



**HAL**  
open science

# Clonage, ingénierie et transfert de grands fragments de génomme chez *Bacillus subtilis*

Julien Barret

► **To cite this version:**

Julien Barret. Clonage, ingénierie et transfert de grands fragments de génome chez *Bacillus subtilis*. Microbiologie et Parasitologie. Université de Bordeaux, 2024. Français. NNT : 2024BORD0458 . tel-04926077

**HAL Id: tel-04926077**

**<https://theses.hal.science/tel-04926077v1>**

Submitted on 3 Feb 2025

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

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

THÈSE PRÉSENTÉE  
POUR OBTENIR LE GRADE DE  
**DOCTEUR DE**  
**L'UNIVERSITÉ DE BORDEAUX**

ÉCOLE DOCTORALE SCIENCES DE LA VIE ET DE LA SANTÉ  
SPÉCIALITÉ MICROBIOLOGIE - IMMUNOLOGIE

Julien BARRET

**CLONAGE, INGÉNIERIE ET TRANSFERT DE GRANDS FRAGMENTS DE  
GÉNOME CHEZ *BACILLUS SUBTILIS***

Sous la direction de : Carole LARTIGUE

Unité Mixte de Recherche 1332 Biologie du Fruit et Pathologie INRAE - Bordeaux

Soutenue le 13/12/24

Membres du jury :

M. JULES Matthieu, Professeur, INRAE, AgroParisTech, Univ. Paris-Saclay	Président
M. DANELON Christophe, Professeur, INSA, Toulouse	Rapporteur
M. CAMBRAY Guillaume, Chargé de Recherche, CBS, Montpellier	Rapporteur
M. DOIGNON François, Professeur, Université de Bordeaux	Examineur
Mme LE RHUN, Chargée de Recherche, CNRS, Bordeaux	Examinatrice
Mme PLANSON Anne-Gaëlle, Chargée de Recherche, INRAE, AgroParisTech, Univ. Paris-Saclay	Examinatrice
Mme LARTIGUE Carole, Chargée de Recherche, INRAE, Villenave d'Ornon	Directrice de thèse



# Résumés



## **Clonage, ingénierie et transfert de grands fragments de génome chez *Bacillus subtilis*.**

L'ingénierie des génomes des micro-organismes est devenue un standard dans les biotechnologies microbiennes. En 2010, des technologies prometteuses de biologie de synthèse utilisant la levure comme plateforme pour l'assemblage et l'ingénierie de génomes synthétiques bactériens suivi de leur transplantation dans une cellule receveuse ont vu le jour. Ces technologies ont conduit à la création des premières cellules synthétiques et ouvert de nouvelles voies vers la construction de cellules aux propriétés biologiques entièrement contrôlées. Le transfert de ces outils à des micro-organismes d'intérêt industriel comme la bactérie Gram+ *Bacillus subtilis* (*Bsu*), modèle dans le secteur des biotechnologies, constituerait une avancée majeure. C'est précisément l'objectif du projet ANR « Bacillus 2.0 », qui réunit deux équipes INRAE et qui se propose d'adapter l'ensemble de ces outils de biologie de synthèse à *Bsu* afin d'être en mesure de partir d'une conception, assistée par ordinateur, de génomes semi-synthétiques de *Bsu* jusqu'à l'obtention de nouvelles souches industrielles. Cependant, les premiers travaux réalisés sur ce projet ont montré que le génome entier de *Bsu* ne pouvait pas être cloné et maintenu en l'état dans la levure. Ces résultats risquaient de remettre en question la faisabilité du projet dans son ensemble et en particulier la pertinence d'utiliser la levure comme plateforme d'assemblage du génome semi-synthétique de *Bsu*.

L'objectif de ma thèse a consisté à démontrer que la levure restait un hôte pertinent pour le projet « Bacillus 2.0 ». Elle s'est déclinée en 3 parties. Dans la première partie, une méthode de clonage de génome récemment développée au laboratoire et dénommée CREasPy-Fusion, a progressivement été adaptée à *Bsu*. Les résultats obtenus ont montré (i) le transfert possible d'ADN plasmidique entre protoplastes bactériens et sphéroplastes de levure, (ii) l'efficacité d'un système CRISPR-Cas9 porté par les cellules de levure pour capturer/modifier cet ADN plasmidique pendant la fusion *Bsu*/levure, puis (iii) l'efficacité de ce même système pour capturer des fragments de génome d'une centaine de kb à partir de trois souches différentes. Des observations en microscopie à fluorescence ont également été réalisées et ont mis en évidence deux types d'interactions qui permettraient de passer d'un contact protoplastes/sphéroplastes à un ADN bactérien cloné dans la levure. Dans la seconde partie de ma thèse, la méthode CREasPy-Fusion a été mise à profit pour tenter de cloner de grands fragments du génome de *Bsu* dans la levure. Des fragments génomiques jusqu'à ~1 Mb ont pu être clonés dans la levure, mais leur capture a nécessité l'ajout préalable d'un grand nombre d'ARS sur le génome de *Bsu* pour stabiliser les constructions génétiques. La dernière partie a été l'adaptation de la méthode RAGE à *Bsu*. Cette méthode permet le transfert, non pas d'un génome entier mais de portions de génomes bactériens depuis la levure vers la bactérie à éditer. Une preuve de concept a été réalisée avec l'échange d'un premier fragment de génome de 155 kb par une version réduite de 44 kb.

En conclusion, les travaux réalisés au cours de cette thèse ont montré la pertinence d'utiliser la levure comme plateforme d'ingénierie dans les modifications à grande échelle du génome de *Bsu*. D'une part, nous avons montré que des fragments d'une centaine de kb peuvent être clonés dans la levure, modifiés et transférés dans une cellule receveuse de façon à générer des *Bsu* mutants. Cette stratégie offre une véritable alternative à la transplantation de génome. D'autre part, nous avons montré que de grands fragments du génome de *Bsu* (jusqu'à 1Mb) peuvent également être clonés dans la levure à condition de contenir de nombreux ARS dans leurs séquences. Grâce à ces résultats, le clonage d'un génome réduit de *Bsu* chez la levure est redevenu un objectif réalisable.

**Mots clés :** *Bacillus subtilis*, Clonage de génome dans la levure, Ingénierie de génome, méthode RAGE, Biologie de Synthèse, ARS

**UMR 1332 Biologie du Fruit et Pathologie – Equipe Mollicutes**

Centre INRAE Nouvelle-Aquitaine Bordeaux, Bâtiment IBVM 71 Avenue Edouard Bourloux, CS20032  
33882 Villenave d'Ornon Cedex, France



## **Cloning, engineering and transfer of large genome fragments in *Bacillus subtilis*.**

Genome engineering of microorganisms has become a standard in microbial biotechnology. In 2010, promising synthetic biology technologies using yeast as a platform for the assembly and engineering of synthetic bacterial genomes followed by their transplantation into a recipient cell have emerged. These technologies have led to the creation of the first synthetic cells and opened new avenues towards the construction of cells with fully controlled biological properties. Transferring these tools to microorganisms of industrial interest such as the Gram+ bacterium *Bacillus subtilis* (*Bsu*), a model in the biotechnology sector, would be a major step forward. This is precisely the aim of the ANR "Bacillus 2.0" project, which brings together two INRAE teams and aims to adapt all these synthetic biology tools to *Bsu* so as to be able to go from computer-aided design of semi-synthetic *Bsu* genomes to the production of new industrial strains. However, initial work on this project showed that the entire *Bsu* genome could not be cloned and maintained in yeast in its current state. These results threatened to call into question the feasibility of the entire project and, in particular, the relevance of using yeast as a platform for assembling the semi-synthetic *Bsu* genome.

The goal of my thesis was to demonstrate that yeast remained a relevant host for the Bacillus 2.0 project. It was divided into 3 parts. In the first part, a genome cloning method recently developed in the laboratory, called CREasPy-Fusion, was progressively adapted to *Bsu*. The results obtained showed (i) the possible transfer of plasmid DNA between bacterial protoplasts and yeast spheroplasts, (ii) the efficiency of a CRISPR-Cas9 system carried by yeast cells to capture/modify this plasmid DNA during *Bsu*/yeast fusion, and then (iii) the efficiency of the same system to capture genomic fragments of about a hundred kb from three different strains. Fluorescence microscopy observations were also carried out revealing two types of interaction that would enable the transition from protoplast/spheroplast contact to cloned bacterial DNA in yeast. In the second part of my thesis, the CREasPy-Fusion method was used in an attempt to clone large *Bsu* genome fragments in yeast. Genomic fragments of up to ~1 Mb could be cloned in yeast, but their capture required the prior addition of a large number of ARS to the *Bsu* genome to stabilize the genetic constructs. The final part was the adaptation of the RAGE method to *Bsu*. This method allow the transfer, not of a whole genome, but of portions of bacterial genomes from yeast to the bacteria to be edited. Proof of concept was achieved by exchanging a 155 kb genome fragment with a reduced 44 kb version.

In conclusion, the work carried out during this thesis has shown the relevance of using yeast as an engineering platform for large-scale modifications of the *Bsu* genome. On the one hand, we have shown that fragments of around 100 kb can be cloned in yeast, modified and transferred into a recipient cell to generate *Bsu* mutants. This strategy offers a real alternative to genome transplantation. On the other hand, we have shown that large fragments of the *Bsu* genome (up to 1 Mb) can also be cloned in yeast, provided they contain numerous ARS in their sequences. Thanks to these results, cloning a reduced *Bsu* genome in yeast has once again become an achievable goal.

**Keywords:** *Bacillus subtilis*, In-yeast genome cloning, Genome engineering, RAGE Method, Synthetic Biology, ARS

**UMR 1332 Biologie du Fruit et Pathologie – Equipe Mollicutes**

Centre INRAE Nouvelle-Aquitaine Bordeaux, Bâtiment IBVM 71 Avenue Edouard Bourlaux, CS20032  
33882 Villenave d'Ornon Cedex, France



# Remerciements



## Remerciements

Tout d'abord je tiens à remercier l'ensemble des membres du jury de m'avoir fait l'honneur d'examiner et juger mes travaux. Je tiens à remercier le Professeur Matthieu Jules, Professeur à l'INRAE, AgroParisTech et à l'Université de Paris Saclay, d'avoir accepté d'évaluer mon travail et de me faire l'honneur de prendre part à ce jury.

Je souhaite remercier le Docteur Guillaume Cambray, chargé de recherche au CBS de Montpellier et le Professeur Christophe Danelon, chef d'équipe à TBI-INSA de Toulouse, pour avoir accepté d'examiner et d'évaluer ce manuscrit ainsi que pour prendre part à ce jury en tant que rapporteurs.

Je tiens à remercier le Professeur François Doignon, Professeur de l'Université de Bordeaux, et au CNRS, Laboratoire de Biogenèse membranaire (LBM), merci d'avoir accepté de prendre part à ce jury. Merci également pour vos enseignements depuis le début du Master et pour votre aide et votre disponibilité.

Je remercie, la Docteure Anaïs Le Rhun cheffe d'équipe à l'ARNA de Bordeaux, d'avoir accepté d'examiner mes travaux. Je tiens aussi à vous remercier pour m'avoir suivi tout le long de ma thèse en tant que tutrice.

Enfin, je souhaite remercier la Docteure Anne-Gaëlle Planson, chargée de recherche à l'INRAE dans l'équipe Syber de Micalis, pour avoir accepté de faire partie de ce jury de thèse en tant qu'examinatrice.

Je tiens à remercier la Docteure Carole Lartigue. Merci pour ton aide, pour tout ce que tu m'as appris pendant près de 4 ans, depuis mon stage de M2. Merci pour tout ce que tu as fait pour moi, pour tout le temps que tu m'as consacré et tout ce que tu m'as apporté d'un point de vue professionnel mais aussi personnel. Malgré des hauts et des bas, c'est un honneur pour moi d'avoir été un de tes étudiants et d'avoir pu travailler avec toi. Je suis fier d'avoir appris auprès de la déesse de la transplantation. Je n'aurais pas pu rêver mieux qu'une directrice de thèse basque.

Je tiens également à remercier l'ensemble de nos collaborateurs de l'équipe Syber de Micalis à Jouy-en-Josas pour votre expertise sur *Bacillus subtilis* et pour tous les échanges que nous avons eus au cours de nos réunions.



Je souhaite aussi remercier tous les membres de l'équipe Molli, ce fut un plaisir de partager presque 4 ans avec vous depuis mon arrivée en tant que stagiaire.

Merci au Professeur Pascal Sirand-Pugnet, notre chef d'équipe, pour ton savoir et ton partage de connaissances. Merci aussi pour ces séances de TP que nous avons faites ensemble durant mon monitorat, c'était très enrichissant d'avoir pu les partager avec toi.

Je tiens à remercier le Professeur Alain Blanchard pour son aide, son soutien et ses conseils d'un point de vue scientifique, et surtout pour nos discussions sur la mer, le bateau et le surf, qui m'ont toujours apporté un rayon de soleil dans la journée.

Merci Yoyo pour tout ce que tu m'as appris, que ce soit en science ou autre (oui les avions c'est trop stylé). Merci pour toutes les bonnes blagues douteuses, les kebabs, la piscine et tant d'autres choses. Merci pour tous ces moments qu'on a partagés malgré ton agenda TRÈS rempli. Pour finir je tiens aussi à te remercier pour m'avoir initié à l'escalade. Au plaisir de s'envoyer une session grimpe de temps en temps.

Bon tant qu'on y est avec les chefs, merci GG la cheffe du bureau. Ces presque 4 ans partagés avec toi dans le même bureau ont été géniaux. Merci pour toute ton aide, pour ta gentillesse et ton soutien. Merci pour tant de choses, mais aussi pour tout ce que tu fais pour la vie de l'équipe, sans toi on serait tous perdus. Merci pour tous les bons moments qu'on a partagés mais aussi pour avoir été un soutien dans les moments un peu plus compliqués. T'avoir rencontrée et avoir pu partager ce bureau magnifiquement repeint avec toi ont été une chance, je n'aurais pas pu tomber dans un meilleur bureau.

Merci à tous les « gens du haut » : MPI, Natalie, Sandrine, Christophe, Sybille, Pascal, Sophie, Xavier, Laure B, Sylvie, Thierry, Jean-Paul, ainsi qu'aux gens des serres : Adelaïde, Laure M, Denis, Marielle, Jean-Saïd, Romain. Bien évidemment merci à notre informaticien hors pair Patrice. Merci aussi à Suzanne de nous permettre de travailler dans un environnement décent.

Même si vous êtes un tout petit peu plus loin géographiquement (oui oui il faut monter des escaliers et traverser tout un couloir c'est énorme !) un grand merci aux cheffes de l'EFA, tout particulièrement Annie et Isa. Sans vous, tout serait tellement plus compliqué. Je n'oublierai pas toutes nos rigolades au coin café et, plus rarement certes, au bar aussi.



Je tiens également à remercier ceux qui sont partis en commençant par Martine. Depuis que tu n'es plus au labo, on se rend bien compte de tout ce que tu nous apportais quotidiennement. Mais surtout merci pour tous les bons moments qu'on a passés, les rigolades et les potins au coin café ou au bar. J'espère que tout va bien à la montagne, il me tarde de pouvoir repasser te voir à Arrens-Marsous.

Gab, Toto et Fab je tiens à vous remercier pour tout ce que vous m'avez apporté, autant d'un point de vue professionnel que personnel. Gab, merci de m'avoir formé et de m'avoir donné envie de travailler sur *Bacillus subtilis*, cette petite bactérie qu'on aime tant. Toto, il me tarde qu'on se fasse une session ski et promis je te ramènerai surfer. Fab, en plus de tout ce que tu m'as apporté au laboratoire et dans la vie, je voulais aussi te remercier pour m'avoir motivé et entraîné pour faire un semi-marathon, ça a été une source de détente et de déconnexion si importante pour moi. J'espère vraiment qu'on pourra bientôt s'en refaire un. Même si on se voit moins maintenant c'est toujours un plaisir de se retrouver au bar et au bad.

Bastos et Francesca, merci d'avoir été présents pendant presque toute ma thèse, ce fut un plaisir de partager cette aventure avec vous. Merci Bastos pour ton efficacité au quiz. Merci à Fra pour nous avoir apporté un peu d'Italie à Bordeaux.

J'espère qu'on continuera encore de se faire des bonnes soirées tous ensemble au fêtes d'Arrossa ou à Bayonne, et qu'on pourra se faire Mimizan 2.0 et vite se retrouver tous ensemble autour d'un bon pouletto.

Merci à mes petites stagiaires Chloé et Gabriella. Ce fut un plaisir de travailler avec vous. C'est avec fierté que je continuerai à suivre vos parcours.

Pour en finir avec le labo, je tiens à remercier les petits jeunes qui sont toujours là (et Arthur même si t'es un peu moins jeune). Vous faites une superbe relève pour le labo et une très belle nouvelle génération.

Merci à Patrick pour avoir été mon compatriote de sport, on en a fait pas mal quand même entre le rugby, le squash (objectif 2028), et un peu de bad, de course, de natation et d'escalade. Comme quoi même un anglais peut être sympa, notamment pour se retourner en plein vol.

Merci au Dark Destroyer (aussi connu sous le nom de Julien numéro 2) pour ta connerie et ton humour.



Merci aux petits loulous du haut, Julien (numéro 3), Yorick et Louis pour vos bonnes blagues bien raffinées.

Merci à Arthur le grand maître du jeu même si tu fais que me nerf. Merci pour ta connerie, ton humour et surtout pour nos magnifiques chants en duo.

Merci aux pinecos Marie (toma) et Pipou pour apporter un peu de douceur et de calme dans ce groupe de fous. Bravo Marie pour avoir supporté ma folie dans le bureau et Pipou je sais que ma paillasse est entre de bonnes mains avec toi.

Un grand merci pour tous les bons moments qu'on a passés ensemble, j'espère qu'on en vivra encore de nombreux autres.

Éloignons-nous maintenant de l'INRAE, mais pas trop quand même. Un grand merci à toute la team du Stag and Loar, tout particulièrement à Seb et Jules pour nous accueillir tous les mercredis (et plus si affinité). Les « pimtade (de merde) » resteront à jamais les premiers champions du Quiz.

Pour finir je voudrais remercier ma famille, vous m'avez toujours soutenu dans tout ce que j'ai entrepris et vous avez toujours voulu continuer à comprendre ce que je faisais même si ce n'est pas forcément très facile. Un grand merci à mes parents Gilles et Keiko, vous qui m'avez toujours permis de suivre les études que je voulais et qui m'avez encouragé dans cette voie, c'est en grande partie grâce à vous que j'en suis là.

Merci à mon frère Hugo. Ton soutien a été vraiment très important pour moi (même si ce n'est parfois qu'un message), je suis très content de pouvoir te compter dans ma vie.

Même si j'ai été très occupé ces dernières années, vous avoir près de moi a été une vraie force. Tous les moments passés avec vous ont été des temps de déconnexion du boulot, ou plutôt de reconnexion à la vraie vie, qui m'ont permis de tenir.

Et enfin merci Charlotte, mon havre de paix, d'avoir été là durant toute cette aventure et d'avoir été un soutien sans faille. Merci de me supporter tous les jours depuis maintenant plus de 7 ans. Tu as toujours été présente dans les moments importants, les moments de joie et ceux plus compliqués, sans toi je ne serais sûrement pas arrivé où j'en suis. Sans oublier que c'est toi qui nous ramènes des petits trucs, pas forcément utiles au premier abord, comme une boule de poils et un petit caillou, mais qui m'ont finalement aussi bien aidé à tenir.



