. Cointegrate formation mediated by Tn9. II. Activity of IS1 is modulated by external DNA sequences. *J. Mol. Biol.* 170, 61–91. [https://doi.org/10.1016/s0022-2836\(83\)80227-7](https://doi.org/10.1016/s0022-2836(83)80227-7)

Chandler, M., Galas, D.J., 1983b. IS1-mediated tandem duplication of plasmid pBR322. Dependence on recA and on DNA polymerase I. *J. Mol. Biol.* 165, 183–190. [https://doi.org/10.1016/s0022-2836\(83\)80249-6](https://doi.org/10.1016/s0022-2836(83)80249-6)

Chandler, M., Mahillon, J., 2000. Insertion sequence nomenclature. *ASM News* 66, 324–324.

Chang, T.-L., Naqvi, A., Anand, S.P., Kramer, M.G., Munshi, R., Khan, S.A., 2002. Biochemical characterization of the *Staphylococcus aureus* PcrA helicase and its role in plasmid rolling circle replication. *J. Biol. Chem.* 277, 45880–45886. <https://doi.org/10.1074/jbc.M207383200>

Chao, L., McBroom, S.M., 1985. Evolution of transposable elements: an IS10 insertion increases fitness in *Escherichia coli*. *Mol. Biol. Evol.* 2, 359–369. <https://doi.org/10.1093/oxfordjournals.molbev.a040356>

Chao, L., Vargas, C., Spear, B.B., Cox, E.C., 1983. Transposable elements as mutator genes in evolution. *Nature* 303, 633–635. <https://doi.org/10.1038/303633a0>

Chater, K.F., Bruton, C.J., Foster, S.G., Tobek, I., 1985. Physical and genetic analysis of IS110, a transposable element of *Streptomyces coelicolor* A3(2). *Mol. Gen. Genet.* MGG 200, 235–239. <https://doi.org/10.1007/BF00425429>

Chelala, C.A., Margolin, P., 1976. Evidence that HT mutant strains of bacteriophage P22 retain an altered form of substrate specificity in the formation of transducing particles in *Salmonella typhimurium*. *Genet. Res.* 27, 315–322. <https://doi.org/10.1017/s0016672300016505>

Chen, J., Quiles-Puchalt, N., Chiang, Y.N., Bacigalupe, R., Fillol-Salom, A., Chee, M.S.J., Fitzgerald, J.R., Penadés, J.R., 2018. Genome hypermobility by lateral transduction. *Science* 362, 207–212. <https://doi.org/10.1126/science.aat5867>

Chen, W., Mandali, S., Hancock, S.P., Kumar, P., Collazo, M., Cascio, D., Johnson, R.C., 2018. Multiple serine transposase dimers assemble the transposon-end synaptic complex during IS607-family transposition. *eLife* 7, e39611. <https://doi.org/10.7554/eLife.39611>

Chen, Y., Braathen, P., Léonard, C., Mahillon, J., 1999. MIC231, a naturally occurring mobile insertion cassette from *Bacillus cereus*. *Mol. Microbiol.* 32, 657–668. <https://doi.org/10.1046/j.1365-2958.1999.01388.x>

Cheng, C., Sun, J., Zheng, F., Lu, W., Yang, Q., Rui, Y., 2016. New structures simultaneously harboring class 1 integron and ISCR1-linked resistance genes in multidrug-resistant Gram-negative bacteria. *BMC Microbiol.* 16, 71. <https://doi.org/10.1186/s12866-016-0683-x>

Cheng, Q., Paszkiet, B.J., Shoemaker, N.B., Gardner, J.F., Salyers, A.A., 2000. Integration and excision of a *Bacteroides* conjugative transposon, CTnDOT. *J. Bacteriol.* 182, 4035–4043. <https://doi.org/10.1128/JB.182.14.4035-4043.2000>

Chiang, Y.N., Penadés, J.R., Chen, J., 2019. Genetic transduction by phages and chromosomal islands: The new and noncanonical. *PLoS Pathog.* 15, e1007878. <https://doi.org/10.1371/journal.ppat.1007878>

Chiura, H.X., Kogure, K., Hagemann, S., Ellinger, A., Velimirov, B., 2011. Evidence for particle-induced horizontal gene transfer and serial transduction between bacteria. *FEMS Microbiol. Ecol.* 76, 576–591. <https://doi.org/10.1111/j.1574-6941.2011.01077.x>

Christie, P.J., Whitaker, N., González-Rivera, C., 2014. Mechanism and structure of the bacterial type IV secretion systems. *Biochim. Biophys. Acta* 1843, 1578–1591. <https://doi.org/10.1016/j.bbamcr.2013.12.019>

Clément, J.M., Wilde, C., Bachellier, S., Lambert, P., Hofnung, M., 1999. IS1397 is active for transposition into the chromosome of *Escherichia coli* K-12 and inserts specifically into palindromic units of bacterial interspersed mosaic elements. *J. Bacteriol.* 181, 6929–6936. <https://doi.org/10.1128/JB.181.22.6929-6936.1999>

Clérot, D., Bernardi, F., 2006. DNA helicase activity is associated with the replication initiator protein rep of tomato yellow leaf curl geminivirus. *J. Virol.* 80, 11322–11330. <https://doi.org/10.1128/JVI.00924-06>

Clewell, D.B., Flannagan, S.E., Jaworski, D.D., 1995. Unconstrained bacterial promiscuity: the Tn916-Tn1545 family of conjugative transposons. *Trends Microbiol.* 3, 229–236. [https://doi.org/10.1016/s0966-842x\(00\)88930-1](https://doi.org/10.1016/s0966-842x(00)88930-1)

- Clewell, D.B., Helinski, D.R., 1969. Supercoiled circular DNA-protein complex in *Escherichia coli*: purification and induced conversion to an open circular DNA form. *Proc. Natl. Acad. Sci. U. S. A.* 62, 1159–1166. <https://doi.org/10.1073/pnas.62.4.1159>
- Cloekaert, A., Baucheron, S., Chaslus-Dancla, E., 2001. Nonenzymatic chloramphenicol resistance mediated by IncC plasmid R55 is encoded by a floR gene variant. *Antimicrob. Agents Chemother.* 45, 2381–2382. <https://doi.org/10.1128/AAC.45.8.2381-2382.2001>
- Cloekaert, A., Baucheron, S., Flaujac, G., Schwarz, S., Kehrenberg, C., Martel, J.L., Chaslus-Dancla, E., 2000. Plasmid-mediated florfenicol resistance encoded by the floR gene in *Escherichia coli* isolated from cattle. *Antimicrob. Agents Chemother.* 44, 2858–2860. <https://doi.org/10.1128/AAC.44.10.2858-2860.2000>
- Cole, S.T., 2014. Who will develop new antibacterial agents? *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 369, 20130430. <https://doi.org/10.1098/rstb.2013.0430>
- Coleman, N.V., Richardson-Harris, J., Wilson, N.L., Holmes, A.J., 2014. Insertion sequence ISPst4 activates pUC plasmid replication in *Pseudomonas stutzeri*. *FEMS Microbiol. Lett.* 356, 242–249. <https://doi.org/10.1111/1574-6968.12417>
- Collis, C.M., Hall, R.M., 1995. Expression of antibiotic resistance genes in the integrated cassettes of integrons. *Antimicrob. Agents Chemother.* 39, 155–162. <https://doi.org/10.1128/AAC.39.1.155>
- Collis, C.M., Hall, R.M., 1992a. Gene cassettes from the insert region of integrons are excised as covalently closed circles. *Mol. Microbiol.* 6, 2875–2885. <https://doi.org/10.1111/j.1365-2958.1992.tb01467.x>
- Collis, C.M., Hall, R.M., 1992b. Site-specific deletion and rearrangement of integron insert genes catalyzed by the integron DNA integrase. *J. Bacteriol.* 174, 1574–1585. <https://doi.org/10.1128/jb.174.5.1574-1585.1992>
- Coluzzi, C., Garcillán-Barcia, M.P., de la Cruz, F., Rocha, E.P.C., 2022. Evolution of Plasmid Mobility: Origin and Fate of Conjugative and Nonconjugative Plasmids. *Mol. Biol. Evol.* 39, msac115. <https://doi.org/10.1093/molbev/msac115>
- Cookson, A.L., Noel, S., Hussein, H., Perry, R., Sang, C., Moon, C.D., Leahy, S.C., Altermann, E., Kelly, W.J., Attwood, G.T., 2011. Transposition of Tn916 in the four replicons of the *Butyrivibrio proteoclasticus* B316(T) genome. *FEMS Microbiol. Lett.* 316, 144–151. <https://doi.org/10.1111/j.1574-6968.2010.02204.x>
- Coros, A.M., Twiss, E., Tavakoli, N.P., Derbyshire, K.M., 2005. Genetic evidence that GTP is required for transposition of IS903 and Tn552 in *Escherichia coli*. *J. Bacteriol.* 187, 4598–4606. <https://doi.org/10.1128/JB.187.13.4598-4606.2005>
- Correia, F.F., Inouye, S., Inouye, M., 1988. A family of small repeated elements with some transposon-like properties in the genome of *Neisseria gonorrhoeae*. *J. Biol. Chem.* 263, 12194–12198.
- Correia, F.F., Inouye, S., Inouye, M., 1986. A 26-base-pair repetitive sequence specific for *Neisseria gonorrhoeae* and *Neisseria meningitidis* genomic DNA. *J. Bacteriol.* 167, 1009–1015. <https://doi.org/10.1128/jb.167.3.1009-1015.1986>
- Correia, S., Poeta, P., Hébraud, M., Capelo, J.L., Igrejas, G., 2017. Mechanisms of quinolone action and resistance: where do we stand? *J. Med. Microbiol.* 66, 551–559. <https://doi.org/10.1099/jmm.0.000475>
- Couto, I., Wu, S.W., Tomasz, A., de Lencastre, H., 2003. Development of methicillin resistance in clinical isolates of *Staphylococcus sciuri* by transcriptional activation of the mecA homologue native to s. *J. Bacteriol.* 185, 645–653. <https://doi.org/10.1128/JB.185.2.645-653.2003>
- Craig, N.L., Chandler, M., Gellert, M., Lambowitz, A., Rice, P.A., Sandmeyer, S. (Eds.), 2015. *Mobile DNA III*. ASM Press, Washington, DC.
- Craig, N.L., Craigie, R., Gellert, M., Lambowitz, A.M., 2002. *Mobile DNA II*, American Society for Microbiology. Wash. DC 1204.
- Craig, R.J., 2023. Replitrans: A major group of eukaryotic transposons encoding HUH endonuclease. *Proc. Natl. Acad. Sci. U. S. A.* 120, e2301424120. <https://doi.org/10.1073/pnas.2301424120>
- Curcio, M.J., Derbyshire, K.M., 2003. The outs and ins of transposition: from mu to kangaroo. *Nat. Rev. Mol. Cell Biol.* 4, 865–877. <https://doi.org/10.1038/nrm1241>
- Darby, E.M., Trampari, E., Siasat, P., Gaya, M.S., Alav, I., Webber, M.A., Blair, J.M.A., 2023. Molecular mechanisms of antibiotic resistance revisited. *Nat. Rev. Microbiol.* 21, 280–295. <https://doi.org/10.1038/s41579-022-00820-y>

- Datta, A.R., Rosner, J.L., 1987. Reduced transposition in rho mutants of *Escherichia coli* K-12. *J. Bacteriol.* 169, 888–890. <https://doi.org/10.1128/jb.169.2.888-890.1987>
- Datta, S., Larkin, C., Schildbach, J.F., 2003. Structural insights into single-stranded DNA binding and cleavage by F factor Tral. *Struct. Lond. Engl.* 1993 11, 1369–1379. <https://doi.org/10.1016/j.str.2003.10.001>
- Davidson, A.R., 2018. A common trick for transferring bacterial DNA. *Science* 362, 152–153. <https://doi.org/10.1126/science.aav1723>
- Davis, M.A., Simons, R.W., Kleckner, N., 1985. Tn10 protects itself at two levels from fortuitous activation by external promoters. *Cell* 43, 379–387. [https://doi.org/10.1016/0092-8674\(85\)90043-1](https://doi.org/10.1016/0092-8674(85)90043-1)
- De Gregorio, E., Silvestro, G., Venditti, R., Carlomagno, M.S., Di Nocera, P.P., 2006. Structural organization and functional properties of miniature DNA insertion sequences in yersiniae. *J. Bacteriol.* 188, 7876–7884. <https://doi.org/10.1128/JB.00942-06>
- de la Cruz, F., Frost, L.S., Meyer, R.J., Zechner, E.L., 2010. Conjugative DNA metabolism in Gram-negative bacteria. *FEMS Microbiol. Rev.* 34, 18–40. <https://doi.org/10.1111/j.1574-6976.2009.00195.x>
- De Palmenaer, D., Siguiet, P., Mahillon, J., 2008. IS4 family goes genomic. *BMC Evol. Biol.* 8, 18. <https://doi.org/10.1186/1471-2148-8-18>
- De Palmenaer, D., Vermeiren, C., Mahillon, J., 2004. IS231-MIC231 elements from *Bacillus cereus* sensu lato are modular. *Mol. Microbiol.* 53, 457–467. <https://doi.org/10.1111/j.1365-2958.2004.04146.x>
- Deatherage, B.L., Cookson, B.T., 2012. Membrane vesicle release in bacteria, eukaryotes, and archaea: a conserved yet underappreciated aspect of microbial life. *Infect. Immun.* 80, 1948–1957. <https://doi.org/10.1128/IAI.06014-11>
- Deceliere, G., Letrillard, Y., Charles, S., Biémont, C., 2006. TESD: a transposable element dynamics simulation environment. *Bioinforma. Oxf. Engl.* 22, 2702–2703. <https://doi.org/10.1093/bioinformatics/btl454>
- deHaseth, P.L., Zupancic, M.L., Record, M.T., 1998. RNA polymerase-promoter interactions: the comings and goings of RNA polymerase. *J. Bacteriol.* 180, 3019–3025. <https://doi.org/10.1128/JB.180.12.3019-3025.1998>
- del Pilar Garcillán-Barcia, M., Bernales, I., Mendiola, M.V., de la Cruz, F., 2001. Single-stranded DNA intermediates in IS91 rolling-circle transposition. *Mol. Microbiol.* 39, 494–501. <https://doi.org/10.1046/j.1365-2958.2001.02261.x>
- Delavat, F., Miyazaki, R., Carraro, N., Pradervand, N., van der Meer, J.R., 2017. The hidden life of integrative and conjugative elements. *FEMS Microbiol. Rev.* 41, 512–537. <https://doi.org/10.1093/femsre/fux008>
- Delihias, N., 2011. Impact of small repeat sequences on bacterial genome evolution. *Genome Biol. Evol.* 3, 959–973. <https://doi.org/10.1093/gbe/evr077>
- Delihias, N., 2008. Small mobile sequences in bacteria display diverse structure/function motifs. *Mol. Microbiol.* 67, 475–481. <https://doi.org/10.1111/j.1365-2958.2007.06068.x>
- Derbyshire, K.M., Grindley, N.D., 1992. Binding of the IS903 transposase to its inverted repeat in vitro. *EMBO J.* 11, 3449–3455. <https://doi.org/10.1002/j.1460-2075.1992.tb05424.x>
- Derbyshire, K.M., Kramer, M., Grindley, N.D., 1990. Role of instability in the cis action of the insertion sequence IS903 transposase. *Proc. Natl. Acad. Sci. U. S. A.* 87, 4048–4052. <https://doi.org/10.1073/pnas.87.11.4048>
- Diaz-Aroca, E., de la Cruz, F., Zabala, J.C., Ortiz, J.M., 1984. Characterization of the new insertion sequence IS91 from an alpha-hemolysin plasmid of *Escherichia coli*. *Mol. Gen. Genet. MGG* 193, 493–499. <https://doi.org/10.1007/BF00382089>
- Diaz-Aroca, E., Mendiola, M.V., Zabala, J.C., de la Cruz, F., 1987. Transposition of IS91 does not generate a target duplication. *J. Bacteriol.* 169, 442–443. <https://doi.org/10.1128/jb.169.1.442-443.1987>
- Dodson, K.W., Berg, D.E., 1989. Factors affecting transposition activity of IS50 and Tn5 ends. *Gene* 76, 207–213. [https://doi.org/10.1016/0378-1119\(89\)90161-3](https://doi.org/10.1016/0378-1119(89)90161-3)
- Dolgin, E.S., Charlesworth, B., 2006. The fate of transposable elements in asexual populations. *Genetics* 174, 817–827. <https://doi.org/10.1534/genetics.106.060434>
- Domingues, S., Nielsen, K.M., 2017. Membrane vesicles and horizontal gene transfer in prokaryotes. *Curr. Opin. Microbiol.* 38, 16–21. <https://doi.org/10.1016/j.mib.2017.03.012>
- Doolittle, W.F., Sapienza, C., 1980. Selfish genes, the phenotype paradigm and genome evolution. *Nature* 284, 601–603. <https://doi.org/10.1038/284601a0>

- Dorward, D.W., Garon, C.F., Judd, R.C., 1989. Export and intercellular transfer of DNA via membrane blebs of *Neisseria gonorrhoeae*. *J. Bacteriol.* 171, 2499–2505. <https://doi.org/10.1128/jb.171.5.2499-2505.1989>
- Doutriaux, M.P., Couteau, F., Bergounioux, C., White, C., 1998. Isolation and characterisation of the RAD51 and DMC1 homologs from *Arabidopsis thaliana*. *Mol. Gen. Genet.* MGG 257, 283–291. <https://doi.org/10.1007/s004380050649>
- Duckett, D.R., Murchie, A.I., Diekmann, S., von Kitzing, E., Kemper, B., Lilley, D.M., 1988. The structure of the Holliday junction, and its resolution. *Cell* 55, 79–89. [https://doi.org/10.1016/0092-8674\(88\)90011-6](https://doi.org/10.1016/0092-8674(88)90011-6)
- Duval-Valentin, G., Normand, C., Khemici, V., Marty, B., Chandler, M., 2001. Transient promoter formation: a new feedback mechanism for regulation of IS911 transposition. *EMBO J.* 20, 5802–5811. <https://doi.org/10.1093/emboj/20.20.5802>
- Dyall-Smith, M.L., Pfeiffer, F., Klee, K., Palm, P., Gross, K., Schuster, S.C., Rampp, M., Oesterhelt, D., 2011. *Haloquadratum walsbyi*: limited diversity in a global pond. *PloS One* 6, e20968. <https://doi.org/10.1371/journal.pone.0020968>
- Dyda, F., Chandler, M., Hickman, A.B., 2012. The emerging diversity of transpososome architectures. *Q. Rev. Biophys.* 45, 493–521. <https://doi.org/10.1017/S0033583512000145>
- Dyda, F., Hickman, A.B., 2003. A mob of reps. *Struct. Lond. Engl.* 1993 11, 1310–1311. <https://doi.org/10.1016/j.str.2003.10.010>
- Džidić, S., Šušković, J., Kos, B., 2008. Antibiotic Resistance Mechanisms in Bacteria: Biochemical and Genetic Aspects: Mehanizmi antibiotičke otpornosti bakterija: biokemijski i genetički aspekti. *Meh. Antibiot. Otpornosti Bakt. Biokem. Genet. Asp.* 46, 11–21.
- Edwards, J.S., Betts, L., Frazier, M.L., Pollet, R.M., Kwong, S.M., Walton, W.G., Ballentine, W.K., Huang, J.J., Habibi, S., Del Campo, M., Meier, J.L., Dervan, P.B., Firth, N., Redinbo, M.R., 2013. Molecular basis of antibiotic multiresistance transfer in *Staphylococcus aureus*. *Proc. Natl. Acad. Sci. U. S. A.* 110, 2804–2809. <https://doi.org/10.1073/pnas.1219701110>
- Eichenbaum, Z., Livneh, Z., 1998. UV light induces IS10 transposition in *Escherichia coli*. *Genetics* 149, 1173–1181. <https://doi.org/10.1093/genetics/149.3.1173>
- El Gharniti, F., Dols-Lafargue, M., Bon, E., Claisse, O., Miot-Sertier, C., Lonvaud, A., Le Marrec, C., 2012. IS30 elements are mediators of genetic diversity in *Oenococcus oeni*. *Int. J. Food Microbiol.* 158, 14–22. <https://doi.org/10.1016/j.ijfoodmicro.2012.06.009>
- Elena, S.F., Ekunwe, L., Hajela, N., Oden, S.A., Lenski, R.E., 1998. Distribution of fitness effects caused by random insertion mutations in *Escherichia coli*. *Genetica* 102–103, 349–358.
- Ellabaan, M.M.H., Munck, C., Porse, A., Imamovic, L., Sommer, M.O.A., 2021. Forecasting the dissemination of antibiotic resistance genes across bacterial genomes. *Nat. Commun.* 12, 2435. <https://doi.org/10.1038/s41467-021-22757-1>
- Ellis, M.J., Trussler, R.S., Haniford, D.B., 2015a. A cis-encoded sRNA, Hfq and mRNA secondary structure act independently to suppress IS200 transposition. *Nucleic Acids Res.* 43, 6511–6527. <https://doi.org/10.1093/nar/gkv584>
- Ellis, M.J., Trussler, R.S., Haniford, D.B., 2015b. Hfq binds directly to the ribosome-binding site of IS10 transposase mRNA to inhibit translation. *Mol. Microbiol.* 96, 633–650. <https://doi.org/10.1111/mmi.12961>
- Emmerich, Rudolf & Löw, Oscar, 1889. Bakteriologische Enzyme als Ursache der erworbenen Immunität und die Heilung von Infektionskrankheiten durch dieselben (pp.1-65). by Emmerich, Rudolf & Löw, Oscar: (1889) | Antiq. F.-D. Söhn - Medicusbooks.Com [WWW Document]. URL <https://www.abebooks.com/Bakteriologische-Enzyme-Ursache-erworbenen-Immunit%C3%A4t-Heilung/630192765/bd> (accessed 10.11.23).
- Epshtein, V., Nudler, E., 2003. Cooperation between RNA polymerase molecules in transcription elongation. *Science* 300, 801–805. <https://doi.org/10.1126/science.1083219>
- Escoubas, J.M., Prère, M.F., Fayet, O., Salvignol, I., Galas, D., Zerbib, D., Chandler, M., 1991. Translational control of transposition activity of the bacterial insertion sequence IS1. *EMBO J.* 10, 705–712. <https://doi.org/10.1002/j.1460-2075.1991.tb08000.x>
- Espinosa, M., del Solar, G., Rojo, F., Alonso, J.C., 1995. Plasmid rolling circle replication and its control. *FEMS Microbiol. Lett.* 130, 111–120. <https://doi.org/10.1111/j.1574-6968.1995.tb07707.x>
- Fayet, O., Ramond, P., Polard, P., Prère, M.F., Chandler, M., 1990. Functional similarities between retroviruses



- and the IS3 family of bacterial insertion sequences? *Mol. Microbiol.* 4, 1771–1777. <https://doi.org/10.1111/j.1365-2958.1990.tb00555.x>
- Feng, X., Colloms, S.D., 2007. In vitro transposition of ISY100, a bacterial insertion sequence belonging to the Tc1/mariner family. *Mol. Microbiol.* 65, 1432–1443. <https://doi.org/10.1111/j.1365-2958.2007.05842.x>
- Filée, J., Siguier, P., Chandler, M., 2007. Insertion sequence diversity in archaea. *Microbiol. Mol. Biol. Rev. MMBR* 71, 121–157. <https://doi.org/10.1128/MMBR.00031-06>
- Filée, Jonathan, Siguier, P., Chandler, M., 2007. I am what I eat and I eat what I am: acquisition of bacterial genes by giant viruses. *Trends Genet. TIG* 23, 10–15. <https://doi.org/10.1016/j.tig.2006.11.002>
- Flannagan, S.E., Zitzow, L.A., Su, Y.A., Clewell, D.B., 1994. Nucleotide sequence of the 18-kb conjugative transposon Tn916 from *Enterococcus faecalis*. *Plasmid* 32, 350–354. <https://doi.org/10.1006/plas.1994.1077>
- Fleming, A., 1929. On the Antibacterial Action of Cultures of a Penicillium, with Special Reference to their Use in the Isolation of *B. influenzae*. *Br. J. Exp. Pathol.* 10, 226–236.
- Fogg, P.C.M., Colloms, S., Rosser, S., Stark, M., Smith, M.C.M., 2014. New applications for phage integrases. *J. Mol. Biol.* 426, 2703–2716. <https://doi.org/10.1016/j.jmb.2014.05.014>
- Forsberg, K.J., Reyes, A., Wang, B., Selleck, E.M., Sommer, M.O.A., Dantas, G., 2012. The shared antibiotic resistome of soil bacteria and human pathogens. *Science* 337, 1107–1111. <https://doi.org/10.1126/science.1220761>
- Foster, P.L., 2007. Stress-induced mutagenesis in bacteria. *Crit. Rev. Biochem. Mol. Biol.* 42, 373–397. <https://doi.org/10.1080/10409230701648494>
- Franke, A.E., Clewell, D.B., 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. *J. Bacteriol.* 145, 494–502. <https://doi.org/10.1128/jb.145.1.494-502.1981>
- Frost, L.S., Leplae, R., Summers, A.O., Toussaint, A., 2005. Mobile genetic elements: the agents of open source evolution. *Nat. Rev. Microbiol.* 3, 722–732. <https://doi.org/10.1038/nrmicro1235>
- Fuller, R.S., Funnell, B.E., Kornberg, A., 1984. The dnaA protein complex with the *E. coli* chromosomal replication origin (oriC) and other DNA sites. *Cell* 38, 889–900. [https://doi.org/10.1016/0092-8674\(84\)90284-8](https://doi.org/10.1016/0092-8674(84)90284-8)
- Fuller, R.S., Kornberg, A., 1983. Purified dnaA protein in initiation of replication at the *Escherichia coli* chromosomal origin of replication. *Proc. Natl. Acad. Sci. U. S. A.* 80, 5817–5821. <https://doi.org/10.1073/pnas.80.19.5817>
- Fulsundar, S., Harms, K., Flaten, G.E., Johnsen, P.J., Chopade, B.A., Nielsen, K.M., 2014. Gene transfer potential of outer membrane vesicles of *Acinetobacter baylyi* and effects of stress on vesiculation. *Appl. Environ. Microbiol.* 80, 3469–3483. <https://doi.org/10.1128/AEM.04248-13>
- Funnell, B.E., Baker, T.A., Kornberg, A., 1986. Complete enzymatic replication of plasmids containing the origin of the *Escherichia coli* chromosome. *J. Biol. Chem.* 261, 5616–5624.
- Funnell, B.E., Phillips, G.J., 2004. Plasmid Biology, in: *Plasmid Biology*. John Wiley & Sons, Ltd, pp. i–xi. <https://doi.org/10.1128/9781555817732.fmatter>
- Fürste, J.P., Pansegrau, W., Ziegelin, G., Kröger, M., Lanka, E., 1989. Conjugative transfer of promiscuous IncP plasmids: interaction of plasmid-encoded products with the transfer origin. *Proc. Natl. Acad. Sci. U. S. A.* 86, 1771–1775. <https://doi.org/10.1073/pnas.86.6.1771>
- Furuyama, N., Sircili, M.P., 2021. Outer Membrane Vesicles (OMVs) Produced by Gram-Negative Bacteria: Structure, Functions, Biogenesis, and Vaccine Application. *BioMed Res. Int.* 2021, 1490732. <https://doi.org/10.1155/2021/1490732>
- Gama, M.J., Toussaint, A., Pato, M.L., 1990. Instability of bacteriophage Mu transposase and the role of host Hfl protein. *Mol. Microbiol.* 4, 1891–1897. <https://doi.org/10.1111/j.1365-2958.1990.tb02038.x>
- Gamas, P., Galas, D., Chandler, M., 1985. DNA sequence at the end of IS1 required for transposition. *Nature* 317, 458–460. <https://doi.org/10.1038/317458a0>
- Garcillán-Barcia, M.P., Alvarado, A., de la Cruz, F., 2011. Identification of bacterial plasmids based on mobility and plasmid population biology. *FEMS Microbiol. Rev.* 35, 936–956. <https://doi.org/10.1111/j.1574-6976.2011.00291.x>
- Garcillán-Barcia, M.P., Bernales, I., Mendiola, M.V., de la Cruz, F., 2002. IS 91 Rolling-Circle Transposition, in: *Mobile DNA II*. John Wiley & Sons, Ltd, pp. 889–904. <https://doi.org/10.1128/9781555817954.ch37>

