

Université de Montréal

**Horus :**  
**Création d'une plateforme CRISPR pour *Vibrio cholerae***

Par Clément Baret-Micaelli

Département de microbiologie, infectiologie et immunologie  
Faculté de médecine

Mémoire présenté  
En vue de l'obtention du grade de Maîtrise des Sciences  
En Microbiologie et Immunologie

Avril, 2022

© Clément Baret, 2022

# RÉSUMÉ

La mutagenèse dirigée est un outil indispensable à toute étude microbiologique, car elle permet d'identifier le rôle de certains locus génétiques identifiés comme acteurs potentiels dans des contextes précis. Cependant, les protocoles de mutagenèse dirigée sont longs et laborieux, et leur mise en œuvre est l'un des points limitants en recherche. L'émergence de CRISPR-Cas9 (clustered regularly interspaced short palindromic repeats) comme outil moléculaire a permis d'accélérer et de faciliter ces procédures de mutagenèse par contre-sélection. La limite de ces protocoles se situe dans la régénération de l'espaceur effectuant la contre-sélection.

Notre plateforme CRISPR, dénommée Horus, offre une solution à cette limitation. Elle utilise du clonage *in vivo* afin de raccourcir autant la durée que la charge de travail du protocole, pour aboutir à l'obtention de mutants en une seule étape. Pour ce faire nous avons conçu *in silico* un ARN guide synthétique capable d'agir comme un interrupteur génétique (porte logique ET) et de performer une contre sélection (discriminant les bactéries de types sauvages des mutants) via le système CRISPR-Cas9.

**Mots-clés:** *Vibrio cholerae*, CRISPR, Lambda Red, Ice Elements, Recombineering, Cas, MGE (Mobile Genetic Elements), MAGE (Multiplex Automated Genome Engineering)

# ABSTRACT

Site-directed mutagenesis is an essential tool for any microbiological study because it makes it possible to identify the role of certain loci identified as potential actors in specific contexts. However, site-directed mutagenesis protocols are long and laborious, and their implementation is one of the limiting points of research. The emergence of CRISPR-Cas9 (clustered regularly interspaced short palindromic repeats) as a molecular tool has accelerated the facilitation of these counter-selection mutagenesis protocols. The limitation of these protocols lies in the regeneration of the protospacer mediating the counter selection.

Our CRISPR platform, called Horus (HOMologous Recombination Using SsDNA), offers a solution to this limitation. It uses *in vivo* cloning to shorten both the duration and the workload of the protocol, allowing to obtain mutants strains in just one step. To do so, we designed *in silico* a synthetic guide RNA capable of acting as a genetic switch (AND Gate) and performing counter-selection (discriminating WT bacteria from mutants) via the CRISPR-Cas9 system.

**Keywords:** *Vibrio cholerae*, CRISPR, Lambda Red, Ice Elements, Recombineering, Cas, MGE, MAGE.

# Table des matières

RÉSUMÉ .....	ii
ABSTRACT.....	iii
Table des matières .....	iv
Liste des tableaux .....	vii
Liste des figures .....	viii
Liste des sigles et abréviations.....	ix
Remerciements.....	x
Chapitre 1 : Introduction.....	1
1.1.1 Historique de CRISPR-Cas .....	1
1.1.2 Architecture des systèmes CRISPR .....	5
1.1.3 Classification des systèmes CRISPR .....	6
1.1.3 Origine des systèmes CRISPR-Cas : Casposons.....	10
1.2 CRISPR comme système immunitaire procaryote.....	13
1.2.1 Adaptation immunitaire.....	13
1.2.1.1 Capture d'un espaceur .....	13
1.2.1.2 Intégration d'un espaceur .....	14
1.2.2 Expression et maturation des ARNcr.....	17
1.2.3 Interférence.....	18
1.2.3.1 Interférence PAM dépendante .....	18
1.2.3.2 Voie PAM indépendante .....	22
1.2.4 Systèmes Anti-CRISPR.....	25
1.2.5 Limites de l'argument immunitaire.....	26
1.2.6 CRISPR comme outil génétique .....	26
1.2.7 Mécanisme d'action de la Cas9 .....	28
1.3 Mutagenèse chez les procaryotes.....	30
1.3.1 Recombinaison homologue .....	31
1.3.2 Échange allélique.....	34
1.3.3 Recombinases Sites spécifiques.....	35
1.3.4 Contre sélection .....	36
1.3.5 Le génie génétique par recombinaison ( <i>Recombineering</i> ) .....	39
1.3.6 Emploi de l'opéron Red en mutagenèse.....	40
1.3.7 Considérations et limitations de l'opéron Red.....	44
1.4 Mutagenèse chez <i>V. cholerae</i> .....	45
1.4.1 CRISPR-Cas comme outil génétique chez les procaryotes .....	46

1.4.2 CRISPR No Scar.....	47
1.5 Problématique.....	52
Chapitre 2 : Matériel et méthodes .....	53
2.1 Souches utilisées.....	53
2.1.2 Réactifs et concentration .....	53
2.1.3 Milieux de cultures .....	53
2.1.4 Amorces .....	54
2.2 ADN Donneur mutagènes simples brins.....	54
2.2.1 ADN Donneurs double brins .....	56
2.2.2 Scaffold synthétiques .....	57
2.3 Transformation de plasmides chez <i>V.cholerae</i> .....	58
2.3.1 Transformation chimique de plasmides chez <i>E. coli</i> : .....	60
2.3.2 Éléctroporation de plasmides chez <i>E. coli</i> :.....	61
2.4 Cycle No-Scar .....	61
2.5 Clonage par enzymes de restrictions :.....	63
2.5.1 CPEC (Circular Polymerase Extension Cloning) .....	64
2.6 Prédiction de repliement .....	66
2.7 Test de dilution en série par gouttes (Spot assay).....	66
2.8 PCR sur colonies .....	67
Chapitre 3 : Résultats.....	68
3.1 Implémentation du protocole No-Scar dans <i>V. cholerae</i> A1552 .....	68
3.2 Optimisation éléctroporation de plasmides .....	69
3.3 Création d'un mutant <i>lacZ</i> de <i>V. cholerae</i> A1552 par le système No-Scar .....	72
3.4 Stabilité du système No-Scar.....	74
3.5 Création de la plateforme Horus (HOMologous Recombinaison Using SsDNA).....	78
3.5.1 Mécanisme de contre sélection de Horus .....	78
3.5.2 Résistance à la spectinomycine par <i>aadA</i> .....	83
3.5.3 Hérité du pKD sans sélection.....	84
3.5.4 Fragment Anubis .....	84
3.6 Création d'une cible suicide.....	87
3.6.1 pRaava.....	87
3.6.2 pKD <sup>2</sup> .....	91
3.6.3 pKD <sup>3</sup> .....	98
3.6.4 pHorusV1 .....	99
3.6.5 pHorusV74 .....	104
3.7 Escapers .....	107
3.8 Implémentation de Horus dans MG1655 .....	111

3.9 Création d'un mutant <i>lacZ</i> de MG1655 par HorusV74.....	112
Chapitre 4 :Discussion.....	117
4.1 Implémentation de la plateforme No-Scar chez <i>V. cholerae</i> .....	117
4.2 Création de la plateforme Horus.....	121
Chapitre 5 : Conclusion.....	133
Annexes .....	135
Bibliographie.....	141

## Liste des tableaux

Tableau 1. Séquences des amorces utilisées pour géotypage.....	54
Tableau 2. Séquence des ADN donneurs simples brins .....	55
Tableau 3. Séquence des ADN donneurs simple brin recombinant sur pRaava .....	55
Tableau 4. Séquence des ADN donneurs simple brin recombinant sur pHorusV1/V74 .....	56
Tableau 5 Séquence des amorces utilisées pour générer des ADN donneurs doubles brins.....	56
Tableau 6. Séquence des scaffold utilisés dans ce mémoire.....	57
Tableau 7. Séquence des amorces utilisées pour CPEC .....	65
Tableau 8. Stabilité du système du système No-Scar depuis un stock glycérol -80°C (Valeurs DO Figure 21).....	77
Tableau 9. Stabilité du système No-Scar depuis une transformation fraîche. (Valeurs DO Figure 22). .....	77
Tableau 10. Efficacité de contre sélection du pRaava .....	89
Tableau 11. Efficacité de contre sélection de pKD <sup>2</sup> .....	98

# Liste des figures

Figure 1. Chronologie de la découverte des systèmes CRISPR-Cas.....	4
Figure 2 Architecture d'un locus CRISPR. ....	6
Figure 3 Classification des systèmes CRISPR-Cas.....	9
Figure 4. Chronologie de l'évolution des systèmes CRISPR-Cas.....	12
Figure 5. Schéma général d'acquisition d'un espaceur CRISPR-Cas .....	16
Figure 6. Interférence PAM dépendante chez les types I, II, III. ....	21
Figure 7. Schéma des deux voies pour l'interférence pour le type I .....	24
Figure 8. Mécanisme d'action de la Cas9.....	29
Figure 9. Résolution des jonctions de Holliday. ....	33
Figure 10. Échange allélique contre sélectionné par <i>sacB</i> .....	38
Figure 11. Utilisation de l'opéron Red dans la procédure MAGE.....	43
Figure 12. Étapes du protocole No-Scar.....	51
Figure 13. Optimisation de l'efficacité de transformation .....	72
Figure 14. Génotypage PCR du mutant A1552 <i>lacZ</i> obtenu.....	73
Figure 15. Stabilité du système No-Scar depuis un stock glycérol -80°C. ....	75
Figure 16. Stabilité du système No-Scar depuis une transformation fraîche. ....	76
Figure 17. Schématisation du protocole par le pKD Horus .....	80
Figure 18. État initial de contre sélection par locus suicide.....	81
Figure 19. État recombiné du locus Horus.....	83
Figure 20. Action du donneur ADN simple brin Anubis sur le pKD LapLapX .....	86
Figure 21. Mécanisme d'action de pRaava .....	88
Figure 22 : Génotypage du plasmide pRaava après CPEC.....	89
Figure 23.pRaava non recombiné.....	90
Figure 24. Repliement du sgARN non modifié.....	92
Figure 25. Repliement du Scaffold modifié pKD <sup>2</sup> .....	95
Figure 26. Repliement du Scaffold de pKD <sup>2</sup> après recombinaison .....	96
Figure 27. Locus du Scaffold synthétique de pKD <sup>2</sup> .....	97
Figure 28. Repliement de Horus V1.....	101
Figure 29. Séquence et structure de HorusV1.....	101
Figure 30. Efficacité de contre-sélection HorusV1 sur DH5 $\alpha$ . ....	103
Figure 31. Repliement de Horus74.....	106
Figure 32. Efficacité de contre sélection de HorusV74 sur MG1655.....	107
Figure 33. Alignement de HorusV1 et un escaper .....	109
Figure 34. Alignement de HorusV1 et un escaper .....	109
Figure 35. Séquence d'une souche escaper HorusV74 après recombinaison.....	110
Figure 36. Séquences du pKD HorusV74 et du chromosome d'un mutant MG1655 $\Delta lacZ$ après recombinaison .....	116
Figure 37. Auto-incompatibilité de pRaava.....	126
Figure 38. Autocomptabilité de Horus .....	130
Figure 39. Repléments illégitimes de HorusV1 .....	136
Figure 40. Colonies satellites .....	137
Figure 41. Hérité du pKD sans sélection.....	138
Figure 42. Conception des ADN donneurs simple brins pour HorusV74.....	140

# Liste des sigles et abréviations

**ACR** : *Anti-CRISPR*

**ADN** : Acide désoxyribonucléique.

**ARN** : Acide ribonucléique.

**aTc** : *Tetracycline anhydre*

**CAS** : *CRISPR associated proteins*

**CASCADE**: *CRISPR-associated complex for antiviral defense*

**CMI**: *Concentration minimal inhibitrice*

**CMR**: *Repeat-associated Mysterious Proteins*

**CPEC**: *Circular Polymerase extension cloning*

**CRISPR**: *Clustered Regularly Interspaced Short Palindromic Repeats*

**DC** : *Direct culture*

**DO** : *Densité optique*

**EMTS** : *méthane sulfonate d'éthyle*

**ePAM**: *Escaper PAM*

**HSEP**: *High Stem Extended + PAM*

**MGEs** : *Éléments mobiles génétiques (Mobile genetic elements)*

**NO-SCAR**: *(Scarless Recombineering)*

**NTG**: *N-méthyl-N'-nitro-N-nitrosoguanidine*

**OD**: *Overday*

**ON**: *Overnight*

**OMV**: *Outer membrane vesicles*

**PAM**: *Proto adjacent motif*

**RAMP**: *Repair associated mysterious proteins*

**RBS**: *Ribosome binding site*

**RED**: *Recombinaison déficient*

**REP** : *Séquences palindromes.*

**SARS** : *Severe acute respiratory syndrome*

**SSRs** : *Site specific recombinases*

# Remerciements

Le cheminement que j'ai entrepris depuis le début de mes études graduées a été tant sur le fond scientifique que sur le plan personnel. Je remercie la Dre. Duperthuy de m'avoir donné la chance de réaliser ce projet.

Je remercie aussi mes collègues de laboratoire d'avoir été mes compagnons dans cette aventure qu'est la recherche scientifique, chacun d'eux ont su être des alliés importants dans mon quotidien. Je remercie également mes collègues étudiants, tant pour les conversations de couloirs tant que pour la richesse des échanges.

Je remercie mon île natale, qui me manque terriblement pour ses paysages comme pour ses habitants. Une nature si sauvage et imprégnée d'une beauté si profonde.

Naturellement, je remercie mes parents, mon frère et mes sœurs, je remercie la fortune de m'avoir donné une famille si haute en couleurs et pourtant si aimante. Je remercie aussi la bienveillance de mes beaux-parents. Vous m'avez permis de trouver de la chaleur dans un pays pourtant parfois si froid.

Je remercie Victor Mullins Dansereau d'être mon phare lorsque les tempêtes s'invitent dans mon esprit. Je remercie les cieux d'avoir fait en sorte que nos chemins se soient un jour croisés. Je remercie Fionna pour ses conseils judicieux et sa sagesse infinie.

Finalement, je salue les membres du jury pour être les derniers acteurs et non les moindres de ma maîtrise. Je vous présente humblement ce travail dans l'espoir de gagner mes gallons en tant que scientifique.

*Only a few know, how much one must know to know  
how little one knows.*

*-Werner Heisenberg*

# Chapitre 1 : Introduction

Quel est le point commun entre le VIH, la couleur des roses, certains mécanismes de résistance aux antibiotiques, et finalement CRISPR ? La réponse : Les transposons (1-4). Ces éléments génétiques mobiles laissent leurs griffes tout autour de nous. Ce mémoire traitera de CRISPR-Cas, le système immunitaire adaptatif bactérien provenant des casposons, une famille de transposons. CRISPR-Cas devenu l’emblème de la biologie moléculaire moderne en 2012 et est le sujet du prix Nobel en chimie en 2020.

Une image valant mille mots, une vidéo « *When phage attack !!* » expliquant le rôle de CRISPR en tant que système immunitaire est disponible

[ici](https://www.cell.com/cell/fulltext/S0092-8674(15)01316-1) ([https://www.cell.com/cell/fulltext/S0092-8674\(15\)01316-1](https://www.cell.com/cell/fulltext/S0092-8674(15)01316-1)).

## 1.1.1 Historique de CRISPR-Cas

Le premier locus CRISPR a été identifié en 1987 par Y. Ishino (Figure 1), chez *Escherichia coli* lors d’une étude portant sur l’un des gènes acteurs dans le métabolisme du phosphate, *iap* (5, 6) . Le séquençage d’un fragment d’ADN de 1.7kb codant pour *iap* retourna, en sus de la région codante, 5 séquences palindromiques (dyade) de 29 pbs séparés par 32bps, du côté 3’ du fragment. La particularité de ces répétitions fut discutée, mais par manque de banques de séquences, il fut impossible de déterminer le rôle de ces dernières.

En 1993 d’autres répétitions seront trouvées chez une archée (*Haloferax mediterranei*) (6, 7). Plus tard, ces mêmes répétitions atypiques ont été identifiées dans les génomes de nombreux procaryotes (Actinobactéries, *Mycobacterium tuberculosis*(8), *Salmonella enterica*, (9) *Shigella dysenteriae*). À cette époque, il a été proposé que ces séquences palindromiques (REP)

retrouvées chez *E. coli* et *S. Typhimurium* aidaient à stabiliser les ARNm (10), bien que ces répétitions n'eussent aucune homologie avec les REP connues à ce moment-là. Ces répétitions furent aussi trouvées dans des MGEs à l'intérieur d'archées, par exemple le plasmide naturel pNOB8 (11).

Le domaine de la biologie moléculaire est entré dans l'ère génomique en 1995, avec la publication du génome de *Haemophilus influenzae* séquencé dans son intégralité. Quelques années plus tard en 2004 s'achevait le projet du Génome humain, pour la toute première fois, l'intégralité du génome humain était à la disposition de la communauté scientifique. La mise au point de méthodes de séquences automatisées issues de ces projets a permis l'introduction d'une nouvelle discipline, celle de la génomique comparée et dans le même temps l'élaboration de banques de séquençage. Via ces banques il est possible de réaliser des études d'approche fonctionnelle pour comprendre la biologie des organismes dans leur ensemble. Il fallut donc attendre l'ère génomique pour mettre en lumière le potentiel d'outil génétique de CRISPR. Ces séquences palindromiques ont eu plusieurs noms, tout au long de leur découverte chez les procaryotes : SRSRs (12) (short regularly spaced repeats), SPIDRs (spacers interspersed direct repeats), et LCTRs (large cluster of tandem repeats) (13).

Le terme CRISPR (Clustered regularly interspaced short palindromic repeat), lui, fut finalement proposé par Jansen et al en 2002 lors de leurs travaux *in silico* sur ces séquences (14). Durant cette étude il apparut qu'il existait une association exclusive de plusieurs gènes aux librairies CRISPR et absentes des organismes procaryotes ne contenant pas ces systèmes. Ces gènes furent nommés Cas pour (Crispr **A**ssociated **P**roteins).

Les Cas ont été initialement identifiées comme protéines impliquées dans le métabolisme de l'ADN uniquement chez des archées thermophiles, les Cas furent nommés RAMPs (**R**epair

Associated Mysterious Proteins) (15). Les Cas 1 et 2 ne partageaient aucune homologie de sites actifs connus rendant la prédiction de leur fonction caduque. En revanche la protéine Cas3 possédait 7 sites actifs caractéristiques de la superfamille d'hélicase. La protéine Cas4 quant à elle possédait des sites semblables aux exonucléases RecB, membres du complexe RecBCD. Ce complexe intervient dans la réparation de l'ADN en digérant l'ADN au point de fracture jusqu'à rencontrer un site *chi* où se déclenche une recombinaison homologue(14).

En 2002, Tang et *al* ont remarqués, lors d'une étude sur l'atténuation par ARN (RNAi) chez une archée, *Archaeoglobus fulgidus*, que plusieurs longs ARN semblaient être codés par des loci CRISPR. Ces ARN étaient par la suite séquentiellement clivés en plus petits fragments pour aboutir à des ARN de 65-77 pb codants en autre pour un seul espaceur (16). Des études subséquentes ont démontré que ces processus existaient chez d'autres organismes, et même au sein d'un locus CRISPR provenant du plasmide conjugatif pKEF9 (17).

En 2005, plusieurs groupes ont pu mettre en évidence le rôle immunitaire qu'occupaient les CRISPR dans les génomes bactériens en pointant l'homologie de séquence entre les CRISPRs et le génome de plusieurs bactériophages. Dans la même année, la première plateforme de génie génétique était décrite par P.D Grégory (18). Cette plateforme employait des nucléases à doigt de Zinc (ZNFs) (19) capable de médier des insertions génétiques. Cette technologie fut finalement adaptée pour amener 10 ans plus tard à la création des TALENs (Transcription activator-like effector nucleases) (20, 21). En 2006 plus de 50 familles de Cas étaient rapportées (22, 23). L'hypothèse de la capacité immunitaire de CRISPR fut confirmée expérimentalement en 2007 chez *S. thermophilus* (24).

En 2012, le système CRISPR-Cas émanant de *Streptococcus pyogenes* fut utilisé *in vitro* pour digérer des ARNs viraux puis reprogrammé afin de démontrer que CRISPR-Cas9 était capable de reconnaître et couper des séquences codées par des espaceurs synthétiques. Les auteurs de ces séries d'expériences, Emmanuelle Charpentier et Jennifer A. Doudna se sont vu attribuer en 2020 le prix Nobel de Chimie (25, 26).

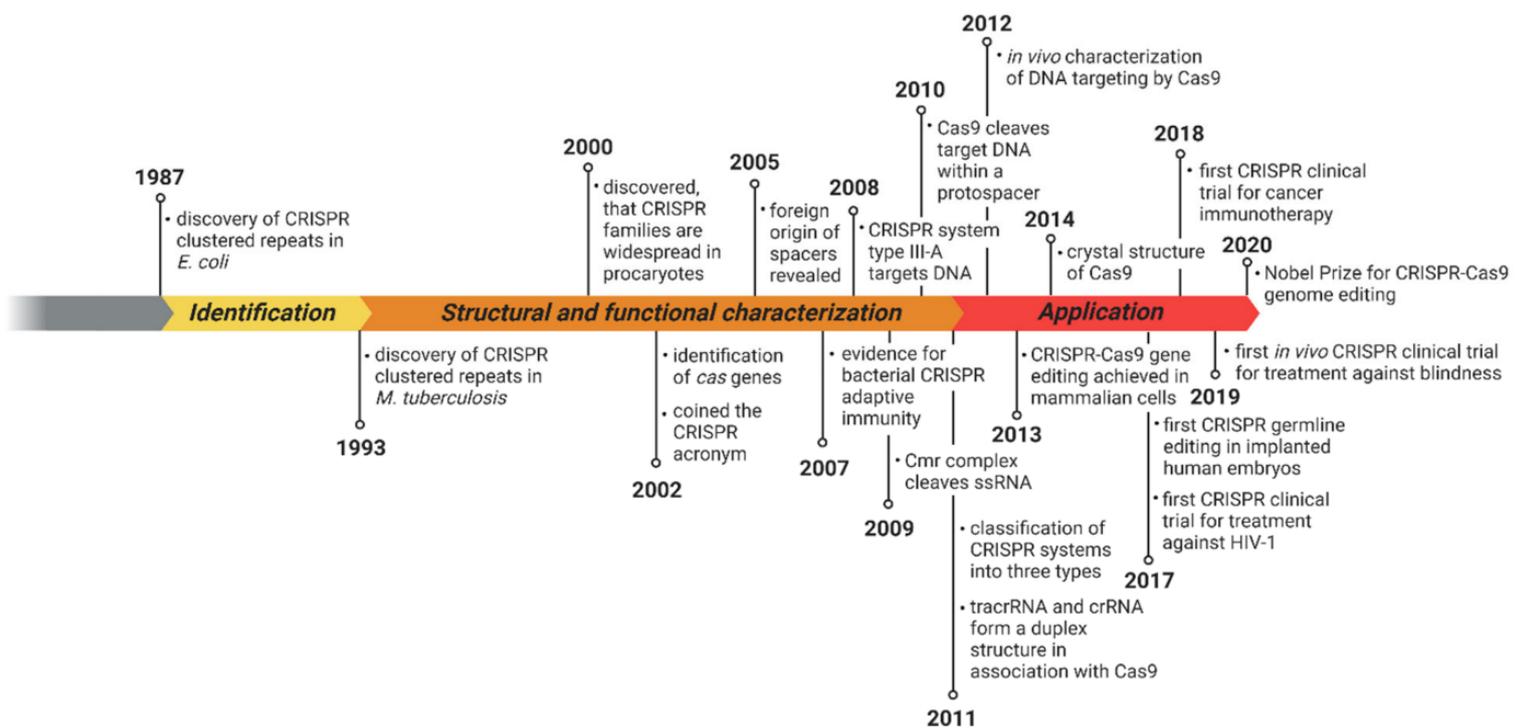


Figure 1. Chronologie de la découverte des systèmes CRISPR-Cas.

La découverte des systèmes CRISPR-Cas est la cumulation de plusieurs découvertes clés. La première étant la signature architecturale des librairies CRISPR (Espaceur-répétitions-espaceurs), l'omniprésence chez les archées et bactéries et finalement l'application de ces systèmes comme outil de génétique. Figure reprise de Vijay Tripathi *et al* 2021.

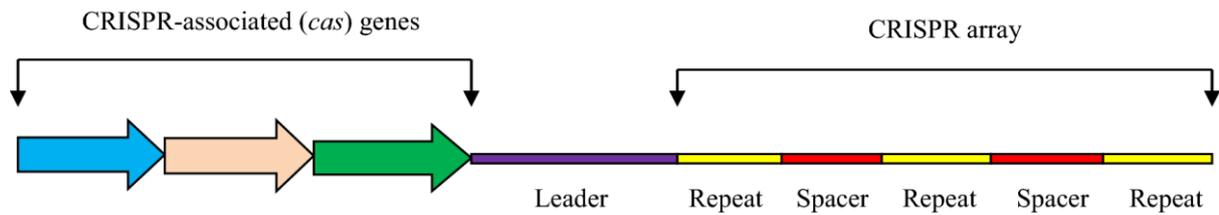
### 1.1.2 Architecture des systèmes CRISPR

Un locus CRISPR est un système comprenant trois acteurs majoritaires : une librairie, des ARN et des protéines Cas. L'organisation de ces trois acteurs est résumée dans la Figure 2.

1) La librairie code pour des séquences variables (*espaceurs*) interrompues par des séquences répétées (*repeats*). Les séquences variables contiennent un fragment d'une trentaine de paires de bases provenant des génomes étrangers (viral, plasmides) que la bactérie a pu rencontrer, vestiges des infections passées par des éléments génétiques mobiles (MGEs). Les librairies sont sous le contrôle d'une séquence directrice. Les séquences variables sont aussi associées à une signature, la séquence PAM (protospacer adjacent motif). Sans cette PAM la reconnaissance ne peut pas s'effectuer. Cette PAM est spécifique à chaque système CRISPR- Cas.

2) Les ARN sont capables de créer des R-Loop (triplex ADN : ARN). Ces R-Loop à l'aide des Cas forment un complexe (RNP) qui reconnaît une séquence cible et entraîne sa dégradation subséquente. Il est à noter que l'homologie entre l'espaceur et sa cible n'a pas besoin d'être parfaite. Ce phénomène assure une certaine plasticité du système immunitaire, mais amène au clivage de cibles secondaires (*off-targets*) lorsque le système CRISPR est utilisé comme outil génétique.

3) Les protéines Cas médient l'activité immunitaire des librairies CRISPR, à savoir la digestion du substrat cible via un ARN guide, ainsi que l'acquisition de nouveaux espaceurs, assurant donc une qualité adaptative au système immunitaire offert par les systèmes CRISPR.



**Figure 2 Architecture d'un locus CRISPR.**

Les locus CRISPR partagent tous une architecture commune. Les gènes Cas, qui selon le type sont chargés de la reconnaissance et/ou la dégradation de leur cible, ainsi que de l'acquisition de nouveaux espaceurs. La séquence directrice, qui régule l'expression des librairies et est le point d'ancrage des espaceurs entrants et, finalement, les librairies, qui sont une alternation d'espaceurs et de séquences répétées. Figure reprise de Supriyo Chakraborty *et al.* 2020.

Les Cas au sein d'un locus CRISPR sont codées par 3 groupes de gènes :  
 Les gènes d'expressions (Cas6, Rnase III), qui sont exprimés en cas d'attaque par un MGE.  
 Les gènes d'interférence (Cas3,5,7,9,12,13) qui lorsqu'associés à un ARN guide reconnaissent et clive leur substrat respectif (sbDNA, dbDNA, sbRNA, etc.) et finalement, les gènes d'adaptations (Cas 1,2,4) impliqués dans l'acquisition de nouveaux espaceurs. (27)

### 1.1.3 Classification des systèmes CRISPR

Due à la quasi-omniprésence des systèmes CRISPR-Cas chez les archées et les bactéries, il existe une grande variabilité de ces systèmes. Cette richesse rend difficile la stricte classification de chacun des systèmes CRISPR. La classification des loci CRISPR actuelle comprend deux grandes classes. Les deux classes sont distinguées l'une de l'autre selon les Cas médiant l'interférence (27). Les classes contiennent 3 types, qui sont différenciés eux, par leur mécanisme de maturation des ARNcr.

Les types I, III, IV entrent dans la classe 1, et les types II, V, VI dans la classe 2. Finalement, chaque type est regroupé en plusieurs sous types. Pour la classe 1, le type I comprend des sous

types allant de A à F et un dernier sous type nommé U. Le type III comprend aussi des sous types allant de A à D. (Figure 3A).

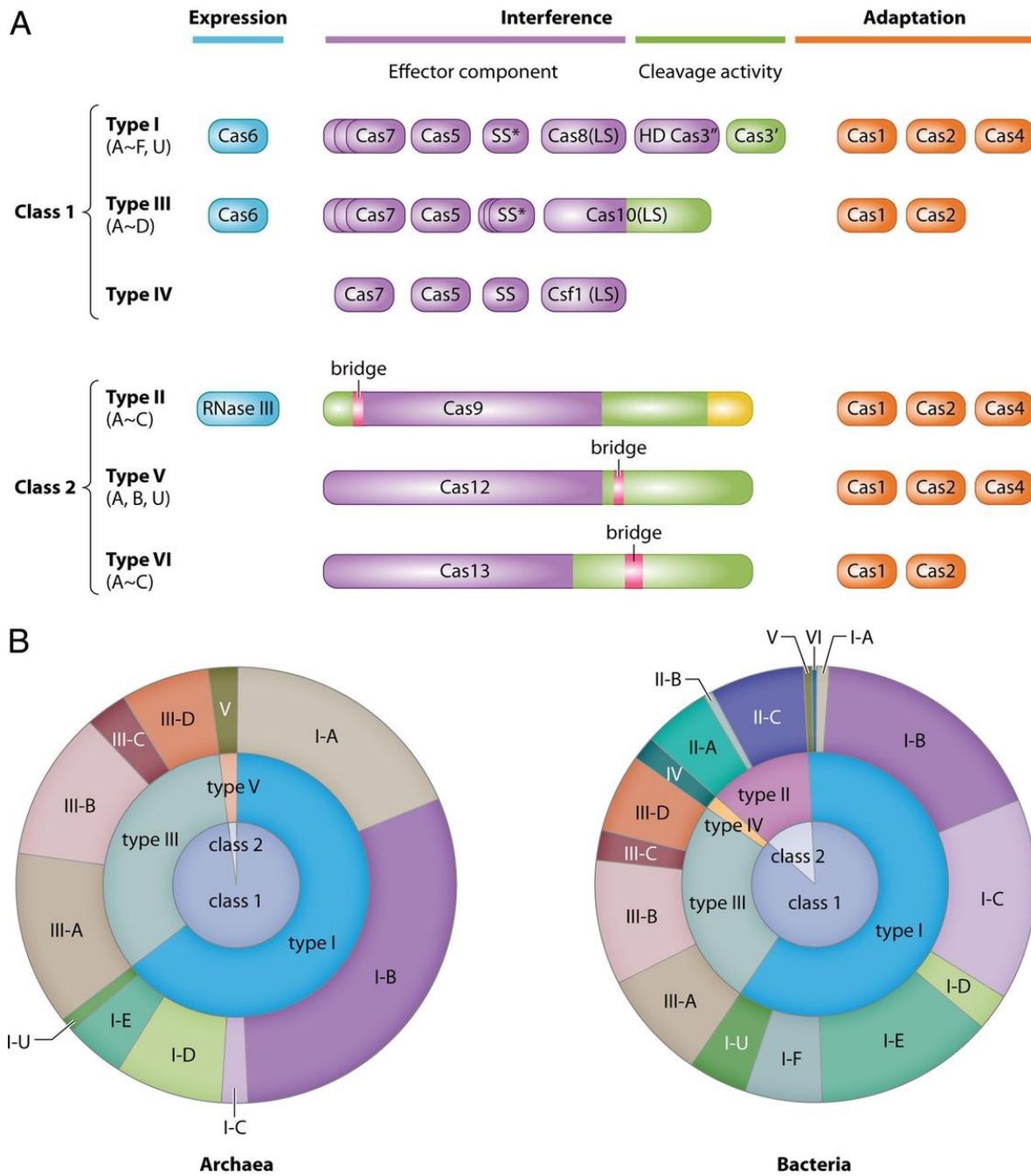
Pour la classe 2, le type II comprend 4 sous types allant de A à C, le type V 3 sous types, A, B, U et le type VI allant de A à C. Le type le plus étudié est le Type II. Le système CRISPR-Cas9 vedette des systèmes CRISPR appartient à la classe 2, type II. (Figure 3A).

La classe 1 représente la plus grande majorité des systèmes CRISPR (90%) (28), cette classe emploie plusieurs Cas (Cascade pour le type 1, Csm pour le type III-A, et Cmr pour le type III-B) comme effecteurs pour médier leur activité de reconnaissance et de clivage de leur substrat (Figure 3B). Le nombre de Cas au sein d'une librairie varie de 4 à 7 protéines par array. Pour le type I, l'assemblage du complexe de reconnaissance s'effectue en deux temps, premièrement un cortège de Cas s'associe à un ARN guide à la recherche d'homologie entre un espaceur et le génome envahissant, ce complexe multimérique s'appelle alors Cascade. Une fois que la séquence variable est reconnue, la Cas3 est recrutée pour finalement digérer le brin non apparié. Pour le type III, le complexe est constitué de Csm/Cmr et la lyse effectuée par Cas10 (29). Le type VI se distingue des 2 autres types par l'absence de Cas 1 et 2. Les librairies peuvent être compatibles entre elles. Un espaceur d'une librairie peut être pris en charge par une autre Cas, par exemple, les espaceurs codés par le type I peuvent être pris en charge par les Cas codés par le sous types III-B.(30, 31).

La classe 2 représente, elle, le reste des systèmes CRISPR établis (10%) (28). Bien que minoritaire la classe 2 est retrouvée principalement chez les bactéries et non les archées (Figure 3B). À l'opposé de la classe 1, la classe 2 utilise une protéine unique pour médier l'activité de reconnaissance et de clivage. Cette architecture plus simple a permis en autres à la Cas9 de

devenir l'outil de génie génétique phare. Il est important de noter qu'au sein d'un même sous type il existe une grande variété. Prenons en exemple le Type II de la classe 2, CRISPR-Cas9. Le type II est retrouvé majoritairement chez les bactéries à GRAM+ (32). Il existe plusieurs versions uniques connues de ce système. Les plus populaires étant respectivement, la spCas9 retrouvée chez *Streptococcus pyogenes*, la saCas9 chez *Staphylococcus aureus* et la nmCas9 chez *Neisseria meningitidis* (33). Bien que ces trois systèmes partagent la même architecture et le même type de Cas9, les deux Cas9 ne reconnaissent pas la même PAM.

Paradoxalement, les systèmes CRISPR sont retrouvés beaucoup plus rarement chez les bactéries environnementales (10%) (28). Les systèmes CRISPR ne sont pas uniquement limités aux chromosomes et peuvent être retrouvés en tant que MGE, au sein de phages, de plasmides, et de transposons. Il existe aussi des systèmes CRISPR atypiques comme il est le cas pour les Cas-solo. À savoir des systèmes contenant uniquement les Cas 1 médiant l'adaptation sans contenir de librairies. Les Cas 1/2 sont les deux Cas les plus conservées au sein de tous les systèmes CRISPR décrits à ce jour (30). Finalement il est important de noter que plus d'un locus CRISPR peut exister au sein d'un génome procaryote.



**Figure 3 Classification des systèmes CRISPR-Cas.**

- A. Les 2 classes et 6 types les plus communs des systèmes CRISPR. Séparés par les Cas effectrices signatures.
- B. Proportion des types retrouvés chez les archées et les bactéries. Le type I est le type le plus prépondérant de tous. Figure reprise de Patrick Forterre *et al.* 2018.

### 1.1.3 Origine des systèmes CRISPR-Cas : Casposons.

La plasticité et l'ubiquité des systèmes CRISPR sont deux facteurs illustrant le caractère évolutif de ces éléments génétiques mobiles. Comme mentionné précédemment les systèmes CRISPR-Cas ne sont pas limités dans les chromosomes, mais peuvent être retrouvés en tant que MGEs. Chez les archées les systèmes Cas-solo sont associés au « dark island matter ». Ces îlots génétiques codent jusqu'à 20% des génomes des archées. Ils encodent des locus dont la fonction est méconnue et sont des hauts lieux d'acquisition des MGEs. Krupovic *et al.* (34) ont noté en 2014 l'association exclusive de ces Cas solo avec des gènes essentiels comme *polB* (Polymérase ADN de la famille B) et entourés de signatures génétiques typique des transposons à savoir des : séquences répétées inverses (TIR) et des sites cibles de duplications (TSD). Ces sites sont le plus souvent associés avec des transposons à ADN(35). Au moment de leur découverte aucune transposase / intégrase connues pouvaient expliquer l'insertion de ces transposons. Cependant, un gène homologue à la Cas1 au sein de ces locus semblait pouvoir médier l'événement d'intégration. Ces locus furent dénommés casposons afin de différencier les Cas1 associées au locus *bone fide* CRISPR de ces systèmes Cas solo (2). La capacité de Cas1 à médier l'événement d'intégration fut démontrée chez *Aciduliprofundum boonei* (2, 36). Le mécanisme ressemble en tout point au mécanisme d'acquisition de nouveau espaceurs dans les systèmes CRISPR établis.

La taille des casposons varie entre 8 Kpb et 20 Kpb (37, 38) et ils ont été retrouvés chez les archées et les bactéries. Des études comparatives et phylogénétiques ont pu dégager 4 familles de casposons (34). Seule la famille 3 est retrouvée chez les bactéries. Les trois autres familles semblent appartenir aux archées. Les casposons ont la particularité d'autoréplication et leur qualité mobile en tant que MGE fut démontrée dans 62 souches chez *Methanosarcina mazei*,

où les casposons sont insérés dans plusieurs sites distinct, indicateur d'événements de mobilité génétique récente (39).

Le mécanisme évolutif ayant conduit à l'apparition des systèmes CRISPR-Cas et leur capacité évolutive a été étudié et plusieurs éléments convergent vers la fusion d'un casposon avec un système Cascade. Bien que le mécanisme exact soit encore méconnu, plusieurs événements génétiques sont au cœur de l'évolution d'un casposon en système CRISPR-Cas. La biogenèse de CRISPR semble représenter en premier un système immunitaire inné ancestral (comme les systèmes de restrictions, et le système argonaute par exemple (40). Krupovic *et al.* présentent le complexe Cascade comme à l'origine des CRISPR (41).

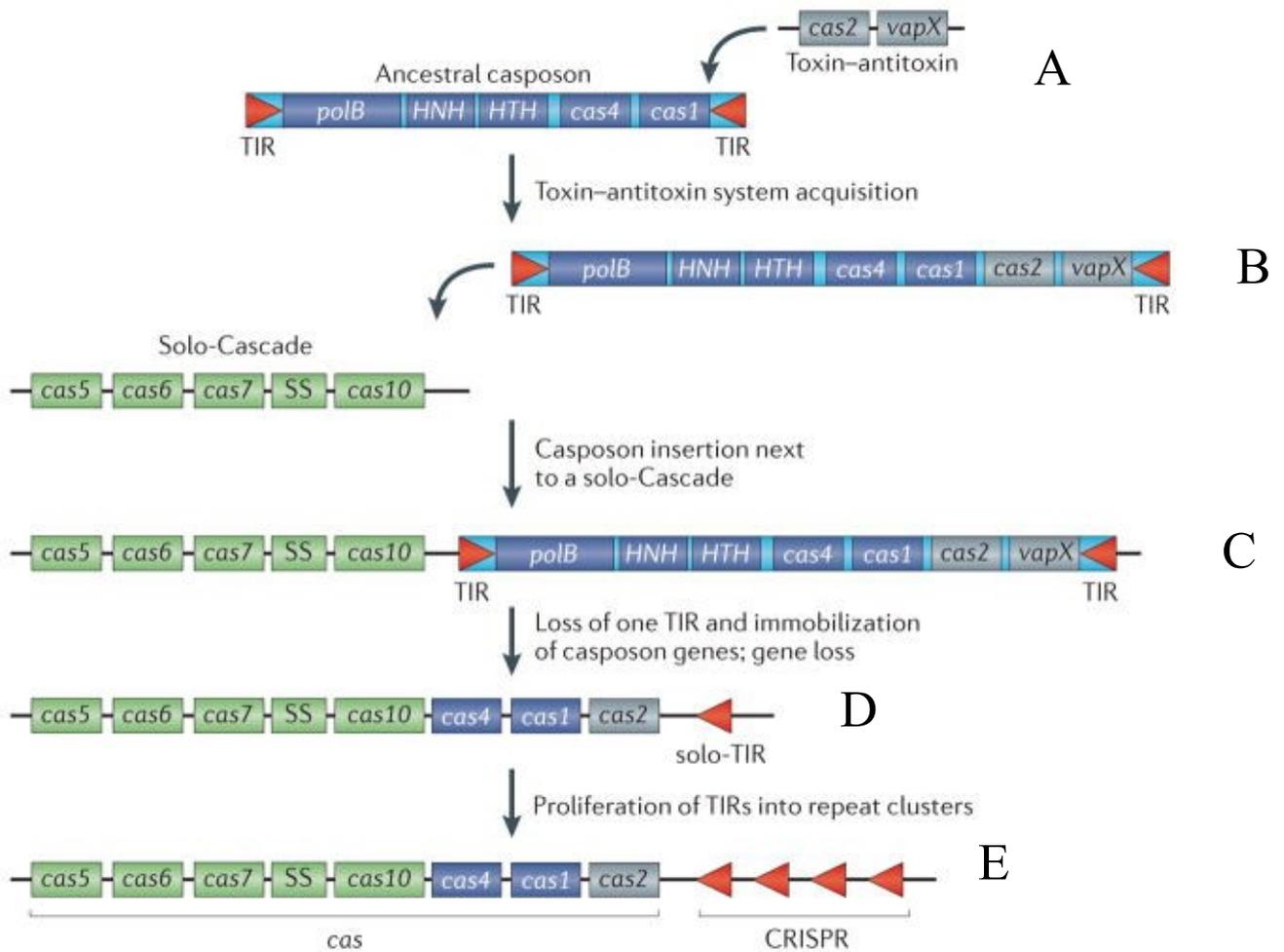
Ce système aurait évolué depuis des protéines effectrices contenant des motifs de reconnaissance des ARN (RRM) (42). Les cas 5,6,7,10 contiennent toutes un ou plusieurs motifs RRM. Les TIRs seraient responsables de l'émergence des répétitions entre chaque espaceur. En effet, l'évènement d'intégration du casposon dans son site cible (TSD) qui est médié par Cas1 ressemble en tout point au mécanisme d'adaptation lorsque Cas1/2 intègrent un nouvel espaceur dans la librairie CRISPR. Les deux événements d'intégrations impliquent la régénération du site cible (respectivement TSD pour le casposon et des séquences répétées pour une librairie CRISPR) et de la séquence directrice. De plus cet évènement d'intégration est polarisé, l'espaceur entrant est toujours intégré derrière la séquence directrice (en 5') et devient le premier espaceur de la librairie CRISPR (43, 44).

La chronologie de l'évolution des casposons en systèmes CRISPR se placerait telle que :

**1) Acquisition du système toxine-antitoxine Cas2/vapX (23) au sein d'un casposon (Figure 4A).**

Le système Cas était un système toxine antitoxine, médiant une activité de répression post traductionnelle par ARN antisens. Cas2 est un analogue de VapD, toxine ciblant les ARNm ou ARNm associés aux ribosomes (45, 46).

- 2) Intégration du casposon ancestral proche d'un locus Cascade solo (Figure 4B et Figure 4C).
- 3) Perte d'une partie des TIRs. Sans ces inversions répétées, le casposon perd sa capacité de mobilité (Figure 4D).
- 4) Prolifération des TIRs devenue maintenant des répétitions CRISP (figure 4E).



**Figure 4. Chronologie de l'évolution des systèmes CRISPR-Cas.**

Modèle proposé pour l'évolution des casposons en systèmes CRISPR *bona fide*. A. Acquisition d'un système toxine antitoxine au sein d'un locus casposon ancestral. B. Fusion d'un système Cas-Solo avec le casposon ancestral. C. Immobilisation du casposon par pertes des répétitions inverses (TIR). E. Évolution finale d'un locus CRISPR *bona fide*. Figure reprise de Marc Krupovic *et al* 2015.

Cette hypothèse permettrait d'expliquer la plasticité des systèmes CRISPR. Particulièrement certains sous-types du type III ne codent pas pour les Cas1/2, mais peuvent être complétés par d'autres locus CRISPR présents au sein du même génome de la même manière que des espaceurs du types I peuvent être pris en charge par le type III (30, 31).

Finalement, l'apparition de la classe 2 serait survenue par le remplacement des Cas formant le complexe Cascade pour une seule protéine effectrice, la Cas9, via ses domaines nucléases, RuvC et HNH (47-49).

## **1.2 CRISPR comme système immunitaire procaryote.**

La capacité de système immunitaire de CRISPR fut confirmée en 2007 par Barrangou *et al.* (24). Ils ont pu démontrer qu'une population de *S. thermophilus* sensible à un phage pouvait développer une résistance à ce dernier si elle survivait à une première infection. Un génotypage de la population survivante pointait vers l'acquisition de nouveaux espaceurs dans la librairie CRISPR codant pour le même phage. Dans la même étude, la capacité reprogrammable des librairies CRISPR a été testée et les résultats ont démontré que la sensibilité de la souche pouvait être modulée en ajoutant ou supprimant des espaceurs dans sa librairie. Le rôle de la Cas7 (renommé *csn2*) dans la capacité immunitaire adaptative, mais non acquise fut aussi identifié (24).