# Table des matières



## Table des matières

<b>Remerciements</b> .....	<b>5</b>
<b>Avant-propos</b> .....	<b>17</b>
1. Le projet Bacillus 2.0 .....	18
2. Avancement du projet à mon arrivée au laboratoire .....	19
3. Objectifs de la thèse .....	24
<b>Etude bibliographique</b> .....	<b>26</b>
1. <i>Bacillus subtilis</i> .....	26
1.1. Taxonomie .....	26
1.2. Historique .....	27
1.3. Description générale .....	28
1.4. Compétence naturelle .....	29
1.5. Systèmes de sécrétion.....	32
1.6. Données omics .....	34
1.6.1. Données de séquençage.....	34
1.7. Ingénierie de génome chez <i>B. subtilis</i> .....	37
1.7.1. Méthodes de transformation .....	37
1.7.2. Outils génétiques disponibles .....	39
1.8. Applications en biotechnologie .....	44
2. Biologie de synthèse.....	46
2.1. Historique .....	46
2.2. La quête de la cellule artificielle.....	49
2.3. Les châssis en Biologie de Synthèse .....	53
2.4. La levure comme plateforme d'ingénierie .....	55
2.4.1. Clonage dans la levure.....	55
2.4.1.1. Méthode « classique » .....	58
2.4.1.2. Méthode de TAR-cloning.....	58
2.4.1.3. Méthode CReasPy-Cloning .....	60
2.4.1.4. Méthode de Fusion .....	61
2.4.1.5. Méthode de CReasPy-Fusion.....	61
2.4.2. Outils génétiques disponibles dans la levure .....	62
2.4.2.1. CRISPR : du système de défense aux outils d'édition du génome .....	63
2.4.3. Méthodes pour « sortir » de la levure.....	65
2.4.3.1. Transplantation de génome .....	65
2.4.3.2. Echange de fragment de génomes.....	67
4. Les systèmes de recombinaison génétique.....	69



4.1.	La recombinaison homologue .....	70
4.1.1.	Description générale .....	70
4.1.2.	La recombinaison homologue chez <i>B. subtilis</i> .....	71
4.1.3.	La recombinaison homologue chez la levure <i>S. cerevisiae</i> .....	72
4.2.	Les systèmes de recombinaison site-spécifique.....	74
4.2.1.	Généralités .....	74
4.2.2.	Le systèmes cre/lox .....	77
<b>Matériels et Méthodes .....</b>		<b>82</b>
1.	Souches microbiennes et conditions de cultures.....	82
1.1.	<i>Escherichia coli</i> .....	82
1.2.	<i>Bacillus subtilis</i> .....	82
1.3.	<i>Saccharomyces cerevisiae</i> .....	83
2.	Techniques de biologie moléculaire.....	84
2.1.	Oligonucléotides.....	84
2.2.	Plasmides.....	84
2.2.1.	Plasmides commerciaux ou précédemment construits .....	84
2.2.2.	Plasmides construits au cours du projet .....	87
2.3.	Gibson Assembly .....	96
2.4.	KLD.....	96
2.5.	Mini-préparation .....	96
2.6.	Midi-préparation .....	97
2.7.	Séquençage .....	97
2.7.1.	Séquençage des plasmides.....	97
2.7.1.1.	Séquençage Sanger .....	97
2.7.1.2.	Séquençage des plasmides complets .....	97
2.7.2.	Séquençage de génomes bactériens complets .....	98
2.8.	Extraction d'ADN génomique de <i>B. subtilis</i> .....	98
2.9.	Extraction d'ADN total de <i>S. cerevisiae</i> .....	99
2.10.	Polymerase Chain Reaction (PCR) simplex .....	100
2.11.	PCR multiplex.....	100
2.12.	Production des pansements de recombinaison .....	101
2.12.1.	Pansements de recombinaison pour l'édition chez la levure .....	101
2.12.2.	Pansements de recombinaison aux jonctions de 90 pb.....	101
2.12.3.	Panement de recombinaison aux jonctions composé des éléments levure ARSH4 CEN6 HIS3.....	102
2.13.	Purification des ADN.....	102



2.14.	Quantification des ADN .....	102
2.15.	Digestion des ADN par restriction enzymatique .....	103
2.16.	Électrophorèse sur gel d'agarose .....	103
2.17.	Génomes en bloc d'agarose .....	103
2.17.1.	Préparation des génomes de <i>B. subtilis</i> en bloc d'agarose .....	103
2.17.2.	Préparation des génomes de <i>S. cerevisiae</i> en bloc d'agarose.....	104
2.17.3.	Digestion des blocs d'agarose par restriction enzymatique .....	105
2.18.	Electrophorèse en champs pulsé (PFGE).....	105
2.18.1.	Analyse des transformants de <i>S. cerevisiae</i> qui propagent des fragments de génome de <i>B. subtilis</i> .....	106
2.18.2.	Analyse des transformants de <i>B. subtilis</i> .....	107
3.	Transformation.....	107
3.1.	Transformation d' <i>E. coli</i> .....	107
3.1.1.	Transformation par choc thermique .....	107
3.1.2.	Transformation par électroporation .....	108
3.2.	Transformation de <i>Bacillus subtilis</i> .....	108
3.2.1.	Protocole MG1/MG2 .....	108
3.2.2.	Protocole « simple » (MC).....	109
3.3.	Transformation de <i>S. cerevisiae</i> .....	109
3.3.1.	Transformation à l'acétate de lithium.....	109
3.3.2.	Transformation des sphéroplastes au PEG : méthode de CReasPy-Fusion (fusion avec <i>B. subtilis</i> ) .....	110
4.	Observation en microscopie.....	112
4.1.	Microscopie à fond noir .....	112
4.2.	Microscopie à fluorescence.....	112
4.2.1.	Microscope à Epifluorescence.....	112
4.2.2.	Microscope Confocal .....	113
<b>Chapitre 1 : Adaptation de la méthode de CReasPy-Fusion à <i>B. subtilis</i> .....</b>		<b>116</b>
1.	Introduction.....	116
2.	Travaux réalisés avant mon arrivée au laboratoire.....	120
2.1.	Etape A: Transfert d'un plasmide navette par contact <i>S. cerevisiae/B. subtilis</i> .....	120
2.2.	Etape B : Transfert et modification d'un plasmide navette depuis <i>B. subtilis</i> vers la levure par CReasPy-Fusion .....	122
3.	Etape C : Capture d'un fragment de génome de <i>B. subtilis</i> marqué avec les éléments levures ARSH4 CEN6 HIS3 .....	123
3.1.	Construction des souches <i>B. subtilis</i> MS, FE220 et MGP192 marquées avec les éléments levures ARSH4 CEN6 HIS3.....	124



3.2.	Préparation des levures avec le système CRISPR-Cas9 .....	127
3.3.	Preuve de concept : Capture d'un premier fragment de génome d'une taille de 125 kb à 150 kb pour 3 souches de <i>B. subtilis</i> différentes.....	128
3.4.	Capture d'un deuxième fragment de génome d'une taille supérieure à 250 kb .....	132
4.	Etape D : Capture d'un fragment de génome de <i>B. subtilis</i> non marqué par contre sélection au 5'FOA .....	134
4.1.	Préparation du pansement de recombinaison contenant les éléments levure.....	134
4.2.	Capture d'un premier fragment de génome par contre sélection au 5'FOA .....	135
5.	Etape E : Observation des évènements de fusion cellulaire entre levure et <i>B. subtilis</i> .....	137
5.1.	Observation des évènements de fusion en microscopie à fond noir.....	138
5.2.	Observation des évènements de fusion en microscopie à épifluorescence. ....	139
5.2.1.	Construction de levures fluorescentes.....	140
5.2.2.	Observation du contact <i>B. subtilis</i> /levure en microscopie à épifluorescence .....	142
5.3.	Observation des évènements de fusion en microscopie confocale.....	144
6.	Conclusion .....	146
<b>Chapitre 2 : Utilisation de la méthode de CReasPy-Fusion pour le clonage de fragments de génome de <i>B. subtilis</i> de grande taille dans la levure .....</b>		<b>153</b>
1.	Introduction.....	153
2.	Design d'une souche <i>B. subtilis</i> possédant 13 ARS dans son génome .....	156
2.1.	Choix de la souche de départ .....	156
2.2.	Conception sur ordinateur d'un premier génome modifié à partir du génome de la souche AG686.....	157
2.3.	Construction des souches de <i>B. subtilis</i> AG686 modifiées.....	158
3.	Premier essai de capture d'un fragment de génome de <i>B. subtilis</i> de 536 kb modifié par l'ajout d'ARS .....	159
3.1.	Construction d'une souche témoin avec une seule ARS.....	160
3.2.	Capture du fragment de 536 kb par CReasPy-Fusion.....	160
4.	Essais de capture d'un fragment de génome de <i>B. subtilis</i> de 925 kb par CReasPy-Fusion ...	162
4.1.	Premier essai de capture du fragment de 925 kb .....	162
4.2.	Ajout de l'ARS 404 avec ou sans le marqueur d'auxotrophie pour la lysine.....	164
4.3.	Deuxième essai de capture du fragment de 925 kb par CReasPy-Fusion .....	166
5.	Capture d'un demi-génome de <i>B. subtilis</i> .....	167
6.	Conception sur ordinateur d'un génome modifié de <i>B. subtilis</i> contenant 19 ARS et 3 marqueurs d'auxotrophie .....	168
7.	Conclusion .....	170
<b>Chapitre 3 : Adaptation de la méthode RAGE à <i>B. subtilis</i> .....</b>		<b>176</b>
1.	Introduction.....	176



2.	Etape 2 du RAGE : Modification du fragment 7.2 cloné dans la levure pour obtenir la molécule donneuse pCC1-LYS-cre-specR frgtBsu 7.2 .....	178
2.1.	Production du pansement de recombinaison contenant toutes les séquences nécessaires au RAGE .....	178
2.2.	Edition du fragment 7.2 cloné dans la levure.....	179
3.	Etape 3 du RAGE : Transformation de la molécule donneuse (pCC1-LYS-cre-specR frgtBsu 7.2) dans <i>E. coli</i> EPI 300 .....	180
4.	Etape 4 : Echange de fragments d'ADN chez <i>B. subtilis</i> .....	183
4.1.	Construction de la cellule receveuse à partir de la souche <i>B. subtilis</i> FE220.....	183
4.1.1.	Ajout du premier landing pad « CmR loxm2/66 » dans le génome de <i>B. subtilis</i> FE220 (borne 3').....	184
4.1.2.	Ajout du second landing pad « TetR lox71 » dans le génome de <i>B. subtilis</i> FE220 CmR loxm2/66 et V3 (borne 5') .....	186
4.2.	Première tentative d'échange de fragments via le système cre/lox.....	188
5.	Validation de la fonctionnalité du système cre/lox chez <i>B. subtilis</i> .....	191
5.1.	Construction d'une cellule receveuse pour un échange de marqueurs .....	191
5.2.	Echange de marqueurs via le système cre/lox chez <i>B. subtilis</i> .....	193
6.	Test de la méthode RAGE avec pré-expression de la CRE.....	196
6.1.	Co-transformation de la cellule receveuse FE220 V3 avec le pDR244 Ery et la molécule donneuse (deuxième essai).....	196
6.2.	Transformation de la molécule donneuse dans une cellule receveuse FE220 V3 portant une copie du gène <i>cre</i> inséré à son génome (troisième essai) .....	197
6.3.	Pré-expression de la CRE dans une cellule receveuse FE220 V3 avec le plasmide répliquatif pDR244 Ery .....	199
6.3.1.	Quatrième tentative d'échange de fragments avec une cellule receveuse pré-exprimant la CRE sur le pDR244 Ery .....	200
6.3.2.	Plasmides répliquatifs chez <i>B. subtilis</i> .....	206
6.4.	Pré expression de la CRE avec le plasmide répliquatif le pJIM CRE .....	207
6.5.	Cinquième tentative d'échange de fragments avec une cellule receveuse hébergeant le plasmide répliquatif pJIM CRE .....	207
7.	Conclusion .....	210
	<b>Discussion et conclusion .....</b>	<b>217</b>
1.	La méthode de CREasPy-Fusion .....	219
2.	La méthode RAGE.....	225
3.	Applications des méthodes de CREasPy-Fusion et RAGE .....	229
3.1.	Construction de mutants de <i>B. subtilis</i> ou autres organismes.....	229
3.2.	Construction de cellules contrôlées par des génomes entièrement synthétiques ou semi-synthétiques .....	230
3.2.1.	Construction de cellules contrôlées par des génomes entièrement synthétiques.....	230



3.2.2. Construction de cellules contrôlées par des génomes semi-synthétiques (projet Bacillus 2.0).....	232
4. Conclusion .....	234
<b>Annexes .....</b>	<b>237</b>
1. Liste des abréviations .....	237
2. Liste des amorces utilisées au cours de ces travaux .....	239
<b>Références .....</b>	<b>242</b>



# Avant-propos



## Avant-propos

La **biologie de synthèse (BS)** est un domaine scientifique émergent aux champs d'application très larges (industrie, santé, agriculture). Elle combine les apports de la biologie et de l'ingénierie et vise à mieux comprendre le vivant. Un enjeu majeur de la BS est d'apporter des solutions nouvelles, et respectueuses de l'environnement, aux problématiques sociétales actuelles.

Un des objectifs de la BS est l'optimisation de souches microbiennes utilisées pour des applications industrielles telles que la production de molécules d'intérêt. Cette approche vise à améliorer leur efficacité tout en minimisant l'impact environnemental du secteur industriel. Dans ce sens, l'un des axes de recherche de la BS est le développement de « **châssis microbiens** » : des micro-organismes modulables dans lesquels des circuits génétiques peuvent être introduits, maintenus stables, et correctement exprimés grâce aux ressources de l'hôte. Des micro-organismes modèles comme les bactéries *Escherichia coli*, ***Bacillus subtilis*** ou encore *Pseudomona putida* sont souvent utilisés comme organisme de départ pour la création de châssis. Une des voies d'améliorations de ces châssis microbiens consiste à utiliser des « châssis simplifiés » contrôlés par des génomes réduits, et des « châssis minimaux » ne disposant que de l'équipement nécessaire à leur multiplication. Minimiser le génome de la souche « hôte » a pour but de réduire au maximum les interférences possibles entre les voies de biosynthèse nouvellement intégrées et le réseau métabolique de l'hôte, et ainsi augmenter la production du produit d'intérêt.

*B. subtilis* est considéré comme le modèle des bactéries à Gram+ à l'instar d'*E. coli* pour les bactéries à Gram-. *B. subtilis* est une bactérie non pathogène retrouvée dans de nombreux environnements comme les sols, les racines des plantes ou le tractus gastro-intestinal. De par ses propriétés (compétence naturelle, croissance rapide dans des milieux simple, sécrétion de protéines, statut GRAS (Generally Recognized As Safe), *B. subtilis* est souvent utilisée dans le secteur industriel pour la production de molécules d'intérêt.

La production de châssis minimaux à partir de *B. subtilis* pourrait constituer une avancée importante dans l'amélioration des souches bactériennes à visée industrielle. Cependant à l'heure actuelle, les outils génétiques permettant la réduction du génome de *B. subtilis* sont trop peu efficaces pour envisager une ingénierie à grande échelle. En effet, il faut



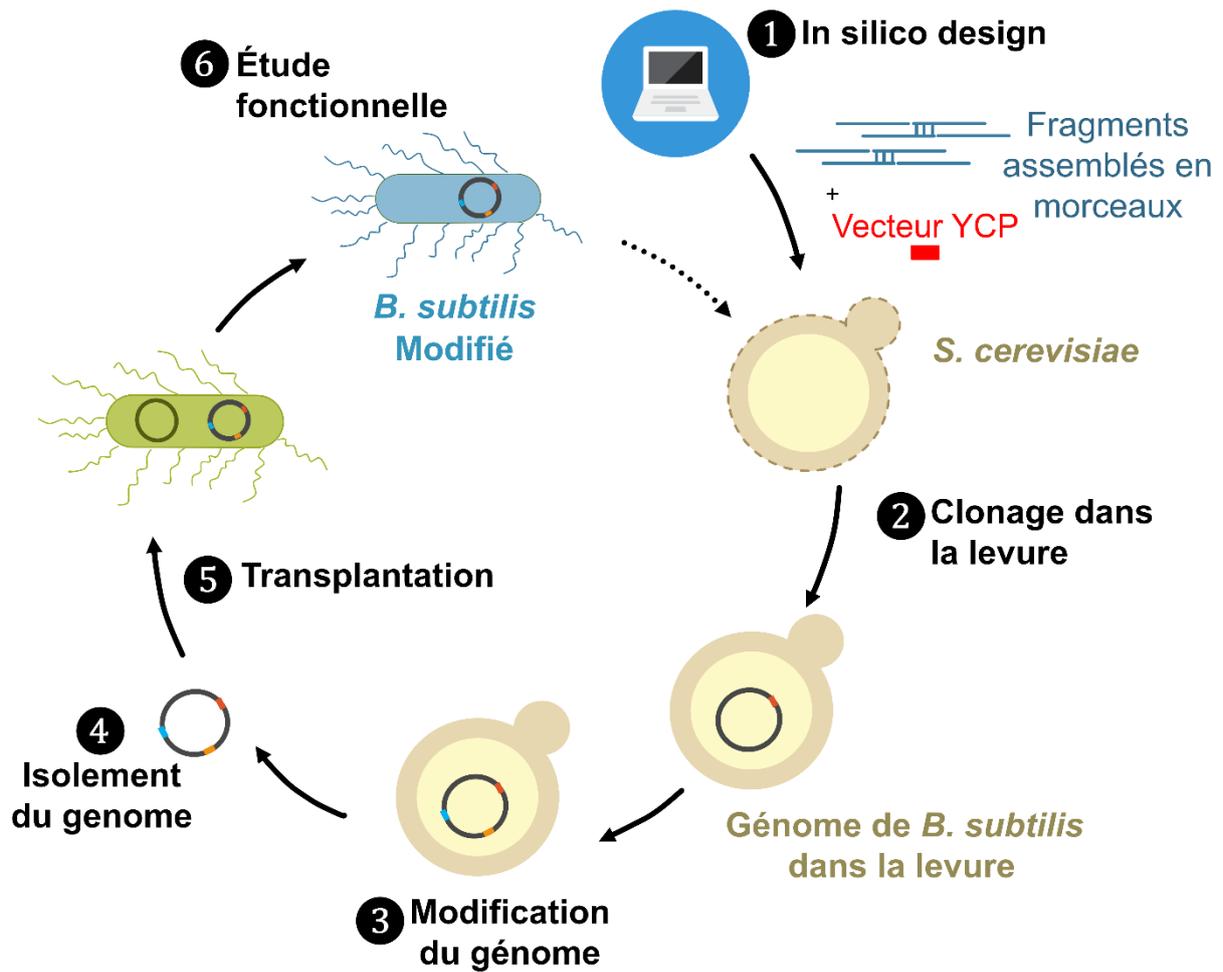
compter environ un mois pour supprimer un groupe de gènes, et l'efficacité du processus décroît dans les souches à génome réduit. Il faudrait ainsi plusieurs mois, voire des années de travail, pour obtenir des souches avec des génomes complètement remaniés, un délai incompatible avec les exigences de l'innovation industrielle. Il paraît donc nécessaire de développer de nouvelles méthodes permettant de réduire, de modifier, voire de redessiner à grande échelle le génome de *B. subtilis*.

Cette thèse s'inscrit dans le projet ANR « **Bacillus 2.0** » dont l'objectif final est le développement de **méthodes efficaces pour la construction rapide de châssis de *B. subtilis* semi-synthétiques, minimaux, programmables et capables de croître dans des milieux simples et peu coûteux**. Pour ce faire, le projet vise à transférer à l'organisme modèle *B. subtilis* des outils de BS, initialement développés chez les mycoplasmes.

## 1. Le projet Bacillus 2.0

Le projet « Bacillus 2.0 » est né d'une collaboration entre l'équipe SyBER (Institut MICALIS, UMR 1319 INRAE / AgroParisTech, Jouy-en-Josas) et l'équipe Mollicutes (Unité BFP, UMR 1332 INRAE / Université de Bordeaux, Villenave d'Ornon). Il est coordonné par le Pr. Matthieu Jules (Institut MICALIS) et s'appuie sur les expertises complémentaires apportées par cette collaboration. L'équipe SyBER est spécialisée dans les approches de biologie des systèmes et de biologie de synthèse pour l'étude et les applications biotechnologiques de *B. subtilis*. L'équipe Mollicutes possède, quant à elle, une solide expertise en génomique fonctionnelle et comparative des mycoplasmes (des bactéries considérées comme les meilleures représentantes du concept de « cellule minimale ») et est à l'avant-garde de développements majeurs dans le domaine de la BS (méthodes de clonage, ingénierie et transplantation de génomes bactériens complets).

Le projet « Bacillus 2.0 » s'inspire d'une approche innovante de BS développée entre 2007 et 2010 au J. Craig Venter Institute (JCVI) sur des mycoplasmes. Il vise à adapter cette approche à *B. subtilis* de la façon suivante : 1) Conception assistée sur ordinateur d'un génome *B. subtilis* semi-synthétique minimal. 2) Assemblage de fragments d'ADN chevauchants d'environ 10 kb en fragments de 100 kb puis en génome complet, chez *S. cerevisiae*. 3) Ingénierie du génome assemblé, à l'aide des outils génétiques disponibles chez la levure. 4) Isolement du génome



**Figure 1 : Cycle d'ingénierie de génome bactérien du projet Bacillus 2.0.** **1)** Etudes bio-informatiques pour concevoir un génome minimal de *B. subtilis*. Synthèse de petits fragments d'ADN chevauchants qui seront assemblés *in vitro* en plus grands fragments puis *in vivo* dans la levure *S. cerevisiae* en présence d'un vecteur YCP (Yeast Centromeric Plasmide). Ce plasmide contient les éléments levures permettant le maintien du génome bactérien dans la levure (un Centromère CEN6, une origine de réplication ARSH4, et un marqueur d'auxotrophie Histidine) ainsi qu'un marqueur de résistance à un antibiotique. **2)** Assemblage du génome par recombinaison homologue dans la levure. **3)** Ingénierie du génome grâce aux outils génétiques disponibles chez la levure (optionnel). **4)** Isolement du génome. **5)** Transplantation du génome entier dans une cellule receveuse. **6)** Sélection et caractérisation de la bactérie synthétique obtenue.

semi-synthétique, 5) Transplantation vers une cellule receveuse. 6) Sélection et caractérisation de nouvelles souches de *B. subtilis* (figure 1).

D'un point de vue organisationnel, l'équipe SyBER de Jouy-en-Josas est en charge de la partie conception et synthèse du génome réduit semi-synthétique de *B. subtilis* d'une taille d'environ 2 Mb. En parallèle, l'équipe Mollicutes de Villenave d'Ornon a pour objectif de cloner et de modifier le génome de *B. subtilis* dans la levure. En effet, l'assemblage du génome semi-synthétique de *B. subtilis* se faisant, en grande partie, dans la levure, il est nécessaire de démontrer que son génome naturel peut être maintenu stable chez cet hôte et modifier en utilisant le répertoire d'outils génétiques disponibles chez la levure. L'équipe a aussi pour rôle de développer les méthodes de transplantation de génomes pour *B. subtilis* (figure 1 étape 5).

## 2. Avancement du projet à mon arrivée au laboratoire

Lors de mon arrivée au laboratoire, le projet « Bacillus 2.0 » avait déjà été entamé, notamment avec les travaux de thèse du Dr Gabrielle Guesdon qui avaient pour objectifs de cloner le génome de *B. subtilis* dans la levure, de le modifier avec les outils d'ingénierie disponibles chez *S. cerevisiae*, puis de le transplanter dans une cellule receveuse de façon à mimer le cycle d'ingénierie présenté figure 1. Or, au cours de ces travaux, de nombreuses difficultés ont été rencontrées et les objectifs initiaux n'ont jamais été atteints.

Afin de présenter le contexte dans lequel mon projet de thèse a démarré (Septembre 2021), une partie des travaux de thèse de Dr. Guesdon (2019-2022) est résumée ci-après.