- Garcillán-Barcia, M.P., de la Cruz, F., 2002. Distribution of IS91 family insertion sequences in bacterial genomes: evolutionary implications. *FEMS Microbiol. Ecol.* 42, 303–313. <https://doi.org/10.1111/j.1574-6941.2002.tb01020.x>
- Garcillán-Barcia, M.P., Francia, M.V., de la Cruz, F., 2009. The diversity of conjugative relaxases and its application in plasmid classification. *FEMS Microbiol. Rev.* 33, 657–687. <https://doi.org/10.1111/j.1574-6976.2009.00168.x>
- Garcillán-Barcia, M.P., Jurado, P., González-Pérez, B., Moncalián, G., Fernández, L.A., de la Cruz, F., 2007. Conjugative transfer can be inhibited by blocking relaxase activity within recipient cells with intrabodies. *Mol. Microbiol.* 63, 404–416. <https://doi.org/10.1111/j.1365-2958.2006.05523.x>
- Garcillán-Barcia, M.P., Redondo-Salvo, S., de la Cruz, F., 2023. Plasmid classifications. *Plasmid* 126, 102684. <https://doi.org/10.1016/j.plasmid.2023.102684>
- Garnier, F., Raked, N., Gassama, A., Denis, F., Ploy, M.-C., 2006. Genetic Environment of Quinolone Resistance Gene *qnrB2* in a Complex *sul1*-Type Integron in the Newly Described *Salmonella enterica* Serovar Keurmassar. *Antimicrob. Agents Chemother.* 50, 3200–3202. <https://doi.org/10.1128/AAC.00293-06>
- Gelpi, A., Gilbertson, A., Tucker, J.D., 2015. Magic bullet: Paul Ehrlich, Salvarsan and the birth of venereology. *Sex. Transm. Infect.* 91, 68–69. <https://doi.org/10.1136/sextrans-2014-051779>
- Genilloud, O., Blázquez, J., Mazodier, P., Moreno, F., 1988. A clinical isolate of transposon Tn5 expressing streptomycin resistance in *Escherichia coli*. *J. Bacteriol.* 170, 1275–1278. <https://doi.org/10.1128/jb.170.3.1275-1278.1988>
- Ghaly, T.M., Gillings, M.R., 2018. Mobile DNAs as Ecologically and Evolutionarily Independent Units of Life. *Trends Microbiol.* 26, 904–912. <https://doi.org/10.1016/j.tim.2018.05.008>
- Gibson, M.K., Forsberg, K.J., Dantas, G., 2015. Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology. *ISME J.* 9, 207–216. <https://doi.org/10.1038/ismej.2014.106>
- Gil, R., Belda, E., Gosalbes, M.J., Delaye, L., Vallier, A., Vincent-Monégat, C., Heddi, A., Silva, F.J., Moya, A., Latorre, A., 2008. Massive presence of insertion sequences in the genome of SOPE, the primary endosymbiont of the rice weevil *Sitophilus oryzae*. *Int. Microbiol. Off. J. Span. Soc. Microbiol.* 11, 41–48.
- Gilbert, C., Cordaux, R., 2013. Horizontal transfer and evolution of prokaryote transposable elements in eukaryotes. *Genome Biol. Evol.* 5, 822–832. <https://doi.org/10.1093/gbe/evt057>
- Gillings, M.R., Labbate, M., Sajjad, A., Giguère, N.J., Holley, M.P., Stokes, H.W., 2009. Mobilization of a Tn402-like class 1 integron with a novel cassette array via flanking miniature inverted-repeat transposable element-like structures. *Appl. Environ. Microbiol.* 75, 6002–6004. <https://doi.org/10.1128/AEM.01033-09>
- Goh, S., Hussain, H., Chang, B.J., Riley, T.V., Mullany, P., 2013. Phage  $\phi$ C2 mediates transduction of Tn6215, encoding erythromycin resistance, between *Clostridium difficile* strains. *mBio* 4, e00840-00813. <https://doi.org/10.1128/mBio.00840-13>
- Goodwin, T.J.D., Butler, M.I., Poulter, R.T.M., 2003. Cryptons: a group of tyrosine-recombinase-encoding DNA transposons from pathogenic fungi. *Microbiol. Read. Engl.* 149, 3099–3109. <https://doi.org/10.1099/mic.0.26529-0>
- Goosen, N., van de Putte, P., 1986. Role of *ner* protein in bacteriophage Mu transposition. *J. Bacteriol.* 167, 503–507. <https://doi.org/10.1128/jb.167.2.503-507.1986>
- Gordon, S.V., Heym, B., Parkhill, J., Barrell, B., Cole, S.T., 1999. New insertion sequences and a novel repeated sequence in the genome of *Mycobacterium tuberculosis* H37Rv. *Microbiol. Read. Engl.* 145 ( Pt 4), 881–892. <https://doi.org/10.1099/13500872-145-4-881>
- Goryshin, I.Y., Miller, J.A., Kil, Y.V., Lanzov, V.A., Reznikoff, W.S., 1998. Tn5/IS50 target recognition. *Proc. Natl. Acad. Sci. U. S. A.* 95, 10716–10721. <https://doi.org/10.1073/pnas.95.18.10716>
- Grabundzija, I., Messing, S.A., Thomas, J., Cosby, R.L., Bilic, I., Miskey, C., Gogol-Döring, A., Kapitonov, V., Diem, T., Dalda, A., Jurka, J., Pritham, E.J., Dyda, F., Izsvák, Z., Ivics, Z., 2016. A Helitron transposon reconstructed from bats reveals a novel mechanism of genome shuffling in eukaryotes. *Nat. Commun.* 7, 10716. <https://doi.org/10.1038/ncomms10716>
- Griffith, F., 1928. The Significance of Pneumococcal Types. *J. Hyg. (Lond.)* 27, 113–159. <https://doi.org/10.1017/s0022172400031879>
- Grindley, N.D.F., Whiteson, K.L., Rice, P.A., 2006. Mechanisms of site-specific recombination. *Annu. Rev. Biochem.* 75, 567–605. <https://doi.org/10.1146/annurev.biochem.73.011303.073908>
- Gross, C.A., Chan, C., Dombroski, A., Gruber, T., Sharp, M., Tupy, J., Young, B., 1998. The functional and

- regulatory roles of sigma factors in transcription. *Cold Spring Harb. Symp. Quant. Biol.* 63, 141–155. <https://doi.org/10.1101/sqb.1998.63.141>
- Groth, A.C., Calos, M.P., 2004. Phage integrases: biology and applications. *J. Mol. Biol.* 335, 667–678. <https://doi.org/10.1016/j.jmb.2003.09.082>
- Gruss, A., Ehrlich, S.D., 1989. The family of highly interrelated single-stranded deoxyribonucleic acid plasmids. *Microbiol. Rev.* 53, 231–241. <https://doi.org/10.1128/mr.53.2.231-241.1989>
- Guasch, A., Lucas, M., Moncalián, G., Cabezas, M., Pérez-Luque, R., Gomis-Rüth, F.X., de la Cruz, F., Coll, M., 2003. Recognition and processing of the origin of transfer DNA by conjugative relaxase TrwC. *Nat. Struct. Biol.* 10, 1002–1010. <https://doi.org/10.1038/nsb1017>
- Guglielmini, J., Quintais, L., Garcillán-Barcia, M.P., de la Cruz, F., Rocha, E.P.C., 2011. The repertoire of ICE in prokaryotes underscores the unity, diversity, and ubiquity of conjugation. *PLoS Genet.* 7, e1002222. <https://doi.org/10.1371/journal.pgen.1002222>
- Guynet, C., Achard, A., Hoang, B.T., Barabas, O., Hickman, A.B., Dyda, F., Chandler, M., 2009. Resetting the site: redirecting integration of an insertion sequence in a predictable way. *Mol. Cell* 34, 612–619. <https://doi.org/10.1016/j.molcel.2009.05.017>
- Guynet, C., Hickman, A.B., Barabas, O., Dyda, F., Chandler, M., Ton-Hoang, B., 2008. In vitro reconstitution of a single-stranded transposition mechanism of IS608. *Mol. Cell* 29, 302–312. <https://doi.org/10.1016/j.molcel.2007.12.008>
- Haas, L.F., 1999. Papyrus of Ebers and Smith. *J. Neurol. Neurosurg. Psychiatry* 67, 578. <https://doi.org/10.1136/jnnp.67.5.578>
- Hacker, J., Knapp, S., Goebel, W., 1983. Spontaneous deletions and flanking regions of the chromosomally inherited hemolysin determinant of an *Escherichia coli* O6 strain. *J. Bacteriol.* 154, 1145–1152. <https://doi.org/10.1128/jb.154.3.1145-1152.1983>
- Haensch, S., Bianucci, R., Signoli, M., Rajerison, M., Schultz, M., Kacki, S., Vermunt, M., Weston, D.A., Hurst, D., Achtman, M., Carniel, E., Bramanti, B., 2010. Distinct Clones of *Yersinia pestis* Caused the Black Death. *PLoS Pathog.* 6, e1001134. <https://doi.org/10.1371/journal.ppat.1001134>
- Halioua, B., 2004. *Histoire de la médecine*. Masson.
- Hall, B.G., 1999. Transposable elements as activators of cryptic genes in *E. coli*. *Genetica* 107, 181–187.
- Hall, R.M., 2012. Integrons and gene cassettes: hotspots of diversity in bacterial genomes. *Ann. N. Y. Acad. Sci.* 1267, 71–78. <https://doi.org/10.1111/j.1749-6632.2012.06588.x>
- Hall, R.M., Brookes, D.E., Stokes, H.W., 1991. Site-specific insertion of genes into integrons: role of the 59-base element and determination of the recombination cross-over point. *Mol. Microbiol.* 5, 1941–1959. <https://doi.org/10.1111/j.1365-2958.1991.tb00817.x>
- Hallet, B., Rezsöhazy, R., Delcour, J., 1991. IS231A from *Bacillus thuringiensis* is functional in *Escherichia coli*: transposition and insertion specificity. *J. Bacteriol.* 173, 4526–4529. <https://doi.org/10.1128/jb.173.14.4526-4529.1991>
- Halling, S.M., Kleckner, N., 1982. A symmetrical six-base-pair target site sequence determines Tn10 insertion specificity. *Cell* 28, 155–163. [https://doi.org/10.1016/0092-8674\(82\)90385-3](https://doi.org/10.1016/0092-8674(82)90385-3)
- Han, H.-J., Kuwae, A., Abe, A., Arakawa, Y., Kamachi, K., 2011. Differential expression of type III effector BteA protein due to IS481 insertion in *Bordetella pertussis*. *PLoS One* 6, e17797. <https://doi.org/10.1371/journal.pone.0017797>
- Hanai, R., Wang, J.C., 1993. The mechanism of sequence-specific DNA cleavage and strand transfer by phi X174 gene A\* protein. *J. Biol. Chem.* 268, 23830–23836.
- Haniford, D.B., Ellis, M.J., 2015. Transposons Tn10 and Tn5. *Microbiol. Spectr.* 3, MDNA3-0002–2014. <https://doi.org/10.1128/microbiolspec.MDNA3-0002-2014>
- Harbeck, M., Seifert, L., Hänsch, S., Wagner, D.M., Birdsell, D., Parise, K.L., Wiechmann, I., Grupe, G., Thomas, A., Keim, P., Zöller, L., Bramanti, B., Riehm, J.M., Scholz, H.C., 2013. *Yersinia pestis* DNA from Skeletal Remains from the 6th Century AD Reveals Insights into Justinianic Plague. *PLOS Pathog.* 9, e1003349. <https://doi.org/10.1371/journal.ppat.1003349>
- Haren, L., Bétermier, M., Polard, P., Chandler, M., 1997. IS911-mediated intramolecular transposition is naturally

- temperature sensitive. *Mol. Microbiol.* 25, 531–540. <https://doi.org/10.1046/j.1365-2958.1997.4951854.x>
- Haren, L., Normand, C., Polard, P., Alazard, R., Chandler, M., 2000. IS911 transposition is regulated by protein-protein interactions via a leucine zipper motif. *J. Mol. Biol.* 296, 757–768. <https://doi.org/10.1006/jmbi.1999.3485>
- Haren, L., Polard, P., Ton-Hoang, B., Chandler, M., 1998. Multiple oligomerisation domains in the IS911 transposase: a leucine zipper motif is essential for activity. *J. Mol. Biol.* 283, 29–41. <https://doi.org/10.1006/jmbi.1998.2053>
- Haren, L., Ton-Hoang, B., Chandler, M., 1999. Integrating DNA: transposases and retroviral integrases. *Annu. Rev. Microbiol.* 53, 245–281. <https://doi.org/10.1146/annurev.micro.53.1.245>
- Harmer, C.J., Holt, K.E., Hall, R.M., 2015. A type 2 A/C2 plasmid carrying the aacC4 apramycin resistance gene and the erm(42) erythromycin resistance gene recovered from two *Salmonella enterica* serovars. *J. Antimicrob. Chemother.* 70, 1021–1025. <https://doi.org/10.1093/jac/dku489>
- Harrison, F., Roberts, A.E.L., Gabrielska, R., Rumbaugh, K.P., Lee, C., Diggle, S.P., 2015. A 1,000-Year-Old Antimicrobial Remedy with Antistaphylococcal Activity. *mBio* 6, e01129. <https://doi.org/10.1128/mBio.01129-15>
- Harrison, P.W., Lower, R.P.J., Kim, N.K.D., Young, J.P.W., 2010. Introducing the bacterial “chromid”: not a chromosome, not a plasmid. *Trends Microbiol.* 18, 141–148. <https://doi.org/10.1016/j.tim.2009.12.010>
- Hashida, S., Kitamura, K., Mikami, T., Kishima, Y., 2003. Temperature shift coordinately changes the activity and the methylation state of transposon Tam3 in *Antirrhinum majus*. *Plant Physiol.* 132, 1207–1216. <https://doi.org/10.1104/pp.102.017533>
- Hawley, D.K., McClure, W.R., 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. *Nucleic Acids Res.* 11, 2237–2255. <https://doi.org/10.1093/nar/11.8.2237>
- Hayashi, J., Hamada, N., Kuramitsu, H.K., 2002. The autolysin of *Porphyromonas gingivalis* is involved in outer membrane vesicle release. *FEMS Microbiol. Lett.* 216, 217–222. <https://doi.org/10.1111/j.1574-6968.2002.tb11438.x>
- Hayes, F., 2003. Toxins-antitoxins: plasmid maintenance, programmed cell death, and cell cycle arrest. *Science* 301, 1496–1499. <https://doi.org/10.1126/science.1088157>
- Yassine, H., 2015. Etude de la séquence d’insertion IS1294b et de son implication dans la dissémination des gènes de résistance aux antibiotiques chez les entérobactéries (These de doctorat). Bordeaux.
- He, S., Corneloup, A., Guynet, C., Lavatine, L., Caumont-Sarcos, A., Siguier, P., Marty, B., Dyda, F., Chandler, M., Ton Hoang, B., 2015. The IS200/IS605 Family and “Peel and Paste” Single-strand Transposition Mechanism. *Microbiol. Spectr.* 3. <https://doi.org/10.1128/microbiolspec.MDNA3-0039-2014>
- He, S., Guynet, C., Siguier, P., Hickman, A.B., Dyda, F., Chandler, M., Ton-Hoang, B., 2013. IS200/IS605 family single-strand transposition: mechanism of IS608 strand transfer. *Nucleic Acids Res.* 41, 3302–3313. <https://doi.org/10.1093/nar/gkt014>
- He, S., Hickman, A.B., Dyda, F., Johnson, N.P., Chandler, M., Ton-Hoang, B., 2011. Reconstitution of a functional IS608 single-strand transpososome: role of non-canonical base pairing. *Nucleic Acids Res.* 39, 8503–8512. <https://doi.org/10.1093/nar/gkr566>
- He, Susu, Hickman, A.B., Varani, A.M., Siguier, P., Chandler, M., Dekker, J.P., Dyda, F., 2015. Insertion Sequence IS26 Reorganizes Plasmids in Clinically Isolated Multidrug-Resistant Bacteria by Replicative Transposition. *mBio* 6, e00762. <https://doi.org/10.1128/mBio.00762-15>
- Headd, B., Bradford, S.A., 2018. Physicochemical Factors That Favor Conjugation of an Antibiotic Resistant Plasmid in Non-growing Bacterial Cultures in the Absence and Presence of Antibiotics. *Front. Microbiol.* 9, 2122. <https://doi.org/10.3389/fmicb.2018.02122>
- Hecht, A., Glasgow, J., Jaschke, P.R., Bawazer, L.A., Munson, M.S., Cochran, J.R., Endy, D., Salit, M., 2017. Measurements of translation initiation from all 64 codons in *E. coli*. *Nucleic Acids Res.* 45, 3615–3626. <https://doi.org/10.1093/nar/gkx070>
- Heidary, M., Khosravi, A.D., Khoshnood, S., Nasiri, M.J., Soleimani, S., Goudarzi, M., 2018. Daptomycin. *J. Antimicrob. Chemother.* 73, 1–11. <https://doi.org/10.1093/jac/dkx349>
- Helmann, J.D., 2002. The extracytoplasmic function (ECF) sigma factors. *Adv. Microb. Physiol.* 46, 47–110. [https://doi.org/10.1016/s0065-2911\(02\)46002-x](https://doi.org/10.1016/s0065-2911(02)46002-x)
- Helmy, Y.A., Taha-Abdelaziz, K., Hawwas, H.A.E.-H., Ghosh, S., AlKafaas, S.S., Moawad, M.M.M., Saied, E.M.,

Kassem, I.I., Mawad, A.M.M., 2023. Antimicrobial Resistance and Recent Alternatives to Antibiotics for the Control of Bacterial Pathogens with an Emphasis on Foodborne Pathogens. *Antibiot. Basel Switz.* 12, 274. <https://doi.org/10.3390/antibiotics12020274>

Hickman, A.B., Chandler, M., Dyda, F., 2010a. Integrating prokaryotes and eukaryotes: DNA transposases in light of structure. *Crit. Rev. Biochem. Mol. Biol.* 45, 50–69. <https://doi.org/10.3109/10409230903505596>

Hickman, A.B., Dyda, F., 2015. Mechanisms of DNA Transposition, in: *Mobile DNA III*. John Wiley & Sons, Ltd, pp. 529–553. <https://doi.org/10.1128/9781555819217.ch25>

Hickman, A.B., James, J.A., Barabas, O., Pasternak, C., Ton-Hoang, B., Chandler, M., Sommer, S., Dyda, F., 2010b. DNA recognition and the precleavage state during single-stranded DNA transposition in *D. radiodurans*. *EMBO J.* 29, 3840–3852. <https://doi.org/10.1038/emboj.2010.241>

Hickman, A.B., Ronning, D.R., Kotin, R.M., Dyda, F., 2002. Structural unity among viral origin binding proteins: crystal structure of the nuclease domain of adeno-associated virus Rep. *Mol. Cell* 10, 327–337. [https://doi.org/10.1016/s1097-2765\(02\)00592-0](https://doi.org/10.1016/s1097-2765(02)00592-0)

Hickman, A.B., Ronning, D.R., Perez, Z.N., Kotin, R.M., Dyda, F., 2004. The nuclease domain of adeno-associated virus rep coordinates replication initiation using two distinct DNA recognition interfaces. *Mol. Cell* 13, 403–414. [https://doi.org/10.1016/s1097-2765\(04\)00023-1](https://doi.org/10.1016/s1097-2765(04)00023-1)

Hickman, A.B., Waninger, S., Scocca, J.J., Dyda, F., 1997. Molecular organization in site-specific recombination: the catalytic domain of bacteriophage HP1 integrase at 2.7 Å resolution. *Cell* 89, 227–237. [https://doi.org/10.1016/s0092-8674\(00\)80202-0](https://doi.org/10.1016/s0092-8674(00)80202-0)

Higgins, P.G., Fluit, A.C., Schmitz, F.J., 2003. Fluoroquinolones: structure and target sites. *Curr. Drug Targets* 4, 181–190. <https://doi.org/10.2174/1389450033346920>

Hirota, Y., Mordoh, J., Jacob, F., 1970. On the process of cellular division in *Escherichia coli*. 3. Thermosensitive mutants of *Escherichia coli* altered in the process of DNA initiation. *J. Mol. Biol.* 53, 369–387. [https://doi.org/10.1016/0022-2836\(70\)90072-0](https://doi.org/10.1016/0022-2836(70)90072-0)

Hirsch, H.J., Starlinger, P., Brachet, P., 1972. Two kinds of insertions in bacterial genes. *Mol. Gen. Genet. MGG* 119, 191–206. <https://doi.org/10.1007/BF00333858>

Hochhut, B., Jahreis, K., Lengeler, J.W., Schmid, K., 1997. CTnscr94, a conjugative transposon found in enterobacteria. *J. Bacteriol.* 179, 2097–2102. <https://doi.org/10.1128/jb.179.7.2097-2102.1997>

Hochhut, B., Lotfi, Y., Mazel, D., Faruque, S.M., Woodgate, R., Waldor, M.K., 2001. Molecular analysis of antibiotic resistance gene clusters in *Vibrio cholerae* O139 and O1 SXT constains. *Antimicrob. Agents Chemother.* 45, 2991–3000. <https://doi.org/10.1128/AAC.45.11.2991-3000.2001>

Hoover, T.A., Vodkin, M.H., Williams, J.C., 1992. A *Coxiella burnetii* repeated DNA element resembling a bacterial insertion sequence. *J. Bacteriol.* 174, 5540–5548. <https://doi.org/10.1128/jb.174.17.5540-5548.1992>

Hu, S.T., Hwang, J.H., Lee, L.C., Lee, C.H., Li, P.L., Hsieh, Y.C., 1994. Functional analysis of the 14 kDa protein of insertion sequence 2. *J. Mol. Biol.* 236, 503–513. <https://doi.org/10.1006/jmbi.1994.1161>

Hu, W.-Y., Derbyshire, K.M., 1998. Target Choice and Orientation Preference of the Insertion Sequence IS903. *J. Bacteriol.* 180, 3039–3048.

Hu, Y., Yang, X., Li, J., Lv, N., Liu, F., Wu, J., Lin, I.Y.C., Wu, N., Weimer, B.C., Gao, G.F., Liu, Y., Zhu, B., 2016. The Bacterial Mobile Resistome Transfer Network Connecting the Animal and Human Microbiomes. *Appl. Environ. Microbiol.* 82, 6672–6681. <https://doi.org/10.1128/AEM.01802-16>

Huber, M., Faure, G., Laass, S., Kolbe, E., Seitz, K., Wehrheim, C., Wolf, Y.I., Koonin, E.V., Soppa, J., 2019. Translational coupling via termination-reinitiation in archaea and bacteria. *Nat. Commun.* 10, 4006. <https://doi.org/10.1038/s41467-019-11999-9>

Huisman, O., Errada, P.R., Signon, L., Kleckner, N., 1989. Mutational analysis of IS10's outside end. *EMBO J.* 8, 2101–2109. <https://doi.org/10.1002/j.1460-2075.1989.tb03619.x>

Iannelli, F., Santagati, M., Santoro, F., Oggioni, M.R., Stefani, S., Pozzi, G., 2014. Nucleotide sequence of conjugative prophage  $\Phi$ 1207.3 (formerly Tn1207.3) carrying the *mef(A)/msr(D)* genes for efflux resistance to macrolides in *Streptococcus pyogenes*. *Front. Microbiol.* 5, 687. <https://doi.org/10.3389/fmicb.2014.00687>

Iguchi, A., Nagaya, Y., Pradel, E., Ooka, T., Ogura, Y., Katsura, K., Kurokawa, K., Oshima, K., Hattori, M., Parkhill, J., Sebaihia, M., Coulthurst, S.J., Gotoh, N., Thomson, N.R., Ewbank, J.J., Hayashi, T., 2014. Genome evolution and plasticity of *Serratia marcescens*, an important multidrug-resistant nosocomial pathogen. *Genome Biol. Evol.*

6, 2096–2110. <https://doi.org/10.1093/gbe/evu160>

Ilangovan, A., Kay, C.W.M., Roier, S., El Mkami, H., Salvadori, E., Zechner, E.L., Zanetti, G., Waksman, G., 2017. Cryo-EM Structure of a Relaxase Reveals the Molecular Basis of DNA Unwinding during Bacterial Conjugation. *Cell* 169, 708–721.e12. <https://doi.org/10.1016/j.cell.2017.04.010>

Ilves, H., Hörak, R., Kivisaar, M., 2001. Involvement of sigma(S) in starvation-induced transposition of *Pseudomonas putida* transposon Tn4652. *J. Bacteriol.* 183, 5445–5448. <https://doi.org/10.1128/JB.183.18.5445-5448.2001>

Ilyina, T.V., Koonin, E.V., 1992. Conserved sequence motifs in the initiator proteins for rolling circle DNA replication encoded by diverse replicons from eubacteria, eucaryotes and archaeobacteria. *Nucleic Acids Res.* 20, 3279–3285.