### **1.2.1 Adaptation immunitaire**

#### **1.2.1.1 Capture d'un espaceur**

L'intégration accidentelle d'espaceurs provenant du génome de l'hôte emmène au phénomène d'auto-immunité et entraîne la mort de l'hôte (50-52). Les systèmes CRISPR-Cas contiennent donc des garde-fous pour prévenir ce phénomène. On retrouve comme mécanismes de

protection l'association de l'espaceur avec une signature PAM (53, 54) qui permet au complexe de reconnaissance de discriminer le soi du non-soi. La source du substrat pour la génération de nouveaux espaceurs joue aussi un rôle dans l'intégrité immunitaire. L'acquisition d'espaceurs provient depuis la digestion d'ADN linéaire, qui n'est pas naturellement présent dans la cellule bactérienne. La digestion d'ADN linéaire est assurée par le complexe RecBCD et les produits digérés sont pris en charge par Cas1/2. La digestion d'un ADN linéaire est continue jusqu'à ce que RecBCD rencontre un site *chi* (55). Ce site de 8pb entraîne un changement de conformation du complexe RecBCD pour amorcer la réparation via RecA. Il indique au complexe RecA où amorcer la recombinaison pour tenter de réparer une cassure double brin, sans causer une trop grande réorganisation chromosomique. Finalement, la limitation de la vitesse d'acquisition d'espaceurs est aussi un mécanisme important. Heler *et al.* ont démontré qu'une mutation, I473F, au sein de Cas9 augmentait considérablement le taux d'acquisition d'espaceurs exogènes et endogènes, augmentant la cytotoxicité du locus CRISPR via auto-immunité (56).

Il est à noter que de nouveaux espaceurs sont générés et intégrés via Cas1/2 dans la librairie CRISPR lors d'une réponse acquise suite à une primo-infection par un MGE/phage.

Finalement si un espaceur est reconnu, mais que sa PAM est mutée (cette PAM est alors dénotée ePAM pour *escaper PAM* (57), le complexe de reconnaissance CRISPR au lieu de cliver son substrat recrute Cas1/2 à la recherche de nouveaux espaceurs (*Cf 1.2.3 Interference*).

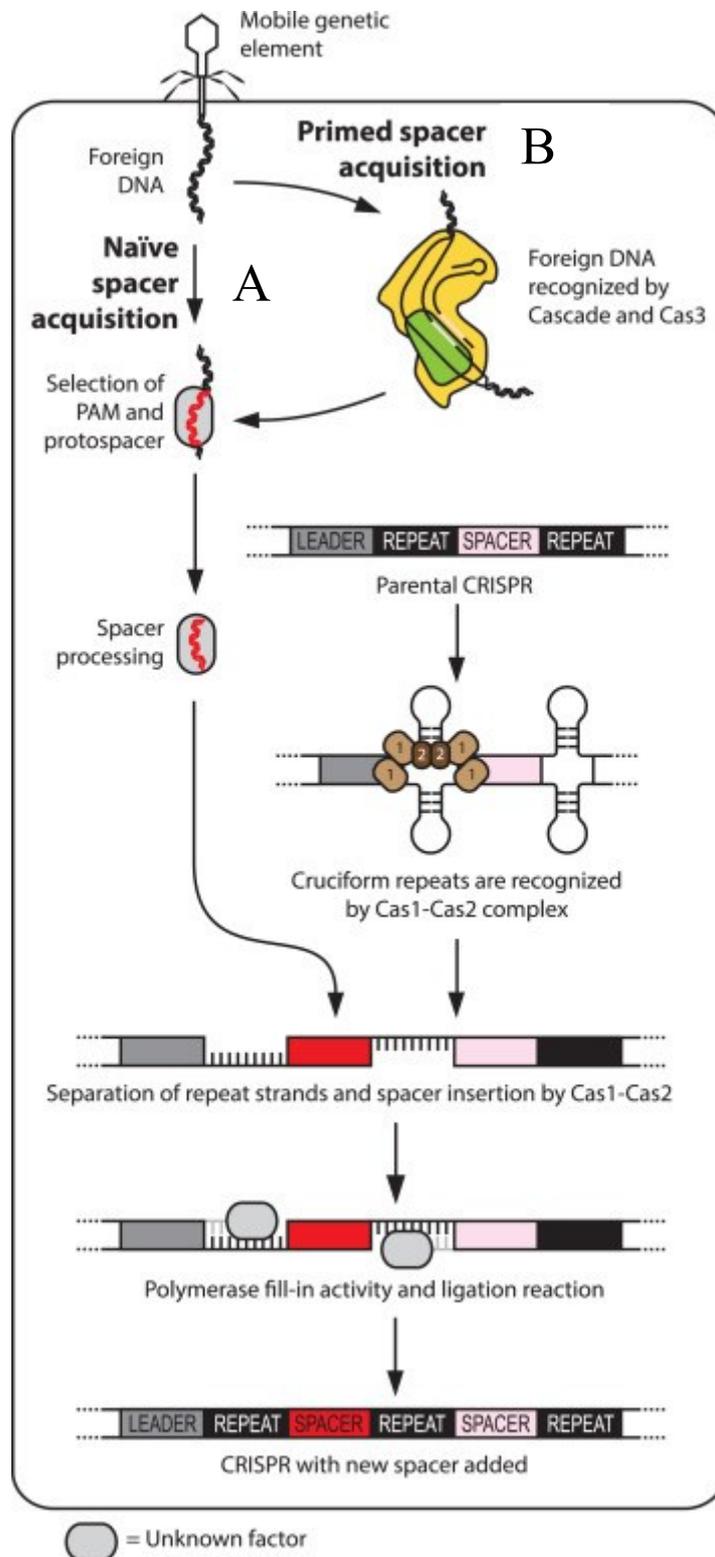
### **1.2.1.2 Intégration d'un espaceur**

Lors d'une infection par un phage, les systèmes CRISPR peuvent engager deux types de réponses immunitaires. On distingue deux types de réponses : naïve et acquise (58). Naïve lorsque la bactérie n'a jamais rencontré ce phage (Figure 5A). Dans cette situation l'acquisition de l'espaceur se déroule principalement par la digestion de l'ADN exogène via le complexe

RecBCD pour les GRAM- et AddAB pour les GRAM+. La réponse acquise elle (Figure 5B), concerne la dégradation de l'ADN exogène par le complexe de reconnaissance CRISPR lorsque la librairie contient un espaceur ayant une homologie pour le phage attaquant.

Les systèmes CRISPR-Cas protègent peu l'hôte lors d'une réponse naïve. L'événement d'acquisition d'un espaceur provenant du génome envahissant est un événement rare (de l'ordre de 1 cellule pour  $10^7$ ) (59-61). Ce mécanisme d'acquisition est retrouvé dans les deux classes et est principalement orchestré par le duo Cas1/2. Un mécanisme de défense auto-immun par l'incorporation accidentelle d'un espaceur provenant de l'hôte serait à l'origine de la rareté de l'évènement. Les Cas1/2 sont des nucléases, leur rôle exclusif dans l'évènement d'intégration fut démontré par Brouns *et al.* (62). Ces deux Cas forment un complexe en forme de papillon qui intègre l'espaceur de manière polarisée (63). L'espaceur est intégré au plus proche de la séquence directrice (*leader*), puis les séquences répétitions sont régénérées. Comme mentionné précédemment le mécanisme d'intégration est commun aux systèmes CRISPR-Cas et aux casposons.

Il reste encore à déterminer si l'espaceur est prélevé depuis le génome digéré ou si la séquence est copiée puis intégrée dans la librairie (64). Cette polarisation sous-entend que l'espaceur le plus proche de la séquence directrice est témoin de l'agression génétique la plus récente. Cette polarisation emmène certains auteurs à considérer les librairies CRISPR-Cas comme des fossiles moléculaires. Il est à noter que plus 90% des espaceurs séquencés à ce jour ne correspondent aucun MGE connu (27). Nombre de ces espaceurs orphelins sont retrouvés dans des librairies CRISPR de certains Type III où Cas1 est associé avec une transcriptase inverse et incorpore des espaceurs depuis des substrats ARN (65). Cette masse d'espaceurs orphelins est dénommée matière noire des espaceurs CRISPR. (dark matter of CRISPR spacers) (27).



**Figure 5. Schéma général d'acquisition d'un espaceur CRISPR-Cas**

Les systèmes CRISPR peuvent enclencher deux réponses. Les deux réponses amènent toutes les deux à la capture d'un nouvel espaceur. L'intégration d'un espaceur est polarisée et se place toujours derrière la séquence directrice. A. Un fragment d'ADN viral est sélectionné aléatoirement, et est inséré dans le locus CRISPR. La réponse naïve offre peu de protection contre le MGE entrant. B. L'ADN entrant est reconnu par le complexe Cascade, en sus de déclencher une dégradation de sa cible, d'autres fragments d'ADN du MGE entrant est ajouté au locus CRISPR. Figure reprise de Magnus Lundgren *et al* 2015.

### 1.2.2 Expression et maturation des ARNcr

Un point crucial de l'immunité des systèmes CRISPR est représenté par les crARN (ARN CRISPR). Le schéma global est le même pour les trois types, les différences dans la prise en charge du crARN est une signature unique à chaque type. Différents types coexistant au sein d'un même génome sont incapables de traiter les crARN d'un type voisin (66).

La réponse immunitaire engagée par les systèmes CRISPR commence par l'expression d'une longue molécule d'ARN précurseur. Les systèmes CRISPR peuvent être activés en réponse à plusieurs facteurs. Le facteur principal est la détection d'une infection virale par un MGE via des facteurs de dommages à l'ADN. Des facteurs secondaires environnementaux des stress (détection du quorum, concentration en glucose et/ou fer, etc.) peuvent aussi activer ces systèmes) (67).

J. McGinn *et al.* (59) ont étudié l'expression de crARN chez *Streptococcus pyogenes* (spCas9, Type II-A). Plus précisément ils ont voulu comparer l'expression entre le premier et le cinquième espaceur au sein d'une librairie CRISPR. Le premier espaceur était, de par sa position relative à la séquence directrice, exprimée deux fois plus que le cinquième espaceur. L'organisation polaire des espaceurs constituerait donc une stratégie de minimisation des risques (*bet-hedging strategy*), ou la priorité serait donnée à une réponse contre le MGE infectieux le plus récent (68). Pour compléter ce mécanisme, plusieurs auteurs ont noté que la délétion d'espaceurs était un phénomène courant. Toujours chez *Streptococcus pyogenes*, la délétion de plusieurs espaceurs proximaux de la séquence directrice peut être sélectionnée pendant une infection par un phage autre que le phage le plus récemment rencontré (59). D'autres mécanismes d'expressions des espaceurs plus ancestraux sont aussi rapportés (69). La faible expression des espaceurs ancestraux permettrait au minimum de déclencher un événement d'acquisition d'espaceurs.

Pour les trois types, la transcription du long crARN précurseur est initiée par la séquence directrice (*leader*) puis est clivée séquentiellement pour former des sous-unités contenant un espaceur et des structures secondaires interagissant avec la ou les Cas. La Cas6 intervient pour les types I (Cas6e) (62) et III (Cas6 III) (70), avec la différence d'une étape supplémentaire pour les types III (71). Le rôle de la Cas6 a été mis en lumière par Carte *et al.* chez *Pyrococcus furiosus* (72) une fois la maturation du crARN terminée ce dernier formera une RNP avec son complexe de reconnaissance respectif (Cascade, Cas9, Csm/Cmr).

Le type II emploie la Rnase III de l'hôte pour le traitement du crARN précurseur. Le type II se distingue aussi par l'expression d'un deuxième type d'ARN : le tracrARN (73). Cet ARN s'associe avec une sous-unité du crARN (après digestion pour la Rnase III) pour former l'ARN guide (*sgARN*). La Cas9 semble jouer un rôle dans la maturation des crARN, bien que le rôle de cette dernière ne soit toujours pas élucidé (74).

### **1.2.3 Interférence**

La base de tout système immunitaire repose sur la discrimination du soi et du non-soi. On retrouve deux voies pour l'interférence : PAM dépendante et PAM indépendante.

#### **1.2.3.1 Interférence PAM dépendante**

La voie PAM dépendante est la voie principale. Le mécanisme est similaire pour le type I et type II. Une fois le complexe de reconnaissance assemblé (Cascade ou Cas9) et appairé à son crARN (complexe mature), il part à la recherche de sa cible. Le complexe mature balaie le génome exogène d'abord pour une séquence PAM puis par diffusion latérale cherche une homologie avec son crARN (75). Le complexe mettrait plusieurs heures à trouver sa cible, cependant, il est estimé que la Cas9 par exemple ne reste que quelques microsecondes sur une

cible n'ayant pas d'homologie suffisante (75, 76). L'ensemble des stratégies d'interférence est résumé en Figure 6.

L'assemblage des Cas en complexe de reconnaissance est indépendant de la présence du crARN. Bien que ce dernier soit essentiel à la reconnaissance de sa cible (77). Une fois les Cas assemblées, le crARN est embarqué dans le complexe de reconnaissance. Si une homologie est trouvée, et que l'ADN cible est détecté comme du non-soi, un changement de conformation du complexe s'opère et la Cas chargée de dégrader sa cible est recrutée (Cas3 pour le type I, Cas9 pour le type II, Cas10 pour le type III). Pour le type I (figure 6A), le brin opposé au brin reconnu est digéré par la Cas3 et son domaine nucléase HD, créant une délétion et laissant le brin reconnu simple brin, plusieurs lésions seraient ainsi répétées sur le MGE envahissant (78). Dans le type II, la Cas9 va créer une lésion double brin en coupant les deux brins 3pb en amont de la séquence PAM (79). Pour le type III, Cas10 possède une activité similaire à la Cas3 (80). La PAM est indispensable pour la dégradation de son substrat, c'est pourquoi, pour éviter les phénomènes d'auto-immunité, les PAM sont absentes au sein des espaceurs dans les bibliothèques CRISPR. L'homologie entre le crARN et sa cible n'a pas besoin d'être parfaite. Elle requiert cependant une forte homologie dans les 8 à 12bp précédant la PAM. Cette région est dénommée région noyau (*seed*). Cette tolérance aux erreurs d'appariement (*mismatch*) permettrait à un seul crARN de reconnaître plus d'une séquence (81, 82).

Le type II emploie une seule Cas (Cas9) pour le complexe de reconnaissance et le clivage de sa cible grâce à ses deux sites nucléase RuvC et HCH (Figure 6B). Cependant un second type d'ARN : le tracrARN est indispensable pour la reconnaissance et le clivage (stabilisation du crARN et Cas9). Ce second ARN a été décrit en premier par Deltcheva *et al* chez *S. pyogenes* (73). Ce tracrARN s'assemble d'abord avec le crRNA pour former le gARN. L'une des premières étapes de l'implémentation des systèmes CRISPR-Cas en tant qu'outil de biologie

moléculaire a été la création d'un ARN chimère, contenant à la fois le crARN et le tracrARN sans impacter l'efficacité enzymatique de la Cas9. Cette chimère s'appelle le sgARN pour le différencier du gARN *bone fide* (74).

Bien que plusieurs types qui coexistent au sein d'un même génome hôte ne peuvent pas prendre en charge un crARN d'un autre type certains sous types le peuvent (83). Le sous-types III-A (Csm) chez *S. epidermidis* et III-B (Cmr) chez *P. furiosus* bien que homologues, reconnaissent tous les deux des substrats différents, ADN et ARN respectivement (84-86). Paradoxalement, Csm a été rapporté être capable de cliver des substrats ARN chez *Thermus thermophilus* et *S. thermophilus* (83, 87). Chez *T. thermophilus* les deux sous types III A et B sont capables de prendre en charge le crARN de l'autre. Peng *et al.* ont pu montrer chez *S. islandicus* que le complexe Cmr pouvait reconnaître un substrat ADN et ARN (88, 89).

À l'origine de ce paradoxe, il y aurait une stratégie immunitaire basée sur la sauvegarde des phages tempérés (lysogénique) et la dégradation des phages virulents. En effet il est reconnu que les phages tempérés peuvent représenter un avantage évolutif en apportant dans leur génome des gènes qui peuvent être bénéfiques pour l'hôte (90, 91). Marraffini *et al* démontrent que le système CRISPR-Cas permet une tolérance conditionnelle aux phages tempérés via une répression de l'activité de dégradation de la cible ADN. Si la même cible ADN n'est pas transcrite en ARN alors il n'y pas d'interférence (transcription dépendante) (92). La tolérance est abolie lorsque le phage induit son cycle lytique. Le type III emploie aussi des ARNase et ADNase, la toxicité et l'efficacité de ce système sont plus élevés que pour les deux autres types (80).

Finalement, deux sous types III A et B ne semblent pas utiliser une PAM pour discerner le soi du non-soi. La signature est assurée par le crARN directement. La maturation d'un crRNA

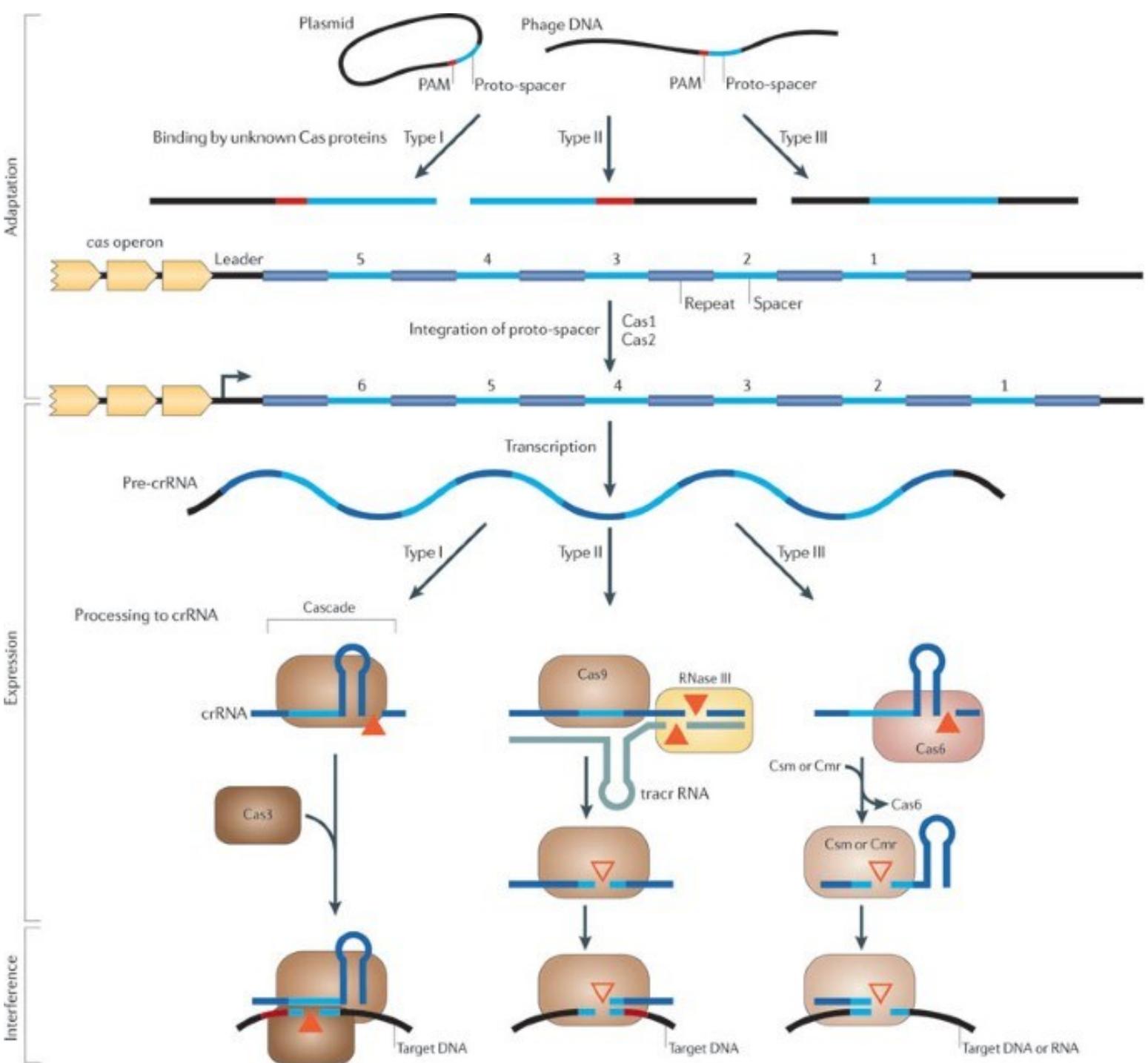


Figure 6. Interférence PAM dépendante chez les types I, II, III.

Tous les types suivent le même modèle d'interférence : Adaptation, Expression et Interférence. Le type III se distingue des autres types par l'absence d'une PAM pour la différenciation du soi et du non-soi. Figure reprise de Eugene V. Koonin *et al* 2011.

laisse une « queue » en 3' en plus de l'espaceur. Si cette queue possède une homologie avec le génome de l'hôte en plus de son espaceur la dégradation est avortée (93).

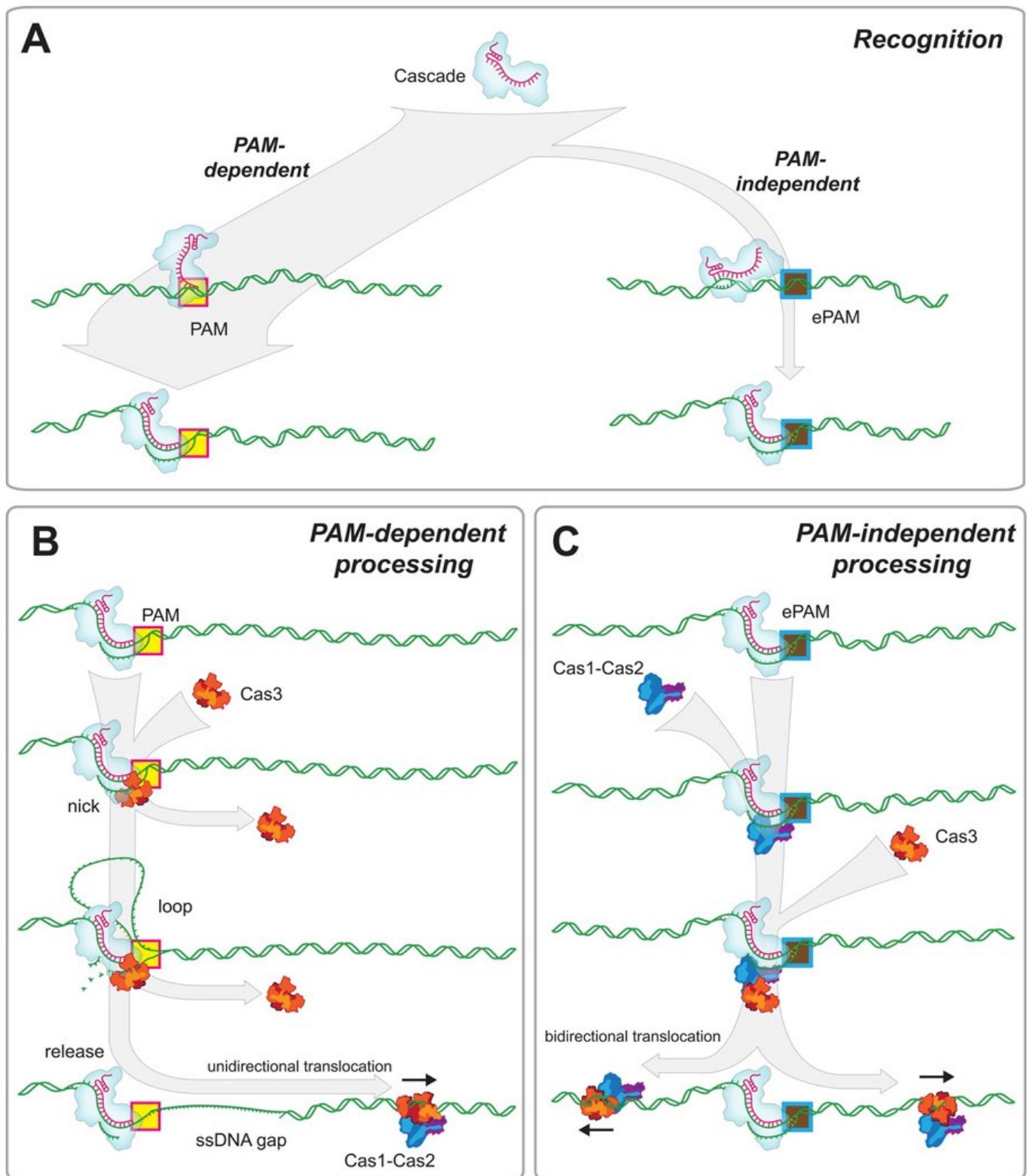
### 1.2.3.2 Voie PAM indépendante

La voie PAM indépendante est la moins empruntée. Même si le type III ne cherche pas de signature PAM, la différenciation repose toujours sur la reconnaissance d'un motif propre à l'hôte. Greene *et al* décrivent une voie PAM indépendante pour le type I qui est empruntée lorsque des mutations dans la cible empêchent la reconnaissance. Cette voie serait donc une voie de dernier recours aux systèmes CRISPR-Cas dans un but de survie (but analogue à la balance-bénéfice/risques produit par la réponse S.O.S via RecA/LexA (94). La PAM mutante est alors dénommée ePAM.

Dans cet article les auteurs décrivent que cette voie est empruntée lorsque la ePAM est suffisamment mutée pour empêcher une réponse immunitaire, mais avec suffisamment d'homologie pour alerter le système immunitaire qu'un MGEs tente de compromettre son génome. Le succès de cette voie est proportionnel aux nombres / types de mutations. À l'image du mécanisme de réparation MMR chez *E. coli*, certaines mutations dans la ePAM sont plus tolérées que d'autres (95). L'augmentation de chances de rencontre entre le complexe Cascade et sa cible mutée requiert une grande concentration intracellulaire de cascade, ce qui traduit que la voie PAM indépendante s'emprunte lors d'une réponse immunitaire après que la voie PAM dépendante a échoué. Les auteurs remarquent que le complexe Cascade cherche sa cible toujours par diffusion latérale.

Chez *E. coli*, sans qu'il enclenche la dégradation de sa cible, le temps que passerait le complexe Cascade sur sa cible mutée serait suffisant pour que Cas1/2 soient recrutées. Cas 1/2 répriment alors l'activité nucléase de Cas3. Dans les deux voies (PAM dépendant et indépendant) la demi-vie des complexes Cascade est significativement plus longue que le temps de doublement (95). Après avoir été recrutée par les Cas1/2, la Cas3 diffuse latéralement sur les deux brins à la recherche d'un nouvel espaceur en prime d'une PAM.

Bien que les auteurs aient pu souligner que Cas1/2 étaient recrutées en premier et qu'elles provoquaient un changement de comportement de la Cas3, le mécanisme reste inconnu. Le mécanisme est résumé en Figure 7.



**Figure 7. Schéma des deux voies pour l'interférence pour le type I**

**A.** Cascade, une fois chargé avec son crARN mature, scanne le chromosome pour une homologie avec sa cible. Si la signature PAM concorde, l'interférence est enclenchée. Si la PAM est mutée, la voie PAM indépendante est enclenchée. **B.** Détail du mécanisme de la voie PAM dépendante, une fois la cible trouvée le complexe Cascade change de conformation et recrute la Cas3 qui va créer une délétion et la digestion du brin non apparié. **C.** Détail du mécanisme de la voie PAM indépendante : l'interférence est avortée si la cible et/ou sa PAM sont mutées. Si l'homologie reste suffisamment fidèle pour que Cascade reconnaisse sa cible, Cas1/2 sont recrutées. Le duo Cas 1/2 cherche de manière bidirectionnelle de nouvelles cibles pour dégrader l'ADN entrant. Figure reprise de Eric C. Greene *et al.*, 2015.

#### 1.2.4 Systèmes Anti-CRISPR

La compétition entre les phages et leur hôte est une lutte permanente. Pour une cellule bactérienne, il existerait 10 virus (96). Les systèmes CRISPR bien que d'une aide précieuse dans cette lutte ne sont pas infaillibles. Comme l'illustre le phénomène de ePAM, il est relativement facile pour un MGE d'échapper à CRISPR via une simple mutation dans la PAM. Les systèmes CRISPR-Cas peuvent aussi rapidement amener à un état d'auto-immunité et compromettre la survie de l'hôte si leur régulation n'est pas contrôlée.

L'hôte possède plusieurs voies de répression, ces répressions dépendent de facteurs environnants, et emploient plusieurs répresseurs qui sont au cœur d'autres voies de signalisation clés. Du côté offensif, les MGEs peuvent coder pour plusieurs facteurs. Bondy *et al.* ont décrit les gènes Acr (anti-CRISPR) chez plusieurs prophages dans le chromosome de *P. aeruginosa*. Ces gènes peuvent perturber la maturation des complexes de reconnaissance et des crARN (67, 97).

Ironiquement, certains phages peuvent utiliser CRISPR à leur avantage comme c'est le cas des phages ICP1 chez *V. cholerae*. Cette famille de phages contient son propre système CRISPR (Type 1-F) qui vise un locus anti-phage. Ces phages peuvent même exploiter la capacité adaptative de CRISPR et prélever des espaceurs provenant de PLE, si les espaceurs originaux venaient à muter (98). En 2020, il a été démontré que les Biggiephages contiennent aussi des systèmes CRISPR-Cas. Ces systèmes sont dits minimaux, ne contenant qu'une seule Cas (Cas Phi) et d'une librairie CRISPR. Ces systèmes permettraient à ces phages d'éliminer d'autres phages en cas de surinfection (infection simultanée par plusieurs phages), et d'éliminer d'autres MGEs, comme des plasmides déjà présents dans la cellule (99).

### **1.2.5 Limites de l'argument immunitaire**

La fonction de CRISPR-Cas a été prédite en 2002 (14) avant d'être plus tard confirmée expérimentalement en 2007 (24). Certains auteurs accusent un biais de confirmation (100). Bien qu'ils ne remettent pas en cause la littérature prouvant la capacité immunitaire de CRISPR, les auteurs pointent les paradoxes et les questions non résolues qui entourent les systèmes CRISPR-Cas tel que le phénomène d'auto-immunité, la faible prévalence des systèmes CRISPR dans les procaryotes environnementaux.

L'un de leur argument principal réside sur la corrélation négative qui devrait exister entre la présence d'un ou plusieurs systèmes CRISPR-Cas et les traces de mobilité récentes par transfert horizontal dans ces mêmes génomes contenant les systèmes CRISPR. Cette corrélation négative est retrouvée et décrite amplement pour les systèmes de restriction (systèmes immunitaires innés bactériens), mais est en revanche discutable pour les systèmes CRISPR.

Les auteurs concluent que plus de validation expérimentale est nécessaire pour saisir l'image de CRISPR-Cas dans sa globalité, que le modèle actuel d'interaction MGE/procaryote de CRISPR-Cas n'explique pas entièrement l'apparition, la transmission et la maintenance de ces systèmes. Finalement, la masse d'espaceurs orphelins suggérait que nous ne connaissons qu'une portion infime des viromes par rapport à la masse totale (100).

### **1.2.6 CRISPR comme outil génétique**

La qualité programmable des systèmes CRISPR-Cas fait d'eux un outil de génétique précieux. En effet, la capacité à pouvoir cibler n'importe quelle séquence (pourvu qu'elle soit accompagnée d'une PAM) mène à la possibilité de modifier le génome de manière ciblée. Chez les eucaryotes, on tire parti du mécanisme de réparation de l'ADN après une cassure double brin. Chez les procaryotes à l'inverse on exploite l'absence de mécanisme de réparation efficace et fidèle pour effectuer une contre sélection (101, 102). La capacité plastique des systèmes CRISPR fut d'abord soulignée en 2007 par Barrangou (24). La conversion réelle des systèmes CRISPR-Cas en outil de génétique fut démontrée indépendamment en 2012 (25) et 2013(103). Plusieurs groupes ont prouvé qu'il était possible de manipuler les espaceurs du type II CRISPR (Cas9) pour induire une cassure double brin. Les types II sont les systèmes les plus compacts des systèmes CRISPR. L'emploi d'une seule Cas pour le complexe de reconnaissance et d'interférence a permis d'élucider et de reprogrammer leur fonction rapidement.

Comme mentionné précédemment, le point tournant fut l'identification et la création d'un sgARN chimérique à partir d'une fusion du crARN et tracrARN en sgARN. La découverte des deux ARN a été réalisée indépendamment avec 1 mois d'écart. Charpentier *et al.* en août 2012 ont pu démontrer *in vitro* que le sgARN chimère permettait de condenser les deux ARN (sans impacter l'efficacité de la Cas9) et de reprogrammer les espaceurs pour digérer des plasmides cibles. Virginijus Šikšnys *et al.* en septembre 2012 (104) ont pu démontrer *in vitro* que le crARN et tracrARN pouvaient être reprogrammés pour dégrader un oligo d'ADN simple brin dénommé SP1.

L'exportation de l'outil CRISPR-Cas chez des cellules humaines fut démontrée par George Church en 2013 (103).

### 1.2.7 Mécanisme d'action de la Cas9

On compte 6 domaines clés dans la Cas9. Rec I, Rec II, le Pont hélice (*Bridge Helix*), le site de reconnaissance de la signature PAM (*PAM Interacting*), HNH et RuvC (105-107).

Le domaine Rec I interagit avec le gARN. Le domaine *PAM Interacting* permet d'assurer la présence de la signature PAM et donc éviter l'auto-immunité (105). L'embarquement du gARN dans la Cas9 entraîne un changement de conformation de la Cas9. La fonction de Rec II est encore mal comprise. Les séquences PAM reconnues par les Cas9 du type II sont riches en pyrimidines, la Cas9 de *S. pyogenes* (SpCas9) reconnaît principalement le motif NGG. Bien que beaucoup plus rarement, le motif NRG peut être aussi reconnu (ou R indique G ou A) (108).

Une fois activée la Cas9 part à la recherche par diffusion latérale d'une séquence homologue au gARN. Une fois trouvés les domaines HNH et RuvC (deux domaines très conservés dans la famille des nucléases) réalisent le clivage exactement 3 paires de bases en amont de la séquence PAM sur les deux brins, amenant à une cassure double brin.

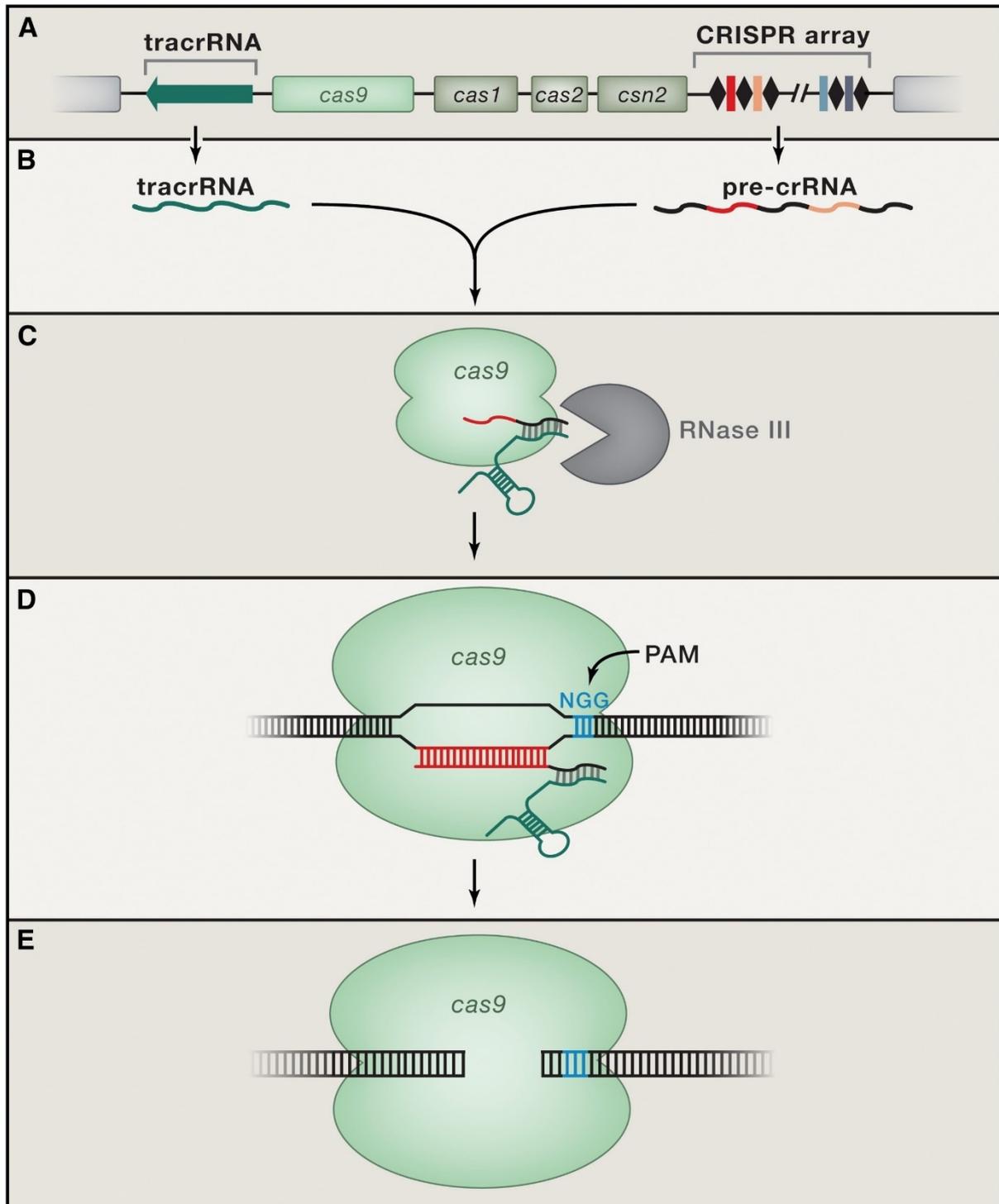


Figure 8. Mécanisme d'action de la Cas9.

La Cas9 emploie un mécanisme similaire au type I. A. Schéma d'un locus CRISPR du type II. B. La transcription de la librairie est amorcée à la suite d'une infection par un phage et des facteurs environnant. C. Le crARN précurseur est pris en charge par l'ARNase III de l'hôte pour former des sous-unités crARN ne contenant qu'un espaceur, ce crARN mature est lié à un tracrARN. D. La Cas9 trouve une homologie entre son crARN et sa cible en sus d'une signature PAM. E. La cible est clivé par les domaines RuvC et HNH.

Figure reprise de Eric S. Lander, 2016.

### 1.3 Mutagenèse chez les procaryotes

La biologie moléculaire est une discipline qui existe depuis moins de 100 ans. Elle trouve ses racines dans les premiers travaux qui tentaient de comprendre quels mécanismes régissaient les lois de ségrégation de Mendel. Morgan Thomas *at al* ont décrit en 1915 et 1926, dans les ouvrages, « *The mechanism of Mendelian Heredity* » et « *The theory of genes* » le lien qui existait entre l'hérédité, les chromosomes et les gènes chez *Drosophila melanogaster* (109-111).

L'un des points marquants fut en 1953, avec l'identification de la structure tridimensionnelle en double hélice de l'ADN, la molécule commune à tout organisme vivant. Cette découverte fut conjointement réalisée par Rosalind Franklin, Francis Crick, James Watson et Maurice Wilkins (112).

À l'instar des plantes, les premières tentatives de manipulation chromosomique sur les procaryotes étaient réalisées de manière non contrôlée, notamment par emploi d'agents extrêmement mutagènes (tant pour l'organisme, modèle tant que pour le manipulateur) comme le NTG (N-méthyl-N'-nitro-N-nitroso-guanidine)(113), l'EMS (Éthyle méthanesulfonate), et les rayons ionisants (114).

La mutagenèse dirigée se définit par la création de mutations précises (de l'ordre de la paire de base) dans un génome donné.

La mutagenèse dirigée est un outil crucial à toute étude microbiologique, car elle permet d'identifier le rôle de certains locus identifiés comme acteurs potentiels dans des contextes précis. Les premiers protocoles de mutagenèse sont apparus avant l'ère génomique, peu de temps avant l'invention de la PCR (*Polymerase Chain Reaction*) en 1985. (115-117)

### 1.3.1 Recombinaison homologue

Les procaryotes se distinguent des eucaryotes par l'absence d'un noyau et leur architecture génétique plus compacte. Le plus souvent le chromosome est unique et circulaire, comme c'est le cas chez *E. coli*. Il existe cependant des exceptions. *V. cholerae*, par exemple, possède deux chromosomes circulaires, *Borrelia burgdorferi*, un spirochète GRAM négatif et *Streptomyces lividans*, deux GRAM positifs maintiennent des chromosomes linéaires (118).

Les chromosomes sont des systèmes dynamiques, tant dans leur mobilité que dans leur évolution. Ce dynamisme s'illustre dans des phénomènes comme la conjugaison, la transformation naturelle et la transduction. Les MGEs peuvent être considérés comme les catalyseurs de ce dynamisme. L'un des grands acteurs retrouvés dans les MGEs sont les plasmides. Comme les chromosomes, les plasmides sont classés comme réplicons, *c.-à-d.* molécules d'ADN initiant leur réplication à partir d'une seule origine de réplication.(119)

Le plus souvent les plasmides contiennent des gènes qui sont bénéfiques à l'hôte codant par exemple pour une résistance à un antibiotique, des facteurs de pathogénicité et/ou de virulence. Les plasmides cryptiques sont des plasmides dont le rôle est encore inconnu ou mal compris. (120)

Les plasmides conjugatifs peuvent être transmis par conjugaison (transfert horizontal) s'ils portent les gènes *tra*, qui code pour le pilus sexuel. Les plasmides ne comprenant pas de gènes de conjugaison sont dits non conjugatifs. Ces plasmides le plus souvent parasitent la machinerie de transfert (Relaxosome + transférasome). La plupart des plasmides dérivés utilisés en biologie moléculaire (on parle alors de vecteurs) sont non conjugatifs (121, 122).

Les plasmides sont capables de s'intégrer au chromosome hôte (on parle alors d'épisomes) et de devenir partie intégrante du patrimoine génétique. L'intégration d'un MGE passe par le phénomène de recombinaison. Des séquences ayant une homologie même partielle vont

chercher à s'apparier. Ce phénomène s'illustre dans la méiose, où des allèles portant chacune leur SNP vont s'échanger pour amener à la création d'un nouvel individu. La recombinaison est aussi cœur des mécanismes de réparation de l'ADN, les cassures doubles brins sont les blessures les plus compromettantes pour les génomes. La mise en place d'une synapse de réparation (RAD51 chez les eucaryotes, RecA pour les procaryotes) permet de récupérer d'une telle blessure. Les mécanismes de réparation sont beaucoup moins exigeants dans l'homologie des sections qu'ils tentent de rattacher et sont plus ou moins mutagènes selon leur degré de fidélité (HDR, MMEJ, NHEJ) (123).

La recombinaison passe par un intermédiaire, la jonction de Holliday, une structure symétrique cruciforme entre 4 brins d'ADN ayant un certain degré d'homologie. Cette structure est mobile et est capable de se déplacer tout au long des 4 brins d'ADN en gardant la même structure. Une jonction de Holiday nécessite des recombinasses appelées les résolvases. Les résolvases clivent la jonction de Holiday selon un axe de symétrie horizontal et/ou vertical (Figure 11). Ainsi selon le type de résolution et la distance parcourue par la jonction, l'événement final de recombinaison ne sera pas le même (patch ou splice) (124).

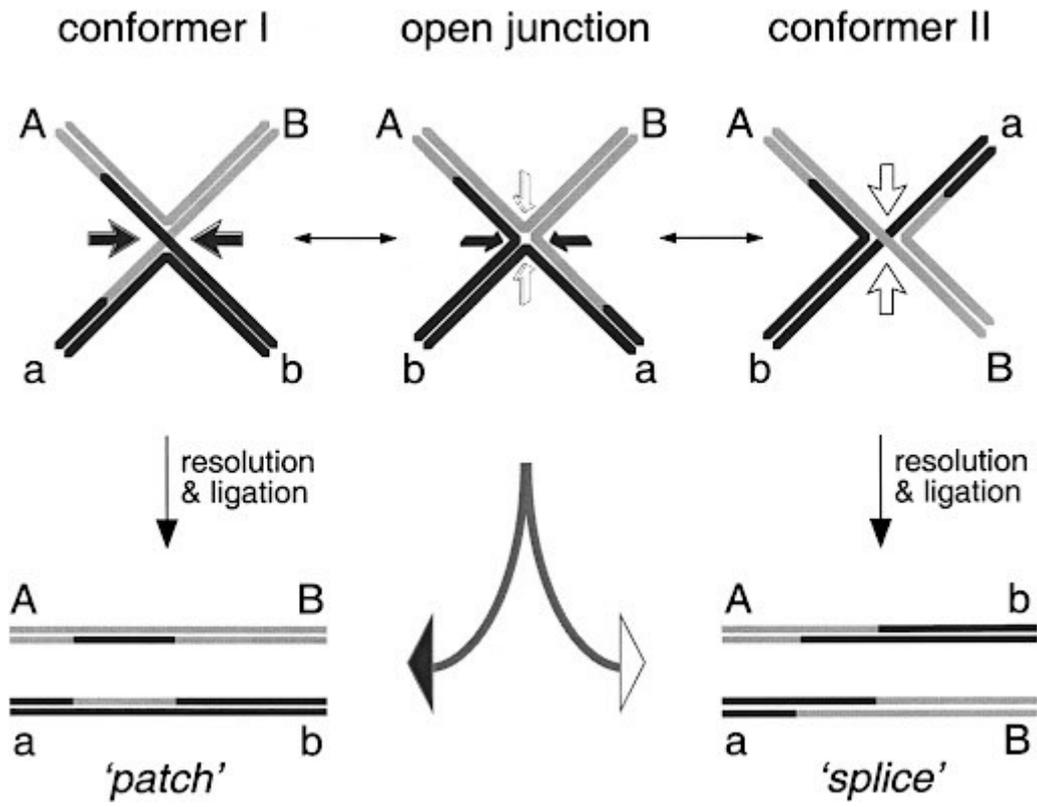


Figure 9. Résolution des jonctions de Holliday.