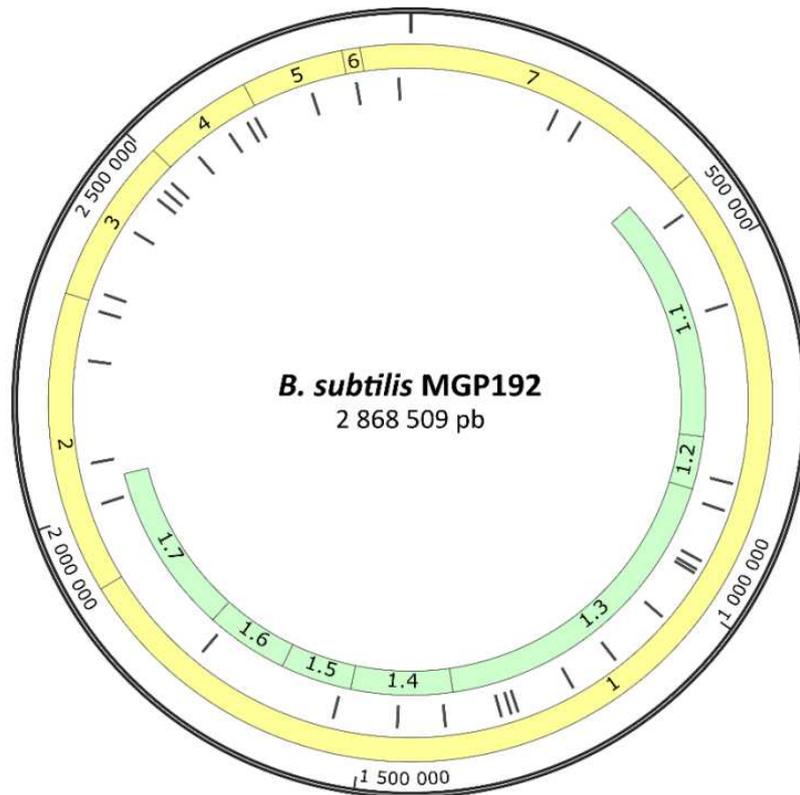
**Le premier objectif de thèse du Dr. Guesdon** était de démontrer qu'un génome de *B. subtilis* natif agrémenté d'une origine de répllication levure (ARS), d'un centromère (CEN), et d'un marqueur d'auxotrophie pouvait être maintenu de façon stable chez la levure et ceci pendant plusieurs passages. Comme aucun génome de taille supérieur à 1,8 Mb et utilisant un code génétique universel n'avait été cloné dans la levure, ce travail était crucial pour attester la faisabilité de la méthode proposée pour construire le futur génome semi-synthétique de *B. subtilis* dans la levure.



Le travail a démarré par la mise en place des stratégies déjà décrites dans la littérature (trois au total, la méthode « classique », la méthode « CReasPy-Cloning » et la méthode par « Fusion »). Malgré de nombreux efforts, aucune des méthodes utilisées n'a permis de cloner avec succès le génome complet de *B. subtilis* souche MS (3.98 Mb), ni même une version réduite de ce génome issue de la souche MGP192 (2.86 Mb). Pour les deux premières méthodes, l'ADN de *B. subtilis* semblait totalement absent des « transformants levure » apparaissant sur milieu de sélection. Pour la troisième méthode, le génome était parfois détecté mais ce dernier ne semblait pas s'y maintenir au cours du temps. Les méthodes utilisées au cours de ce travail sont détaillées dans la partie Introduction de ce manuscrit, (cf paragraphes 2.4.1.1, 2.4.1.3 et 2.4.1.4 et figures 15, 17 et 18 de l'introduction) ainsi que dans la partie introductive du chapitre 1 (cf partie 1 du chapitre 1).

Suite à ces expériences, plusieurs hypothèses ont été formulées pour expliquer ces échecs :

- **La taille du génome (et/ou l'état de condensation du génome au moment du clonage).** La taille du génome de *B. subtilis* (3.98 Mb), y compris celle de la souche à génome réduit (2,86 Mb), sont bien supérieures à celle des plus grands génomes bactériens clonés dans la levure (*Haemophilus influenzae* 1,8 Mb (Karas *et al.*, 2013) et *Spiroplasma citri* 1,8 Mb (Labroussaa *et al.*, 2016)). A l'heure actuelle, aucune donnée n'indique que des génomes de taille supérieure à 1.8Mb peuvent entrer et être maintenus dans la levure.
- **Le manque d'ARS fonctionnelle (Autonomously Replicating Sequence).** Les ARS sont des séquences qui fonctionnent comme des origines de réplication chromosomique chez *S. cerevisiae* et qui jouent un rôle essentiel dans la maintenance des chromosomes. Elles ont une longueur d'environ 100-200 pb dont la séquence varie considérablement, mais elles ont toutes en commun (i) une teneur élevée en A-T et (ii) un motif de 11 pb appelé ACS (5'-WTTTATRTTTW-3'). Une conservation substantielle de la séquence a été observée dans les 3 pb de part et d'autre de l'ACS essentiel, ce qui a permis d'identifier un ACS étendu de 17 pb (EACS), 5'-WWWTTTAYRTTTWGTT-3'. La recherche manuelle de motifs ACS 5'-WTTTATRTTTW-3' (11 nt) ou 5'-WWWTTTAYRTTTWGTT-3' (17 nt) a permis de comptabiliser 145 ACS dans le génome MGP192, mais la présence de ces motifs ne garantit pas la présence d'une ARS fonctionnelle à ces sites. Cette hypothèse est renforcée par l'outil de prédiction Ori-Finder3 qui permet d'identifier les ARS putatives par comparaison avec une banque de

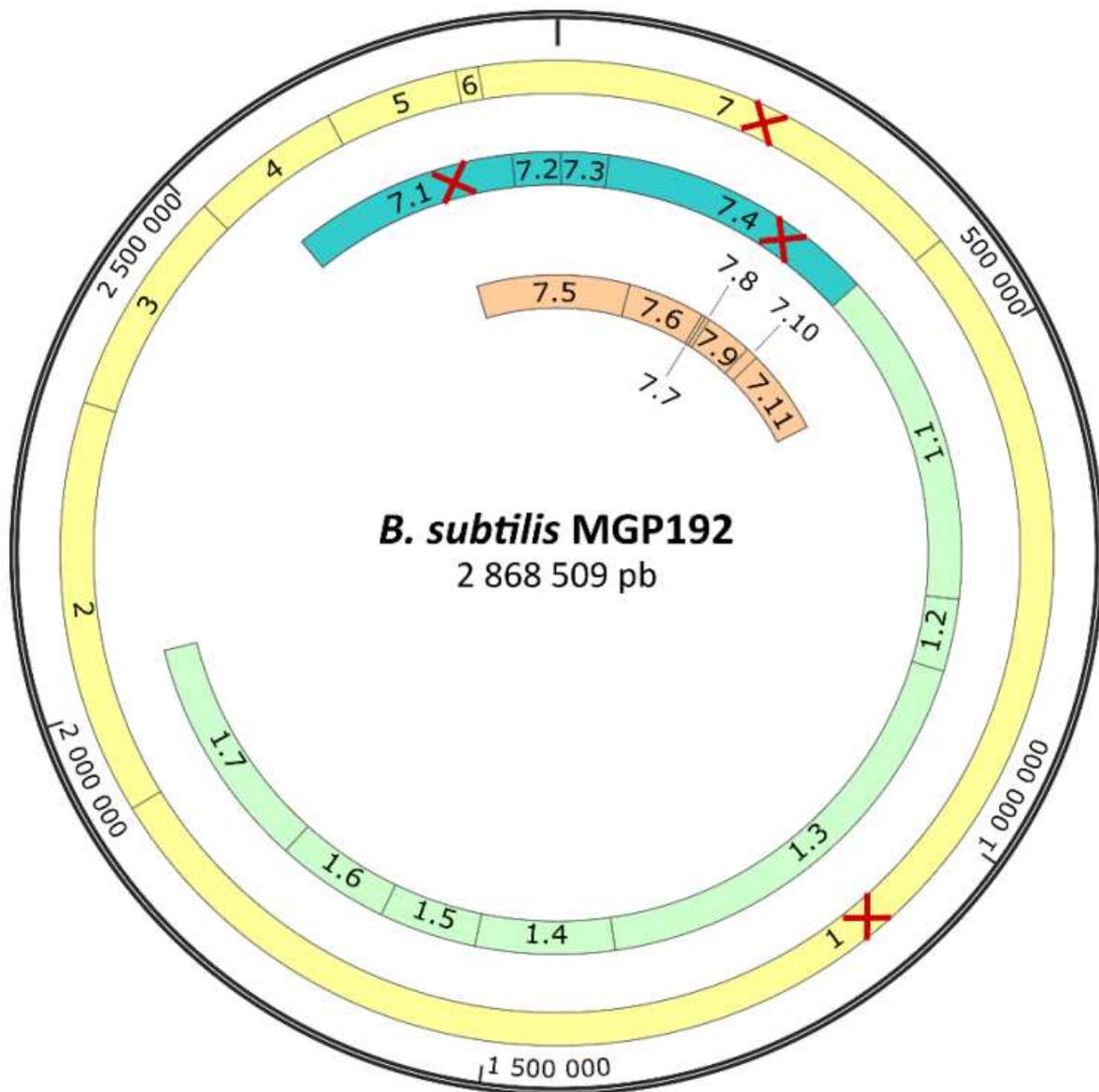


**Figure 2 : Carte du génome de *B. subtilis* MGP192 (figure issue de la thèse du Dr. Guesdon).** Les différents fragments issus de la digestion par l'enzyme de restriction SgrDI sont présentés en jaune, ceux issus de la digestion par l'enzyme FseI sont présentés en vert. Les ARS prédites par l'outil Ori-finder3 sont indiquées par les traits noirs.

données qui répertorie toutes les séquences ARS fonctionnelles chez la levure et qui a permis de prédire la présence de 36 ARS fonctionnelles le long de ce génome. La répartition de ces motifs sur le génome de *B. subtilis* MGP192 est représentée figure 2. Chez *S. cerevisiae*, une ARS forte permet la répllication d'une région de 200 à 300 kb (Dershowitz and Newlon, 1993).

- **Le pourcentage en G+C du génome (en lien avec les ARS).** De manière générale, les génomes bactériens possédant un pourcentage en G+C inférieur ou égal à 40% peuvent être clonés et maintenus dans la levure simplement en ajoutant un centromère (CEN), un marqueur de sélection (HIS3 par exemple) et une ARS levure (voire pas d'ARS). L'hypothèse la plus probable est que la teneur en A-T de ces génomes est suffisamment élevée pour contenir « naturellement » des éléments dits ARS-like, jouant le rôle d'ARS dans la levure. Il a été montré, en revanche, que l'ajout artificiel d'ARS provenant de *S. cerevisiae* était nécessaire pour cloner des génomes ou fragments d'ADN avec un pourcentage en G+C supérieur à 40% (Venetz *et al.*, 2019; Zhou *et al.*, 2016; Noskov *et al.*, 2012). Le pourcentage en G+C du génome de *B. subtilis* est d'environ 43,5%, soit légèrement au-dessus du seuil des 40%, indiquant que l'ajout de quelques ARS est probablement nécessaire.
- **La présence de gènes dont l'expression serait toxique pour la levure.** Il est en effet possible que l'expression de certains gènes ou régions du génome de *B. subtilis* soit délétère pour la levure.
- **La présence d'éléments répétés.** La présence de nombreux éléments répétés peut stimuler la machinerie de recombinaison homologue chez *S. cerevisiae* et être à l'origine de remaniements chromosomiques.

Pour valider ou invalider certaines de ces hypothèses, **la deuxième partie de la thèse du Dr. Guesdon** a consisté à mettre en place des expériences de TAR-cloning (Transformation-Associated Recombination) (Natalay Kouprina and Vladimir Larionov, 2016). Cette méthode, détaillée dans partie 1.4.1.2 de l'introduction de ce manuscrit, permet de capturer individuellement des fragments de génomes et d'évaluer leur stabilité et/ou leur toxicité. Pour y parvenir, le génome réduit de *B. subtilis* (2.86 Mb) a été extrait en bloc d'agarose puis hydrolysé *in situ* à l'aide de trois enzymes de restriction au cours de trois expériences indépendantes. La première expérience a mis en jeu l'enzyme de restriction SgrDI qui



**Figure 3 : Bilan des essais de clonage des fragments du génome de *B. subtilis* par TAR-cloning (figure issue de la thèse du Dr. Guesdon).** Carte du génome de *B. subtilis* MGP192 avec les fragments issus de la digestion par l'enzyme SgrDI (en jaune), ou par l'enzyme FseI (en vert ou bleu) ou par l'enzyme AsiSI (en orange). Les fragments et sous fragments qui n'ont pas pu être clonés dans la levure sont marqués d'une croix rouge.

hydrolyse le génome en 7 fragments (figure 3, fragments jaunes). La deuxième expérience a mis en jeu l'enzyme de restriction FseI qui hydrolyse le génome en 17 fragments (figure 3, fragments verts ou bleus), et la troisième, l'enzyme de restriction AsiSI qui hydrolyse le génome en 32 fragments (figure 3, fragments oranges). Suite à ces digestions, plusieurs tentatives de clonage de fragments de génome de *B. subtilis* MGP 192 ont été réalisées.

Si la combinaison des 3 expériences a indiqué, à la grande surprise de l'équipe, que l'ensemble du génome de *B. subtilis* MGP192 pouvait être cloné dans la levure en 21 morceaux allant de 6 kb à 515 kb, l'analyse des résultats de chaque expérience séparément a révélé que certaines zones du génome ne se prêtaient pas au clonage (figure 3), en particulier :

- Le « fragment 7 » d'une taille de 482 kb (en jaune, figure 3), qui correspond à la région située entre les gènes *sacX* et *swrC* kb après digestion du génome par l'enzyme SgrDI, et qui n'a jamais pu être cloné entier dans la levure. Des analyses de séquences ont permis de mettre en évidence dans cette région plusieurs séquences répétées pouvant être responsables de l'instabilité du fragment dans la levure (7 opérons ribosomiques), ainsi que des *loci* pouvant être à l'origine d'une certaine toxicité (ex : l'opéron *sfrAA-AD* codant la surfactine ou peut-être aussi l'expression des gènes ribosomiques). Ce fragment a été découpé en 4 sous-fragments lors de la digestion avec l'enzyme FseI (7.1 à 7.4, en bleu, figure 3) et en 7 sous-fragments avec l'enzyme AsiSI (7.5 à 7.11, en orange, figure 3). Tous ces sous-fragments ont pu être clonés dans la levure à l'exception des fragments 7.4 et 7.1 (figure 3).
- Le « fragment 1 » d'une taille de 1500 kb (en jaune, figure 3), qui correspond à la région entre les gènes *swrC* et *nadR*, et qui lui non plus n'a jamais pu être cloné entier dans la levure. Ce fragment a été découpé en 7 sous-fragments lors de la digestion avec FseI (fragments 1.1 à 1.7, en vert, figure 3). Chacun de ces sous-fragments a pu être cloné sans difficulté dans la levure, suggérant que le manque de stabilité du fragment 1 était vraisemblablement dû à la taille du fragment, au manque de séquence ARS fonctionnelles ou au ratio nombre d'ARS fonctionnelles/taille du fragment.

Ainsi, bien que le génome entier de *B. subtilis* n'ait pas été cloné avec succès, les expériences de TAR-cloning ont permis de mettre en lumière plusieurs éléments critiques. Il est apparu (i) que l'efficacité du clonage diminuait en fonction de l'augmentation de la taille



des fragments, (ii) que la présence de gènes/régions spécifiques sur certains fragments empêchait le clonage de ces derniers dans la levure (si cela se confirmait, ces régions devraient être éliminées ou modifiées), et (iii) que la taille du génome et la proportion de séquences dites « ARS-like » sont des facteurs qui, combinés, pourraient mettre en difficulté le clonage du génome de *B. subtilis* dans la levure.

Bien que puissante, la méthode de TAR-cloning impose certaines contraintes, notamment celles de devoir travailler avec les sites de restriction présents sur le génome, ce qui laisse peu de flexibilité à l'expérimentateur. **La troisième partie de la thèse du Dr. Guesdon** a donc consisté à développer une nouvelle méthode de clonage, appelée CReasPy-Fusion, qui permet la capture (et l'édition si nécessaire) de génomes entiers ou de fragments de génome de façon spécifique en choisissant le site de coupure grâce au système CRISPR-Cas9. Le génome de *B. subtilis* n'étant pas stable dans la levure, cette méthode a été mise au point avec des mycoplasmes dont les génomes avaient déjà été clonés dans la levure par d'autres méthodes. Au cours de ce travail, les génomes de 6 espèces de mycoplasmes ainsi que de grands fragments de génome ont pu être clonés/édités dans la levure (Guesdon *et al.*, 2023).

**Mon projet de thèse a donc débuté dans le contexte suivant :**

- Le génome de *B. subtilis* MGP192 ne peut pas être cloné en l'état dans la levure *S. cerevisiae*, remettant en question la faisabilité du projet.
- Les expériences de TAR-cloning ont mis en évidence (i) des barrières ou facteurs mettant en échec le clonage du génome de *B. subtilis* dans la levure (ex : le manque d'ARS, la présence de séquences répétées, des problèmes de toxicité liés à l'expression de certains gènes dans la levure, etc...) et également (ii) les régions du génome réfractaires au clonage.
- Une nouvelle méthode de clonage de génomes ou fragments de génomes, plus flexible, est disponible.
- Aucune expérience de transplantation de génomes n'a encore été amorcée.



### 3. Objectifs de la thèse

Dans un tel contexte, mon projet de thèse a eu pour objectif de démontrer que la levure *S. cerevisiae* restait un hôte pertinent pour la réalisation du projet « Bacillus 2.0 ». Pour cela, mon travail a été divisé en 3 parties qui correspondent aux chapitres 1, 2 et 3 de la partie « Résultats » de ce manuscrit.

- (i) Le premier objectif de ma thèse (chapitre 1) a été d'adapter la méthode de CReasPy-Fusion initialement développée chez des mycoplasmes à *B. subtilis*.
- (ii) Le second objectif (chapitre 2) a consisté à utiliser cette méthode dans le but de tenter de cloner des fragments de plus en plus grands du génome de *B. subtilis* dans la levure. Cette partie a été réalisée en étroite collaboration avec nos collègues de l'équipe SyBER (Institut MICALIS, UMR 1319 INRAE / AgroParisTech, Jouy-en-Josas).
- (iii) Enfin, le dernier objectif de ma thèse (chapitre 3) a été de développer une méthode alternative à la transplantation de génome de façon à pouvoir produire des mutants de *B. subtilis* ayant subi de grands remaniements.



# Etude bibliographique

**Tableau 1 : Taxonomie de *Bacillus subtilis*.**

Domaine	Procaryote
Règne	Bactérie
Embranchement/Phylum	<i>Bacillota</i> (anciennement <i>Firmicutes</i> )
Classe	<i>Bacilli</i>
Ordre	<i>Caryophanales</i> (anciennement <i>Bacillales</i> )
Famille	<i>Bacillaceae</i>
Genre	<i>Bacillus</i>
Espèce	<i>Bacillus subtilis</i>

## Etude bibliographique

Dans cette étude bibliographique, trois thèmes en lien direct avec mon projet de thèse sont abordés, à savoir (i) l'organisme modèle *Bacillus subtilis*, (ii) le domaine de la BS et enfin (iii) les systèmes de recombinaison.

### 1. *Bacillus subtilis*

#### 1.1. Taxonomie

*Bacillus subtilis* Cohn 1872 est considérée comme la bactérie modèle des bactéries Gram+, à l'image d'*Escherichia coli* chez les bactéries Gram-. *B. subtilis* fait partie de l'embranchement *Bacillota* (Oren and Garrity, 2021), autrefois connu sous le nom des Firmicutes, ou encore des *Firmacutes* (Gibbons and Murray, 1978). Cet embranchement comporte sept classes dont la classe des *Bacilli* qui, à son tour, regroupe deux ordres, les *Lactobacillales* et les *Bacillales* (Vos *et al.*, 2009). Suite aux travaux de B.J Tindall en 2019 (Tindall, 2019), l'ordre des *Bacillales* a été renommé *Caryophanales*, un nom proposé dès 1939 par Peshkoff. La famille des *Bacillaceae* fait partie de cet ordre, elle comprend 62 genres et au moins 457 espèces (Mandic-Mulec *et al.*, 2016). Le genre *Bacillus* en fait partie et regroupe, à lui seul, 226 espèces.

Les bactéries du genre *Bacillus* sont des bacilles Gram+, possédant une paroi, composée d'une épaisse couche de peptidoglycane qui réagit à la coloration de Gram en se colorant en violet. Elles ont la capacité de produire des endospores extrêmement résistantes (phénomène de sporulation) leur permettant de survivre dans des conditions extrêmes. Lorsque les conditions environnementales s'améliorent, la spore est capable de germer, entraînant la reprise de la prolifération de la bactérie (Cohn, 1872). Certaines espèces de ce genre ont fait l'objet d'études dans le domaine médical (exemple : *Bacillus anthracis* responsable de la maladie du charbon) (Wright and Mandell, 1986), dans celui des biotechnologies (exemple : *B. subtilis* pour la production d'enzymes) (Schallmeyer *et al.*, 2004), ou dans le secteur agricole (exemple : *Bacillus thuringiensis* pour combattre les insectes nuisibles).

En résumé, *B. subtilis* est une espèce bactérienne qui appartient au genre *Bacillus*, à la famille des *Bacillaceae*, à l'ordre des *Caryophanales* (anciennement *Bacillales*), à la classe des *Bacilli* et à l'embranchement des *Bacillota* (anciennement *Firmacutes*) (tableau 1).



## 1.2. Historique

*B. subtilis* a été découvert pour la première fois dans les années 1830 par Ehrenberg qui a nommé cet organisme *Vibrio subtilis*. Trente-quatre ans plus tard Cohn le renomma *B. subtilis*, créant en même temps le genre *Bacillus* (Cohn, 1872).

A l'origine, toutes les souches de laboratoire dérivait de la souche sauvage appelée « Marburg » décrite par Conn en 1930 (Conn, 1930). Une trentaine d'années après, une seconde souche pris le pas sur la souche Marburg, la souche 168, qui est aujourd'hui considérée à travers le monde comme la souche de référence. Cette souche a été produite en 1947 par Burkholder et Giles, deux botanistes de l'université de Yales. Au cours d'expériences de mutagénèse, ces chercheurs ont utilisé la souche Marburg pour produire des mutants par irradiation aux rayons X (Burkholder and Giles, 1947). Au cours de ces travaux, ils ont généré une souche auxotrophe pour le tryptophane, la souche 168. Pendant une longue période, cette souche n'a suscité que peu d'intérêt pour la communauté et il a fallu attendre les années 1960 et une série de travaux menés par Spizizen et Anagnostopoulos pour que cette souche soit reconnue et devienne, grâce à des caractéristiques qui lui sont propres, la souche de référence.