Iranzo, J., Gómez, M.J., López de Saro, F.J., Manrubia, S., 2014. Large-scale genomic analysis suggests a neutral punctuated dynamics of transposable elements in bacterial genomes. *PLoS Comput. Biol.* 10, e1003680. <https://doi.org/10.1371/journal.pcbi.1003680>

Isberg, R.R., Syvanen, M., 1982. DNA gyrase is a host factor required for transposition of Tn5. *Cell* 30, 9–18. [https://doi.org/10.1016/0092-8674\(82\)90006-x](https://doi.org/10.1016/0092-8674(82)90006-x)

Iyer, A., Barbour, E., Azhar, E., Salabi, A.A.E., Hassan, H.M.A., Qadri, I., Chaudhary, A., Abuzenadah, A., Kumosani, T., Damanhour, G., Alawi, M., Na'was, T., Nour, A.M.A., Harakeh, S., 2013. Transposable elements in *Escherichia coli* antimicrobial resistance. *Adv. Biosci. Biotechnol.* 4, 415–423. <https://doi.org/10.4236/abb.2013.43A055>

Iyer, S., Le, D., Park, B.R., Kim, M., 2018. Distinct mechanisms coordinate transcription and translation under carbon and nitrogen starvation in *Escherichia coli*. *Nat. Microbiol.* 3, 741–748. <https://doi.org/10.1038/s41564-018-0161-3>

Janoir, C., Podglajen, I., Kitzis, M.D., Poyart, C., Gutmann, L., 1999. In vitro exchange of fluoroquinolone resistance determinants between *Streptococcus pneumoniae* and viridans streptococci and genomic organization of the parE-parC region in *S. mitis*. *J. Infect. Dis.* 180, 555–558. <https://doi.org/10.1086/314888>

Jansson, C., Sköld, O., 1991. Appearance of a new trimethoprim resistance gene, dhfrIX, in *Escherichia coli* from swine. *Antimicrob. Agents Chemother.* 35, 1891–1899. <https://doi.org/10.1128/AAC.35.9.1891>

Jellen-Ritter, A.S., Kern, W.V., 2001. Enhanced expression of the multidrug efflux pumps AcrAB and AcrEF associated with insertion element transposition in *Escherichia coli* mutants Selected with a fluoroquinolone. *Antimicrob. Agents Chemother.* 45, 1467–1472. <https://doi.org/10.1128/AAC.45.5.1467-1472.2001>

Jin, M., Liu, L., Wang, D.-N., Yang, D., Liu, W.-L., Yin, J., Yang, Z.-W., Wang, H.-R., Qiu, Z.-G., Shen, Z.-Q., Shi, D.-Y., Li, H.-B., Guo, J.-H., Li, J.-W., 2020. Chlorine disinfection promotes the exchange of antibiotic resistance genes across bacterial genera by natural transformation. *ISME J.* 14, 1847–1856. <https://doi.org/10.1038/s41396-020-0656-9>

Johnson, C.M., Grossman, A.D., 2015. Integrative and Conjugative Elements (ICEs): What They Do and How They Work. *Annu. Rev. Genet.* 49, 577–601. <https://doi.org/10.1146/annurev-genet-112414-055018>

Johnson, R.C., Reznikoff, W.S., 1983. DNA sequences at the ends of transposon Tn5 required for transposition. *Nature* 304, 280–282. <https://doi.org/10.1038/304280a0>

Johnston, C., Martin, B., Fichant, G., Polard, P., Claverys, J.-P., 2014. Bacterial transformation: distribution, shared mechanisms and divergent control. *Nat. Rev. Microbiol.* 12, 181–196. <https://doi.org/10.1038/nrmicro3199>

Johnston, N.J., Mukhtar, T.A., Wright, G.D., 2002. Streptogramin antibiotics: mode of action and resistance. *Curr. Drug Targets* 3, 335–344. <https://doi.org/10.2174/1389450023347678>

Jové, T., Da Re, S., Denis, F., Mazel, D., Ploy, M.-C., 2010. Inverse correlation between promoter strength and excision activity in class 1 integrons. *PLoS Genet.* 6, e1000793. <https://doi.org/10.1371/journal.pgen.1000793>

Juhas, M., Power, P.M., Harding, R.M., Ferguson, D.J.P., Dimopoulou, I.D., Elamin, A.R. e, Mohd-Zain, Z., Hood, D.W., Adegbola, R., Erwin, A., Smith, A., Munson, R.S., Harrison, A., Mansfield, L., Bentley, S., Crook, D.W., 2007. Sequence and functional analyses of *Haemophilus* spp. genomic islands. *Genome Biol.* 8, R237. <https://doi.org/10.1186/gb-2007-8-11-r237>

Juiz-Río, S., Osorio, C.R., de Lorenzo, V., Lemos, M.L., 2005. Subtractive hybridization reveals a high genetic diversity in the fish pathogen *Photobacterium damsela* subsp. *piscicida*: evidence of a SXT-like element. *Microbiol. Read. Engl.* 151, 2659–2669. <https://doi.org/10.1099/mic.0.27891-0>

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek,

- A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S.A.A., Ballard, A.J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A.W., Kavukcuoglu, K., Kohli, P., Hassabis, D., 2021. Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589. <https://doi.org/10.1038/s41586-021-03819-2>
- Kado, C.I., 2014. Historical account on gaining insights on the mechanism of crown gall tumorigenesis induced by *Agrobacterium tumefaciens*. *Front. Microbiol.* 5, 340. <https://doi.org/10.3389/fmicb.2014.00340>
- Kahne, D., Leimkuhler, C., Lu, W., Walsh, C., 2005. Glycopeptide and lipoglycopeptide antibiotics. *Chem. Rev.* 105, 425–448. <https://doi.org/10.1021/cr030103a>
- Kapitonov, V.V., Jurka, J., 2007. Helitrons on a roll: eukaryotic rolling-circle transposons. *Trends Genet. TIG* 23, 521–529. <https://doi.org/10.1016/j.tig.2007.08.004>
- Kapitonov, V.V., Jurka, J., 2006. Self-synthesizing DNA transposons in eukaryotes. *Proc. Natl. Acad. Sci. U. S. A.* 103, 4540–4545. <https://doi.org/10.1073/pnas.0600833103>
- Kapitonov, V.V., Jurka, J., 2001. Rolling-circle transposons in eukaryotes. *Proc. Natl. Acad. Sci. U. S. A.* 98, 8714–8719. <https://doi.org/10.1073/pnas.151269298>
- Kapitonov, V.V., Jurka, J., 1999. Molecular paleontology of transposable elements from *Arabidopsis thaliana*. *Genetica* 107, 27–37.
- Karvelis, T., Druteika, G., Bigelyte, G., Budre, K., Zedaveinyte, R., Silanskas, A., Kazlauskas, D., Venclovas, Č., Siksnyš, V., 2021. Transposon-associated TnpB is a programmable RNA-guided DNA endonuclease. *Nature* 599, 692–696. <https://doi.org/10.1038/s41586-021-04058-1>
- Kehrenberg, C., Schwarz, S., 2005. Plasmid-borne florfenicol resistance in *Pasteurella multocida*. *J. Antimicrob. Chemother.* 55, 773–775. <https://doi.org/10.1093/jac/dki102>
- Keller, M.P., 1996. Regulation of transposition activity in the procaryotic insertion sequences IS30 and IS186.
- Kersulyte, D., Akopyants, N.S., Clifton, S.W., Roe, B.A., Berg, D.E., 1998. Novel sequence organization and insertion specificity of IS605 and IS606: chimaeric transposable elements of *Helicobacter pylori*. *Gene* 223, 175–186. [https://doi.org/10.1016/s0378-1119\(98\)00164-4](https://doi.org/10.1016/s0378-1119(98)00164-4)
- Kersulyte, D., Kalia, A., Zhang, M., Lee, H.-K., Subramaniam, D., Kiuduliene, L., Chalkauskas, H., Berg, D.E., 2004. Sequence organization and insertion specificity of the novel chimeric ISHp609 transposable element of *Helicobacter pylori*. *J. Bacteriol.* 186, 7521–7528. <https://doi.org/10.1128/JB.186.22.7521-7528.2004>
- Kersulyte, D., Mukhopadhyay, A.K., Shirai, M., Nakazawa, T., Berg, D.E., 2000. Functional organization and insertion specificity of IS607, a chimeric element of *Helicobacter pylori*. *J. Bacteriol.* 182, 5300–5308. <https://doi.org/10.1128/JB.182.19.5300-5308.2000>
- Kersulyte, D., Velapatiño, B., Dailide, G., Mukhopadhyay, A.K., Ito, Y., Cahuayme, L., Parkinson, A.J., Gilman, R.H., Berg, D.E., 2002. Transposable element ISHp608 of *Helicobacter pylori*: nonrandom geographic distribution, functional organization, and insertion specificity. *J. Bacteriol.* 184, 992–1002. <https://doi.org/10.1128/jb.184.4.992-1002.2002>
- Kesty, N.C., Mason, K.M., Reedy, M., Miller, S.E., Kuehn, M.J., 2004. Enterotoxigenic *Escherichia coli* vesicles target toxin delivery into mammalian cells. *EMBO J.* 23, 4538–4549. <https://doi.org/10.1038/sj.emboj.7600471>
- Kireeva, M.L., Kashlev, M., 2009. Mechanism of sequence-specific pausing of bacterial RNA polymerase. *Proc. Natl. Acad. Sci. U. S. A.* 106, 8900–8905. <https://doi.org/10.1073/pnas.0900407106>
- Kleckner, N., 1981. Transposable elements in prokaryotes. *Annu. Rev. Genet.* 15, 341–404. <https://doi.org/10.1146/annurev.ge.15.120181.002013>
- Kleckner, N., Chalmers, R.M., Kwon, D., Sakai, J., Bolland, S., 1996. Tn10 and IS10 transposition and chromosome rearrangements: mechanism and regulation in vivo and in vitro. *Curr. Top. Microbiol. Immunol.* 204, 49–82. [https://doi.org/10.1007/978-3-642-79795-8\\_3](https://doi.org/10.1007/978-3-642-79795-8_3)
- Klieve, A.V., Yokoyama, M.T., Forster, R.J., Ouwkerk, D., Bain, P.A., Mawhinney, E.L., 2005. Naturally occurring DNA transfer system associated with membrane vesicles in cellulolytic *Ruminococcus* spp. of ruminal origin. *Appl. Environ. Microbiol.* 71, 4248–4253. <https://doi.org/10.1128/AEM.71.8.4248-4253.2005>
- Klug, A., Schwabe, J.W., 1995. Protein motifs 5. Zinc fingers. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* 9, 597–604.

- Kluska, K., Adamczyk, J., Krężel, A., 2018. Metal binding properties, stability and reactivity of zinc fingers. *Coord. Chem. Rev.* 367, 18–64. <https://doi.org/10.1016/j.ccr.2018.04.009>
- Knapp, S., Hacker, J., Then, I., Müller, D., Goebel, W., 1984. Multiple copies of hemolysin genes and associated sequences in the chromosomes of uropathogenic *Escherichia coli* strains. *J. Bacteriol.* 159, 1027–1033. <https://doi.org/10.1128/jb.159.3.1027-1033.1984>
- Kohler, R., Mooney, R.A., Mills, D.J., Landick, R., Cramer, P., 2017. Architecture of a transcribing-translating expressome. *Science* 356, 194–197. <https://doi.org/10.1126/science.aal3059>
- Kojima, K.K., 2020. Structural and sequence diversity of eukaryotic transposable elements. *Genes Genet. Syst.* 94, 233–252. <https://doi.org/10.1266/ggs.18-00024>
- Koonin, E.V., Ilyina, T.V., 1993. Computer-assisted dissection of rolling circle DNA replication. *Biosystems* 30, 241–268. [https://doi.org/10.1016/0303-2647\(93\)90074-m](https://doi.org/10.1016/0303-2647(93)90074-m)
- Kornberg, A., Baker, T.A., 1992. *DNA Replication*. W.H. Freeman.
- Korswagen, H.C., Durbin, R.M., Smits, M.T., Plasterk, R.H., 1996. Transposon Tc1-derived, sequence-tagged sites in *Caenorhabditis elegans* as markers for gene mapping. *Proc. Natl. Acad. Sci. U. S. A.* 93, 14680–14685. <https://doi.org/10.1073/pnas.93.25.14680>
- Koskella, B., 2013. Phage-mediated selection on microbiota of a long-lived host. *Curr. Biol. CB* 23, 1256–1260. <https://doi.org/10.1016/j.cub.2013.05.038>
- Krebs, M.P., Reznikoff, W.S., 1986. Transcriptional and translational initiation sites of IS50. Control of transposase and inhibitor expression. *J. Mol. Biol.* 192, 781–791. [https://doi.org/10.1016/0022-2836\(86\)90028-8](https://doi.org/10.1016/0022-2836(86)90028-8)
- Kretschmer, P.J., Cohen, S.N., 1979. Effect of temperature on translocation frequency of the Tn3 element. *J. Bacteriol.* 139, 515–519. <https://doi.org/10.1128/jb.139.2.515-519.1979>
- Kuan, C.T., Tessman, I., 1991. LexA protein of *Escherichia coli* represses expression of the Tn5 transposase gene. *J. Bacteriol.* 173, 6406–6410. <https://doi.org/10.1128/jb.173.20.6406-6410.1991>
- Kuno, S., Yoshida, T., Kamikawa, R., Hosoda, N., Sako, Y., 2010. The distribution of a phage-related insertion sequence element in the cyanobacterium, *Microcystis aeruginosa*. *Microbes Environ.* 25, 295–301. <https://doi.org/10.1264/jsme2.me10125>
- Lagier, J.-C., Hugon, P., Khelaifia, S., Fournier, P.-E., La Scola, B., Raoult, D., 2015. The rebirth of culture in microbiology through the example of culturomics to study human gut microbiota. *Clin. Microbiol. Rev.* 28, 237–264. <https://doi.org/10.1128/CMR.00014-14>
- Lallement, C., 2018. Caractérisation des séquences d'insertions ISCR bactériennes impliquées dans la résistance aux antibiotiques (These de doctorat). Limoges.
- Lallement, C., Pasternak, C., Ploy, M.-C., Jové, T., 2018. The Role of ISCR1-Borne POUT Promoters in the Expression of Antibiotic Resistance Genes. *Front. Microbiol.* 9, 2579. <https://doi.org/10.3389/fmicb.2018.02579>
- Lam, S., Roth, J.R., 1986. Structural and functional studies of insertion element IS200. *J. Mol. Biol.* 187, 157–167. [https://doi.org/10.1016/0022-2836\(86\)90225-1](https://doi.org/10.1016/0022-2836(86)90225-1)
- Lam, S., Roth, J.R., 1983. IS200: a *Salmonella*-specific insertion sequence. *Cell* 34, 951–960. [https://doi.org/10.1016/0092-8674\(83\)90552-4](https://doi.org/10.1016/0092-8674(83)90552-4)
- Lanka, E., Wilkins, B.M., 1995. DNA processing reactions in bacterial conjugation. *Annu. Rev. Biochem.* 64, 141–169. <https://doi.org/10.1146/annurev.bi.64.070195.001041>
- Larkin, C., Datta, S., Harley, M.J., Anderson, B.J., Ebie, A., Hargreaves, V., Schildbach, J.F., 2005. Inter- and intramolecular determinants of the specificity of single-stranded DNA binding and cleavage by the F factor relaxase. *Struct. Lond. Engl.* 1993 13, 1533–1544. <https://doi.org/10.1016/j.str.2005.06.013>
- Larkin, C., Datta, S., Nezami, A., Dohm, J.A., Schildbach, J.F., 2003. Crystallization and preliminary X-ray characterization of the relaxase domain of F factor Tral. *Acta Crystallogr. D Biol. Crystallogr.* 59, 1514–1516. <https://doi.org/10.1107/s0907444903012964>
- Larkin, C., Haft, R.J.F., Harley, M.J., Traxler, B., Schildbach, J.F., 2007. Roles of active site residues and the HUH motif of the F plasmid Tral relaxase. *J. Biol. Chem.* 282, 33707–33713. <https://doi.org/10.1074/jbc.M703210200>
- Lartigue, M.-F., Poirel, L., Aubert, D., Nordmann, P., 2006. In vitro analysis of ISEcp1B-mediated mobilization of naturally occurring beta-lactamase gene blaCTX-M of *Kluyvera ascorbata*. *Antimicrob. Agents Chemother.* 50, 1282–1286. <https://doi.org/10.1128/AAC.50.4.1282-1286.2006>



- Lascols, C., Podglajen, I., Verdet, C., Gautier, V., Gutmann, L., Soussy, C.-J., Collatz, E., Cambau, E., 2008. A plasmid-borne *Shewanella* algae Gene, *qnrA3*, and its possible transfer in vivo between *Kluyvera ascorbata* and *Klebsiella pneumoniae*. *J. Bacteriol.* 190, 5217–5223. <https://doi.org/10.1128/JB.00243-08>
- Lauf, U., Müller, C., Herrmann, H., 1999. Identification and characterisation of IS1383, a new insertion sequence isolated from *Pseudomonas putida* strain H. *FEMS Microbiol. Lett.* 170, 407–412. <https://doi.org/10.1111/j.1574-6968.1999.tb13401.x>
- Lawrence, J.G., Hendrix, R.W., Casjens, S., 2001. Where are the pseudogenes in bacterial genomes? *Trends Microbiol.* 9, 535–540. [https://doi.org/10.1016/s0966-842x\(01\)02198-9](https://doi.org/10.1016/s0966-842x(01)02198-9)
- Laxminarayan, R., Duse, A., Wattal, C., Zaidi, A.K.M., Wertheim, H.F.L., Sumpradit, N., Vlieghe, E., Hara, G.L., Gould, I.M., Goossens, H., Greko, C., So, A.D., Bigdeli, M., Tomson, G., Woodhouse, W., Ombaka, E., Peralta, A.Q., Qamar, F.N., Mir, F., Kariuki, S., Bhutta, Z.A., Coates, A., Bergstrom, R., Wright, G.D., Brown, E.D., Cars, O., 2013. Antibiotic resistance—the need for global solutions. *Lancet Infect. Dis.* 13, 1057–1098. [https://doi.org/10.1016/S1473-3099\(13\)70318-9](https://doi.org/10.1016/S1473-3099(13)70318-9)
- Le Rouzic, A., Boutin, T.S., Capy, P., 2007. Long-term evolution of transposable elements. *Proc. Natl. Acad. Sci. U. S. A.* 104, 19375–19380. <https://doi.org/10.1073/pnas.0705238104>
- Le Rouzic, A., Capy, P., 2005. The First Steps of Transposable Elements Invasion. *Genetics* 169, 1033–1043. <https://doi.org/10.1534/genetics.104.031211>
- Le Rouzic, A., Deceliere, G., 2005. Models of the population genetics of transposable elements. *Genet. Res.* 85, 171–181. <https://doi.org/10.1017/S0016672305007585>
- Leclercq, S.O., Wang, C., Zhu, Y., Wu, H., Du, X., Liu, Z., Feng, J., 2016. Diversity of the Tetracycline Mobilome within a Chinese Pig Manure Sample. *Appl. Environ. Microbiol.* 82, 6454–6462. <https://doi.org/10.1128/AEM.01754-16>
- Lederberg, J., Tatum, E.L., 1946. Gene recombination in *Escherichia coli*. *Nature* 158, 558. <https://doi.org/10.1038/158558a0>
- Lee, C.A., Auchtung, J.M., Monson, R.E., Grossman, A.D., 2007. Identification and characterization of *int* (integrase), *xis* (excisionase) and chromosomal attachment sites of the integrative and conjugative element ICEBs1 of *Bacillus subtilis*. *Mol. Microbiol.* 66, 1356–1369. <https://doi.org/10.1111/j.1365-2958.2007.06000.x>
- Lee, H.H., Yoon, J.Y., Kim, H.S., Kang, J.Y., Kim, K.H., Kim, D.J., Ha, J.Y., Mikami, B., Yoon, H.J., Suh, S.W., 2006. Crystal structure of a metal ion-bound IS200 transposase. *J. Biol. Chem.* 281, 4261–4266. <https://doi.org/10.1074/jbc.M511567200>
- Leelaporn, A., Firth, N., Byrne, M.E., Roper, E., Skurray, R.A., 1994. Possible role of insertion sequence IS257 in dissemination and expression of high- and low-level trimethoprim resistance in staphylococci. *Antimicrob. Agents Chemother.* 38, 2238–2244. <https://doi.org/10.1128/AAC.38.10.2238>
- Lerminiaux, N.A., Cameron, A.D.S., 2019. Horizontal transfer of antibiotic resistance genes in clinical environments. *Can. J. Microbiol.* 65, 34–44. <https://doi.org/10.1139/cjm-2018-0275>
- Lévesque, C., Brassard, S., Lapointe, J., Roy, P.H., 1994. Diversity and relative strength of tandem promoters for the antibiotic-resistance genes of several integrons. *Gene* 142, 49–54. [https://doi.org/10.1016/0378-1119\(94\)90353-0](https://doi.org/10.1016/0378-1119(94)90353-0)
- Levings, R.S., Partridge, S.R., Djordjevic, S.P., Hall, R.M., 2007. SGI1-K, a variant of the SGI1 genomic island carrying a mercury resistance region, in *Salmonella enterica* serovar Kentucky. *Antimicrob. Agents Chemother.* 51, 317–323. <https://doi.org/10.1128/AAC.01229-06>
- Levy, S.B., Marshall, B., 2004. Antibacterial resistance worldwide: causes, challenges and responses. *Nat. Med.* 10, S122–129. <https://doi.org/10.1038/nm1145>
- Lewis, L.A., Astatke, M., Umekubo, P.T., Alvi, S., Saby, R., Afrose, J., 2011. Soluble expression, purification and characterization of the full length IS2 Transposase. *Mob. DNA* 2, 14. <https://doi.org/10.1186/1759-8753-2-14>
- Lewis, L.A., Cylin, E., Lee, H.K., Saby, R., Wong, W., Grindley, N.D.F., 2004. The left end of IS2: a compromise between transpositional activity and an essential promoter function that regulates the transposition pathway. *J. Bacteriol.* 186, 858–865. <https://doi.org/10.1128/JB.186.3.858-865.2004>
- Li, J., Cao, J., Zhu, Y.-G., Chen, Q.-L., Shen, F., Wu, Y., Xu, S., Fan, H., Da, G., Huang, R.-J., Wang, J., de Jesus, A.L., Morawska, L., Chan, C.K., Peccia, J., Yao, M., 2018. Global Survey of Antibiotic Resistance Genes in Air. *Environ. Sci. Technol.* 52, 10975–10984. <https://doi.org/10.1021/acs.est.8b02204>

- Li, J., Lan, R., Xiong, Y., Ye, C., Yuan, M., Liu, X., Chen, X., Yu, D., Liu, B., Lin, W., Bai, X., Wang, Yan, Sun, Q., Wang, Yiting, Zhao, H., Meng, Q., Chen, Q., Zhao, A., Xu, J., 2014. Sequential isolation in a patient of *Raoultella planticola* and *Escherichia coli* bearing a novel ISCR1 element carrying bla<sub>NDM-1</sub>. *PLoS One* 9, e89893. <https://doi.org/10.1371/journal.pone.0089893>
- Li, Z., Clarke, A.J., Beveridge, T.J., 1998. Gram-negative bacteria produce membrane vesicles which are capable of killing other bacteria. *J. Bacteriol.* 180, 5478–5483. <https://doi.org/10.1128/JB.180.20.5478-5483.1998>
- Liu, G., Thomsen, L.E., Olsen, J.E., 2022. Antimicrobial-induced horizontal transfer of antimicrobial resistance genes in bacteria: a mini-review. *J. Antimicrob. Chemother.* 77, 556–567. <https://doi.org/10.1093/jac/dkab450>
- Llosa, M., Gomis-Rüth, F.X., Coll, M., de la Cruz Fd, F., 2002. Bacterial conjugation: a two-step mechanism for DNA transport. *Mol. Microbiol.* 45, 1–8. <https://doi.org/10.1046/j.1365-2958.2002.03014.x>
- Llosa, M., Zunzunegui, S., de la Cruz, F., 2003. Conjugative coupling proteins interact with cognate and heterologous VirB10-like proteins while exhibiting specificity for cognate relaxosomes. *Proc. Natl. Acad. Sci. U. S. A.* 100, 10465–10470. <https://doi.org/10.1073/pnas.1830264100>
- Loot, C., Turlan, C., Chandler, M., 2004. Host processing of branched DNA intermediates is involved in targeted transposition of IS911. *Mol. Microbiol.* 51, 385–393. <https://doi.org/10.1046/j.1365-2958.2003.03850.x>
- López de Felipe, F., Magni, C., de Mendoza, D., López, P., 1996. Transcriptional activation of the citrate permease P gene of *Lactococcus lactis* biovar diacetylactis by an insertion sequence-like element present in plasmid pCIT264. *Mol. Gen. Genet.* MGG 250, 428–436. <https://doi.org/10.1007/BF02174031>
- Lu, J., Wang, Y., Zhang, S., Bond, P., Yuan, Z., Guo, J., 2020. Triclosan at environmental concentrations can enhance the spread of extracellular antibiotic resistance genes through transformation. *Sci. Total Environ.* 713, 136621. <https://doi.org/10.1016/j.scitotenv.2020.136621>
- Lynch, M., Conery, J.S., 2003. The origins of genome complexity. *Science* 302, 1401–1404. <https://doi.org/10.1126/science.1089370>
- Ma, J., Campbell, A., Karlin, S., 2002. Correlations between Shine-Dalgarno sequences and gene features such as predicted expression levels and operon structures. *J. Bacteriol.* 184, 5733–5745. <https://doi.org/10.1128/JB.184.20.5733-5745.2002>
- Machida, Y., Machida, C., Ohtsubo, H., Ohtsubo, E., 1982. Factors determining frequency of plasmid cointegration mediated by insertion sequence IS1. *Proc. Natl. Acad. Sci. U. S. A.* 79, 277–281. <https://doi.org/10.1073/pnas.79.2.277>
- Maekawa, T., Yanagihara, K., Ohtsubo, E., 1996. Specific nicking at the 3' ends of the terminal inverted repeat sequences in transposon Tn3 by transposase and an *E. coli* protein ACP. *Genes Cells Devoted Mol. Cell. Mech.* 1, 1017–1030. <https://doi.org/10.1046/j.1365-2443.1996.d01-221.x>
- Mahillon, J., Chandler, M., 1998. Insertion sequences. *Microbiol. Mol. Biol. Rev.* MMBR 62, 725–774. <https://doi.org/10.1128/MMBR.62.3.725-774.1998>
- Makris, J.C., Nordmann, P.L., Reznikoff, W.S., 1988. Mutational analysis of insertion sequence 50 (IS50) and transposon 5 (Tn5) ends. *Proc. Natl. Acad. Sci. U. S. A.* 85, 2224–2228. <https://doi.org/10.1073/pnas.85.7.2224>
- Malachowa, N., DeLeo, F.R., 2010. Mobile genetic elements of *Staphylococcus aureus*. *Cell. Mol. Life Sci. CMLS* 67, 3057–3071. <https://doi.org/10.1007/s00018-010-0389-4>
- Mammeri, H., 2013. Mode d'action des antibiotiques. *Serv. Bactériologie CHU Amiens P 2.*
- Mammeri, H., Van De Loo, M., Poirel, L., Martinez-Martinez, L., Nordmann, P., 2005. Emergence of plasmid-mediated quinolone resistance in *Escherichia coli* in Europe. *Antimicrob. Agents Chemother.* 49, 71–76. <https://doi.org/10.1128/AAC.49.1.71-76.2005>
- Markley, J.L., Wenczewicz, T.A., 2018. Tetracycline-Inactivating Enzymes. *Front. Microbiol.* 9, 1058. <https://doi.org/10.3389/fmicb.2018.01058>
- Martiel, J.-L., Blot, M., 2002. Transposable elements and fitness of bacteria. *Theor. Popul. Biol.* 61, 509–518. <https://doi.org/10.1006/tpbi.2002.1603>
- Matos, R.C., Lapaque, N., Rigottier-Gois, L., Debarbieux, L., Meylheuc, T., Gonzalez-Zorn, B., Repoila, F., Lopes, M. de F., Serror, P., 2013. *Enterococcus faecalis* prophage dynamics and contributions to pathogenic traits. *PLoS Genet.* 9, e1003539. <https://doi.org/10.1371/journal.pgen.1003539>
- Mazel, D., 2006. Integrons: agents of bacterial evolution. *Nat. Rev. Microbiol.* 4, 608–620.