Les jonctions de Holliday, par leur nature symétrique, peuvent amener à plusieurs événements de recombinaison. Dans le cas de *patch* la résolution n'entraîne pas un réarrangement significatif, seul un « patch » est échangé entre les allèles. Dans le cas de *splice* y a un échange significatif entre les allèles autour du site de recombinaison. Figure reprise de Stephen C. West *et al.* 1999

### 1.3.2 Échange allélique

L'échange allélique est l'exploitation à des fins de mutagenèse de la recombinaison homologue et concerne l'échange d'allèle entre un chromosome cible (l'allèle sauvage) et un ADN donneur (l'allèle mutée). L'ADN donneur est le plus souvent retrouvé au sein d'un vecteur de type plasmidique. L'ADN présent comme insert au sein d'un vecteur permet une certaine persistance du code génétique entrant, l'ADN linéaire étant rapidement digéré par la cellule hôte. L'ADN dans sa structure peut être considéré comme un code informatique (125). Comme dans le cas d'un algorithme, on reconnaît 4 types d'opérations qu'il est possible d'opérer sur un génome : l'excision/insertion, l'inversion, la translocation et l'échange de cassette.

L'origine de répllication des vecteurs suicide est conditionnelle, et ne permet la répllication que dans des conditions précises. La plus étudiée de ces origines est R6K. Cette origine de répllication a d'abord été décrite dans un but d'élucider les mécanismes de répllication du chromosome bactérien. R6K ne peut se répliquer que si deux éléments minimaux sont présents. *L'ori*  $\gamma$ , une région riche en adénosine et thymine, et le gène *pir*, provenant du phage  $\lambda$ , qui code pour la protéine  $\pi$ . *pir* peut être présent en cis (sur le vecteur), ou en trans (porté par un autre plasmide ou le chromosome hôte). Ainsi si la bactérie transformée ne contient pas le gène *pir*, le vecteur est dit « suicide ». Seule la population ayant intégré le plasmide peut survivre à la sélection. L'échange allélique est une manipulation qui se réalise en deux temps, d'abord l'intégration du plasmide dans le chromosome hôte. Puis, l'excision du plasmide qui donne deux populations : les révertants au génotype sauvages et les mutants (126, 127).

La recombinaison est déclenchée via la construction de l'insert dans le vecteur. L'insert possède les régions homologues aux régions adjacentes au gène cible, à la fois en amont et en aval. Après transformation du plasmide, la population est sélectionnée par l'antibiotique codé pour le vecteur. Puis, lors du deuxième événement de recombinaison, l'excision du vecteur est

sélectionnée par la perte de la résistance et par le gain du phénotype de la mutation. La fréquence et la résolution des mutants obtenus par cette technique est faible. De plus il est extrêmement difficile de sélectionner l'évènement d'excision créant la mutation souhaitée. (127).

### **1.3.3 Recombinases Sites spécifiques**

Dans l'ensemble, les génomes sont suffisamment résilients pour supporter des mutations insérées.

L'un des moyens pour augmenter la fréquence et la qualité de recombinaison est l'emploi de recombinases. Les recombinases sont des enzymes capables de catalyser la recombinaison. Certaines recombinases comme RecA interviennent dans la réparation et ne discriminent pas réellement de sites. Les recombinases sites spécifiques sont des recombinases qui requièrent un site spécifique (AttB, AttP, IR, etc.), afin de catalyser la recombinaison selon un mécanisme précis. Les sites de reconnaissance sont généralement plutôt courts, d'une trentaine de paires de bases (128).

Ces enzymes sont employées en biologie moléculaire depuis maintenant 30 ans (129). Les recombinases sont divisées en deux familles, les recombinases Tyrosine (T-SSR) et les recombinases Sérine (S-SSR), en référence aux acides aminés impliqués dans la catalyse de l'enzyme. Au-delà de la composition différente en acide aminé de la poche catalytique, les deux familles de SSR se distinguent par leur mécanisme de recombinaison.

Les enzymes FLP FRT ont été d'abord décrites chez les levures (130). Les levures sont un organisme de choix dans l'étude de chromosomes eucaryotes plus compacte, car elles se transforment facilement.

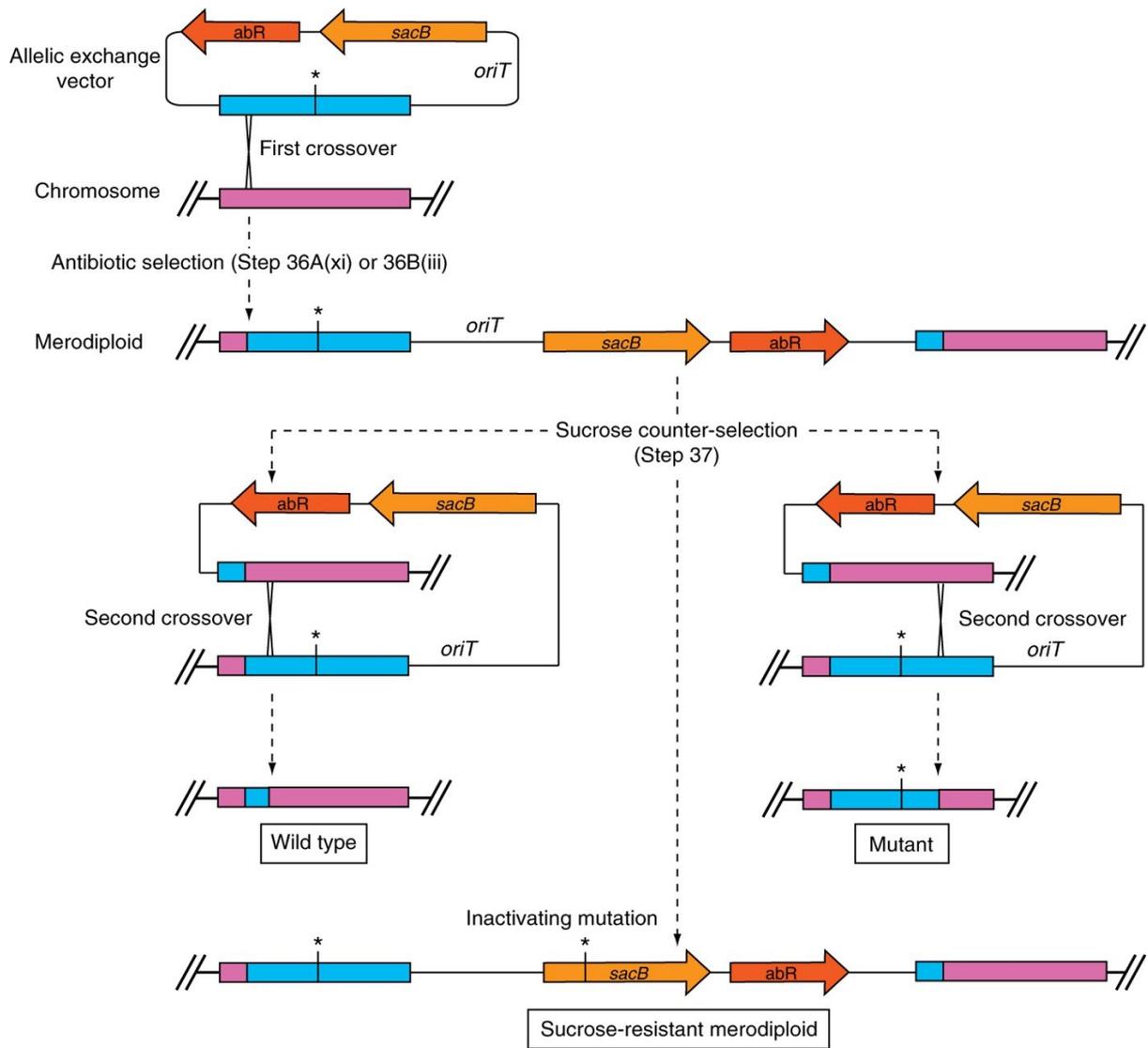
Les SSRs sont appliquées tant chez les eucaryotes, comme c'est le cas de CreLox et chez les procaryotes, avec FLP-FRT. Les SSR offre la possibilité de réaliser les 4 types d'opérations(131) sur le génome cible, mais laissent une « cicatrice » de recombinaison au locus modifié, qui peuvent créer des effets de bords (mutations non désirées) et rend difficile la création de mutations simultanée, car la présence du site de recombinaison résiduel va entraîner des recombinaisons à cet endroit (132).

XerC et XerD sont deux recombinasses tyrosine impliquée dans la séparation des dimères de chromosome à la fin de la réplication (133). Ces SSRs sont aussi impliquées dans l'intégration du phage CXT au sein de chromosome de *V. cholerae* (134).

#### **1.3.4 Contre sélection**

La contre sélection consiste à l'élimination des bactéries n'ayant pas vécu un évènement génétique précis et n'ayant donc pas le génotype souhaité. La contre sélection est un procédé naturel que l'on retrouve au sein des systèmes de partitionnement. En effet les plasmides bien que bénéfiques peuvent représenter un fardeau pour l'hôte, car ils parasitent la machinerie de réplication de l'hôte. Il existe donc des mécanismes pour assurer que les deux cellules filles reçoivent le même nombre de plasmides ou du moins une seule copie au minimum. Certains plasmides assurent le partitionnement par des gènes qui vont aider physiquement à la ségrégation des plasmides. D'autres plasmides emploient, eux des moyens plus drastiques à l'aide de couples toxine / antitoxine, chez *Enterococcus faecalis* par exemple, où le plasmide naturel pAD1 est maintenu par l'élément de stabilité ARN *par*. Ce locus code à la fois pour *fst* pour une toxine qui réprime la traduction et ARNII, un petit ARN qui est capable, par répression antisens de réprimer la traduction de *fst*. L'efficacité du système est dû à la longue demi-vie de la toxine et à la courte demi-vie de ARNII, qui rend donc *par* essentiel à la survie de la cellule fille (135, 136).

En outil de biologie moléculaire la contre sélection peut être utilisée pour « forcer » l'événement d'excision du plasmide. Le gène *sacB* code pour la levansucrase(137), cette enzyme polymérise le saccharose en levan. Ce composé osidique est une source de carbone et de réserve de sucre pour la cellule, en revanche la levansucrase peut être toxique chez d'autres organismes. Bien que les mécanismes de toxicité soient encore mal compris, l'accumulation de levan dans le périplasma et le cytoplasme de la cellule en serait à l'origine. La contre sélection par *sacB* est utilisée surtout chez les GRAM- qui sont de manière plus générale sensible que les GRAM+. Les cellules deviennent alors « sensibles » au saccharose. L'efficacité de la contre sélection par *sacB* peut être améliorée par la présence d'un second agent létal. En effet, certaines cellules peuvent échapper à la contre sélection si elles portent une mutation dans *sacB*. Court *et al.* ont créé un système *sacB-tetA*. Ou *tetA*, qui codant pour la résistance à la tétracycline provoque aussi une sensibilité à l'acide fusarique (138).



**Figure 10. Échange allélique contre sélectionné par *sacB*.**

Dans le contexte d'échange allélique par vecteur suicide, le vecteur s'intègre par un premier évènement de recombinaison dans le chromosome hôte. Cet évènement est sélectionné par la résistance que le plasmide confère. Une fois intégrée, la cellule contient deux copies d'un allèle et est dite Mero-diploïde

. Le deuxième évènement de recombinaison qui amène à l'excision du plasmide est contre sélectionné par *sacB*. Les cellules n'ayant pas excisée le plasmide codent pour un gène qui leur est létal. Figure reprise de Hmelo *et al.*, 2015.

### 1.3.5 Le génie génétique par recombinaison (*Recombineering*)

Le phage  $\lambda$  est probablement le phage le plus étudié, sa première description remonte à 1951 (139). Le phage  $\lambda$  est un phage tempéré, car il est capable de réaliser à la fois un cycle lytique qui résulte par la morte de l'hôte, et un cycle lysogénique qui résulte par l'intégration du génome du phage dans l'hôte. À l'instar du phage CXT (140), le phage  $\lambda$  se circularise peu après son entrée dans la cellule, afin de se protéger des nucléases de l'hôte (128). Suivant une infection, le phage entame la plupart du temps par la phase lytique, puis selon la concentration du répresseur *Cl* le phage peut passer en phase lysogénique, le phage intégré dans le génome est dit lysogène(141). Comme dans le cas de CRISPR, la transition d'une phase à l'autre est aussi dépendante des conditions environnementales, par exemple la présence de radicaux réactifs de l'oxygène (ROS) tend à pousser le phage à amorcer sa phase lytique (142). Le phage  $\lambda$  joue un rôle crucial en biologie moléculaire.

Le phage  $\lambda$  connaît deux mécanismes par lesquels il peut intégrer son génome au sein du chromosome de *E. coli*. Le premier passe est RecA dépendant, ou deux sites, *attB* dans le génome de *E. coli* et *attP* dans le génome du phage permettent la recombinaison et donc l'intégration. La jonction de Holliday résultante est résolue par la recombinase Int du phage et IHF de l'hôte (143).

Le second mécanisme lui fait appel à l'opéron Red (REcombination Deficient). Cet opéron permet d'emprunter une voie de recombinaison sans utiliser la protéine RecA, il fut découvert par Clark *et al.* (144). Dans leur article, les auteurs rapportent que le phage  $\lambda$  était capable de se recombiner dans le génome de mutants d'*E. coli* pour les gènes *recA*, *B*, *C*. Ces gènes étaient jusque-là considérés comme essentiels pour l'évènement d'intégration. Paradoxalement, l'opéron Red n'est pas essentiel pour le cycle de vie du phage  $\lambda$ .

### 1.3.6 Emploi de l'opéron Red en mutagenèse

L'opéron Red contient 3 gènes clés : Gam (pour Gamma), bloquant le complexe RecBCD, empêchant la dégradation du génome alors linéaire du phage lambda. Beta, une protéine capable de lier l'ADN simple brin et de catalyser la « recombinaison ». (Le terme ici peut prêter à confusion, car il n'y a pas apparition de Jonction de Holliday, structure pourtant centrale de la recombinaison). Exo, une exonucléase dégradant un des deux brins d'un homoduplex d'ADN, fonctionnant de concert avec Beta. L'homoduplex une fois digéré est inséré comme un fragment d'Okazaki au sein de la fourche de réplication. L'opéron Red accepte donc deux types de substrats : i) des fragments ADN double brins, qui nécessitent d'être digérés par Exo puis appariés à Beta et ii) des fragments ADN simple brins, qui peuvent être obtenus par synthèse chimique. En écho à son mécanisme de recombinaison, si un fragment simple brin est utilisé dans l'opéron Red il faut identifier le replichore. En effet, la réplication du chromosome bactérien est symétrique, elle commence à l'origine de réplication et se termine au site *dif*. Comme l'ADNsb est intégré comme fragment d'Okazaki il est nécessaire de savoir dans quel replichore se situe le locus que l'on souhaite modifier, afin d'identifier le brin tardif qui est répliqué de manière non continue et implique des fragments d'Okazaki.

Le mécanisme précis de l'opéron Red fut élucidé en 2010, soit 12 ans après sa première utilisation comme outil de mutagenèse chez *E. coli*. Preuve qu'il n'est pas toujours nécessaire de comprendre en intégralité un mécanisme pour l'exploiter.

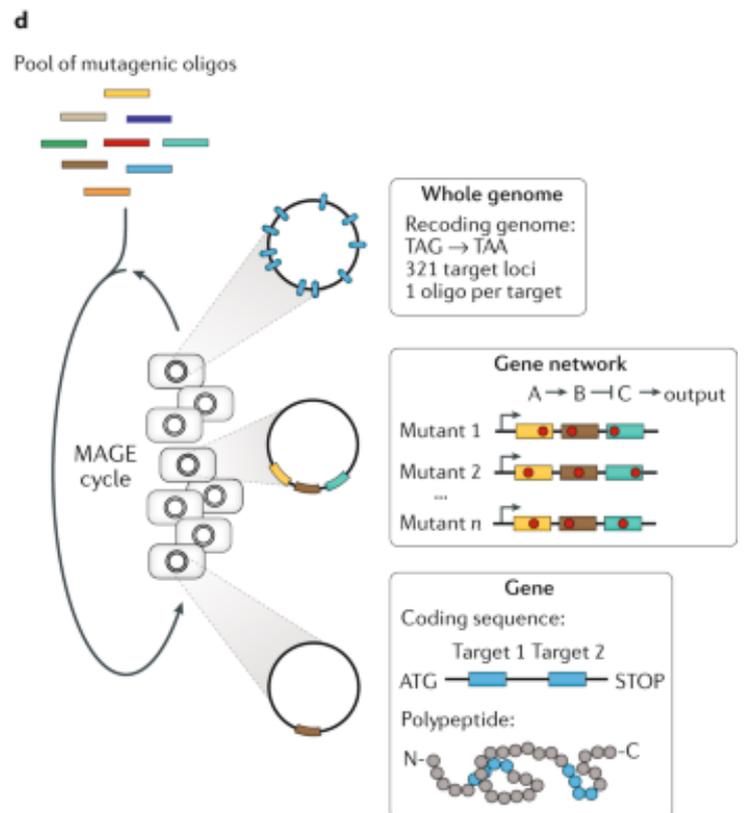
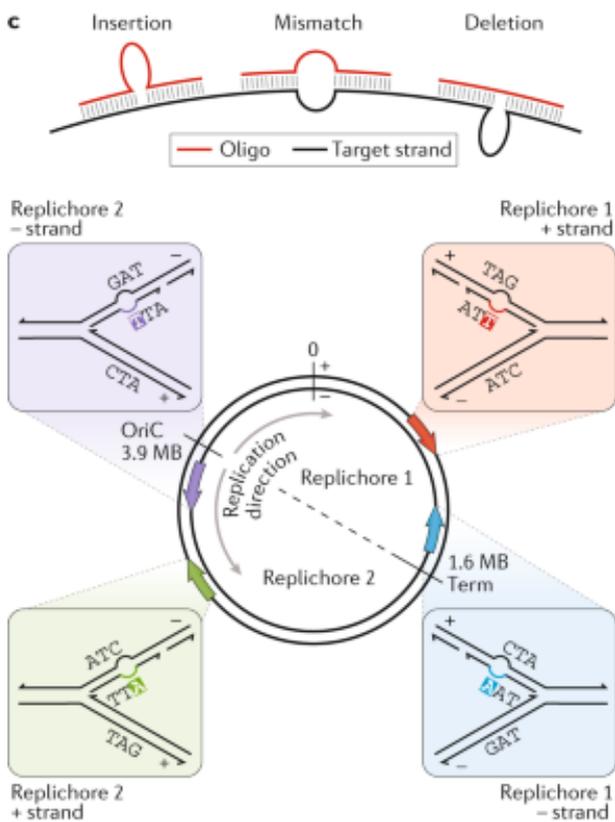
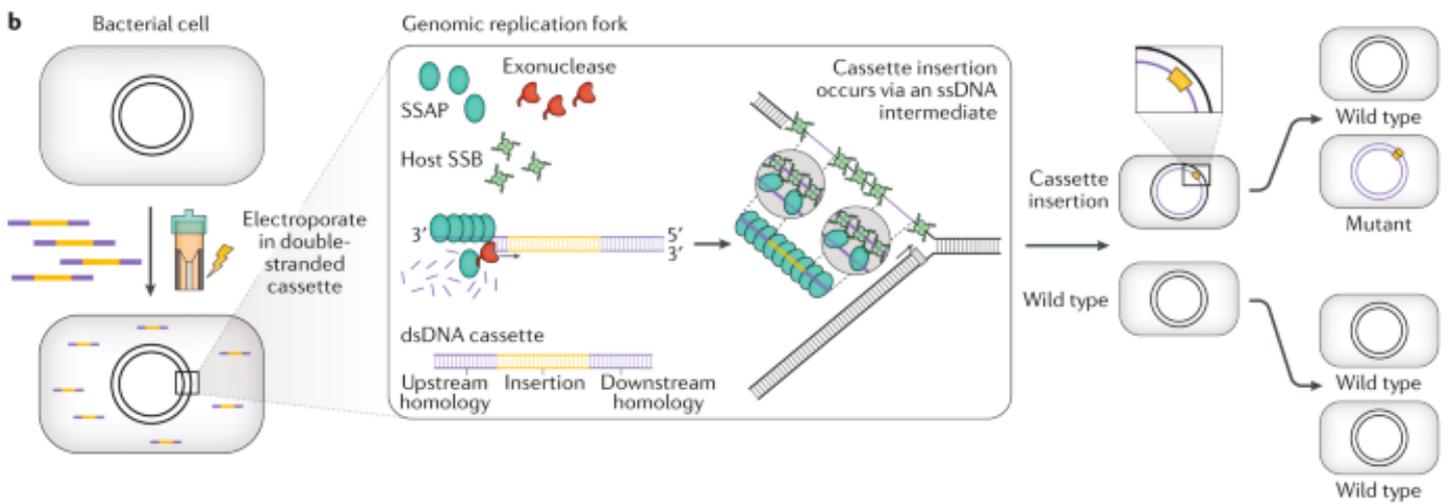
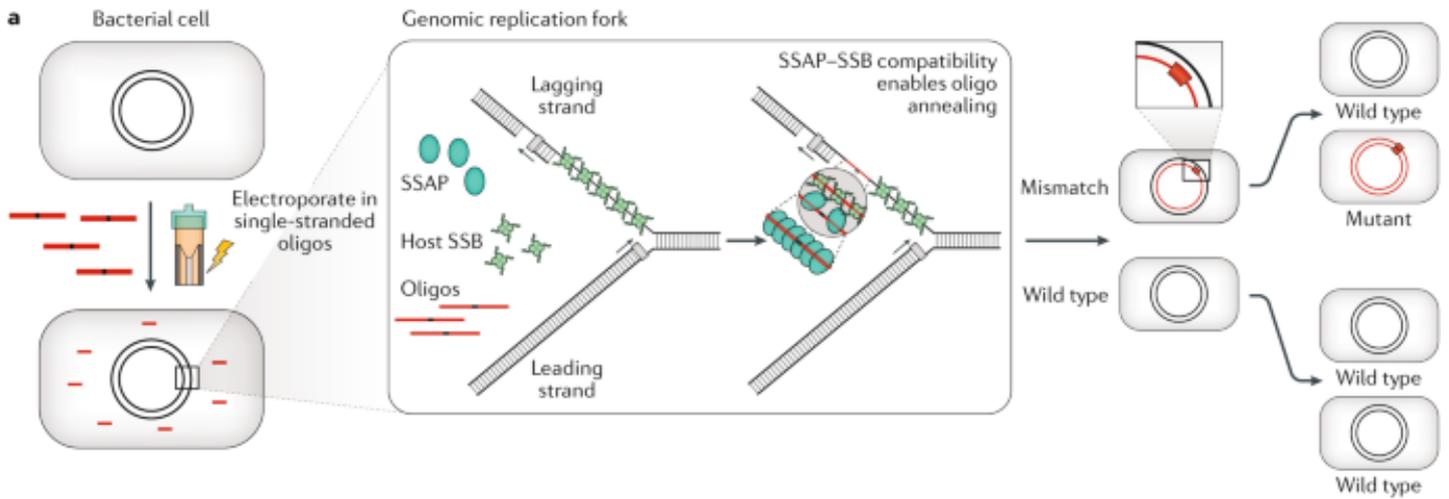
En l'an 2000, les auteurs Datsenko et Wanner adaptent une méthode déjà existante employant l'opéron Red- $\lambda$ , décrite elle en 1998 (145). Leur plateforme varie par l'emploi d'un amplicon PCR pour médier la recombinaison au lieu d'un fragment d'ADN double brin obtenu depuis la digestion d'un vecteur.(132)

Datsenko et Wanner ont créé une famille de vecteurs : Les pKD, capables de médier des insertions, délétions et mutations via des amplicons PCR contenant à la fois une cassette antibiotique pour la sélection et les régions homologues au gène visée pour la recombinaison. Les pKD sont associés à une origine de répllication, *pSC101*, qui ne permet au plasmide d'être maintenu dans la cellule qu'en petit nombre (*low copy plasmid*), les auteurs argumentent qu'avec un trop grand nombre de plasmides par cellules l'efficacité de la recombinaison est compromise. De plus l'origine de répllication est dépendante d'une protéine, Rep101Ts, qui est vitale pour le maintien du plasmide, mais qui est instable a des températures au-dessus de 33 ° C., conférant ainsi au plasmide une thermo sensibilité permettant de recycler le plasmide une fois la mutation obtenue. L'opéron Red est maintenu sous un promoteur *pAraB*, qui peut être induit par l'isomère L de l'Arabinose.

Finalement dans leur méthodologie la cassette antibiotique au sein de l'amplicon servant à la sélection est recyclée via les recombinaisons sites spécifiques FLP-FRT (Recombinases flippase, et sites de reconnaissances flippase).

Le *Recombineering* a pu évoluer pour donner naissance à des techniques comme le M.A.G.E (Multiplex Automated Genome Engineering), dans lequel les cellules sont transformées en cycle, pour faire d'abord apparaître un allèle mutant dans une population puis enrichir cette mutation au sein de cette même population afin d'obtenir une librairie. En augmentant la fréquence de l'allèle mutante, permet de cribler plus facilement la mutation.

Cependant l'opéron Red possède une utilité limitée si l'on tente de l'exporter dans d'autres espèces GRAM négatif. Church *et al.* (146) argumentent que cette portabilité limitée viendrait d'une incompatibilité entre Beta de l'opéron Red et les protéines SSB de l'hôte.



### **Figure 11. Utilisation de l'opéron Red dans la procédure MAGE**

- A. L'ADN donneur consiste en des fragments ADNsb. Une fois électroporés les fragments sont intégrés dans la fourche de réplication correspondant au brin tardif. Beta couvre les fragments ADNsb pour les protéger de la dégradation par les nucléases puis par interaction avec les SSB de l'hôte intègrent l'ADN donneur comme un fragment d'Okazaki. Suivant l'événement d'intégration et de ségrégation, 25% des cellules porteront la mutation désirée.
- B. L'ADN donneur est un fragment ADN double brin, le même raisonnement s'applique que pour un fragment ADN simple brin à la seule différence que le donneur se doit d'être digéré par Exo pour devenir un intermédiaire ADN simple brin.
- C. Opérations réalisables par un ADN donneur. La figure souligne les considérations du replichore à prendre selon la position du locus que l'on souhaite modifier.
- D. Automatisation de la procédure de Recombineering, le MAGE pour obtenir des bibliothèques complexes de souches mutantes. Figure reprise de George M. Church *et al.*, 2021.

### 1.3.7 Considérations et limitations de l'opéron Red.

L'ADN donneur remplace un fragment d'Okazaki ce dernier est donc soumis aux mécanismes de réparation des mutations qui peuvent être introduits naturellement. La réparation est médiée entre autres par MutS. Les considérations à prendre en compte sont le nombre de mésappariement successif. Il a été montré expérimentalement que l'efficacité des mécanismes BER et MMR (Base Excision Repair, Mismatch Repair) étaient grandement réduites si 4 ou plus mésappariement successif étaient présent dans la séquence de l'ADN donneur (147).

Le type de mésappariement est aussi à prendre en compte, les mésappariement T/T sont les mutations les mieux prises en charge et ont la plus grande fréquence de remplacement. À l'inverse les mutations C/C sont rarement corrigées (148).

La taille du donneur ADNsb doit être entre 70 et 90 paires de bases, en dessous de cette taille la fréquence d'intégration chute énormément, au-dessus de cette taille il existe un trop grand risque de mutations intégrées par le procédé de synthèse chimique. Pour un donneur double brin, il est recommandé d'utiliser des régions homologues d'au moins une 100aine de paires de bases pour *E. coli* (149). La position de la mutation au sein du donneur est aussi importante, la mutation doit être au minimum 9 à 15 paires de bases de l'extrémité 3' et/ou 5'. Il est à noter que la mutation doit être préférentiellement placée du côté 3', car la synthèse chimique d'oligo d'ADN est polarisée de 3' vers 5'.

Le panel de mutations qu'il est possible de créer via l'opéron Red est limité. Les insertions sont restreintes à une 30 aines de paires de bases via un donneur ADN simple brin et a 2-3Kb pour un ADN double brin. Les délétions sont aussi limitées à plusieurs kilobases. Bien que des délétions jusqu'à 42 Kilobase ont été rapportées, la fréquence d'intégration des délétions tend à diminuer proportionnellement avec la taille de ces dernières (101, 102).

Enfin, l'ADN donneur simple brin doit respecter des critères thermodynamiques. Toute séquence d'ADN va chercher à s'apparier, en s'appariant un fragment d'ADN va former des structures secondaires stables. Il est possible de quantifier la stabilité de ces structures par leur Enthalpie libre ( $\Delta G$ ), qui s'exprime en kcal/mol. Dans le cadre de l'opéron Red la limite est située à -12.5 kcal. Wang *et al* ont montrés expérimentalement que les donneurs ADN ayant une enthalpie plus faible que cette valeur perdaient en efficacité de recombinaison. Cependant cette valeur n'est pas une limite absolue, dans d'autres papiers ce groupe emploie des ADN donneur ayant une bonne efficacité de recombinaison qui dépassent cette limite, les auteurs argumentent que la synthèse chimique d'oligonucléotides est le facteur le plus limitant, car le nombre d'erreur augmente énormément à chaque paire de bases ajoutées.

#### **1.4 Mutagenèse chez *V. cholerae***

*V. cholerae* est l'agent étiologique du choléra. Le choléra est une maladie d'intérêt clinique affectant encore plusieurs millions de personnes et entraînant le décès de plusieurs centaines de milliers d'entre elles par an(150)). Sept pandémies consécutives ont été décrites au cours des 200 dernières années et la septième, ayant débuté au début des années 1960, est toujours en cours. De nouveaux foyers épidémiques sont rapportés régulièrement (aout 2021, Niger). Outre sa capacité a provoqué le choléra, *V. cholerae* est une souche d'intérêt au vu de la plasticité de son génome. Les membres de l'espèce *V. cholerae* contiennent deux chromosomes. Le deuxième chromosome aurait été d'abord un plasmide et ayant reçu des gènes essentiels au cours de l'évolution de *V. cholerae* (151, 152). On retrouve chez *V. cholerae* un phage, SXT, capable de s'intégrer dans le génome via un mécanisme extrêmement similaire au phage  $\lambda$  (153, 154).

La pathogénicité, la virulence et les résistances de *V. cholerae* proviennent majoritairement de l'interaction de la bactérie avec des MGEs. La toxine cholérique par exemple, provient de deux

événements clés, l'acquisition du pili TCP, provenant de l'infection par un phage, puis l'infection par un second phage, via ce même pilus : le phage CXT (155).

*V. cholerae* possède la capacité remarquable d'accepter et d'intégrer naturellement des fragments d'ADN environnants par transformation naturelle. Cette capacité s'illustre notamment par la capacité de *Vibrio* à changer de biotype (156-158).

Concernant la mutagenèse chez *V. cholerae*, on rapporte l'emploi de SSRs comme FRP-FLT. Comme mentionné précédemment l'opéron Red est difficilement transposable à d'autres espèces qu'*E. coli*, des auteurs avancent que cette portabilité est probablement due au fait que le phage  $\lambda$  a coévolué avec *E. coli*. L'efficacité de recombinaison de l'opéron Red est reconnue comme être 100 à 1000 fois moins efficace chez *V. cholerae* que *E. coli*. (159)

#### **1.4.1 CRISPR-Cas comme outil génétique chez les procaryotes**

Comme mentionné plus haut, les systèmes CRISPR Cas tirent profit, chez les bactéries, d'une absence de mécanisme fidèle pour la réparation des cassures doubles brins dans l'ADN. Les bactéries ne possèdent qu'une seule copie de leur chromosome, et ne disposent pas de substrat sur lequel se baser pour coordonner une réparation (160). Tandis que CRISPR-Cas fut une véritable révolution en tant qu'outil de mutagenèse chez les eucaryotes, il eut un impact plus modeste chez les bactéries, car pléthore de méthodes existaient déjà pour médier la mutagenèse dirigée chez ces organismes. Toujours chez *E. coli* plusieurs auteurs (52, 101, 102, 161) ont pu démontrer qu'il était possible de médier une contre sélection en important une librairie CRISPR accompagnée de la spCas9.

Plusieurs déclinaisons du système existent, que cela concerne le nombre de plasmides nécessaires pour médier la contre sélection, la nature de l'ADN donneur, et le type de sélection

(avec ou sans antibiotique) via ce même donneur ADN, le point commun à cette méthode reste un raisonnement de contre sélection.

Une population est d'abord générée, généralement par recombinaison homologe. Cette population contient une vaste majorité de souches sauvage et révertants et une fraction minime de bactéries contenant la mutation désirée. CRISPR-Cas9 est alors utilisé pour contre-sélectionner la population sauvage en ciblant la version non mutée du gène, afin de réduire sa taille et de pouvoir identifier le clone mutant rapidement et avec facilité.

#### 1.4.2 CRISPR No Scar

Le système No-Scar (Scarless Cas9 Assisted Recombineering) décrit en 2015 par Reisch *et al.* utilise une stratégie élégante pour créer une librairie via l'opéron Red, puis de contre-sélectionner cette librairie via CRISPR-Cas9. Au lieu d'importer l'intégralité d'une librairie, un seul espaceur est utilisé via le sgARN chimérique créé par Charpentier *et al.* (105). Ce système emploie 3 plasmides pour générer une population mutante puis pour recycler les plasmides:

Jour 1, Figure 12. La souche sauvage est transformée avec le pCas9-cr4. Ce plasmide encode pour plusieurs éléments clés : La spCas9, qui provient originellement de *Streptococcus pyogenes* reconnaissant une signature 5'-NGG-3' comme PAM (ou N peut être n'importe quel nucléotide). Le plasmide est construit dans un effort de réprimer au possible l'expression de la Cas9 en absence d'inducteur afin d'en limiter sa cytotoxicité. Cette stratégie se traduit par le clonage en amont de la Cas9 d'un RBS de faibles expressions, limitant le taux de traduction de la Cas9 et de la présence d'un tag ssRA. Ce tag est reconnu par le protéasome bactérien ClpP et catalyse la protéolyse de la Cas9. Ces efforts ont été déployés au sein de ce système, car aucun promoteur connu ne peut offrir une répression parfaite. L'auteur argumente que le

phénomène de fuite transcriptionnelle (leaky transcription) amène une faible expression basale de la Cas9 qui pourrait être toxique, et ce même sans la présence d'un espaceur. Finalement le plasmide contient le promoteur TetA/TetR qui code à la fois pour le gène TetR et pour des motifs de 19 paires de bases appelés TRE (*Tetracycline Response Element*). (162)

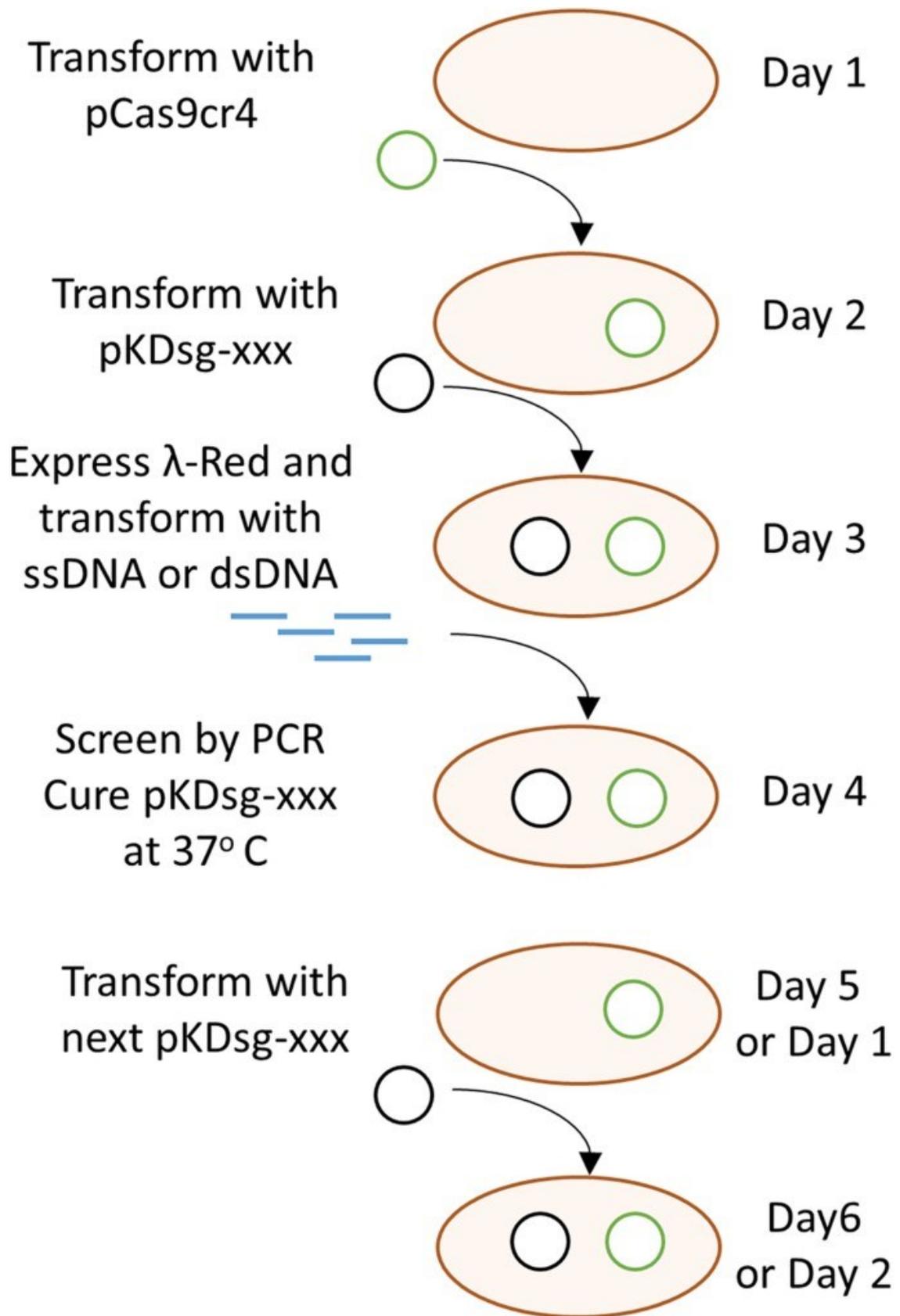
En absence d'inducteur, TetR est autorépressible et lie les TRE en bloquant physiquement l'accès à l'ARN Pol au locus. La Cas9 est placée sous le contrôle de ce promoteur et permet donc de réprimer l'expression de cette dernière. Ce promoteur provient du transposon Tn10 et offre un excellent compromis entre répression et induction (163). Malgré cette stratégie il reste impossible de cotransformer les plasmides, l'auteur argumente que TetR doit atteindre une concentration intracellulaire cible afin de pouvoir médier son action de répresseur. La molécule permettant de relaxer le promoteur est l'aTc, un dérivé de la tétracycline ne possédant pas de réelle activité antimicrobienne.

Les deux autres plasmides sont tous les deux types de pKD. Au sein de ce système, ces plasmides sont appelés « pKDsgRNA ». La structure de base de ces plasmides provient du pKD46 décrit par Datsenko et Wanner (132), les pKD conservent leur capacité thermosensible pour recyclage à la fin du protocole. Le protocole se base sur la mutagenèse via l'opéron Red comme décrit précédemment et utilise donc le pKD46 comme base pour générer une librairie. Le pKD46 a été modifié pour contenir l'ARN chimérique (sgARN+espaceur) placé sous le contrôle du promoteur pLTetO1, cet ARN code pour une région qui n'existe que dans le chromosome de la souche sauvage et qui est supprimé par la mutation apportée. Ce promoteur est aussi réprimé par TetR et peut être relaxé par la présence d'aTc. L'intégration de la séquence cible (N-20) dans le plasmide est réalisée par CPEC (**C**ircular **P**olymerase **E**xtension **C**loning), où un pKD est utilisé comme substrat pour générer deux fragments par PCR. Ces fragments possèdent deux régions homologues, une région « d'ancrage » de 300 paires de bases, et une région variable contenant le nouvel N-20. Après digestion par DpnI, une enzyme de restriction

capable de cliver uniquement des substrats méthylés retrouvés seulement dans le plasmide substrat et pas au niveau des fragments générés par PCR, les fragments sont purifiés sur gel puis assemblés lors d'une dernière PCR dans un cycle de Gibson assembly. La construction du plasmide néoformée est alors transformée dans DH5 $\alpha$ , purifiée et transformée dans la souche d'intérêt contenant déjà le pCas9-cr4 (Jour 2 Figure 12)

Ensuite, un fragment d'ADN double brin recombinant contenant la mutation désirée est transformée dans la souche d'intérêt contenant les 2 plasmides et la recombinaison par le système  $\lambda$ -Red est induite (*Cf. 2.4 Cycle No-Scar*). Ensuite, l'expression de la spCas9 et de l'ARN chimérique est activée afin de contre-sélectionner les souches mutantes (Jour 3, Figure 12).

Finalement, une fois qu'un cycle de contre sélection est réalisé et qu'un mutant a été isolé, le pKD contenant l'espaceur ayant permis la contre sélection est recyclée par plusieurs passages successifs à 37-42°C (Jour 4, Figure 12). Le pCas9-cr4 ne possède pas de propriété inhérente pouvant servir à le recycler, le système fait alors appel à un second pKD, le pKD-p15A, dont le N-20 vise l'origine de réplication (p15A) du pCas9-cr4. Une fois le pKD-p15A transformé, un nouveau cycle de contre sélection est réalisé afin de vérifier que les souches ont bien perdu le pCas9-cr4. Une fois le plasmide perdu plusieurs passages à 37-42°C sont réalisés pour recycler le pKD-p15A. (Jour 5 et 6, Figure 12).



### **Figure 12. Étapes du protocole No-Scar.**

Les plasmides pCas9-cr4 et pKDsgRNA sont transformés séquentiellement dans une souche d'intérêt. Une fois que la souche a reçu les deux plasmides, elle est cultivée en présence de L-Arabinose pour induire l'opéron Red, puis un ADN donneur est transformé (ici des oligo ADNsb), l'ADN donneur est recombiné au sein de la fourche de réplication comme fragment d'Okazaki. Après une période de récupération variable dans du milieu riche la réaction de transformation est étalée sur du milieu contenant de l'aTc. Cet inducteur amené à l'expression de la Cas9 et du sgARN, ces deux éléments forment un complexe de reconnaissance CRISPR. Si une cible est reconnue par ce complexe, une cassure double brin est introduite et entraîne la mort bactérienne, contre sélectionnant alors la population sauvage et ne permettant uniquement la survie de la population mutante au sein d'une librairie. Figure reprise de Kristala L. J. Prather *et al.*, 2015.

## 1.5 Problématique

Bien que *V. cholerae* soit capable de transformation naturelle il est difficile de manipuler de *V. cholerae* pour y apporter des mutations précises. Ainsi, malgré son importance clinique et son rôle d'organisme modèle évolutif les méthodes permettant de créer aisément des mutations dans cet organisme demeurent fastidieuses et de relativement faible efficacité. Bien que l'emploi de SSRs ait été rapporté, ces derniers comme discuté précédemment n'offrent qu'un panel de possibilité limité. L'opéron Red se présente comme une alternative prometteuse, cependant il n'est que partiellement compatible avec des bactéries autres qu'*E. coli*. Les bactéries sont souvent considérées à tort comme des organismes primitifs bien qu'elles possèdent des capacités qui ont une application presque infinie dans plusieurs secteurs variés (dégradation de métaux lourds, de plastique, secteur agroalimentaire, application clinique, microbiome). Bien que plus simples comparé aux eucaryotes, les bactéries nous ont donné des outils précieux, comme les antibiotiques et plus récemment des outils de mutagenèse avec la découverte de CRISPR.

La pandémie depuis 2020, causée par le Coronavirus SARS-CoV-2 responsable de la COVID-19, nous rappelle à quel point il est important pour nos sociétés modernes de comprendre l'évolution de souches pathogéniques et leur interaction entre chaque règne du vivant. Pour ce projet nous nous sommes penchés sur l'implémentation d'une plateforme CRISPR pour *V. cholerae* afin de faciliter la construction de mutant. Notre objectif était alors d'optimiser deux points cruciaux : diminuer la cytotoxicité de la Cas9 et augmenter l'efficacité de recombinaison de l'opéron Red.

# Chapitre 2 : Matériel et méthodes

## 2.1 Souches utilisées

<b>Souches</b>	<b>Génotype d'intérêt</b>
A1552 ( <i>Vibrio cholerae</i> )	<i>lacZ+</i> , <i>SXT-</i>
DH5A ( <i>Escherichia coli</i> )	<i>lacZ-</i> , $\Delta$ <i>recA</i>
MG1655 ( <i>Escherichia coli</i> )	<i>lacZ+</i> ,
V52 ( <i>Vibrio cholerae</i> )	<i>lacZ+</i> , <i>SXT-</i>
MO10 ( <i>Vibrio cholerae</i> )	<i>lacZ+</i> , <i>SXT+</i>
A1552 $\Delta$ <i>ompV</i> ( <i>Vibrio cholerae</i> )	$\Delta$ <i>ompV</i>
A1552 $\Delta$ <i>ompW</i> ( <i>Vibrio cholerae</i> )	$\Delta$ <i>ompW</i>
A1552 $\Delta$ <i>lacZ</i> ( <i>Vibrio cholerae</i> )	$\Delta$ <i>lacZ</i>
MG1655 $\Delta$ <i>lacZ</i> ( <i>Escherichia coli</i> )	$\Delta$ <i>lacZ</i>

### 2.1.2 Réactifs et concentration

<b>Antibiotique / Réactif</b>	<b>Concentration finale <i>V. cholerae</i></b>	<b>Concentration finale <i>E. coli</i></b>
Kanamycine ( $\mu$ g/mL)	100	100
Chloramphénicol ( $\mu$ g/mL)	5	35
Spectinomycine ( $\mu$ g/mL)	100-200	50-120
Streptomycine ( $\mu$ g/mL)	75-150	50
Tétracycline anhydrique (aTc) ng/mL	100-200	100-200
L-Arabinose (m/v)	0.2%	0.2%
D-Glucose (m/v)	1%	1%

### 2.1.3 Milieux de cultures

-LB/LBA : Les souches décrites dans ce mémoire ont été cultivées dans du milieu Lysogeny-Broth / Luria-Bertani, contenant 10 g/L de Tryptone, 5 g/L de NaCl et extraits de levure. Pour des milieux solides de l'Agar (LBA) est ajouté pour une concentration finale de 20g/L. Après ajout de 1 litre d'eau désionisé, le milieu est autoclavé.