Deux expériences clés lui ont conféré ce statut. Tout d'abord, celle de John Spizizen en 1958 démontra que la souche 168 était capable d'intégrer de l'ADN exogène beaucoup plus facilement que sa souche mère, notamment lorsqu'il est isolé à partir d'autres souches sauvages de *B. subtilis* (Spizizen, 1958). Ensuite, celle d'Anagnostopoulos et Spizizen, trois ans plus tard, permis de développer un protocole de transformation efficace basé sur la compétence naturelle de cette souche (Anagnostopoulos and Spizizen, 1961).

Avec un tel protocole, il devenait facile d'introduire des molécules d'ADN étrangères dans cette souche et de générer des modifications au niveau génétique. L'ensemble de ces travaux a contribué à faire de *B. subtilis* 168 la souche de référence pour *B. subtilis* et un modèle d'étude en génétique.



### 1.3. Description générale

*B. subtilis* est une bactérie à paroi Gram+. Elle a une forme de bâtonnet (bacille) et une taille qui varie de 2 à 4 µm en longueur, et de 0,5 à 2 µm en largeur. Elle est mobile grâce à un système de flagelles recouvrant toute sa surface (ciliature péritriche). Le génome *B. subtilis* souche 168 a été séquencé en 1997 (Kunst *et al.*, 1997). Il a une taille de 4,2 Mb avec environ 4100 gènes codant des protéines et un pourcentage en G+C de 43.5%.

*B. subtilis* a une croissance assez rapide avec un temps de génération de moins de 30 min en milieu LB. A titre de comparaison le temps de génération d'*E. coli* est de 20 min. Elle peut croître de façon aérobie et anaérobie facultative. En absence d'accepteur terminal d'électron, elle est capable de croître par fermentation, en présence d'un mélange de glucose et de pyruvate (Nakano and Zuber, 1998). Cette capacité de *B. subtilis* est notamment utilisée pour la production de natto, un aliment japonais traditionnel à base de graines de soja fermentées (Kubo *et al.*, 2011). *B. subtilis* est non pathogène et possède le statut GRAS (Generally Recognized as Safe) délivré par la Food and Drug Administration (administration américaine chargée de la surveillance des denrées alimentaires et des médicaments).

*B. subtilis* est retrouvé dans de nombreux environnements différents, c'est une bactérie ubiquitaire et saprophyte. Dans la nature, elle est principalement retrouvée dans le sol en interaction avec les plantes (Ashlee M. Earl, 2008), le tractus digestif d'animaux ou des humains (Hong *et al.*, 2009), ou encore dans l'air et sur les plantes. Dans le sol, elle se développe notamment au niveau de la rhizosphère, qui correspond à la partie du sol pénétrée par les racines et les micro-organismes associés. La plante libère des molécules au niveau de ses racines qui agissent comme des signaux attirant ou repoussant les microorganismes présents dans les sols (Zhang *et al.*, 2014). Ce mécanisme permet d'attirer au niveau des racines un ensemble de bactéries favorisant la croissance et la santé des plantes, les rhizobactéries. *B. subtilis* fait partie de ces bactéries bénéfiques pour la plante, elle agit principalement en protégeant la plante contre l'attaque de pathogènes et en lui donnant accès à des nutriments présents dans les sols et essentiels à sa croissance (Blake *et al.*, 2021).



Le fait qu'elle soit présente dans tous ces environnements différents témoigne d'une forte capacité d'adaptation à des contraintes physico-chimiques et nutritionnelles multiples. En effet, lorsque les conditions environnementales se dégradent, *B. subtilis* est capable d'intégrer de l'ADN exogène. Cette faculté, appelée compétence, lui permet d'acquérir de nouveaux traits. Dans les cas extrêmes, afin de survivre, elle peut également produire des endospores très résistantes (Cohn, 1872). Une fois que les conditions environnementales s'améliorent, ces endospores peuvent à nouveau entrer en germination et reprendre un cycle de vie végétatif. De plus, dans les milieux naturels où elle évolue, elle est en constante compétition avec les autres micro-organismes présents. Pour s'en défendre, elle produit des biofilms qui correspondent à des communautés de micro-organismes associés entre eux et attachés à une surface, le tout étant enfermé dans une matrice extracellulaire autoproduite (Vlamakis *et al.*, 2013). Ce biofilm permet donc de protéger les cellules de *B. subtilis* des autres micro-organismes ainsi que des contraintes extérieures. Enfin, *B. subtilis* a la capacité de sécréter un grand nombre d'enzymes extracellulaires comme par exemple des composés antimicrobiens comme la subtilisine A ou la surfactine (Caulier *et al.*, 2019). La formation de biofilms ainsi que la sécrétion d'enzymes extracellulaires sont des mécanismes permettant à *B. subtilis* de se protéger mais aussi de protéger la plante avec laquelle elle interagit (Bais *et al.*, 2004).

Pour résumer, sa robustesse, son statut GRAS, mais également ses capacités de fermentation, de compétence naturelle et de sécrétion de protéines font de *B. subtilis* une bactérie avec un fort potentiel pour la production industrielle de molécules d'intérêts (Schallmey *et al.*, 2004). Les mécanismes de compétence et de sécrétion de protéines seront détaillés dans la suite de cette étude bibliographique.

#### 1.4. Compétence naturelle

La compétence est la capacité d'une cellule à prélever et importer une molécule d'ADN exogène libre dans l'environnement. Chez *B. subtilis*, ce processus peut être naturel et lui permet de s'adapter à différentes conditions environnementales. Les cellules de *B. subtilis* ne sont cependant pas constamment en état de compétence. Le passage dans cet état n'est possible que lors de la transition entre la phase de croissance exponentielle et la phase stationnaire (Hahn *et al.*, 1994). Ce phénomène est finement régulé et contrôlé par le facteur

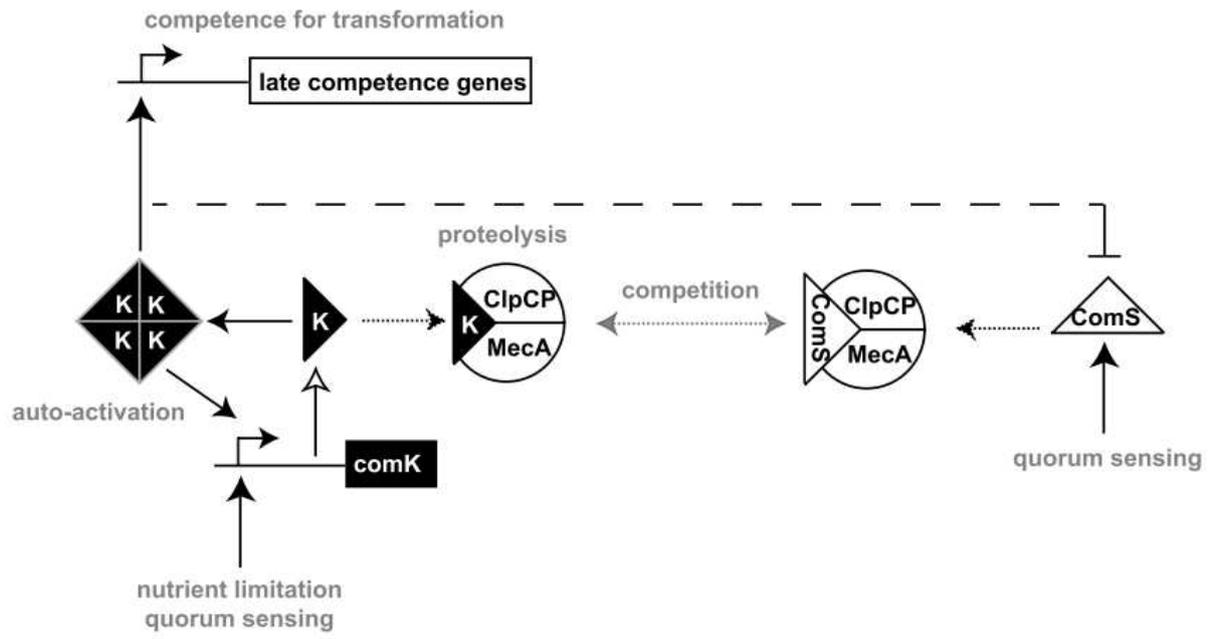


Figure 4 : Schéma simplifié du réseau de régulation de la compétence chez *B. subtilis* d'après Maier, 2020. Le facteur ComK est symbolisé par le triangle noir avec la lettre K.

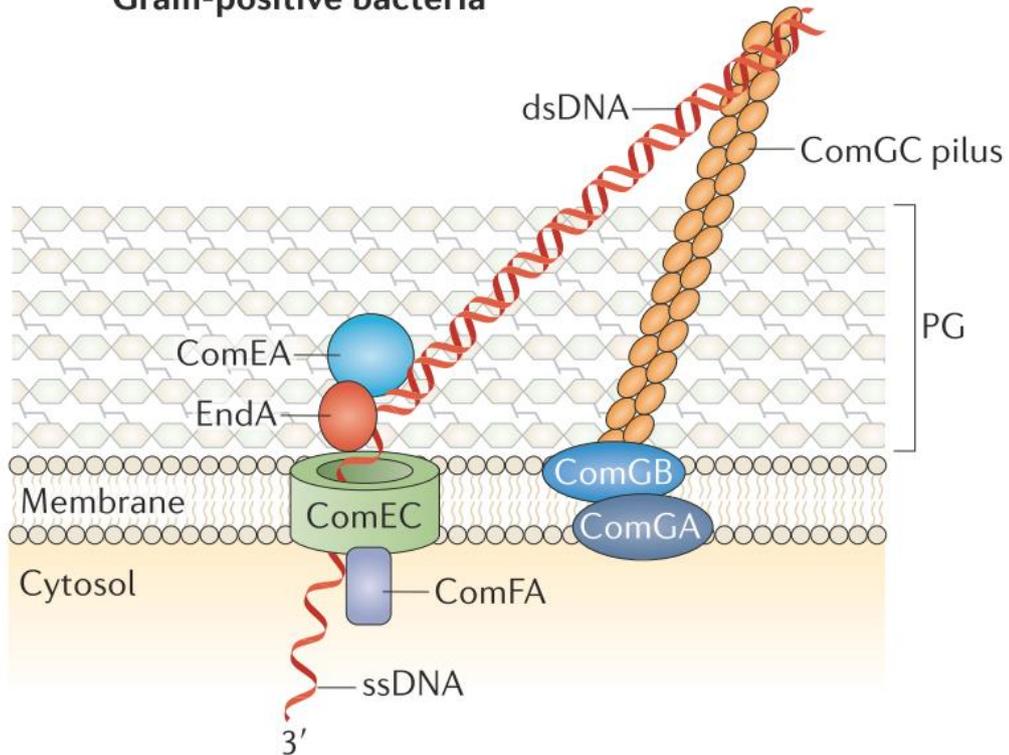
de transcription ComK (Van Sinderen *et al.*, 1995). C'est la concentration de ComK qui définit le passage en état de compétence. Lorsque ComK dépasse un seuil de concentration, les cellules deviennent compétentes. Ce phénomène n'est pas synchrone et, à l'échelle d'une population, seulement environ 15% des cellules deviennent compétentes (NESTER and STOCKER, 1963), parce que toutes les cellules ne disposent pas d'une quantité suffisante de ComK.

Au sein de la cellule, la régulation de ce mécanisme est très complexe et le facteur ComK, qui est un activateur de très nombreux gènes de la compétence, est central dans cette régulation (figure 4). Au niveau transcriptionnel, différents facteurs vont influencer l'expression du gène *comk*. Tout d'abord, le facteur ComK agit sur sa propre transcription par une boucle de contrôle positive. Il se fixe au promoteur de son propre gène, induisant sa transcription (Van Sinderen and Venema, 1994). Au début de la transition vers l'état de compétence, lorsque la concentration en ComK est insuffisante pour provoquer cette boucle de contrôle positif, le facteur DegU facilite la liaison de ComK à son promoteur. DegU est considéré comme le facteur d'initiation de la compétence chez *B. subtilis* (Hamoen *et al.*, 2000). Si ComK et le régulateur de réponse DegU activent la transcription du gène *comK*, d'autres facteurs l'inhibent ; c'est le cas des trois répresseurs AbrB, CodY (Serror and Sonenshein, 1996) et Rok (repressor of ComK) (Hoa *et al.*, 2002) qui sont eux aussi finement régulés. Par exemple, la transcription du gène *rok* est inhibée par ComK mais aussi par son propre produit, Rok (Maier, 2020; Hamoen *et al.*, 2003). L'expression de ComK est aussi contrôlée négativement au niveau de son ARNm par la protéine Kre qui réduit sa stabilité (Gamba *et al.*, 2015).

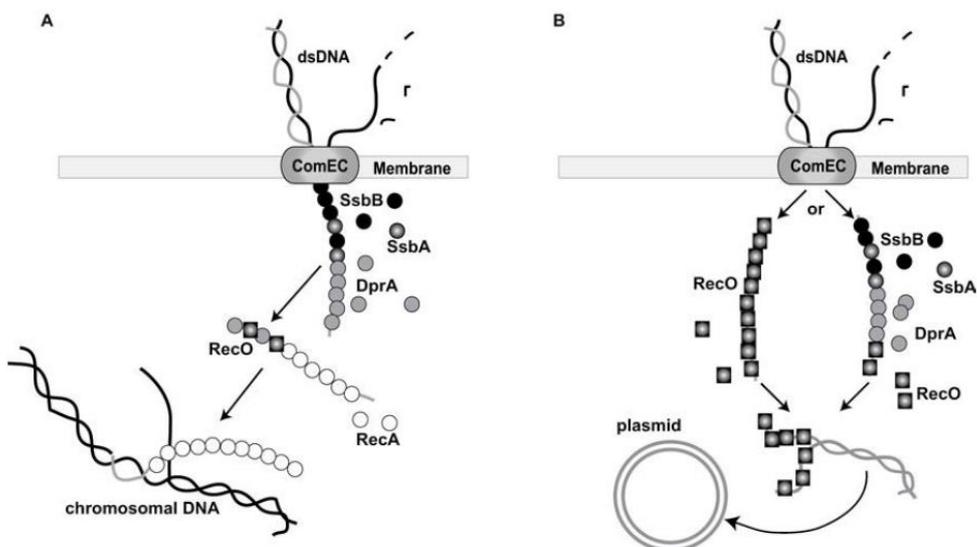
Au niveau post-traductionnel, la régulation se fait par protéolyse via la facteur MecA. En phase exponentielle de croissance, le facteur de transcription ComK lié à MecA est dégradé par le complexe protéasique ClpCP (Turgay *et al.*, 1997; Turgay *et al.*, 1998), le maintenant ainsi à un niveau basal. Puis, lorsque la densité cellulaire augmente, la quantité de protéine ComS augmente (voir ci-dessous) et entre en compétition pour la liaison avec la protéase MecA-ClpCP, protégeant ainsi ComK de la dégradation (Maier, 2020) (figure 4).

L'entrée en compétence est due à la densité cellulaire et à la limite en nutriments dans le milieu. En effet, les cellules sécrètent des phéromones dans le milieu, dont l'accumulation est le signal d'une forte densité cellulaire. La détection par les cellules de ces phéromones

## Gram-positive bacteria



**Figure 5 : Représentation schématique de la compétence naturelle chez les bactéries à Gram positif, d'après Johnston *et al.*, 2014.** Le complexe protéique permettant l'entrée de l'ADN exogène par compétence naturelle est représenté. L'ADN double brin est capté par le pseudopilus composé principalement de ComGC ainsi que des protéines ComGB et ComGA. L'ADN double brin est ensuite fixé par ComEA qui le guide vers le complexe ComEC. Ce complexe qui correspond à un pore membranaire, s'associe à la translocase ATP-dépendante ComFA pour permettre le passage à travers la membrane, de l'ADN sous forme simple brin (le second brin étant dégradé). PG : peptidoglycane. dsDNA : double strand DNA (ADN double brin). ssDNA : single strand DNA (ADN simple brin).



**Figure 6 : Schéma représentant le devenir intracellulaire d'une molécule d'ADN importée d'après Maier, 2020. [A] Import et recombinaison homologue. [B] Import et re-circularisation d'un plasmide.**

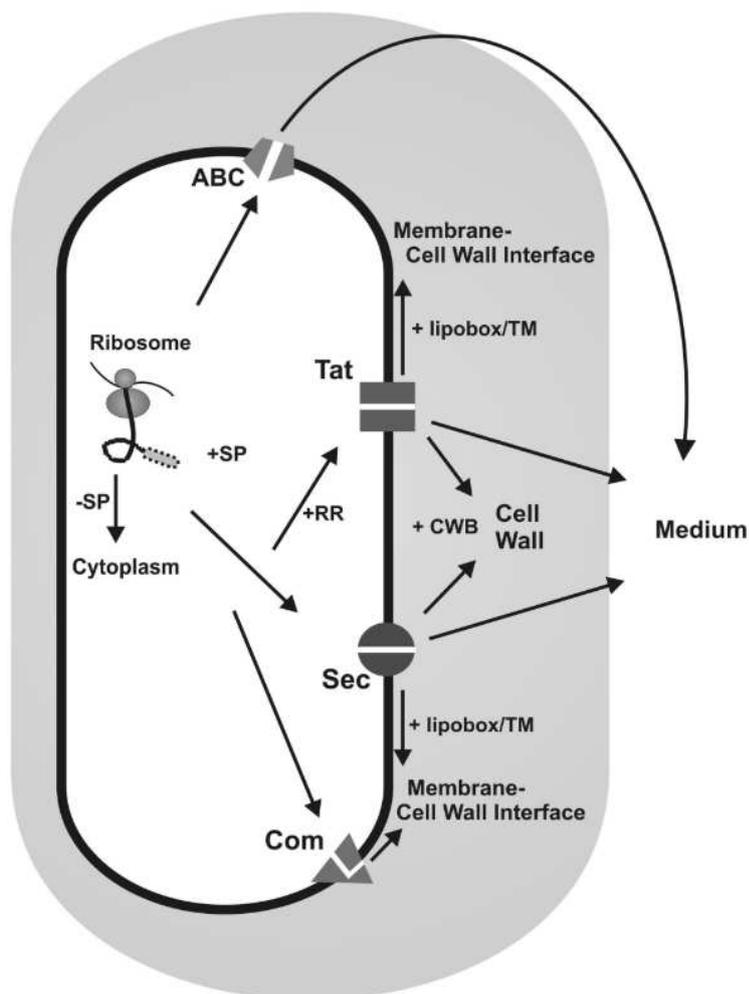
présentes dans le milieu est appelée « quorum sensing ». Deux phéromones induisant la compétence ont été identifiées. La première, ComX, détectée par la protéine membranaire ComP, entraîne, entre autres, l'expression de ComS. La seconde, CSF (Competence Stimulating Factor), pénètre dans la cellule via la perméase Spo0K et induit, elle aussi, *in fine* l'expression de ComS (Hamoen *et al.*, 2003) (figure 4).

Une fois l'état de compétence atteint, les cellules sont capables d'intégrer de l'ADN exogène, ce phénomène s'appelle la transformation. Chez *B. subtilis*, ce mécanisme est basé sur plusieurs protéines « Com » dont l'expression est induite par ComK. Il est divisé en 3 grandes étapes : 1) ancrage de l'ADN à la surface des cellules et fragmentation, 2) import de l'ADN et 3) recombinaison et insertion dans le génome ou re-circularisation.

La première étape de la transformation est l'accroche de l'ADN exogène double brin (ADNdb) à la surface de la bactérie. Cet ancrage se fait par l'intermédiaire de la protéine ComEA fixée à la membrane plasmique. Afin d'avoir accès à cette protéine, l'ADN doit passer l'épaisse couche de peptidoglycane de la paroi de la bactérie (figure 5). Un pseudopilus formant un canal permet à l'ADN exogène de traverser cette couche de peptidoglycane. Il est composé majoritairement de protéines ComGC et minoritairement de protéines ComGD, ComGE et ComGG. Ces protéines s'assemblent pour former le pseudopilus par l'intermédiaire de l'ATPase ComGA et de la protéine membranaire ComGB. Le pseudopilus en lui-même n'est pas capable de fixer l'ADN, il donne simplement l'accès à la protéine membranaire réceptrice, ComEA (Provvedi and Dubnau, 1999). L'ADNdb exogène est ainsi ancré à la surface de la cellule (figure 5).

La deuxième étape consiste à importer dans le cytoplasme l'ADN sous forme simple brin (ADNsb). Cette étape met en jeu 3 protéines, le récepteur ComEA, la perméase ComEC et l'ATPase ComFA. ComEA apporte l'ADNdb au canal membranaire formé par ComEC (Dubnau, 1999). Ce canal est associé à l'ATPase ComFA qui sert de moteur pour l'entrée de l'ADN sous forme simple brin dans le cytoplasme. Un des brins est donc dégradé au cours de ce transport (figure 5).

Une fois l'ADNsb entré dans la cellule, la dernière étape consiste soit à l'intégrer dans le génome de *B. subtilis* (figure 6A), soit à le re-circulariser (figure 6B). L'intégration au chromosome se fait par recombinaison homologue par l'intermédiaire de la protéine RecA. La re-circularisation de la molécule d'ADN se fait, quant à elle, par l'intermédiaire de la protéine



**Figure 7 : Représentation schématique des voies de sécrétion des protéines chez *B. subtilis* d'après Tjalsma *et al.*, 2004.** Le devenir des protéines synthétisées est dicté par la présence (+SP) ou l'absence (-SP) d'un peptide signal spécifique. Les protéines qui doivent être retenue du côté extra cytoplasmique de la membrane contiennent soit un segment transmembranaire (TM) soit une modification lipidique (lipobox). Elles sont exportées par la voie Sec. En présence d'un peptide signal twin-arginine (RR), les protéines sont exportées par la voie Tat. Les protéines qui forment le pseudopilus du complexe de la compétence naturelle sont exportées par la voie Com. Les protéines qui contiennent des séquences de liaison à la paroi cellulaire (CWB) sont retenues dans la paroi. Enfin certaines protéines peuvent être sécrétées dans le milieu extérieur en passant par les voies Sec, Tat ou par les transporteurs ABC.

RecO qui favorise la recombinaison intramoléculaire (Kidane *et al.*, 2009) et requiert l'action de RecU qui bloque l'activité de recombinaison homologue de RecA (Carrasco *et al.*, 2005).