<https://doi.org/10.1038/nrmicro1462>

McBroom, A.J., Johnson, A.P., Vemulapalli, S., Kuehn, M.J., 2006. Outer membrane vesicle production by *Escherichia coli* is independent of membrane instability. *J. Bacteriol.* 188, 5385–5392. <https://doi.org/10.1128/JB.00498-06>

McBroom, A.J., Kuehn, M.J., 2007. Release of outer membrane vesicles by Gram-negative bacteria is a novel envelope stress response. *Mol. Microbiol.* 63, 545–558. <https://doi.org/10.1111/j.1365-2958.2006.05522.x>

McCarthy, J.E., Gualerzi, C., 1990. Translational control of prokaryotic gene expression. *Trends Genet. TIG* 6, 78–85. [https://doi.org/10.1016/0168-9525\(90\)90098-q](https://doi.org/10.1016/0168-9525(90)90098-q)

McEvoy, C.R.E., Falmer, A.A., Gey van Pittius, N.C., Victor, T.C., van Helden, P.D., Warren, R.M., 2007. The role of IS6110 in the evolution of *Mycobacterium tuberculosis*. *Tuberc. Edinb. Scotl.* 87, 393–404. <https://doi.org/10.1016/j.tube.2007.05.010>

Medini, D., Donati, C., Tettelin, H., Massignani, V., Rappuoli, R., 2005. The microbial pan-genome. *Curr. Opin. Genet. Dev.* 15, 589–594. <https://doi.org/10.1016/j.gde.2005.09.006>

Mell, J.C., Redfield, R.J., 2014. Natural competence and the evolution of DNA uptake specificity. *J. Bacteriol.* 196, 1471–1483. <https://doi.org/10.1128/JB.01293-13>

Mendiola, M.V., Bernales, I., de la Cruz, F., 1994. Differential roles of the transposon termini in IS91 transposition. *Proc. Natl. Acad. Sci. U. S. A.* 91, 1922–1926. <https://doi.org/10.1073/pnas.91.5.1922>

Mendiola, M.V., de la Cruz, F., 1992. IS91 transposase is related to the rolling-circle-type replication proteins of the pUB110 family of plasmids. *Nucleic Acids Res.* 20, 3521. <https://doi.org/10.1093/nar/20.13.3521>

Mendiola, M.V., de la Cruz, F., 1989. Specificity of insertion of IS91, an insertion sequence present in alpha-haemolysin plasmids of *Escherichia coli*. *Mol. Microbiol.* 3, 979–984. <https://doi.org/10.1111/j.1365-2958.1989.tb00247.x>

Mendiola, M.V., Jubete, Y., de la Cruz, F., 1992. DNA sequence of IS91 and identification of the transposase gene. *J. Bacteriol.* 174, 1345–1351. <https://doi.org/10.1128/jb.174.4.1345-1351.1992>

Menecier, S., Servant, P., Coste, G., Bailone, A., Sommer, S., 2006. Mutagenesis via IS transposition in *Deinococcus radiodurans*. *Mol. Microbiol.* 59, 317–325. <https://doi.org/10.1111/j.1365-2958.2005.04936.x>

Messier, N., Roy, P.H., 2001. Integron integrases possess a unique additional domain necessary for activity. *J. Bacteriol.* 183, 6699–6706. <https://doi.org/10.1128/JB.183.22.6699-6706.2001>

Messing, S.A.J., Ton-Hoang, B., Hickman, A.B., McCubbin, A.J., Peaslee, G.F., Ghirlando, R., Chandler, M., Dyda, F., 2012. The processing of repetitive extragenic palindromes: the structure of a repetitive extragenic palindrome bound to its associated nuclease. *Nucleic Acids Res.* 40, 9964–9979. <https://doi.org/10.1093/nar/gks741>

Meyer, J., Iida, S., Arber, W., 1980. Does the insertion element IS1 transpose preferentially into A+T-rich DNA segments? *Mol. Gen. Genet. MGG* 178, 471–473. <https://doi.org/10.1007/BF00270502>

Mihajlovic, S., Lang, S., Sut, M.V., Strohmaier, H., Gruber, C.J., Koraimann, G., Cabezón, E., Moncalián, G., de la Cruz, F., Zechner, E.L., 2009. Plasmid r1 conjugative DNA processing is regulated at the coupling protein interface. *J. Bacteriol.* 191, 6877–6887. <https://doi.org/10.1128/JB.00918-09>

Mira, A., Ochman, H., Moran, N.A., 2001. Deletional bias and the evolution of bacterial genomes. *Trends Genet. TIG* 17, 589–596. [https://doi.org/10.1016/s0168-9525\(01\)02447-7](https://doi.org/10.1016/s0168-9525(01)02447-7)

Miseta, A., Csutora, P., 2000. Relationship between the occurrence of cysteine in proteins and the complexity of organisms. *Mol. Biol. Evol.* 17, 1232–1239. <https://doi.org/10.1093/oxfordjournals.molbev.a026406>

Mitra, K., Schaffitzel, C., Shaikh, T., Tama, F., Jenni, S., Brooks, C.L., Ban, N., Frank, J., 2005. Structure of the *E. coli* protein-conducting channel bound to a translating ribosome. *Nature* 438, 318–324. <https://doi.org/10.1038/nature04133>

Mohapatra, S.S., Biondi, E.G., 2017. DNA Methylation in Prokaryotes: Regulation and Function, in: Krell, T. (Ed.), *Cellular Ecophysiology of Microbe, Handbook of Hydrocarbon and Lipid Microbiology*. Springer International Publishing, Cham, pp. 1–21. [https://doi.org/10.1007/978-3-319-20796-4\\_23-1](https://doi.org/10.1007/978-3-319-20796-4_23-1)

Montaño, S.P., Rice, P.A., 2011. Moving DNA around: DNA transposition and retroviral integration. *Curr. Opin. Struct. Biol.* 21, 370–378. <https://doi.org/10.1016/j.sbi.2011.03.004>

Mooney, R.A., Darst, S.A., Landick, R., 2005. Sigma and RNA polymerase: an on-again, off-again relationship? *Mol. Cell* 20, 335–345. <https://doi.org/10.1016/j.molcel.2005.10.015>

- Moran, N.A., 1996. Accelerated evolution and Muller's ratchet in endosymbiotic bacteria. *Proc. Natl. Acad. Sci. U. S. A.* 93, 2873–2878. <https://doi.org/10.1073/pnas.93.7.2873>
- Moran, N.A., Plague, G.R., 2004. Genomic changes following host restriction in bacteria. *Curr. Opin. Genet. Dev.* 14, 627–633. <https://doi.org/10.1016/j.gde.2004.09.003>
- Morse, M.L., Lederberg, E.M., Lederberg, J., 1956. Transduction in *Escherichia Coli* K-12. *Genetics* 41, 142–156. <https://doi.org/10.1093/genetics/41.1.142>
- Mugnier, P.D., Poirel, L., Nordmann, P., 2009. Functional analysis of insertion sequence ISAb1, responsible for genomic plasticity of *Acinetobacter baumannii*. *J. Bacteriol.* 191, 2414–2418. <https://doi.org/10.1128/JB.01258-08>
- Murai, N., Kamata, H., Nagashima, Y., Yagisawa, H., Hirata, H., 1995. A novel insertion sequence (IS)-like element of the thermophilic bacterium PS3 promotes expression of the alanine carrier protein-encoding gene. *Gene* 163, 103–107. [https://doi.org/10.1016/0378-1119\(95\)00384-i](https://doi.org/10.1016/0378-1119(95)00384-i)
- Murakami, K.S., Masuda, S., Campbell, E.A., Muzzin, O., Darst, S.A., 2002. Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science* 296, 1285–1290. <https://doi.org/10.1126/science.1069595>
- Murray, C.J.L., Ikuta, K.S., Sharara, F., Swetschinski, L., Aguilar, G.R., Gray, A., Han, C., Bisignano, C., Rao, P., Wool, E., Johnson, S.C., Browne, A.J., Chipeta, M.G., Fell, F., Hackett, S., Haines-Woodhouse, G., Hamadani, B.H.K., Kumaran, E.A.P., McManigal, B., Achalapong, S., Agarwal, R., Akech, S., Albertson, S., Amuasi, J., Andrews, J., Aravkin, A., Ashley, E., Babin, F.-X., Bailey, F., Baker, S., Basnyat, B., Bekker, A., Bender, R., Berkley, J.A., Bethou, A., Bielicki, J., Boonkasidecha, S., Bukosia, J., Carvalheiro, C., Castañeda-Orjuela, C., Chansamouth, V., Chaurasia, S., Chiurchiù, S., Chowdhury, F., Donatien, R.C., Cook, A.J., Cooper, B., Cressey, T.R., Criollo-Mora, E., Cunningham, M., Darboe, S., Day, N.P.J., Luca, M.D., Dokova, K., Dramowski, A., Dunachie, S.J., Bich, T.D., Eckmanns, T., Eibach, D., Emami, A., Feasey, N., Fisher-Pearson, N., Forrest, K., Garcia, C., Garrett, D., Gastmeier, P., Giref, A.Z., Greer, R.C., Gupta, V., Haller, S., Haselbeck, A., Hay, S.I., Holm, M., Hopkins, S., Hsia, Y., Iregbu, K.C., Jacobs, J., Jarovsky, D., Javanmardi, F., Jenney, A.W.J., Khorana, M., Khusuwan, S., Kissoon, N., Kobeissi, E., Kostyanev, T., Krapp, F., Krumkamp, R., Kumar, A., Kyu, H.H., Lim, C., Lim, K., Limmathurotsakul, D., Loftus, M.J., Lunn, M., Ma, J., Manoharan, A., Marks, F., May, J., Mayxay, M., Mturi, N., Munera-Huertas, T., Musicha, P., Musila, L.A., Mussi-Pinhata, M.M., Naidu, R.N., Nakamura, T., Nanavati, R., Nangia, S., Newton, P., Ngoun, C., Novotney, A., Nwakanma, D., Obiero, C.W., Ochoa, T.J., Olivas-Martinez, A., Olliaro, P., Ooko, E., Ortiz-Brizuela, E., Ounchanum, P., Pak, G.D., Paredes, J.L., Peleg, A.Y., Perrone, C., Phe, T., Phommasone, K., Plakkal, N., Ponce-de-Leon, A., Raad, M., Ramdin, T., Rattanavong, S., Riddell, A., Roberts, T., Robotham, J.V., Roca, A., Rosenthal, V.D., Rudd, K.E., Russell, N., Sader, H.S., Saengchan, W., Schnall, J., Scott, J.A.G., Seekaew, S., Sharland, M., Shivamallappa, M., Sifuentes-Osorio, J., Simpson, A.J., Steenkeste, N., Stewardson, A.J., Stoeva, T., Tasak, N., Thaiprakong, A., Thwaites, G., Tigo, C., Turner, C., Turner, P., Doorn, H.R. van, Velaphi, S., Vongpradith, A., Vongsouvath, M., Vu, H., Walsh, T., Walson, J.L., Waner, S., Wangrangsimakul, T., Wannapini, P., Wozniak, T., Sharma, T.E.M.W.Y., Yu, K.C., Zheng, P., Sartorius, B., Lopez, A.D., Stergachis, A., Moore, C., Dolecek, C., Naghavi, M., 2022. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *The Lancet* 399, 629–655. [https://doi.org/10.1016/S0140-6736\(21\)02724-0](https://doi.org/10.1016/S0140-6736(21)02724-0)
- Naas, T., Poirel, L., Nordmann, P., 1999. Molecular characterisation of In51, a class 1 integron containing a novel aminoglycoside adenyltransferase gene cassette, *aadA6*, in *Pseudomonas aeruginosa*. *Biochim. Biophys. Acta* 1489, 445–451. [https://doi.org/10.1016/s0167-4781\(99\)00202-x](https://doi.org/10.1016/s0167-4781(99)00202-x)
- Nagel, M., Reuter, T., Jansen, A., Szekat, C., Bierbaum, G., 2011. Influence of ciprofloxacin and vancomycin on mutation rate and transposition of IS256 in *Staphylococcus aureus*. *Int. J. Med. Microbiol. IJMM* 301, 229–236. <https://doi.org/10.1016/j.ijmm.2010.08.021>
- Nagy, Z., Chandler, M., 2004. Regulation of transposition in bacteria. *Res. Microbiol., Genome plasticity and the evolution of microbial genomes* 155, 387–398. <https://doi.org/10.1016/j.resmic.2004.01.008>
- Nazarian, P., Tran, F., Boedicker, J.Q., 2018. Modeling Multispecies Gene Flow Dynamics Reveals the Unique Roles of Different Horizontal Gene Transfer Mechanisms. *Front. Microbiol.* 9, 2978. <https://doi.org/10.3389/fmicb.2018.02978>
- Nojiri, H., Shintani, M., Omori, T., 2004. Divergence of mobile genetic elements involved in the distribution of xenobiotic-catabolic capacity. *Appl. Microbiol. Biotechnol.* 64, 154–174. <https://doi.org/10.1007/s00253-003-1509-y>
- Normand, C., Duval-Valentin, G., Haren, L., Chandler, M., 2001. The terminal inverted repeats of IS911: requirements for synaptic complex assembly and activity. *J. Mol. Biol.* 308, 853–871. <https://doi.org/10.1006/jmbi.2001.4641>

- Novick, R.P., 1989. Staphylococcal plasmids and their replication. *Annu. Rev. Microbiol.* 43, 537–565. <https://doi.org/10.1146/annurev.mi.43.100189.002541>
- Novick, R.P., 1987. Plasmid incompatibility. *Microbiol. Rev.* 51, 381–395. <https://doi.org/10.1128/mr.51.4.381-395.1987>
- Novick, R.P., Clowes, R.C., Cohen, S.N., Curtiss, R., Datta, N., Falkow, S., 1976. Uniform nomenclature for bacterial plasmids: a proposal. *Bacteriol. Rev.* 40, 168–189. <https://doi.org/10.1128/br.40.1.168-189.1976>
- Nowotny, M., 2009. Retroviral integrase superfamily: the structural perspective. *EMBO Rep.* 10, 144–151. <https://doi.org/10.1038/embor.2008.256>
- Obeng, N., Pratama, A.A., Elsas, J.D. van, 2016. The Significance of Mutualistic Phages for Bacterial Ecology and Evolution. *Trends Microbiol.* 24, 440–449. <https://doi.org/10.1016/j.tim.2015.12.009>
- Odegrip, R., Haggård-Ljungquist, E., 2001. The two active-site tyrosine residues of the  $\alpha$  protein play non-equivalent roles during initiation of rolling circle replication of bacteriophage  $\phi$ 2. *J. Mol. Biol.* 308, 147–163. <https://doi.org/10.1006/jmbi.2001.4607>
- Odegrip, R., Schoen, S., Haggård-Ljungquist, E., Park, K., Chattoraj, D.K., 2000. The interaction of bacteriophage P2 B protein with *Escherichia coli* DnaB helicase. *J. Virol.* 74, 4057–4063. <https://doi.org/10.1128/jvi.74.9.4057-4063.2000>
- O'Donoghue, E.J., Krachler, A.M., 2016. Mechanisms of outer membrane vesicle entry into host cells. *Cell. Microbiol.* 18, 1508–1517. <https://doi.org/10.1111/cmi.12655>
- O'Hara, J.A., McGann, P., Snesrud, E.C., Clifford, R.J., Waterman, P.E., Lesho, E.P., Doi, Y., 2013. Novel 16S rRNA methyltransferase RmtH produced by *Klebsiella pneumoniae* associated with war-related trauma. *Antimicrob. Agents Chemother.* 57, 2413–2416. <https://doi.org/10.1128/AAC.00266-13>
- Ohta, S., Yoshimura, E., Ohtsubo, E., 2004. Involvement of two domains with helix-turn-helix and zinc finger motifs in the binding of IS1 transposase to terminal inverted repeats. *Mol. Microbiol.* 53, 193–202. <https://doi.org/10.1111/j.1365-2958.2004.04103.x>
- Olasz, F., Farkas, T., Kiss, J., Arini, A., Arber, W., 1997. Terminal inverted repeats of insertion sequence IS30 serve as targets for transposition. *J. Bacteriol.* 179, 7551–7558. <https://doi.org/10.1128/jb.179.23.7551-7558.1997>
- Olliver, A., Vallé, M., Chaslus-Dancla, E., Cloeckaert, A., 2005. Overexpression of the multidrug efflux operon *acrEF* by insertional activation with IS1 or IS10 elements in *Salmonella enterica* serovar typhimurium DT204 *acrB* mutants selected with fluoroquinolones. *Antimicrob. Agents Chemother.* 49, 289–301. <https://doi.org/10.1128/AAC.49.1.289-301.2005>
- O'Neill, J., 2016. Tackling drug-resistant infections globally: final report and recommendations (Report). Government of the United Kingdom.
- O'Neill, J., 2014. Antimicrobial resistance : tackling a crisis for the health and wealth of nations / the Review on Antimicrobial Resistance chaired by Jim O'Neill. [WWW Document]. Wellcome Collect. URL <https://wellcomecollection.org/works/rdpck35v> (accessed 10.11.23).
- Orench-Rivera, N., Kuehn, M.J., 2016. Environmentally controlled bacterial vesicle-mediated export. *Cell. Microbiol.* 18, 1525–1536. <https://doi.org/10.1111/cmi.12676>
- Orgel, L.E., Crick, F.H., 1980. Selfish DNA: the ultimate parasite. *Nature* 284, 604–607. <https://doi.org/10.1038/284604a0>
- Orozco, B.M., Hanley-Bowdoin, L., 1996. A DNA structure is required for geminivirus replication origin function. *J. Virol.* 70, 148–158. <https://doi.org/10.1128/JVI.70.1.148-158.1996>
- Osborn, A.M., Böltner, D., 2002. When phage, plasmids, and transposons collide: genomic islands, and conjugative- and mobilizable-transposons as a mosaic continuum. *Plasmid* 48, 202–212. [https://doi.org/10.1016/s0147-619x\(02\)00117-8](https://doi.org/10.1016/s0147-619x(02)00117-8)
- Pachori, P., Gothwal, R., Gandhi, P., 2019. Emergence of antibiotic resistance *Pseudomonas aeruginosa* in intensive care unit; a critical review. *Genes Dis.* 6, 109–119. <https://doi.org/10.1016/j.gendis.2019.04.001>
- Padjasek, M., Kocyla, A., Kluska, K., Kerber, O., Tran, J.B., Krężel, A., 2020. Structural zinc binding sites shaped for greater works: Structure-function relations in classical zinc finger, hook and clasp domains. *J. Inorg. Biochem.* 204, 110955. <https://doi.org/10.1016/j.jinorgbio.2019.110955>
- Paget, M.S.B., Helmann, J.D., 2003. The sigma70 family of sigma factors. *Genome Biol.* 4, 203.

<https://doi.org/10.1186/gb-2003-4-1-203>

Pal, C., Bengtsson-Palme, J., Kristiansson, E., Larsson, D.G.J., 2016. The structure and diversity of human, animal and environmental resistomes. *Microbiome* 4, 54. <https://doi.org/10.1186/s40168-016-0199-5>

Pal, C., Bengtsson-Palme, J., Kristiansson, E., Larsson, D.G.J., 2015. Co-occurrence of resistance genes to antibiotics, biocides and metals reveals novel insights into their co-selection potential. *BMC Genomics* 16, 964. <https://doi.org/10.1186/s12864-015-2153-5>

Paquola, A.C.M., Asif, H., Pereira, C.A. de B., Feltus, B.C., Bonatto, D., Lima, W.C., Menck, C.F.M., 2018. Horizontal Gene Transfer Building Prokaryote Genomes: Genes Related to Exchange Between Cell and Environment are Frequently Transferred. *J. Mol. Evol.* 86, 190–203. <https://doi.org/10.1007/s00239-018-9836-x>

Parkhill, J., Sebaihia, M., Preston, A., Murphy, L.D., Thomson, N., Harris, D.E., Holden, M.T.G., Churcher, C.M., Bentley, S.D., Mungall, K.L., Cerdeño-Tárraga, A.M., Temple, L., James, K., Harris, B., Quail, M.A., Achtman, M., Atkin, R., Baker, S., Basham, D., Bason, N., Cherevach, I., Chillingworth, T., Collins, M., Cronin, A., Davis, P., Doggett, J., Feltwell, T., Goble, A., Hamlin, N., Hauser, H., Holroyd, S., Jagels, K., Leather, S., Moule, S., Norberczak, H., O'Neil, S., Ormond, D., Price, C., Rabinowitsch, E., Rutter, S., Sanders, M., Saunders, D., Seeger, K., Sharp, S., Simmonds, M., Skelton, J., Squares, R., Squares, S., Stevens, K., Unwin, L., Whitehead, S., Barrell, B.G., Maskell, D.J., 2003. Comparative analysis of the genome sequences of *Bordetella pertussis*, *Bordetella parapertussis* and *Bordetella bronchiseptica*. *Nat. Genet.* 35, 32–40. <https://doi.org/10.1038/ng1227>

Partridge, S.R., 2011. Analysis of antibiotic resistance regions in Gram-negative bacteria. *FEMS Microbiol. Rev.* 35, 820–855. <https://doi.org/10.1111/j.1574-6976.2011.00277.x>

Partridge, S.R., Hall, R.M., 2005. Evolution of transposons containing *bla*<sub>TEM</sub> genes. *Antimicrob. Agents Chemother.* 49, 1267–1268. <https://doi.org/10.1128/AAC.49.3.1267-1268.2005>

Partridge, S.R., Hall, R.M., 2003a. The IS1111 family members IS4321 and IS5075 have subterminal inverted repeats and target the terminal inverted repeats of Tn21 family transposons. *J. Bacteriol.* 185, 6371–6384. <https://doi.org/10.1128/JB.185.21.6371-6384.2003>

Partridge, S.R., Hall, R.M., 2003b. In34, a complex In5 family class 1 integron containing *orf513* and *dfrA10*. *Antimicrob. Agents Chemother.* 47, 342–349. <https://doi.org/10.1128/AAC.47.1.342-349.2003>

Partridge, S.R., Kwong, S.M., Firth, N., Jensen, S.O., 2018. Mobile Genetic Elements Associated with Antimicrobial Resistance. *Clin. Microbiol. Rev.* 31, e00088-17. <https://doi.org/10.1128/CMR.00088-17>

Partridge, S.R., Recchia, G.D., Scaramuzzi, C., Collis, C.M., Stokes, H.W., Hall, R.M., 2000. Definition of the *attI1* site of class 1 integrons. *Microbiol. Read. Engl.* 146 ( Pt 11), 2855–2864. <https://doi.org/10.1099/00221287-146-11-2855>

Pasternak, C., Dulermo, R., Ton-Hoang, B., Debuchy, R., Siguier, P., Coste, G., Chandler, M., Sommer, S., 2013. ISDra2 transposition in *Deinococcus radiodurans* is downregulated by TnpB. *Mol. Microbiol.* 88, 443–455. <https://doi.org/10.1111/mmi.12194>

Pasternak, C., Ton-Hoang, B., Coste, G., Bailone, A., Chandler, M., Sommer, S., 2010. Irradiation-induced *Deinococcus radiodurans* genome fragmentation triggers transposition of a single resident insertion sequence. *PLoS Genet.* 6, e1000799. <https://doi.org/10.1371/journal.pgen.1000799>

Pato, M.L., Reich, C., 1984. Stoichiometric use of the transposase of bacteriophage Mu. *Cell* 36, 197–202. [https://doi.org/10.1016/0092-8674\(84\)90089-8](https://doi.org/10.1016/0092-8674(84)90089-8)

Paul, J.H., 2008. Prophages in marine bacteria: dangerous molecular time bombs or the key to survival in the seas? *ISME J.* 2, 579–589. <https://doi.org/10.1038/ismej.2008.35>

Pereira, J., 2021. GCsnap: Interactive Snapshots for the Comparison of Protein-Coding Genomic Contexts. *J. Mol. Biol.* 433, 166943. <https://doi.org/10.1016/j.jmb.2021.166943>

Pérez-Cruz, C., Carrión, O., Delgado, L., Martínez, G., López-Iglesias, C., Mercade, E., 2013. New type of outer membrane vesicle produced by the Gram-negative bacterium *Shewanella vesiculosa* M7T: implications for DNA content. *Appl. Environ. Microbiol.* 79, 1874–1881. <https://doi.org/10.1128/AEM.03657-12>

Perkins-Balding, D., Duval-Valentin, G., Glasgow, A.C., 1999. Excision of IS492 requires flanking target sequences and results in circle formation in *Pseudoalteromonas atlantica*. *J. Bacteriol.* 181, 4937–4948. <https://doi.org/10.1128/JB.181.16.4937-4948.1999>

Peters, J.E., Craig, N.L., 2001a. Tn7: smarter than we thought. *Nat. Rev. Mol. Cell Biol.* 2, 806–814. <https://doi.org/10.1038/35099006>

- Peters, J.E., Craig, N.L., 2001b. Tn7 recognizes transposition target structures associated with DNA replication using the DNA-binding protein TnsE. *Genes Dev.* 15, 737–747. <https://doi.org/10.1101/gad.870201>
- Pfeifer, F., Blaseio, U., 1990. Transposition burst of the ISH27 insertion element family in *Halobacterium halobium*. *Nucleic Acids Res.* 18, 6921–6925. <https://doi.org/10.1093/nar/18.23.6921>
- Plague, G.R., Dunbar, H.E., Tran, P.L., Moran, N.A., 2008. Extensive proliferation of transposable elements in heritable bacterial symbionts. *J. Bacteriol.* 190, 777–779. <https://doi.org/10.1128/JB.01082-07>
- Plikaytis, B.B., Crawford, J.T., Shinnick, T.M., 1998. IS1549 from *Mycobacterium smegmatis* forms long direct repeats upon insertion. *J. Bacteriol.* 180, 1037–1043. <https://doi.org/10.1128/JB.180.5.1037-1043.1998>
- Poirel, L., Carrère, A., Pitout, J.D., Nordmann, P., 2009. Integron mobilization unit as a source of mobility of antibiotic resistance genes. *Antimicrob. Agents Chemother.* 53, 2492–2498. <https://doi.org/10.1128/AAC.00033-09>
- Poirel, L., Decousser, J.-W., Nordmann, P., 2003. Insertion sequence ISEcp1B is involved in expression and mobilization of a bla(CTX-M) beta-lactamase gene. *Antimicrob. Agents Chemother.* 47, 2938–2945. <https://doi.org/10.1128/AAC.47.9.2938-2945.2003>
- Post, V., Hall, R.M., 2009. Insertion sequences in the IS1111 family that target the attC recombination sites of integron-associated gene cassettes. *FEMS Microbiol. Lett.* 290, 182–187. <https://doi.org/10.1111/j.1574-6968.2008.01412.x>
- Pradeep Ram, A.S., Sime-Ngando, T., 2010. Resources drive trade-off between viral lifestyles in the plankton: evidence from freshwater microbial microcosms. *Environ. Microbiol.* 12, 467–479. <https://doi.org/10.1111/j.1462-2920.2009.02088.x>
- Prère, M.-F., Canal, I., Wills, N.M., Atkins, J.F., Fayet, O., 2011. The interplay of mRNA stimulatory signals required for AUU-mediated initiation and programmed -1 ribosomal frameshifting in decoding of transposable element IS911. *J. Bacteriol.* 193, 2735–2744. <https://doi.org/10.1128/JB.00115-11>
- Prescott, J.F., 2014. The resistance tsunami, antimicrobial stewardship, and the golden age of microbiology. *Vet. Microbiol.* 171, 273–278. <https://doi.org/10.1016/j.vetmic.2014.02.035>
- Preston, A., Parkhill, J., Maskell, D.J., 2004. The bordetellae: lessons from genomics. *Nat. Rev. Microbiol.* 2, 379–390. <https://doi.org/10.1038/nrmicro886>
- Pritham, E.J., Feschotte, C., 2007. Massive amplification of rolling-circle transposons in the lineage of the bat *Myotis lucifugus*. *Proc. Natl. Acad. Sci. U. S. A.* 104, 1895–1900. <https://doi.org/10.1073/pnas.0609601104>
- Pritham, E.J., Putliwala, T., Feschotte, C., 2007. Mavericks, a novel class of giant transposable elements widespread in eukaryotes and related to DNA viruses. *Gene* 390, 3–17. <https://doi.org/10.1016/j.gene.2006.08.008>
- Proshkin, S., Rahmouni, A.R., Mironov, A., Nudler, E., 2010. Cooperation between translating ribosomes and RNA polymerase in transcription elongation. *Science* 328, 504–508. <https://doi.org/10.1126/science.1184939>
- Prudhomme M, Attaiech L, Sanchez G, Martin B, Claverys JP. Antibiotic stress induces genetic transformability in the human pathogen *Streptococcus pneumoniae*. *Science*. 2006 Jul 7;313(5783):89-92. doi: 10.1126/science.1127912.
- Quiroga, M.P., Arduino, S.M., Merkier, A.K., Quiroga, C., Petroni, A., Argentinian Integron Study Group, Roy, P.H., Centrón, D., 2013. Distribution and functional identification of complex class 1 integrons. *Infect. Genet. Evol. J. Mol. Epidemiol. Evol. Genet. Infect. Dis.* 19, 88–96. <https://doi.org/10.1016/j.meegid.2013.06.029>
- Rådström, P., Swedberg, G., 1988. RSF1010 and a conjugative plasmid contain sullI, one of two known genes for plasmid-borne sulfonamide resistance dihydropteroate synthase. *Antimicrob. Agents Chemother.* 32, 1684–1692. <https://doi.org/10.1128/AAC.32.11.1684>
- Rajeev, L., Malanowska, K., Gardner, J.F., 2009. Challenging a paradigm: the role of DNA homology in tyrosine recombinase reactions. *Microbiol. Mol. Biol. Rev. MMBR* 73, 300–309. <https://doi.org/10.1128/MMBR.00038-08>
- Rak, B., Lusky, M., Hable, M., 1982. Expression of two proteins from overlapping and oppositely oriented genes on transposable DNA insertion element IS5. *Nature* 297, 124–128. <https://doi.org/10.1038/297124a0>
- Ramsay, J.P., Sullivan, J.T., Stuart, G.S., Lamont, I.L., Ronson, C.W., 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. *Mol. Microbiol.* 62, 723–734. <https://doi.org/10.1111/j.1365-2958.2006.05396.x>
- Rankin, D.J., Bichsel, M., Wagner, A., 2010. Mobile DNA can drive lineage extinction in prokaryotic populations. *J. Evol. Biol.* 23, 2422–2431. <https://doi.org/10.1111/j.1420-9101.2010.02106.x>