-SoB/SoC : Pour les cycles No-Scar les souches ont été cultivées dans du milieu SoB (Super optimal broth, catabolic repression), contenant les concentrations suivantes : 20g/L de Tryptone, 5g/L d'extrait de levure, 0.5g/L de NaCl. Avant autoclavage 10mL d'une solution de KCl 1.86% est ajouté au milieu. Après autoclavage 5mL d'une solution stérile de MgCl<sub>2</sub> 19% est ajouté au milieu. Pour obtenir du SoC 50μL d'une solution de glucose 20% stérile est ajouté pour 1 mL de SoB.

### 2.1.4 Amorces

Amorces employées pour la vérification du génotype après un cycle No-Scar.

Tableau 1. Séquences des amorces utilisées pour génotypage

<u>Séquence</u>	<u>Nom</u>
cgcgccacataatacgcaact	Vérif <i>ctxAB</i> F1 (A1552)
tgggcgacagggtcatcacatttac	Vérif <i>ctxAB</i> R1 (A1552)
gaccaagagtccattggggcg	Vérif <i>ompV</i> F1 (A1552)
cggtcagaagtgctagccga	Vérif <i>ompV</i> R1 (A1552)
acaaccattcttcagaaggaatc	Vérif <i>lapLapX</i> F1 (A1552)
ggtatggatgtggtagaagtg	Vérif <i>lapLapX</i> R1 (A1552)
acaggcgatgactaacctac	Vérif <i>lacZ</i> F1 (A1552)
ggagagtcaaagcgcagaac	Vérif <i>lacZ</i> R1 (A1552)
tgcaccacagatgaaacgcc	Vérif <i>lacZ</i> F1 (MG1655)
cggctcgtatgtgtgtgga	Vérif <i>lacZ</i> R1 (MG1655)

## 2.2 ADN Donneur mutagène simples brins

Les ADN donneurs simple brin suivant permettent de médier une mutation au sein des deux chromosomes de *V. cholerae* et du chromosome de MG1655

Tableau 2. Séquence des ADN donneurs simples brins

<u>Séquence</u>	<u>Gène cible</u>	<u>Réplicore</u>
tacagtgtcacatccaacataagctcttaattggaaggacatatcggctagcacttctgaacgataccgccttcagacaggagctcct	<i>ompV</i>	+
ttattgtgggatgacgcttcacacgtaaggtgagctgatactgatcggagaagttgcgcataagatcctcttgaatacgcgaagcaac	<i>lacZ</i>	-
ttaattgtttgatcaattattttctgttaacaaaggagcattatgatataaaaaagccacctca	<i>ctxAB</i>	+
attgtttgatcaattattttctgttaacaaaggagcattatgatataaaaaagccacctcagtgggctttttgtggtcgtatga	<i>ctxAB</i>	+
aattgcattacttttattaacaaccattcttcagaaggaatccttgggaaaagcgatatacacaacaactggagttgctggtcga	<i>laplapX</i>	-
tgtgagcggataacaattcacacaggaaacagctaggccagacgcgaattattttgatggcgtaact	<i>lacZ</i> (MG1655)	+

Les ADN donneurs simple brin ci-dessous se recombinent sur le pKD, ils remplacent la cible suicide et la région plastique. Le plasmide cible est pRaava, l'organisme visé est *V. cholerae*. Le dernier donneur de cette liste est l'ADN donneur Anubis, il se recombine sur le pKD non-suicide.

Tableau 3. Séquence des ADN donneurs simple brin recombinant sur pRaava

<u>Séquence</u>	<u>Cible</u>
<u>ccttattttaactgctatttctagctctaaaaccaacacacaaatatacgcggtgctcagtatctctatcactgatagggatgtaa</u>	<i>ctxAB</i>
<u>gccttattttaactgctatttctagctctaaaactcagcaagcgttagcgttcGGtgctcagtatctctatcactgatagggatgtaa</u>	<i>lacZ</i>
<u>gccttattttaactgctatttctagctctaaaacgactttgtcttcagcattgggtgctcagtatctctatcactgatagggatgtaa</u>	<i>laplapX</i>
<u>gccttattttaactgctatttctagctctaaaacaaccagtgttacactcgtgtgctcagtatctctatcactgatagggatgtaa</u>	<i>ompV</i>
<u>gccttattttaactgctatttctagctctaaaacggttggactcaagacgataggtgctcagtatctctatcactgatagggatgtaa</u>	<i>ori p15A</i>

Les ADN donneurs simple brin ci-dessous se recombinent sur le pKD\_Horus, ils remplacent la cible suicide et la région plastique. Le plasmide cible est pHorusV1 et pHorusV74, les organismes visés sont *V. cholerae* et MG1655.

**Tableau 4. Séquence des ADN donneurs simple brin recombinant sur pHorusV1/V74**

<b><u>Séquence</u></b>	<b><u>Cible</u></b>
<u>agccttattttaacttccccaacgaccgggtctaaaactcagcaagcgtagcgcttcgtgctcagtatctctatcaactgatagggat</u>	<i>lacZ (E.coli)</i>
<u>agccttattttaacttccccaacgaccgggtctaaaactcagcaagcgtagcgcttcgtgctcagtatctctatcaactgatagggat</u>	<i>lacZ (A1552)</i>

### 2.2.1 ADN Donneurs doubles brins

Ci-dessous sont présenté les séquences des amorces utilisées pour générer des ADN donneurs doubles brins.

**Tableau 5 Séquence des amorces utilisées pour générer des ADN donneurs doubles brins.**

<b><u>Séquence</u></b>	<b><u>Nom</u></b>
gatgcctgatttcegcgtac	dsDNA FR <i>lacZ</i> F1
tgcttcgcgtattcaagaggatcttgccagagagccttaaggctctctttt	dsDNA FR <i>lacZ</i> R1
aaaaagagagccttaaggctctctggcaagatcctcttgaatacgcgaagca	dsDNA FR <i>lacZ</i> F2
aacgtttacacaattggtgagtggtc	dsDNA FR <i>lacZ</i> R2
catgccgaactcagaagtgaaacg	pKD Up F1
atcttgtcaatcattgtcctgtcagcagaccaagt	pKD Up R1
gctgacaggacaatgattgaacaagatggattg	KanR F1
ttgaaaaggaagagttcagaagaactcgtcaagaagg	Kan R1
tgacgagttctctgaactcttcttttcaatattattgaagcatttatcagg	pKD Down F1
caggtggcacttttcgggga	pKD Down R1

## 2.2.2 Scaffold synthétiques

Ci-dessous sont présentées les séquences des scaffold construits pendant ce mémoire. **ALT** désigne le Scaffold protégé par remplacement de la séquence plastique (*Cf. 4.2 Création de la plateforme Horus*). La région plastique est inscrite en majuscule.

Tableau 6. Séquence des scaffold utilisés dans ce mémoire

<u>Séquence</u>	<u>Nom</u>
gtttagagctagaaatagcaagttaaaataaggctagtcggtatcaactgaaaaagtgccaccgagtcggtgc	Scaffold original
gtttagacctagaaataggaagttaaaataaggctagtcggtatcaactgaaaaagtgccaccgagtcggtgc	pKD
gtttagacctaTCCctaggaagttaaaataaggctagtcggtatcaactgaaaaagtgccaccgagtcggtgc	pKD_ALT
gtttagacccgggaaattggggaagttaaaataaggctagtcggtatcaactgaaaaagtgccaccgagtcggtgc	HorusV1
gtttagacccggTCGTtggggaagttaaaataaggctagtcggtatcaactgaaaaagtgccaccgagtcggtgc	HorusV1_ALT
gtttagaccacgggacatcgtggaagttaaaataaggctagtcggtatcaactgaaaaagtgccaccgagtcggtgc	HorusV74
gtttagaccacgACTGCtcgtggaagttaaaataaggctagtcggtatcaactgaaaaagtgccaccgagtcggtgc	HorusV74_ALT

### 2.3 Transformation de plasmides chez *V. cholerae*

Une colonie de *V. cholerae* est utilisée pour ensemençer un volume de 5mL de milieu LB ou SoB puis est placé à 37°C avec une agitation à 200RPM. La culture est laissée overnight pendant 16 heures dans ces conditions d'incubation, il est vivement conseillé de respecter cette fenêtre de temps. La compétence semble décliner une fois les 16 heures atteintes. 1.5mL de cette culture en phase plateau est transférée dans un Eppendorf 1.5mL puis est placé sur glace pendant environ 30 min.

Une fois l'incubation sur glace terminée les cellules sont centrifugées une première fois à 6000RCF, pendant 2 minutes, dans une centrifugeuse aussi réfrigérée à 4°C. Le surnageant est enlevé par retournement. Le culot est alors resuspendu dans 1mL de glycérol 10% (maintenu aussi à 4°C). Cette étape est répétée deux fois pour un total de 3 lavages.

L'étape de lavage consiste à nettoyer les sels contenus dans le milieu LB pour éviter la formation d'arcs électriques pendant l'électroporation et de maintenir les cellules dans un milieu isotonique. Après le troisième lavage les cellules sont resuspendues dans un volume final de 50 à 100uL de glycérol 10% (Le rapport volume culture/nombre de réactions est de 1.5mL pour 1 réaction) et sont alors réputées compétentes.

50 à 100µL de ces cellules compétentes est mélangé avec 200ng de plasmide, il est important que le volume de vecteur ne dépasse pas 2.5µL. Le vecteur doit être resuspendu dans de l'eau désionisée pour éviter l'introduction d'agent chimique conducteur. Le mélange cellules compétentes/vecteur est placé dans une cuvette d'électroporation 1mm, le voltage est réglé à 1.25kV pour *V. cholerae*. La constante de temps doit être de 5ms. L'efficacité de transformation est grandement diminuée, voire nulle, pour les constantes plus courtes. Immédiatement après électroporation 1mL de SoC préchauffé est ajouté directement dans la cuvette puis replacé dans un tube Falcon 50mL. Il est important de récupérer le plus de cellules

compétentes, mais il est nécessaire de manipuler les cellules avec soin, car ces dernières sont fragilisées par l'électroporation. Les cellules sont incubées une heure à 37°C ou 30°C si le plasmide est thermosensible et avec agitation 200RPM, il s'agit de l'étape de récupération, cette étape est essentielle, car les cellules expriment le gène de résistance codé par le vecteur qui permet de sélectionner les cellules ayant reçu le plasmide.

Une fois le délai de récupération terminé 100µL du volume de récupération est étalé sur milieu contenant l'agent antibiotique permettant la sélection des cellules transformée. Les 900µL restants sont centrifugés 2 minutes à 6000RCF puis sont resuspendus dans 100µL de PBS/LB/SoB pour être finalement étalé sur milieu contenant le même agent antibiotique. Le temps et température d'incubation de ces milieux est variable selon le vecteur transformé, par exemple, si la population contient déjà un plasmide, la nature de l'insert etc. Pour une souche sauvage de *V. cholerae* transformé avec un vecteur simple comme pUC18 ou pBAD24, le temps d'incubation est de 15-16h.

### 2.3.1 Transformation chimique de plasmides chez *E. coli* :

La transformation chimique reprend le même principe que l'électroporation. L'efficacité de ce protocole est inférieure par rapport à l'électroporation, cependant les cellules traitées par ce protocole sont plus stables lorsque stockées au  $-80^{\circ}\text{C}$  :

Dans ce protocole l'entrée d'ADN exogène se réalise via un choc thermique (passage de  $4^{\circ}\text{C}$  à  $42^{\circ}\text{C}$ ). Les cellules sont récoltées au début de leur phase exponentielle ( $\text{DO}_{600}$  entre 0.3-0.4) puis sont incubées 10 minutes sur glace.

Les cellules sont culotées une première fois et resuspendues dans une solution contenant 10% de glycérol et 2g/L de  $\text{MgCl}_2$ . Les ions  $\text{Mg}^{2+}$  et  $\text{Cl}^-$  jouent un rôle isotonique et sont réputés comme acteurs potentiels pour la formation de pore pendant un changement brutal de température (étape de choc thermique). Le ratio volume culture / réaction est de 1.5mL pour une réaction. Usuellement 50mL de culture est traité en une fois pour faire 10 stocks glycérols contenant 600 $\mu\text{L}$  de cellules compétentes. Après avoir été resuspendues les cellules sont incubées de nouveau sur glace pendant 20 minutes suivit d'une étape de lavage finale. La présence de sels est ici de moindre importance que pour l'électroporation. Les cellules sont alors resuspendues dans un volume final de 200 $\mu\text{L}$  et réputé chimio compétent.

100 à 200 ng de plasmide est ajouté aux cellules compétentes. Que le plasmide provienne d'une miniprep ou d'une construction *in vitro* (ligation ou CPEC), il est préférable de limiter le volume maximal à 5 $\mu\text{L}$ . après ajout de l'ADN la mixture est placée sur glace pendant 30 minutes. Passé ce délai les cellules sont placées dans un bain sec à  $42^{\circ}\text{C}$  pendant 1 minute. Il est crucial de respecter scrupuleusement ce temps d'incubation. Après 1 minute les cellules sont replacées sur glace pendant deux minutes.

Après cette incubation sur glace, 1mL de milieu SoC préchauffé est ajouté aux cellules qui sont incubées avec agitation (200-300RPM) et à température dépendante selon la nature thermosensible du plasmide.

### **2.3.2 Électroporation de plasmides chez *E. coli* :**

Le protocole d'électroporation de *E. coli* est similaire au protocole de transformation de *V. cholerae*. La principale différence est une plus grande flexibilité des conditions de culture. La culture peut être démarrée depuis l'ensemencement d'un bouillon (culture directe) ou depuis une culture over-day, ensemencée depuis une culture primaire (ON). Dans les deux cas la DO<sub>600</sub> de départ avoisine 0.05. La culture est placée sur glace dès que le DO atteint 0.4-0.5, signe que les cellules entament la phase de croissance exponentielle. Le ratio Volume culture /réaction est aussi différent. Pour *E. coli* il est de 2.5mL de culture par réaction. Les cellules sont électroporées à 1.8kV pour *E. coli*.

Pour l'étape de sélection, il est recommandé de limiter le volume étalé ou de réaliser des dilutions sans quoi un tapis de cellule est obtenu.

### **2.4 Cycle No-Scar**

Le protocole de transformation d'ADN donneur (simple brin ou doubles brins) est similaire aux protocoles de transformation de plasmides. Nous appelons un cycle No-Scar toutes les manipulations permettant d'obtenir des mutants via recombinaison par l'opéron Red et contre sélection par le système CRISPR-Cas9.

Un cycle No-Scar emploie une souche contenant déjà le pCas9-cr4 transformée récemment avec un pKD. Le cycle No-Scar requiert aussi des cellules en phase de croissance exponentielle. Les cultures overnight sont incompatibles avec un cycle No-Scar.

Le ratio volume culture est aussi changé. Pour *E. coli* 1mL de culture est employé par réaction, pour *V. cholerae* le volume est de 2.5mL par réaction.

Pendant leur croissance les cellules sont incubées en présence de L-Arabinose pour induire l'opéron Red. La concentration finale de L-Arabinose est toujours 0.2%. Pour *E. coli* l'induction est réalisée pendant 15 minutes dès que la culture atteint une DO<sub>600</sub> de 0.4. Pour *V. cholerae* l'induction est réalisée dès le début de la culture.

Une fois la DO<sub>600</sub> atteinte les cellules sont placées sur glace pendant 20 à 30 minutes. Les conditions de lavages restent inchangées pour *E. coli*. Pour *V. cholerae* la durée de la centrifugation est de 5 minutes par lavage. La solution de lavage peut être du glycérol 10% ou du Sorbitol 1M. Le sorbitol est conseillé si l'opéron Red a été remplacé par l'orthologue SXT. Il est à noter qu'une portion de cellule reste toujours en suspension pour *V. cholerae* et est perdue à chaque lavage.

La concentration d'ADN donneur est de 1µM par ADN donneur pour *E. coli*, ainsi dans un cycle No-Scar via la construction Horus la concentration finale est de 2µM. La concentration pour *V. cholerae* est de 5µM par ADN donneur, donc une concentration finale de 10µM si deux types d'ADN donneurs sont présents.

Dans le cas d'un ADN donneur double brin la concentration est de 100 à 200ng pour *E. coli* et de 100 à 1000ng pour *V. cholerae*.

Pour un cycle No-Scar nous dénommons réaction la population ayant reçu de l'ADN d'intérêt, le témoin lui concerne la population ayant reçu un ADN de contrôle ou uniquement de l'eau.

La récupération aussi est plus longue dans ce protocole. Pour les plasmides une heure est suffisante, pour la ségrégation de clones mutants la période est plus longue. De par son fonctionnement, l'opéron Red ne recombine que sur le brin tardif en se comportant comme un fragment d'Okazaki, le brin opposé lui, conserve un génotype sauvage. Il est ainsi nécessaire de laisser les cellules récupérer plus longtemps afin que l'allèle mutante soit sur les deux brins. Le temps minimal correspond donc à deux cycles de réplication du chromosome. En sus de la ségrégation, la contre sélection est amorcée directement dans le volume de récupération, par ajout d'aTc à une concentration finale de 100 à 200 ng/mL après au minimum 3 heures de récupération.

La réaction de transformation est laissée 2 heures supplémentaires après ajout de l'aTc puis un spot assay est réalisé. La concentration en aTc et chloramphénicol reste identique dans le milieu de sélection solide. Si pHorus est utilisé dans un cycle No-Scar, la concentration en spectinomycine peut varier de 50 µg/mL à 120 µg/mL, selon le degré de contre sélection choisi.

## **2.5 Clonage par enzymes de restrictions :**

Le clonage par enzyme de restriction est l'une des stratégies de clonage les plus employées, elle peut se retrouver dans la littérature sous le nom de clonage classique.

Le principe repose sur la digestion d'un amplicon PCR possédant des sites de restrictions et la digestion d'un vecteur d'accueil. On reconnaît deux types de digestion : Extrémité cohésive et franche. Pour les constructions des gènes *ompV* et *ompW* un site HindIII (5'- AAGCTT -3') a été rajouté en aval de deux gènes, et un site EcoR1 (5'- GAATTC -3') en amont des deux gènes. Les amorces utilisées, en plus de reconnaître leur gène et de contenir les sites de

restrictions les amorces contiennent aussi 6 paires de bases supplémentaires, ces motifs sont nécessaires pour que les enzymes puissent s'accrocher sur leurs cibles.

Les amplicons sont nettoyés sur colonne d'affinité puis 1µg de chaque amplicon est doublement digéré 1 heure à 37°C avec un mix contenant 1 unité des deux enzymes de restrictions. Le plasmide d'accueil (ici pBAD24) est digéré dans les mêmes conditions.

Après digestion, les amplicons et le plasmide sont migrés sur gel d'agarose 1% par électrophorèse (90V pendant 45minutes ou 120V pendant 25 minutes) avec 1X de tampon TAE.

Après purification sur gel, les produits sont ligués par la ligase T4 pendant 20 minutes à température pièce. 5µL de ce produit de ligation est utilisé pour transformer des cellules chimio compétentes.

### **2.5.1 CPEC (Circular Polymerase Extension Cloning)**

La CPEC est un protocole de biologie moléculaire ayant pour but de régénérer un plasmide *in vitro*. Elle emploie 2 paires d'amorces. Ces amorces possèdent usuellement de longues régions homologues, dans le cas du protocole No-Scar les amorces de CPEC fournies comprennent une région homologue d'ancrage, longue de 300 paires de bases et une région mutante de 20 paires de bases, qui permet de créer une nouvelle cible par PCR mutagène.

Premièrement, deux réactions PCR séparées sont réalisées afin de créer deux fragments capables de se fusionner en plasmide via homologie respective. Deux amorces sont communes à chaque réaction PCR : pKDa\_R1 et pKDb\_F1. Les fragments sont par la suite migrés puis sont purifiés sur gel d'agarose 1% afin de sélectionner le produit attendu. La CPEC, comme toute réaction d'amplification PCR est suspecte de créer des fragments non spécifiques, ces

derniers peuvent être par exemple : une bande parasite issue d'une mauvaise conception d'amorces ou bien des dimères d'amorces, surtout si ces dernières présentent un degré d'homologie. Une fois les fragments purifiés sur gel ils sont traités par DpnI. Cette enzyme de restriction reconnaît le motif GATC, cependant ce motif doit être méthylé pour que l'enzyme soit capable de le digérer. Cette particularité permet de digérer le vecteur ayant servi de *gabarit* tout en laissant les fragments amplifiés par PCR intacts (ces derniers ne sont pas méthylés).

Une fois les fragments purifiés et digérés, ils sont utilisés dans une réaction PCR sans amorces. Les régions d'homologies permettent à la polymérase de circulariser les deux fragments. Uniquement 15 cycles sont nécessaires. Ci-dessous la séquence des amorces utilisées pour la génération de pKD codant pour un espaceur ciblant le chromosome. Les séquences « Lien F1/R1 » désignent les séquences rajoutées en amont et aval des espaceurs pour la conception des amorces régénérant une cible.

**Tableau 7. Séquence des amorces utilisées pour CPEC**

<b><u>Séquence</u></b>	<b><u>Nom</u></b>
ccaattgtccatattgcatca	pKDaR1
tttataacctccttagagctcga	pKDbF1
gttttagagctagaaatagcaag	Lien F1
gtgctcagtatctctactga	Lien R1
gaagcgctaacgcttgctga	<i>lacZ</i>
aataatctcgccctatcca	<i>laplapX</i>
aataatctcgccctatcca	<i>ompV</i>
gcgtatatattgtgtgtg	<i>ctxAB</i>
agtcctttccttgagttg	Raava
gttttagacctagaaataggaagttaaaataaggctagt	pKD F0
ctgagcacttttagacctagaaatgttttagacctagaaataggaagttaaaa	pKD F1
gttttagacccgggaaattggggaagttaaaataaggctagtcggtatc	HorusV1 F0
cacttttagacccgggaaattggtttagacccgggaaattggggaagttaaaataag	Horus V1 F1
gcctgcagtctagactcgag	Vérification générique R1

## 2.6 Prédiction de repliement

Les ARN ou ADN ont été conçus *in silico* à l'aide d'outils bio-informatiques. Les structures secondaires ont été prédites par l'algorithme mFold, décrit par Zuker *et al.* (164). Les réglages de bases ont été conservés. Le programme est accessible à l'adresse : <http://18.219.218.27/>

Pour l'ADN, le formulaire « DNA Folding » a été utilisé de manière à observer les repliements aux extrémités 5' et 3' ainsi que pour estimer l'enthalpie des donneurs.

Pour l'ARN, le formulaire « RNA Folding » a été utilisée afin d'observer les structures émergeant des séquences construites via le logiciel de visualisation de séquence ADN Snapgene 5.3.2

## 2.7 Test de dilution en série par gouttes (Spot assay)

Le spot assay désigne une stratégie d'isolement de clones après une réaction de transformation ou un cycle No-Scar. Cette stratégie est plus adéquate lorsque l'organisme est *E. coli*, pour *V. cholerae* les résultats sont plus variables et l'efficacité de transformation est réduite considérablement.

Une dilution en série est réalisée dans des tubes Eppendorf 1.5mL, partant de  $10^0$  et allant de  $10^{-4}$  à  $10^{-6}$  selon la taille de la population estimée. 90  $\mu$ L de PBS/LB/SoC (le choix du diluant est ajusté selon l'expérience) est ajouté dans chaque tube Eppendorf. Après la période de récupération, la totalité du volume de réaction est placée dans le tube correspondant à la dilution  $10^0$ . Le tube est vortexé puis 10 $\mu$ L est transféré au tube  $10^{-1}$ . La manipulation est répétée jusqu'à la dernière dilution où les derniers 10 $\mu$ L sont jetés. Il est nécessaire de changer d'embout entre chaque dilution.

Une goutte de 10 $\mu$ L de chaque dilution est déposée sur milieu contenant les agents de sélection, bien que ce test ne puisse pas de déterminer le nombre exact d'UFC il permet d'estimer l'ordre de grandeur des UFC. Les milieux sont laissés entre ouvert dans une hotte de biosécurité pour sécher les gouttes puis sont scellés avec du parafilm pour être finalement incubés. Après incubation le nombre de colonies est estimé depuis la goutte la plus diluée qui contient des colonies tel que :

$$\text{UFC/mL} = \text{Nombre de colonies sur la dilution la plus faible} * 100 *$$

*Facteur de dilution*

Dans le cadre d'un cycle No-Scar on exprime l'efficacité de contre sélection s'exprime en différence de magnitude.

## **2.8 PCR sur colonies**

Afin de génotyper les clones obtenus après transformation (plasmide ou ADN donneur) les colonies sont traitées de la manière suivante :

1 colonie est dissociée complètement dans 25  $\mu$ L d'eau MilliQ dans un Eppendorf 1.5mL. Pour *V. cholerae* le tube Eppendorf est placé à 95°C pendant 5 à 10 minutes. Le tube est centrifugé par une centrifugeuse de paillasse avant et après l'étape de lyse par la chaleur. La température permet d'éclater les cellules et de rendre l'ADN libre en solution. 1 $\mu$ L de cette solution de lyse est utilisé comme *gabarit* pour une amplification PCR. Cette étape n'est pas nécessaire pour *E. coli*. Il est conseillé de conserver la première étape de lyse cellule (5 minutes à 98°C) dans un programme de PCR classique.

# Chapitre 3 : Résultats

## 3.1 Implémentation du protocole No-Scar dans *V. cholerae* A1552

Afin de créer nos propres souches mutantes de *V. cholerae*, nous avons décidé d'implémenter la plateforme No-Scar décrite par Reisch (101, 102) dans *V. cholerae* A1552. Cette plateforme offre des résultats séduisants chez *E. coli*. que nous n'avons pas été en mesure de reproduire chez *V. cholerae*.

Dans nos premiers essais, il nous a été impossible de transformer de manière séquentielle les plasmides pCas9-cr4 et pKDsgRNA, tous deux commandés chez Addgene (Watertown, États Unis) et extraits de *E. coli* DH5 $\alpha$ . Le protocole No-Scar requiert de transformer d'abord le pCas9, car il contient *TetR*. Sans *TetR* la répression de l'expression de la Cas9 et du sgARN est impossible.

La transformation du pCas9-cr4 dans A1552 échouait systématiquement. Il nous a été possible d'obtenir rarement quelques colonies isolées, mais ces dernières étaient soit des mutants spontanés résistants au chloramphénicol soit possédaient un phénotype macroscopique mucoïde qui est très différent du phénotype usuel de *V. cholerae*. Au microscope les bactéries présentaient une forme filamenteuse, typique de bactéries en réponse SOS.

En parallèle, nous avons créé le pKD *laplapX*. Le locus *laplapX* (VCA\_0812-13), encode pour deux leucyl-aminopeptidases Lap et LapX dont le rôle chez *V. cholerae* reste à déterminer (165). Après avoir généré le pKD LapLapX via CPEC (Gibson Assembly) et transformation subséquente dans *E. coli* DH5 $\alpha$ , nous avons alors transformé le plasmide pKD LapLapX dans *V. cholerae* A1552 sauvage avec succès.

Nous avons donc raisonné de la manière suivante : A1552 ne possède pas de systèmes CRISPR-Cas natifs, sans le sgARN apporté par le pKD, la Cas9 ne possède aucune cible pour médier une quelconque dégradation, aucune raison apparente justifie l'impossibilité de transformer le pCas9-cr4 dans *V. cholerae* A1552.

Notre laboratoire possède plusieurs autres souches de *V. cholerae*, comme V52, V6, et MO10. Ces souches se différencient les unes des autres notamment par différents événements d'acquisition de matériel génétique exogène. Ainsi, nous avons essayé de transformer le pCas-cr4 dans chacune de ces souches. V52 et V6 comme A1552 ne pouvaient pas maintenir le plasmide. En revanche la souche MO10, était capable d'accepter le plasmide avec toutefois une efficacité extrêmement faible. Si le plasmide était par la suite extrait de MO10 puis transformé dans A1552 il nous était finalement possible d'obtenir des colonies ne présentant pas de phénotype mucoïde (pas de filamentation en microscopie) et possédant effectivement le plasmide.

Une fois le pCas9-cr4 "installé" dans A1552, il nous fut possible de reprendre la méthode No-Scar selon le protocole de transformation séquentielle préconisé. Nous supposons que des systèmes de restrictions propres à A1552 n'existant pas chez MO10 reconnaissent une signature de méthylation (166, 167) lorsque le plasmide pCas9-cr4 provient directement de DH5 $\alpha$ .

### **3.2 Optimisation électroporation de plasmides**

Nous avons souhaité améliorer le protocole de transformation de plasmide dans *V. cholerae*. La littérature recommande de manière générale de cultiver 10mL de *V. cholerae* depuis 100uL d'une culture overnight (dilution 1 :100), puis de laisser la culture croître pendant 1h30-2h00

jusqu'à atteindre une  $DO_{600}$  de 0.5, signe que les cellules sont entrées dans leur phase de croissance exponentielle.

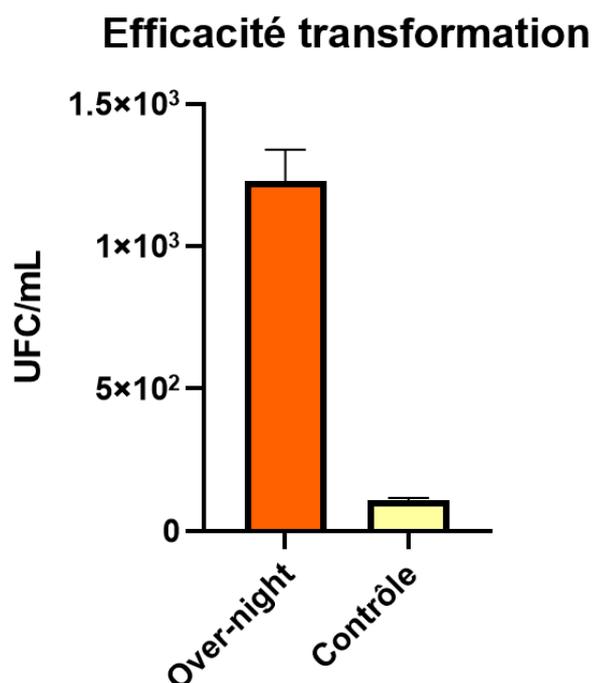
Par réaction de transformation 2,5mL de cellules sont utilisées afin d'obtenir environ  $10^7$  cellules(source). La culture est placée sur glace pendant 15 à 30 minutes. Une succession de lavages à froid avec une solution aqueuse contenant 10% de glycérol permet de rendre les cellules électro-compétentes. Un aliquot de 100uL de ces cellules compétentes est mélangé avec 100ng de plasmide purifié, puis électroporés puis incubé pendant au moins 1h00 à 37°C avec agitation, le temps que les cellules expriment le gène de résistance codée par le vecteur transformé.

Ce protocole permet de transformer certains plasmides avec une efficacité modérée (une dizaine de cellules au total sont obtenues). Cependant, entre nos mains ce protocole ne permet pas la transformation de certains plasmides pourtant rapportés comme étant compatible avec A1552 (comme pMMB66EH).

Étant donné que la compétence est liée à la détection du quorum et à la forte densité cellulaire chez *V. cholerae* (168) nous avons alors émis l'hypothèse qu'une électroporation avec des cellules à plus forte  $DO_{600}$  pourrait améliorer l'efficacité de transformation. Nous avons alors tenté de réaliser une transformation d'un plasmide, pBAD24 dans A1552 depuis une culture overnight, ayant récemment atteint la phase plateau ( $\approx 16h$ ).

Nous avons donc comparé le nombre d'unités formatrices de colonies (UFC) obtenues par transformation sur une culture en phase plateau (ON) et une culture en phase exponentielle (OD). Afin de comparer l'efficacité pour le même nombre initial de cellules données (environ  $10^7$  bactéries) nous avons utilisés 1.5mL de bouillon pour l'ON et 2.5 mL pour l'OD.

Nous avons utilisé le plasmide pBAD24 comme vecteur de contrôle, contenant le gène *bla* codant pour une bêta-lactamase. Après transformation les cellules ont récupéré une 1h00 dans du SOC puis étalée sur un milieu contenant 100 µg/mL de Carbénicilline. Nos résultats montrent que la culture ON permet d'obtenir plus de 1000 UFC à une culture en phase exponentielle (Figure 19). Ainsi, l'efficacité de transformation est grandement augmentée en électroporant des cellules en début de phase stationnaire de croissance chez *V. cholerae*. L'efficacité de transformation sur une ON diminue avec l'âge de la culture, il semble que les cellules doivent atteindre une DO<sub>600</sub> pour activer leur détection du quorum.



**Figure 13. Optimisation de l'efficacité de transformation**

Comparaison du nombre de colonies (UFC) obtenues sur milieu contenant 100 µg/mL de Carbénicilline entre le protocole de transformation « classique » pour *Vibrio cholerae* A1552 via la transformation du plasmide pBAD24. Over-night concerne la transformation sur une culture ayant atteint la phase plateau, le contrôle concerne la transformation sur une culture en phase de croissance exponentielle).

### 3.3 Création d'un mutant *lacZ* de *V. cholerae* A1552 par le système No-Scar

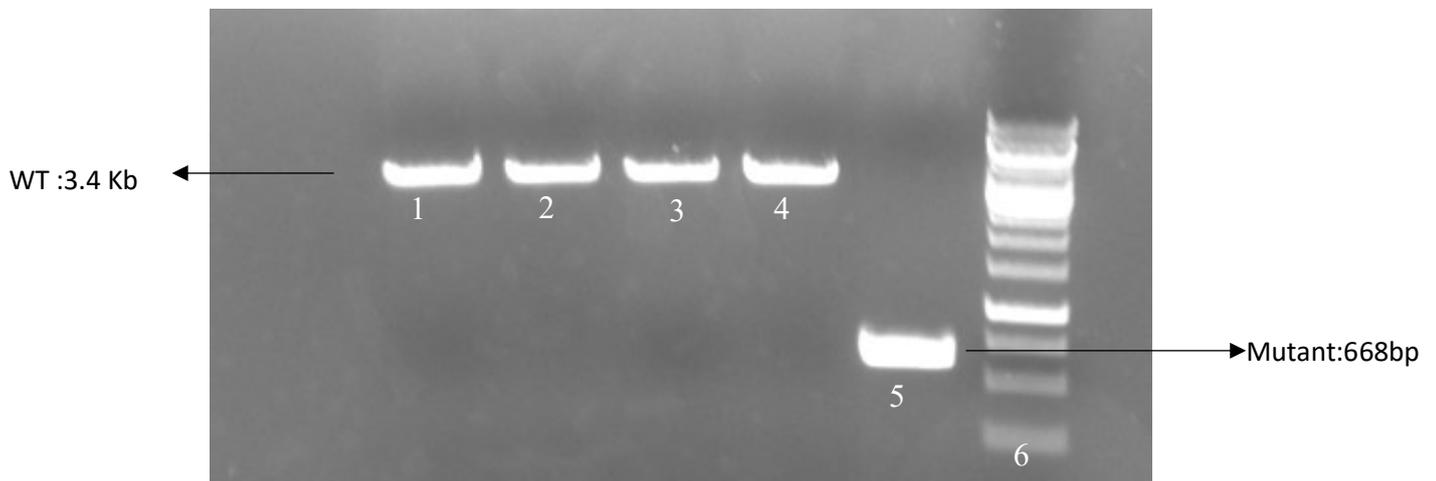
La littérature rapporte que l'efficacité de l'opéron Red est 1000 fois inférieure chez choléra comparé à *E. coli*. L'opéron Red est compatible à la fois avec de l'ADN double brin (dsDNA) et simple brin (ssDNA). Nous avons voulu observer quel type d'ADN présentait la meilleure efficacité de recombinaison apparente par la plateforme No-Scar en ciblant le gène *lacZ* de *V. cholerae* A1552. Les ADN donneurs portaient une modification phosphorothioates en 5' afin de le protéger de la digestion par les exonucléases de *V. cholerae*.

La mutation *lacZ* est aisément visible si le milieu contient du X-Gal (composé chromogène devenant bleu s'il est dégradé) et de l'IPTG (pour l'induction de l'opéron *lacZ*). Les colonies blanches sont analysées par PCR pour confirmer le génotype.

Brièvement, nous avons générés le pKD\_ *lacZ* par CPEC qui cible le gène *lacZ*. Après un cycle No-Scar aucune différence significative du nombre d'UFC n'a été observée entre le contrôle (ADN donneur ne supprimant pas la cible) et la réaction (cellules ayant reçue l'ADN donneur créant la mutation supprimant la cible). De plus, le temps de croissance d'une souche de *V. cholerae* contenant à la fois le pCas9-cr4 et le pKD a 30°C est considérablement rallongée par rapport à la croissance de *V. cholerae* à 30°C contenant uniquement le pKD.

Il nous a été possible d'isoler une colonie blanche. Toutefois, l'efficacité de la contre sélection était faible voire nulle, alors que l'efficacité de l'opéron Red est dans les normes rapportées par la littérature (Cf. 4.1 Implémentation de la plateforme NoScar chez *V. cholerae*). Le génotype de colonies ayant un phénotype bleu clair à blanc fut vérifié par amplification PCR puis par séquençage (Figure 14).

Le génotype de plusieurs colonies a été vérifié par PCR Colonie (Figure 14).



**Figure 14. Génotypage PCR du mutant A1552 *lacZ* obtenu**

Génotypage PCR de 4 colonies prélevées au hasard obtenues après un cycle No-Scar sur A1552 pour créer le mutant *lacZ*. Le génotypage a été réalisé sur une souche sauvage (Puit 1), sur 3 colonies présentant une couleur bleu pâle (puits 2 à 4) et une seule colonie totalement blanche (puit 5), l'échelle moléculaire (puit 6) est Gene-O-ruler 1Kb de Thermofisher.

L'efficacité de recombinaison était très faible. Nous trouvions en moyenne 1 colonie blanche pour 1000 cellules escapers bleues pour un donneur double brin, pour un ADN donneur simple brin il nous était possible de trouver 3 colonies blanches pour 1000 cellules escapers.

### 3.4 Stabilité du système No-Scar

Nous avons voulu évaluer si le système No-Scar était capable de contre sélection après une culture prolongée chez *V. cholerae*. Nous avons découvert qu'il était possible d'estimer si le système No-Scar était encore opérationnel pour une colonie donnée. Cette estimation est réalisée par ensemencement de culture liquide en limitant la taille de l'inoculum, au lieu de dissocier les colonies dans le bouillon, afin de limiter au maximum l'inoculum.

Après 16 à 20 heures la turbidité des cultures fut mesurée (Tableau 9 et 10). Le critère de réussite était l'absence de croissance dans le milieu contenant de la tétracycline anhydre. (aTc)

Nous avons donc comparé deux conditions :

1) Figure 20 : Une condition concernait des colonies contenant le pCas9-cr4 et fraîchement transformées avec le pKD LapLapX, la moitié de la transformation avait été étalée sur un milieu ne contenant pas d'aTc et l'autre contenant 200 ng/mL d'aTc.

2) Figure 21 : La seconde condition concernait des colonies obtenues après repiquage d'un stock glycérol gardé à -80 °C d'une souche de A1552 contenant les deux plasmides du système No-Scar (pCas9-cr4 et pKD LapLapX). La souche fut repiquée sur milieu avec ou sans aTc. Aucune croissance n'était observée sur la boîte contenant de l'aTc, nous laissant donc songer que le système était stable.

Tel quel, ces résultats donnaient l'impression que le système était stable pour plusieurs passages. Cependant nos résultats de transformations laissaient songer une narrative différente, l'efficacité de contre sélection semble s'effondrer après un certain temps. De plus, nous avons

contacté l'auteur directement pour le questionner sur la stabilité du système. Ce dernier nous renseigna que le système devait être utilisé au plus frais jamais à partir d'une overnight, sans quoi la proportion de cellules échappant à la contre sélection (*escapers*) devenait trop importante.

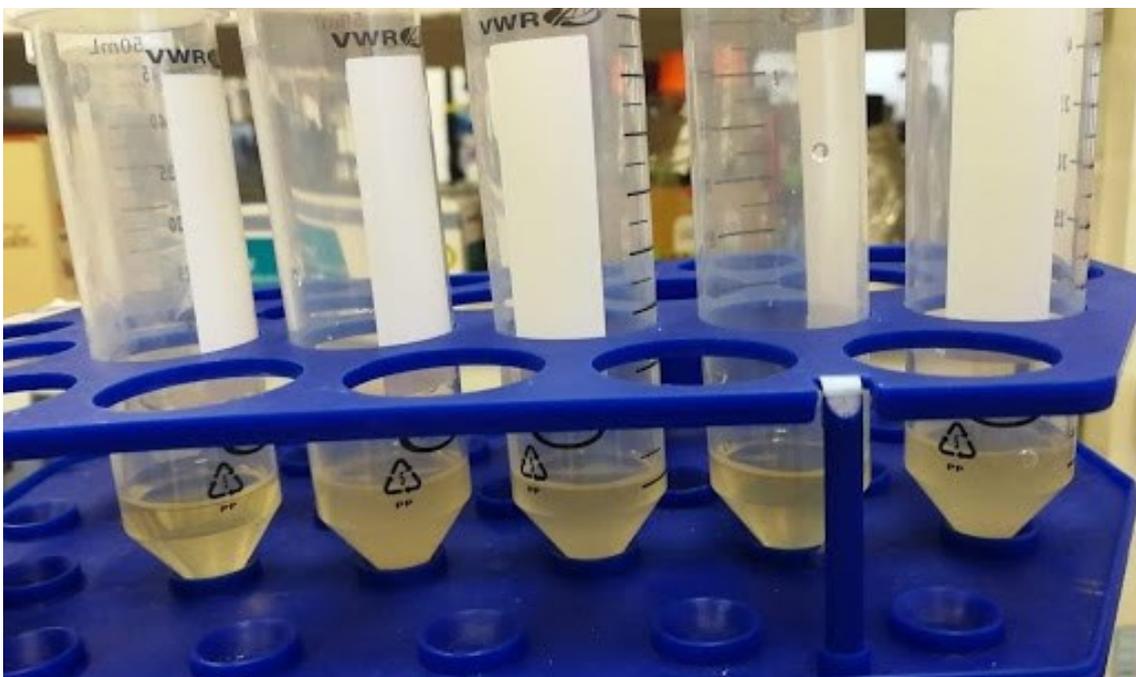
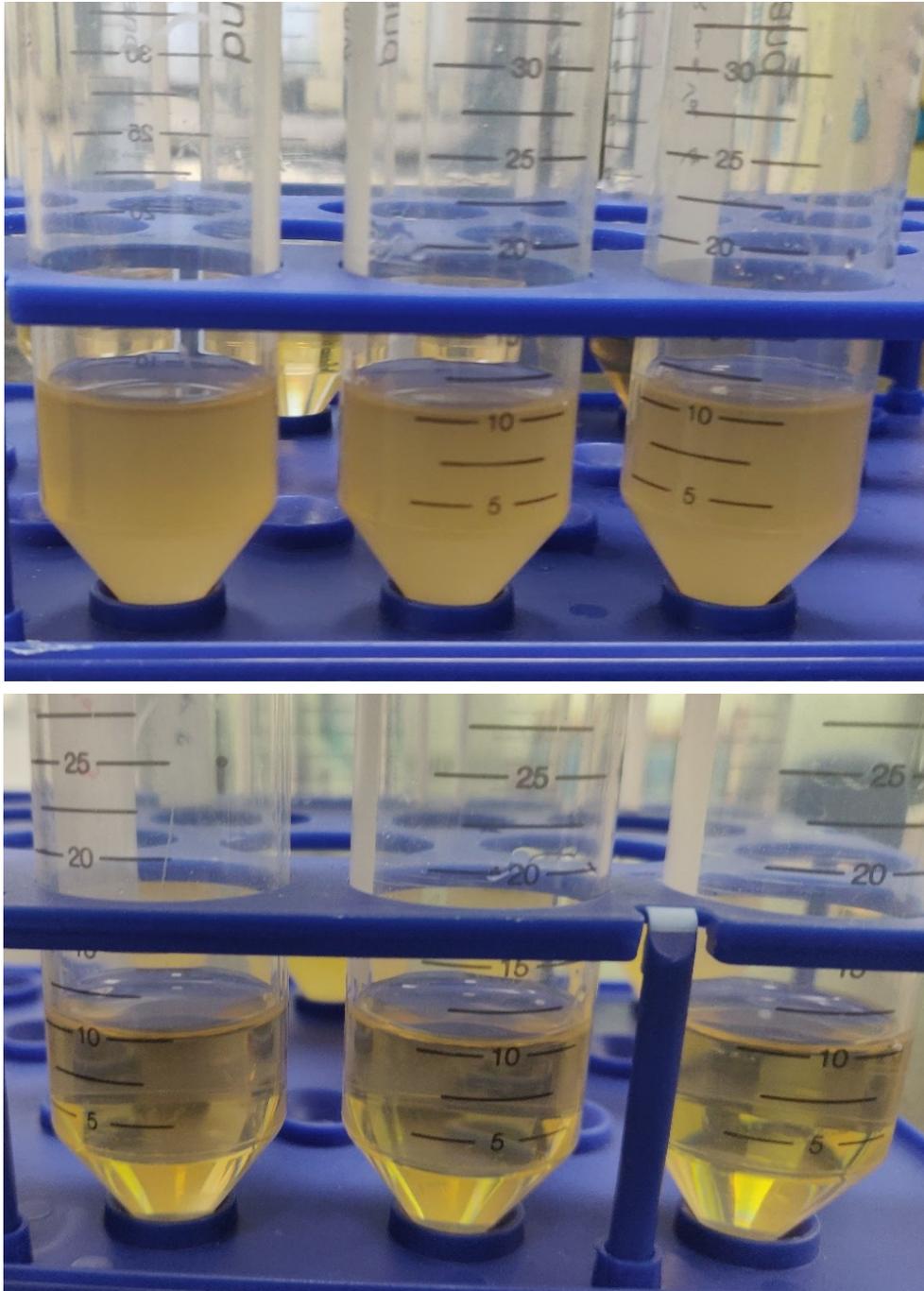


Figure 15. Stabilité du système No-Scar depuis un stock glycérol  $-80^{\circ}\text{C}$ .