C'est grâce à cette capacité de compétence naturelle que Spizizen et Anagnostopoulos ont pu développer leur protocole de transformation de *B. subtilis* en 1961 (Anagnostopoulos and Spizizen, 1961). Plus récemment, des cellules de *B. subtilis* « super-compétentes » ont pu être obtenues grâce à la surexpression de ComK et ComS, augmentant de 6 à 7 fois leur efficacité de transformation par rapport à la souche de référence *B. subtilis* 168 (Rahmer *et al.*, 2015)

### 1.5. Systèmes de sécrétion

Lorsqu'une protéine est synthétisée, elle peut rester dans le cytoplasme ou être dirigée, via différents systèmes de sécrétion, vers d'autres localisations subcellulaires comme la membrane plasmique, la paroi cellulaire ou le milieu extracellulaire. Chez *B. subtilis* il en existe 4 : (i) la voie Sec, (ii) la voie twin-arginine translocation (Tat), (iii) les transporteurs liés à l'ATP, appelés aussi transporteurs ABC (ATP-Binding Cassette) et (iv) la voie d'export par le pseudopilus Com pour la mise en place de la compétence naturelle (Ling Lin Fu *et al.*, 2007) (figure 7).

Dans cette partie ne seront détaillés que les systèmes de sécrétion principaux de *B. subtilis*, à savoir les voies Sec et Tat. La différence entre ces deux voies est que la voie Sec permet la sécrétion de protéines en cours de synthèse alors que la voie Tat permet la sécrétion de protéines déjà synthétisées et repliées.

La voie Sec est un des mécanismes d'export de protéines les plus conservés et constitue la voie de sécrétion la plus utilisée chez *B. subtilis*. Elle est divisée en 3 étapes : (i) ciblage, (ii) translocation et (iii) repliement et libération. Au cours de la première étape, le SRP (Signal recognition particle) reconnaît et cible le peptide signal positionné en N-terminal de la protéine à sécréter lorsque cette dernière émerge du ribosome. Le SRP va ensuite fixer la protéine d'amarrage GTPase FtsY qui va diriger l'ensemble du complexe vers le canal de sécrétion SecYEG. Ce dernier est composé des protéines membranaires SecY, SecE, SecG et SecDF. La seconde étape (étape de translocation) consiste à faire passer la protéine à travers le pore SecYEG. L'ATPase SecA associée à la membrane est capable de fixer à la fois le

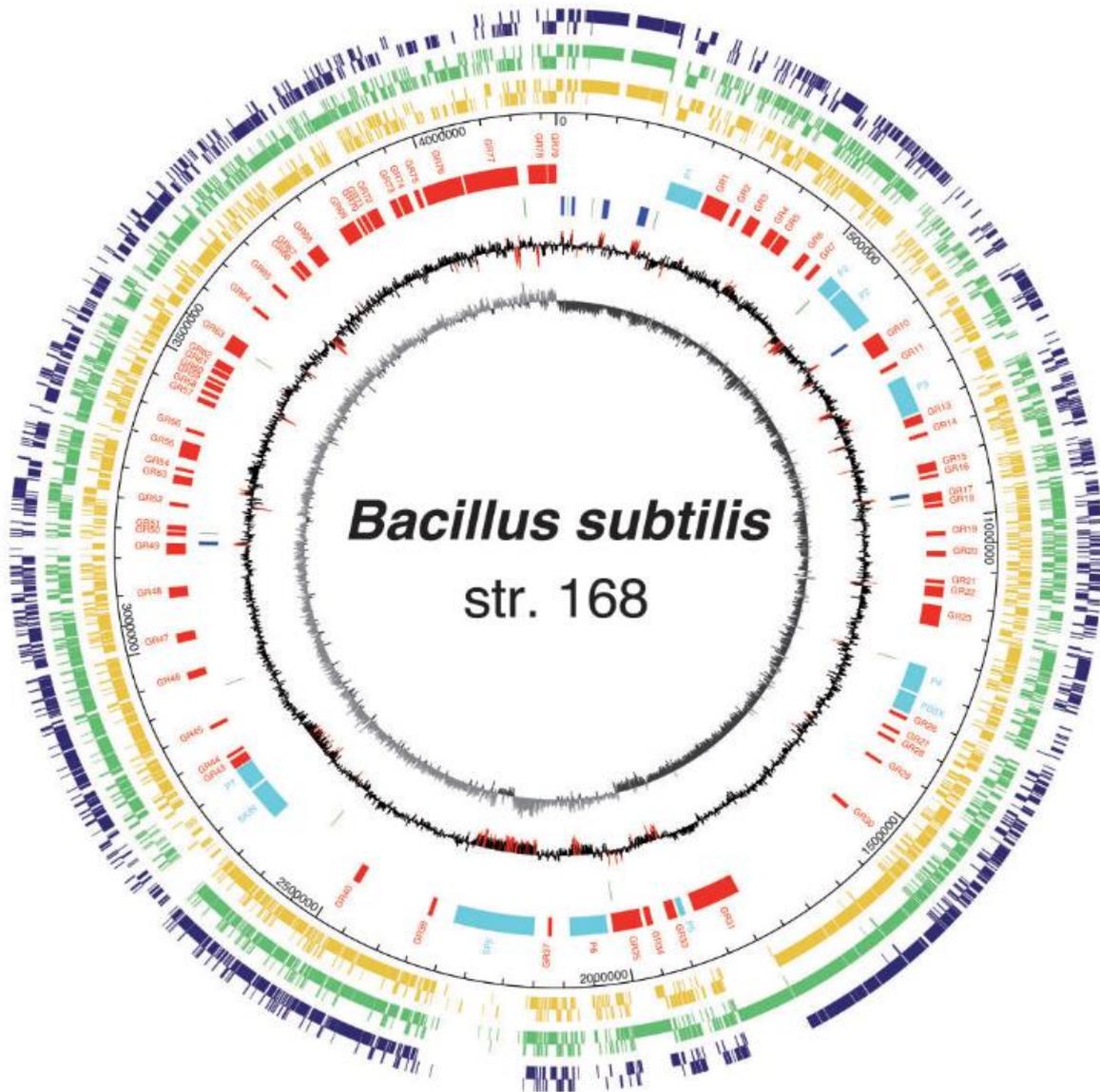


complexe SRP/protéine et le canal SecYEG. Elle agit comme un moteur pour la translocation en utilisant l'ATP comme source d'énergie. Après translocation, la troisième étape consiste à cliver le peptide signal par l'intermédiaire de peptide signal peptidase (SPases). Chez *B. subtilis*, cinq SPases de type I (SipS, SipT, SipU, SipV et SipW) ainsi qu'une de type II (LspA) sont disponibles pour ce clivage (Ling Lin Fu *et al.*, 2007).

La voie Tat permet le transport de protéines entièrement synthétisées et fermement repliées ainsi que celui de complexes enzymatiques multimériques (Van Dijl *et al.*, 2002). Ce système est composé d'un complexe d'ancrage et d'un pore. Il nécessite deux sous-unités transmembranaires, TatC et TatA. Chez *B. subtilis*, il existe deux gènes *tatC* codant les protéines TatCd et TatCy ainsi que trois gènes *tatA* codant TatAc, TatAd et TatAy. Ainsi, deux complexes protéiques formant des pores dans la membrane permettent d'avoir deux voies indépendantes TatAy-TatCy et TatAd-TatCd. La protéine à sécréter, présentant un peptide signal spécifique de cette voie, va interagir et s'ancrer au complexe Tat. La sous-unité TatC va insérer la protéine dans la membrane qui va ensuite pouvoir passer à travers le pore formé par la sous unité TatA. La translocation des protéines est effectuée ici en suivant un gradient de pH (Goosens *et al.*, 2014). Le peptide signal est ensuite clivé de la même manière que dans la voie de sécrétion Sec.

Pour que les protéines en cours de synthèse, ou complètement synthétisées, adoptent une bonne conformation, des protéines chaperonnes intracellulaires et extracytoplasmiques interviennent. Au niveau intracellulaire, les protéines chaperonnes de type « heat shock » synthétisées par l'opéron *dnak*, ainsi que les protéines GroES/GroEL synthétisées par l'opéron *groESL* et CsaA permettent de maintenir les protéines en cours de synthèse dans une conformation nécessaire à leur translocation, stimulant ainsi leur sécrétion. Une fois sécrétées, les protéines chaperonnes extracytoplasmiques entrent en jeu : (i) le facteur de repliement PrsA ainsi que les thiol-disulfure oxidoreductase BdbB/C/D permettent le bon repliement des protéines matures (Sarvas *et al.*, 2004), et (ii) les protéines HtrA et HtrB ainsi que WprA permettent le contrôle qualité de la conformation des protéines sécrétées et si besoin leur repliement, ou leur dégradation (Tjalsma *et al.*, 2004).

Après passage par ces systèmes et export des protéines à travers la membrane cytoplasmique, les nombreuses protéines sont finalement libérées dans l'environnement



**Figure 8 : Représentation du génome de *B. subtilis* souche 168 d'après Barbe *et al.*, 2009.** De l'intérieur vers l'extérieur, les données affichées en cercle sont : **(1)** le biais en GC, **(2)** la déviation en GC (contenu moyen en GC dans une fenêtre de 1 Kb par rapport à la moyenne en GC), **(3)** les gènes codant les ARNr (vert) et les ADNr (bleu), **(4)** les régions génomiques présentant des caractéristiques spécifiques les différenciant de la séquence moyenne. Les rectangles colorés en bleu clair indiquent les régions provenant de phages, **(5)** l'échelle utilisée et **(6, 7, 8)** les gènes ayant un orthologue présumé chez d'autres espèces de *Bacillus* (respectivement *B. licheniformis*, *B. amyloliquefaciens* et *B. pumilus*).

extracellulaire par diffusion passive à travers la couche de peptidoglycane (Green and Mecsas, 2016).

Une connaissance fine de ces systèmes de sécrétion permet de les améliorer en les modifiant par ingénierie. Par exemple en 2010, une étude a montré que la délétion des 61 acides aminés en C-terminal de SecA permet d'augmenter la production de plusieurs protéines hétérologues chez *B. subtilis* (Kakeshtia *et al.*, 2010). Plus récemment, l'ingénierie de la voie de sécrétion Tat, chez *Bacillus licheniformis*, pour la production d'arginase a permis une augmentation de l'activité enzymatique extracellulaire de 5,2 fois après modification (Ji *et al.*, 2024).

Cette capacité de sécrétion fait de *B. subtilis* l'une des bactéries les plus utilisées en industrie pour la production de molécules d'intérêt. Son importance dans le milieu industriel est détaillée partie 1.8 de cette étude bibliographique.

## 1.6. Données omics

### 1.6.1. Données de séquençage

En 1997, un consortium, constitué de 32 laboratoires et 2 entreprises en biotechnologie, a publié la séquence du génome complet de *B. subtilis* à partir de la souche 168 (Kunst *et al.*, 1997) (figure 8) . Ce travail eut un gros retentissement car ce fut le premier génome d'une bactérie Gram+ à être séquencé. Grâce à cet effort colossal, les premières caractéristiques du génome de *B. subtilis* ont pu être établies. Il a une taille de 4 214 810 pb avec un pourcentage en G+C de 43,5% mais qui varie considérablement le long du chromosome. En effet, plusieurs îlots riches en A + T ont été mis en évidence, pouvant être la signature de régions génomiques issues de bactériophages. En prenant en compte l'ensemble de ces régions ainsi que les bactériophages connus à cette époque, les auteurs décrivent que 10 de ces éléments seraient présents dans le génome de *B. subtilis*. Sur ce génome, environ 4 100 CDS (Coding DNA Sequence) ont pu être identifiées, couvrant 87% du génome. Cependant seulement 30% des gènes (~1200) avaient une fonction connue à l'époque. Des fonctions putatives ont pu être attribuées à certains gènes et une classification fonctionnelle des gènes codant des protéines a été réalisée, permettant de mieux comprendre les systèmes de régulations, les voies de



signalisation, le quorum sensing, ainsi que les voies de sécrétion. Ce premier séquençage a aussi permis d'obtenir des données sur l'organisation du génome et la synténie des gènes. La séquence du génome de *B. subtilis* issue de ce travail a été mise à disposition dans la base de données SubtiList.

Bien que primordial, ce premier travail de séquençage réalisé par une trentaine d'équipes, ayant chacune des expertises en matière de séquençage très variées et à une époque où le séquençage des génomes était un travail extrêmement difficile, présentait de nombreuses disparités. A partir du milieu des années 2000, de nouvelles techniques de séquençage (Margulies *et al.*, 2005) ainsi que des nouveaux systèmes d'annotation (Vallenet *et al.*, 2006) ont vu le jour. Ces méthodes ont ouvert la voie au re-séquençage et à la ré-annotation du génome de *B. subtilis*, qui ont été entrepris en 2009 par une équipe française (Barbe *et al.*, 2009). Ce travail a permis à la fois de préciser et de valider/invalidier les données obtenues au cours du premier séquençage en 1997, mais également de définir la fonction de plus de 400 nouveaux gènes. L'ensemble des gènes a été classé en catégories fonctionnelles, ce qui a permis une meilleure compréhension des gènes et fonctions essentiels à la vie (le paléome) ainsi que des gènes et fonctions liant un organisme à sa niche écologique (le cénome).

Suite à ce travail, les données présentes dans SubtiList furent mises à jour et la base de données SubtiList servit de modèle pour la création de GenoList (<http://genolist.pasteur.fr/>), un environnement intégré dédié à l'interrogation et à l'analyse des données génomiques d'espèces bactériennes, qui regroupait des informations sur plus de 700 génomes procaryotes (Lechat *et al.*, 2008), qui a été interrompu récemment. Parallèlement à SubtiList, est née également la base de données SubtiWiki, créée en 2009 par l'équipe de Jörg Stülke (Flórez *et al.*, 2009) et qui est aujourd'hui la base de données de référence pour *B. subtilis* (<https://subtiwiki.uni-goettingen.de/v4/>). Depuis sa création, cette base de données communautaire a été mise à jour continuellement. En 2018, elle contenait plus de 6000 pages d'informations sur les gènes, les structures opéroniques, les ADN<sub>r</sub>, les ARN<sub>t</sub>, les protéines, ainsi que des liens vers plus de 6800 références publiées. Des cartes interactives de 49 voies métaboliques et d'autres voies de régulation sont accessibles, ainsi que la séquence complète du génome (Zhu and Stülke, 2018).

Cependant, il est important de noter que les souches bactériennes évoluent rapidement en laboratoire. Il a aussi été observé qu'une même souche peut évoluer différemment entre



laboratoire, ce qui a été le cas avec la souche MG1655 de *E. coli* (Soupene *et al.*, 2003). Ces observations faites, le génome de *B. subtilis* a par la suite été ré-annoté plusieurs fois afin de le maintenir à jour.

L'obtention de toutes ces données de génomique grâce aux différents séquençages et aux ré-annotations du génome de *B. subtilis*, parallèlement au développement de nouvelles technologies à haut débit, a ouvert la voie à d'autres études « omiques » (transcriptomique, protéomique et métabolomique) et aboutissent aujourd'hui à une description de plus en plus précise du contenu d'une cellule de *B. subtilis*.

Les études transcriptomiques permettent de mieux comprendre les mécanismes qui régulent de façon dynamique les systèmes biologiques pour répondre aux signaux externes. Le projet Européen BaSysBio (Bacillus System Biology) lancé en 2006 a pour but de comprendre ces mécanismes chez *B. subtilis*. En 2012 un travail mené dans le cadre de ce projet rapporte les transcriptomes de *B. subtilis* dans différentes conditions environnementales et nutritionnelles (Nicolas *et al.*, 2012). Le site web dédié au projet BaSysBio regroupe les données acquises au cours de ce travail. De plus, l'ensemble des données transcriptomiques de *B. subtilis* est disponible dans les banques de données SubtiWiki et BDGatlas (Geissler *et al.*, 2021).

L'étude du protéome permet de mieux comprendre les mécanismes liés aux capacités de *B. subtilis* (comme celles décrites précédemment cf parties 1.3 ; 1.4 ; 1.5 de l'introduction). Un travail important a été mené en 2018 sur l'ensemble du protéome de *B. subtilis* permettant d'obtenir une carte quasi-complète de 3 159 protéines, qui couvre 75 % du protéome théorique de *B. subtilis* (Ravikumar *et al.*, 2018). L'ensemble des données générées a été déposé auprès du Consortium ProteomeXchange, une base de données protéomique mondiale. De plus, il existe d'autres bases de données comme par exemple PDB (Burley *et al.*, 2019) et UNIPROT (Consortium, 2007) qui sont aussi des outils puissants pour l'étude des protéines.

Récemment, une étude multi-omics alliant métabolomique et protéomique, a été réalisée chez *B. subtilis*. Elle a permis de mettre en évidence des différences dans les compositions métabolique et protéique entre les cellules végétatives et les spores de *B. subtilis*. Lors de ce



travail, une méthode d'extraction a été développée pour obtenir simultanément les profils du métabolome et du protéome intracellulaire chez *B. subtilis* (Huang *et al.*, 2024).

L'ensemble de ces ressources permettent de mieux comprendre la biologie de *B. subtilis* et facilite l'étude de cet organisme. La mise à disposition de tous ces outils permet aussi de placer *B. subtilis* comme le modèle des bactéries Gram+ en recherche fondamentale mais aussi avec un fort potentiel en recherche appliquée.

## 1.7. Ingénierie de génome chez *B. subtilis*

### 1.7.1. Méthodes de transformation

La transformation bactérienne est un phénomène qui permet à une bactérie de d'incorporer de l'ADN libre présent dans son environnement. Pour prélever de l'ADN dans le milieu, la bactérie doit être en état de « compétence », cet état peut être naturel ou « induit/forcé ».

Chez *B. subtilis* plusieurs protocoles de transformation existent : (i) le protocole de transformation basé sur la compétence naturelle, (ii) le protocole de transformation artificielle basé sur l'utilisation de poly-éthylène glycol (PEG), un agent fusogène (Boni and Hui, 1987), et (iii) le protocole de transformation par électroporation.

La méthode de transformation basée sur la compétence naturelle est la plus répandue, c'est également la première à avoir été mise au point suite aux travaux pionniers d'Anagnostopoulos et Spizizen (Spizizen, 1958, Anagnostopoulos and Spizizen, 1961). (cf partie 1.4 de l'introduction). Pour mettre au point ce protocole, ces chercheurs ont d'abord tenté de définir la phase de croissance au cours de laquelle les cellules sont en état de compétence et ont déterminé que celle-ci était associée à la fin de la phase exponentielle. Parallèlement, ils ont travaillé sur le développement de deux milieux de culture (MG1 et MG2) critiques pour leur protocole : le milieu MG1 est un milieu minimum enrichi, dans lequel les cellules sont mises en culture pendant 4 à 6 heures, et le milieu MG2 est un milieu carencé mimant des conditions environnementales pauvres en nutriments, dans lequel les cellules sont ensuite diluées et laissées pendant 1h30. C'est lors de cette deuxième culture que les



cellules entrent en état de compétence et peuvent être transformées avec de l'ADN (Anagnostopoulos and Spizizen, 1961).

En 2013, un second protocole basé sur la compétence naturelle est développé. Il est encore plus simple que celui proposé par Anagnostopoulos and Spizizen, et ne nécessite l'utilisation que d'un seul milieu, appelé le milieu MC (Modified Competence). Au cours de ce protocole, les cellules sont cultivées dans le milieu MC jusqu'à entrer en état de compétence, puis elles sont transformées avec de l'ADN (Konkol *et al.*, 2013).

Bien que les méthodes de transformation basées sur la compétence naturelle soient les plus utilisées, des protocoles de transformation artificiels ont également été développés. En 1979, Chang et Cohen mettent au point un protocole de transformation utilisant des protoplastes de *B. subtilis* et du PEG (Chang and Cohen, 1979). Les protoplastes sont des cellules dépourvues de paroi. Chez *B. subtilis*, ils sont obtenus par action du lysozyme dans un milieu hypertonique afin d'éviter la lyse des cellules. Plus précisément, au cours de ce protocole, les protoplastes de *B. subtilis* sont mis au contact d'ADN plasmidique en présence de PEG 40% pendant deux minutes. Les protoplastes sont ensuite étalés sur un milieu de régénération sans sélection pendant 1h30. Après régénération, les transformants sont sélectionnés par transfert des cellules sur un milieu gélosé contenant un antibiotique. Dans cette étude, les auteurs ont utilisé de l'ADN plasmidique, mais de l'ADN chromosomique peut également être utilisé (Lévi-Meyrueis *et al.*, 1980). La fusion de protoplastes différemment marqués en présence de PEG a également été rapportée (Schaeffer and Hotchkiss, 1978). Plus récemment, Kazuki Ohta et Norikazu Ichihashi ont décrit une méthode permettant de transformer *B. subtilis* avec plusieurs plasmides simultanément. Dans cette étude, des protoplastes de *B. subtilis* sont fusionnés, à l'aide de PEG, avec des liposomes contenant plusieurs plasmides permettant l'intégration simultanée de tous les plasmides (Ohta and Ichihashi, 2019).