- Rankin, D.J., Rocha, E.P.C., Brown, S.P., 2011. What traits are carried on mobile genetic elements, and why? *Heredity* 106, 1–10. <https://doi.org/10.1038/hdy.2010.24>
- Rauch, P.J., De Vos, W.M., 1992. Characterization of the novel nisin-sucrose conjugative transposon Tn5276 and its insertion in *Lactococcus lactis*. *J. Bacteriol.* 174, 1280–1287. <https://doi.org/10.1128/jb.174.4.1280-1287.1992>
- Redondo-Salvo, S., Fernández-López, R., Ruiz, R., Vielva, L., de Toro, M., Rocha, E.P.C., Garcillán-Barcia, M.P., de la Cruz, F., 2020. Pathways for horizontal gene transfer in bacteria revealed by a global map of their plasmids. *Nat. Commun.* 11, 3602. <https://doi.org/10.1038/s41467-020-17278-2>
- Reif, H.J., Saedler, H., 1975. IS1 is involved in deletion formation in the gal region of *E. coli* K12. *Mol. Gen. Genet.* MGG 137, 17–28. <https://doi.org/10.1007/BF00332538>
- Reimann, C., Haas, D., 1987. Mode of replicon fusion mediated by the duplicated insertion sequence IS21 in *Escherichia coli*. *Genetics* 115, 619–625. <https://doi.org/10.1093/genetics/115.4.619>
- Reimann, C., Moore, R., Little, S., Savioz, A., Willetts, N.S., Haas, D., 1989. Genetic structure, function and regulation of the transposable element IS21. *Mol. Gen. Genet.* MGG 215, 416–424. <https://doi.org/10.1007/BF00427038>
- Reisenauer, A., Shapiro, L., 2002. DNA methylation affects the cell cycle transcription of the CtrA global regulator in *Caulobacter*. *EMBO J.* 21, 4969–4977. <https://doi.org/10.1093/emboj/cdf490>
- Renelli, M., Matias, V., Lo, R.Y., Beveridge, T.J., 2004. DNA-containing membrane vesicles of *Pseudomonas aeruginosa* PAO1 and their genetic transformation potential. *Microbiol. Read. Engl.* 150, 2161–2169. <https://doi.org/10.1099/mic.0.26841-0>
- Reynolds, P.E., 1989. Structure, biochemistry and mechanism of action of glycopeptide antibiotics. *Eur. J. Clin. Microbiol. Infect. Dis. Off. Publ. Eur. Soc. Clin. Microbiol.* 8, 943–950. <https://doi.org/10.1007/BF01967563>
- Reznikoff, W.S., 1993. The Tn5 transposon. *Annu. Rev. Microbiol.* 47, 945–963. <https://doi.org/10.1146/annurev.mi.47.100193.004501>
- Rhodes, G., Huys, G., Swings, J., McGann, P., Hiney, M., Smith, P., Pickup, R.W., 2000. Distribution of oxytetracycline resistance plasmids between aeromonads in hospital and aquaculture environments: implication of Tn1721 in dissemination of the tetracycline resistance determinant tet A. *Appl. Environ. Microbiol.* 66, 3883–3890. <https://doi.org/10.1128/AEM.66.9.3883-3890.2000>
- Rice, P.A., Baker, T.A., 2001. Comparative architecture of transposase and integrase complexes. *Nat. Struct. Biol.* 8, 302–307. <https://doi.org/10.1038/86166>
- Richardson, J.P., 2002. Rho-dependent termination and ATPases in transcript termination. *Biochim. Biophys. Acta* 1577, 251–260. [https://doi.org/10.1016/s0167-4781\(02\)00456-6](https://doi.org/10.1016/s0167-4781(02)00456-6)
- Richter, G.Y., Björklöf, K., Romantschuk, M., Mills, D., 1998. Insertion specificity and trans-activation of IS801. *Mol. Gen. Genet.* MGG 260, 381–387. <https://doi.org/10.1007/s004380050907>
- Rinkel, M., Hubert, J.C., Roux, B., Lett, M.C., 1994. Identification of a new transposon Tn5403 in a *Klebsiella pneumoniae* strain isolated from a polluted aquatic environment. *Curr. Microbiol.* 29, 249–254. <https://doi.org/10.1007/BF01577436>
- Roberts, A.P., Chandler, M., Courvalin, P., Guédon, G., Mullany, P., Pembroke, T., Rood, J.I., Smith, C.J., Summers, A.O., Tsuda, M., Berg, D.E., 2008. Revised nomenclature for transposable genetic elements. *Plasmid* 60, 167–173. <https://doi.org/10.1016/j.plasmid.2008.08.001>
- Roberts, A.P., Davis, I.J., Seville, L., Villedieu, A., Mullany, P., 2006. Characterization of the ends and target site of a novel tetracycline resistance-encoding conjugative transposon from *Enterococcus faecium* 664.1H1. *J. Bacteriol.* 188, 4356–4361. <https://doi.org/10.1128/JB.00129-06>
- Roberts, A.P., Johanesen, P.A., Lyras, D., Mullany, P., Rood, J.I., 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. *Microbiol. Read. Engl.* 147, 1243–1251. <https://doi.org/10.1099/00221287-147-5-1243>
- Roberts, A.P., Mullany, P., 2009. A modular master on the move: the Tn916 family of mobile genetic elements. *Trends Microbiol.* 17, 251–258. <https://doi.org/10.1016/j.tim.2009.03.002>
- Roberts, D., Hoopes, B.C., McClure, W.R., Kleckner, N., 1985. IS10 transposition is regulated by DNA adenine methylation. *Cell* 43, 117–130. [https://doi.org/10.1016/0092-8674\(85\)90017-0](https://doi.org/10.1016/0092-8674(85)90017-0)



- Roche, D., Fléchar, M., Lallier, N., Répérant, M., Brée, A., Pascal, G., Schouler, C., Germon, P., 2010. ICEEc2, a new integrative and conjugative element belonging to the pKLC102/PAGI-2 family, identified in *Escherichia coli* strain BEN374. *J. Bacteriol.* 192, 5026–5036. <https://doi.org/10.1128/JB.00609-10>
- Rodriguez-Martinez, J.-M., Poirel, L., Canton, R., Nordmann, P., 2006. Common region CR1 for expression of antibiotic resistance genes. *Antimicrob. Agents Chemother.* 50, 2544–2546. <https://doi.org/10.1128/AAC.00609-05>
- Roier, S., Zingl, F.G., Cakar, F., Schild, S., 2016. Bacterial outer membrane vesicle biogenesis: a new mechanism and its implications. *Microb. Cell Graz Austria* 3, 257–259. <https://doi.org/10.15698/mic2016.06.508>
- Romantschuk, M., Richter, G.Y., Mukhopadhyay, P., Mills, D., 1991. IS801, an insertion sequence element isolated from *Pseudomonas syringae* pathovar phaseolicola. *Mol. Microbiol.* 5, 617–622. <https://doi.org/10.1111/j.1365-2958.1991.tb00732.x>
- Ronning, D.R., Guynet, C., Ton-Hoang, B., Perez, Z.N., Ghirlando, R., Chandler, M., Dyda, F., 2005. Active site sharing and subterminal hairpin recognition in a new class of DNA transposases. *Mol. Cell* 20, 143–154. <https://doi.org/10.1016/j.molcel.2005.07.026>
- Rosario, K., Duffy, S., Breitbart, M., 2012. A field guide to eukaryotic circular single-stranded DNA viruses: insights gained from metagenomics. *Arch. Virol.* 157, 1851–1871. <https://doi.org/10.1007/s00705-012-1391-y>
- Ross, J.A., Ellis, M.J., Hossain, S., Haniford, D.B., 2013. Hfq restructures RNA-IN and RNA-OUT and facilitates antisense pairing in the Tn10/IS10 system. *RNA N. Y. N* 19, 670–684. <https://doi.org/10.1261/rna.037747.112>
- Ross, K., Varani, A.M., Snesrud, E., Huang, H., Alvarenga, D.O., Zhang, J., Wu, C., McGann, P., Chandler, M., 2021. TnCentral: a Prokaryotic Transposable Element Database and Web Portal for Transposon Analysis. *mBio* 12, e0206021. <https://doi.org/10.1128/mBio.02060-21>
- Ross, W., Ernst, A., Gourse, R.L., 2001. Fine structure of *E. coli* RNA polymerase-promoter interactions: alpha subunit binding to the UP element minor groove. *Genes Dev.* 15, 491–506. <https://doi.org/10.1101/gad.870001>
- Rousseau, P., Gueguen, E., Duval-Valentin, G., Chandler, M., 2004. The helix-turn-helix motif of bacterial insertion sequence IS911 transposase is required for DNA binding. *Nucleic Acids Res.* 32, 1335–1344. <https://doi.org/10.1093/nar/gkh276>
- Rowe-Magnus, D.A., Guérout, A.M., Mazel, D., 1999. Super-integrans. *Res. Microbiol.* 150, 641–651. [https://doi.org/10.1016/s0923-2508\(99\)00127-8](https://doi.org/10.1016/s0923-2508(99)00127-8)
- Roychowdhury, T., Mandal, S., Bhattacharya, A., 2015. Analysis of IS6110 insertion sites provide a glimpse into genome evolution of *Mycobacterium tuberculosis*. *Sci. Rep.* 5, 12567. <https://doi.org/10.1038/srep12567>
- Rozental, T., Mascarenhas, L.F., Rozenbaum, R., Gomes, R., Mattos, G.S., Magno, C.C., Almeida, D.N., Rossi, M.I.D., Favacho, A.R.M., de Lemos, E.R.S., 2012. Coxiella burnetii, the agent of Q fever in Brazil: its hidden role in seronegative arthritis and the importance of molecular diagnosis based on the repetitive element IS1111 associated with the transposase gene. *Mem. Inst. Oswaldo Cruz* 107, 695–697. <https://doi.org/10.1590/s0074-02762012000500021>
- Rozhon, W., Petutschnig, E., Khan, M., Summers, D.K., Poppenberger, B., 2010. Frequency and diversity of small cryptic plasmids in the genus *Rahnella*. *BMC Microbiol.* 10, 56. <https://doi.org/10.1186/1471-2180-10-56>
- Rudant, E., Courvalin, P., Lambert, T., 1998. Characterization of IS18, an element capable of activating the silent *aac(6′)-Ij* gene of *Acinetobacter* sp. 13 strain BM2716 by transposition. *Antimicrob. Agents Chemother.* 42, 2759–2761. <https://doi.org/10.1128/AAC.42.10.2759>
- Ruiz-Masó, J.A., Lurz, R., Espinosa, M., del Solar, G., 2007. Interactions between the RepB initiator protein of plasmid pMV158 and two distant DNA regions within the origin of replication. *Nucleic Acids Res.* 35, 1230–1244. <https://doi.org/10.1093/nar/gkl1099>
- Russo, L., Palmieri, M., Caso, J.V., D’Abrosca, G., Diana, D., Malgieri, G., Baglivo, I., Isernia, C., Pedone, P.V., Fattorusso, R., 2015. Towards understanding the molecular recognition process in prokaryotic zinc-finger domain. *Eur. J. Med. Chem.* 91, 100–108. <https://doi.org/10.1016/j.ejmech.2014.09.040>
- Sabaté, M., Navarro, F., Miró, E., Campoy, S., Mirelis, B., Barbé, J., Prats, G., 2002. Novel complex *sul1*-type integron in *Escherichia coli* carrying *bla*(CTX-M-9). *Antimicrob. Agents Chemother.* 46, 2656–2661. <https://doi.org/10.1128/AAC.46.8.2656-2661.2002>
- Sacerdot, C., Fayat, G., Dessen, P., Springer, M., Plumbridge, J.A., Grunberg-Manago, M., Blanquet, S., 1982. Sequence of a 1.26-kb DNA fragment containing the structural gene for *E. coli* initiation factor IF3: presence of an AUU initiator codon. *EMBO J.* 1, 311–315. <https://doi.org/10.1002/j.1460-2075.1982.tb01166.x>

- Sanderson, A., Mitchell, J.E., Minchin, S.D., Busby, S.J.W., 2003. Substitutions in the *Escherichia coli* RNA polymerase sigma70 factor that affect recognition of extended -10 elements at promoters. *FEBS Lett.* 544, 199–205. [https://doi.org/10.1016/s0014-5793\(03\)00500-3](https://doi.org/10.1016/s0014-5793(03)00500-3)
- Sasakawa, C., Lowe, J.B., McDivitt, L., Berg, D.E., 1982. Control of transposon Tn5 transposition in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* 79, 7450–7454. <https://doi.org/10.1073/pnas.79.23.7450>
- Sasakawa, C., Uno, Y., Yoshikawa, M., 1981. The requirement for both DNA polymerase and 5' to 3' exonuclease activities of DNA polymerase I during Tn5 transposition. *Mol. Gen. Genet.* MGG 182, 19–24. <https://doi.org/10.1007/BF00422761>
- Schaefer, M.R., Kahn, K., 1998. Cyanobacterial transposons Tn5469 and Tn5541 represent a novel noncomposite transposon family. *J. Bacteriol.* 180, 6059–6063. <https://doi.org/10.1128/JB.180.22.6059-6063.1998>
- Schlör, S., Riedl, S., Blass, J., Reidl, J., 2000. Genetic rearrangements of the regions adjacent to genes encoding heat-labile enterotoxins (eltAB) of enterotoxigenic *Escherichia coli* strains. *Appl. Environ. Microbiol.* 66, 352–358. <https://doi.org/10.1128/AEM.66.1.352-358.2000>
- Schmieger, H., 1982. Packaging signals for phage P22 on the chromosome of *Salmonella typhimurium*. *Mol. Gen. Genet.* MGG 187, 516–518. <https://doi.org/10.1007/BF00332637>
- Schwechheimer, C., Rodriguez, D.L., Kuehn, M.J., 2015. NlpI-mediated modulation of outer membrane vesicle production through peptidoglycan dynamics in *Escherichia coli*. *MicrobiologyOpen* 4, 375–389. <https://doi.org/10.1002/mbo3.244>
- Sekine, Y., Ohtsubo, E., 1989. Frameshifting is required for production of the transposase encoded by insertion sequence 1. *Proc. Natl. Acad. Sci. U. S. A.* 86, 4609–4613. <https://doi.org/10.1073/pnas.86.12.4609>
- Sengstag, C., Iida, S., Hiestand-Nauer, R., Arber, W., 1986. Terminal inverted repeats of prokaryotic transposable element IS186 which can generate duplications of variable length at an identical target sequence. *Gene* 49, 153–156. [https://doi.org/10.1016/0378-1119\(86\)90395-1](https://doi.org/10.1016/0378-1119(86)90395-1)
- Seputiene, V., Povilonis, J., Suziedeliene, E., 2012. Novel variants of AbaR resistance islands with a common backbone in *Acinetobacter baumannii* isolates of European clone II. *Antimicrob. Agents Chemother.* 56, 1969–1973. <https://doi.org/10.1128/AAC.05678-11>
- Sharkey, L.K.R., Edwards, T.A., O'Neill, A.J., 2016. ABC-F Proteins Mediate Antibiotic Resistance through Ribosomal Protection. *mBio* 7, e01975. <https://doi.org/10.1128/mBio.01975-15>
- Sharma, V., Firth, A.E., Antonov, I., Fayet, O., Atkins, J.F., Borodovsky, M., Baranov, P.V., 2011. A pilot study of bacterial genes with disrupted ORFs reveals a surprising profusion of protein sequence recoding mediated by ribosomal frameshifting and transcriptional realignment. *Mol. Biol. Evol.* 28, 3195–3211. <https://doi.org/10.1093/molbev/msr155>
- Sharma, V., Prère, M.-F., Canal, I., Firth, A.E., Atkins, J.F., Baranov, P.V., Fayet, O., 2014. Analysis of tetra- and hepta-nucleotides motifs promoting -1 ribosomal frameshifting in *Escherichia coli*. *Nucleic Acids Res.* 42, 7210–7225. <https://doi.org/10.1093/nar/gku386>
- Shiga, Y., Sekine, Y., Ohtsubo, E., 1999. Transposition of IS1 circles. *Genes Cells Devoted Mol. Cell. Mech.* 4, 551–561. <https://doi.org/10.1046/j.1365-2443.1999.00281.x>
- Shigemoto, N., Kuwahara, R., Kayama, S., Shimizu, W., Onodera, M., Yokozaki, M., Hisatsune, J., Kato, F., Ohge, H., Sugai, M., 2012. Emergence in Japan of an imipenem-susceptible, meropenem-resistant *Klebsiella pneumoniae* carrying blaIMP-6. *Diagn. Microbiol. Infect. Dis.* 72, 109–112. <https://doi.org/10.1016/j.diagmicrobio.2011.09.019>
- Shine, J., Dalgarno, L., 1975. Determinant of cistron specificity in bacterial ribosomes. *Nature* 254, 34–38. <https://doi.org/10.1038/254034a0>
- Shine, J., Dalgarno, L., 1974. The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proc. Natl. Acad. Sci. U. S. A.* 71, 1342–1346. <https://doi.org/10.1073/pnas.71.4.1342>
- Shultzaberger, R.K., Chen, Z., Lewis, K.A., Schneider, T.D., 2007. Anatomy of *Escherichia coli* sigma70 promoters. *Nucleic Acids Res.* 35, 771–788. <https://doi.org/10.1093/nar/gkl956>
- Signon, L., Kleckner, N., 1995. Negative and positive regulation of Tn10/IS10-promoted recombination by IHF: two distinguishable processes inhibit transposition off of multicopy plasmid replicons and activate chromosomal events that favor evolution of new transposons. *Genes Dev.* 9, 1123–1136. <https://doi.org/10.1101/gad.9.9.1123>
- Siguier, P., Gourgouyere, E., Chandler, M., 2014. Bacterial insertion sequences: their genomic impact and diversity.

FEMS Microbiol. Rev. 38, 865–891. <https://doi.org/10.1111/1574-6976.12067>

Siguier, P., Gourbeyre, E., Varani, A., Ton-Hoang, B., Chandler, M., 2015. Everyman's Guide to Bacterial Insertion Sequences. *Microbiol. Spectr.* 3, MDNA3-0030–2014. <https://doi.org/10.1128/microbiolspec.MDNA3-0030-2014>

Simons, R.W., Hoopes, B.C., McClure, W.R., Kleckner, N., 1983. Three promoters near the termini of IS10: pIN, pOUT, and pIII. *Cell* 34, 673–682. [https://doi.org/10.1016/0092-8674\(83\)90400-2](https://doi.org/10.1016/0092-8674(83)90400-2)

Sittka, A., Lucchini, S., Papenfort, K., Sharma, C.M., Rolle, K., Binnewies, T.T., Hinton, J.C.D., Vogel, J., 2008. Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator, Hfq. *PLoS Genet.* 4, e1000163. <https://doi.org/10.1371/journal.pgen.1000163>

Sköld, O., 2000. Sulfonamide resistance: mechanisms and trends. *Drug Resist. Updat. Rev. Comment. Antimicrob. Anticancer Chemother.* 3, 155–160. <https://doi.org/10.1054/drup.2000.0146>

Slater, F.R., Bailey, M.J., Tett, A.J., Turner, S.L., 2008. Progress towards understanding the fate of plasmids in bacterial communities. *FEMS Microbiol. Ecol.* 66, 3–13. <https://doi.org/10.1111/j.1574-6941.2008.00505.x>

Smillie, C., Garcillán-Barcia, M.P., Francia, M.V., Rocha, E.P.C., Fernando de la Cruz, 2010. Mobility of plasmids. *Microbiol. Mol. Biol. Rev. MMBR* 74, 434–452. <https://doi.org/10.1128/MMBR.00020-10>

Sobecky, P.A., Hazen, T.H., 2009. Horizontal gene transfer and mobile genetic elements in marine systems. *Methods Mol. Biol. Clifton NJ* 532, 435–453. [https://doi.org/10.1007/978-1-60327-853-9\\_25](https://doi.org/10.1007/978-1-60327-853-9_25)

Sohn, S.G., Lee, J.J., Song, J.S., Lee, J.H., Sun, H.I., Park, K.S., Bae, I.K., Lee, J.-H., Jeong, B.C., Lee, S.H., 2009. Nomenclature of ISCR1 elements capable of mobilizing antibiotic resistance genes present in complex class 1 integrons. *J. Microbiol. Seoul Korea* 47, 514–516. <https://doi.org/10.1007/s12275-009-0054-5>

Sørum, H., Roberts, M.C., Crosa, J.H., 1992. Identification and cloning of a tetracycline resistance gene from the fish pathogen *Vibrio salmonicida*. *Antimicrob. Agents Chemother.* 36, 611–615. <https://doi.org/10.1128/AAC.36.3.611>

Sota, M., Top, E.M., 2008. Horizontal Gene Transfer Mediated by Plasmids [WWW Document]. URL <https://www.caister.com/hsp/abstracts/pla/05.html> (accessed 10.11.23).

Soucy, S.M., Huang, J., Gogarten, J.P., 2015. Horizontal gene transfer: building the web of life. *Nat. Rev. Genet.* 16, 472–482. <https://doi.org/10.1038/nrg3962>

Spellberg, B., Hansen, G.R., Kar, A., Cordova, C.D., Price, L.B., Johnson, and J.R., 2016. Antibiotic Resistance in Humans and Animals. *NAM Perspect.* <https://doi.org/10.31478/201606d>

Stalder, R., Caspers, P., Olsasz, F., Arber, W., 1990. The N-terminal domain of the insertion sequence 30 transposase interacts specifically with the terminal inverted repeats of the element. *J. Biol. Chem.* 265, 3757–3762.

Stevenson, C., Hall, J.P., Harrison, E., Wood, Aj., Brockhurst, M.A., 2017. Gene mobility promotes the spread of resistance in bacterial populations. *ISME J.* 11, 1930–1932. <https://doi.org/10.1038/ismej.2017.42>

Stoebel, D.M., Dorman, C.J., 2010. The effect of mobile element IS10 on experimental regulatory evolution in *Escherichia coli*. *Mol. Biol. Evol.* 27, 2105–2112. <https://doi.org/10.1093/molbev/msq101>

Stokes, H.W., Hall, R.M., 1989. A novel family of potentially mobile DNA elements encoding site-specific gene-integration functions: integrons. *Mol. Microbiol.* 3, 1669–1683. <https://doi.org/10.1111/j.1365-2958.1989.tb00153.x>

Stokes, H.W., Holmes, A.J., Nield, B.S., Holley, M.P., Nevalainen, K.M., Mabbutt, B.C., Gillings, M.R., 2001. Gene cassette PCR: sequence-independent recovery of entire genes from environmental DNA. *Appl. Environ. Microbiol.* 67, 5240–5246. <https://doi.org/10.1128/AEM.67.11.5240-5246.2001>

Stokes, H.W., O'Gorman, D.B., Recchia, G.D., Parsekhian, M., Hall, R.M., 1997. Structure and function of 59-base element recombination sites associated with mobile gene cassettes. *Mol. Microbiol.* 26, 731–745. <https://doi.org/10.1046/j.1365-2958.1997.6091980.x>

Stokes, H.W., Tomaras, C., Parsons, Y., Hall, R.M., 1993. The partial 3'-conserved segment duplications in the integrons In6 from pSa and In7 from pDGO100 have a common origin. *Plasmid* 30, 39–50. <https://doi.org/10.1006/plas.1993.1032>

St-Pierre, F., Endy, D., 2008. Determination of cell fate selection during phage lambda infection. *Proc. Natl. Acad. Sci. U. S. A.* 105, 20705–20710. <https://doi.org/10.1073/pnas.0808831105>

Streisinger, G., Edgar, R.S., Denhardt, G.H., 1964. CHROMOSOME STRUCTURE IN PHAGE T4. I. CIRCULARITY OF THE LINKAGE MAP. *Proc. Natl. Acad. Sci. U. S. A.* 51, 775–779. <https://doi.org/10.1073/pnas.51.5.775>

- Streisinger, G., Emrich, J., Stahl, M.M., 1967. Chromosome structure in phage t4, iii. Terminal redundancy and length determination. *Proc. Natl. Acad. Sci. U. S. A.* 57, 292–295. <https://doi.org/10.1073/pnas.57.2.292>
- Su, Z., Dai, X., Chen, J., Kong, F., Wang, H., Li, Y., Peng, S., Wang, S., Shao, Q., Lv, L., Xu, H., 2008. The bla(CTX-M-1) gene located in a novel complex class I integron bearing an ISCR1 element in *Escherichia coli* isolates from Zhenjiang, China. *J. Antimicrob. Chemother.* 62, 1150–1151. <https://doi.org/10.1093/jac/dkn300>
- Sun, D., 2018. Pull in and Push Out: Mechanisms of Horizontal Gene Transfer in Bacteria. *Front. Microbiol.* 9, 2154. <https://doi.org/10.3389/fmicb.2018.02154>
- Sun, D., 2016. Two different routes for double-stranded DNA transfer in natural and artificial transformation of *Escherichia coli*. *Biochem. Biophys. Res. Commun.* 471, 213–218. <https://doi.org/10.1016/j.bbrc.2016.01.137>
- Suzuki, M.M., Bird, A., 2008. DNA methylation landscapes: provocative insights from epigenomics. *Nat. Rev. Genet.* 9, 465–476. <https://doi.org/10.1038/nrg2341>
- Szuplewska, M., Bartosik, D., 2009. Identification of a mosaic transposable element of *Paracoccus marcusii* composed of insertion sequence ISPmar4 (ISAs1 family) and an IS1247a-driven transposable module (TMO). *FEMS Microbiol. Lett.* 292, 216–221. <https://doi.org/10.1111/j.1574-6968.2009.01495.x>
- Tada, T., Shrestha, S., Shimada, K., Ohara, H., Sherchand, J.B., Pokhrel, B.M., Kirikae, T., 2017. PER-8, a Novel Extended-Spectrum  $\beta$ -Lactamase PER Variant, from an *Acinetobacter baumannii* Clinical Isolate in Nepal. *Antimicrob. Agents Chemother.* 61, e02300-16. <https://doi.org/10.1128/AAC.02300-16>
- Tarasova, E., Khayat, R., 2021. A Structural Perspective of Reps from CRESS-DNA Viruses and Their Bacterial Plasmid Homologues. *Viruses* 14, 37. <https://doi.org/10.3390/v14010037>
- Tato, I., Matilla, I., Arechaga, I., Zunzunegui, S., de la Cruz, F., Cabezon, E., 2007. The ATPase activity of the DNA transporter TrwB is modulated by protein TrwA: implications for a common assembly mechanism of DNA translocating motors. *J. Biol. Chem.* 282, 25569–25576. <https://doi.org/10.1074/jbc.M703464200>
- Tavakoli, N., Comanducci, A., Dodd, H.M., Lett, M.C., Albiger, B., Bennett, P., 2000. IS1294, a DNA element that transposes by RC transposition. *Plasmid* 44, 66–84. <https://doi.org/10.1006/plas.1999.1460>
- Tavakoli, N.P., DeVost, J., Derbyshire, K.M., 1997. Defining functional regions of the IS903 transposase. *J. Mol. Biol.* 274, 491–504. <https://doi.org/10.1006/jmbi.1997.1410>
- Taylor, S.D., Palmer, M., 2016. The action mechanism of daptomycin. *Bioorg. Med. Chem.* 24, 6253–6268. <https://doi.org/10.1016/j.bmc.2016.05.052>
- Teo, S.S., Cheng, A., Melvin, D., Ng, K., Harris, J., 2022. Lateral airways X-ray with a history of choking. *J. Paediatr. Child Health.* <https://doi.org/10.1111/jpc.14595>
- Tettelin, H., Riley, D., Cattuto, C., Medini, D., 2008. Comparative genomics: the bacterial pan-genome. *Curr. Opin. Microbiol.* 11, 472–477. <https://doi.org/10.1016/j.mib.2008.09.006>
- Tetu, S.G., Holmes, A.J., 2008. A family of insertion sequences that impacts integrons by specific targeting of gene cassette recombination sites, the IS1111-attC Group. *J. Bacteriol.* 190, 4959–4970. <https://doi.org/10.1128/JB.00229-08>
- Thierauf, A., Perez, G., Maloy, A.S., 2009. Generalized transduction. *Methods Mol. Biol.* Clifton NJ 501, 267–286. [https://doi.org/10.1007/978-1-60327-164-6\\_23](https://doi.org/10.1007/978-1-60327-164-6_23)
- Thomas, J., Pritham, E.J., 2015. Helitrons, the Eukaryotic Rolling-circle Transposable Elements. *Microbiol. Spectr.* 3. <https://doi.org/10.1128/microbiolspec.MDNA3-0049-2014>
- Tobes, R., Pareja, E., 2006. Bacterial repetitive extragenic palindromic sequences are DNA targets for Insertion Sequence elements. *BMC Genomics* 7, 62. <https://doi.org/10.1186/1471-2164-7-62>
- Tobiason, D.M., Lenich, A.G., Glasgow, A.C., 1999. Multiple DNA binding activities of the novel site-specific recombinase, Piv, from *Moraxella lacunata*. *J. Biol. Chem.* 274, 9698–9706. <https://doi.org/10.1074/jbc.274.14.9698>
- Toleman, M.A., Bennett, P.M., Walsh, T.R., 2006a. ISCR elements: novel gene-capturing systems of the 21st century? *Microbiol. Mol. Biol. Rev. MMBR* 70, 296–316. <https://doi.org/10.1128/MMBR.00048-05>
- Toleman, M.A., Bennett, P.M., Walsh, T.R., 2006b. Common regions e.g. orf513 and antibiotic resistance: IS91-like elements evolving complex class 1 integrons. *J. Antimicrob. Chemother.* 58, 1–6. <https://doi.org/10.1093/jac/dkl204>
- Toleman, M.A., Walsh, T.R., 2011. Combinatorial events of insertion sequences and ICE in Gram-negative bacteria. *FEMS Microbiol. Rev.* 35, 912–935. <https://doi.org/10.1111/j.1574-6976.2011.00294.x>