Photographie des cultures overnight. Une souche contenant les deux plasmides du système No-Scar a été repiquée sur milieu sans aTc. Les colonies obtenues ont été cultivées en présence (A) ou en absence (B) d'aTc. La turbidité des bouillons fut observée après 16h.



**Figure 16. Stabilité du système No-Scar depuis une transformation fraîche.**

Photographie des cultures overnight. Une souche fraîchement transformée les deux plasmides du système No-Scar a été étalée sur milieu LB sans aTc. Les colonies obtenues ont été cultivées en présence (A) ou en absence (B) d'aTc. La turbidité des bouillons fut observée après 16h.

**Tableau 8. Stabilité du système du système No-Scar depuis un stock glycérol -80°C (Valeurs DO Figure 21).**

aTc -	aTc+
1.89	0.3
1.88	1.6
1.64	1.13
1.85	0.6
2	1.4

**Tableau 9. Stabilité du système No-Scar depuis une transformation fraîche. (Valeurs DO Figure 22).**

aTc-	aTc+
1.80	0.02
1.78	0.01
2	0.01

### 3.5 Création de la plateforme Horus (HOMologous Recombinaison Using SsDNA)

#### 3.5.1 Mécanisme de contre sélection de Horus

Au vu des résultats obtenus depuis l'implémentation du protocole No-Scar nous avons voulu créer notre propre plateforme de contre sélection CRISPR. Il apparaît que la coexistence du pKD et du pCas9 crée une pression de sélection par stress génotoxique sur le chromosome (*Cf. 3.7 Escapers*). L'auteur de la méthode précise que le promoteur pLTetO1, bien que capable d'une forte répression est tout de même victime de fuite transcriptionnelle.

Afin d'adresser ce problème nous avons conçu une stratégie de contre sélection alternative :

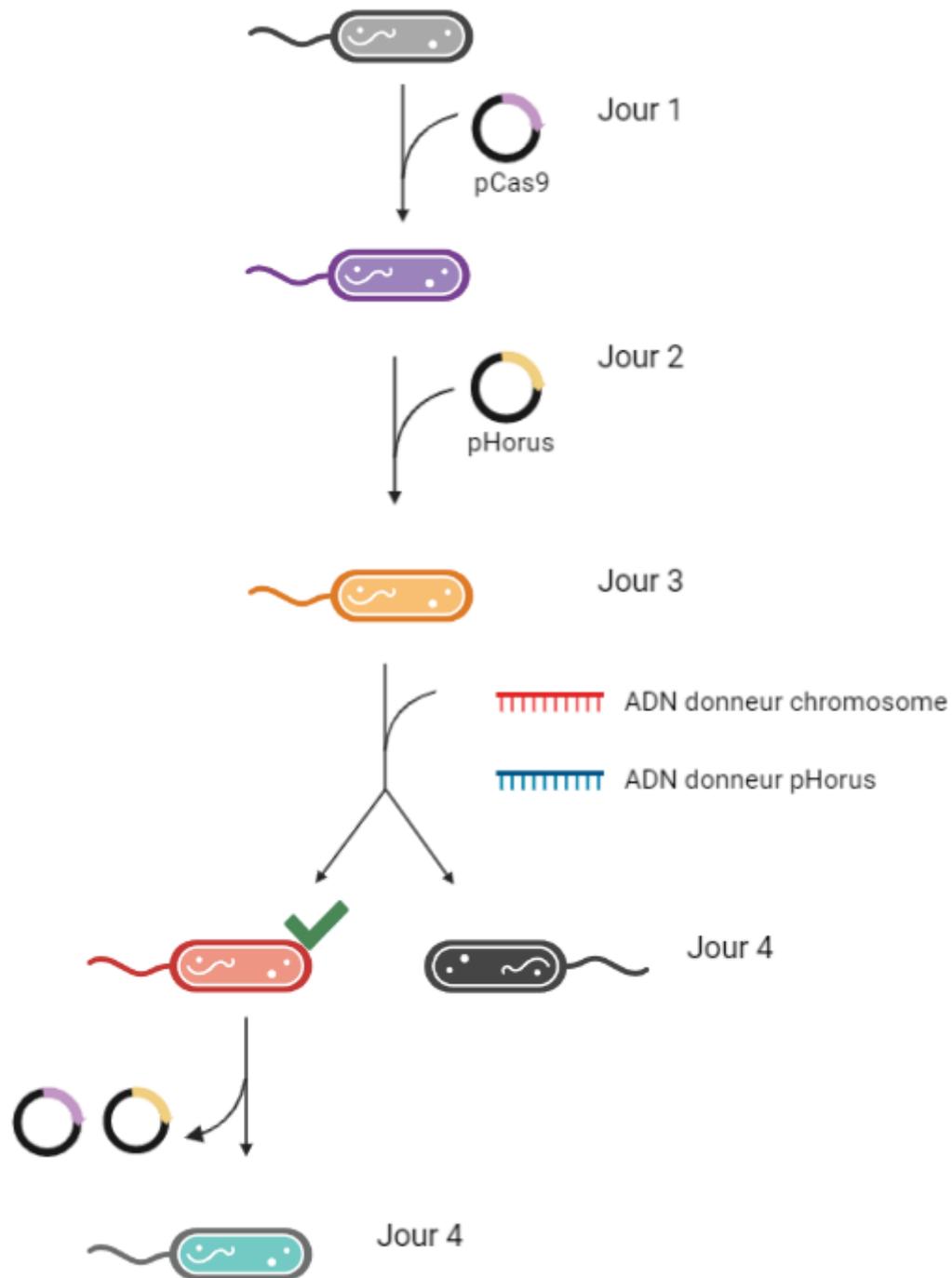
Au lieu que la contre sélection s'opère sur le chromosome, notre plateforme fonctionne en deux temps. Dans son état initial (Figure 17) la contre sélection s'effectue contre le pKD. Sans recombinaison, la cible reconnaît sa propre séquence (cible suicide, en noir sur la Figure 17). La dégradation de cette cible provoque une blessure double brin dans le plasmide et résulte donc par un phénotype Spectinomycine sensible (Spectinomycine S).

La plateforme tire parti du maintien de l'hérédité de certains plasmides même en absence de sélection (*Cf. 3.5.3 Hérité du pKD sans sélection*), ainsi si l'agent de sélection est retiré du milieu pendant la culture de la souche, cette dernière ne subit plus de pression de sélection par la fuite d'expression du promoteur pLTetO1. L'agent de sélection n'est appliqué qu'à la fin, après la période de récupération. Une fois le pCas9-cr4 et le pHorus transformés séquentiellement la souche est mise en croissance dans un milieu riche (sans glucose pour éviter une interférence avec le promoteur AraBD) (Figure 17).

Une fois que la souche atteint une  $DO_{600}$  d'environ 0.5 elle est électroporée en présence de deux types d'ADN donneurs simple brins : Un premier donneur vient créer la mutation souhaitée sur le ou les chromosomes de la souche d'intérêt, le deuxième donneur vient changer la cible suicide sur le pKD (Figure 23C). Le deuxième ADN donneur change deux séquences sur la cible suicide, l'espaceur de 20 paires de bases qui code pour une nouvelle cible, ainsi qu'un motif de 5 paires de bases, dénommé la région plastique (Cf. 4.2, Création de la plateforme Horus). Cette région plastique est vitale pour le fonctionnement de notre stratégie de contre sélection.

Après recombinaison la cible suicide perd sa capacité d'auto-reconnaissance et gagne la capacité de cibler un nouvel espaceur qui permettra de faire la contre sélection sur le chromosome sauvage. La région plastique elle sert à protéger le pKD recombiné des autres pKD co-existant dans la cellule, mais n'ayant pas recombiné. Après électroporation les cellules récupèrent pendant plusieurs heures afin de ségréguer les allèles recombinés puis sont étalées sur milieu contenant de la spectinomycine (sélection du pKD) et de l'aTc (activation du promoteur pLTet-O1).

Via ce modèle, seule la sous-population ayant recombiné sur le chromosome et au moins sur un pKD est capable de survivre à la contre sélection. Cette plateforme permet de réduire considérablement le temps qu'une souche ayant reçu le pCas9-cr4 et le pKD prend à atteindre une  $DO_{600}$  de 0,5. Notre plateforme permet aussi de ne plus générer le pKD par CPEC.



**Figure 17. Schématisation du protocole par le pKD Horus**

Une souche sauvage est transformée séquentiellement avec le pCas9-cr4 et le pKD Horus. Au jour 3 les cellules portant les deux plasmides sont mises en bouillon jusqu'à  $DO_{600}$  de 0.5, l'opéron Ara est induit puis les ADN donneurs sont électroporés. Deux ADN donneurs sont transformés, un ADN donneur recombine sur le pKD, et le second ADN donneur créer une mutation sur le chromosome.

Après une période de récupération, les cellules sont étalées sur un milieu contenant de l'aTc et antibiotique. Le mécanisme de contre sélection cible à la fois le chromosome et le pCas9-cr4. Au jour 4 les colonies obtenues sont repiquées à 37°C sur milieu sans sélection et sur milieu contenant l'agent de sélection du pCas9-cr4. Les colonies ne poussant pas sur milieu avec sélection sont réputées mutantes et sans plasmide.

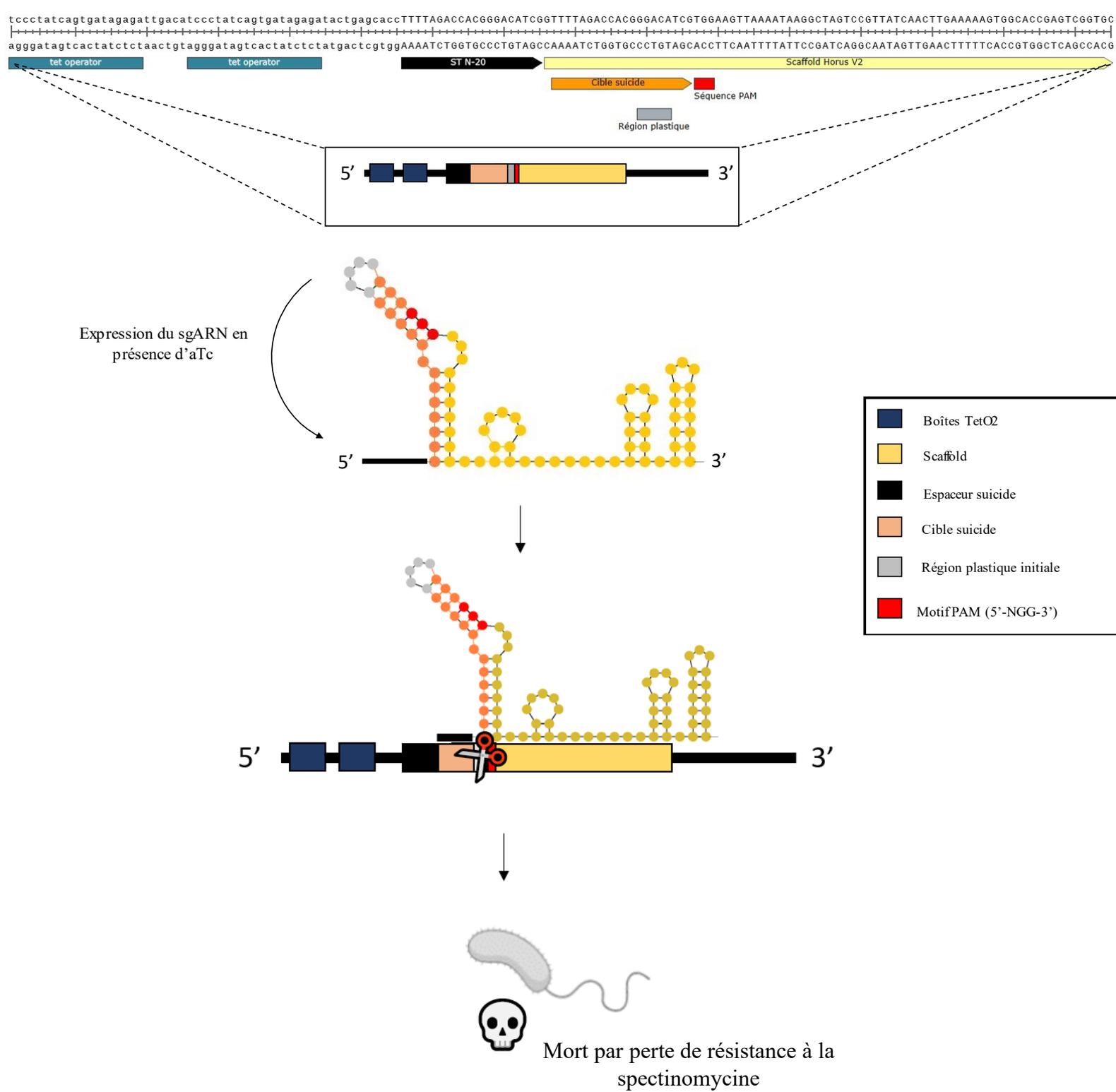
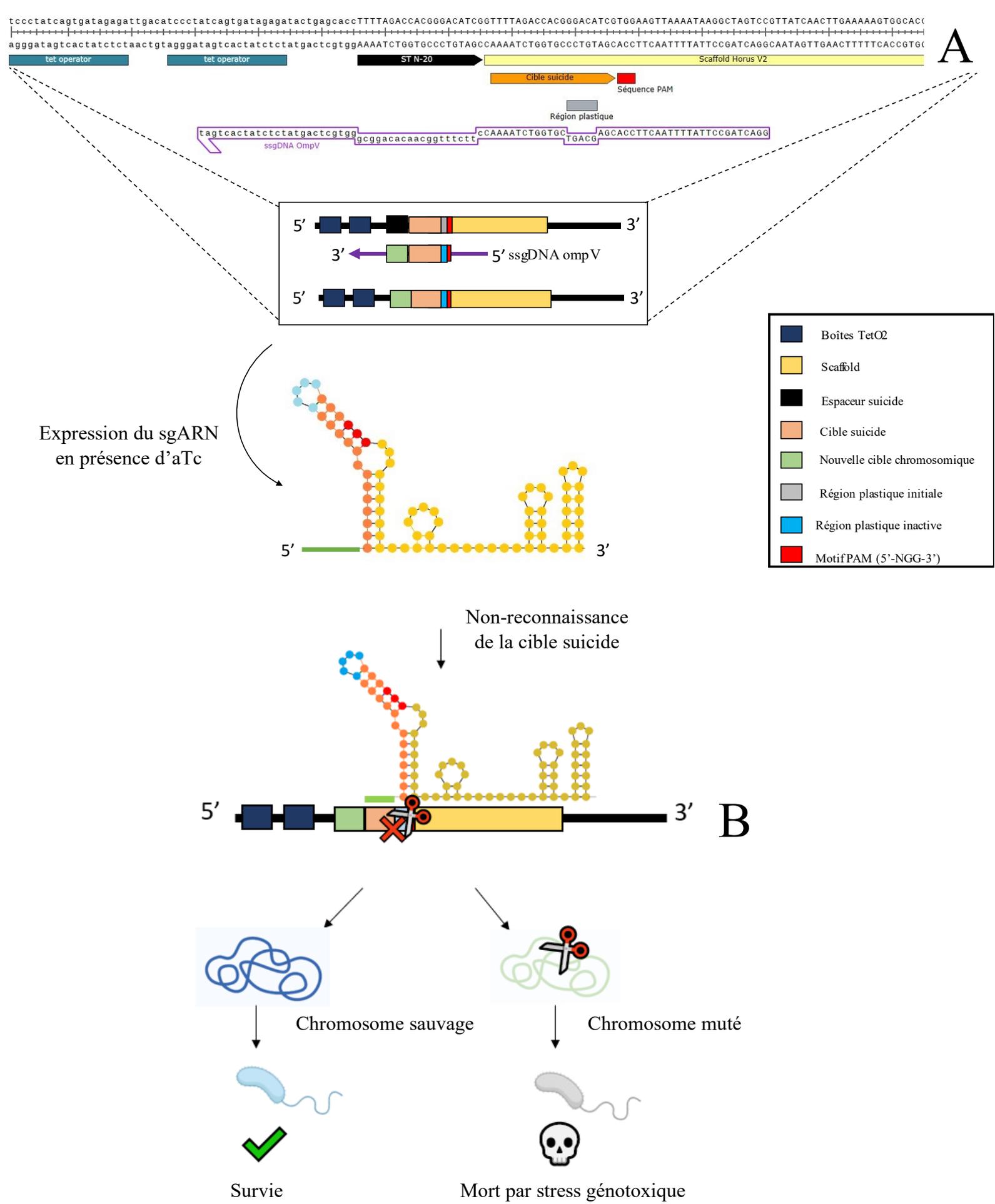


Figure 18. État initial de contre sélection par locus suicide

B. Le locus suicide est constitué de plusieurs éléments clés : En bleu, les deux boîtes TetO2 qui permettent la répression du locus en absence d'aTc, le ST-N20, en noir (espaceur suicide) qui reconnaît la cible suicide en orange. La reconnaissance repose sur la présence du motif 5'- GACAT-3' en amont du motif PAM (en rouge). Après induction par aTc la reconnaissance par le sgARN néo transcrit entraîne une blessure double brin dans le pKD qui entraîne la perte de résistance à la spectinomycine.



### Figure 19. État recombiné du locus Horus

A. Après recombinaison par un ADN donneur simple brin, la séquence du locus est changée. L'espaceur suicide est remplacé par un espaceur qui reconnaît une cible chromosomique (en vert). Une région de 5 paires de bases est aussi changée (région plastique) qui est en amont du motif PAM (en rouge).

B. Les mutations introduites par le donneur ADN simple brin empêchent la reconnaissance de la cible suicide par les pKD co-existant dans la même cellule qui n'ont pas recombinés. En parallèle, la présence d'un second ADN donneur simple brin vient introduire une mutation dans le chromosome. Après induction par aTc le sgARN néo transcrit cherche la nouvelle cible sur le chromosome, si le chromosome n'a pas recombiné il est coupé au niveau de la nouvelle cible.

### 3.5.2 Résistance à la spectinomycine par *aadA*

*aadA* est une adénylyl-transférase. Ce gène est réputé être capable d'inactiver deux substrats : la spectinomycine et la streptomycine. La spectinomycine est bio statique et la streptomycine est biocide. Les deux antibiotiques visent la même cible ribosomale. De fait, dans un cadre de contre sélection la Streptomycine est plus attirante par son action létale.

Pour les essais suivants nous avons réalisé que nous pouvions moduler la contre sélection en fonction de la concentration en antibiotique.

Nous avons testé plusieurs concentrations de spectinomycine et streptomycine. Entre nos mains le pKD tolère moins la streptomycine que la spectinomycine. À 50 µg/mL la Spectinomycine et Streptomycine produisent des colonies jusqu'à la dilution  $10^{-4}$  dans un *spot assay*, bien que dans le cas de la streptomycine ce résultat est variable. À 120 µg/mL la Spectinomycine produit marginalement la même quantité d'UFC/mL, pour la streptomycine le nombre d'UFC est réduit est de  $10^{-1}$  UFC/mL à cette concentration.

Nous avons donc décidé pour les essais suivants d'utiliser la spectinomycine, malgré sa qualité d'agent bio statique cet antibiotique était mieux géré que la streptomycine. Pour les essais suivants, le plasmide pRaava était utilisé comme contrôle. Les concentrations pour la contre sélection du plasmide fut alors placée à 120 µg/mL de spectinomycine

### 3.5.3 Hérité du pKD sans sélection

Nous avons voulu savoir si le pKD était stable (i.e. capable d'être transmis aux cellules filles) en absence de sélection. Ce point était crucial pour la modification du système No-Scar. En effet, nous voulions créer une plateforme de contre sélection dirigée contre le pKD. En déplaçant la contre sélection du chromosome au plasmide sans sélection par antibiotique nous émettions l'hypothèse que la même pression de sélection amenant lentement le système No-Scar d'un état opérationnel à un état escaper serait grandement réduite. En absence de sélection du pKD par la spectinomycine l'évènement « bruit de fond » de dégradation du plasmide par une fuite transcriptionnelle ne présentait aucune cytotoxicité apparente.

Pour tester l'hérité du plasmide en absence de sélection, nous avons procédé tels que :

Une souche de DH5 $\alpha$  contenant un pKD fut cultivée en présence et absence de Spectinomycine, après une overnight un *spot-assay* fut réalisé jusqu'à la dilution 10<sup>-5</sup> sur milieu sans spectinomycine et avec spectinomycine.

Le pKD est extrêmement stable même sans son agent de sélection. Après 16h d'incubation les spots assay ont poussés jusqu'à la dilution 10<sup>-5</sup> sur milieu sans sélection comme sur milieu contenant de la Spectinomycine. Les spots présentaient la même confluence, le nombre d'UFC précis fut impossible à déterminer, car la dilution la plus faible était confluyente à 100% (*Cf. Annexe*).

### 3.5.4 Fragment Anubis

Nous avons voulu observer s'il était possible d'apporter une nouvelle cible (N-20) non pas par CPEC, mais par recombinaison sur le pKD via l'opéron Red.

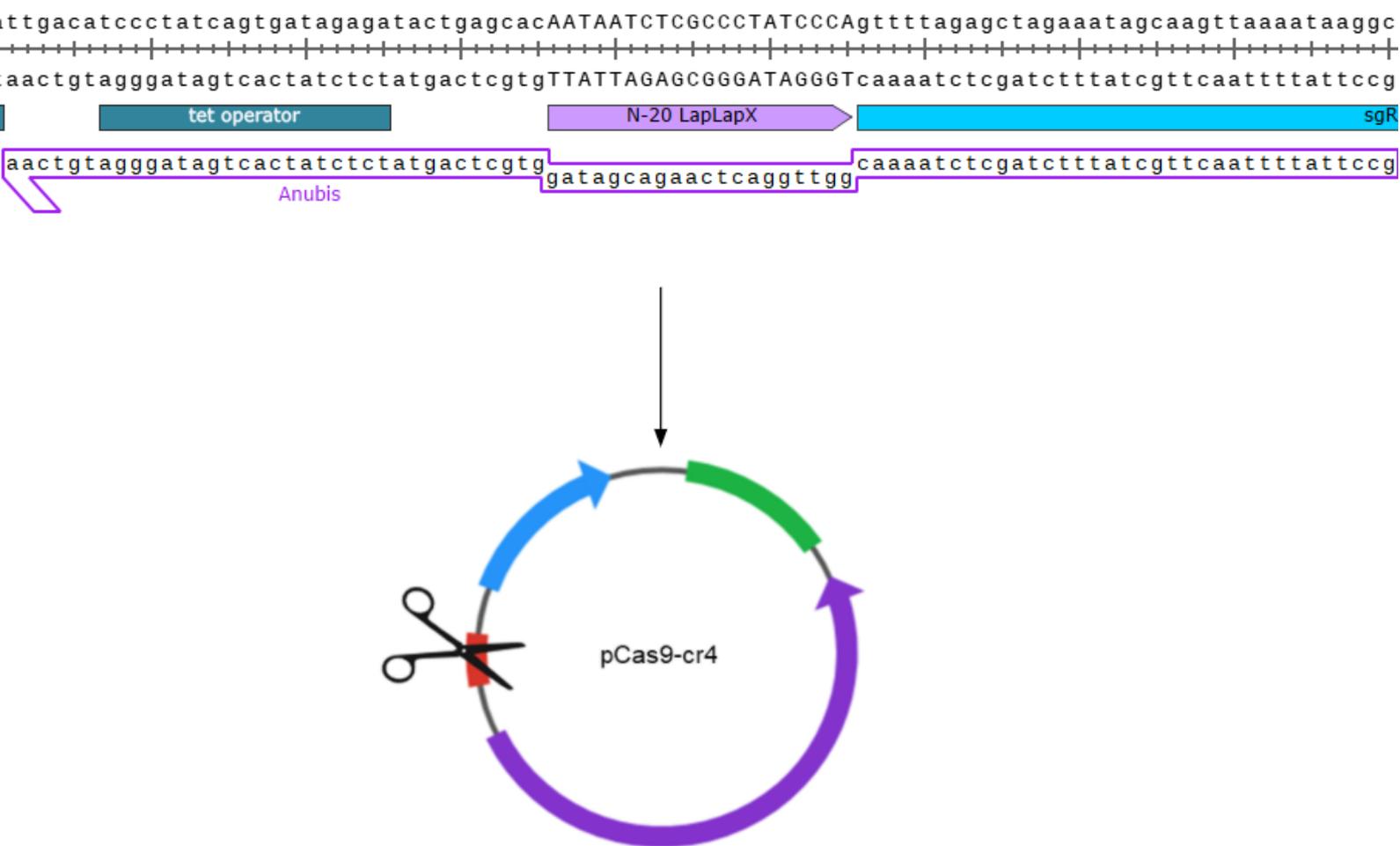
Nous avons conçu un ADN donneur ssDNA qui venait se recombiner sur la mini-librairie CRISPR du pKD (Figure 26). La mini librairie est située sur le réplichoire +, le brin tardif est donc le brin antisens (3'→5').

D'une longueur de 90 paires de bases cet ADN donneur possédait 35 paires de bases d'homologie avec le promoteur pLTetO1 et 35 paires de bases avec le Scaffold (tracrARN + crARN) en sus d'une région de 20 paires de bases contenant un espaceur ciblant l'origine de réplication du pCas9-cr4. Ce fragment fut dénommé Anubis.

Nous avons donc utilisé une souche de A1552 contenant le pKD LapLapX et le pCas9-cr4. L'ADN donneur fut transformé à raison d'une concentration de 5µM. Les cellules ont récupéré une heure dans du milieu SoC préchauffé puis 1µL d'aTc fut rajouté au bout d'une heure pour induire la pCas9-cr4 et le pKD pour une autre heure. Les cellules furent alors étalées sur un milieu contenant 50µg/mL de Spectinomycine et 200ng/mL d'aTc. Les colonies obtenues furent repiquées sur un milieu contenant 10 µg/mL de chloramphénicol pour conclure sur le recyclage du pCas9-cr4.

À notre surprise 100% des colonies repiquées sur milieu contenant du chloramphénicol étaient incapables de croître. Paradoxalement, les colonies obtenues depuis la transformation d'Anubis, si testées par amplification PCR via les amorces pCas9-cr4 F1/R1 ciblant le pCas9-cr4 retournait la bande escomptée de 500pb, ces mêmes colonies étaient tout de même incapables de croître sur milieu du chloramphénicol, si la même PCR était effectuée sur les souches repiquées sans chloramphénicol aucune bande n'était retournée, preuve de la perte du pCas9-cr4. Lorsque Anubis avait été testé pour une seconde fois dans une souche de A1552Δ*lacZ* obtenue précédemment par le système No-Scar, contenant toujours le pKD *lacZ*

et le pCas9-cr4 l'efficacité de recombinaison retombait à 17.8% (5 de 28 colonies choisies au hasard n'ont pas poussé lorsque repiquées sur un milieu contenant 10µg/mL de chloramphénicol). Encore une fois, l'amplification PCR de ces colonies retournait une bande correspondant au pCas9-cr4 qui disparaissait après que les colonies étaient repiquées.



**Figure 20. Action du donneur ADN simple brin Anubis sur le pKD LapLapX**

Anubis, encadré en violet, est un ADN donneur simple brin capable de venir se recombinaison sur le pKD. Il est complémentaire au brin sens (+), car le locus qu'il vise se situe dans le réplique de droite. En se recombinaison Anubis est capable de changer la cible d'un pKD sans que cette dernière ne soit créée par CPEC.

## 3.6 Création d'une cible suicide

### 3.6.1 pRaava

Dans nos premiers essais nous avons simplement créer une version suicide du pKD. Pour ce faire nous avons cloné un espaceur codant pour l'origine de réplication du plasmide (ori SC101). Dans ce modèle, en absence de recombinaison sur le pKD, ce dernier va créer un phénotype Spectinomycine S (sensible) par coupure double brin a son origine de réplication (Figure 20).

Le plasmide portant cette construction fût dénommé « pRaava ». pRaava fut construit par CPEC et vérifié par génotypage PCR (Figure 21). Afin de vérifier l'intégrité de la cible, nous avons utilisé une amorce ayant la même séquence que l'espaceur de pRaava. De fait, comme pRaava cible sa propre origine de réplication, cette séquence est présente en double au sein du plasmide pRaava (1 fois au niveau de l'origine de réplication et 1 fois au niveau de la séquence guide), nous attendions donc deux bandes, une bande de 3.6Kb et une bande de 2.2Kb.

La construction validée par PCR fût transformée dans une souche de A1552 contenant le pCas9-cr4. Après une période de récupération d'1h dans du SoC, la réaction de transformation fût est divisée en deux, une moitié étant étalée sur milieu contenant de l'aTc et l'autre sans aTc. Typiquement, nous obtenions plusieurs colonies sur le milieu ne contenant pas d'aTc pour 2 ou moins colonies sur un milieu contenant de l'aTc (Tableau 11). Ce résultat laisse entendre qu'une activité de contre sélection était présente.

La stabilité du système fut estimée par mise en culture des colonies portant le pKD Raava et le pCas9-cr4, en présence de Spectinomycine (50µg/mL) et d'aTc (200ng/mL). Aucune croissance ne fut notée dans les bouillons contenant de l'aTc signe que le système était stable sur une overnight.

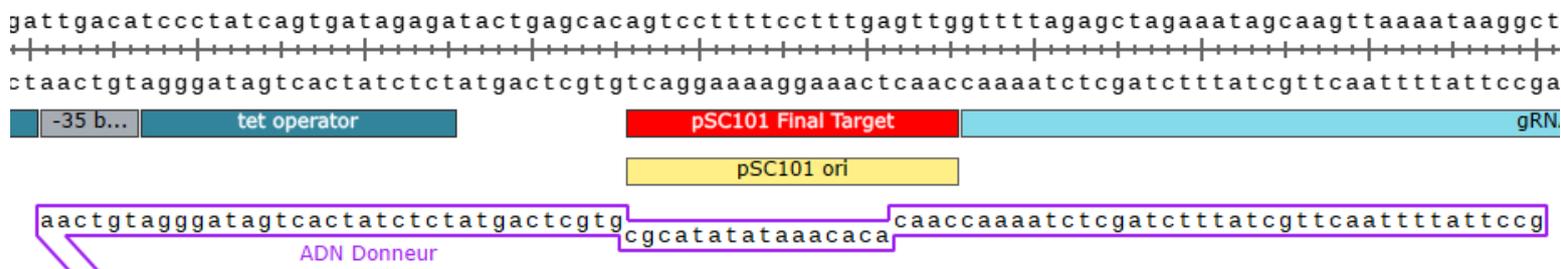
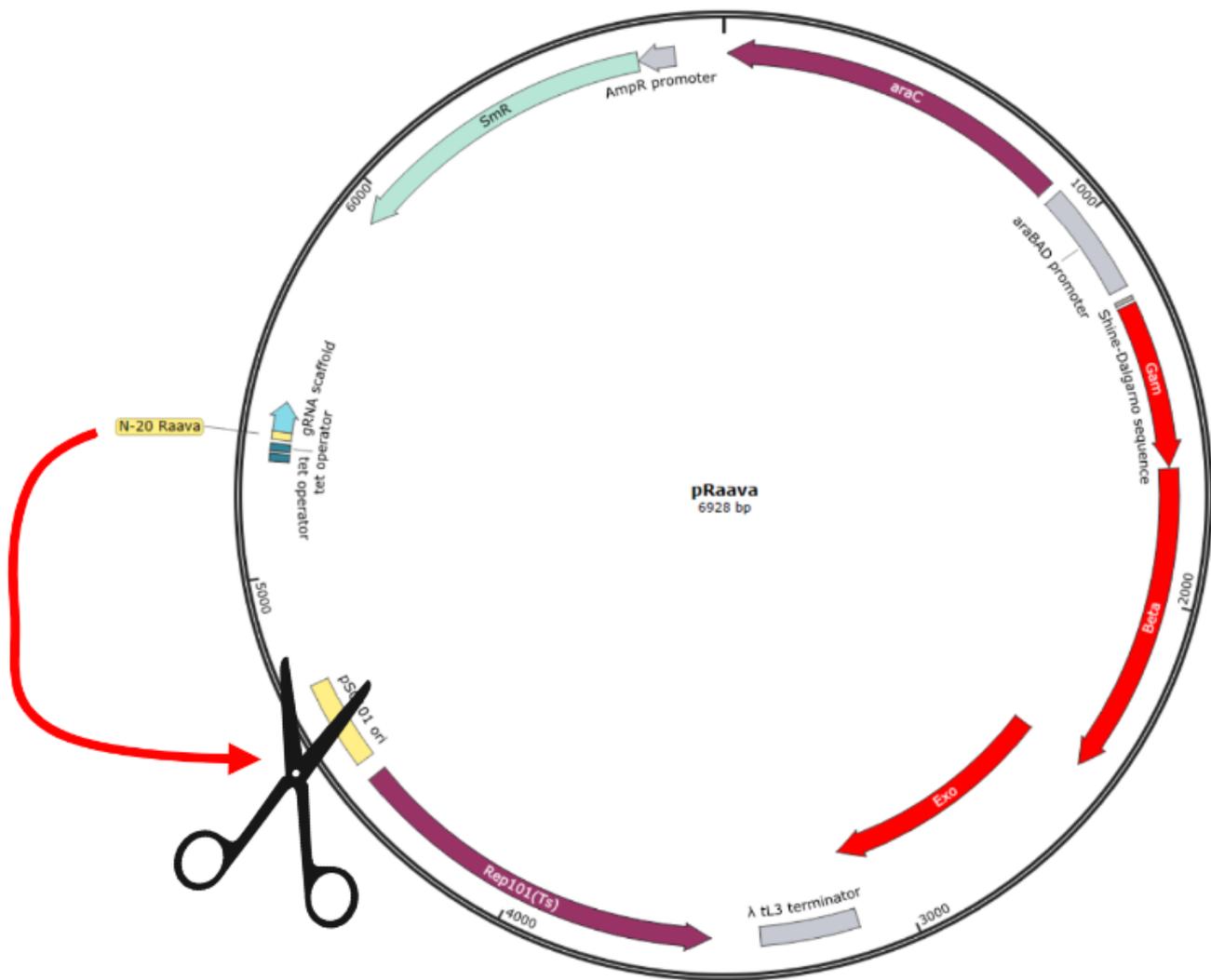
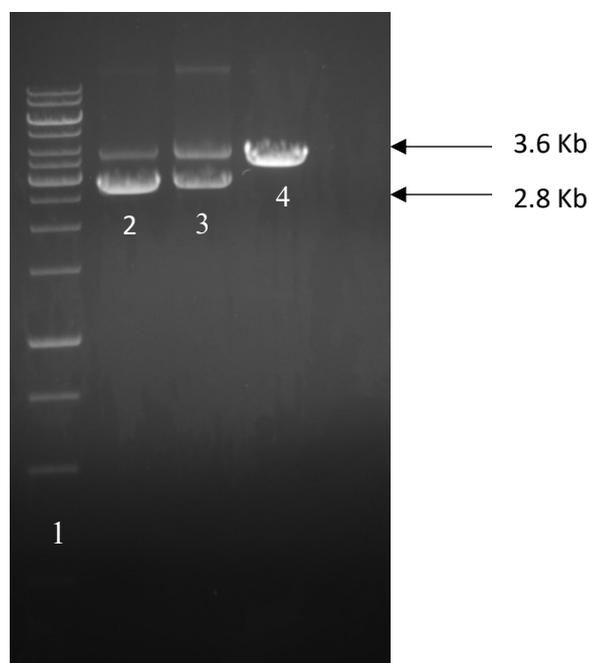


Figure 21. Mécanisme d'action de pRaava

Le plasmide pRaava est capable de médier sa propre dégradation en ciblant le clivage de sa propre origine de réplication (flèche rouge). Il encode pour une cible visant *pSC101* (en rouge dans le schéma de séquence). Dans ce concept un ADN donneur (séquence encadrée en violet), vient remplacer la cible suicide par une nouvelle cible visant le chromosome. pRaava **perd** alors la capacité d'auto-dégradation et **gagne** la capacité de viser une nouvelle cible.



**Figure 22 : Génotypage du plasmide pRaava après CPEC**

Génotypage du plasmide pRaava obtenu dans DH5Aα par CPEC. Nous nous attendions donc à obtenir deux bandes lors de la PCR de vérification, une bande de 3.6Kb correspondant à l'origine de réplication et une seconde bande de 2.8Kb preuve de la présence de la séquence cible. 1 : Échelle moléculaire, 2 et 3 : Deux colonies de DH5A obtenues après transformation de la CPEC pRaava. 3 : Plasmide pKD-p15A utilisé comme référence.

**Tableau 10. Efficacité de contre sélection du pRaava**

aTc-	aTc+
15	0
23	2
36	1

La capacité de contre sélection une fois établie nous avons tenté de créer les 3 mutants, *ompV*, *ctxAB*, *laplapX* en utilisant la souche contenant la pCas9-cr4 et le pKD Raava. Le cycle No-Scar restait principalement inchangé à l'exception de la présence d'un second ADN donneur lors de l'étape de transformation pour une concentration finale en ADN donneur de 10 $\mu$ M. En sus de l'ADN donneur permettant de créer la mutation souhaitée dans le locus cible, un second ADN allait se recombiner sur le pRaava pour remplacer la cible suicide pour une nouvelle cible visant le chromosome.

De manière similaire à nos essais via le protocole No-Scar de base, des bandes correspondant au génome sauvage étaient retournées en plus de bandes correspondant à la cible mutée au sein d'une même colonie. En outre des bandes sauvages, il nous était impossible de trouver des plasmides recombinés par amplification PCR (Figure 22). Paradoxalement, les conditions témoins comptaient produisaient plus d'UFC que les essais avec le bon ADN donneur. Ceci suggérait une coexistence, dans une cellule du pKD recombinaison ayant intégré la nouvelle cible et du pKD Raava n'ayant pas recombinaison. Ce dernier étant capable de cliver sa propre origine de réplication, mais aussi celle du pKD recombinaison, tous les pKD étaient éliminés lors de l'activation de la Cas9 et du guide par l'ajout d'aTc (Cf. 4.2 *Création de la plateforme Horus*).



**Figure 23. pRaava non recombinaison**

Le même patron de bande est observé sur les plasmides obtenus après un cycle No-Scar que le plasmide pRaava de référence, la présence de la bande de 2.8Kb pour les puits 2 à 8 indique l'absence de recombinaison. Puit 1 = pKD-p15A.

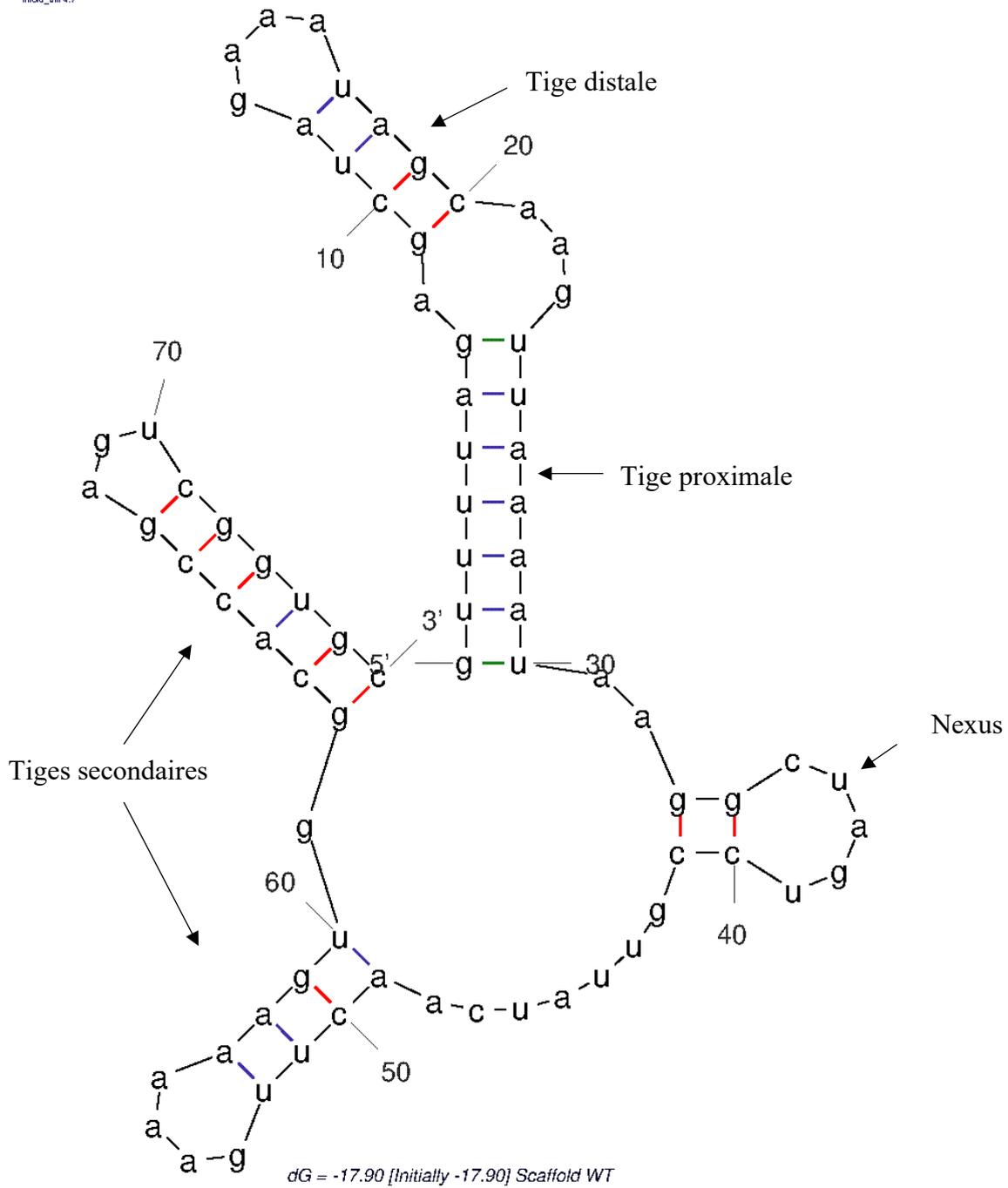
### 3.6.2 pKD<sup>2</sup>

L'impossibilité de trouver un plasmide pRaava recombinant suggère que le plasmide capable de coder pour sa propre dégradation était incapable de se protéger de lui-même. En effet, l'origine de répllication *pSC101* permet de maintenir le pKD à environ 5 copies par cellules. L'opéron Red est généralement incapable de recombiner toutes les copies du plasmide. Ainsi s'il demeure un plasmide non recombiné celui-ci va médier la dégradation de tout pKD au sein de la cellule, sans discriminer les recombinés des non recombinés (Cf. 4.2 *Création de la plateforme Horus*).

Dans le but de combler cette faiblesse nous avons entrepris de créer une cible qui était capable de se reconnaître elle-même, à l'image d'espaceurs codants pour leur propre séquence, créant un état d'auto-immunité (Figure 25).

La région qui code pour le N-20 se constitue comme une mini-librairie CRISPR et comporte 3 éléments clés : i) Le promoteur pLTetO1, qui module l'activité du locus via l'inducteur aTc. ii) Un espaceur, appelé aussi N-20, codant pour une cible de 20 paires de bases, il est important de noter que le PAM associée à la Cas9 est absente au sein du N-20. iii) Le sgARN ou « *Scaffold* » long de 76 paires de bases, ce sgARN est l'ARN chimérique créer par Doudna et *al.*, il forme avec le N-20 un ARN qui est pris en charge par la Cas9.

Nous avons décidé de manipuler la séquence du sgARN pour créer une cible suicide (Figure 23 et 24). Le but était de créer une cible capable de provoquer sa propre dégradation tout en étant capable, après recombinaison de se protéger des autres pKD n'ayant pas recombinés. Pour cela nous avons exploité deux mécanismes intrinsèques de la Cas9 et de l'opéron Red.



**Figure 24. Repliement du sgARN non modifié**

La séquence primaire du Scaffold a été analysée via l'algorithme mFold afin d'observer la structure secondaire de l'ARN. Les structures importantes sont marquées d'une flèche.

En effet, comme mentionné plus haut dans l'introduction la reconnaissance de la cible par le complexe Cas9/sgARN est dépendante d'une séquence cœur (*seed*), une mutation dans cette région diminue grandement l'efficacité de clivage de la cible, et plusieurs mutations consécutives abolissent la dégradation. Dans le même temps la recombinaison par l'opéron Red requiert plusieurs erreurs d'appariement consécutives qui sont nécessaires pour éviter la réparation de la mutation introduite par les mécanismes MMR et BER.