Enfin, des protocoles de transformation de *B. subtilis* par électroporation sont également disponibles dans la littérature. Les premiers essais d'électroporation de *B. subtilis* ont été réalisés en 1989. Ces essais ont permis de définir les premières conditions nécessaires pour l'entrée d'ADN dans les cellules (Kusaoke *et al.*, 1989). Cependant, l'efficacité de transformation de cette méthode était assez faible, avec seulement  $1,5 \times 10^3$  transformants



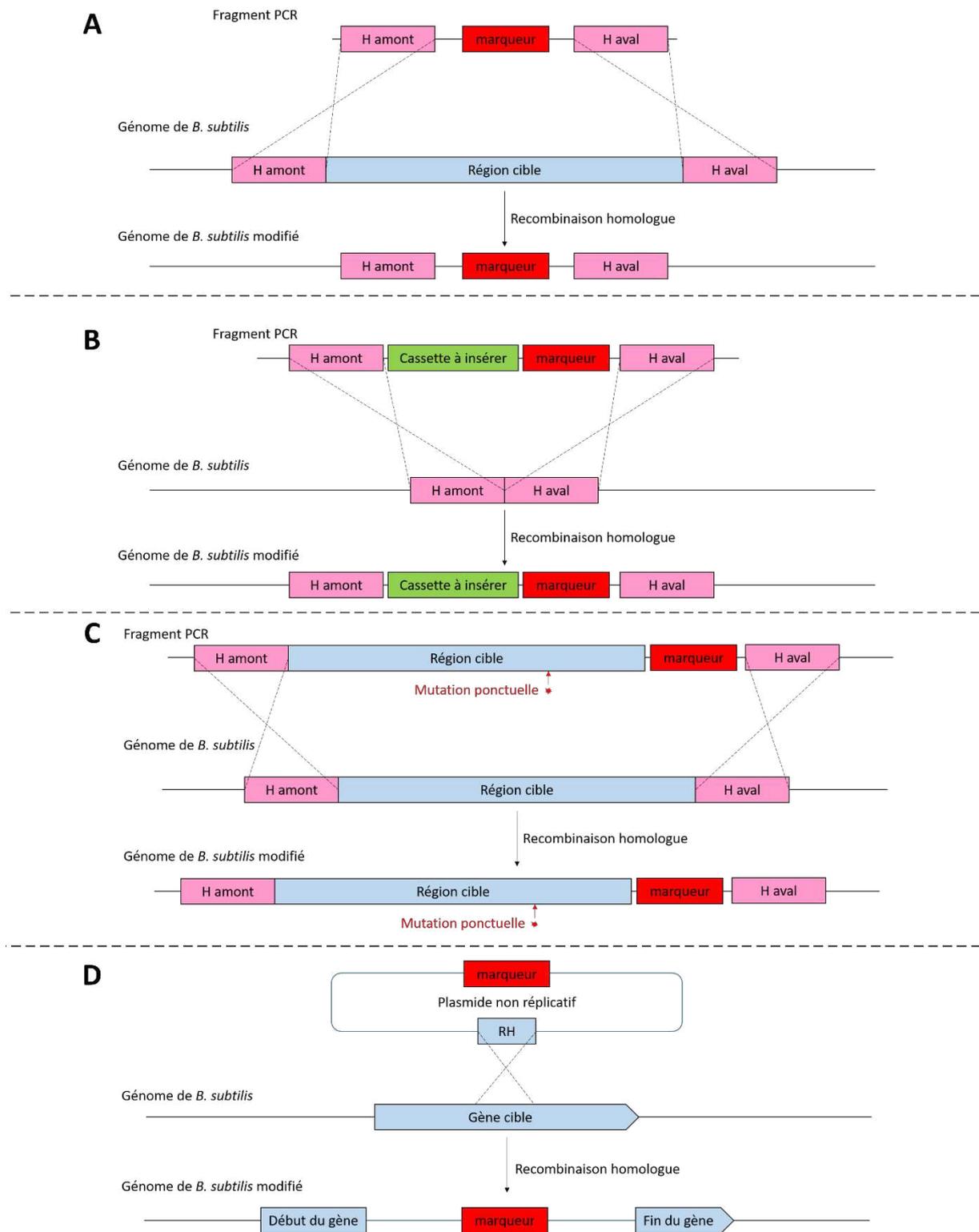
par  $\mu\text{g}$  d'ADN. Par la suite, plusieurs travaux ont été réalisés pour tenter d'augmenter l'efficacité de cette méthode : (i) en se plaçant dans des conditions avec une haute osmolarité dans les milieux d'électroporation et de régénération (Xue *et al.*, 1999), (ii) en utilisant des protoplastes, ou encore (iii) en protégeant au maximum les cellules des chocs électriques (Lu *et al.*, 2012). La transformation basée sur l'électroporation reste très peu utilisée chez *B. subtilis*.

L'ensemble de ces méthodes permettent de transformer *B. subtilis* avec différents types d'ADN, plasmidique, chromosomique ou même directement par fusion entre protoplastes. Cette capacité à pouvoir transformer de l'ADN chez *B. subtilis* a ouvert la voie à l'ingénierie génétique de cette bactérie.

### 1.7.2. Outils génétiques disponibles

**Mutagenèse aléatoire.** La mutagenèse aléatoire aux UV et aux rayons X a été la première méthode utilisée pour induire des mutations génétiques dans le génome de *B. subtilis*. Ce travail, réalisé en 1947 par Burkholder et Giles, a abouti à l'obtention de plusieurs mutants, dont la souche *B. subtilis* 168, auxotrophe pour le tryptophane et maintenant considérée comme la souche de référence (Burkholder and Giles, 1947). Plus tard, la mutagenèse aléatoire par traitement aux agents physico-chimiques a laissé place à la mutagenèse par transposition. Cette approche se base sur l'insertion d'éléments génétiques transposables, porteurs d'un gène de résistance à un antibiotique au hasard dans le génome pour inactiver des gènes. Plusieurs transposons issus de bactéries Gram+ se sont révélés fonctionnels chez *B. subtilis* (Tn916, Tn1545, Tn917) (Franke and Clewell, 1981; Courvalin and Carlier, 1987; Youngman *et al.*, 1983), mais également des transposons issus de bactéries Gram- (Tn10) (Petit *et al.*, 1990). Plus récemment, ce sont des transposons (famille des TnYLB) dérivant de l'élément transposable *mariner Himar 1*, isolé de la mouche *Haematobia irritans*, qui se sont révélés très efficace chez *B. subtilis*, montrant d'une part une fréquence d'insertion au génome élevée et d'autre part ne semblant pas avoir de sites d'insertion préférentiel (pas de « hotspot » d'insertion) (Pozsgai *et al.*, 2012; Le Breton *et al.*, 2006, ).

**Mutagenèse dirigée.** Comme exposé précédemment, *B. subtilis* est une bactérie capable d'importer de l'ADN exogène de manière naturelle. Mais c'est une bactérie également dotée



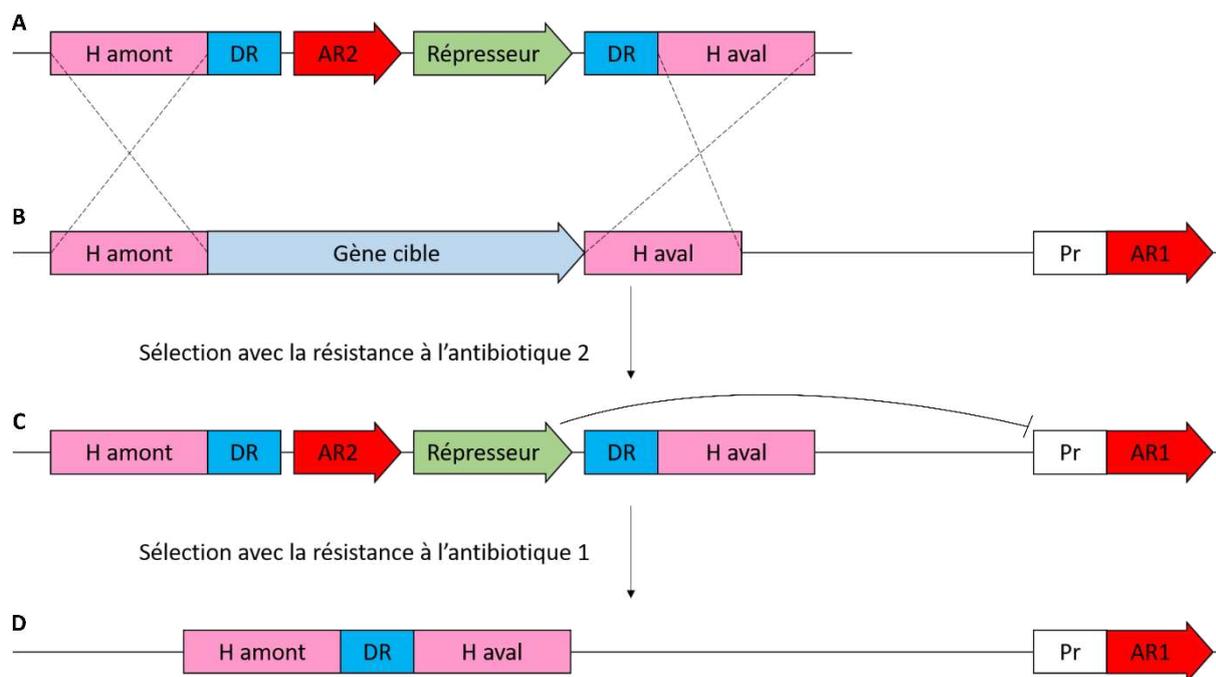
**Figure 9 : Mutagènesse diriégée chez *B. subtilis* basée sur sa capacité de recombinaison homologue.** [A] Délétion d'une région cible. [B] Insertion d'une cassette d'intérêt. [C] Mutation ponctuelle. [D] Inactivation d'un gène par intégration complète d'un plasmide via simple crossing over. « H amont » et « H aval » : région d'homologie entre la cassette à insérer et le génome cible, respectivement, en amont et en aval du gène cible. « RH » : région d'homologie.

d'une machinerie de recombinaison homologue très efficace. Ces deux propriétés ont permis la mise en œuvre de travaux de mutagenèse dirigée dont le but n'est plus d'induire des mutations au hasard dans le génome (cf. mutagenèse aléatoire), mais de modifier un gène ou une région particulière du génome. La méthode classique consiste à produire une cassette d'ADN (sous forme d'un fragment PCR ou bien d'une cassette portée par un plasmide) composée d'un marqueur de résistance à un antibiotique flanqué de séquences (500 pb à 1000 pb) (Westbrook *et al.*, 2016) identiques aux bornes du gène cible ou de la région cible. Après transformation, l'intégration de la cassette de recombinaison au génome par double crossing-over (échange allélique) peut conduire à la délétion d'un gène ou d'une région cible, à une insertion, ou encore produire des mutations précises (mutations ponctuelles) (figure 9A-C).

Chez *B. subtilis*, il est également possible d'inactiver un gène par recombinaison homologue via un simple crossing over (mutagenèse par insertion-duplication). Cette approche requiert généralement l'utilisation d'un plasmide non répliquatif et à faible nombre de copies (ex : pBR322 ou pSC101) portant un fragment interne du gène cible. Comme illustrée figure 9D, l'intégration complète d'un tel plasmide au gène cible (via un crossing-over) conduit à la formation de deux copies du gène cible dont l'une est tronquée en 5', et l'autre en 3', et donc à son inactivation. Chez *B. subtilis*, cette méthode a notamment été utilisée en 2003 pour définir l'ensemble minimal de gènes nécessaires au maintien de la vie de la bactérie en conditions de laboratoire (Kobayashi *et al.*, 2003).

Quelle que soit la méthode utilisée, chaque cycle de mutagenèse requiert l'utilisation d'un marqueur de sélection qui s'intègre au génome. Cela peut rapidement devenir une contrainte, d'une part parce que le nombre de marqueurs de sélection disponible n'est pas illimité, et d'autre part parce que les marqueurs de sélection qui sont, pour la plupart, des marqueurs de résistance aux antibiotiques, s'accumulent dans les génomes. Pour pallier ces problèmes, de nombreux efforts ont donc été réalisés dans la communauté pour développer des systèmes permettant de recycler les marqueurs de sélection et ainsi effectuer des modifications génétiques sans trace ni marqueur (Dong and Zhang, 2014; Wozniak and Simmons, 2022). Quelques stratégies couramment utilisées sont décrites ci-après :

**(1) Les stratégies d'ingénierie génétique qui reposent sur un système de sélection basé sur un système opérateur-répresseur.** Ces stratégies utilisent des cassettes



**Figure 10 : Stratégie d'ingénierie de génome basée sur un système opérateur-répresseur. [A]** Cassette d'intégration. **[B]** Génome cible. **[C] Etape 1** : Après recombinaison homologue entre la cassette d'intégration et le génome cible grâce aux bras d'homologie (H amont et H aval), la cassette est intégrée à la place du gène cible. Les transformants sont sélectionnés en utilisant la résistance à l'antibiotique 2 (AR2). A cette étape, le répresseur inhibe l'activité de promoteur (Pr) et les transformants sont sensibles à l'antibiotique 2. **[D] Etape 2** : Dans les conditions de sélection avec l'antibiotique 1, les transformants obtenus à l'étape 1 éliminent le marqueur de sélection 2 ainsi que le gène codant le répresseur via une recombinaison homologue entre les deux régions répétées (DR). Après recombinaison, le promoteur (Pr) du gène de résistance à l'antibiotique 1 (AR1) n'est plus inhibé, les transformants sont donc résistants à l'antibiotique 1. « AR » : gène de résistance à un antibiotique. « DR » : régions répétées. « H amont » et « H aval » : région d'homologie entre la cassette à insérer et le génome cible, respectivement, en amont et en aval du gène cible. « Pr » : promoteur régulé par le répresseur. La barre T indique une régulation négative.

d'ADN composites conçues comme illustré figure 10A. Brièvement, le principe est le suivant. Dans un premier temps, les transformants, dans lesquels la cassette complète s'est insérée au génome par double crossing-over, sont sélectionnés en utilisant le marqueur de résistance antibiotique AR2 sous le contrôle d'un promoteur constitutif. Dans ces cellules, le gène codant le répresseur est exprimé, ce qui bloque l'expression du gène AR1. Le marqueur AR2 peut être recyclé dans un second temps, en mettant les cellules dans un milieu contenant de l'antibiotique 1. En effet, dans un tel milieu, seules les cellules dans lesquelles un crossing-over entre les séquences DR (séquences répétées directes) aura eu lieu (=élimination du marqueur AR2 et du gène codant le répresseur) seront en capacité de se multiplier. Ce système permet *in fine* de produire une modification sans marqueur, en recyclant l'AR2 (figure 10). L'inconvénient majeur de ces méthodes est qu'elles requièrent une modification de la cellule en amont (intégration du marqueur AR1 sous le contrôle du promoteur Pr). Comme exemple de cette stratégie on peut citer quelques méthodes :

- La méthode basée sur le répresseur *cl* (du phage lambda) comme marqueur de contre sélection. Ce répresseur est capable de se fixer au promoteur ( $P_r$ ) du phage lambda pour bloquer son activité. C'est en 1999 qu'Itaya (Itaya, 1999) utilise ce principe en utilisant un gène de résistance à la néomycine sous le contrôle de  $P_r$ . Lorsque le gène *cl* est présent, les cellules sont sensibles à la néomycine. En revanche lorsque le gène est perdu, les cellules deviennent résistantes à la néomycine.
- La méthode basée sur l'enzyme UPRTase codée par le gène *upp*. Dans la voie de biosynthèse de la pyrimidine, cette enzyme catalyse la réaction de conversion de l'uracile en UMP, permettant d'utiliser l'uracile exogène. D'un autre côté cette enzyme, peut aussi convertir l'analogue toxique de la pyrimidine, le 5-fluorouracile (5-FU), en 5-fluoro-UMP, qui est un inhibiteur de la thymidylate synthétase. En 2002, Fabret *et al.*, (Fabret *et al.*, 2002) développent une méthode basée sur l'utilisation du gène *upp*. Ils créent une cassette contenant le



gène *upp* ainsi que le gène de résistance à la phléomycine (cassette *upp*), qu'ils lient avec un fragment PCR constitué du gène cible avec la mutation voulue. Cette cassette est ensuite intégrée dans le génome d'une souche de *B. subtilis*  $\Delta upp$  par recombinaison homologue avec double crossing over. A ce stade, les transformants sont sélectionnés par la résistance à la phléomycine. Ensuite, le gène de résistance à la phléomycine ainsi que *upp* sont excisés par recombinaison (simple crossing over) de séquences répétées présentes de part et d'autre de la « cassette *upp* », intégrées dans la construction de la molécule d'ADN transformée. L'ajout de 5-FU dans le milieu permet de sélectionner seulement les transformants dans lesquelles la « cassette *upp* » a été éliminée.

- Ces deux méthodes ont été combinées en 2013 pour créer une version améliorée de la stratégie d'ingénierie génétique basée sur un système opérateur-répresseur (Tanaka *et al.*, 2013).
- Les méthodes basées sur l'auxotrophie conditionnelle pour la lysine ou pour l'histidine couplée avec un système opérateur-répresseur (Dong and Zhang, 2014).

**(2) Les stratégies d'ingénierie qui reposent sur une sélection basée sur des systèmes toxine-antitoxine.** Brièvement, les systèmes toxine-antitoxine sont composés de deux gènes qui codent respectivement une toxine qui va entraîner la mort de la cellule et une antitoxine qui inhibe l'activité ou l'expression de la toxine. Ces systèmes sont classés en 6 types en fonction du mode d'action de la toxine. Le type II est le plus étudié et le plus répandu (Bordes and Genevaux, 2017). Le système *mazEF* identifié dans le génome de *B. subtilis* en fait partie. Dans ce système, la toxine *mazF* est capable de cliver les ARN messagers (ARNm), induisant la mort de la cellule. En 2006, Zhang *et al.*, l'utilisent comme marqueur de contre sélection chez *B. subtilis*. Dans leur méthode, ils utilisent la « cassette *mazF* » contenant le gène *mazF* sous le contrôle du promoteur inductible à l'IPTG ( $P_{spac}$ ) et le gène de résistance à la spectinomycine, le tout bordé de deux régions répétées. Cette cassette est ensuite flanquée



d'extrémités identiques aux bornes du gène cible. Une fois intégrée dans le gène cible par double crossing over, la « cassette *mazF* » peut ensuite être excisée par simple crossing-over entre les deux régions répétées qui la bordent. L'ajout d'IPTG dans le milieu permet de sélectionner les transformants qui l'ont bien éliminée (Zhang *et al.*, 2006).

**(3) Les stratégies d'ingénierie basées sur la recombinaison site-spécifique.** Ces systèmes utilisent des recombinases qui catalysent une recombinaison entre deux sites de reconnaissance spécifiques. Les systèmes de recombinaison site-spécifique, très utilisés dans la troisième partie de mon projet de thèse, seront détaillés dans la partie 3.2 de cette étude bibliographique (cf partie 3.2 de l'introduction). Chez *B. subtilis*, deux systèmes ont particulièrement été utilisés, les systèmes cre/lox et Xer/dif (Dong and Zhang, 2014). Nous focaliserons notre étude sur le système cre/lox, qui a notamment été utilisé pour la création de deux banques complètes de mutants par délétion de gènes, chacune d'elles correspondant à la délétion de plus de 3 900 gènes (Koo *et al.*, 2017). Dans cette étude, le gène cible est remplacé, par recombinaison homologue, par une cassette contenant un marqueur de résistance (érythromycine ou kanamycine) bordé de deux sites lox (lox66 et lox71). Cette cassette est ensuite éliminée par recombinaison site-spécifique en utilisant la Cre recombinase exprimée par le plasmide pDR244 thermosensible. La recombinase entraîne une recombinaison entre les sites lox66 et lox71 aboutissant à une cicatrice de 34 pb (lox72) et à l'élimination du gène de résistance. Le lox72 est dit « inactif » car la Cre recombinase a une très faible affinité avec ce site (Yan *et al.*, 2008). Après excision du gène de résistance, le plasmide pDR244 est éliminé par incubation des transformants à une température ne permettant pas sa réplication (42°C) (Koo *et al.*, 2017).

Enfin, plus récemment, l'application du système CRISPR (clustered regularly interspaced short palindromic repeat) à *B. subtilis* est venu enrichir sa boîte à outils génétiques. Le système CRISPR-Cas9 a été adapté à *B. subtilis* pour la première fois en 2016 par Westbrook *et al* (Westbrook *et al.*, 2016). Chez *B. subtilis* il existe trois types de stratégies d'ingénierie de



génomique basées sur le système CRISPR-Cas9 : (i) le premier est basé sur l'utilisation d'un seul plasmide pour l'expression à la fois de la Cas9 et de l'ARN guide (ARNg), (ii) le second est basé sur l'utilisation de deux plasmides permettant l'expression séparément de la Cas9 et de l'ARNg. Ce système est plus flexible que le premier et (iii) le troisième est lui basé sur l'intégration dans le génome du gène pour l'expression de la Cas9. Ce système est plus stable et plus efficace que les deux premiers mais nécessite une étape de pré-modification de la cellule pour l'insertion du gène *cas9* (Su *et al.*, 2020). En plus du système CRISPR-Cas9, d'autres outils ont été développés pour *B. subtilis*. On peut citer des outils permettant la régulation de l'expression des gènes comme les systèmes CRISPR interference (CRISPRi) et CRISPR activation (CRISPRa) (Song *et al.*, 2022), qui peuvent réguler l'expression des gènes positivement ou négativement. Ces systèmes sont composés d'une Cas9 dont l'activité endonucléase a été supprimée (dCas9) et d'un ARNg qui permet de cibler le gène à réguler. Des systèmes pour l'édition simultanée de plusieurs gènes ont aussi été développés, comme par exemple les systèmes CRISPR-Cas9n (D. Liu *et al.*, 2019) ou CRISPR-Cpf1 (Wu *et al.*, 2020).

Comme exposés dans cette partie, de nombreux outils existent pour modifier le génome de *B. subtilis*. Ils ont ouvert la voie à l'ingénierie de cette bactérie à des fins industrielles, notamment dans le but d'augmenter les rendements de production d'une molécule d'intérêt.

## 1.8. Applications en biotechnologie

De par ses nombreuses propriétés, *B. subtilis* est une bactérie qui suscite un fort intérêt chez les industriels : croissance rapide dans des milieux simples et peu coûteux, fermentation, systèmes de sécrétion de protéines efficaces, statut GRAS, outils génétiques disponibles, etc. Elle est d'ailleurs déjà utilisée dans des secteurs comme les biotechnologies, l'agriculture ou encore le milieu médical.

Dans le secteur des biotechnologies, *B. subtilis* est régulièrement utilisée pour la production d'enzymes et de protéines hétérologues. Les enzymes produites par *B. subtilis* représentent près de 50% du marché total des enzymes (Schallmey *et al.*, 2004). On peut citer par exemple des amylases, la lichenase, la  $\beta$ -galactosidase (Watzlawick and Altenbuchner, 2019), des cellulases (Thanh Bien *et al.*, 2014), ainsi que des protéases (Wang *et al.*, 2006). On peut noter aussi la production de lipases utilisées dans les secteurs de l'alimentation, du



biodiesel (biocarburant, alternative renouvelable au carburant classique) (Adetunji and Olaniran, 2021), et des détergents.