- Toleman, M.A., Walsh, T.R., 2010. ISCR elements are key players in IncA/C plasmid evolution. *Antimicrob. Agents Chemother.* 54, 3534; author reply 3534. <https://doi.org/10.1128/AAC.00383-10>
- Toleman, M.A., Walsh, T.R., 2008. Evolution of the ISCR3 group of ISCR elements. *Antimicrob. Agents Chemother.* 52, 3789–3791. <https://doi.org/10.1128/AAC.00479-08>
- Tomcsanyi, T., Berg, D.E., 1989. Transposition effect of adenine (Dam) methylation on activity of O end mutants of IS50. *J. Mol. Biol.* 209, 191–193. [https://doi.org/10.1016/0022-2836\(89\)90271-4](https://doi.org/10.1016/0022-2836(89)90271-4)
- Tompkins, K.J., Houtti, M., Litzau, L.A., Aird, E.J., Everett, B.A., Nelson, A.T., Pornschloegl, L., Limón-Swanson, L.K., Evans, R.L., Evans, K., Shi, K., Aihara, H., Gordon, W.R., 2021. Molecular underpinnings of ssDNA specificity by Rep HUH-endonucleases and implications for HUH-tag multiplexing and engineering. *Nucleic Acids Res.* 49, 1046–1064. <https://doi.org/10.1093/nar/gkaa1248>
- Tomsic, M., Tsujikawa, L., Panaghie, G., Wang, Y., Azok, J., deHaseth, P.L., 2001. Different roles for basic and aromatic amino acids in conserved region 2 of Escherichia coli sigma(70) in the nucleation and maintenance of the single-stranded DNA bubble in open RNA polymerase-promoter complexes. *J. Biol. Chem.* 276, 31891–31896. <https://doi.org/10.1074/jbc.M105027200>
- Ton-Hoang, B., Bétermier, M., Polard, P., Chandler, M., 1997. Assembly of a strong promoter following IS911 circularization and the role of circles in transposition. *EMBO J.* 16, 3357–3371. <https://doi.org/10.1093/emboj/16.11.3357>
- Ton-Hoang, B., Guynet, C., Ronning, D.R., Cointin-Marty, B., Dyda, F., Chandler, M., 2005. Transposition of ISHp608, member of an unusual family of bacterial insertion sequences. *EMBO J.* 24, 3325–3338. <https://doi.org/10.1038/sj.emboj.7600787>
- Ton-Hoang, B., Pasternak, C., Siguier, P., Guynet, C., Hickman, A.B., Dyda, F., Sommer, S., Chandler, M., 2010. Single-stranded DNA transposition is coupled to host replication. *Cell* 142, 398–408. <https://doi.org/10.1016/j.cell.2010.06.034>
- Ton-Hoang, B., Siguier, P., Quentin, Y., Onillon, S., Marty, B., Fichant, G., Chandler, M., 2012. Structuring the bacterial genome: Y1-transposases associated with REP-BIME sequences. *Nucleic Acids Res.* 40, 3596–3609. <https://doi.org/10.1093/nar/gkr1198>
- Touchon, M., Rocha, E.P.C., 2007. Causes of insertion sequences abundance in prokaryotic genomes. *Mol. Biol. Evol.* 24, 969–981. <https://doi.org/10.1093/molbev/msm014>
- Traglia, G.M., Place, K., Dotto, C., Fernandez, J.S., Montaña, S., Bahiense, C.D.S., Soler-Bistue, A., Iriarte, A., Perez, F., Tolmasky, M.E., Bonomo, R.A., Melano, R.G., Ramírez, M.S., 2019. Interspecies DNA acquisition by a naturally competent Acinetobacter baumannii strain. *Int. J. Antimicrob. Agents* 53, 483–490. <https://doi.org/10.1016/j.ijantimicag.2018.12.013>
- Trocter, M., Felisberto-Rodrigues, C., Christie, P.J., Waksman, G., 2014. Recent advances in the structural and molecular biology of type IV secretion systems. *Curr. Opin. Struct. Biol.* 27, 16–23. <https://doi.org/10.1016/j.sbi.2014.02.006>
- Tsujikawa, L., Tsodikov, O.V., deHaseth, P.L., 2002. Interaction of RNA polymerase with forked DNA: evidence for two kinetically significant intermediates on the pathway to the final complex. *Proc. Natl. Acad. Sci. U. S. A.* 99, 3493–3498. <https://doi.org/10.1073/pnas.062487299>
- Turlan, C., Chandler, M., 2000. Playing second fiddle: second-strand processing and liberation of transposable elements from donor DNA. *Trends Microbiol.* 8, 268–274. [https://doi.org/10.1016/s0966-842x\(00\)01757-1](https://doi.org/10.1016/s0966-842x(00)01757-1)
- Turnbull, L., Toyofuku, M., Hynen, A.L., Kurosawa, M., Pessi, G., Petty, N.K., Osvath, S.R., Cárcamo-Oyarce, G., Gloag, E.S., Shimoni, R., Omasits, U., Ito, S., Yap, X., Monahan, L.G., Cavaliere, R., Ahrens, C.H., Charles, I.G., Nomura, N., Eberl, L., Whitchurch, C.B., 2016. Explosive cell lysis as a mechanism for the biogenesis of bacterial membrane vesicles and biofilms. *Nat. Commun.* 7, 11220. <https://doi.org/10.1038/ncomms11220>
- Turner, A.K., De La Cruz, F., Grinsted, J., 1990. Temperature sensitivity of transposition of class II transposons. *J. Gen. Microbiol.* 136, 65–67. <https://doi.org/10.1099/00221287-136-1-65>
- Twiss, E., Coros, A.M., Tavakoli, N.P., Derbyshire, K.M., 2005. Transposition is modulated by a diverse set of host factors in Escherichia coli and is stimulated by nutritional stress. *Mol. Microbiol.* 57, 1593–1607. <https://doi.org/10.1111/j.1365-2958.2005.04794.x>
- Ur Rahman, S., Ali, T., Ali, I., Khan, N.A., Han, B., Gao, J., 2018. The Growing Genetic and Functional Diversity of Extended Spectrum Beta-Lactamases. *BioMed Res. Int.* 2018, 9519718. <https://doi.org/10.1155/2018/9519718>

- Valentine, C.R., Heinrich, M.J., Chissoe, S.L., Roe, B.A., 1994. DNA sequence of direct repeats of the *sull* gene of plasmid pSa. *Plasmid* 32, 222–227. <https://doi.org/10.1006/plas.1994.1059>
- van Mansfeld, A.D., van Teeffelen, H.A., Baas, P.D., Jansz, H.S., 1986. Two juxtaposed tyrosyl-OH groups participate in phi X174 gene A protein catalysed cleavage and ligation of DNA. *Nucleic Acids Res.* 14, 4229–4238. <https://doi.org/10.1093/nar/14.10.4229>
- van Raaij, M.J., Abrahams, J.P., Leslie, A.G., Walker, J.E., 1996. The structure of bovine F1-ATPase complexed with the antibiotic inhibitor aurovertin B. *Proc. Natl. Acad. Sci. U. S. A.* 93, 6913–6917.
- Vannuffel, P., Cocito, C., 1996. Mechanism of action of streptogramins and macrolides. *Drugs* 51 Suppl 1, 20–30. <https://doi.org/10.2165/00003495-199600511-00006>
- Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., Laydon, A., Židek, A., Green, T., Tunyasuvunakool, K., Petersen, S., Jumper, J., Clancy, E., Green, R., Vora, A., Lutfi, M., Figurnov, M., Cowie, A., Hobbs, N., Kohli, P., Kleywegt, G., Birney, E., Hassabis, D., Velankar, S., 2022. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* 50, D439–D444. <https://doi.org/10.1093/nar/gkab1061>
- Vary, P.H., Andersen, P.R., Green, E., Hermon-Taylor, J., McFadden, J.J., 1990. Use of highly specific DNA probes and the polymerase chain reaction to detect *Mycobacterium paratuberculosis* in Johne's disease. *J. Clin. Microbiol.* 28, 933–937. <https://doi.org/10.1128/jcm.28.5.933-937.1990>
- Vilei, E.M., Nicolet, J., Frey, J., 1999. IS1634, a novel insertion element creating long, variable-length direct repeats which is specific for *Mycoplasma mycoides* subsp. *mycoides* small-colony type. *J. Bacteriol.* 181, 1319–1323. <https://doi.org/10.1128/JB.181.4.1319-1323.1999>
- Villa, L., Carattoli, A., 2005. Integrons and transposons on the *Salmonella enterica* serovar typhimurium virulence plasmid. *Antimicrob. Agents Chemother.* 49, 1194–1197. <https://doi.org/10.1128/AAC.49.3.1194-1197.2005>
- Vogel, U., Jensen, K.F., 1994. The RNA chain elongation rate in *Escherichia coli* depends on the growth rate. *J. Bacteriol.* 176, 2807–2813. <https://doi.org/10.1128/jb.176.10.2807-2813.1994>
- Waddell, C.S., Craig, N.L., 1988. Tn7 transposition: two transposition pathways directed by five Tn7-encoded genes. *Genes Dev.* 2, 137–149. <https://doi.org/10.1101/gad.2.2.137>
- Wagner, A., 2009. Transposable elements as genomic diseases. *Mol. Biosyst.* 5, 32–35. <https://doi.org/10.1039/b814624c>
- Wang, G., Wilson, T.J., Jiang, Q., Taylor, D.E., 2001. Spontaneous mutations that confer antibiotic resistance in *Helicobacter pylori*. *Antimicrob. Agents Chemother.* 45, 727–733. <https://doi.org/10.1128/AAC.45.3.727-733.2001>
- Wang, R., Yu, D., Yue, J., Kan, B., 2016. Variations in SXT elements in epidemic *Vibrio cholerae* O1 El Tor strains in China. *Sci. Rep.* 6, 22733. <https://doi.org/10.1038/srep22733>
- Wang, X., Thompson, C.D., Weidenmaier, C., Lee, J.C., 2018. Release of *Staphylococcus aureus* extracellular vesicles and their application as a vaccine platform. *Nat. Commun.* 9, 1379. <https://doi.org/10.1038/s41467-018-03847-z>
- Ward, C.M., Wardle, S.J., Singh, R.K., Haniford, D.B., 2007. The global regulator H-NS binds to two distinct classes of sites within the Tn10 transpososome to promote transposition. *Mol. Microbiol.* 64, 1000–1013. <https://doi.org/10.1111/j.1365-2958.2007.05708.x>
- Wardle, S.J., Chan, A., Haniford, D.B., 2009. H-NS binds with high affinity to the Tn10 transpososome and promotes transpososome stabilization. *Nucleic Acids Res.* 37, 6148–6160. <https://doi.org/10.1093/nar/gkp672>
- Wardle, S.J., O'Carroll, M., Derbyshire, K.M., Haniford, D.B., 2005. The global regulator H-NS acts directly on the transpososome to promote Tn10 transposition. *Genes Dev.* 19, 2224–2235. <https://doi.org/10.1101/gad.1338905>
- Waters, J.L., Salyers, A.A., 2013. Regulation of CTnDOT conjugative transfer is a complex and highly coordinated series of events. *mBio* 4, e00569-00513. <https://doi.org/10.1128/mBio.00569-13>
- Weinreich, M.D., Makris, J.C., Reznikoff, W.S., 1991. Induction of the SOS response in *Escherichia coli* inhibits Tn5 and IS50 transposition. *J. Bacteriol.* 173, 6910–6918. <https://doi.org/10.1128/jb.173.21.6910-6918.1991>
- Wells, J.N., Feschotte, C., 2020. A Field Guide to Eukaryotic Transposable Elements. *Annu. Rev. Genet.* 54, 539–561. <https://doi.org/10.1146/annurev-genet-040620-022145>
- Whitfield, C.R., Wardle, S.J., Haniford, D.B., 2009. The global bacterial regulator H-NS promotes transpososome formation and transposition in the Tn5 system. *Nucleic Acids Res.* 37, 309–321. <https://doi.org/10.1093/nar/gkn935>

- Whittle, G., Shoemaker, N.B., Salyers, A.A., 2002. The role of *Bacteroides* conjugative transposons in the dissemination of antibiotic resistance genes. *Cell. Mol. Life Sci. CMLS* 59, 2044–2054. <https://doi.org/10.1007/s000180200004>
- Whittle, Gabrielle, Shoemaker, N.B., Salyers, A.A., 2002. Characterization of genes involved in modulation of conjugal transfer of the *Bacteroides* conjugative transposon CTnDOT. *J. Bacteriol.* 184, 3839–3847. <https://doi.org/10.1128/JB.184.14.3839-3847.2002>
- WHO, 2014. Antimicrobial resistance: global report on surveillance [WWW Document]. URL <https://www.who.int/publications-detail-redirect/9789241564748> (accessed 10.11.23).
- Wiegand, T.W., Reznikoff, W.S., 1994. Interaction of Tn5 transposase with the transposon termini. *J. Mol. Biol.* 235, 486–495. <https://doi.org/10.1006/jmbi.1994.1008>
- Wilde, C., Escartin, F., Kokeguchi, S., Latour-Lambert, P., Lectard, A., Clément, J.-M., 2003. Transposases are responsible for the target specificity of IS1397 and ISKpn1 for two different types of palindromic units (PUs). *Nucleic Acids Res.* 31, 4345–4353. <https://doi.org/10.1093/nar/gkg494>
- Winter, M., Buckling, A., Harms, K., Johnsen, P.J., Vos, M., 2021. Antimicrobial resistance acquisition via natural transformation: context is everything. *Curr. Opin. Microbiol.* 64, 133–138. <https://doi.org/10.1016/j.mib.2021.09.009>
- Wise, R., 1999. A review of the mechanisms of action and resistance of antimicrobial agents. *Can. Respir. J.* 6 Suppl A, 20A–2A.
- Wolf, M.K., de Haan, L.A., Cassels, F.J., Willshaw, G.A., Warren, R., Boedeker, E.C., Gaastra, W., 1997. The CS6 colonization factor of human enterotoxigenic *Escherichia coli* contains two heterologous major subunits. *FEMS Microbiol. Lett.* 148, 35–42. <https://doi.org/10.1111/j.1574-6968.1997.tb10263.x>
- Wolfe, S.A., Nekludova, L., Pabo, C.O., 2000. DNA recognition by Cys2His2 zinc finger proteins. *Annu. Rev. Biophys. Biomol. Struct.* 29, 183–212. <https://doi.org/10.1146/annurev.biophys.29.1.183>
- Wood, M.M., Gardner, J.F., 2015. The Integration and Excision of CTnDOT. *Microbiol. Spectr.* 3, MDNA3-0020–2014. <https://doi.org/10.1128/microbiolspec.MDNA3-0020-2014>
- Wood, M.S., Byrne, A., Lessie, T.G., 1991. IS406 and IS407, two gene-activating insertion sequences for *Pseudomonas cepacia*. *Gene* 105, 101–105. [https://doi.org/10.1016/0378-1119\(91\)90519-h](https://doi.org/10.1016/0378-1119(91)90519-h)
- Wösten, M.M., 1998. Eubacterial sigma-factors. *FEMS Microbiol. Rev.* 22, 127–150. <https://doi.org/10.1111/j.1574-6976.1998.tb00364.x>
- Wozniak, R.A.F., Waldor, M.K., 2010. Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow. *Nat. Rev. Microbiol.* 8, 552–563. <https://doi.org/10.1038/nrmicro2382>
- Wu, C.A., Zechner, E.L., Marians, K.J., 1992. Coordinated leading- and lagging-strand synthesis at the *Escherichia coli* DNA replication fork. I. Multiple effectors act to modulate Okazaki fragment size. *J. Biol. Chem.* 267, 4030–4044.
- Xue, H., Cordero, O.X., Camas, F.M., Trimble, W., Meyer, F., Guglielmini, J., Rocha, E.P.C., Polz, M.F., 2015. Eco-Evolutionary Dynamics of Episomes among Ecologically Cohesive Bacterial Populations. *mBio* 6, e00552-00515. <https://doi.org/10.1128/mBio.00552-15>
- Yanisch-Perron, C., Vieira, J., Messing, J., 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33, 103–119. [https://doi.org/10.1016/0378-1119\(85\)90120-9](https://doi.org/10.1016/0378-1119(85)90120-9)
- Yarnell, W.S., Roberts, J.W., 1999. Mechanism of intrinsic transcription termination and antitermination. *Science* 284, 611–615. <https://doi.org/10.1126/science.284.5414.611>
- Yoneyama, H., Katsumata, R., 2006. Antibiotic resistance in bacteria and its future for novel antibiotic development. *Biosci. Biotechnol. Biochem.* 70, 1060–1075. <https://doi.org/10.1271/bbb.70.1060>
- Young, R., Bremer, H., 1976. Polypeptide-chain-elongation rate in *Escherichia coli* B/r as a function of growth rate. *Biochem. J.* 160, 185–194. <https://doi.org/10.1042/bj1600185>
- Yu, Z., Wang, Y., Henderson, I.R., Guo, J., 2022. Artificial sweeteners stimulate horizontal transfer of extracellular antibiotic resistance genes through natural transformation. *ISME J.* 16, 543–554. <https://doi.org/10.1038/s41396-021-01095-6>
- Zabala, J.C., de la Cruz, F., Ortiz, J.M., 1982. Several copies of the same insertion sequence are present in alpha-hemolytic plasmids belonging to four different incompatibility groups. *J. Bacteriol.* 151, 472–476.

<https://doi.org/10.1128/jb.151.1.472-476.1982>

Zapun, A., Contreras-Martel, C., Vernet, T., 2008. Penicillin-binding proteins and beta-lactam resistance. *FEMS Microbiol. Rev.* 32, 361–385. <https://doi.org/10.1111/j.1574-6976.2007.00095.x>

Zechner, E.L., Wu, C.A., Marians, K.J., 1992. Coordinated leading- and lagging-strand synthesis at the *Escherichia coli* DNA replication fork. III. A polymerase-primase interaction governs primer size. *J. Biol. Chem.* 267, 4054–4063.

Zeng, L., Skinner, S.O., Zong, C., Sippy, J., Feiss, M., Golding, I., 2010. Decision making at a subcellular level determines the outcome of bacteriophage infection. *Cell* 141, 682–691. <https://doi.org/10.1016/j.cell.2010.03.034>

Zerbib, D., Gamas, P., Chandler, M., Prentki, P., Bass, S., Galas, D., 1985. Specificity of insertion of IS1. *J. Mol. Biol.* 185, 517–524. [https://doi.org/10.1016/0022-2836\(85\)90068-3](https://doi.org/10.1016/0022-2836(85)90068-3)

Zerbib, D., Jakowec, M., Prentki, P., Galas, D.J., Chandler, M., 1987. Expression of proteins essential for IS1 transposition: specific binding of InsA to the ends of IS1. *EMBO J.* 6, 3163–3169. <https://doi.org/10.1002/j.1460-2075.1987.tb02627.x>

Zerbib, D., Prentki, P., Gamas, P., Freund, E., Galas, D.J., Chandler, M., 1990. Functional organization of the ends of IS 1: specific binding site for an IS1-encoded protein. *Mol. Microbiol.* 4, 1477–1486. <https://doi.org/10.1111/j.1365-2958.1990.tb02058.x>

Zhang, X.-X., Zhang, T., Fang, H.H.P., 2009. Antibiotic resistance genes in water environment. *Appl. Microbiol. Biotechnol.* 82, 397–414. <https://doi.org/10.1007/s00253-008-1829-z>

Zhou, L., Mitra, R., Atkinson, P.W., Hickman, A.B., Dyda, F., Craig, N.L., 2004. Transposition of hAT elements links transposable elements and V(D)J recombination. *Nature* 432, 995–1001. <https://doi.org/10.1038/nature03157>

Zhu, M., Dai, X., Wang, Y.-P., 2016. Real time determination of bacterial *in vivo* ribosome translation elongation speed based on LacZ $\alpha$  complementation system. *Nucleic Acids Res.* 44, e155. <https://doi.org/10.1093/nar/gkw698>

Zinder, N.D., Lederberg, J., 1952. Genetic exchange in *Salmonella*. *J. Bacteriol.* 64, 679–699. <https://doi.org/10.1128/jb.64.5.679-699.1952>

Zrimec, J., Lapanje, A., 2018. DNA structure at the plasmid origin-of-transfer indicates its potential transfer range. *Sci. Rep.* 8, 1820. <https://doi.org/10.1038/s41598-018-20157-y>

Zweiger, G., Marczynski, G., Shapiro, L., 1994. A *Caulobacter* DNA methyltransferase that functions only in the predivisional cell. *J. Mol. Biol.* 235, 472–485. <https://doi.org/10.1006/jmbi.1994.1007>



## Annexes

---

Annexe 1. Liste des communications scientifiques.....	273
Annexe 2. Number accession des séquences d'IS décrites (IS91, IS1294b et IS801).....	275
Annexe 2.1. Number accession IS91.....	275
Annexe 2.2. Number accession IS1294b.....	276
Annexe 2.3. Number accession IS801.....	277
Annexe 3. Description simplifiée des isoformes de la base de données GenBank® .....	278
Annexe 3.1. Isoforme de l'élément IS91 .....	278
Annexe 3.2. Isoforme de l'élément IS1294b .....	279
Annexe 3.3. Isoforme de l'élément IS801 .....	281
Annexe 4. Agencement des réplicons à plusieurs copies d'IS.....	282
Annexe 4.1. Réplicon à plusieurs copies de d'élément IS91 .....	282
Annexe 4.2. Réplicon à plusieurs copies de d'élément IS1294b .....	283
Annexe 4.3. Réplicon à plusieurs copies de d'élément IS801 .....	284
Annexe 5. Schéma de la PCR arbitraire .....	285
Annexe 6. Liste des amorces utilisées lors de cette étude .....	286

## **Annexe 1. Liste des communications scientifiques**

### **Communications orales :**

#### ***Nationale***

**Fauconnier A.**, Jové T., Pasternak C. and Ploy M.C. Impact d'un ORF sur l'activité d'une séquence d'insertion impliquée dans la propagation de la résistance aux antibiotiques. OmegaLive 2021 - institut Omega-Health (Limoges, France)

### **Communications affichées :**

#### ***Nationale***

**Fauconnier A.**, Gaschet M., Pasternak C. and Ploy M.C. Orf121 carried by the IS91 insertion sequence negatively influences its mobility. 4th Microbiology Day 2023 (Bordeaux, France)

**Fauconnier A.**, Jové T., Pasternak C. and Ploy M.C. Une petite ORF présente dans la séquence d'insertion IS91 influence négativement sa transposition. Société Française de Microbiologie (SFM) : Microbes 2022 (Montpellier, France)

**Fauconnier A.**, Jové T., Pasternak C. and Ploy M.C. Impact d'un ORF sur l'activité d'une séquence d'insertion impliquée dans la propagation de la résistance aux antibiotiques. Journées de l'interdisciplinarité 2021 – université de Limoges (Limoges, France)

#### ***Internationale***

**Fauconnier A.**, Gaschet M., Da Re S., Jové T., Ploy M.C. and Pasternak C. The IS91-encoded Orf121 negatively influences IS mobility. Federation of European Microbiological Societies 2023 (FEMS) (Hamburg, Allemagne)

**Fauconnier A.**, Gaschet M., Pasternak C. and Ploy M.C. Orf121 carried by the IS91 insertion sequence negatively influences its mobility. Antimicrobial Resistance in Animals and the Environment 2023 (ARAE) (Tours, France)

**Fauconnier A.**, Ploy M.C., Jové T. and Pasternak C. A small ORF carried by the IS91 insertion sequence negatively influences its transposition.

International Symposium on Plasmid Biology 2022 (ISPB) (Toulouse, France)



# Annexe 2.2. Number accession IS1294b

AATXKU01000031.1	NZ_CXER01000203.1	NZ_CAJFAZ01000004.1	AAWUNV010000053.1	NZ_QQVW01000038.1	DABUZI010000066.1	NZ_CP044189.1	KP017243.1
AATXKU01000044.0	NZ_CXEX01000218.1	NZ_CAJFBA01000005.1	AAKELL01000032.1	AASMGU010000054.1	DABULX010000083.1	NZ_CP091038.1	CP019255.1
AATXKV01000030.1	NZ_CXJK01000236.1	NZ_CAJFB8010000009.1	ABAIUB010000043.1	NZ_CP060750.1	NZ_UVWA01000002.1	AAKAVP010000109.1	NC_014383.1
AATXKW010000147.1	NZ_CXLM01000266.1	CP055125.1	ABCUEU010000131.1	AAARME010000037.1	DADGMW010000302.1	AAMCCB010000032.1	NC_019116.1
AATXKX01000034.1	NZ_CXLN01000221.1	NZ_JAALAC010000001.1	ABCBCW010000050.1	CP012496.1	DAEWL010000025.1	AAMCCB010000061.1	NC_019118.1
AATXKY01000028.1	NZ_CP050866.1	NZ_CP050866.1	ABCCRA010000048.1	ABFEYD010000050.1	NZ_JAAMTU010000051.1	AAOTPO010000020.1	NC_021622.1
AATXLA01000028.1	NZ_CXLV01000218.1	NZ_CP050866.1	ABCTJO010000056.1	NZ_PVMG010000003.1	AZAFZ010000029.1	AAOTPO010000035.1	NC_023315.1
AATXMW010000058.1	NZ_CXLM01000153.1	NZ_CP055137.1	ABCOTK010000054.1	NZ_PVZZ010000003.1	AZAFZ010000071.1	AAMVWL010000078.1	NC_023239.1
AATXOE01000031.1	NZ_CXMB010000008.1	NZ_CP026794.1	ABFCY010000052.1	NZ_QQVX010000039.1	ABCQBA010000049.1	ABCRHY010000069.1	NC_023899.1
AATXOF010000144.1	NZ_CXMI01000211.1	NZ_CP026794.1	ABDGEM010000079.1	NZ_CP012492.1	ABDGCW010000046.1	ABEJX010000020.1	NC_024978.1
AATXOG01000033.1	NZ_FTSW01000020.1	NZ_CP026794.1	ABGVVW010000073.1	NZ_CP011417.1	DADGMW010000021.1	ABEJX010000128.1	NC_024979.1
AATXOH01000032.1	NZ_FTUX01000166.1	NZ_CP026794.1	ABGQY010000049.1	AATIKR010000060.1	DAEWAC010000024.1	DAEBAF010000141.1	NC_024980.1
AATXCZ01000012.1	NZ_FTXX01000189.1	CP045785.1	ABGPMQ010000100.1	NZ_UGGI01000007.1	DAEWAC010000098.1	DAEVMX010000060.1	NC_025142.1
AAUGMF01000037.1	NZ_FTZX01000254.1	CP045785.1	ABHIEW010000065.1	NZ_LSGC01000022.1	DAEWAD010000032.1	DAEVMX010000060.1	NC_025143.1
AAUGMW01000036.1	NZ_FTYO01000179.1	NZ_CABVZF010000123.1	DAIVIV010000044.1	NZ_UGDX01000001.1	DAEWAD010000163.1	NZ_JAGYU010000122.1	NC_025144.1
AAUNPA01000035.1	NZ_CABVWA010000117.1	NZ_CABVWA010000117.1	DAIVVL010000045.1	AAEIVW010000062.1	DAEWAE010000023.1	AAMPZH010000023.1	NC_025176.1
CP00039.1	AAVBYV010000135.1	AAVBYV010000135.1	JABWQ020000092.1	NZ_JALPZ010000050.1	DAEWAG010000034.1	AAMPZH010000048.1	NZ_KU355873.1
NC_007385.1	NZ_CP024472.1	NZ_CP024472.1	NZ_VRWZ010000061.1	NZ_UGZG010000067.1	DAEWAR010000024.1	NZ_CP083487.1	NZ_KU355873.1
NZ_CXGA01000292.1	UGFN01000002.1	UGFN01000002.1	NZ_JAFHAC010000057.1	ABEFFM010000139.1	DAEWAR010000124.1	NZ_CP045519.1	NZ_LC520276.1
NZ_CXIO01000004.1	AATIEV010000141.1	AATIEV010000141.1	UFYX01000005.1	NZ_LRKL01000148.1	DAEWBK010000034.1	AAMVWL010000040.1	NZ_LC520276.1
NZ_CXJO01000292.1	NZ_UGF01000002.1	NZ_UGF01000002.1	NZ_CYEX01000025.1	NZ_LRYN01000003.1	DAEWBK010000159.1	AAMVWL010000078.1	NZ_LC520276.1
NZ_CXJT010000290.1	UYVTO000010.1	UYVTO000010.1	ABBWOP010000039.1	NZ_CP065612.1	DAEWCN010000026.1	DAQAAR010000011.1	NZ_LC520276.1
NZ_CXKO01000024.1	NZ_UGDX01000001.1	NZ_UGDX01000001.1	AAAGIQ010000041.1	NZ_LR828998.1	DAEWCN010000096.1	DAQAAR010000030.1	NZ_UGLGF010000046.1
NZ_CXKQ010000262.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	DALFYS010000015.1	NZ_LR828998.1	DAEWDA010000028.1	DAQAAR010000033.1	NZ_AZEL02000004.1
NZ_CXKX010000294.1	NZ_CWTV01000138.1	NZ_CWTV01000138.1	NZ_CZKX010000046.1	NZ_LR828998.1	DAEWDA010000049.1	DAQAAR010000029.1	NZ_AZEL02000004.1
NZ_CXLM010000317.1	NZ_CWZD01000227.1	NZ_CWZD01000227.1	NZ_CZKX010000023.1	NZ_LR828998.1	DAEWDG010000028.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXMD01000001.1	NZ_CWZR01000263.1	NZ_CWZR01000263.1	NZ_CZYS010000045.1	NZ_LR828998.1	DAEWDG010000028.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXMR01000009.1	NZ_CXAK01000270.1	NZ_CXAK01000270.1	NZ_CYFC01000013.1	NZ_LR828998.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXNF01000001.1	NZ_CXB010000276.1	NZ_CXB010000276.1	NZ_CYFO010000042.1	NZ_LR828998.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXNG01000009.1	NZ_CXBN01000183.1	NZ_CXBN01000183.1	NZ_CYGD010000026.1	NZ_PNSU01000131.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXNH01000002.1	NZ_CXBW01000110.1	NZ_CXBW01000110.1	NZ_CYJN01000147.1	NZ_PNSY01000113.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXNJ010000001.1	NZ_JAAXCM010000037.1	NZ_JAAXCM010000037.1	NZ_CP011417.1	NZ_PNSY01000113.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXNK01000001.1	NZ_FTUC01000050.1	NZ_FTUC01000050.1	NZ_UGDX01000002.1	NZ_PNSZ01000118.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXNL01000004.1	NZ_FUDU01000016.1	NZ_FUDU01000016.1	NZ_UGYQ01000002.1	NZ_PNBT01000119.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXNM01000001.1	NZ_FUEO01000016.1	NZ_FUEO01000016.1	NZ_UGYQ01000002.1	NZ_PNBZ01000124.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXNO01000002.1	NZ_CP053752.1	NZ_CP053752.1	NZ_UGYQ01000002.1	NZ_POAM01000128.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXNP01000001.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXNQ01000002.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXNR01000003.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXNS01000004.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXNT01000005.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXNV01000006.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXNW01000007.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXNX01000008.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXNY01000009.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXNZ01000010.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXOA01000011.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXOB01000012.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXOC01000013.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXOD01000014.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXOE01000015.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXOF01000016.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXOG01000017.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXOH01000018.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXOI01000019.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXOJ01000020.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXOK01000021.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXOL01000022.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXOM01000023.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXON01000024.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXOO01000025.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXOP01000026.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXOQ01000027.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXOR01000028.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXOS01000029.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXOT01000030.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXOU01000031.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXOV01000032.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXOW01000033.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXOX01000034.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXOY01000035.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXOZ01000036.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXPA01000037.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXPB01000038.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXPC01000039.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXPD01000040.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXPE01000041.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXPF01000042.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXPG01000043.1	NZ_CVWSY01000107.1	NZ_CVWSY01000107.1	NZ_UGYQ01000002.1	NZ_QAFF01000002.1	DAEWEA010000026.1	DAQAAR010000030.1	NZ_ANVA010000084.1
NZ_CXPH01000044							