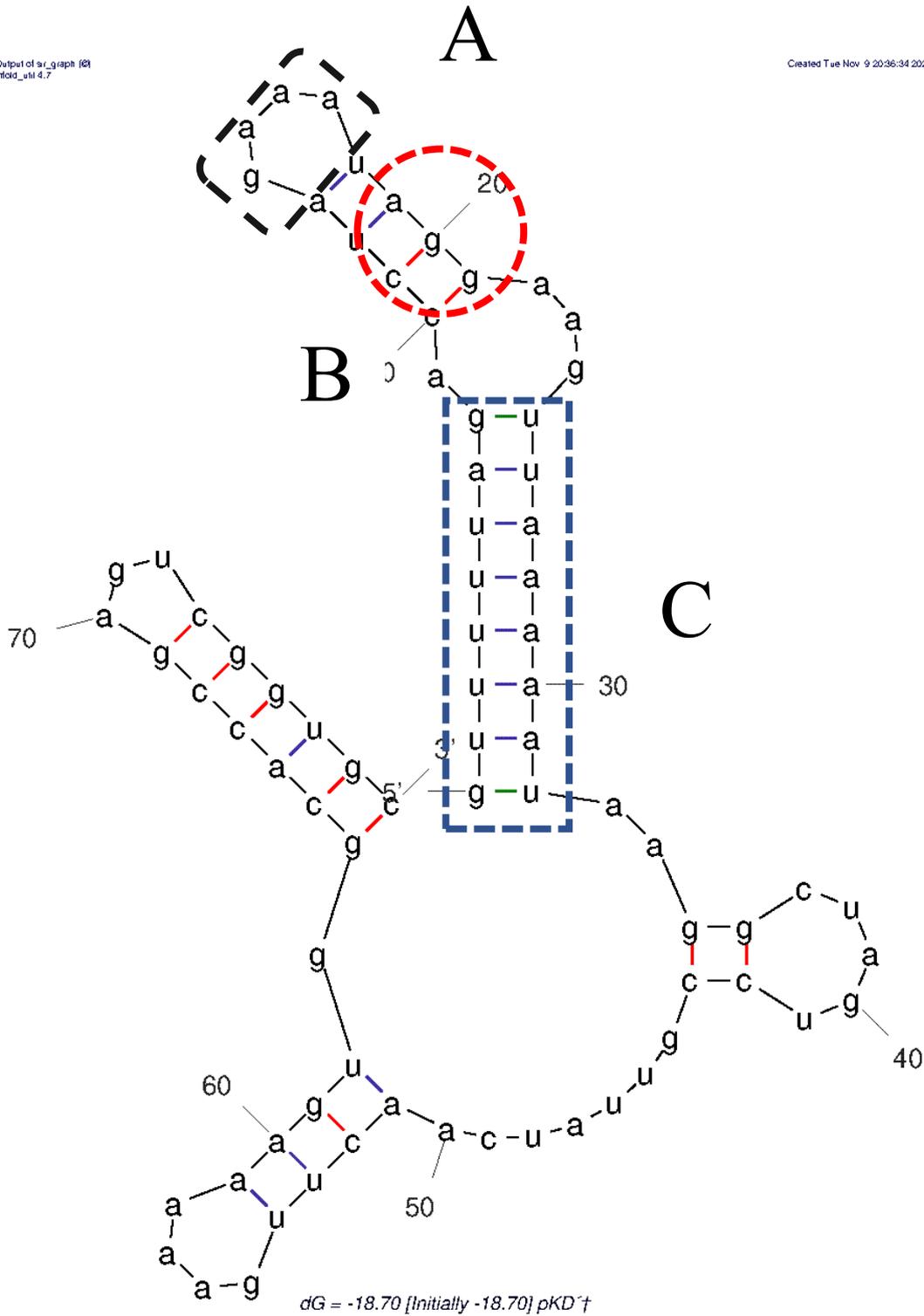
Le Scaffold est capable d'accueillir plusieurs mutations, tant que ces dernières permettent de conserver la structure de repliement prédite par des outils informatiques comme l'algorithme mFold. Le but ici était donc de créer des mutations tout en respectant le repliement du Scaffold.

Nous avons alors modifié la séquence du Scaffold afin d'introduire plusieurs éléments :

- i) Création d'un motif PAM en position relative 19-21 (encadré en rouge Figure 24).
- ii) Création d'une région plastique : Nous avons tiré profit de la séquence 5'-GAAA-3' en position relative 13-16 au sein du Scaffold (encadrée en noir Figure 24). Cette région artificielle permet de fusionner le crARN et le tracrARN et ne participe pas dans l'interaction du complexe Cas9/sgARN. Cette région peut donc accueillir presque n'importe quelle séquence sans perturber l'activité de dégradation par la Cas9. L'idée est d'exploiter cette région, l'ADN donneur en se recombinant sur le pKD viendrait alors changer cette séquence en sus de l'espaceur suicide du pKD. Cette région fut dénommée la région plastique de par sa capacité à accueillir n'importe quelle séquence.
- iii) Modification et allongement de la tige proximale afin de produire une cible contenant au minimum 17 paires de bases, afin de médier une activité de contre sélection. La Cas9 a

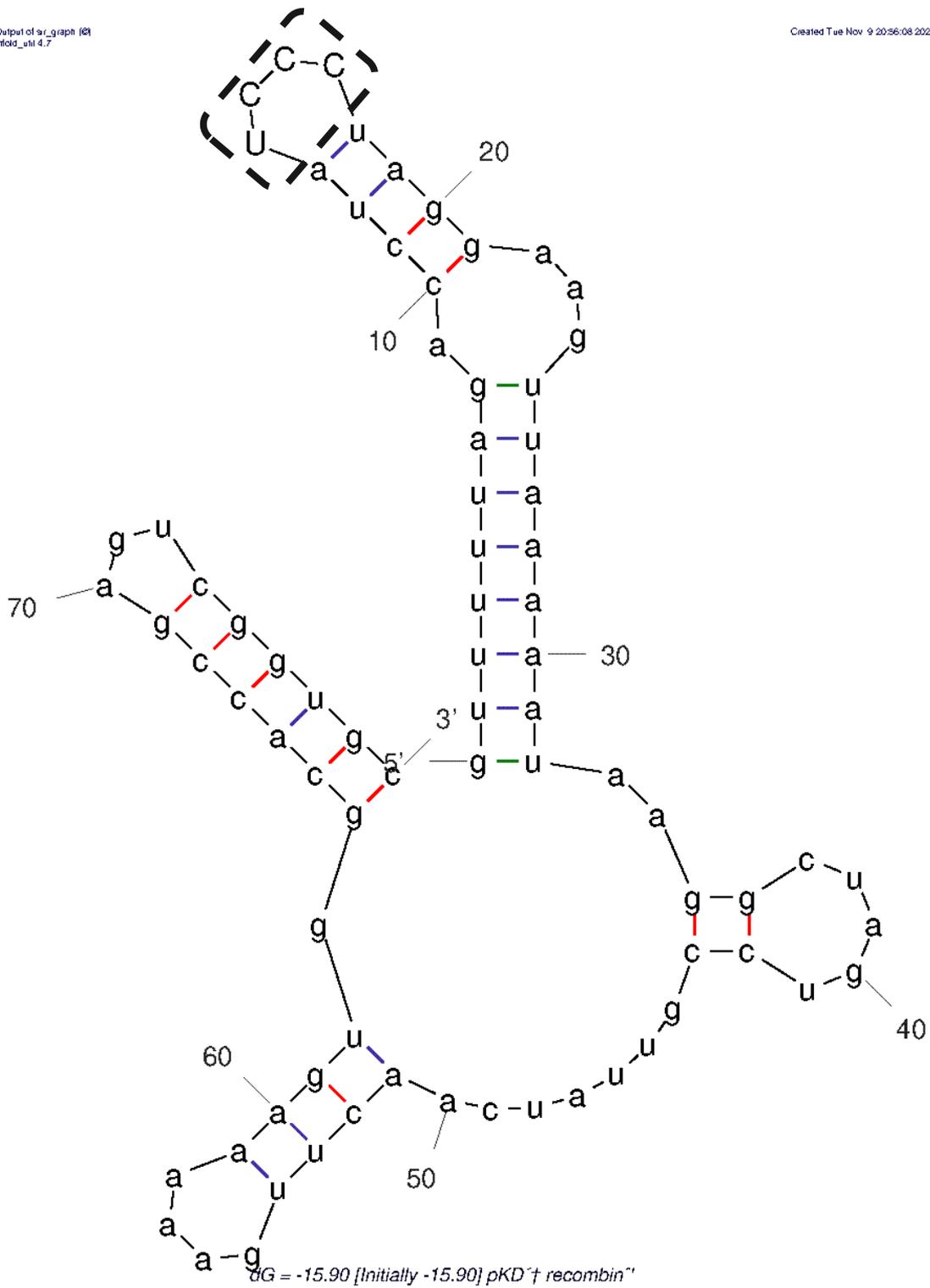
été rapportée être compatible avec des espaceurs allant de 22 à 17 paires de bases (encadré en bleu, Figure 24).

La construction du Scaffold une fois validée *in silico* fut construite par PCR mutagénique, puis insérée dans un pKD par CPEC. Le pKD portant ce Scaffold synthétique fut dénommé pKD<sup>2</sup>



**Figure 25. Repliement du Scaffold modifié pKD<sup>2</sup>**

La séquence du Scaffold original d a été modifiée afin de lui conférer une propriété d'auto-reconnaissance (pKD<sup>2</sup>) (A) Région plastique permettant la mise place d'un interrupteur génétique qui sera changé pour une autre séquence après recombinaison, empêchant la reconnaissance par les autres pKD non recombinés. (B) Création d'une séquence PAM (5'-AGG-3') (C) Allongement et modification de la tige proximale afin d'accommoder 17 paires de bases.



**Figure 26. Repliement du Scaffold de pKD<sup>2</sup> après recombinaison**

Séquence et repliement du Scaffold de pKD<sup>2</sup> après recombinaison. La région plastique est encadrée en noir, cette région via un ADN donneur simple brin sera permuté de (5'-GAAA-3') à (5'-TCCC-3') afin de protéger le pKD<sup>2</sup> recombinant des autres plasmides n'ayant pas recombinés.

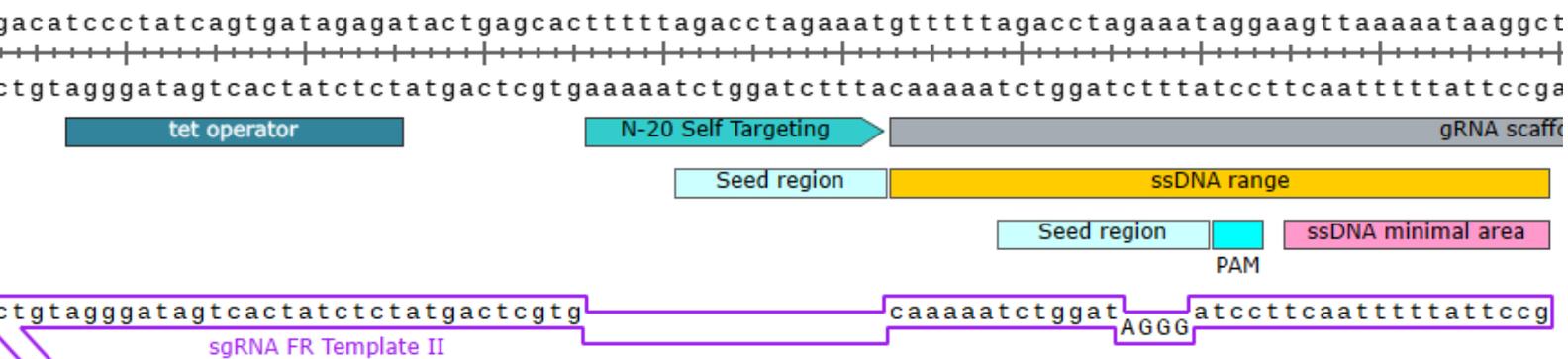


Figure 27. Locus du Scaffold synthétique de pKD<sup>2</sup>

Séquence du locus de pKD<sup>2</sup>. Les régions clés sont annotées. Un exemple d'ADN donneur (sgRNA FR Template II) venant recombinaison le pKD est encadré en violet. Cet ADN donneur remplacera la cible de pKD pour une cible visant le chromosome, dans un second temps, l'ADN donneur va aussi venir changer la région plastique, afin de protéger ce plasmide recombinant des autres plasmides coexistant au sein de la même cellule, n'ayant, eux pas recombinaison.

La construction par CPEC du pKD<sup>2</sup> fut extrêmement laborieuse. En effet, la PCR de fragment contenant des répétitions tend à former des produits non spécifiques. Ce phénomène est aggravé lorsque le succès de cette PCR dépend de l'appariement de ces régions répétées. Une fois obtenu dans DH5 $\alpha$  le pKD<sup>2</sup> fut extrait et transformé dans une souche de *V. cholerae* A1552 contenant le pCas9-cr4. La plupart du temps il nous fut impossible d'obtenir des colonies après transformation du pKD<sup>2</sup>. Pour les transformations réussies, une activité de contre sélection fut observée (Tableau 12). Cette dernière était cependant plus faible que celle obtenue pour pRaava ou n'importe quel pKD visant le chromosome. De plus cette dernière était aussi trop instable pour pouvoir être utilisée telle quelle au sein d'une plateforme CRISPR. Nous suspectons que la taille de l'espaceur était trop courte pour médier une activité comparable au système No-Scar.

**Tableau 11. Efficacité de contre sélection de pKD<sup>2</sup>**

aTc-	aTc+
65	1
25	3
2	0

En effet, les colonies obtenues après transformation étaient capables de pousser en présence d'aTc (200ng/mL) et de Spectinomycine (50 µg/mL), suggérant que la stabilité du système était suboptimale. De plus le pourcentage en GC de la cible codée par pKD était uniquement de 24%.

### 3.6.3 pKD<sup>3</sup>

Bien que le plasmide pRaava offrait une bonne efficacité de contre sélection de la résistance à la Spectinomycine, nous avons tenté de remplacer la cassette de résistance du pKD<sup>2</sup>. Originellement le pKD est sélectionnable par Spectinomycine, grâce au gène *aadA*, une cassette codant pour la résistance à la Spectinomycine et la Streptomycine. Nous avons souhaité remplacer ce gène par un autre gène codant pour la résistance à la Néomycine/Kanamycine, afin d'observer si un autre antibiotique pouvait améliorer la contre sélection sans pour autant changer la cible du pKD. En effet, *V. cholerae* a une forte tendance à former des mutants spontanés résistants à la spectinomycine /streptomycine, ce qui compliquait davantage le criblage dans nos expérimentations. Ce remplacement a été effectué par la génération d'un ADN double brin possédant 100 paires de bases d'homologies avec les régions adjacentes de la cassette *aadA* du pKD<sup>2</sup>. Cet ADN donneur nous permettait de remplacer la cassette de résistance et en même temps d'estimer l'efficacité de recombinaison

chez DH5 $\alpha$ . Une fois la cassette créée par OE-PCR, plusieurs colonies de DH5 $\alpha$  contenant pKD<sup>2</sup> furent cultivées dans du SoB avec 50 $\mu$ g/mL de Spectinomycine et 25 $\mu$ g/mL de Chloramphénicol à 30°C, 300RPM, une fois que la culture atteignait une DO<sub>600</sub> de 0.3-0.4. 0.2% de L-Arabinose fut ajouté au milieu et laissé incubé pendant 15 minutes avec rotation pour l'expression de l'opéron Red. 100ng de la cassette de la cassette purifiée fut électroporée dans 50 $\mu$ L de cellules compétentes. Immédiatement après électroporation, les cellules furent suspendues dans 1mL de SoC, et incubées à 30°C avec agitation à 300RPM pendant 1h, le temps que les cellules expriment le gène de résistance KanR.

Après 20 heures de croissance à 30°C, nous avons obtenu 10<sup>3</sup> UFC. 3 colonies furent choisies aléatoirement pour confirmer la présence de la cassette KanR. 2 colonies sur 3 possédaient bien le gène de résistance.

Le plasmide obtenu fut purifié depuis DH5 $\alpha$  puis transformé dans une souche d'*E. coli* DH5 $\alpha$  et de *V. cholerae* A1552 possédant déjà le pCa9-cr4. À notre surprise aucune activité de contre-sélection ne fut observée, qu'importe la concentration de Kanamycine employée (entre 50 et 100  $\mu$ g/mL).

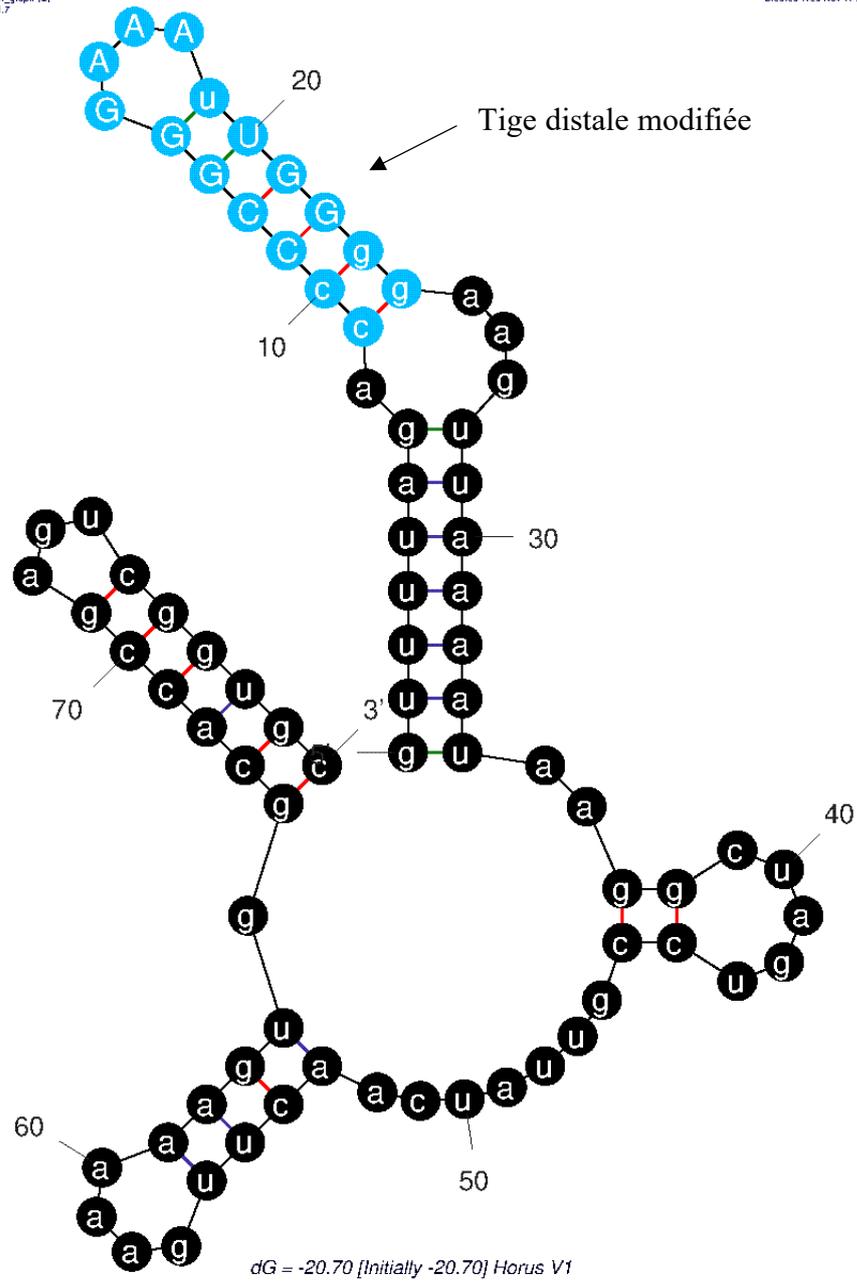
### 3.6.4 pHorusV1

L'efficacité de contre-sélection de pKD<sup>2</sup> laissait suggérer qu'il était possible de médier la contre-sélection du pKD sur une cible non létale (i.e. une cible intergénique ne codant pas un gène vital), cette capacité est un autre point crucial de notre plateforme CRISPR. En effet, aucune littérature ne définit si la coupure d'un plasmide dans une région intergénique est létale. Au sein d'un plasmide, seul le locus de répllication et le gène antibiotique sont véritablement essentiels à l'hérédité et au maintien du plasmide. Nous argumentons que la coupure codée par le sgARN suicide est suffisamment proche de l'origine de répllication pour engendrer la

dégradation de l'origine de réplication et cassette antibiotiques pour être tout de même considérée létale (l'origine de réplication *pSC101* et la cassette de résistance *aadA* se situent respectivement à 565 et 493 paires de bases de la mini-librairie CRISPR).

Nous avons souhaité modifier à nouveau la séquence du Scaffold. Nous avons concentré nos efforts sur une autre structure du sgARN de pKD<sup>2</sup> : la tige distale (nucléotides en bleu, Figure 27). La tige la plus proximale de la position 5' joue un rôle d'interaction avec la Cas9 tant de par sa structure que par sa séquence. Cette tige est une structure chimérique (Doudna *et al.*) et ne participe pas de manière séquence dépendante à l'interaction Cas9/sgARN. Nous avons rallongé la taille de la tige distale du sgARN pour créer un espaceur de 20 paires de bases. Nous avons aussi muté la séquence de la tige distale pour ramener le pourcentage de %GC à au moins 40.

La première itération de ces modifications fut dénommée H.S.E.P, de l'anglais (Higher Stem Extended + PAM), puis pas soucis de communication fut renommée Horus.



**Figure 28. Repliement de Horus V1**

Structure secondaire formée par le Scaffold synthétique Horus. Les changements par rapport au Scaffold original sont marqués en bleu. La structure est prédite par l'algorithme mFold.

Nous avons tenté de créer Horus par CPEC. Comme pour pKD<sup>2</sup> la construction de Horus par CPEC fut extrêmement difficile pour les mêmes raisons citées plus haut. La première itération de cet ARN, dénommée HorusV1 possédait un espaceur de 20 paires de bases et un % de GC de 45. Nous avons employé 5 amorces pour générer pHorus depuis p15A. La première paire d'amorces permettait de créer un fragment d'environ 5Kb, comprenant l'origine de répllication, l'opéron Red. Pour la création de la cible suicide, nous avons réalisé deux PCR mutagène consécutive. Via l'amorce F0 et R1, le scaffold est muté à plusieurs positions en respect de la séquence conçue *in silico*. Un fragment d'environ 3Kb est généré, contenant le promoteur araC et ayant deux régions homologues avec le fragment de 5Kb, une région d'ancrage de 300 paires de bases et une région correspondant à la cible suicide. Une fois le scaffold muté, une deuxième paire d'amorces est utilisée pour ajouter la cible suicide en amont du scaffold muté.

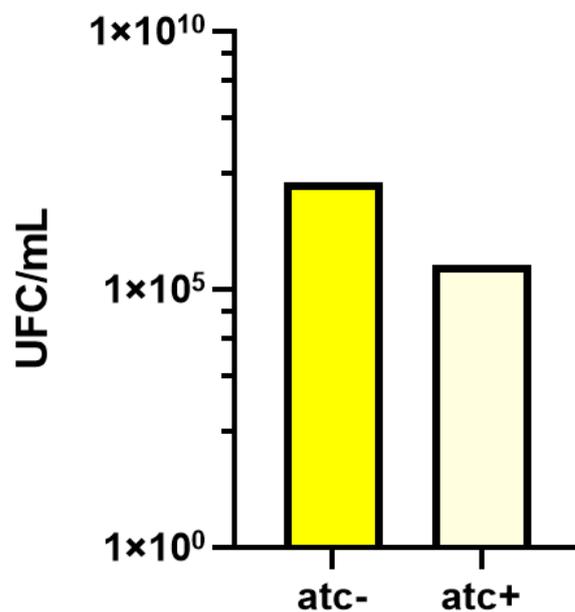
À ce point du projet nous avons voulu tester le système sur *E. coli* DH5 $\alpha$  et MG1655, le nombre de variable inconnues étant trop grand sur *V. cholerae* A1552, il nous aurait été impossible de conclure si un problème donné venait de notre construction ou d'une incompatibilité de A1552.

L'activité de contre sélection a été testée sur DH5 $\alpha$  puis sur *V. cholerae* A1552. Le choix et génotype de la souche est tout aussi crucial que le développement de l'ARN. Par exemple, lorsque testé chez DH5 $\alpha$  tous les plasmides (pKD, pRaava, pHorusV1) retournaient une efficacité de contre sélection de 100%, résultats directement en contradiction avec les résultats obtenus plus tôt dans le projet. En nous penchant un peu plus sur le génome de DH5A nous avons réalisés que cette dernière est  $\Delta recA$ . Cette mutation fait de cet organisme un choix intéressant comme souche pour maintenir des plasmides, mais rend la souche extrêmement sensible au stress génotoxique. Dans le même temps l'absence de *recA* diminue l'efficacité de

recombinaison. Ainsi l'efficacité de contre sélection a été testée sur MG1655 (Figure 29), 100 fois moins de colonies étaient observées sur milieu avec aTc par rapport au milieu sans aTc.

Lorsque testé sur *V. cholerae* A1552 le système retournait des résultats en concordance avec nos résultats obtenus plus tôt dans le projet. Quelques colonies étaient observées sur le milieu sans aTc et aucune sur milieu contenant l'aTc.

Nous avons voulu vérifier par séquençage Horus. Comme ce dernier comporte deux répétitions, les amplifications PCR de vérifications furent difficiles à réaliser. Après séquençage, à notre surprise il existait une mutation non souhaitée au sein du Scaffold de Horus, concernant la délétion d'une thymine au tout début de la séquence (*Cf. Annexes*)



**Figure 30. Efficacité de contre-sélection HorusV1 sur DH5α.**

La capacité suicide de Horus fut estimée via la transformation du pKD HorusV1 dans une souche de DH5A contenant le pCas9-cr4. Le nombre d'UFC fut calculé depuis la dilution la plus faible du spot assay ayant produit des colonies.

### 3.6.5 pHorusV74

Bien que la tige distale ne participe pas de manière séquence spécifique à l'assemblage Cas9/sgARN un respect de la symétrie et du repliement est vital. Cette délétion entraînait un changement de repliement conséquent impactant la capacité d'auto contre sélection de Horus, de plus cette délétion diminuait la taille de l'espaceur à 19 paires de bases. (*Cf. Annexe*).

Nous avons décidé alors de passer par synthèse chimique pour la création de Horus via la technologie gBlock. Cette technologie permet de synthétiser de longs fragments d'ADN double avec grande fidélité. Nous avons conçu un fragment de 469 paires de bases. Le fragment a été conçu de manière à ce que des régions uniques assurent l'appariement durant la CPEC, étape qu'il nous était impossible de réaliser par amplification PCR mutagène.

À ce point du projet, nous avons découvert un nouvel outil bio-informatique permettant d'estimer l'efficacité d'un N-20 à médier sa coupure. Cet outil est l'algorithme On-score décrit par Doench *et al.* (169) Plusieurs critères existent, comme l'identification de certains motifs connus pour provoquer une baisse d'activité significative si ces motifs sont présents à la fin de la séquence *seed* (GGCC, TATT, TTTC, TTGC). HorusV1 possède un score prédit de 25. Un score relativement faible.

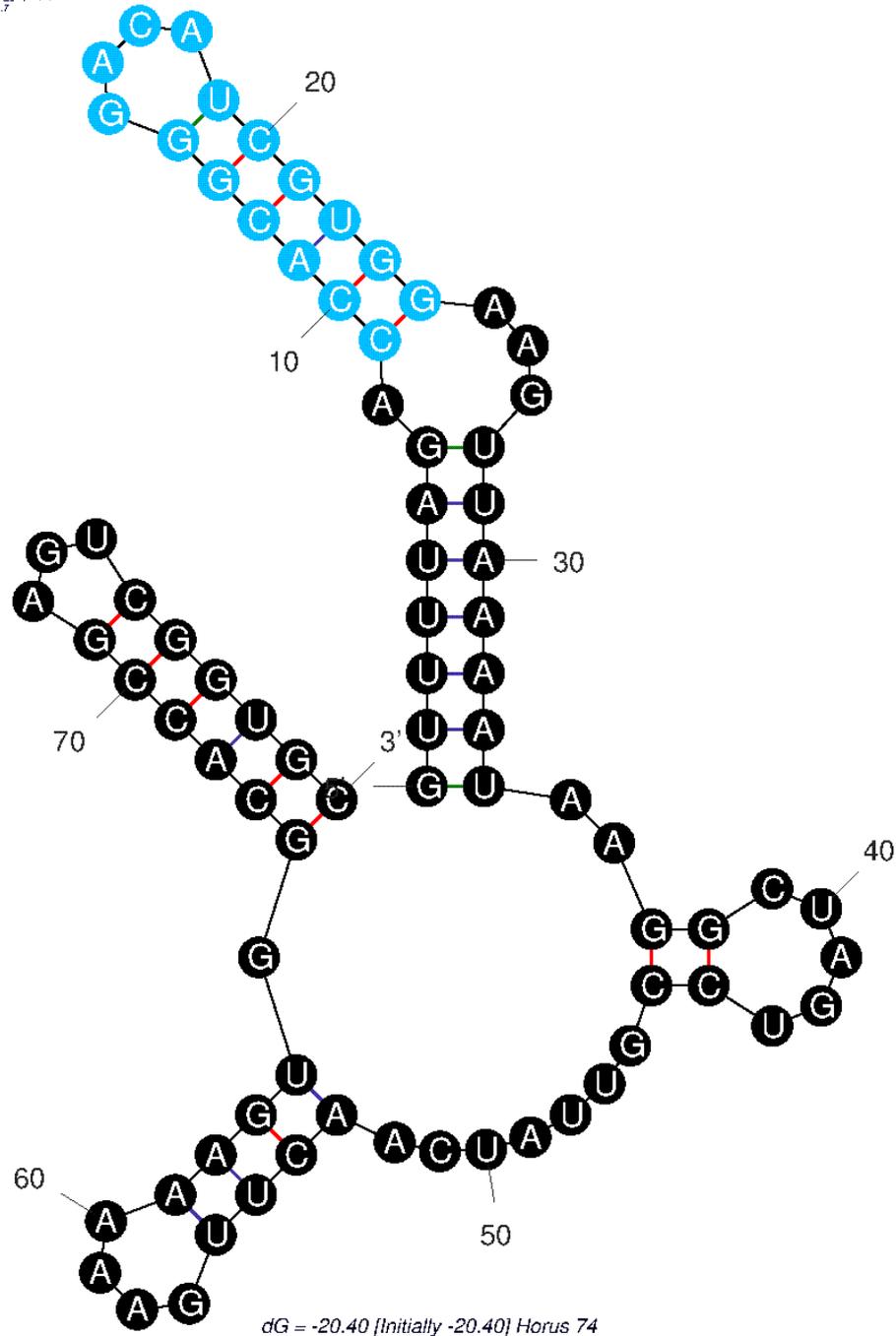
Avant commande nous avons apporté des dernières modifications à la séquence pour augmenter le On-Score. En sus de mFold et du On-score, nous avons utilisé des outils comme EteRNA. Cet outil de manipulation d'ARN en ligne permet de simplifier la méthodologie de mutagenèse par un support visuel.

Les dernières modifications sur Horus furent la permutation de 5 paires de bases, dans sa première version Horus (pHorus V1) contenant l'enchaînement de 4 résidus de guanine, cet enchaînement particulier tend à introduire des stress topologiques au sein de la structure secondaire d'un ARN (170), les nucléotides modifiés de la tige distale sont visibles en bleu sur

la Figure 30. La séquence plastique fut aussi changée de manière à stabiliser la boucle (baisser son énergie libre), la séquence de la région plastique pour Horus74 est 5'-GACAT-3'. Nos modifications ont permis d'augmenter le On-Score de Horus de 25 à 74.

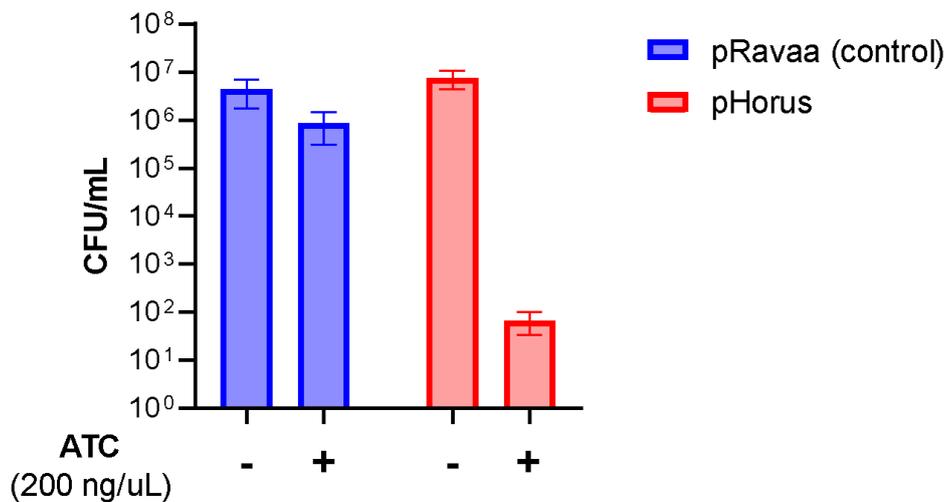
Dans un second temps nous avons voulu tester une version alternative du promoteur pLTetO1, dans nos mains l'évènement génétique amenant une souche à devenir escaper se produisant toujours au sein de ce promoteur (*Cf. 3.7 Escapers*). Pour contrer l'évènement génétique, nous avons souhaité créer un promoteur alternatif contenant un homologue de la séquence TetO2. Ce promoteur emploie un tandem TetO1-TetO2. TetO1 est une séquence reconnue par TetR mais possédant une séquence différente à TetO2. En ajoutant TetO2. La construction s'est révélée être une impasse. TetO1 possède une affinité avec TetR mais celle-ci est plus faible que l'affinité TetO2/TetR, s'en résulte alors une fuite d'expression grandement augmentée. L'efficacité de contre sélection était présente, mais inférieure à celle de pLTetO1.

En testant le système sur *E. coli* nous avons pu agrandir notre résolution pour estimer la contre sélection. La transformation par électroporation offre en effet plusieurs ordres de magnitudes d'accroissement du nombre d'UFC que pour *V. cholerae* permettant d'apprécier la capacité de contre sélection de chaque plasmide plus précisément. Nous nous attendions à obtenir une contre sélection proche de celle de pRaava, de manière surprenante HorusV74 possédait une efficacité de contre sélection 10000 fois supérieure à celle de pRaava. (Figure 31). Nous argumentons qu'en changeant la séquence du Scaffold pour créer Horus nous avons dans le même temps stabilisé l'ARN en lui-même. En effet l'énergie de repliement de Horusv74 est de -20.40 kcal/mol, celui de pKD est de -18.70 kcal/mol, le Scaffold original est de 17.90 kcal/mol.



**Figure 31. Repliement de Horus74**

Structure secondaire formée par le Scaffold synthétique Horus. Les changements par rapport au Scaffold original sont marqués en bleu. La structure est prédite par l'algorithme mFold. Nous avons permuté plusieurs paires de bases au sein de la tige supérieure pour stabiliser la structure et augmenté le On-score



**Figure 32. Efficacité de contre sélection de HorusV74 sur MG1655**

La capacité suicide de Horusv74 fut estimée via la transformation du pKD HorusV74 dans une souche de MG1655 contenant le pCas9-cr4. Le nombre d'UFC fut calculé depuis la dilution la plus faible du spot assay ayant produit des colonies.

Via le système Horus, une souche de MG1655 prend 2h pour atteindre une DO<sub>600</sub> de 0.5, A1552 met 3h pour atteindre cette DO<sub>600</sub>. Ces temps de croissance sont extrêmement proches d'une souche sauvage à 30°C.

### 3.7 Escapers

Les escapers sont un élément central des systèmes de contre sélection. Comme mentionné dans l'introduction, certaines cellules échappent à la contre sélection sans avoir vécu l'évènement de recombinaison souhaité. Dans le système No-Scar l'auteur émet l'hypothèse que les escapers proviennent de mutations ponctuelles au sein de la Cas9. Nos résultats montrent que d'autres évènements génétiques peuvent être à l'origine des escapers. Nous avons séquencé le pKD de plusieurs souches escapers obtenues. Tous les pKD séquencés portaient la même signature où le promoteur pLTetO1 portait une délétion de la boîte -35 et d'un site TetO2

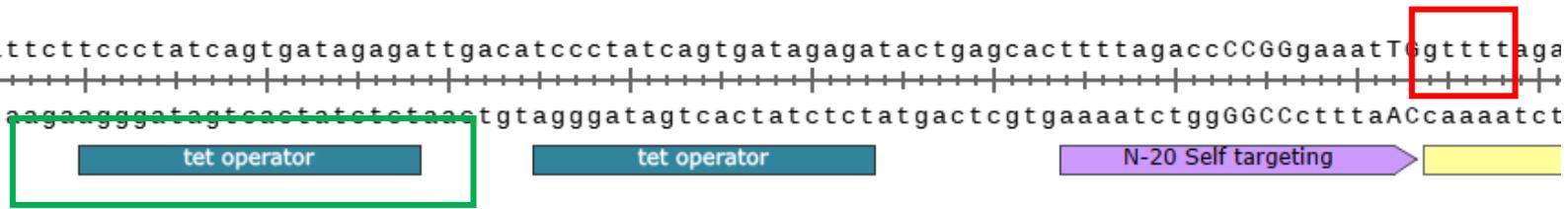
(Encadré en vert, Figure 34B). Cette mutation abolie complètement l'expression du sgARN en présence d'aTc protégeant les cellules portant cette mutation de la contre sélection par la Cas9. En sus de la disparition du premier site TetO2 le Scaffold porte une délétion d'une thymine (encadré en rouge, Figure 34B).

Ce promoteur consiste à un tandem de deux sites palindromiques, les sites TetO2. Le premier site de ces sites TetO2 (5'-tcctatcagtgatagaga-3') est répété une seconde fois, séparé du premier par 6 paires de bases. Ces 6 paires de bases représentent par la boîte -35, élément essentiel à la transcription.

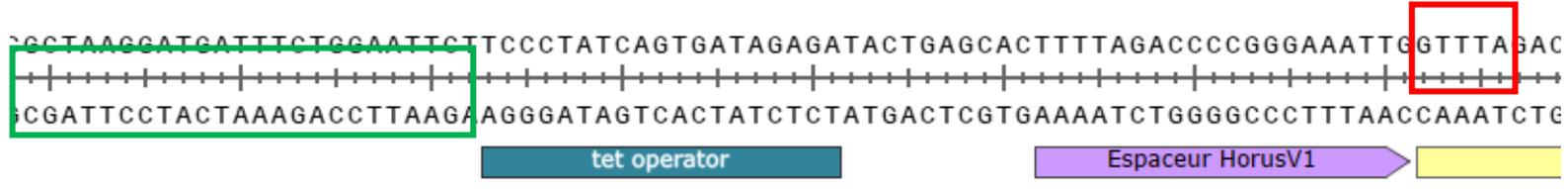
Dans cette stratégie de désactivation du système le premier site TetO2 saute et recombine sur le second site, entraînant la délétion des 6 paires de bases, cette délétion élégante serait une stratégie simple pour la bactérie de se protéger de la contre sélection, puisque l'expression du sgARN est avortée.

Les escaper sont des bactéries ayant survécu à l'événement de contre sélection, mais ne possédant pas le génotype désiré. La présence d'une Cas9 sur sa cible permet par sa seule présence physique sur le locus perturber l'expression d'un gène. Cette propriété est réimmiscence de l'acquisition de nouveaux espaceurs dans la réponse PAM indépendante. Jinek et al décrivent deux mutations (D10A et H841A qui abolissent la capacité de la Cas9 à cliver sa cible, sans perturber la fonction de reconnaissance, on parle alors de dCas9 qui permet de faire de l'atténuation d'expression (RNAi) (171).

# A



# B



**Figure 34. Alignement de HorusV1 et un escaper**

Résultat du séquençage de plasmide de référence (A) et d'une colonie ayant poussé sur aTc+. Nous observons une délétion du premier site TetO2 provoquant l'inhibition totale de l'expression du locus contenant la cible (mini-librairie). La mutation commune de la thymine est encadrée en rouge. La délétion provoquant le phénotype escaper est encadrée en vert.

Pour HorusV74 (construction expliquée ci-après, sous-section pHorus V74) une autre voie semble apparaître (Figure 35). En effet lorsque nous avons testé HorusV74 au sein d'un cycle No-Scar, nous avons séquençé un escaper provenant de la réaction, l'amplicon généré nous apprend un événement de délétion que nous n'avions pas observé jusqu'ici. Usuellement, les cellules deviennent escapers par la délétion d'un site TetO2 au sein du promoteur pLTetO1. Dans cette colonie, l'ADN simple brin a réussi à remplacer la cible et la région plastique, comme pour le mutant, mais nous observons aussi une délétion du motif 5'-ACGTTTTAGACC-3', un motif de 12 paires de bases présentes au début du scaffold synthétique. Ce résultat est surprenant, car l'ADN donneur joue le rôle de matrice de réparation, le remplacement de la cible et de la région plastique démontre que l'événement d'intégration s'est produit. La délétion a dû se produire après recombinaison. Néanmoins, le fait que même sur une souche escaper l'événement d'intégration s'est produit nous amène à penser que la recombinaison sur le pKD n'est pas l'évènement limitant.



**Figure 35. Séquence d'une souche escaper HorusV74 après recombinaison**

Résultat du séquençage d'une souche escaper après un cycle No-Scar via la construction pHorusV74. La délétion du motif provoquant le phénotype escaper est encadrée en vert.

### 3.8 Implémentation de Horus dans MG1655

Une fois la capacité de Horus à médier sa propre dégradation établie nous avons entrepris de l'implémenter comme élément de contre sélection dans le système No-Scar. Nous avons testé nos deux versions de Horus obtenues, HorusV1 qui avait été obtenu par CPEC et HorusV74 par gBlock. Dans les deux cas le vecteur d'accueil était le plasmide pKD\_p15A. Nous avons choisi ce plasmide, car il s'agit du plasmide reçu avec le système No-Scar. Bien que la CPEC soit une méthode fidèle et utilise pour la création *in vitro* de plasmides, cette dernière contiendra toujours un degré d'infidélité de par l'introduction d'erreur de la réplication avec la Q5. Ce phénomène reste bien entendu rare, la Q5 étant une enzyme conçue expressément pour diminuer la chance d'apparition de telles mutations.

- Conception expérimentale

Pour Horus V1 nous avons choisi de tester le système dans *V. cholerae* A1552. L'ADN donneur pour créer la mutation était le même pour l'implémentation de la plateforme No-Scar. Le second ADN donneur recombinait sur le pKD Horus pour remplacer la cible suicide avec une cible visant le gène *LacZ*. Cet ADN venait aussi remplacer la séquence 5'-GAAA-3' par la séquence 5'-TCGT-3' au sein de la région plastique, dans le but de protéger le pHorus recombiné des autres pKD non recombinés au sein de la cellule. Le protocole de transformation des ADN donneurs était le même que les protocoles décrit précédemment. La réaction et les témoins ont récupéré 2h dans du SoB à 30°C puis étalés sur milieu contenant 120µg/mL de Spectinomycine et 200 ng/mL d'aTc. Le milieu contenait aussi du X-Gal 1X pour observer la capacité des clones à dégrader le composé chromogène.

Pour *V. cholerae* A1552 nous n'arrivions pas à isoler des colonies blanches. Cependant il nous était possible de trouver que 3 plasmides sur 8 avaient bien recombinaison pour produire pKD\_ *lacZ* en employant une paire d'amorces cherchant la cible recombinaison sur le pKD. Nous avons donc observé la présence d'une bande de 500 paires de bases, témoignant de l'évènement de recombinaison sur le pKD.

Un point important à préciser est l'apparition de colonies satellites que nous n'avions pas observés jusqu'ici. Des grosses colonies bleues étaient entourées de micro-colonies. (Cf. *Annexes*), ces colonies n'étaient présentes que sur la réaction et pas le témoin.

### **3.9 Création d'un mutant *lacZ* de MG1655 par HorusV74**

Afin de nous assurer que notre plateforme était bien capable de générer des mutants, nous avons réalisé un essai sur *E. coli* MG1655. Nous avons choisi cet organisme, car il nous permettait d'avoir la meilleure efficacité de recombinaison par l'opéron Red. Le phage lambda ayant coévolué avec le genre *Escherichia* la littérature concernant l'opéron Red emploie cet organisme comme organisme modèle.

- Conception expérimentale :

Nous avons donc conçu deux ADN donneurs, tous deux simples brins et long de 90 paires de bases. Les deux ADN donneurs sont aussi modifiés chimiquement par des liens phosphorothioates afin de les protéger des nucléases de *E. coli* MG1655. Le premier donneur : ssDNA  $\Delta lacZ$ , permet de créer une délétion de 409 paires de bases au début du gène *lacZ*. Abolissant l'expression de la  $\beta$  galactosidase. Cette mutation permet de réaliser un test bleu-blanc, vérifiable aisément par X-Gal.

Le second donneur, ssgDNA\_ *lacZ* se recombine dans le pKD portant la construction HorusV74 (pHorus), il permet de remplacer la cible suicide 5'-3' par une cible 5'-3' visant le locus *lacZ* de MG1655. Cette cible est absente des clones ayant recombines avec le premier ADN donneur. En sus du remplacement de la cible suicide, le second ADN donneur remplace aussi le motif 5'-GACAT-3' par le motif 5'-ACTGC-3' dans la séquence plastique.

Dans notre stratégie de contre sélection, seule la sous-population ayant recombine les deux ADN donneurs peut survivre à l'évènement de contre sélection. Nous avons réalisé un cycle No-Scar avec une souche de MG1655 contenant les deux plasmides du système. La concentration finale d'ADN donneur était de 2µM. Le témoin a reçu 2µM d'un ADN donneur simple brin ne protégeant pas de la contre sélection.