Le développement d'une telle industrie a été rendu possible en partie grâce aux outils génétiques disponibles pour modifier le génome de *B. subtilis* (cf partie 1.7.2 de l'introduction). C'est ainsi que des gènes ou groupes de gènes provenant d'autres organismes ont pu être intégrés au génome et ouvrir la voie à l'expression de protéines hétérologues chez cette bactérie (Cui *et al.*, 2018). De manière générale, la quantité de protéines hétérologues produites dépend du niveau d'expression du gène codant la protéine d'intérêt. Elle dépend également des systèmes qui peuvent dégrader la protéine. C'est pourquoi de nombreuses études ont été réalisées sur les promoteurs d'une part et sur des enzymes de dégradation comme les protéases (extracellulaires) d'autre part. L'ingénierie des promoteurs a permis de créer des gammes de promoteurs de telle sorte que différents objectifs de production peuvent désormais être atteints (expression forte, expression contrôlée, expression faible). Des mutants dépourvus de ces protéases ont également été construits, permettant de mieux préserver les protéines sécrétées dans le milieu extracellulaire. Outre les protéases, des travaux ont également été menés sur l'ingénierie des systèmes de repliement des protéines afin de pallier aux problèmes de dégradation des protéines hétérologues et d'augmenter les rendements (Schallmeyer *et al.*, 2004).

*B. subtilis* produit naturellement quelques antibiotiques, plus précisément des peptides microbiens à faible poids moléculaire et des bactériocines comme la surfactine, la bacilysine et la subtiline (Su *et al.*, 2020). Ces produits ont un fort potentiel dans le secteur de la santé humaine comme outils thérapeutiques pour la protection contre des micro-organismes pathogènes (Sumi *et al.*, 2015). *B. subtilis* produit aussi des vitamines comme par exemple la riboflavine et l'acide hyaluronique.

*B. subtilis* est également utilisée dans le secteur de l'agriculture, dans lequel elle est employée comme agent de biocontrôle. Elle agit en tant que biopesticide pour la protection de plantes contre des pathogènes par compétition avec d'autres micro-organismes et pour la synthèse d'antibiotiques. Elle est aussi exploitée en tant que biofertilisant : elle peut en effet favoriser la croissance de la plante en lui donnant accès à certains nutriments présents dans le sol (Akinsemolu *et al.*, 2024).



Grâce à son statut GRAS, elle est aussi utilisée en tant que probiotique. Les probiotiques sont des micro-organismes vivants ajoutés à certains produits alimentaires et bénéfiques pour la santé. En tant que probiotique, *B. subtilis* est souvent employée pour améliorer la digestion chez les animaux, ce qui favorise leur croissance et prévient les maladies (Lee *et al.*, 2019). Elle est aussi utilisée pour la restauration de l'eau et comme probiotique pour certaines espèces comme les crevettes ou les poissons (Olmos *et al.*, 2020).

Enfin, *B. subtilis* présente d'excellentes aptitudes pour la bioremédiation, qui consiste à décontaminer un milieu pollué par l'activité d'organismes vivants. En effet, *B. subtilis* possède plusieurs voies métaboliques lui permettant de dégrader des substances polluantes, comme par exemple des hydrocarbures, des composés pétroliers ou des pesticides, en composés moins polluants (Elenga-Wilson *et al.*, 2021).

Pour conclure, outre l'utilisation de *B. subtilis* comme hôte bactérien pour la production de molécules d'intérêts, cette bactérie est utile pour de nombreuses autres applications, par exemple en tant que probiotique, agent de biocontrôle, ou encore comme agent de décontamination de milieux pollués.

## 2. Biologie de synthèse

### 2.1. Historique

Le terme « Biologie de Synthèse » (BS) est apparu en 1912. Il a été proposé par le chimiste français Stéphane Leduc dans son ouvrage « La biologie synthétique » (Leduc, 1912) dans lequel il est écrit que : « Quand on est arrivé à connaître le mécanisme physique de la production d'un objet ou d'un phénomène, (...) il devient possible (...) de reproduire l'objet ou le phénomène, la science est devenue de synthèse. La biologie est une science comme les autres, (...) elle doit être successivement descriptive, analytique et de synthèse. ». D'après l'auteur, la capacité de reproduction ou de « synthèse » d'un système biologique, permet de valider la compréhension de ce dernier.

Mais ce n'est que dans les années 1960-1970 que se font sentir les prémices de ce qu'allait devenir, des années plus tard le domaine de la BS, avec dès 1961 les travaux de Jacob et



Monod sur la régulation de l'activité de gènes (Jacob and Monod, 1961). Dix ans plus tard, en 1971, Stanley Cohen construit un premier plasmide hybride et le transforme chez *E. coli*. L'année suivante, en 1972, les premières expériences de production d'ADN recombinant sont menées par Paul Berg et ses collègues de l'Université de Stanford. Enfin, en 1973, le premier gène étranger est cloné chez *E. coli* par Stanley Cohen et Herbert Boyer. Ces travaux marquent le début du génie génétique (BERCHE, 2016).

La notion de synthèse est ensuite introduite en 1979 avec la production du premier gène synthétique par H. Gobind Khorana (Kresge *et al.*, 2009). Dans les années 1980, des outils puissants facilitant l'ingénierie génétique sont mis au point comme par exemple, en 1987, la technique de la polymérase chain reaction (PCR) développée par Mullis (Mullis, 1987). Dans les années 1990, ce sont les méthodes de séquençage qui voient le jour et qui permettent, entre autres, le séquençage des premiers génomes bactériens en 1995 avec celui de *Haemophilus influenzae* (Fleischmann *et al.*, 1995) et *Mycoplasma genitalium* (Fraser *et al.*, 1995), du génome de *S. cerevisiae* par un consortium de plusieurs laboratoires en 1996 (Goffeau *et al.*, 1996) et d'*E. coli* en 1997 (Blattner *et al.*, 1997). L'accès à ces séquences offre une meilleure compréhension de l'organisation des gènes et des génomes.

La manipulation des systèmes biologiques par génie génétique est progressivement reconnue comme une approche permettant d'étudier les systèmes naturels et de créer des réseaux de régulation artificiels pour de potentielles applications en biotechnologie et en santé. A la fin des années 1990 certains ingénieurs, physiciens, informaticiens et mathématiciens saisissent cette opportunité et rejoignent la discipline (Cameron *et al.*, 2014). Grâce à l'ensemble de ces nouvelles expertises, les fondements du domaine de la BS sont réellement établis.

Depuis l'ingénierie des premiers circuits génétiques dans les années 2000 (McAdams and Arkin, 2000), le domaine de la BS n'a cessé de se développer. On peut notamment souligner l'importance de l'ingénierie des voies métaboliques qui offre un aspect appliqué à la BS. En effet, en 2003, des travaux décrivent la synthèse hétérologue d'un précurseur de l'artémisinine (médicament contre la malaria) par *E. coli* (Martin V *et al.*, 2003), permettant la production du médicament en grande quantité. Par la suite, ce médicament a aussi pu être produit en utilisant des levures modifiées (Paddon *et al.*, 2013). On peut citer comme autres exemples, l'intégration de voies métaboliques chez *E. coli* pour la production de bio-carburant

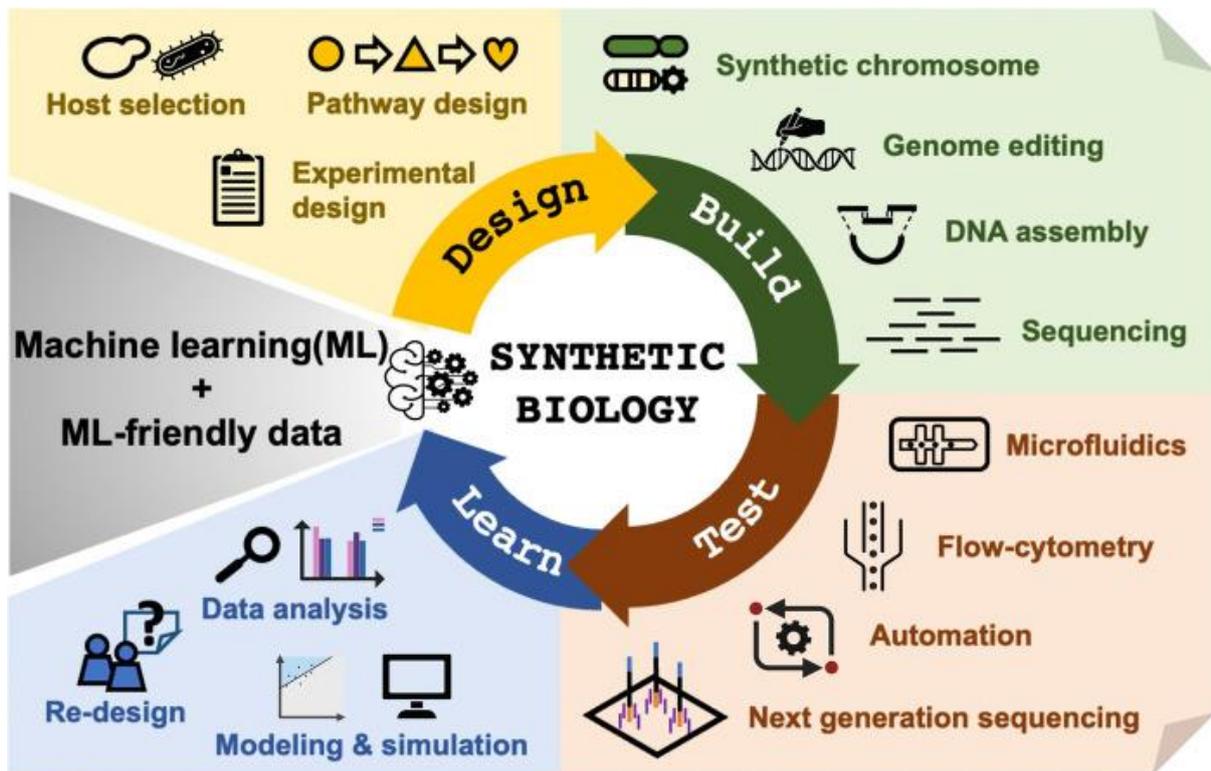


Figure 11 : Cycle DBTL appliquée à la Biologie de Synthèse d'après Kitano *et al.*, 2023. On y retrouve un panel de techniques réparties dans les quatre étapes du cycle. Un petit encart est ajouté montrant l'importance des approches de « machine learning » (ML) dans l'étape d'apprentissage (learn).

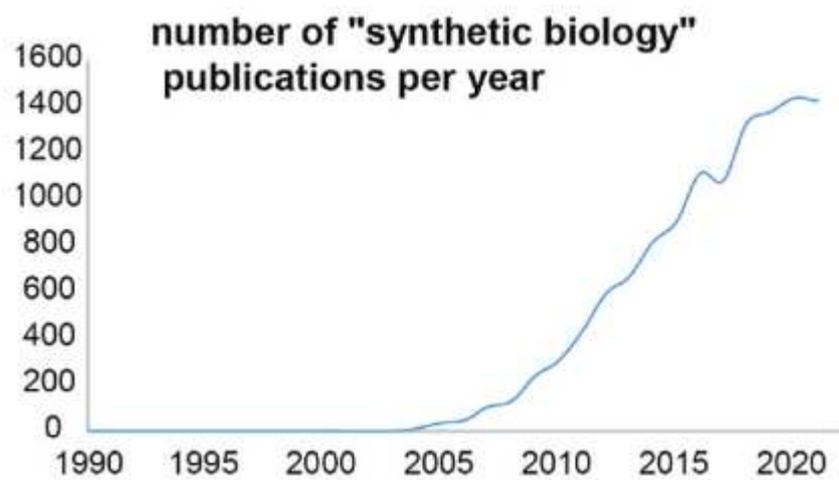
(Atsumi *et al.*, 2008) ou encore, chez la levure, la synthèse d'opiacés (Galanie *et al.*, 2015) ou de précurseurs pour la production de vinblastine, un médicament anti-cancéreux (Zhang *et al.*, 2022).

D'un point de vue plus fondamental, en 2010, une équipe du JCVI (John Craig Venter Institute) a construit la première cellule contrôlée par un génome synthétique, JCVI Syn1.0 (1,078 Mb) (Gibson *et al.*, 2010). Quelques années plus tard, en 2016, la construction d'une cellule contrôlée par un génome synthétique quasi-minimal JCVI Syn3.0, d'une taille de 531 Mb, est réalisée par la même équipe (Hutchison *et al.*, 2016). Par la suite, en 2019, est décrite la construction d'une cellule de *E. coli* contrôlée par un génome synthétique réduit et recodé, dans lequel 3 des 64 codons du code génétique standard ont été éliminés (Fredens *et al.*, 2019). Ces travaux ont joué un rôle central dans le développement d'une des disciplines phares du domaine de la BS, à savoir la construction de cellules artificielles.

Parallèlement à ces avancées, le développement de la puissance de calcul informatique a permis d'améliorer les capacités de prédiction et de modélisation, aboutissant notamment en 2012 à un modèle prédictif du phénotype de cellule entière de *M. genitalium* (Karr *et al.*, 2012).

L'acquisition des connaissances et les expériences menées en BS sont souvent basées sur un cycle « Design Built Test Learn » (DBTL). Ce cycle correspond à une approche d'apprentissage itératif basée sur des données bio-informatiques et expérimentales pour la résolution d'un problème. Cette méthode permet l'amélioration du système étudié en fonction des résultats expérimentaux obtenus. En BS, ce cycle s'appuie sur l'ensemble des techniques et méthodes développées, comme par exemple la synthèse d'ADN et le séquençage. De plus, des approches de « machine learning » (ML), rendues possibles grâce à la puissance de calcul en informatique, prennent une part de plus en plus importante dans ce cycle, notamment dans la prédiction et la modélisation de phénomènes d'intérêts (Kitano *et al.*, 2023) (figure 11).

Depuis les années 2000, les scientifiques se sont regroupés, formant ainsi une réelle communauté de chercheurs en BS. On peut retracer l'organisation de la première conférence internationale de BS en 2004 au MIT (Massachusetts Institute of Technology), qui rassemblait toute la communauté, à l'époque très hétérogène, composée de biologistes, chimistes,



**Figure 12 : Graphique représentant le nombre de publications de biologie de synthèse par an de 1985 à 2021 d'après Donati *et al.*, 2022.** Les données ont été obtenues en cherchant dans Web of Science les publications qui incluent « synthetic biology » dans le titre, le résumé ou les mots clés.

physiciens, informaticiens, mathématiciens et ingénieurs. La même année, est également organisée la première compétition internationale de BS, l'IGEM (International Genetically Engineered Machine) (Cameron *et al.*, 2014). Cette compétition dédiée aux étudiants (Master/licence) a ensuite été renouvelée tous les ans. Enfin, l'engouement pour le domaine de la BS est aussi marqué au niveau des publications, avec depuis 2005, une forte augmentation du nombre de publications en lien avec le domaine de la BS (Donati *et al.*, 2022) (figure 12).

La définition de la BS n'est pas claire et dépend des instances, de la réglementation, des personnes (Arkin *et al.*, 2009), et évolue en fonction de la période et des connaissances. Cependant, deux grands objectifs se démarquent : (i) fabriquer ou modifier des systèmes biologiques existants dans la nature, et (ii) concevoir ou créer de nouveaux systèmes ou composants qui ne sont pas trouvés dans la nature. La BS est un domaine scientifique au champ d'application très large (industrie, santé, agriculture). Elle combine un ensemble d'approches issues de la biologie et de l'ingénierie dans le but d'étudier et d'améliorer les systèmes biologiques. L'ambition majeure de la BS est d'apporter des solutions nouvelles (et respectueuses de l'environnement) aux problématiques sociétales actuelles, telles que la lutte contre pollution, la crise énergétique, et les maladies.

## 2.2. La quête de la cellule artificielle

Les scientifiques sont captivés par la perspective de créer une cellule entièrement synthétique, qui pourrait révolutionner la biologie, la médecine et l'économie. On peut imaginer ces cellules sous une variété de formes infinie : allant de structures micrométriques semblables à nos cellules et capables de reproduire les fonctions essentielles des cellules vivantes, jusqu'à des matériaux d'ingénierie qui n'imiteraient que certaines des propriétés de la cellule comme la forme ou la motilité. Cependant, même les plus petites unités de vie que nous trouvons aujourd'hui sur notre planète sont encore terriblement complexes. Par conséquent, certains scientifiques se sont attachés à définir « l'essence » d'une cellule, c'est-à-dire à comprendre la vie sous sa forme la plus basique, qui pourrait être plus facile à construire.



Une cellule minimale est une cellule qui ne retiendrait que les fonctions essentielles à sa croissance en milieu axénique (dans des conditions de laboratoire idéales, c'est-à-dire dans un environnement sans stress et avec tous les nutriments nécessaires). Ainsi définie, la cellule minimale ne possède que la machinerie nécessaire à la vie sans redondance inutile et sans fonction surnuméraire. Ce concept est déjà ancien, il a vu le jour dans les années 1930 lorsque le physicien Max Delbrück fonde l'American Phage Group composé de physiciens, chimistes, et biologistes, dont l'objectif était la compréhension des principes de base de la vie par l'étude de systèmes biologiques simples, les bactériophages. Ils adoptèrent une approche réductionniste pour comprendre le fonctionnement de la vie en tentant d'identifier et de déterminer le fonctionnement de chaque élément de leur système modèle (Glass *et al.*, 2017; Morange, 2000).

Dans la même idée, Morowitz propose en 1984 d'utiliser les mycoplasmes comme modèle pour la compréhension des lois fondamentales du vivant (Morowitz, 1984). En effet, les mycoplasmes sont des bactéries appartenant à la classe Mollicutes. Elles sont définies par un génome de taille réduite (0.58 Mb à 1.35 Mb), une absence de paroi cellulaire, et un métabolisme simple, ce qui en fait les formes de vie les plus petites et plus simples capables d'autoréplication sur la planète. Bien que ces bactéries aient un mode de vie parasitaire, la plupart peuvent être isolées et cultivées en laboratoire (Citti and Blanchard, 2013). Ces caractéristiques font des mycoplasmes des modèles idéaux pour définir l'ensemble minimal des gènes et/ou fonctions nécessaires à la vie. Parmi les mycoplasmes, *M. genitalium* avec une taille de génome environ égale à 580 kb, est le meilleur représentant du concept de cellule minimale.

Dans le but de construire une cellule minimale, la première étape consiste à identifier les gènes et les classer. Trois catégories ont été définies: (i) les gènes essentiels (E) dont l'inactivation entraîne la mort de la cellule, (ii) les gènes non essentiels (NE) dont l'inactivation n'a aucun effet sur la viabilité ou la croissance des cellules dans un environnement donné, et (iii) les gènes quasi-essentiels (QE) dont l'absence affecte la croissance des cellules mais qui ne sont pas strictement essentiels (Glass *et al.*, 2017; Hutchison *et al.*, 2016; Lluch-Senar *et al.*, 2015).

Avec le développement des méthodes de séquençage, les séquences complètes des génomes de plusieurs bactéries avaient été obtenus, en particulier celles d'*H. influenzae*, une



bactérie Gram- (Fleischmann *et al.*, 1995) et *M. genitalium* une bactérie qui dérive d'un ancêtre commun aux bactéries Gram+ (Fraser *et al.*, 1995). En 1996, la comparaison de ces deux génomes par une approche de bio-informatique, établit la liste d'un premier ensemble de 256 gènes essentiels conservés entre ces deux espèces éloignées d'un point de vue phylogénétique. En 1999, des approches expérimentales d'inactivation de gènes par transposition sont menées chez *M. genitalium* et *Mycoplasma pneumoniae*, deux mycoplasmes très proches. Elles ont permis d'estimer un nombre de gènes non-essentiels chez ces espèces qui se situerait entre environ 180 et 215, et entre 265 et 350 pour le nombre de gènes essentiels (Hutchison *et al.*, 1999). Dans les années qui ont suivi, des approches similaires d'inactivation de gènes ont été réalisées chez de nombreuses autres bactéries, toujours dans le but de définir les gènes essentiels à la vie. Par exemple, chez *Staphylococcus aureus*, une approche expérimentale utilisant l'ARN anti-sens pour inhiber l'expression des gènes a permis d'identifier 150 gènes essentiels (Ji *et al.*, 2001). Chez *B. subtilis*, une méthode d'inactivation de gènes par intégration de plasmide a conduit à l'identification de 271 gènes essentiels (Kobayashi *et al.*, 2003). Enfin, dans la cellule synthétique JCVI Syn1.0, 432 gènes ont été identifiés comme non essentiels, 240 comme essentiels et 229 comme quasi-essentiels. Ce travail a notamment permis la construction de la cellule synthétique réduite JCVI Syn3.0 contrôlée par un génome quasi-minimal de 531 kb et 473 gènes, ce qui en fait la cellule ayant le plus petit génome parmi toutes les cellules naturelles capables de réplication autonome en milieu axénique (Hutchison *et al.*, 2016).

Après avoir identifié les gènes essentiels, deux approches opposées peuvent être utilisées pour l'étape de construction à proprement parler de la cellule avec un génome réduit :

- L'approche « top down » qui consiste à modifier un système biologique complexe existant dans la nature pour obtenir un système plus simple en retirant tous les composants non essentiels. Comme exemple, on peut citer le travail de réduction itérative du génome de *B. subtilis* par « Pop-in/pop-out » qui a abouti à l'obtention de 298 souches à génome réduit. Au cours de ce travail, le génome de *B. subtilis* a pu être réduit à une taille de 2.73 Mb soit une réduction de 1.48 Mb par rapport à la souche de référence *B. subtilis* 168 (4.215 Mb) (Dervyn *et al.*, 2023). D'autres travaux ont permis d'atteindre une réduction du génome de *B. subtilis* d'environ 36 % par rapport



à la souche de référence *B. subtilis* 168 avec l'élimination de 1605 gènes en 94 délétions, aboutissant à un génome de 2.68 Mb (Reuß *et al.*, 2017).