### Annexe 2.3. Number accession IS801

NZ_JAANQA01000003.1	X57269.1	NC_005773.3	CP000058.1
NZ_JAANQA01000009.1	NZ_JAANQA01000004.1	CP000058.1	CP000058.1
NZ_JAANQA01000009.1	NZ_JAANQA01000004.1	CP000059.1	CP000058.1
NZ_ODAL01000163.1	NC_007274.1	NZ_JAANPZ01000004.1	NZ_JANAKG010000139.1
NZ_LILJ01000099.1	CP000059.1	NC_005773.3	NZ_JAANQA01000009.1
NZ_LILK01000294.1	NZ_JAANPZ01000001.1	CP000058.1	NZ_JANAKG010000099.1
NZ_LILL01000185.1	NZ_JAANQA01000004.1	NZ_JAANPZ01000004.1	NZ_JAANPZ01000004.1
NZ_ODAM01000159.1	NZ_JAANQA01000009.1	NC_005773.3	NC_005773.3
NZ_OLMQ01000001.1	NZ_JAANQA01000004.1	CP000058.1	NZ_JAANPZ01000001.1
NZ_ODAM01000163.1	NZ_JAANPZ01000001.1	NZ_JAANPZ01000004.1	NZ_JAANPZ01000004.1
NZ_OLMQ01000001.1	NZ_JAANPZ01000001.1	NZ_JAANPZ01000004.1	NZ_LT963402.1
NZ_ODAM01000187.1	NC_007274.1	NC_005773.3	NZ_LT963402.1
NZ_LT985191.1	CP000059.1	NC_005773.3	NC_005773.3
NZ_LT985191.1	NZ_JAANQA01000003.1	NC_005773.3	NC_005773.3
NZ_LT963405.1	NZ_JAANQA01000009.1	CP000058.1	NZ_LT963405.1
NZ_LT963403.1	NZ_JAANQA01000009.1	CP000058.1	NC_007274.1
NZ_CP026565.1	NZ_JAANQA01000009.1	CP000058.1	NZ_JAANPZ01000004.1
NZ_CP026564.1	NC_007274.1	NZ_CP085686.1	
NZ_LT963402.1	CP000059.1	NZ_JAANPZ01000004.1	

# Annexe 3. Description simplifiée des isoformes de la base de données GenBank®

## Annexe 3.1. Isoforme de l'élément IS91

Variant	Accession length	# of	Year
IS91A1	1829	1	1991
IS91A2	1829	1	1991
IS91A3	1829	1	1991
IS91A4	1829	1	1991
IS91A5	1829	1	1991
IS91A6	1829	1	1991
IS91A7	1829	1	1991
IS91A8	1829	1	1991
IS91A9	1829	1	1991
IS91A10	1829	1	1991
IS91A11	1829	1	1991
IS91A12	1829	1	1991
IS91A13	1829	1	1991
IS91A14	1829	1	1991
IS91A15	1829	1	1991
IS91A16	1829	1	1991
IS91A17	1829	1	1991
IS91A18	1829	1	1991
IS91A19	1829	1	1991
IS91A20	1829	1	1991
IS91A21	1829	1	1991
IS91A22	1829	1	1991
IS91A23	1829	1	1991
IS91A24	1829	1	1991
IS91A25	1829	1	1991
IS91A26	1829	1	1991
IS91A27	1829	1	1991
IS91A28	1829	1	1991
IS91A29	1829	1	1991
IS91A30	1829	1	1991
IS91A31	1829	1	1991
IS91A32	1829	1	1991
IS91A33	1829	1	1991
IS91A34	1829	1	1991
IS91A35	1829	1	1991
IS91A36	1829	1	1991
IS91A37	1829	1	1991
IS91A38	1829	1	1991
IS91A39	1829	1	1991
IS91A40	1829	1	1991
IS91A41	1829	1	1991
IS91A42	1829	1	1991
IS91A43	1829	1	1991
IS91A44	1829	1	1991
IS91A45	1829	1	1991
IS91A46	1829	1	1991
IS91A47	1829	1	1991
IS91A48	1829	1	1991
IS91A49	1829	1	1991
IS91A50	1829	1	1991
IS91A51	1829	1	1991
IS91A52	1829	1	1991
IS91A53	1829	1	1991
IS91A54	1829	1	1991
IS91A55	1829	1	1991
IS91A56	1829	1	1991
IS91A57	1829	1	1991
IS91A58	1829	1	1991
IS91A59	1829	1	1991
IS91A60	1829	1	1991
IS91A61	1829	1	1991
IS91A62	1829	1	1991
IS91A63	1829	1	1991
IS91A64	1829	1	1991
IS91A65	1829	1	1991
IS91A66	1829	1	1991
IS91A67	1829	1	1991
IS91A68	1829	1	1991
IS91A69	1829	1	1991
IS91A70	1829	1	1991
IS91A71	1829	1	1991
IS91A72	1829	1	1991
IS91A73	1829	1	1991
IS91A74	1829	1	1991
IS91A75	1829	1	1991
IS91A76	1829	1	1991
IS91A77	1829	1	1991
IS91A78	1829	1	1991
IS91A79	1829	1	1991
IS91A80	1829	1	1991
IS91A81	1829	1	1991
IS91A82	1829	1	1991
IS91A83	1829	1	1991
IS91A84	1829	1	1991
IS91A85	1829	1	1991
IS91A86	1829	1	1991
IS91A87	1829	1	1991
IS91A88	1829	1	1991
IS91A89	1829	1	1991
IS91A90	1829	1	1991
IS91A91	1829	1	1991
IS91A92	1829	1	1991
IS91A93	1829	1	1991
IS91A94	1829	1	1991
IS91A95	1829	1	1991
IS91A96	1829	1	1991
IS91A97	1829	1	1991
IS91A98	1829	1	1991
IS91A99	1829	1	1991
IS91A100	1829	1	1991







### Annexe 3.3. Isoforme de l'élément IS801

Variet	séquence IS801		Étrémucitité de charge Ex801.5 max	Type IS801							localisation			Variet				
	%Identité nucléotidique	Identité nt		nt. nucléotidique	%Identité nucléotidique	Identité nt	nt. acide aminé	%Identité acide aminé	Identité aa	Zinc Finger (Cy54, Cy56, Cy71, H73, Cy78, Cy81, Cy85, Cy89)	HH (Cy87, Cy93, Cy97, H73, Cy88, Cy89)	HH (Cy97, Cy98, Cy99)	Accession length		nr/s	br/s		
IS801-V1	100.00%	1512/1512	67C	1233	100.00%	1233/1233	410	100.00%	410/410	C/C/C/H/C/C/C/C	HH	V/GRV	C/C	5517	2	1513	+	IS801-V1
IS801-V2	99.47%	1504/1512	67C	1233	99.53%	1235/1233	410	98.05%	402/410	C/C/C/H/C/C/C/C	HH	V/GRV	C/C	117325	68092	69603	+	IS801-V2
IS801-V3	99.40%	1503/1512	67C	1233	99.27%	1234/1233	410	98.34%	404/410	C/C/C/H/C/C/C/C	HH	V/GRV	C/C	131950	71545	73056	+	IS801-V3
IS801-V4	99.40%	1503/1512	67C	1233	99.27%	1234/1233	410	97.80%	403/410	C/C/C/H/C/C/C/C	HH	V/GRV	C/C	117325	19981	19470	-	IS801-V4
IS801-V5	99.40%	1503/1512	67C	1233	99.27%	1234/1233	410	97.80%	403/410	C/C/C/H/C/C/C/C	HH	V/GRV	C/C	117325	60157	58632	-	IS801-V5
IS801-V6	99.33%	1502/1512	67C	471	99.79%	468/469	156	99.35%	155/156	C/C/C/H/C/C/C/C	Absent	Absent	Absent	134282	11387	12898	+	IS801-V6
IS801-V7	99.33%	1502/1512	67C	1233	99.19%	1233/1233	410	97.56%	409/410	C/C/C/H/C/C/C/C	HH	V/GRV	C/C	8071	17759	16248	-	IS801-V7
IS801-V8	99.33%	1502/1512	67C	1233	99.27%	1234/1233	410	97.80%	403/410	C/C/C/H/C/C/C/C	HH	V/GRV	C/C	451218	62174	63885	+	IS801-V8
IS801-V9	99.33%	1502/1512	67C	1233	99.19%	1233/1233	410	97.80%	403/410	C/C/C/H/C/C/C/C	HH	V/GRV	C/C	451218	144306	144517	+	IS801-V9
IS801-V10	99.27%	1501/1512	67C	471	99.79%	468/469	156	99.36%	155/156	C/C/C/H/C/C/C/C	Absent	Absent	Absent	131950	39116	37605	-	IS801-V10
IS801-V11	99.21%	1500/1512	67C	471	99.79%	468/469	156	99.36%	155/156	C/C/C/H/C/C/C/C	Absent	Absent	Absent	134282	87383	85872	-	IS801-V11
IS801-V12	99.21%	1500/1512	67C	471	99.79%	467/469	156	98.72%	154/156	C/C/C/H/C/C/C/C	Absent	Absent	Absent	6116208	532506	5326517	+	IS801-V12
IS801-V13	99.21%	1500/1512	67C	471	99.79%	468/469	156	99.36%	155/156	C/C/C/H/C/C/C/C	Absent	Absent	Absent	131950	66609	65098	-	IS801-V13
IS801-V14	99.14%	1489/1512	67C	471	99.36%	466/469	156	98.72%	154/156	C/C/C/H/C/C/C/C	Absent	Absent	Absent	6116208	2991607	2990082	-	IS801-V14
IS801-V15	99.07%	1488/1512	67C	471	99.36%	466/469	156	98.68%	153/156	C/C/C/H/C/C/C/C	Absent	Absent	Absent	593287	5901507	5930018	+	IS801-V15
IS801-V16	99.01%	1497/1512	67C	1233	99.03%	1231/1233	410	98.05%	402/410	C/C/C/H/C/C/C/C	HH	V/GRV	C/C	6116208	272984	274495	+	IS801-V16
IS801-V17	99.01%	1497/1512	67C	1233	98.78%	1238/1233	410	97.32%	399/410	C/C/C/H/C/C/C/C	HH	V/GRV	C/C	119174	12500	10989	-	IS801-V17
IS801-V18	98.61%	1491/1512	67C	471	99.79%	468/469	156	99.36%	155/156	C/C/C/H/C/C/C/C	Absent	Absent	Absent	6116208	294417	289328	+	IS801-V18
IS801-V19	97.69%	1477/1512	67C	1230	97.24%	1199/1233	409	95.63%	393/410	C/C/C/H/C/C/C/C	HH	V/GRV	C/C	8719	5945	3537	-	IS801-V19
IS801-V20	97.49%	1475/1512	67C	471	99.57%	467/469	156	98.72%	154/156	C/C/C/H/C/C/C/C	Absent	Absent	Absent	451218	1878897	1894006	+	IS801-V20
IS801-V21	97.29%	1472/1512	67C	1233	96.68%	1193/1234	410	96.08%	394/410	C/C/C/H/C/C/C/C	HH	V/GRV	C/C	16710	12573	14084	+	IS801-V21
IS801-V22	96.90%	1467/1514	67C	468	98.72%	463/469	155	97.44%	152/156	C/C/C/H/C/C/C/C	Absent	Absent	Absent	6116208	2888674	2867167	-	IS801-V22
IS801-V23	96.84%	1469/1517	67C	1233	96.13%	1193/1241	410	95.12%	390/410	C/C/C/H/C/C/C/C	HH	V/GRV	C/C	8071	43702	45213	+	IS801-V23
IS801-V24	96.77%	1468/1517	67C	1233	96.13%	1193/1241	410	95.12%	390/410	C/C/C/H/C/C/C/C	HH	V/GRV	C/C	451218	4364245	4365756	+	IS801-V24
IS801-V25	95.97%	1452/1513	67C	1233	96.52%	1191/1234	410	96.34%	395/410	C/C/C/H/C/C/C/C	HH	V/GRV	C/C	6973	3895	5406	+	IS801-V25
IS801-V26	95.90%	1450/1512	67C	1230	96.09%	1186/1233	409	95.61%	392/410	C/C/C/H/C/C/C/C	HH	V/GRV	C/C	4482	2771	1265	-	IS801-V26
IS801-V27	95.57%	1446/1513	67C	1233	96.43%	1190/1234	410	94.63%	388/410	C/C/C/H/C/C/C/C	HH	V/GRV	C/C	10776	47	1558	+	IS801-V27
IS801-V28	95.44%	1445/1514	67C	1233	96.43%	1190/1234	410	95.12%	390/410	C/C/C/H/C/C/C/C	HH	V/GRV	C/C	265614	126071	124559	-	IS801-V28
IS801-V29	95.38%	1444/1514	67C	1233	96.11%	1186/1234	410	94.88%	389/410	C/C/C/H/C/C/C/C	HH	V/GRV	C/C	8297	5331	3309	-	IS801-V29
IS801-V30	95.24%	1442/1514	67C	1233	95.54%	1179/1234	410	95.77%	391/410	C/C/C/H/C/C/C/C	HH	V/GRV	C/C	265614	114145	112833	-	IS801-V30
IS801-V31	95.18%	1440/1513	67C	1233	95.95%	1194/1234	410	94.63%	389/410	C/C/C/H/C/C/C/C	HH	V/GRV	C/C	11064	3310	1799	-	IS801-V31
IS801-V32	95.11%	1440/1514	67C	1233	95.88%	1177/1234	410	95.12%	390/410	C/C/C/H/C/C/C/C	HH	V/GRV	C/C	67760	10679	9167	-	IS801-V32
IS801-V33	95.11%	1440/1514	67C	1233	96.03%	1185/1234	410	94.88%	389/410	C/C/C/H/C/C/C/C	HH	V/GRV	C/C	612038	262415	263927	+	IS801-V33
IS801-V34	95.05%	1439/1514	67C	1233	95.30%	1176/1234	410	95.12%	390/410	C/C/C/H/C/C/C/C	HH	V/GRV	C/C	108842	33149	31637	-	IS801-V34
IS801-V35	95.05%	1439/1514	67C	1233	95.46%	1178/1234	410	95.37%	391/410	C/C/C/H/C/C/C/C	HH	V/GRV	C/C	612038	4474733	4476245	+	IS801-V35

## Annexe 4. Agencement des réplicons à plusieurs copies d'IS

### Annexe 4.1. Réplicon à plusieurs copies de d'élément IS91

Genre	Support Génétique	replication initiator	Nombre copies	Premier variant		Second variant		Troisième variant		Quatrième variant		Représentation
				Variant	Nb. de copies	Variant	Nb. de copies	Variant	Nb. de copies	Variant	Nb. de copies	
<i>E. coli</i>	plasmid			IS91-V10	2							←; ←
<i>E. coli</i>	plasmid	IncFII		IS91-V10	2							→; →
<i>E. coli</i>	plasmid	IncFII		IS91-V24	2							←; →
<i>E. coli</i>	plasmid	IncFII		IS91-V8	2							→; →
<i>E. coli</i>	plasmid	IncFII		IS91-V22	1	IS91-V35	1					→; ←
<i>E. coli</i>	plasmid	IncFII		IS91-V11	1	IS91-V10	1					→; →
<i>E. coli</i>	plasmid	IncN		IS91-V10	2							→; ←
<i>E. coli</i>	plasmid	IncFIB		IS91-V9	2							←; →
<i>E. coli</i>	plasmid	IncFIB		IS91-V9	2							←; →
<i>E. coli</i>	plasmid	IncFII	2	IS91-V24	1	IS91-V10	1					←; ←
<i>E. coli</i>	plasmid	IncFII		IS91-V24	1	IS91-V10	1					←; ←
<i>E. coli</i>	plasmid	IncY		IS91-V24	2							→; ←
<i>E. coli</i>	plasmid	IncY		IS91-V24	1	IS91-V9	1					→; →
<i>E. coli</i>	plasmid	IncY		IS91-V24	1	IS91-V9	1					→; →
<i>E. coli</i>	plasmid	IncY		IS91-V24	1	IS91-V9	1					→; →
<i>E. coli</i>	plasmid	IncFII		IS91-V24	1	IS91-V9	1					→; →
<i>E. coli</i>	plasmid	IncFII		IS91-V10	2							←; →
<i>E. coli</i>	plasmid	IncFII		IS91-V24	2							←; →
<i>E. coli</i>	plasmid	IncFII		IS91-V108	1	IS91-V111	1					→; ←
<i>E. coli</i>	plasmid	IncFII		IS91-V25	3							←; ←; →
<i>E. coli</i>	plasmid	IncFII		IS91-V29	3							←; ←; ←
<i>E. coli</i>	plasmid	IncFII		IS91-V29	3							→; →; →
<i>E. coli</i>	plasmid	IncFII		IS91-V8	3							←; ←; →
<i>E. coli</i>	plasmid	IncFII	3	IS91-V8	2	IS91-V18	1					←; ←; →
<i>E. coli</i>	plasmid	IncFII		IS91-V24	2	IS91-V5	1					←; →; →
<i>E. coli</i>	plasmid	IncFII		IS91-V24	3							←; →; →
<i>E. coli</i>	plasmid	IncFII		IS91-V24	3							←; →; →
<i>E. coli</i>	chromosome			IS91-V9	4							←; →; →
<i>E. coli</i>	chromosome			IS91-V9	4							←; →; →
<i>E. coli</i>	chromosome			IS91-V10	4							←; ←; ←; →
<i>E. coli</i>	plasmid	IncFII	4	IS91-V8	4							←; ←; →
<i>E. coli</i>	plasmid	IncFII		IS91-V9	4							←; ←; →
<i>E. coli</i>	plasmid	IncFII		IS91-V9	4							←; ←; →
<i>E. coli</i>	plasmid	IncFII	5	IS91-V9	4	IS91-V101	1					←; ←; ←; ←
<i>E. coli</i>	plasmid	IncFII		IS91-V9	4	IS91-V101	1					←; ←; ←; ←
<i>E. coli</i>	plasmid	IncFII	7	IS91-V10	1	IS91-V77	3	IS91-V24	3			→; ←; ←; ←; →
<i>E. coli</i>	plasmid	IncFII		IS91-V8	7							←; →; →; →; →
<i>E. coli</i>	plasmid	IncFII	9	IS91-V7	3	IS91-V23	2	IS91-V11	2	IS91-V10	2	→; ←; ←; →; ←; ←; →

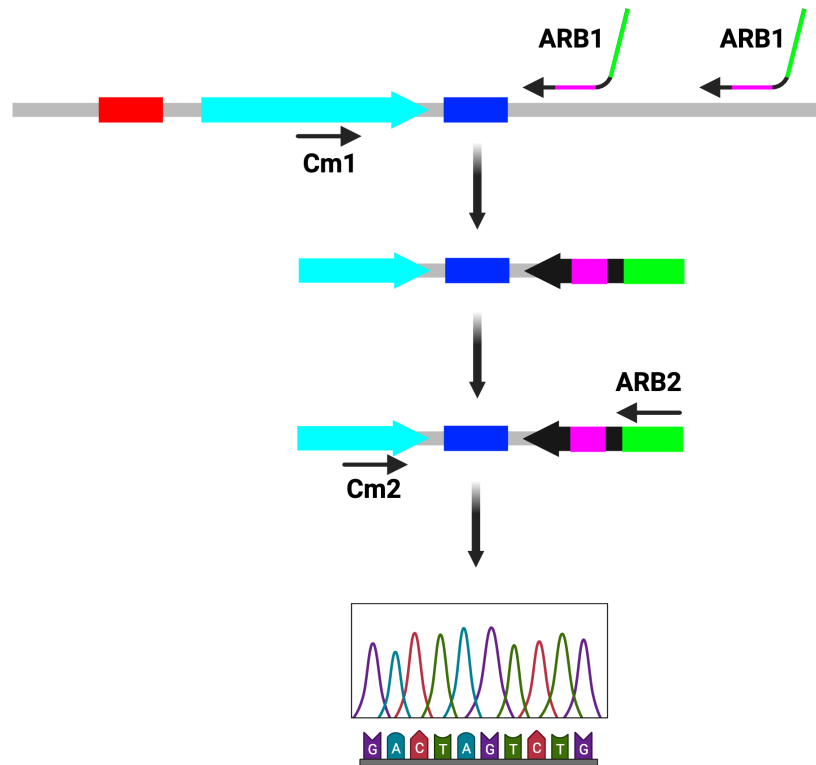
## Annexe 4.2. Réplicon à plusieurs copies de d'élément IS1294b