Après électroporation, les cellules ont récupéré 3 heures dans du milieu SoC pour ségréguer les allèles mutantes. Après 3 heures 1µL d'aTc, pour une concentration finale de 200 ng/mL a été ajouté au milieu puis incubé pour encore 2 heures. Après un total de 5 heures un spot assay fut réalisé jusqu'à la dilution 10<sup>-4</sup> sur un milieu contenant 100µg/mL de Spectinomycine, 200 ng/mL d'aTc et 1X de X-Gal. 50uL des dilutions 10<sup>-4</sup> furent étalés sur le même type de milieu. Après ensemencements des milieux, les réactions de transformation ont été conservées à 4°C. Après incubation à 30°C pendant 19 heures, le spot assay retourna une population de 6x10<sup>6</sup> UFC/mL pour la réaction et 1.6x10<sup>4</sup> UFC/mL pour le témoin (une différence de population de 37.5). Aucune croissance n'était observable sur les milieux ensemencés par étalement.

Nous avons donc réalisé un second spot assay depuis les cultures réfrigérées sur le même type de milieu, en incluant cette fois-ci des milieux ne contenant que 50µg/mL de Spectinomycine. 50uL des dilutions 10<sup>-2</sup> furent aussi étalées sur le même type de milieu.

Nous observons que la contre-sélection perd en efficacité après un passage à 4°C. Pour le même milieu contenant 100 µg/mL de Spectinomycine le témoin donnait 24 UFC sur la dilution 10<sup>-</sup>

<sup>2</sup>. La réaction a produit 1 UFC sur la dilution  $10^{-4}$ . Il nous a été possible d'isoler une colonie blanche parmi depuis les 50 $\mu$ L de la dilution  $10^{-2}$  étalés sur milieu contenant 50 $\mu$ g/mL. Pour 156 colonies bleues, il nous est possible d'isoler une colonie blanche. Nous amenant à une efficacité de 0.64%.

Nous avons testé par la suite plusieurs conditions de récupération en présence ou non d'agent de sélection (récupération de 4h, contre sélection sur une O/N en présence d'une combinaison d'aTc et de spectinomycine). Nous obtenons une efficacité de 6.6% (40 colonies pour 600 escapers) si la réaction récupère pendant 4 heures puis est étalée directement sur milieu contenant de l'aTc et de la spectinomycine. L'efficacité est nulle si la récupération dure uniquement une heure.

Nous avons voulu vérifier le génotype des colonies blanches isolées ainsi que d'une souche escaper. Pour cela nous avons réalisés une amplification PCR depuis les colonies directement (Cf. 2.8 PCR sur colonies) à l'aide d'une paire d'amorces cherchant le gène *lacZ* sur le chromosome et une paire d'amorces cherchant sur la construction Horus sur le pKD. Les amplicons provenant du chromosome de MG1655, retournaient bien une bande de 100 paires de bases la taille attendue pour le mutant *ΔlacZ*.

Nous avons fait séquencer les amplicons obtenus depuis l'amplification du pKD et du chromosome pour vérifier à la fois l'évènement d'intégration et la fidélité de cet évènement.

Pour le pKD nous avons observé que la souche mutante portait exactement les mutations prévues elles aussi *in silico*. Aucune délétion créée par l'ADN donneur est présente. La cible suicide 5'-TTTTAGACCACGGGACATCG-3' a bien été remplacée par la cible 5'-TGTTGATGAAAGCTGGCTAC-3', qui vise elle, le locus *lacZ*. La région plastique a bien été remplacée correctement par le donneur ADN (Figure 36A).

Pour le chromosome nous avons obtenu une faible résolution de séquençage, mais suffisant pour observer que la mutation créée était celle prévue *in silico* (Figure 36B).

A



B



Figure 36. Séquences du pKD HorusV74 et du chromosome d'un mutant MG1655  $\Delta lacZ$  après recombinaison

- (A) Résultat du séquençage et visualisation de la séquence pour le pKD après recombinaison, la cible suicide de pHorusV74 à bien été changée pour un espaceur ciblant le locus *lacZ* de MG1655 (N-20 *lacZ*). La région plastique a aussi recombiné (région seed).
- (B) Résultat du séquençage et Visualisation du chromosome de MG1655 mutant. La souche porte la mutation souhaitée qui la protège de la contre sélection par la nouvelle cible (N-20 *lacZ*).

## Chapitre 4 : Discussion

### 4.1 Implémentation de la plateforme No-Scar chez *V. cholerae*.

La mutagenèse dirigée est la méthode de choix pour l'étude de la fonction des gènes bactériens. Bien que plusieurs techniques existent pour la mutagenèse dirigée chez *V. cholerae*, celles-ci peuvent s'avérer longues et peu efficaces. En effet, celles-ci nécessitent le plus souvent de réaliser 2 étapes, la première consistant en le remplacement du locus d'intérêt par une cassette de résistance à un antibiotique, et la seconde en l'excision de la cassette de résistance. D'autre part, *V. cholerae* est capable de compétence naturelle qui peut être employée pour créer des mutations (172). Cependant, l'efficacité de cette technique étant très faible (de l'ordre de 1 transformant sur  $10^9$  bactéries), sa résolution reste modérée et se limite principalement à l'insertion d'une cassette de résistance au sein d'un gène.

La présence de cette cassette de résistance facilite la sélection des mutants, mais ampute de nombreuses études surtout si ces dernières concernent l'étude des mécanismes de résistance en raison principalement des possibilités de résistance croisées entre antimicrobiens (173). D'autre part, pour certaines souches de *V. cholerae* (N16961) l'emploi de la compétence naturelle est impossible, de par une mutation dans le gène *hapR*, impactant à la fois la compétence naturelle et la détection du quorum (174).

Dans ce contexte, la technologie CRISPR/Cas9 s'avère très attrayante puisqu'il est théoriquement possible de s'affranchir de la cassette de résistance et donc de diminuer le nombre d'étapes nécessaires à l'obtention d'un mutant. D'autre part, elle offre la possibilité de construire facilement des mutations sur un codon spécifique, permettant ainsi d'étudier le rôle d'un acide aminé dans la fonction d'une protéine (102). Au début de notre étude, aucune technologie CRISPR/Cas9 n'avait été rapportée chez *V. cholerae*. Depuis, seuls des études de

dCas9, n'impliquant pas la délétion du gène, mais une forte répression de l'expression de celui-ci a été rapporté (175).

Le choix de la plateforme CRISPR/Cas 9 s'est porté sur le système No-Scar publié par Reisch *et al* (102). Ce système offrait l'une des meilleures efficacités publiées. D'autre part, il offrait la possibilité de pouvoir contrôler l'expression de la Cas9 via le promoteur *tetO/tetR*, un avantage indéniable chez *V. cholerae* puisque les systèmes utilisant le promoteur natif de la Cas9 de *Streptococcus pyogenes* se sont avérés extrêmement toxique chez *V. cholerae* A1552 (résultats préliminaires du laboratoire, non publiés). Finalement, la thermo sensibilité du plasmide pKD est un atout pour son élimination après obtention de la mutation.

L'implémentation de cette plateforme dans *V. cholerae* fut un travail de longue haleine. Nous avons rencontré des défis dans de nombreuses étapes du protocole, comme la transformation du pCas9-cr4, la création d'un pKD contenant un espaceur ciblant le chromosome de A1552 via CPEC et finalement l'isolement des clones recombinants.

La majorité de nos tentatives avec le système No-Scar natif se sont révélées infructueuses. Bien qu'une réduction de la population fût observée, cette dernière restait non significative. De plus, aucun mutant ne fut retrouvé par génotypage. D'autre part, pour l'essai de la création du mutant *ctxAB* nous obtenions un profil multi bandes après amplification par PCR. Ces bandes ne proviennent pas d'une mauvaise conception d'amorce, car une amplification PCR de l'ADN génomique de A1552 retournait une seule bande. Il nous est actuellement impossible de conclure sur l'origine de ces bandes. Une piste possible serait la présence du locus excisé sous forme de plasmide au sein de la cellule (176).

Via le protocole No-Scar nous avons pu obtenir un mutant de A1552 pour le gène *lacZ*, après séquençage la souche contenant bien la délétion souhaitée. Les donneurs ADN simples brins protégés par 4 liens phosphorothioates offrent une fréquence de recombinaison plus élevée,

particulièrement si la cible est un plasmide et non le chromosome. Cependant la colonie blanche était « noyée » par des colonies bleu escapers au génotype sauvage. Si ce faible taux d'efficacité est acceptable pour la construction de mutant visiblement ciblables, il n'en est pas de même pour des mutants criblables uniquement par PCR pour lesquels de nombreuses colonies devraient être testées (de l'ordre de plusieurs centaines) pour identifier un mutant. Ainsi la méthode n'est pas nativement utilisable chez *V. cholerae*.

Entre nos mains deux phénomènes apparaissent pour *V. cholerae* :

Premièrement, le système de contre sélection CRISPR-Cas9 présente une cytotoxicité importante. Cette cytotoxicité est particulièrement perceptible par le temps qu'une souche portant les deux plasmides du système prend à atteindre une DO<sub>600</sub> de 0.5. Ce temps varie de 6 heures à 8 heures pour A1552, alors que sans ces plasmides, A1552 atteint cette DO<sub>600</sub> en 2 à 3 heures. Cette cytotoxicité n'est cependant en rien comparable avec celle retrouvée avec les systèmes utilisant le promoteur natif de *S. pyogenes* pour lesquels une lyse bactérienne totale est observée après seulement 2 à 3 heures de croissance (résultats non publiés). La cytotoxicité du système No-Scar peut s'expliquer par la fuite d'expression du sgRNA et de la Cas9. Cette expression, bien que faible crée, une pression de sélection. Face à cette pression, la population qui accueille un tel système finit par désactiver le système. Les cassures double brin sont extrêmement toxiques pour les bactéries. Il a été rapporté que les cassures doubles brins perturbent le cycle cellulaire et particulièrement la division cellulaire : la bactérie stoppe sa division cellulaire si un dommage au chromosome est détecté et tente une réparation d'urgence, qui résulte le plus souvent dans la mort de l'hôte (177).

Ainsi, le système No-Scar natif ne permet pas d'obtenir aisément des mutants portant des mutations ponctuelles chez *V. cholerae*. Il serait donc nécessaire de recourir à un phénotype afin d'isoler ces clones mutants, limitant ainsi grandement les avantages d'un tel système.

Dans un autre temps, l'opéron Red présente une efficacité de recombinaison plus modeste chez *V. cholerae* que chez *E. coli*. La littérature rapporte, pour *V. cholerae* une efficacité diminuée 1000fois plus faible par rapport à MG1655, résultats en concordance avec nos essais de transformation.

D'autre part, la littérature décrivant l'emploi de l'opéron Red chez A1552 est parcimonieuse (159).

L'efficacité de l'opéron Red peut être augmentée par l'emploi d'ADN donneurs simples brins, capables d'évader les mécanismes de réparation MMR, BER, et protégés des exonucléases, car ils offrent la plus grande ARF absolue. Il existe peu de possibilités de manipuler le système au même degré que le système CRISPR-Cas9. Il reste possible d'employer des gènes secondaires comme *dam*, qui permettent un état d'hyperméthylation mettant en échec les systèmes MMR et BER. La qualité de l'ADN donneur reste un point déterminant. La qualité de synthèse chimique d'ADN donneurs simple brin est virtuellement limitée à 90 paires de bases et dépend de l'efficacité de couplement. Par exemple, la synthèse d'un fragment de 90 paires de bases, avec un rendement théorique de 99.5%, permettra d'obtenir 70% d'oligonucléotide ayant la bonne séquence (taille + séquence) et 30% d'oligonucléotides portant une délétion. Notre fournisseur d'oligonucléotides (Invitrogen) nous permet d'obtenir uniquement 30% d'oligo désiré.

Toutefois, une fréquence de recombinaison amoindrie n'est pas le point limitant des techniques de mutagenèse dirigée puisque la contre sélection permet d'augmenter la fréquence de sélection de l'évènement de recombinaison (ARF). Toute méthode de contre-sélection, qu'elle soit par CRISPR-Cas9 ou par *sacB* cherche à créer la plus grande population sensible à la contre sélection possible.

Selon la méthode de sélection employée, le même événement génétique retournera une ARF différente. On peut alors concevoir une ARF absolue de l'ordre de  $10^{-4}$  à  $10^{-5}$  (1 clones mutants pour  $10^4$  à  $10^5$  cellules négative) et une ARF relative qui peut approcher les 100%.

## 4.2 Création de la plateforme Horus

Le succès des systèmes CRISPR peut se retrouver dans la pléthore de plateformes commerciales disponibles. Ces plateformes se concentrent plus sur les eucaryotes. Il est possible de commander des sgARN contenant des espaceurs provenant de bio banques ayant démontré *in situ* une grande efficacité. Pour des cibles non répertoriées, il est aussi possible de créer *de novo* des sgARN. Plusieurs sgARN candidats sont générés en respect de plusieurs modèles de prédiction (On Score, Off-Target score). Il est aussi possible de commander des sgARN dont la structure primaire (suppression terminateur de transcription) et secondaire (enthalpie libre) a été optimisée.

Pour les procaryotes les plateformes commerciales CRISPR sont moins nombreuses.

La littérature rapporte tout de même plusieurs stratégies alternatives au système No-Scar. On retrouve le système CRMAGE (178): ce système repose sur le même principe de recombinaison via l'opéron Red puis contre sélection. Dans ce protocole un plasmide : PMA7CR\_2.0 portes la Cas9 et la protéine SSB  $\beta$  de l'opéron Red seule représentante de l'opéron usuel. Gam et Exo sont nécessaire uniquement pour les ADN donneurs double brin.

Le plasmide pMAZ-SK lui, encode la librairie CRIRSPR contenant 3 espaceurs. Deux de ces espaceurs ciblent l'origine de répllication du plasmide (*ColA*) et le gène de résistance (Kanamycine), et assurent donc le recyclage du plasmide. Le troisième espaceur, lui, cible une

séquence au sein du chromosome pour réaliser une contre sélection. Il reste cependant nécessaire de régénérer l'espaceur de ce plasmide *in vitro*.

On peut aussi retrouver la plateforme Pro-AG (179) cette plateforme s'oriente dans le recyclage de plasmide à grandes copies. Son mécanisme d'action est similaire au protocole No-Scar à l'exception d'un plasmide portant une cassette ayant des bras d'homologies avec la cassette antibiotique AmpR du plasmide cible et contient un sgARN. L'induction de l'opéron Red intègre ce sgARN dans la cassette de résistance du plasmide. Après induction du sgARN par l'aTc la souche clive son propre plasmide. Ce premier événement d'intégration permet de créer une boucle *rétroactive* positive ayant pour conséquence d'augmenter la concentration de sgARN dans la cellule (on parle de dosage sgRNA). Cependant cette technique présente un intérêt limité pour les modifications chromosomiques, la charge de travail moléculaire était très élevée par rapport à d'autres protocoles existants.

L'emploi d'orthologue de l'opéron Red peut être une possibilité, la littérature retourne le remplacement de l'opéron Red par l'élément ICE SXT et l'orthologue W3 chez (153, 154, 180).

Dans nos mains SXT ne semble pas augmenter l'efficacité, et cause à l'inverse une augmentation de la cytotoxicité. Cependant le locus de SXT que nous avons employé est le locus natif, Burrus et al décrivent l'emploi de SXT comme recombinaise et décrivent un atténuateur de transcription en aval d'Exo (181).

La nature modulaire des systèmes CRISPR permet une grande liberté pour apporter des changements au système. Le plus gros changement que nous avons entrepris dans la stratégie de contre sélection du protocole No-Scar est le déplacement de la contre sélection. Dans le protocole No-Scar la contre sélection s'opère toujours contre le chromosome. Cependant, lors de l'étape de recyclage le protocole No-Scar est capable de cibler un plasmide, le pCas9-cr4.

Dans notre stratégie le plasmide pKD médie d'abord sa propre dégradation, puis, après recombinaison par un ADN donneur, il perd cette capacité suicide pour gagner la capacité de cibler le chromosome. La stratégie entend que seules les cellules ayant reçue deux types d'ADN donneurs peuvent survivre à la contre sélection : les cellules doivent recombiner le pKD pour conserver la résistance codée par le pKD ET recombiner sur le chromosome pour survivre à la contre sélection sur le chromosome. Cette condition de survivre à l'évènement pour être comparé à une porte logique AND.

Cette idée de déplacement un système suicide se suggère par la possibilité de changer la séquence d'un espaceur d'un pKD donné en introduisant un ADN donneur simple brin (Anubis) conçu comme si l'on traitait un chromosome (respect MMR du réplicore, etc.)

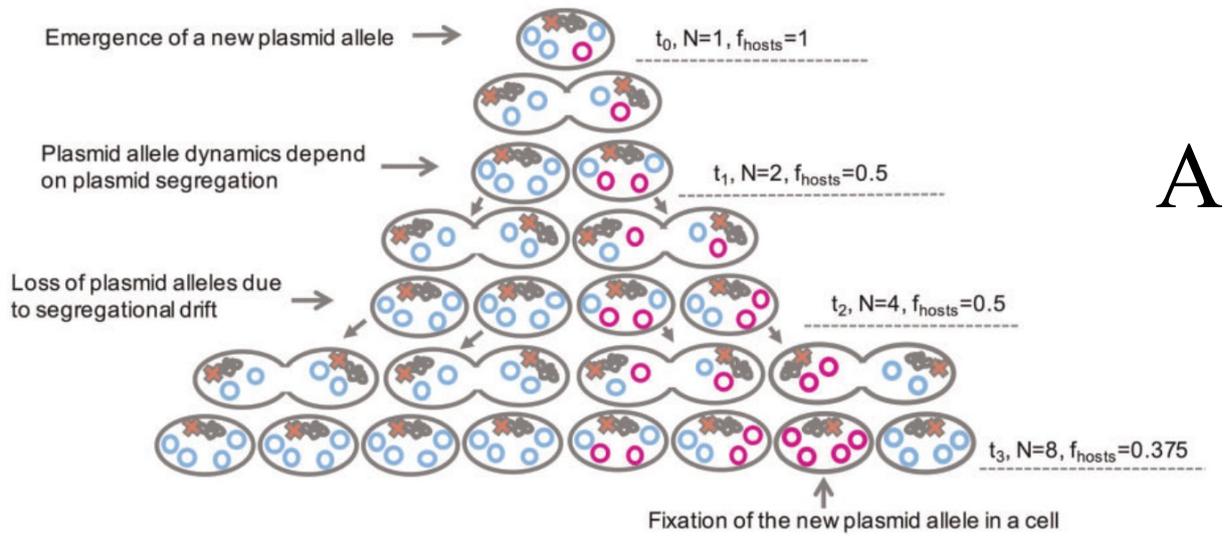
Dans un autre temps, nous avons conçu un pKD, pRaava capable de cibler sa propre origine de réplication (*pSC101*) avec une efficacité de contre sélection extrêmement efficace et comparable celles obtenues avec une cible chromosomique.

Ensemble ces deux éléments permettent de rendre la plateforme No-Scar indépendante de la régénération d'un pKD par CPEC / Assemblage Gibson. En sus de cette indépendance la plateforme est plus stable. La fuite d'expression du promoteur pLTetO1 entraîne toujours une pression de sélection, mais celle-ci est beaucoup moins toxique pour la cellule, car l'évènement n'est pas létal pour la bactérie tant que de la spectinomycine est absente du milieu de culture, le pKD étant stable même en absence de sélection (*Cf. 3.5.3 Hérité du pKD sans sélection*).

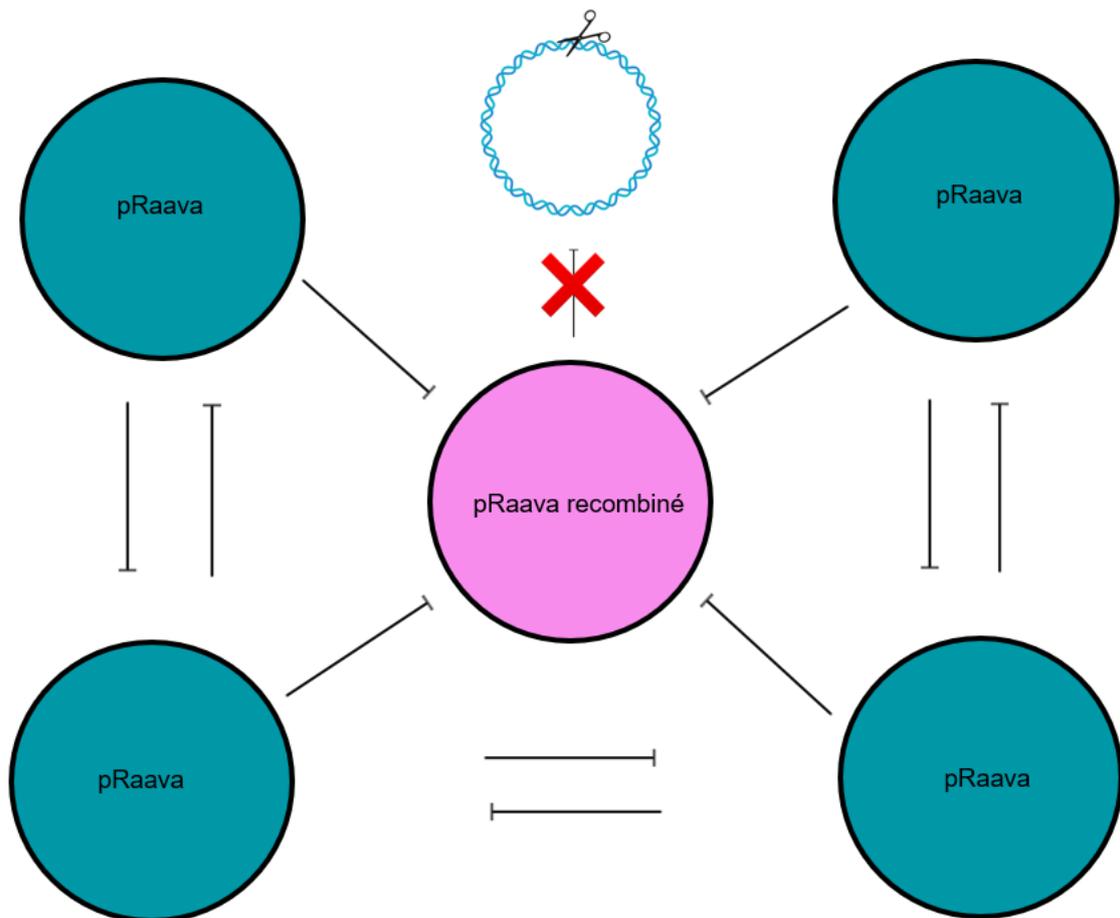
Dans notre premier modèle nous avons sous-estimé le nombre de copies du plasmide pKD. Notre premier modèle était basé que pRaava était maintenu en copie unique. Bien que pRaava fut capable de créer une contre sélection contre lui-même il nous était impossible d'isoler un plasmide contenant le nouvel N-20 après transformation du donneur (Figure 37A).

L'origine *pSC101* est une origine de réplication minimale, l'ori *pSC101* canonique contient un gène, *par* qui assure que le plasmide soit transmis également après la division. pKD46 ne contient pas ce gène, ce qui réduit le nombre de plasmides par cellules. Le pKD du protocole No-Scar se maintiendrait donc entre 3-5 copies par cellules.

La coexistence de plusieurs copies signifie que la ségrégation du pRaava doit être complète à travers les 5 copies du plasmide. Si une copie du pKD non recombinaison demeure, le pRaava recombinaison sera dégradé sans discrimination de l'évènement de recombinaison. Il est possible de prolonger la ségrégation (Figure 37B).



A



B

### Figure 37. Auto-incompatibilité de pRaava

- (A) Co ségrégation d'allèles mutantes introduites à la fois sur un plasmide et le chromosome (en violet), l'auto-incompatibilité de pRaava ne peut être résolue qu'après trois événements de ségrégation ( $t_3$ ). S'il demeure ne serait-ce qu'un plasmide non recombiné (en bleu), pRaava sera dégradé. Une grande portion de la population est alors perdue. Figure adaptée de Tal Dagan *et al.*, 2018.
- (B) Schématisation des allèles du pKD existant après le temps  $t_1$ . Après introduction d'une nouvelle allèle le plasmide pRaava coexiste avec d'autres plasmides pRaava non recombines. Le pKD est maintenu autour de 5 copies par cellule. Même après changement de l'espaceur par un donneur ADNsb pRaava est encore soumis à sa propre contre sélection. La cible originelle de pRaava vise sa propre origine de répllication *pSC101*. Même après recombinaison sur le plasmide, cette cible est toujours présente amenant à la perte du pKD recombiné.

Au vu de ces résultats nous avons décidé de créer un espaceur capable de se cibler lui-même ET de capable après recombinaison d'être protégé des espaceurs cohabitants n'ayant pas recombinés (Figure 38). Pour ce faire nous avons tiré profit de la plasticité de la chimère sgARN pour créer une telle cible. Cette chimère est capable plusieurs mutations dans sa séquence pourvu que ces dernières respectent la structure secondaire. Cependant cette construction devait être suffisamment proche de l'espaceur de 20 paires de bases pour pouvoir être recombinée en même temps que l'espaceur. Nous avons tiré parti de la stratégie de conception du mécanisme d'évasion de MMR et BER pour un ADN donneur simple brin pour créer une région plastique de 5 paires de bases.

Bien que cette stratégie soit hypothétiquement viable, la littérature ne rapportait pas d'information sur la capacité létale d'une coupure sur un plasmide dans une séquence non vitale pour la réplication du même plasmide.

Notre premier ARN suicide, pKD<sup>2</sup> nous permis de confirmer cette hypothèse. Bien que la contre sélection fût initialement faible, elle demeurait existante. Après analyse de la séquence de l'espaceur, il s'avérait que celle-ci ne répondait pas aux critères d'un espaceur de bonne efficacité, le %GC était trop faible à 24. L'espaceur était aussi très court, ne contenant que 17 paires de bases. Cette taille fut rapportée comme la taille minimale pour une dégradation significative confirmée *in situ* pour un espaceur. Il est difficile de réellement apprécier l'efficacité d'un espaceur donné sans le tester *in situ* bien que certains algorithmes puissent aiguiller dans l'étape de conception *in silico*.

En prenant en compte les résultats obtenus via notre premier ARN, nous avons conçu de nouvelles itérations, la version finale fut dénommée HorusV74. À notre surprise Horus possédait une efficacité de contre sélection supérieure à celle du pRaava, bien que Horus vise une cible non vitale. Il est possible d'attribuer cette augmentation de l'efficacité depuis la

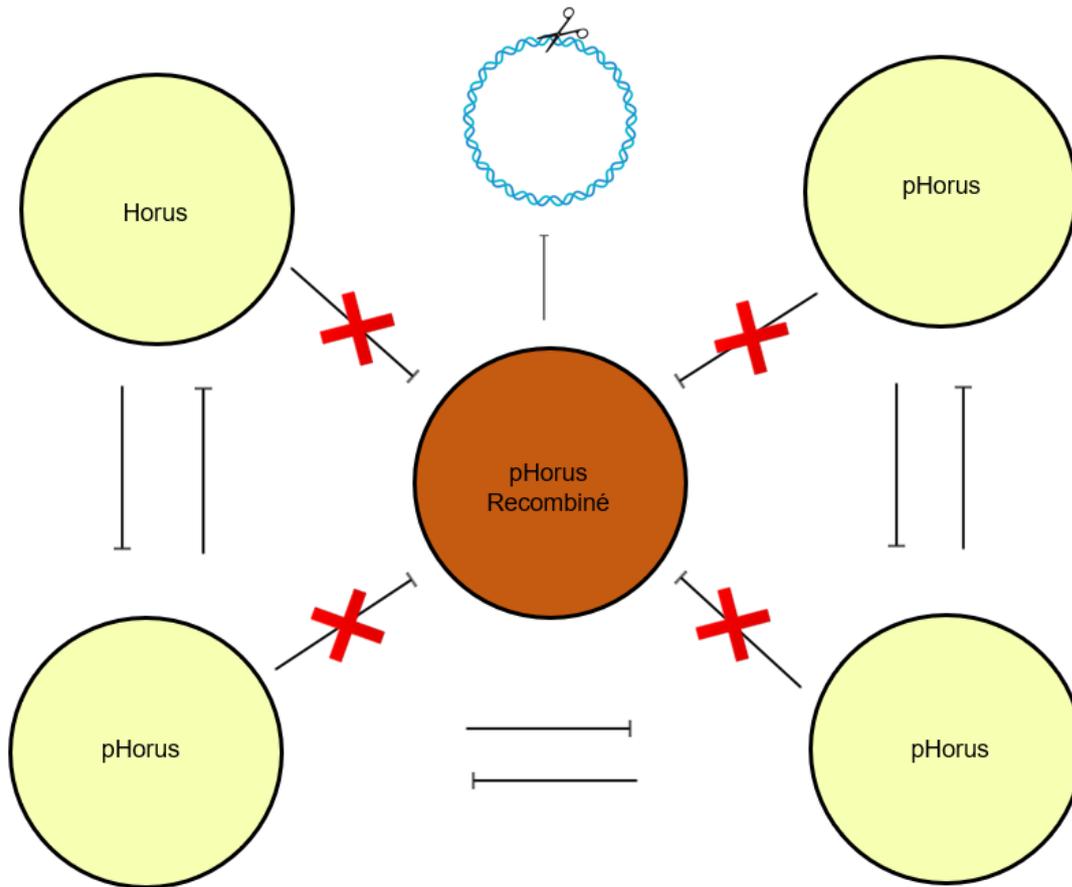
stabilisation de la structure secondaire de l'ARN. En effet, Horus possède une énergie de repliement -20.40 kcal/moL comparé a -17.90kcal/mol de l'ARN du système No-Scar.

La construction Horus peut donc être incluse dans un cycle No-Scar en tant que médiateur de la contre sélection. En employant une seule souche de référence portant, le pCas9-cr4 et le pKD permet *in fine* un meilleur contrôle de la plateforme, en effet plusieurs mutants peuvent être créés de manière parallèle en augmentant le volume de culture employé. Pour *E. coli* il est possible de créer des stocks glycérol -80°C de cellules sensibles à la contre sélection et compétentes pour la réception d'ADN donneurs.

Via Horus il n'est plus nécessaire de créer un pKD pour chaque mutant par CPEC / Assemblage de Gibson. L'étape de CPEC reste laborieuse et cette technique peut se révéler capricieuse du fait de l'homologie de séquence et des risques d'intégration de mutation lors de la PCR sur des longs fragments. Dans un autre temps, il est plus aisé de caractériser la plateforme. Via Horus on emploie toujours les deux mêmes plasmides, ainsi le comportement de la souche, comme le temps de croissance et l'efficacité de contre sélection sont tous deux connus avant de procéder à un cycle NoScar.

Il est donc possible d'estimer quelle dilution sera nécessaire pour obtenir le plus grand ratio mutant / escapers après transformation et récupération des cellules. L'étape de recyclage du pCas9-cr4 peut être aussi facilitée si le N-20 de Horus est rajouté dans la séquence du pCas9-cr4. Cette insertion permettrait alors la reconnaissance de cette cible sur le pCas9-cr4 et le clivage de celui-ci simultanément avec celui des pRaava non recombines lors de l'étape de contre sélection. Par la suite, une simple élimination du pHorus thermosensible serait nécessaire pour s'affranchir totalement de la présence de plasmides. Ainsi via cette insertion il est possible de tirer profit des pHorus non recombines, qui en entraînant leur lyse viendront aussi cliver le pCas9-cr4. Les escapers, cellules ayant désactivées le système par une délétion

au sein du locus du sgARN (que cela soit par une délétion au début du Scaffold ou délétion d'un site TetO2) conserverons eux, un phénotype SmR. D'autre part, cette double contre-sélection permettrait aussi de conférer aux bactéries mutantes (ayant recombinaison sur le pKD et le chromosome) un phénotype unique. En effet par la perte du pCas9-cr4, la souche mutante devient sensible à l'agent de sélection du pCas9-cr4 (originellement du 25 µg/mL de chloramphénicol). Ce point est crucial chez *V. cholerae* en raison de l'efficacité plus modeste de l'opéron Red par rapport à *E. coli*. Ainsi au lieu de cibler un grand nombre de bactéries (plus de 100 colonies) par amplification PCR sur colonies il devient possible de simplement repiquer les colonies sur milieu contenant du chloramphénicol et sur milieu sans antibiotique pour sélectionner les mutants. Il existera néanmoins toujours une sous-population ayant perdu le pCas9-cr4 sans porter la mutation, mais celle-ci représente une part moins importante que les escapers ayant désactivé le système et ayant donc gardé une résistance au chloramphénicol.



**Figure 38. Auto-comptabilité de Horus**

Dans le même contexte que présenté dans la Figure 44, pHorus est capable, après recombinaison de se protéger des autres pKD coexistants n'ayant pas recombinés. Cette protection permet d'augmenter la proportion de cellules capables de survivre à la contre sélection.

Bien qu'un mutant *lacZ* de MG1655 ai été obtenu via la plateforme Horus plusieurs étapes d'optimisation restent nécessaires. À l'heure où ce texte est rédigé, nous manquons encore d'informations précieuses. Par exemple, il serait judicieux de comprendre la dynamique entre le nombre de copies d'un plasmide et la résistance à l'agent de sélection de ce même plasmide. Nous émettons l'hypothèse qu'avec notre raisonnement de contre sélection, le plasmide Horus recombinant et la perte des plasmides non recombinés offre une résistance moindre qu'avec la présence de 5 plasmides par cellules. Le remplacement de la cassette au sein du pKD est aussi un axe intéressant, la spectinomycine, comme mentionné plus haut étant un antibiotique de type bio statique. Un antibiotique de type biocide serait plus intéressant, car il permettrait une élimination plus rapide des clones non recombinés. De plus la littérature rapporte que la spectinomycine n'est pas un antibiotique de choix même pour la sélection d'un vecteur, le nombre de résistants spontanés étant élevé (182). De plus, la concentration de l'antibiotique est différente pour chaque souche et construction.

L'implémentation de la plateforme Horus dans A1552 requiert au minimum d'interchanger les cassettes antibiotiques, A1552 est très sensible au chloramphénicol présentant une CMI de de 2µg/mL. Ainsi si le pKD Horus portait la résistance au chloramphénicol et que le pCas9, possédait, inversement la résistance à la spectinomycine, l'efficacité du système serait grandement améliorée.

L'insertion de deux gènes, *recX* en amont de la Cas9 et *dam* en amont de l'opéron Red serait une suite logique. RecX permet de créer un phénotype RecA transitoire (178). La présence de RecA reste essentielle pendant l'étape de recombinaison par l'opéron Red, mais sa présence devient un inconvénient du moment où une coupure double brin est créé (178). L'ajout de RecX en amont de la Cas9 permettrait donc une expression spatio-temporelle de RecX. Dam quant à lui serait un allié précieux dans la recombinaison, car son expression transitoire permet de bloquer les mécanismes de MMR et BER (178). Le maintien de ce mécanisme est nécessaire

pour éviter d'augmenter le nombre de mutations spontanées, leur répression transitoire pendant la recombinaison permettrait donc d'augmenter l'efficacité de cette dernière sans avoir à subir un génotype hypermutagène.

La plateforme aussi pourrait être adaptée pour du multiplex, mais requière un grand changement dans la plateforme. Pour que Horus soit compatible avec du multiplex il serait nécessaire que le pCas9 porte l'opéron Red. Le pKD linéarisé serait co-transformé avec les deux ADN donneurs simples brins. La plateforme utiliserait ici le mécanisme Gap Repair. Dans cette stratégie seules les bactéries ayant réparée le pKD via les ADN donneurs seront capables de survivre au cycle NoScar.

Une solution complètement alternative à Horus serait la capacité de supprimer totalement le génome d'une bactérie tout en maintenant l'intégrité cellulaire de cette dernière. Après quoi il serait possible de transformer un génome complètement artificiel. Notre plateforme Horus est aussi grandement limitée par l'état de la technologie de synthèse chimique d'ADN.

En réflexion plus lointaine, il serait possible de créer complètement *de novo* une nouvelle porte logique AND sous forme de circuits de contre sélection. Le projet open source Biobrick permet de créer synthétiquement des circuits biologiques. Il est possible de commander des séquences régulatrices (insulateur, promoteur, terminateur séquence RBS). Il est également possible d'accéder à des biobanques contenant des constructions préétablies et validées expérimentalement. La conception de circuits génétiques de contre sélection et éventuellement de régulation n'est pas sans rappeler la conception d'un circuit électronique (183).

## Chapitre 5 : Conclusion

La création de souches mutantes est au cœur de toute étude microbiologie et est cruciale pour élucider les mécanismes de virulence, pathogénicité et résistance de *V. cholerae*. Bien que des méthodes aient été décrites pour générer des mutants pour A1552 ces dernières incluent soit une cassette de résistance (compétence naturelle) et/ou des cicatrices de recombinaisons par l'emploi de SSRs (FLP/FRT).

Par ce projet de maitrise nous avons pu démontrer que le système de contre sélection CRISPR-Cas9 pouvait être implémenté dans *V. cholerae* mais que l'efficacité était grandement réduite comparativement à *E. coli*, nécessitant donc de nombreuses optimisations.

À l'issue de ma Maitrise, nous avons pu prouver qu'il était possible de créer une plateforme CRISPR avec une cytotoxicité réduite et CPEC indépendante. Via ces travaux nous avons pu souligner la littérature existante, mais parcimonieuse sur certains éléments et découvrir de nouvelles informations qui sont applicables au-delà de notre plateforme. Par exemple, la création d'une cible techniquement non létale sur un plasmide (ne ciblant pas le gène de résistance et/ou l'origine de répllication) peut devenir létale pourvu que cette dernière soit suffisamment proche d'au moins l'un de ces deux éléments vitaux. Dans un autre temps nous avons démontré la capacité d'un donneur ADN simple brin à médier plus d'un type de mutations (ici deux permutations).

Nous avons ainsi démontré la fonctionnalité et le gain d'efficacité de contre sélection de notre plateforme Horus comparativement à la plateforme No-Scar native, chez *E. coli* MG1655. Bien que l'efficacité de contre sélection initiale soit supérieure à celle offerte par le système No-Scar notre ARF est plus basse que dans le cycle No-Scar. Cependant en changeant simplement la durée de récupération il nous fut possible d'augmenter notre ARF de 0.6% à 6%. Une

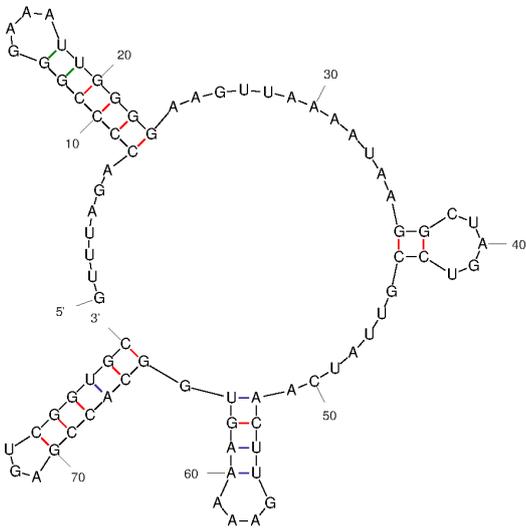
construction similaire (*lacZ*) est actuellement à l'étude chez *V. cholerae* afin de déterminer l'efficacité de notre plateforme chez cette bactérie.

Des améliorations sont encore possibles et implique notamment l'insertion d'une séquence N20 de Horus dans le pCas9-cr4 (élimination facilitée du pCAs-cr4), la modification de la cassette de résistance du pKD par un antibiotique bactéricide (élimination plus rapide des clones non recombinés) et, l'insertion de *recX* en amont de la Cas9 et de *dam* en amont de l'opéron Red (amélioration de l'efficacité de recombinaison).

# Annexes

Output of *trajopt* IR  
IR94\_L01.7

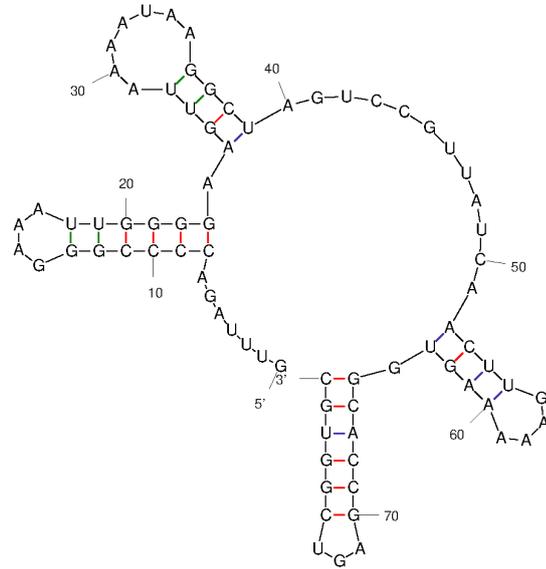
Created Thu Nov 18 20:28:06 2021



dG = -19.50 [Initially -19.50] 21Nov18-20-28-05

Output of *trajopt* IR  
IR94\_L01.7

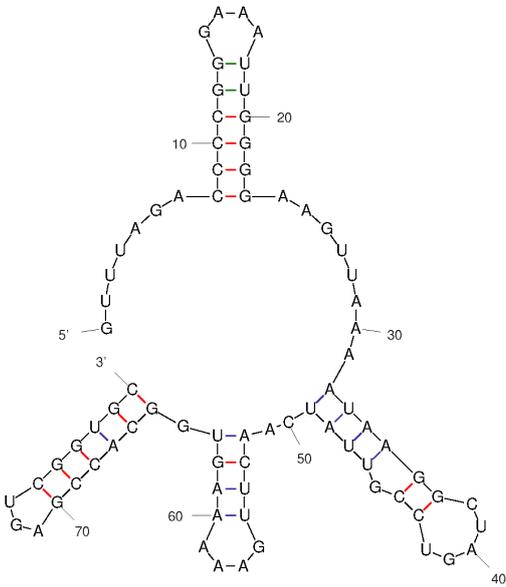
Created Thu Nov 18 20:28:06 2021



dG = -19.60 [Initially -19.30] 21Nov18-20-28-05

Output of *trajopt* IR  
IR94\_L01.7

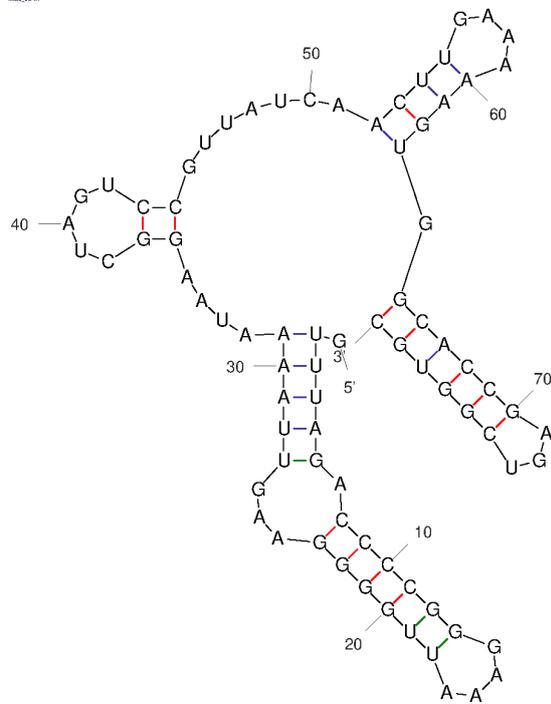
Created Thu Nov 18 20:28:06 2021



dG = -19.10 [Initially -19.10] 21Nov18-20-28-05

Output of *trajopt* IR  
IR94\_L01.7

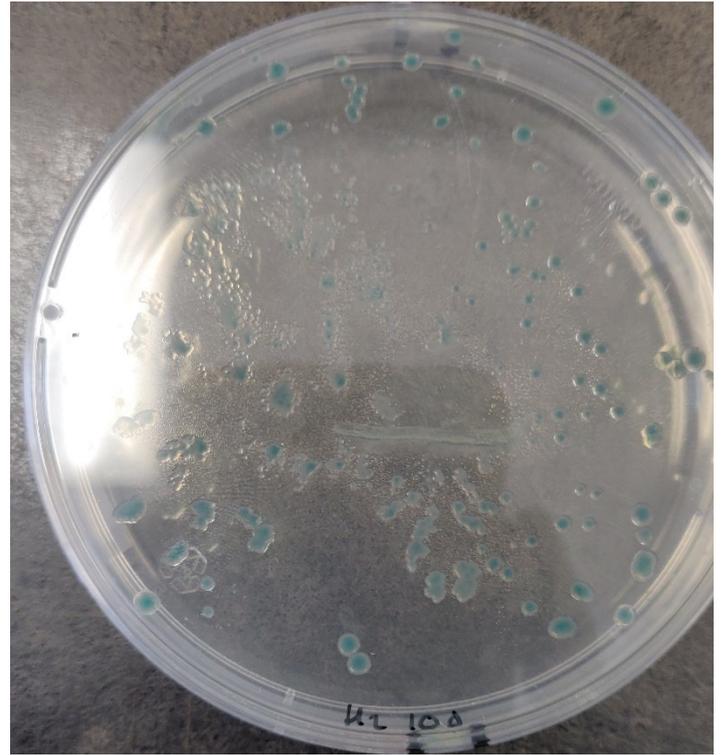
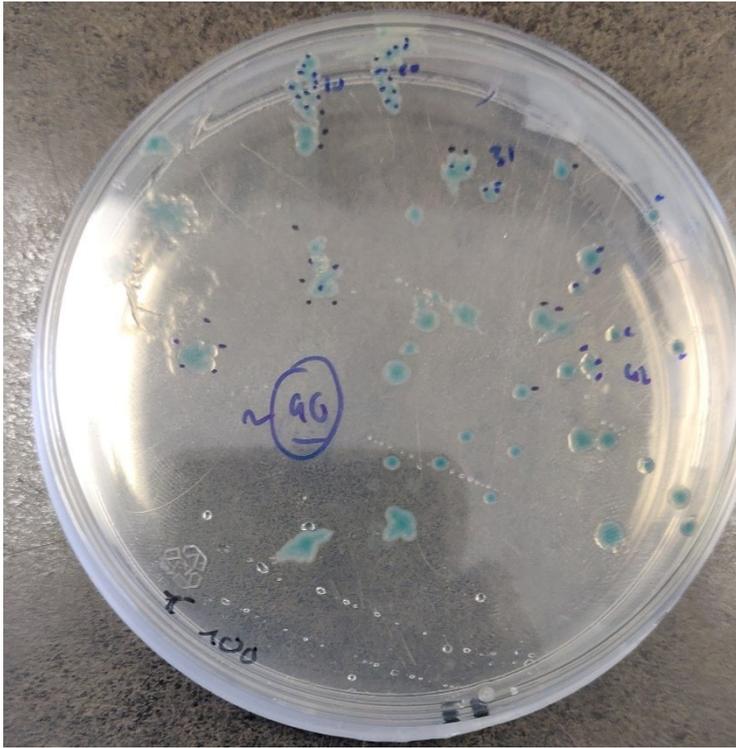
Created Thu Nov 18 20:28:06 2021



dG = -18.70 [Initially -18.70] 21Nov18-20-28-05

### **Figure 39. Repléments illégitimes de HorusV1**

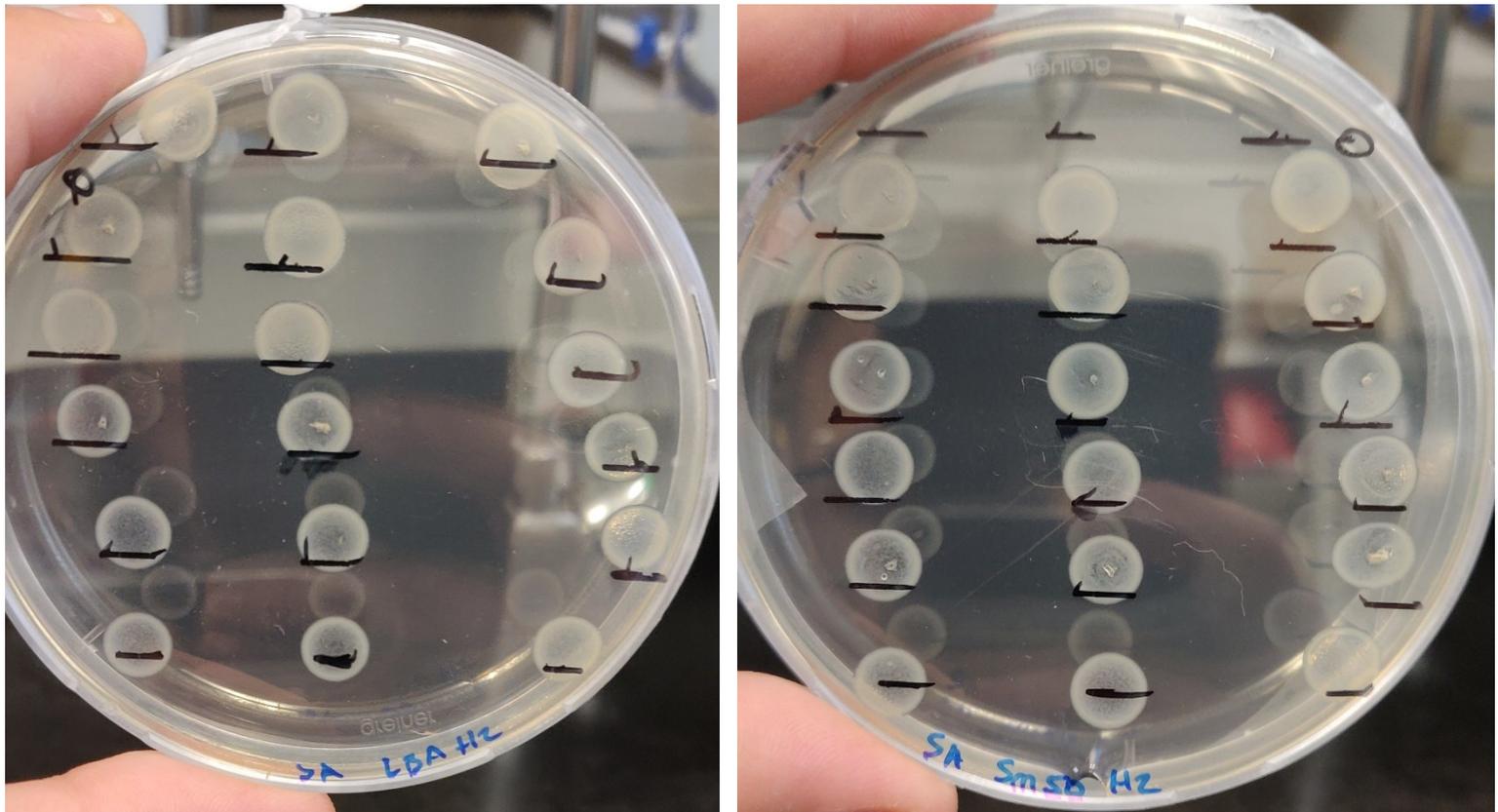
Repléments prédits par mFold. La délétion d'une thymine au début du Scaffold de HorusV1 entraîne un changement conséquent dans la structure de sgARN, cette perte de structure se traduit expérimentalement par une diminution de l'efficacité.



**Figure 40. Colonies satellites**

Photographie de deux boîtes de pétri avec un cycle No-Scar réalisé par Horus Vi1. À gauche, 100 $\mu$ L de la réaction témoin n'ayant pas reçu un ADN pouvant protéger de la contre sélection. Pas de colonies satellites observables

À droite la réaction, ayant reçu l'ADN donneur créant la mutation *lacZ* et l'ADN donneur se recombinant sur le pKD Horus. De nombreuses colonies satellites sont observables Les deux réactions ont été étalées sur milieu contenant 120 $\mu$ g/mL de Spectinomycine et 200 ng/mL d'aTc.



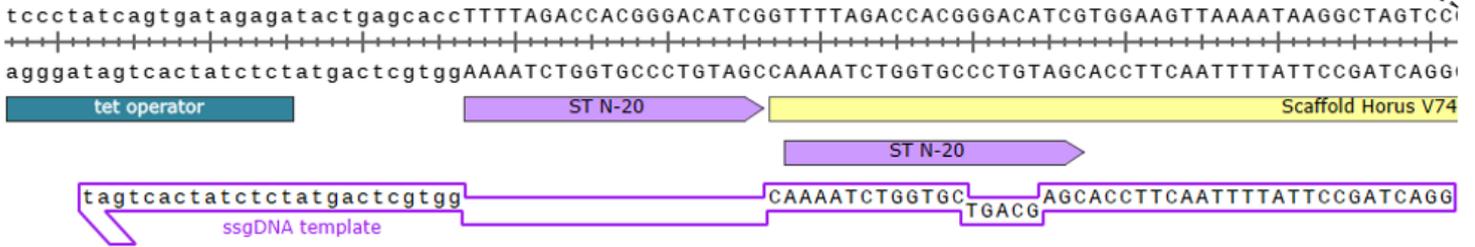
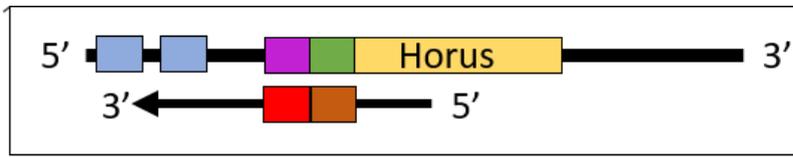
**Figure 41. Hérité du pKD sans sélection**

Une souche contenant le pKD Horus fut cultivée en absence de Spectinomycine, antibiotique sélectionnant pour le pKD. Après une overnight un spot-assay fut réalisé jusqu'à la dilution  $10^{-5}$  sur milieu sans Spectinomycine et avec Spectinomycine.

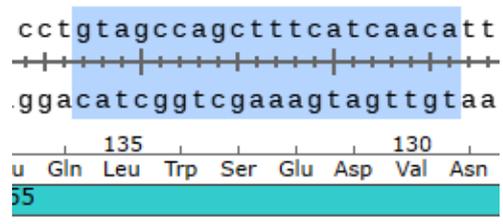
À gauche, milieu LBA sans antibiotique, on observe une croissance jusqu'à la dilution  $10^{-5}$

À droite, milieu LBA avec  $50\mu\text{g/mL}$  de Spectinomycine la croissance est identique au milieu sans sélection.

A

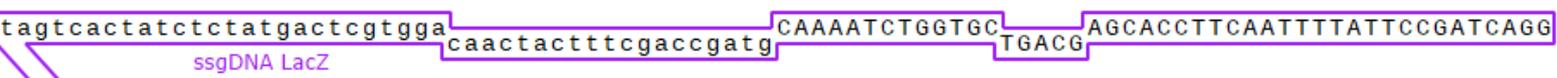


B



**Séquence de la cible sur le brin - :**

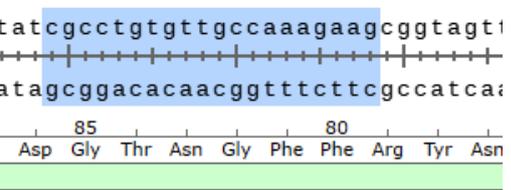
5'-tgttgatgaaagctggctac-3'



**Séquence de l'ADN recombinant sur le pKD correspondant :**

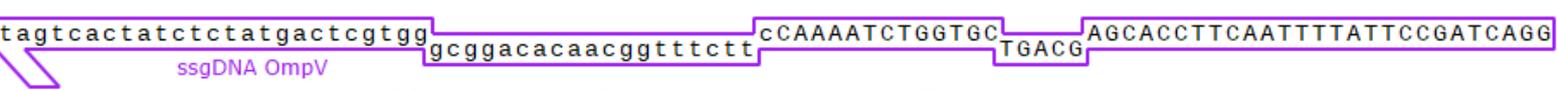
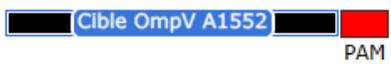
5'-ggactagccttatttaactccacgaGCAGTcgtggtctaaaacttctttggcaacacaggcgggtgctcagtatctctatcactgat -3'

C



**Séquence de la cible sur le brin + :**

5'-cgctgtgttgccaaagaag-3'



**Séquence de l'ADN recombinant sur le pKD correspondant :**

5'-ggactagccttatttaactccacgaGCAGTcgtggtctaaaacttctttggcaacacaggcgggtgctcagtatctctatcactgat -3'

**Figure 42. Conception des ADN donneurs simple brins pour HorusV74**

- A. En haut, schématisation de l'évènement de recombinaison, en bleu les sites régulateurs TetO2, en violet la cible suicide, en vert la région plastique originelle. En rouge la nouvelle cible apportée par le donneur ADN, en marron la région plastique alternative pour protection. En bas, séquence du locus HorusV74 et de l'ADN donneur cible (violet), sans espaceur.
- B. Conception d'un ssgADN donneur pour une cible située sur le brin **négatif**. Les nucléotides en majuscules soulignés correspondent à la région plastique alternative. Les nucléotides en gras correspondent au nouvel espaceur.
- C. Conception d'un ssgADN donneur pour une cible située sur le brin **positif**. Les nucléotides en majuscules soulignés correspondent à la région plastique alternative. Les nucléotides en gras correspondent au nouvel espaceur.

# Bibliographie

1. Eickbush T, Malik H. Mobile DNA II, Vol. 49. ASM Press, Washington, DC, USA; 2002.
2. Hickman AB DF. The casposon-encoded Cas1 protein from *Aciduliprofundum boonei* is a DNA integrase that generates target site duplications. . Nucleic Acids Res. 2015;43:10576-87.
3. Clegg MT, Durbin ML. Flower color variation: a model for the experimental study of evolution. Proc Natl Acad Sci U S A. 2000;97(13):7016-23.
4. Babakhani S, Oloomi M. Transposons: the agents of antibiotic resistance in bacteria. J Basic Microbiol. 2018;58(11):905-17.
5. Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata A. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. J Bacteriol. 1987;169(12):5429-33.
6. Ishino Y, Krupovic M, Forterre P. History of CRISPR-Cas from Encounter with a Mysterious Repeated Sequence to Genome Editing Technology. J Bacteriol. 2018;200(7).
7. Wenyuan Han QS. CRISPR History: Discovery Characterization, and Prosperity. Progressin Molecular BiologyandTranslational Science. 2017;152:1877-173.
8. Hermans PW vSD, Bik EM, de Haas PE, Dale JW, van Embden JD. . Insertion element IS987 from *Mycobacteriumbovis* BCG is located in a hot-spot integration region for insertion elements in *Mycobacterium tuberculosis* complex strains. . Infect Immun. 1991(59):2695–705.
9. Nakata A, Amemura M, Makino K. Unusual nucleotide arrangement with repeated sequences in the *Escherichia coli* K-12 chromosome. J Bacteriol. 1989;171(6):3553-6.
10. Newburg SF SN, Robinson EC, Hiles IE, Higgins CF. Stabilization of translationally active mRNA by prokaryotic REP sequences. . Cell 1987;48:297–310.
11. She Q, Phan H, Garrett RA, Albers SV, Stedman KM, Zillig W. Genetic profile of pNOB8 from Sulfolobus: the first conjugative plasmid from an archaeon. Extremophiles. 1998;2(4):417-25.
12. Mojica MJ JG, Rodríguez-Valera F. Transcription at different salinities of *Haloferox mediterranei* sequences adjacent to partially modified PstI sites. Mol Microbiol. 1993.;9:613–21.
13. She Q, Singh RK, Confalonieri F, Zivanovic Y, Allard G, Awayez MJ, et al. The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. Proc Natl Acad Sci U S A. 2001;98(14):7835-40.
14. Jansen R EJ, Gaastra W, Schouls LW. Identification of genes that are associated with DNA repeats in prokaryotes. Mol Microbiol 2002;43:1565–75.
15. Makarova KS, Aravind L, Grishin NV, Rogozin IB, Koonin EV. A DNA repair system specific for thermophilic Archaea and bacteria predicted by genomic context analysis. Nucleic Acids Res. 2002;30(2):482-96.
16. Tang TH PN, Zywicki M, Huber H, Brugger K, Garrett R, Bachellerie JP., A. H. Identification of novel non-coding RNAs as potential antisense regulators in the archaeon *Sulfolobussolfataricus*. . Mol Microbiol. 2005;55:469-81.
17. Lillestol RK SS, Brugger K, Redder P, Phan H, Christiansen J, Garrett RA. CRISPR families of the crenarchaeal genus Sulfolobus: bidirectional transcription and dynamic properties. . Mol Microbiol. 2009;72(259-272).
18. F.D. Urnov EJR, M.C. Holmes, H.S. Zhang, P.D. Gregory. Genome editing with engineered zinc finger nucleases. Nat Rev Genet. 2010; 11 636-46.
19. Bibikova M GM, Golic KG, Carroll D. . Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases. Genetics. 2002(161):1169-75.
20. Bibikova M, Beumer K, Trautman JK, Carroll D. Enhancing gene targeting with designed zinc finger nucleases. Science. 2003;300(5620):764.
21. Chandrasegaran S, Carroll D. Origins of Programmable Nucleases for Genome Engineering. J Mol Biol. 2016;428(5 Pt B):963-89.

22. Haft DH, Selengut J, Mongodin EF, Nelson KE. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Comput Biol.* 2005;1(6):e60.
23. Makarova KS, Shabalina SA, Wolf YI, Koonin EV. A putative RNA interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol Direct.* 2006(1):7.
24. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, et al. CRISPR provides acquired resistance against viruses in prokaryotes. *Science.* 2007;315(5819):1709-12.
25. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science.* 2012;337(6096):816-21.
26. Katherine E. Uyhazi JB. A CRISPR view of the 2020 Nobel Prize in Chemistry. *J Clin Invest.* 2020;131:e145214.
27. Shmakov S, Smargon A, Scott D, Cox D, Pyzocha N, Yan W, et al. Diversity and evolution of class 2 CRISPR-Cas systems. *Nat Rev Microbiol.* 2017;15(3):169-82.
28. Burstein D, Strutt SC, Probst AJ, Anantharaman K, Thomas BC, Doudna JA, Banfield JF. New CRISPR-Cas systems from uncultivated microbes. *Nature* 542:237–241. *Nature.* 2017;542:237-41.
29. C. V. Structure of Csm2 elucidates the relationship between small subunits of CRISPR-Cas effector complexes. *FEBS Lett.* 2016;590:1521-9.
30. Makarova KS, Wolf YI, Alkhnbashi OS, Costa F, Shah SA, Saunders SJ, et al. An updated evolutionary classification of CRISPR-Cas systems. *Nat Rev Microbiol.* 2015;13(11):722-36.
31. Garrett RA, Vestergaard G, Shah SA. Archaeal CRISPR-based immune systems: exchangeable functional modules. *Trends Microbiol.* 2011;19(11):549-56.
32. Lee H, Zhou Y, Taylor DW, Sashital DG. Cas4-Dependent Prespacer Processing Ensures High-Fidelity Programming of CRISPR Arrays. *Mol Cell.* 2018;70(1):48-59 e5.
33. Hou Z, Zhang Y, Propson NE, Howden SE, Chu LF, Sontheimer EJ, et al. Efficient genome engineering in human pluripotent stem cells using Cas9 from *Neisseria meningitidis*. *Proc Natl Acad Sci U S A.* 2013;110(39):15644-9.
34. Krupovic M, Forterre P, Prangishvili D, Koonin EV. Casposons: a new superfamily of self-synthesizing DNA transposons at the origin of prokaryotic CRISPR-Cas immunity. *BMC Biology* 2014;12:36.
35. Hickman AB, Dyda F. Mechanisms of DNA Transposition. *Microbiol Spectr.* 2015;3(2):MDNA3-0034-2014.
36. Beguin P, Charpin N, Koonin EV, Forterre P, Krupovic M. Casposon integration shows strong target site preference and recapitulates protospacer integration by CRISPR-Cas systems. *Nucleic Acids Res.* 2016;44(21):10367-76.
37. Krupovic M, Beguin P, Koonin EV. Casposons: mobile genetic elements that gave rise to the CRISPR-Cas adaptation machinery. *Curr Opin Microbiol.* 2017;38:36-43.
38. Krupovic M, Koonin EV. Self-synthesizing transposons: unexpected key players in the evolution of viruses and defense systems. *Curr Opin Microbiol.* 2016;31:25-33.
39. Krupovic M, Shmakov S, Makarova KS, Forterre P, Koonin EV. Recent Mobility of Casposons, Self-Synthesizing Transposons at the Origin of the CRISPR-Cas Immunity. *Genome Biol Evol.* 2016;8(2):375-86.
40. Makarova KS, Wolf YI, van der Oost J, Koonin EV. Prokaryotic homologs of Argonaute proteins are predicted to function as key components of a novel system of defense against mobile genetic elements. *Biol Direct.* 2009;4:29.
41. Koonin EV, Makarova KS. Origins and evolution of CRISPR-Cas systems. *Philos Trans R Soc Lond B Biol Sci.* 2019;374(1772):20180087.
42. Koonin EV, Krupovic M. Evolution of adaptive immunity from transposable elements combined with innate immune systems. *Nat Rev Genet.* 2015;16(3):184-92.

43. Clare Rollie SS, Anna Sophie Brinkmann, Edward L Bolt, Malcolm F White. Intrinsic sequence specificity of the Cas1 integrase directs new spacer acquisition. *eLife* 2015;eLife 2015;4:e08716.
44. Yosef I, Goren MG, Qimron U. Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. *Nucleic Acids Res.* 2012;40(12):5569-76.
45. Anantharaman V, Makarova KS, Burroughs AM, Koonin EV, Aravind L. Comprehensive analysis of the HEPN superfamily: identification of novel roles in intra-genomic conflicts, defense, pathogenesis and RNA processing. *Biol Direct.* 2013;8(1):15.
46. Makarova KS, Wolf YI, Koonin EV. Comprehensive comparative-genomic analysis of type 2 toxin-antitoxin systems and related mobile stress response systems in prokaryotes. *Biol Direct.* 2009;4(1):19.
47. Makarova KS, Haft DH, Barrangou R, Brouns SJ, Charpentier E, Horvath P, et al. Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol.* 2011;9(6):467-77.
48. Chylinski K, Le Rhun A, Charpentier E. The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems. *RNA Biol.* 2013;10(5):726-37.
49. Chylinski K, Makarova KS, Charpentier E, Koonin EV. Classification and evolution of type II CRISPR-Cas systems. *Nucleic Acids Res.* 2014;42(10):6091-105.
50. Bikard D, Hatoum-Aslan A, Mucida D, Marraffini LA. CRISPR interference can prevent natural transformation and virulence acquisition during in vivo bacterial infection. *Cell Host Microbe.* 2012;12(2):177-86.
51. Edgar R, Qimron U. The *Escherichia coli* CRISPR system protects from lambda lysogenization, lysogens, and prophage induction. *J Bacteriol.* 2010;192(23):6291-4.
52. Jiang W, Bikard D, Cox D, Zhang F, Marraffini LA. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol.* 2013;31(3):233-9.
53. Deveau H, Barrangou R, Garneau JE, Labonte J, Fremaux C, Boyaval P, et al. Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J Bacteriol.* 2008;190(4):1390-400.
54. Mojica FJM, Diez-Villasenor C, Garcia-Martinez J, Almendros C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology (Reading).* 2009;155(Pt 3):733-40.
55. Levy A, Goren MG, Yosef I, Auster O, Manor M, Amitai G, et al. CRISPR adaptation biases explain preference for acquisition of foreign DNA. *Nature.* 2015;520(7548):505-10.
56. Heler R, Wright AV, Vucelja M, Bikard D, Doudna JA, Marraffini LA. Mutations in Cas9 Enhance the Rate of Acquisition of Viral Spacer Sequences during the CRISPR-Cas Immune Response. *Mol Cell.* 2017;65(1):168-75.
57. Redding S, Sternberg SH, Marshall M, Gibb B, Bhat P, Guegler CK, et al. Surveillance and Processing of Foreign DNA by the *Escherichia coli* CRISPR-Cas System. *Cell.* 2015;163(4):854-65.
58. Fineran PC, Charpentier E. Memory of viral infections by CRISPR-Cas adaptive immune systems: acquisition of new information. *Virology.* 2012;434(2):202-9.
59. McGinn J, Marraffini LA. CRISPR-Cas Systems Optimize Their Immune Response by Specifying the Site of Spacer Integration. *Mol Cell.* 2016;64(3):616-23.
60. Heler R, Samai P, Modell JW, Weiner C, Goldberg GW, Bikard D, et al. Cas9 specifies functional viral targets during CRISPR-Cas adaptation. *Nature.* 2015;519(7542):199-202.
61. Hynes AP, Villion M, Moineau S. Adaptation in bacterial CRISPR-Cas immunity can be driven by defective phages. *Nat Commun.* 2014;5(1):4399.
62. Brouns SJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJ, Snijders AP, et al. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science.* 2008;321(5891):960-4.
63. Wiedenheft B, Zhou K, Jinek M, Coyle SM, Ma W, Doudna JA. Structural basis for DNase activity of a conserved protein implicated in CRISPR-mediated genome defense. *Structure.* 2009;17(6):904-12.
64. Dupuis ME, Villion M, Magadan AH, Moineau S. CRISPR-Cas and restriction-modification systems are compatible and increase phage resistance. *Nat Commun.* 2013;4:2087.
65. Silas S, Mohr G, Sidote DJ, Markham LM, Sanchez-Amat A, Bhaya D, et al. Direct CRISPR spacer acquisition from RNA by a natural reverse transcriptase-Cas1 fusion protein. *Science.* 2016;351(6276):aad4234.

66. Carte J, Christopher RT, Smith JT, Olson S, Barrangou R, Moineau S, et al. The three major types of CRISPR-Cas systems function independently in CRISPR RNA biogenesis in *Streptococcus thermophilus*. *Mol Microbiol.* 2014;93(1):98-112.
67. Shivram H, Cress BF, Knott GJ, Doudna JA. Controlling and enhancing CRISPR systems. *Nat Chem Biol.* 2021;17(1):10-9.
68. Weinberger AD, Sun CL, Plucinski MM, Deneff VJ, Thomas BC, Horvath P, et al. Persisting viral sequences shape microbial CRISPR-based immunity. *PLoS Comput Biol.* 2012;8(4):e1002475.
69. Richter C, Dy RL, McKenzie RE, Watson BN, Taylor C, Chang JT, et al. Priming in the Type I-F CRISPR-Cas system triggers strand-independent spacer acquisition, bi-directionally from the primed protospacer. *Nucleic Acids Res.* 2014;42(13):8516-26.
70. Hatoum-Aslan A, Maniv I, Marraffini LA. Mature clustered, regularly interspaced, short palindromic repeats RNA (crRNA) length is measured by a ruler mechanism anchored at the precursor processing site. *Proc Natl Acad Sci U S A.* 2011;108(52):21218-22.
71. Nam KH, Haitjema C, Liu X, Ding F, Wang H, DeLisa MP, et al. Cas5d protein processes pre-crRNA and assembles into a cascade-like interference complex in subtype I-C/Dvulg CRISPR-Cas system. *Structure.* 2012;20(9):1574-84.
72. Carte J, Wang R, Li H, Terns RM, Terns MP. Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. *Genes Dev.* 2008;22(24):3489-96.
73. Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature.* 2011;471(7340):602-7.
74. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science.* 2012;337(6096):816-21.
75. Globy V, Lee SH, Bae T, Kim JS, Joo C. CRISPR/Cas9 searches for a protospacer adjacent motif by lateral diffusion. *EMBO J.* 2019;38(4):e99466.
76. Jones DL, Leroy P, Unoson C, Fange D, Curic V, Lawson MJ, et al. Kinetics of dCas9 target search in *Escherichia coli*. *Science.* 2017;357(6358):1420-4.
77. Beloglazova N, Kuznedelov K, Flick R, Datsenko KA, Brown G, Popovic A, et al. CRISPR RNA binding and DNA target recognition by purified Cascade complexes from *Escherichia coli*. *Nucleic Acids Res.* 2015;43(1):530-43.
78. Sinkunas T, Gasiunas G, Fremaux C, Barrangou R, Horvath P, Siksnys V. Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. *EMBO J.* 2011;30(7):1335-42.
79. Cucinotta FA, Pluth JM, Anderson JA, Harper JV, O'Neill P. Biochemical kinetics model of DSB repair and induction of gamma-H2AX foci by non-homologous end joining. *Radiat Res.* 2008;169(2):214-22.
80. Lin J, Feng M, Zhang H, She Q. Characterization of a novel type III CRISPR-Cas effector provides new insights into the allosteric activation and suppression of the Cas10 DNase. *Cell Discov.* 2020;6(1):29.
81. Sashital DG, Wiedenheft B, Doudna JA. Mechanism of foreign DNA selection in a bacterial adaptive immune system. *Mol Cell.* 2012;46(5):606-15.
82. Wiedenheft B, van Duijn E, Bultema JB, Waghmare SP, Zhou K, Barendregt A, et al. RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. *Proc Natl Acad Sci U S A.* 2011;108(25):10092-7.
83. Staals RH, Zhu Y, Taylor DW, Kornfeld JE, Sharma K, Barendregt A, et al. RNA targeting by the type III-A CRISPR-Cas Csm complex of *Thermus thermophilus*. *Mol Cell.* 2014;56(4):518-30.
84. Marraffini LA, Sontheimer EJ. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science.* 2008;322(5909):1843-5.
85. Hale CR, Zhao P, Olson S, Duff MO, Graveley BR, Wells L, et al. RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell.* 2009;139(5):945-56.
86. Staals RHJ, Agari Y, Maki-Yonekura S, Zhu Y, Taylor DW, van Duijn E, et al. Structure and activity of the RNA-targeting Type III-B CRISPR-Cas complex of *Thermus thermophilus*. *Mol Cell.* 2013;52(1):135-45.

87. Tamulaitis G, Kazlauskienė M, Manakova E, Venclovas C, Nwokeoji AO, Dickman MJ, et al. Programmable RNA shredding by the type III-A CRISPR-Cas system of *Streptococcus thermophilus*. *Mol Cell*. 2014;56(4):506-17.
88. Deng L, Garrett RA, Shah SA, Peng X, She Q. A novel interference mechanism by a type IIIB CRISPR-Cmr module in *Sulfolobus*. *Mol Microbiol*. 2013;87(5):1088-99.
89. Peng W, Feng M, Feng X, Liang YX, She Q. An archaeal CRISPR type III-B system exhibiting distinctive RNA targeting features and mediating dual RNA and DNA interference. *Nucleic Acids Res*. 2015;43(1):406-17.
90. Brussow H, Canchaya C, Hardt WD. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol Mol Biol Rev*. 2004;68(3):560-602, table of contents.
91. Cumby N, Davidson AR, Maxwell KL. The moron comes of age. *Bacteriophage*. 2012;2(4):225-8.
92. Goldberg GW, Jiang W, Bikard D, Marraffini LA. Conditional tolerance of temperate phages via transcription-dependent CRISPR-Cas targeting. *Nature*. 2014;514(7524):633-7.
93. Marraffini LA, Sontheimer EJ. Self versus non-self discrimination during CRISPR RNA-directed immunity. *Nature*. 2010;463(7280):568-71.
94. Baharoglu Z, Mazel D. SOS, the formidable strategy of bacteria against aggressions. *FEMS Microbiol Rev*. 2014;38(6):1126-45.
95. Szczelkun MD, Tikhomirova MS, Sinkunas T, Gasiunas G, Karvelis T, Pschera P, et al. Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. *Proc Natl Acad Sci U S A*. 2014;111(27):9798-803.
96. A. SC. Viruses: unlocking the greatest biodiversity on Earth. *Genome*. 2013;56(10):542-4.
97. Bondy-Denomy J, Pawluk A, Maxwell KL, Davidson AR. Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. *Nature*. 2013;493(7432):429-32.
98. Seed KD, Lazinski DW, Calderwood SB, Camilli A. A bacteriophage encodes its own CRISPR/Cas adaptive response to evade host innate immunity. *Nature*. 2013;494(7438):489-91.
99. Pausch P, Al-Shayeb B, Bisom-Rapp E, Tsuchida CA, Li Z, Cress BF, et al. CRISPR-CasPhi from huge phages is a hypercompact genome editor. *Science*. 2020;369(6501):333-7.
100. Westra ER, Levin BR. It is unclear how important CRISPR-Cas systems are for protecting natural populations of bacteria against infections by mobile genetic elements. *Proc Natl Acad Sci U S A*. 2020;117(45):27777-85.
101. Reisch CR, Prather KLJ. Scarless Cas9 Assisted Recombineering (no-SCAR) in *Escherichia coli*, an Easy-to-Use System for Genome Editing. *Curr Protoc Mol Biol*. 2017;117:31 8 1- 8 20.
102. Reisch CR, Prather KL. The no-SCAR (Scarless Cas9 Assisted Recombineering) system for genome editing in *Escherichia coli*. *Sci Rep*. 2015;5:15096.
103. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-guided human genome engineering via Cas9. *Science*. 2013;339(6121):823-6.
104. Gasiunas G, Barrangou R, Horvath P, Siksnys V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci U S A*. 2012;109(39):E2579-86.
105. Jinek M, Jiang F, Taylor DW, Sternberg SH, Kaya E, Ma E, et al. Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science*. 2014;343(6176):1247997.
106. Garrity C. CRISPR Mechanism. In: University T, editor.: CRISPR/Cas9; 2014.
107. Nishimasu H, Ran FA, Hsu PD, Konermann S, Shehata SI, Dohmae N, et al. Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell*. 2014;156(5):935-49.
108. Zhang XH, Tee LY, Wang XG, Huang QS, Yang SH. Off-target Effects in CRISPR/Cas9-mediated Genome Engineering. *Mol Ther Nucleic Acids*. 2015;4(11):e264.
109. Morgan TH. *The theory of the gene*: New Haven, CT: Yale University Press; 1926.
110. Morgan TH, Alfred H. Sturtevant, Hermann J. Muller, and C. B. Bridges. *The Mechanism of Mendelian Heredity*: New York: Henry Holt and Company.; 1915.
111. Tabery J, Monika Piotrowska, Lindley Darden, . *Molecular Biology Spring 2021 Edition ed: The Stanford Encyclopedia of Philosophy 2021*.
112. Watson JD, Crick FH. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*. 1953;171(4356):737-8.

113. Nestmann ER. Mutagenesis by nitrosoguanidine, ethyl methanesulfonate, and mutator gene mutH in continuous cultures of *Escherichia coli*. *Mutat Res.* 1975;28(3):323-30.
114. Stadler LJ. Mutations in Barley Induced by X-Rays and Radium. *Science.* 1928;68(1756):186-7.
115. Ruvkun GB, Ausubel FM. A general method for site-directed mutagenesis in prokaryotes. *Nature.* 1981;289(5793):85-8.
116. Hemsley A, Arnheim N, Toney MD, Cortopassi G, Galas DJ. A simple method for site-directed mutagenesis using the polymerase chain reaction. *Nucleic Acids Res.* 1989;17(16):6545-51.
117. Mullis K, Faloona F, Scharf S, Saiki R, Horn G, Erlich H. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb Symp Quant Biol.* 1986;51 Pt 1:263-73.
118. Hinnebusch J, Tilly K. Linear plasmids and chromosomes in bacteria. *Mol Microbiol.* 1993;10(5):917-22.
119. Johnson TJ, Nolan LK. Plasmid replicon typing. *Methods Mol Biol.* 2009;551:27-35.
120. Bills TD, Marking LL, Chandler Jr JH. Formalin: its toxicity to nontarget aquatic organisms, persistence, and counteraction. Report. 1977. Report No.: 73.
121. Raleigh EA, Low KB. Conjugation. In: Maloy S, Hughes K, editors. *Brenner's Encyclopedia of Genetics.* San Diego: Academic Press; 2013. p. 144-51.
122. Lang S, Gruber CJ, Raffl S, Reisner A, Zechner EL. Common requirement for the relaxosome of plasmid R1 in multiple activities of the conjugative type IV secretion system. *J Bacteriol.* 2014;196(11):2108-21.
123. Seol JH, Shim EY, Lee SE. Microhomology-mediated end joining: Good, bad and ugly. *Mutat Res.* 2018;809:81-7.
124. Liberi G, Foiani M. The double life of Holliday junctions. *Cell Res.* 2010;20(6):611-3.
125. Hood L, Galas D. The digital code of DNA. *Nature.* 2003;421(6921):444-8.
126. Rakowski SA, Filutowicz M. Plasmid R6K replication control. *Plasmid.* 2013;69(3):231-42.
127. Ortiz-Martin I, Macho AP, Lambersten L, Ramos C, Beuzon CR. Suicide vectors for antibiotic marker exchange and rapid generation of multiple knockout mutants by allelic exchange in Gram-negative bacteria. *J Microbiol Methods.* 2006;67(3):395-407.
128. Landy A. Dynamic, structural, and regulatory aspects of lambda site-specific recombination. *Annu Rev Biochem.* 1989;58(1):913-49.
129. Sauer B. Functional expression of the cre-lox site-specific recombination system in the yeast *Saccharomyces cerevisiae*. *Mol Cell Biol.* 1987;7(6):2087-96.
130. Sadowski PD. The Flp recombinase of the 2-microns plasmid of *Saccharomyces cerevisiae*. *Prog Nucleic Acid Res Mol Biol.* 1995;51:53-91.
131. Nern A, Pfeiffer BD, Svoboda K, Rubin GM. Multiple new site-specific recombinases for use in manipulating animal genomes. *Proc Natl Acad Sci U S A.* 2011;108(34):14198-203.
132. Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A.* 2000;97(12):6640-5.
133. Blakely GW, Sherratt DJ. Interactions of the site-specific recombinases XerC and XerD with the recombination site dif. *Nucleic Acids Res.* 1994;22(25):5613-20.
134. Das B, Kumari R, Pant A, Sen Gupta S, Saxena S, Mehta O, et al. A novel, broad-range, CTXPhi-derived stable integrative expression vector for functional studies. *J Bacteriol.* 2014;196(23):4071-80.
135. Greenfield TJ, Ehli E, Kirshenmann T, Franch T, Gerdes K, Weaver KE. The antisense RNA of the par locus of pAD1 regulates the expression of a 33-amino-acid toxic peptide by an unusual mechanism. *Mol Microbiol.* 2000;37(3):652-60.
136. Hayes F. Toxins-antitoxins: plasmid maintenance, programmed cell death, and cell cycle arrest. *Science.* 2003;301(5639):1496-9.
137. Meng G, Futterer K. Structural framework of fructosyl transfer in *Bacillus subtilis* levansucrase. *Nat Struct Biol.* 2003;10(11):935-41.
138. Li XT, Thomason LC, Sawitzke JA, Costantino N, Court DL. Positive and negative selection using the tetA-sacB cassette: recombineering and P1 transduction in *Escherichia coli*. *Nucleic Acids Res.* 2013;41(22):e204.

139. Lederberg EM, editor Lysogenicity in *E-coli* K-12. Genetics; 1951: 428 EAST PRESTON ST, BALTIMORE, MD 21202.
140. Pant A, Das B, Bhadra RK. CTX phage of *Vibrio cholerae*: Genomics and applications. *Vaccine*. 2020;38 Suppl 1:A7-A12.
141. Lewis D, Le P, Zurla C, Finzi L, Adhya S. Multilevel autoregulation of lambda repressor protein CI by DNA looping in vitro. *Proc Natl Acad Sci U S A*. 2011;108(36):14807-12.
142. Keen EC. Paradigms of pathogenesis: targeting the mobile genetic elements of disease. *Front Cell Infect Microbiol*. 2012;2:161.
143. Schindler D, Echols H. Retroregulation of the int gene of bacteriophage lambda: control of translation completion. *Proc Natl Acad Sci U S A*. 1981;78(7):4475-9.
144. Unger RC, Clark AJ. Interaction of the recombination pathways of bacteriophage lambda and its host *Escherichia coli* K12: effects on exonuclease V activity. *J Mol Biol*. 1972;70(3):539-48.
145. Murphy KC. Use of bacteriophage lambda recombination functions to promote gene replacement in *Escherichia coli*. *J Bacteriol*. 1998;180(8):2063-71.
146. Wannier TM, Ciaccia PN, Ellington AD, Filsinger GT, Isaacs FJ, Javanmardi K, et al. Recombineering and MAGe. *Nature Reviews Methods Primers*. 2021;1(1):7.
147. Wang HH, Church GM. Multiplexed genome engineering and genotyping methods applications for synthetic biology and metabolic engineering. *Methods Enzymol*. 2011;498:409-26.
148. Costantino N, Court DL. Enhanced levels of lambda Red-mediated recombinants in mismatch repair mutants. *Proc Natl Acad Sci U S A*. 2003;100(26):15748-53.
149. Yu T, Bao X, Piao W, Peng J, Li W, Yang C, et al. Recent patents on oligonucleotide synthesis and gene synthesis. *Recent Pat DNA Gene Seq*. 2012;6(1):10-21.
150. Ali M, Nelson AR, Lopez AL, Sack DA. Updated global burden of cholera in endemic countries. *PLoS Negl Trop Dis*. 2015;9(6):e0003832.
151. Trucksis M, Michalski J, Deng YK, Kaper JB. The *Vibrio cholerae* genome contains two unique circular chromosomes. *Proc Natl Acad Sci U S A*. 1998;95(24):14464-9.
152. Bruhn M, Schindler D, Kemter FS, Wiley MR, Chase K, Koroleva GI, et al. Functionality of Two Origins of Replication in *Vibrio cholerae* Strains With a Single Chromosome. *Front Microbiol*. 2018;9(2932):2932.
153. Burrus V, Quezada-Calvillo R, Marrero J, Waldor MK. SXT-related integrating conjugative element in New World *Vibrio cholerae*. *Appl Environ Microbiol*. 2006;72(4):3054-7.
154. Roy D, Huguet KT, Grenier F, Burrus V. IncC conjugative plasmids and SXT/R391 elements repair double-strand breaks caused by CRISPR-Cas during conjugation. *Nucleic Acids Res*. 2020;48(16):8815-27.
155. Ghosh C, Nandy RK, Dasgupta SK, Nair GB, Hall RH, Ghose AC. A search for cholera toxin (CT), toxin coregulated pilus (TCP), the regulatory element ToxR and other virulence factors in non-01/non-0139 *Vibrio cholerae*. *Microb Pathog*. 1997;22(4):199-208.
156. Marvig RL, Blokesch M. Natural transformation of *Vibrio cholerae* as a tool--optimizing the procedure. *BMC Microbiol*. 2010;10:155.
157. Suckow G, Seitz P, Blokesch M. Quorum sensing contributes to natural transformation of *Vibrio cholerae* in a species-specific manner. *J Bacteriol*. 2011;193(18):4914-24.
158. Karlsson SL, Thomson N, Mutreja A, Connor T, Sur D, Ali M, et al. Retrospective Analysis of Serotype Switching of *Vibrio cholerae* O1 in a Cholera Endemic Region Shows It Is a Non-random Process. *PLoS Negl Trop Dis*. 2016;10(10):e0005044.
159. Yamamoto S, Izumiya H, Morita M, Arakawa E, Watanabe H. Application of lambda Red recombination system to *Vibrio cholerae* genetics: simple methods for inactivation and modification of chromosomal genes. *Gene*. 2009;438(1-2):57-64.
160. Bowater R, Doherty AJ. Making ends meet: repairing breaks in bacterial DNA by non-homologous end-joining. *PLoS Genet*. 2006;2(2):e8.
161. Jiang Y, Chen B, Duan C, Sun B, Yang J, Yang S. Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system. *Appl Environ Microbiol*. 2015;81(7):2506-14.
162. Das AT, Tenenbaum L, Berkhout B. Tet-On Systems For Doxycycline-inducible Gene Expression. *Curr Gene Ther*. 2016;16(3):156-67.
163. Beck CF, Mutzel R, Barbe J, Muller W. A multifunctional gene (tetR) controls Tn10-encoded tetracycline resistance. *J Bacteriol*. 1982;150(2):633-42.

164. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 2003;31(13):3406-15.
165. Mathieu-Denoncourt A, Giacomucci S, Duperthuy M. The Secretome of *Vibrio cholerae*. 2021.
166. Marinus MG, Morris NR. Isolation of deoxyribonucleic acid methylase mutants of *Escherichia coli* K-12. *J Bacteriol.* 1973;114(3):1143-50.
167. Geier GE, Modrich P. Recognition sequence of the dam methylase of *Escherichia coli* K12 and mode of cleavage of Dpn I endonuclease. *J Biol Chem.* 1979;254(4):1408-13.
168. Haycocks JRJ, Warren GZL, Walker LM, Chlebek JL, Dalia TN, Dalia AB, et al. The quorum sensing transcription factor AphA directly regulates natural competence in *Vibrio cholerae*. *PLoS Genet.* 2019;15(10):e1008362.
169. Doench JG, Fusi N, Sullender M, Hegde M, Vaimberg EW, Donovan KF, et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol.* 2016;34(2):184-91.
170. Fay MM, Lyons SM, Ivanov P. RNA G-Quadruplexes in Biology: Principles and Molecular Mechanisms. *J Mol Biol.* 2017;429(14):2127-47.
171. Park JJ, Dempewolf E, Zhang W, Wang ZY. RNA-guided transcriptional activation via CRISPR/dCas9 mimics overexpression phenotypes in *Arabidopsis*. *PLoS One.* 2017;12(6):e0179410.
172. Meibom KL, Blokesch M, Dolganov NA, Wu CY, Schoolnik GK. Chitin induces natural competence in *Vibrio cholerae*. *Science.* 2005;310(5755):1824-7.
173. Duperthuy M, Sjostrom AE, Sabharwal D, Damghani F, Uhlin BE, Wai SN. Role of the *Vibrio cholerae* matrix protein Bap1 in cross-resistance to antimicrobial peptides. *PLoS Pathog.* 2013;9(10):e1003620.
174. Stutzmann S, Blokesch M. Circulation of a Quorum-Sensing-Impaired Variant of *Vibrio cholerae* Strain C6706 Masks Important Phenotypes. *mSphere.* 2016;1(3).
175. Zhang S, Voigt CA. Engineered dCas9 with reduced toxicity in bacteria: implications for genetic circuit design. *Nucleic Acids Res.* 2018;46(20):11115-25.
176. Kim EJ, Yu HJ, Lee JH, Kim JO, Han SH, Yun CH, et al. Replication of *Vibrio cholerae* classical CTX phage. *Proc Natl Acad Sci U S A.* 2017;114(9):2343-8.
177. White MA, Darmon E, Lopez-Vernaza MA, Leach DRF. DNA double strand break repair in *Escherichia coli* perturbs cell division and chromosome dynamics. *PLoS Genet.* 2020;16(1):e1008473.
178. Ronda C, Pedersen LE, Sommer MO, Nielsen AT. CRMAGE: CRISPR Optimized MAGE Recombineering. *Sci Rep.* 2016;6:19452.
179. Valderrama JA, Kulkarni SS, Nizet V, Bier E. A bacterial gene-drive system efficiently edits and inactivates a high copy number antibiotic resistance locus. *Nat Commun.* 2019;10(1):5726.
180. Corts AD, Thomason LC, Gill RT, Gralnick JA. A new recombineering system for precise genome-editing in *Shewanella oneidensis* strain MR-1 using single-stranded oligonucleotides. *Sci Rep.* 2019;9(1):39.
181. Garriss G. RÔLE ET RÉGULATION DU SYSTÈME DE RECOMBINAISON DES ICES DE LA FAMILLE SXT/R391. *Savoirs UDeS.* 2012.
182. Anderson P. Sensitivity and Resistance to Spectinomycin in *Escherichia coli*. *J Bacteriol.* 1969;100(2):939-47.
183. Micali G, Grilli J, Osella M, Cosentino Lagomarsino M. Concurrent processes set *E. coli* cell division. *Sci Adv.* 2018;4(11):eaau3324.