	Genre	Support Génétique/replication initiator	Nombre copies	Premier variant		Second variant		Troisième variant		Quatrième variant		Cinquième variant		Sixième variant		Représentation
				Variant	Nb. de copies	Variant	Nb. de copies	Variant	Nb. de copies	Variant	Nb. de copies	Variant	Nb. de copies	Variant	Nb. de copies	
AY87342.1	<i>Shigella flexneri</i>	plasmid		IS1294b-V118	2											→;→
CABE12010000170.1	<i>Klebsiella pneumoniae</i>	plasmid		IS1294b-V199	2											→;→
CP000039.1	<i>Shigella sonnei</i>	plasmid	<i>incFII</i>	IS1294b-V176	1	IS1294b-V186	1									→;←
CP045785.1	<i>Shigella dysenteriae</i>	plasmid	<i>incFII</i>	IS1294b-V126	2											←;→
CP091399.1	<i>Escherichia coli</i>	plasmid		IS1294b-V82	1	IS1294b-V175	1									→;→
UGLH01000003.1	<i>Klebsiella pneumoniae</i>	plasmid		IS1294b-V208	2											←;←
UGLI01000002.1	<i>Klebsiella pneumoniae</i>	plasmid		IS1294b-V206	2											←;←
NZ_ULGF01000046.1	<i>Escherichia coli</i>	plasmid		IS1294b-V8	1	IS1294b-V1	1									→;→
UGLV01000002.1	<i>Klebsiella pneumoniae</i>	plasmid		IS1294b-V219	2											→;→
UGMD01000001.1	<i>Klebsiella pneumoniae</i>	plasmid		IS1294b-V206	2											←;←
UGLN01000002.1	<i>Klebsiella pneumoniae</i>	chromosome		IS1294b-V204	1	IS1294b-V217	1									←;←
NZ_UGS01000007.1	<i>Escherichia coli</i>	plasmid	<i>incFII</i>	IS1294b-V82	1	IS1294b-V87	1									←;←
NZ_KU355873.1	<i>Escherichia coli</i>	plasmid	<i>incFII</i>	IS1294b-V1	2											→;→
DAEBF001000042.1	<i>Escherichia coli</i>	plasmid	<i>incFII</i>	IS1294b-V28	2											←;←
AF348706.1	<i>Shigella flexneri</i>	plasmid	<i>incFII</i>	IS1294b-V118	2											→;←
NC_002698.1	<i>Shigella flexneri</i>	plasmid	<i>incFII</i>	IS1294b-V118	2											→;←
F0203500.1	<i>Klebsiella pneumoniae</i>	plasmid		IS1294b-V199	2											→;→
NC_019197.1	<i>Shigella flexneri</i>	plasmid	<i>incFII</i>	IS1294b-V118	2											→;←
NC_021231.1	<i>Klebsiella pneumoniae</i>	plasmid	<i>ColE1</i>	IS1294b-V199	2											→;→
NC_016833.1	<i>Shigella sonnei</i>	plasmid	<i>incFII</i>	IS1294b-V159	1	IS1294b-V165	1									→;←
NC_007385.1	<i>Shigella sonnei</i>	plasmid	<i>incFII</i>	IS1294b-V176	1	IS1294b-V186	1									→;←
NZ_AP022354.1	<i>Escherichia coli</i>	plasmid	<i>incFII</i>	IS1294b-V1	2											→;←
NZ_AP022355.1	<i>Escherichia coli</i>	plasmid	<i>incFII</i>	IS1294b-V1	2											←;→
NC_024996.1	<i>Shigella flexneri</i>	plasmid	<i>incFII</i>	IS1294b-V104	1	IS1294b-V118	1									←;←
NZ_CALSCT01000004.1	<i>Escherichia coli</i>	chromosome		IS1294b-V1	2											←;→
NZ_CP010830.1	<i>Shigella sonnei</i>	plasmid	<i>incFII</i>	IS1294b-V176	2											←;→
NZ_CALSCT01000001.1	<i>Escherichia coli</i>	plasmid		IS1294b-V1	2											→;→
NZ_CA_FBB01000009.1	<i>Shigella flexneri</i>	plasmid	<i>incFII</i>	IS1294b-V121	1	IS1294b-V177	1									←;→
NZ_CA_FAZ01000008.1	<i>Shigella flexneri</i>	plasmid	<i>incFII</i>	IS1294b-V121	1	IS1294b-V177	1									←;→
NZ_BNY01000006.1	<i>Escherichia coli</i>	plasmid	<i>incFII</i>	IS1294b-V1	2											←;←
NZ_CP026787.1	<i>Shigella dysenteriae</i>	plasmid	<i>incFII</i>	IS1294b-V215	1	IS1294b-V221	1									←;←
NZ_CP026789.1	<i>Shigella flexneri</i>	plasmid	<i>incFII</i>	IS1294b-V125	2											→;←
NZ_CP026794.1	<i>Shigella flexneri</i>	plasmid	<i>incFII</i>	IS1294b-V125	2											←;→
NZ_CP026800.1	<i>Shigella flexneri</i>	plasmid	<i>incFII</i>	IS1294b-V104	2											←;→
NZ_CP053752.1	<i>Shigella sonnei</i>	plasmid	<i>incFII</i>	IS1294b-V181	1	IS1294b-V176	1									→;←
NZ_CP037924.1	<i>Shigella flexneri</i>	plasmid	<i>incFII</i>	IS1294b-V118	2											→;←
NZ_CP04189.1	<i>Salmonella enterica</i>	plasmid	<i>IncC_1</i>	IS1294b-V1	2											→;←
NZ_CP026829.1	<i>Shigella dysenteriae</i>	plasmid	<i>incFII</i>	IS1294b-V153	2											→;→
NZ_JANKV01000002.1	<i>Shigella sonnei</i>	plasmid	<i>incFII</i>	IS1294b-V176	1	IS1294b-V149	1									←;←
NZ_JACKX010000030.1	<i>Escherichia coli</i>	plasmid		IS1294b-V113	1	IS1294b-V144	1									→;→
NZ_JANKX01000002.1	<i>Shigella flexneri</i>	plasmid	<i>incFII</i>	IS1294b-V177	1	IS1294b-V121	1									←;→
NZ_CP085203.1	<i>Shigella flexneri</i>	plasmid	<i>incFII</i>	IS1294b-V142	2											←;←
NZ_JANKZ01000002.1	<i>Shigella flexneri</i>	plasmid	<i>incFII</i>	IS1294b-V118	2											→;←
NZ_CP046285.1	<i>Shigella sonnei</i>	plasmid	<i>incFII</i>	IS1294b-V159	1	IS1294b-V149	1									→;←
NZ_CP050863.1	<i>Escherichia coli</i>	plasmid	<i>incFII</i>	IS1294b-V124	1	IS1294b-V170	1									←;←
NZ_CP050866.1	<i>Escherichia coli</i>	plasmid	<i>incFII</i>	IS1294b-V124	1	IS1294b-V170	1									←;←
NZ_CP026767.1	<i>Shigella boydii</i>	plasmid	<i>incFII</i>	IS1294b-V99	1	IS1294b-V147	1									←;←
NZ_CP013026.1	<i>Escherichia coli</i>	plasmid	<i>incFII</i>	IS1294b-V1	2											→;→
NZ_CP019255.1	<i>Escherichia coli</i>	plasmid	<i>ColE1</i>	IS1294b-V45	1	IS1294b-V1	1									←;←
NZ_CP019696.1	<i>Shigella sonnei</i>	plasmid	<i>incFII</i>	IS1294b-V176	2											←;→
NZ_CP022068.1	<i>Salmonella enterica</i>	plasmid	<i>Col</i>	IS1294b-V9	2											←;←
NZ_CA_FAT01000011.1	<i>Shigella flexneri</i>	plasmid	<i>incFII</i>	IS1294b-V121	1	IS1294b-V177	2									←;→;←
NZ_CA_FAW01000006.1	<i>Shigella flexneri</i>	plasmid	<i>incFII</i>	IS1294b-V121	1	IS1294b-V177	2									←;→;←
NZ_CA_FAW01000004.1	<i>Shigella flexneri</i>	plasmid	<i>incFII</i>	IS1294b-V121	1	IS1294b-V177	2									←;→;←
NZ_CA_FAZ01000004.1	<i>Shigella flexneri</i>	plasmid	<i>incFII</i>	IS1294b-V121	1	IS1294b-V177	2									←;→;←
NZ_CA_FBA01000005.1	<i>Shigella flexneri</i>	plasmid	<i>incFII</i>	IS1294b-V121	1	IS1294b-V177	2									←;→;←
NZ_CA_FAY01000005.1	<i>Shigella flexneri</i>	plasmid	<i>incFII</i>	IS1294b-V177	2	IS1294b-V121	1									←;→;→
NZ_CALSCT01000003.1	<i>Escherichia coli</i>	plasmid	<i>ColE1</i>	IS1294b-V177	3											←;←;←
NZ_CP023646.1	<i>Shigella sonnei</i>	plasmid	<i>incFII</i>	IS1294b-V176	2	IS1294b-V149	1									←;→;←
NZ_CP024472.1	<i>Shigella flexneri</i>	plasmid	<i>incFII</i>	IS1294b-V128	1	IS1294b-V116	1	IS1294b-V139	1							←;→;→
NZ_CP026826.1	<i>Shigella dysenteriae</i>	plasmid	<i>incFII</i>	IS1294b-V141	3											←;→;→
NZ_CP109776.1	<i>Shigella sonnei</i>	plasmid	<i>incFII</i>	IS1294b-V159	2	IS1294b-V149	1									←;→;←
NZ_JANKU01000001.1	<i>Shigella sonnei</i>	plasmid	<i>incFII</i>	IS1294b-V149	1	IS1294b-V176	2									←;→;→
NZ_LR878366.1	<i>Shigella flexneri</i>	plasmid	<i>incFII</i>	IS1294b-V121	1	IS1294b-V177	2									←;→;←
NZ_UGY001000002.1	<i>Shigella flexneri</i>	plasmid	<i>incFII</i>	IS1294b-V113	1	IS1294b-V152	1	IS1294b-V139	1							←;→;→
NZ_CP011417.1	<i>Escherichia coli</i>	plasmid	<i>incFII</i>	IS801-V87	1	IS801-V78	1	IS801-V137	1	IS801-V80	1					←;←;←;→
NZ_CA_FAP01000003.1	<i>Shigella flexneri</i>	plasmid	<i>incFII</i>	IS801-V121	2	IS801-V177	2									←;→;←;←
NZ_CP04199.1	<i>Salmonella enterica</i>	plasmid		IS801-V71	4											→;←;←;←
NZ_CP025857.1	<i>Escherichia coli</i>	plasmid	<i>incFII</i>	IS801-V9	4											←;→;→;→
NZ_CP065614.1	<i>Escherichia coli</i>	plasmid	<i>incFII</i>	IS801-V109	1	IS801-V108	1	IS801-V111	1	IS801-V112	1					←;→;→;←
NZ_LCS20276.1	<i>Escherichia coli</i>	plasmid	<i>IncFIIb</i>	IS801-V1	4											→;←;→;←
UGFN01000002.1	<i>Escherichia coli</i>	plasmid	<i>incFII</i>	IS801-V130	1	IS801-V183	1	IS801-V129	1	IS801-V174	1					→;←;→;→
NZ_UGY01000002.1	<i>Shigella flexneri</i>	plasmid	<i>incFII</i>	IS801-V139	1	IS801-V188	1	IS801-V139	1	IS801-V105	1					←;←;→;→
NZ_UGFA01000001.1	<i>Escherichia coli</i>	plasmid	<i>incFII</i>	IS801-V111	1	IS801-V112	1	IS801-V109	1	IS801-V108	1					→;←;→;→
UFYX01000004.1	<i>Escherichia coli</i>	plasmid	<i>incFII</i>	IS801-V156	1	IS801-V92	1	IS801-V172	1	IS801-V157	1	IS801-V477	1			←;←;→;→
NZ_LR882988.1	<i>Escherichia coli</i>	plasmid	<i>incFII</i>	IS801-V71	5											→;←;←;→;→
NZ_JAAR001000001.1	<i>Escherichia coli</i>	plasmid	<i>incFII</i>	IS801-V108	2	IS801-V75	1	IS801-V146	1	IS801-V123	1					→;→;←;→;→
NZ_CP065612.1	<i>Escherichia coli</i>	plasmid	<i>incFII</i>	IS801-V74	1	IS801-V150	1	IS801-V137	1	IS801-V70	1	IS801-V179	1	IS801-V181	1	→;←;→;→;→;→
NZ_UFYN01000003.1	<i>Escherichia coli</i>	plasmid	<i>incFII</i>	IS801-V74	1	IS801-V203	1	IS801-V179	1	IS801-V70	1					

### Annexe 4.3. Réplicon à plusieurs copies de d'élément IS801

	Genre	Support Génétique	replication initiator	Nombre copies	Premier variant		Représentation
					Variant	Nb. de copies	
NZ_JAANQA010000004.12	Pseudomonas savastanoi	plasmid		2	IS801-V2	2	→ ; →
NZ_JAANPZ010000004.12	Pseudomonas savastanoi	chromosome			IS801-V16	2	→ ; →
CP000058.1	Pseudomonas savastanoi	chromosome		3	IS801-V16	3	← ; ← ; ←
NC_005773.3	Pseudomonas savastanoi	chromosome			IS801-V16	3	← ; ← ; ←

## Annexe 5. Schéma de la PCR arbitraire



Created with BioRender.com

La mini-IS est représenté par un rectangle bleu, un flèche cyan et un rectangle rouge indiquant l'extrémité oriS, le gène *cmR* et l'extrémité *terIS*, respectivement. Le trait indique par exemple la matrice d'ADN double brin du plasmide rapporteur pOX38Km. Un premier round de PCR est effectué avec une amorce dans le gène *cmR* (Cm1) et une seconde amorce appelée ARB1 qui contient un hexamère fortement représenté dans pOX38Km, 10 nucléotides « dégénérés » : toutes les possibilités pour chaque nucléotide (rose), 20 nucléotides conservés (vert). Ce premier round permet de cartographier les insertions dans le plasmide rapporteur. Ensuite, un second round est effectué pour favoriser le produit obtenu. Pour ce faire, une amorce plus externe dans le gène *cmR* (Cm2) est utilisée et une seconde amorce appelée ARB2 qui reconnaît la séquence de 20 nucléotides conservée (vert). Ce nouveau produit favorisé est ensuite envoyé au séquençage Sanger afin d'identifier les sites cibles d'insertion.

## Annexe 6. Liste des amorces utilisées lors de cette étude

Les sites de restriction sont en gras.

Oligonucléotide	Séquences d'oligonucléotides (5' → 3')	Utilisation
<b>Formation des linkers</b>		
CP21	<b>AATCCGAGTAGGCAGCCTGGCGGCTGCGGCTTGT</b> ATGGTCTGGAATTACCGTTATAAAAAAGATAATGTCA TTGTCTTTCAGGTAGT <b>G</b>	Formation du linker <i>terIS91</i> (construction de pCP6 et de pAF24)
CP22	<b>GATCCACTACCTGAAAGACAATGACATTATCTTTTT</b> ATAACGGTAATCCAGACCATGACAAGCCGCAGCCGC CAGGCTGCCTACT <b>CGG</b>	
CP60	<b>GCCGGCAGTTAAAATTGGGGCTTGAGCGGTAAACGA</b> GTGAGGGAATTCAGGTAAGATACTTCGGATGAGGAG CAAAAAGGTGGTTTATACTTCTATACCC <b>CAAGA</b>	Formation du linker <i>oriISCR1</i> (construction de pAF24)
CP61	<b>AGTCTTGGGGTATAGGAAGTATAAACACCTTTTTG</b> CTCCTCATCCGAAGTATCTTACCTGAAATTCCTCACT CGTTTACCGCTCAAGCCCAATTTAACTGCCGG <b>CTG</b> <b>CA</b>	
CP34	<b>AATCCGGGAGTGACGGGCACTGGCTGGCAATGTCTA</b> GCAACGGCAGGCATTTTCGGCTGAGGGTAAAAGAACTT TCCGCTAAGCGATAGACTGTATGTAACACAGTATTGC AAGGACGCGGA <b>ACG</b>	Formation du linker <i>terISCR2</i> (construction de pCP67)
CP35	<b>GATCCGTTCCGCGTCCTTGCAACTGTGTTTACATAC</b> AGTCTATCGCTTAGCGGAAAGTTCTTTTACCCTCAGCC GAAATGCCTGCCGTTGCTAGACATTGCCAGCCAGTGC CCGTC <b>ACTCCG</b>	
CP117	<b>AGTCTTGGACGTATAGGAAGAATAAACGCCCTTTTCA</b> CCCAAGTCCAACAGCTTTGGACCGCAGTTGACTCTTT CGACACCCCTGCGATGCAACCCAATCCGG <b>CTGCTGC</b> <b>A</b>	Formation du linker <i>oriISCR2</i> (construction de pCP67)
CP118	<b>GCAGCCGGATTGGGTTGCATCGCAGGGGTGTCGAAA</b> GAGTCAACTGCCGTTCAAAGCTGTTGGACTTGGGTGA AAAGGGCGTTTATTCTTCTATACGTGA <b>ACA</b>	
<b>Construction des plasmides</b>		
<b>Étude des domaines I et VI de TnpA<sub>IS91</sub></b>		
AF30	TGACAAAATCATGGCAAGTGGGACCACGCTGATGGG CTAC	Insertion d'une mutation ponctuelle dans le domaine I C41S (construction de pAF13)
AF31	CCATCAGCGTGGTCCCACCTGCCATGATTTTTG <b>CTACT</b> GAC	
AF40	TGGGCTACACGCAATGGAGTTGTTGTCACCGGACTG TTGCC	Insertion d'une mutation ponctuelle dans le domaine I C52S (construction de pAF18)
AF41	CAGTCCGGTGACGAACA <b>ACTCCATTGCGTGTAGCCCA</b> TCAGC	
AF42	TGATGGGCTACACGCAATGGAGTAGTTCGTCACCGGA CTGTTGCC	Insertion d'une double mutation ponctuelle dans le domaine I C52/53S (construction de pAF19)
AF43	CAGTCCGGTGACGA <b>ACTACTCCATTGCGTGTAGCCCA</b> TCAGCGTG	
AF26	GCAACAGTCCGGTGACGA <b>ACTGCACCATTGCGTGTAG</b> CCCATC	Insertion d'une mutation ponctuelle dans le domaine I C53S (construction de pAF11)
AF27	GGCTACACGCAATGGTGCAGTT <b>CGTCACCGGACTGTT</b> GCCAC	
AF28	AGACCTTTTTGGTGTGGCA <b>ACTGTCCGGTGACGAACA</b> GCACCA	Insertion d'une mutation ponctuelle dans le domaine I C58S (construction de pAF12)
AF29	GCTGTT <b>CGTCACCGGACAGTTGCCACACCAAAAAGGT</b> CTGCTTCC	
AF34	TGTT <b>CGTCACCGGACAGTAGTCACACCAAAAAGGTCT</b> GC	Insertion d'une double mutation ponctuelle dans le domaine I C58/59S (construction de pAF15)
AF35	CCTTTTTGGTGTGACTACTGTCCGGTGACGAACAGCA CC	
AF32	TTCGTCACCGGACTGTAGTCACACCAAAAAGGTCTGC TTCCGG	Insertion d'une mutation ponctuelle dans le domaine I C59S (construction de pAF14)
AF33	AAGCAGACCTTTTTGGTGTGACTACAGTCCGGTGACG AACAGC	
AF48	TTCGTCACCGGACTGTTGCGCAACCAAAAAGGTCTGC TTCCGG	Insertion d'une mutation ponctuelle dans le domaine I H60A (construction de pAF27)
AF49	AAGCAGACCTTTTTGGTGTGCGCAACAGTCCGGTGACG AACAGC	

AF36	GTTGCCACACCAAAAAGGTCAGTTTCCGGTGTAAAAG CCGC	Insertion d'une mutation ponctuelle dans le domaine I C65S (construction de pAF16)
AF37	TTTTACACCGGAAACTGACCTTTTTGGTGTGGCAACAG TCCG	
AF15	AAAAGGTCTGCTTCCGGAGTAAAAGCCGCTCCTGTCC GCAC	Insertion d'une mutation ponctuelle dans le domaine I C68S (construction de pAF8)
AF16	GGACAGGAGCGGCTTTTACTCCGGAAGCAGACCTTTT TGG	
AF19	TCCGGTGTAAAAGCCGCTCCAGTCCGCACTGCGGGG TGAAG	Insertion d'une mutation ponctuelle dans le domaine I C73S (construction de pAF9)
AF20	TCACCCCGCAGTGC GGACTGGAGCGGCTTTTACACC GGAAGC	
AF38	CTCCTGTCCGCACAGTGGGGTGAAGGCTGGCGCACA G	Insertion d'une mutation ponctuelle dans le domaine I C76S (construction de pAF17)
AF39	CCAGCCTTACCCCACTGTGCGGACAGGAGCGGCTT TTAC	
AF50	TGAAGGTTGACCCGCTGAAGAGTATTCTGTGCGGAAG TCAGATGCG	Insertion d'une mutation ponctuelle dans le domaine VI C360S (construction de pAF28)
AF51	TGACTTCCGCACAGAATACTCTTACGCGGGTCAACCT TCAGTAAC	
AF52	TAAAACGCATCTGACTTCCACTCAGAATGCACTTCAGC GGGTCAACC	Insertion d'une mutation ponctuelle dans le domaine VI C363S (construction de pAF29)
AF53	ACCCGCTGAAGTGCATTCTGAGTGGAAGTCAGATGCG TTTTACGGG	
AF83	TGACTTCCACTCAGAATACTCTTACGCGGGTCAACCTT CAGTAAC	Insertion d'une double mutation ponctuelle dans le domaine VI C360/363S (construction de pAF40)
AF84	TGAAGGTTGACCCGCTGAAGAGTATTCTGAGTGGAAG TCAGATGCG	
CP15 CP16	TATACATATGCTTCCCCGTTTTGCCGAC GCGTGGATCCTATTAACAGGATGACACTG	Amplification du fragment de fusion Junc des différentes mutations ponctuelles des domaines I et VI dans la CDS de la <i>tnpA<sub>IS91</sub></i>
<b>Mutation du motif CxxC de Orf121 et TnpA<sub>IS91</sub> (domaine VI)</b>		
AF60	CAGCGCCGGCGGCATGTACAGTCCGGAGTGTGGCAC GGTTCAC	Insertion d'une mutation ponctuelle dans le motif CxxC de Orf121 C105S (dérivé du pCP40 ; construction de pAF30)
AF61	ACCGTGCCACACTCCGGACTGTACATGCCGCCGGCG CTGATAC	
AF62	GGCATGTACTGCCCGGAGAGTGGCACGGTTCCTG CCGGACGG	Insertion d'une mutation ponctuelle dans le motif CxxC de Orf121 C108S (dérivé du pCP40 ; construction de pAF31)
AF63	CGGCCAGTGAACCGTCCACTCTCCGGGCAGTACAT GCCGCCG	
AF50	TGAAGGTTGACCCGCTGAAGAGTATTCTGTGCGGAAG TCAGATGCG	Insertion d'une mutation ponctuelle dans le domaine VI C360S (dérivé du pCP40 ; construction de pAF32)
AF51	TGACTTCCGCACAGAATACTCTTACGCGGGTCAACCT TCAGTAAC	
AF52	TAAAACGCATCTGACTTCCACTCAGAATGCACTTCAGC GGGTCAACC	Insertion d'une mutation ponctuelle dans le domaine VI C363S (dérivé du pCP40 ; construction de pAF33)
AF53	ACCCGCTGAAGTGCATTCTGAGTGGAAGTCAGATGCG TTTTACGGG	
CP58 CP16	TATACATATGGCCGTTTCAAGCTAAACC GCGTGGATCCTATTAACAGGATGACACTG	Amplification du CDS de <i>orf121-tnpA<sub>IS91</sub></i> contenant les mutations ponctuelles CxxC (dérivé du pCP40)
CP40	GCGCAAGCTTGATCGATAGGAATTTAAATCC	
<b>Étude des mutations ponctuelles des éléments ISK<sub>pn22</sub>, ISCR1 et ISCR2</b>		
CP63 CP64	TCTGCATATGGACAACTGACACTAC TATAGGATCCTCATTAAAGCGGGGCCG	Amplification de la CDS <i>tnpA</i> de ISK <sub>pn22</sub>
AF17	TATTGTAGTAACGGTCATTGTTGCGGAGTCTTTATAA CTC	Insertion d'une mutation ponctuelle dans le domaine I L54C (construction de pAF7)
AF18	TTATAAAAGACTCCGCAACAATGACCGTTACTACAATA TTG	
CP63 CP64	TCTGCATATGGACAACTGACACTAC TATAGGATCCTCATTAAAGCGGGGCCG	Amplification du fragment de fusion Junc de la mutation ponctuelle L54C des domaines I de la CDS de la <i>tnpA<sub>ISKpn22</sub></i>
CP1 CP2	TGTACATATGTCGCTGGCAAGGAACGCC TGGAGCATGCTCAGTCAAAGAGACGACTC	Amplification de la CDS <i>tnpA</i> de ISCR1
AF11	CAGCGGGATAAGCTTGAGTACTTGTGTCGTACATTA GTCG	Insertion d'une mutation ponctuelle dans le domaine IV R325Y (construction de pAF5)
AF12	ACTAATGTAGCGACACAAGTACTCAAGCTTATCCCGCT GATG	



AF44	GCTAAATCTCAATGTCCACATTCACATGCTGTTTCTCG ATGG	Insertion d'une mutation ponctuelle dans le domaine III Y190I (construction de pAF21)
AF45	GAGAAACAGCATGTGAATGTGGACATTGAGATTTAGC GCGG	
CP1	TGTACATATGTCGCTGGCAAGGAACGCC	Amplification du fragment de fusion Junc contenant la ou les mutations ponctuelles des domaines III et/ou IV de la CDS de la <i>tnpA<sub>rrc1</sub></i>
CP2	TGGAGCATGCTCAGTCAAAGAGACGACTC	
CP5	TAGTCATATGCCTCATGTGGCGGCCAGGAC	Amplification de la CDS <i>tnpA</i> de ISCR2
CP6	TCCAGCATGCTCAGTCAAACAGACCCAGG	
AF13	GAACGCAAGAAGCTCGAATACCTGTGCCGGTACATCA GCCG	Insertion d'une mutation ponctuelle dans le domaine IV R322Y (construction de pAF6)
AF14	CTGATGTACCGGCACAGGTATTGAGCTTCTTGCCTT CATC	
AF46	CTCAATCTGAATGTTACATTCACATGCTGTTTCTCGA CGG	Insertion d'une mutation ponctuelle dans le domaine III F189I (construction de pAF23)
AF47	GAGAAACAGCATGTGAATGTGAACATTCAGATTGAGC GCCG	
CP5	TAGTCATATGCCTCATGTGGCGGCCAGGAC	Amplification du fragment de fusion Junc contenant la ou les mutations ponctuelles des domaines III et/ou IV de la CDS de la <i>tnpA<sub>rrc2</sub></i>
CP6	TCCAGCATGCTCAGTCAAACAGACCCAGG	
<b>Amorce arbitraire (AP)-PCR</b>		
ARB1	GGCCTCGCGTCGACTACTTCNNNNNNNNNCCGGCA	Première série de PCR arbitraires pour cartographier les sites d'insertion IS91 dans pOX38Km du côté <i>oriS</i>
ARB2	GGCCTCGCGTCGACTACTTC	Deuxième série de PCR arbitraires pour cartographier les sites d'insertion IS91 dans pOX38Km et dans le chromosome de <i>E. coli</i> du côté <i>oriS</i>
Cm-1005	GCTGATGCCGCTGGCGATTC	Première série de PCR arbitraires du côté <i>oriS91</i> , <i>oriISCR1</i> et <i>oriISCR2</i> dans pOX38Km avec ARB1
Cm-1049	GCTTCCATGTCCGCAGATGC	Deuxième série de PCR arbitraires du côté <i>oriS91</i> , <i>oriISCR1</i> et <i>oriISCR2</i> dans pOX38Km avec ARB2
Ori91-1104	GCAGTCCGCTCATATGGTGC	Pour le séquençage Sanger du côté <i>oriS91</i>
OriCR1-1129	AGCGGTAACGAGTGAGG	Pour le séquençage Sanger du côté <i>oriISCR1</i>
Cm-473	ATCAACGGTGGTATATCCAGTG	Première série de PCR arbitraires du côté <i>terIS91</i> dans pOX38Km avec ARB1
Ter91-337	CAACGTCTCAGGATCCACTACCTG	Deuxième série de PCR arbitraires du côté <i>terIS91</i> dans pOX38Km avec ARB2 (pCP7)
Ter91-276	TCCAGACCATGACAAGCCGC	Pour le séquençage Sanger du côté <i>terIS91</i> (pCP7)
Ter91-318	CTACCTGAAAGACAATGAC	Pour le séquençage Sanger du côté <i>terIS91</i> (pCP39)
Ter2A	CGTTCGCGTCCTTGCAATACTGTG	Première série de PCR arbitraires du côté <i>terISCR2</i> dans pOX38Km avec ARB1
Ter2B	CTCAGCCGAAATGCCTGC	Deuxième série de PCR arbitraires du côté <i>terISCR2</i> dans pOX38Km avec ARB2

## Étude des modalités de transposition des séquences d'insertion bactériennes des familles IS91-ISCR

---

Les éléments de la famille IS91/ISCR constituent une famille hautement associée à des gènes de virulence et de résistance aux antibiotiques. Cette famille d'IS atypique code une transposase HUH médiant les événements de transposition en utilisant un mécanisme de transposition en cercle roulant. Contrairement aux autres membres de la famille, IS91 contient un petit ORF codant potentiellement pour un polypeptide de 121 acides aminés, Orf121, en amont du gène de la transposase, *tnpA*. La première partie de ce travail a consisté à une étude *in silico* de deux éléments de la famille IS91, IS91 et IS1294b présentent des séquences conservées au niveau de la transposase et des régions *terIS* et *oriIS*. La deuxième partie du travail a été expérimentale concernant la régulation et l'efficacité de la transposition. Nous avons montré que i) la transcription du gène *tnpA* provient principalement du promoteur P<sub>orf121</sub> et que la transcription des deux gènes est couplée, ii) l'expression de la protéine Orf121 en cis ou en trans a un effet inhibiteur sur la transposition *in vivo* de IS91, iii) Orf121 est nécessaire à la reconnaissance et au clivage précis de l'extrémité *terIS91* pour limiter la mobilisation de l'ADN adjacent. En ce qui concerne la transposition de IS91, nous avons montré que seuls les intermédiaires circulaires ADN simple brin peuvent être insérés dans une nouvelle séquence cible et avons identifié deux doigts de zinc essentiels pour l'activité de la transposition, appelés ZF1 impliquant les cystéines 41, 68, 73, 76 et ZF2 impliquant les cystéines, 53, 58, 360 et 363. Enfin, nous avons démontré que les transposases des éléments de la famille IS91 (IS91, ISKnp22) et de la famille ISCR (ISCR1) sont capables de se mobiliser et de reconnaître et cliver l'extrémité *oriIS* de IS91, ISCR1 et ISCR2 et l'extrémité *terIS* de IS91 et ISCR2.

---

Mots-clés : Séquences d'insertions, IS91/ISCR, transposition, régulation, Orf121

### Study of transposition modalities for bacterial insertion sequences of the IS91-ISCR families

---

The IS91/ISCR family is highly associated with virulence and antibiotic resistance genes. This atypical IS family encodes a HUH transposase mediating transposition events using a rolling circle transposition mechanism. Unlike the other members of the family, IS91 contains a small ORF potentially encoding a 121 amino acid polypeptide, Orf121, upstream of the transposase gene, *tnpA*. The first part of this work consisted of an *in silico* study of two members of the IS91 family, IS91 and IS1294b, that have conserved sequences of the transposase and *terIS* and *oriIS* regions. The second part of this work was experimental and focused on the transposition efficiency and regulation. We have shown that i) transcription of the *tnpA* gene originates mainly from the P<sub>orf121</sub> promoter and that transcription of the two genes is coupled, ii) expression of the Orf121 protein *in cis* or *in trans* has an inhibitory effect on *in vivo* transposition of IS91, iii) Orf121 is required for recognition and precise cleavage of the *terIS91* end to limit mobilization of the adjacent DNA. With regard to IS91 transposition, we showed that only single-stranded circular DNA intermediates can be inserted into a new target sequence and identified two zinc fingers essential for the transposase activity, called ZF1 involving cysteines 41, 68, 73 and 76 and ZF2 involving cysteines 53, 58, 360 and 363. Finally, we demonstrated that transposases from the IS91 family (IS91, ISKnp22) and the ISCR family (ISCR1) are able to mobilize and recognize and cleave the *oriIS* end of IS91, ISCR1 and ISCR2 and the *terIS* end of IS91 and ISCR2.

---

Keywords: Insertion sequences, IS91/ISCR, transposition, regulation, Orf121