- L'approche « bottom up » est opposée et consiste à assembler des composants de base ayant des fonctions bien définies (bio-briques) afin de reconstruire un système complet et recréer les fonctions du vivant. En théorie, la construction d'une cellule artificielle nécessite l'apport de 3 ingrédients majeurs : (i) un génome artificiel capable d'assurer les fonctions essentielles de la cellule, comme la réplication et la transcription/traduction (TX/TL) pour initier l'expression du génome, (ii) un métabolisme qui favorise la croissance et l'homéostasie, et (iii) un compartiment artificiel mimant une membrane cellulaire. Les résultats qui se rapprochent le plus de cet objectif correspondent à la création de la première cellule contrôlée par un génome synthétique, JCVI Syn1.0 (Gibson *et al.*, 2010), puis la construction de la première cellule contrôlée par un génome synthétique réduit, JCVI Syn3.0 (Hutchison *et al.*, 2016). Ces travaux montrent qu'il est possible de synthétiser un génome complet et de le transplanter dans une cellule dite « receveuse » pour obtenir, par sélection, une cellule contrôlée par le génome synthétique. La mise au point de cette méthode a ouvert la voie à de l'ingénierie de génome à grande échelle. Il est cependant trop tôt pour parler de cellules complètement artificielles car les cellules créées lors de ces travaux sont certes contrôlées par un génome synthétique mais ont été obtenues à partir de cellules receveuses déjà vivantes.

Une fois obtenues, les cellules minimales sont aussi capables d'évolution. En effet, des études ont été menées sur l'évolution de la cellule minimale JCVI Syn3B. Cette cellule a un génome de 543 kb contenant 493 gènes et diffère de JCVI Syn3.0 (531 kb) par la présence de 19 gènes supplémentaires non essentiels, ce qui la rend plus facilement manipulable (Breuer *et al.*, 2019). Dans ce travail, la capacité d'évolution de JCVI Syn3B a été comparée à celle de la souche parentale *Mycoplasma mycoides*. Les résultats de ces expériences ont montré que, dans les conditions utilisées, la cellule minimale évoluait 40 fois plus vite que la cellule parentale, et que la sélection naturelle permettait à la cellule minimale de regagner une *fitness* correcte en 2 000 générations (Moger-Reischer *et al.*, 2023).



### 2.3. Les châssis en Biologie de Synthèse

Au sens premier du terme, un châssis est un cadre rigide, fixe ou mobile, destiné à entourer ou supporter quelque chose. Par exemple, un châssis en automobile est destiné à supporter les mécanismes (moteur, direction, etc...), les roues, les suspensions et la carrosserie du véhicule.

En BS, on peut distinguer deux types de châssis, chacun ayant des fonctions différentes et ses propres avantages et inconvénients :

- Les châssis à base de nanoparticules qui correspondent à des compartiments artificiels très minimaux, principalement utilisés pour l'administration de produits biologiques et de petites molécules thérapeutiques (Batty *et al.*, 2021).
- Les châssis microbiens qui sont des hôtes cellulaires modulables utilisés pour héberger et maintenir stable des circuits génétiques qui permettent de réaliser ou programmer des opérations plus ou moins complexes. Ils sont nécessaires à la propagation de ces circuits ainsi qu'à leur activité en fournissant les ressources nécessaires à leur fonctionnement (Kim *et al.*, 2016).

Dans cette partie, nous nous concentrerons spécifiquement sur les châssis microbiens. Pour qu'un châssis de ce type soit considéré comme idéal, l'organisme choisi doit posséder certaines propriétés : (i) il doit avoir un mode de vie simple nécessitant des besoins nutritionnels simple, (ii) être robuste notamment grâce à une paroi cellulaire solide, (iii) être apte à résister à des stress physico-chimiques et à des conditions extrêmes (température élevée, pH, etc...) et enfin (iv) être capable de se multiplier rapidement. De plus, il est nécessaire d'avoir à disposition (i) des connaissances suffisantes sur la biologie de cet organisme, notamment au niveau de son réseau de régulation ainsi que sur ses potentielles interactions avec des systèmes exogènes, et (ii) des outils de manipulation génétique afin de pouvoir le modifier si besoin.

Certains organismes comme *E. coli* (Ahmadi and Pfeifer, 2016), *B. subtilis* (Y. Liu *et al.*, 2019) ou *S. cerevisiae* (Billingsley *et al.*, 2016) sont couramment utilisés en BS pour de nombreuses applications. Cependant plusieurs autres espèces peuvent également être utilisées et offrir des avantages, comme par exemple *Pseudomonas putida* (Kim *et al.*, 2016) ou certaines espèces de *Streptomyces* (Komatsu *et al.*, 2010). Une notion importante à



prendre en compte est qu'il n'existe pas, du moins pour le moment, de châssis microbien idéal pour des applications universelles. Tous les châssis microbiens possèdent leurs avantages et leurs inconvénients, il est donc impossible de considérer un châssis comme étant « le meilleur » pour la BS en général. Un organisme peut tout de même être considéré comme modèle dans un domaine spécifique ou pour un objectif spécifique (Kim *et al.*, 2016).

De cette notion de châssis microbien est née la notion de « châssis simplifiés » qui correspond à un organisme contrôlé par un génome réduit, et de « châssis minimaux » qui correspond à une cellule minimale (cf partie 2.2 de l'introduction). L'idée derrière la minimisation du génome du châssis est de pouvoir réduire au maximum les interférences potentielles entre les systèmes biologiques nouvellement intégrés et le métabolisme de l'hôte.

*B. subtilis* est considérée comme un châssis microbien à fort potentiel pour plusieurs raisons. Tout d'abord, c'est une bactérie modèle pour laquelle nous disposons de connaissances fondamentales considérables mais également de connaissances techniques poussées (méthodes et outils de manipulation génétique bien établis). Ensuite, c'est une des rares bactéries à avoir le statut GRAS conféré par l'Agence américaine des produits alimentaires et médicamenteux (FDA). Enfin, ses capacités notamment à importer de l'ADN exogène (compétence naturelle), et à sécréter des protéines hétérologues directement à l'extérieur de la cellule (ce qui permet d'éviter les coûts associés à la purification) lui donnent des avantages que d'autres bactéries n'ont pas (cf parties 1.4 et 1.5 de l'introduction). C'est pourquoi *B. subtilis* est déjà largement utilisée dans l'industrie pour la production de molécules d'intérêt (Schallmey *et al.*, 2004). Les avancées dans l'amélioration de châssis de *B. subtilis* tendent à réduire son génome afin de produire des châssis simplifiés de cette bactérie. En effet, en 2008 une étude rapporte qu'une délétion de 874 kb (~20%) du génome de *B. subtilis* a permis une production de cellulases et de protéases respectivement 1.7 et 2.5 fois supérieure à la souche de référence 168 (Morimoto *et al.*, 2008). Les résultats de ce travail, combinés avec les études sur l'essentialité des gènes chez *B. subtilis* (Kobayashi *et al.*, 2003) et les méthodes de réduction de son génome (Reuß *et al.*, 2017; Dervyn *et al.*, 2023), permettent maintenant d'imaginer la construction de châssis ultra-simplifiés voir minimaux pour l'amélioration des souches de *B. subtilis* à visée industrielle.

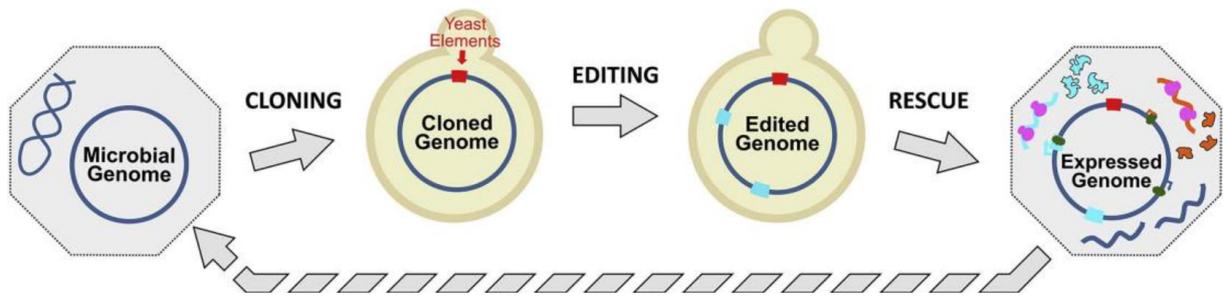


Figure 13 : Cycle d'ingénierie de génome bactérien en utilisant la levure comme plateforme de modification d'après Vashee *et al.*, 2020. Cloning : clonage, Editing : édition, Rescues : sauvetage.

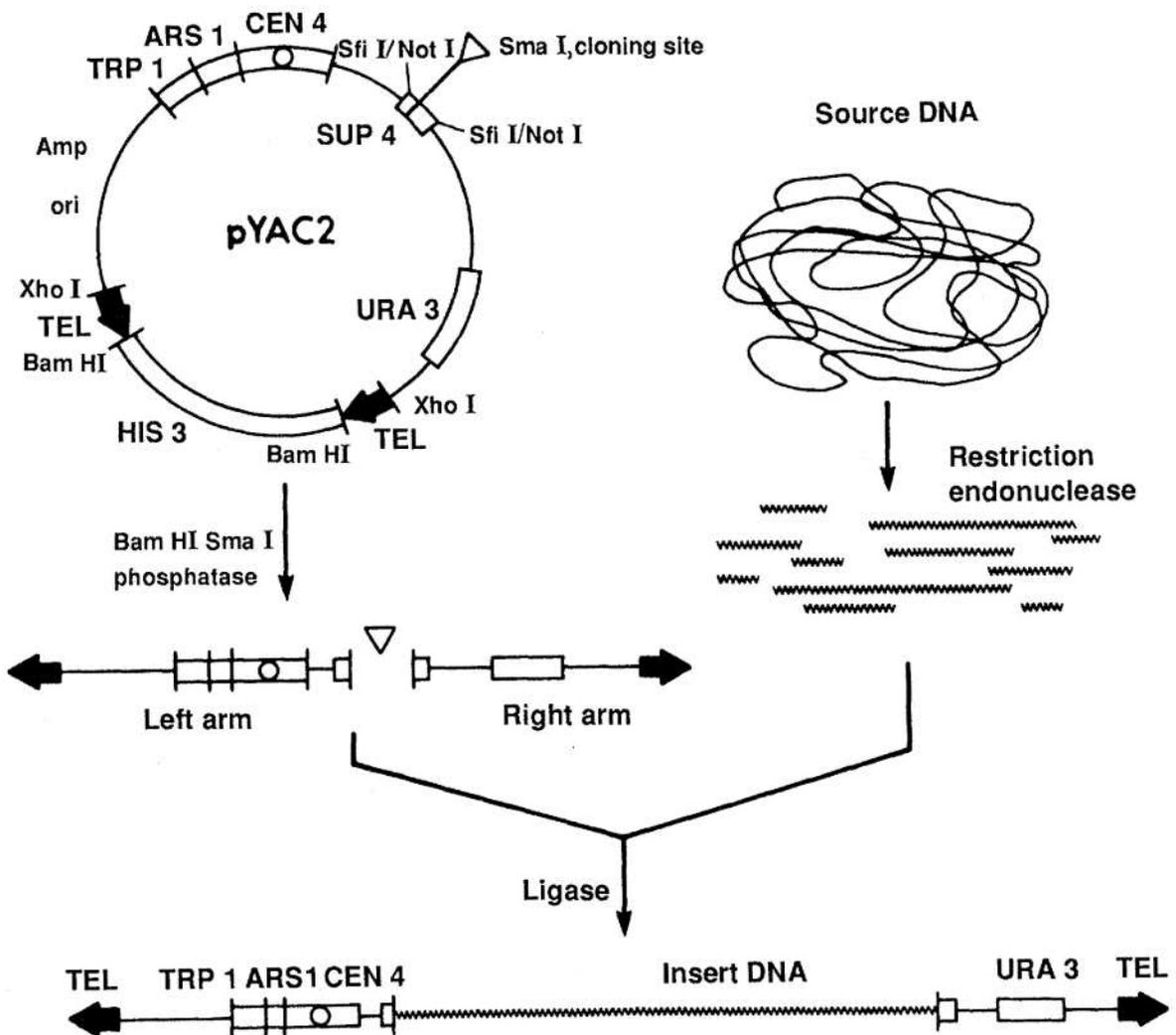


Figure 14 : Système de clonage d'un fragment d'ADN dans un YAC, d'après Burke *et al.*, 1987. Le vecteur pYAC2 contient une ARS, un centromère (CEN), un site de clonage pour l'ADN à capturer, trois marqueurs d'auxotrophie (HIS3, URA3 et TRP1) ainsi que deux séquences télomérique (TEL). Les deux bras du YAC sont obtenus par digestion de ce plasmide par les enzymes de restriction BamHI et SmaI. Les sites de digestion des enzymes de restriction BamHI et SmaI sont représentés sur la carte du plasmide. Le bras gauche (left arm) est composé d'une séquence de télomère, du marqueur TRP1, ainsi qu'ARS et CEN. Le bras droit (Right arm) est composé du marqueur URA3 et d'une séquence de télomère. Le fragment d'ADN d'intérêt obtenu par digestion enzymatique de l'ADN source, est capturé entre ces deux bras par l'action d'une ligase.

## 2.4. La levure comme plateforme d'ingénierie

La levure *S. cerevisiae* est un eucaryote unicellulaire, de forme ovoïde ou arrondie, d'une longueur de 6 à 12  $\mu\text{m}$  et d'une largeur de 6 à 8  $\mu\text{m}$ . Elle possède un génome d'une taille de 12 Mb, composé de 16 chromosomes linéaires pour environ 6 600 gènes. Son plus petit chromosome a une taille de 0.230 Mb et son plus grand 1.531 Mb (Engel *et al.*, 2014). Son génome est bien répertorié dans la banque de donnée SGD (Saccharomyces Genome Database) (<https://www.yeastgenome.org/>). *S. cerevisiae* est utilisée en biologie cellulaire et génétique en tant que modèle pour les eucaryotes. C'est également un hôte pour la propagation de molécules d'ADN et même de génomes complets.

Certains micro-organismes sont difficiles à cultiver en laboratoire ou à modifier sur le plan génétique. Or, il a été montré que le clonage de tels génomes dans la levure peut permettre de les manipuler plus facilement (Benders *et al.*, 2010) Mais pour que l'utilisation de la levure *S. cerevisiae* en tant que « plateforme d'ingénierie de génomes (ou fragments d'ADN) » exogènes ait un réel intérêt, il est en réalité nécessaire de passer par trois phases : (i) le clonage ou l'assemblage dans la levure de la molécule d'ADN ou du génome bactérien d'intérêt, (ii) l'édition de cette molécule d'ADN en utilisant les outils génétiques disponibles chez la levure, et (iii) le « sauvetage », qui correspond à la sortie de la molécule d'ADN de la levure pour la ré-introduire dans un environnement cellulaire adéquat à son expression, soit par transplantation de génome entier ou par transformation dans la cellule bactérienne receveuse (figure 13). Ces étapes sont détaillées dans la suite de cette partie.

### 2.4.1. Clonage dans la levure

Depuis longtemps, la levure *S. cerevisiae* est largement utilisée comme hôte pour le clonage et la propagation de molécules d'ADN sous forme de chromosomes artificiels de levure (YAC : Yeast Artificial Chromosome) ou de plasmides centromériques de levure (YCps). En 1987, un plasmide appelé « vecteur YAC » a été développé. Il contient un site de clonage pour l'ADN à capturer, une origine de répllication levure (ARS : Autonomously Replicative Sequence), un centromère (CEN), deux marqueurs de sélection de chaque côté du centromère (ici TRP1 et URA3) ainsi que deux séquences télomériques (TEL) (figure 14). Ce plasmide, une fois digéré par des enzymes de restriction, permet la capture de fragments d'ADN sous forme

Tableau 2 : Exemples de génomes clonés ou édités dans la levure d'après Vashee *et al.*, 2020.

Table 1						
Key examples of microbial genomes cloned or edited in yeast <sup>a</sup> .						
Organism	Size (Mbp)	% G + C	Genetic code	Cloning strategy <sup>b</sup>	Rescue strategy	References <sup>c</sup>
<b>Mollicutes</b>						
<i>Mycoplasma genitalium</i>	0.58	32	Nonstandard	Synthesis and assembly	N/A	[66]
<i>Mycoplasma mycoides</i> subsp. <i>capri</i>	1.1	24	Nonstandard	Cloning	Transplantation	[10]
<i>Mycoplasma pneumoniae</i>	0.81	41	Nonstandard	Cloning	N/A	[11,17]
JCVI Syn 1.0	1.1	24	Nonstandard	Synthesis and assembly	Transplantation	[1]
<i>Acholeplasma laidlawii</i>	1.5	32	Universal	Cloning	N/A	[51]
JCVI Syn 3.0	0.53	24	Nonstandard	Synthesis and assembly	Transplantation	[2]
<i>Mycoplasma mycoides</i> subsp. <i>mycoides</i>	1.2	24	Nonstandard	Cloning	Transplantation	[45]
<i>Mycoplasma capricolum</i> subsp. <i>capricolum</i>	1.1	25	Nonstandard	Cloning	Transplantation	[45]
<i>Mycoplasma leachii</i>	1	24	Nonstandard	Cloning	Transplantation	[45]
<i>Mycoplasma putrefaciens</i>	0.8	27	Nonstandard	Cloning	Transplantation	[45]
<i>Spiroplasma citri</i>	1.8	26	Nonstandard	Cloning	N/A	[45]
<i>Mycoplasma hominis</i>	0.66	27	Nonstandard	Cloning	N/A	[67]
<i>Mesoplasma florum</i>	0.79	27	Nonstandard	Cloning	Transplantation	[46]
<i>Mycoplasma capricolum</i> subsp. <i>capripneumoniae</i>	1	24	Nonstandard	Cloning	Transplantation	Pers. com 2020
<i>Mycoplasma ferriurinatoris</i>	1.2	24	Nonstandard	Cloning	Transplantation	Pers. com. 2019
<b>Proteobacteria</b>						
<i>Haemophilus influenzae</i>	1.8	38	Universal	Cloning	N/A	[12]
<i>Escherichia coli</i> (reduced genome)	1.03	51	Universal	Synthesis and assembly	N/A	[48]
<i>Escherichia coli</i> (recoded genome)	3.98	N/A	Universal	Synthesis and assembly	Partial replacement	[68]
<i>Salmonella typhimurium</i> (recoded genome)	4.47	N/A	Universal	Synthesis and assembly	Partial replacement	[47]
<i>Escherichia coli</i> (recoded genome)	3.98	N/A	Universal	Synthesis and assembly	Partial replacement	[3]
<i>Caulobacter ethensis 2.0</i> (reduced/recoded genome)	0.78	57	Universal	Synthesis and assembly	N/A	[4]
<b>Cyanobacteria</b>						
<i>Prochlorococcus marinus</i>	1.6	31	Universal	Cloning	N/A	[69]
<i>Synechococcus elongatus</i> (fragments)	2.7	55	Universal	Cloning	N/A	[70]
<b>Diatoms</b>						
<i>Phaeodactylum tricornutum</i> Chromosome 25	0.5	48	Universal	Cloning	N/A	[71]
<i>Phaeodactylum tricornutum</i> Chromosome 26	0.44	48	Universal	Cloning	N/A	[71]
<b>Viruses</b>						
Dengue virus type 2	0.011	46	Universal	Cloning	RNA transfection	[72]
MERS-CoV	0.029	41	Universal	Synthesis and assembly	DNA transfection	[73]
AcMNPV	0.14	45	Universal	Synthesis and assembly	DNA transfection	[74]
HCMV	0.23	49	Universal	TAR cloning and assembly	DNA transfection	[35]
Herpes simplex virus type 1	0.15	68	Universal	TAR cloning and assembly	DNA transfection	[36]
Horsepox virus	0.21	33	Universal	Synthesis and assembly <sup>c</sup>	DNA transfection	[42]
SARS-CoV-2	0.03	38	Universal	Synthesis and assembly	RNA transfection	[39]
MHV	0.032	42	Universal	Synthesis and assembly	RNA transfection	[39]
MERS-CoV	0.03	41	Universal	Synthesis and assembly	N/A	[39]
HCoV-229E	0.027	38	Universal	Synthesis and assembly	N/A	[39]
ZIKA virus	0.011	51	Universal	Synthesis and assembly	N/A	[39]
Human RSV-B	0.015	34	Universal	Synthesis and assembly	N/A	[39]

<sup>a</sup> Sorting is done by year of publication.

<sup>b</sup> Cloning refers to any method described in the text.

<sup>c</sup> The assembly was performed in mammalian cells.

linéaire dans un YAC (Burke *et al.*, 1987) (figure 14). Depuis, un grand nombre de molécules d'ADN provenant d'eucaryotes (Larionov *et al.*, 1996), de procaryotes (Kuspa *et al.*, 1989) et de virus (Ketner *et al.*, 1994) ont pu être clonées dans la levure (Vashee *et al.*, 2020) (tableau 2).

En 1997, Larionov utilise la méthode TAR (Transformation-Associated Recombinaison) (Larionov *et al.*, 1996) pour cloner le gène *brca2* dans la levure sous forme de YAC circulaire (Larionov *et al.*, 1997). Par la suite, plusieurs travaux ont décrit le clonage de génomes bactériens entiers sous forme circulaires de YCps comme par exemple la capture des génomes de *M. genitalium*, *Mycoplasma pneumoniae* et *Mycoplasma mycoides* subsp. *capri* (Benders *et al.*, 2010). On peut aussi citer le travail de Karas *et al.*, en 2013, au cours duquel le génome de *H. influenzae* (1.8 Mb), a été cloné sous forme circulaire dans la levure. Il convient de souligner que les plus grands génomes bactériens clonés dans la levure à ce jour sont ceux de *H. influenzae* et *Spiroplasma citri*, tous deux d'une taille de 1.8 Mb. (Karas *et al.*, 2013; Labroussaa *et al.*, 2016).

Plusieurs génomes synthétiques ont aussi pu être assemblés dans la levure grâce à sa capacité de recombinaison homologe. C'est notamment le cas du premier génome synthétique *M. genitalium* (Gibson *et al.*, 2008a), mais aussi celui de *M. mycoides* subsp. *capri* ensuite utilisé pour la construction de la première cellule synthétique JCVI Syn1.0 (Gibson *et al.*, 2010). Cette méthode d'assemblage de génomes synthétiques a aussi été utilisée pour la construction de la version minimale de la cellule synthétique, JCVI Syn3.0 (Hutchison *et al.*, 2016), ainsi que pour la construction des cellules synthétiques d'*E. coli* (Zhou *et al.*, 2016) et de *Caulobacter crescentus* (Venetz *et al.*, 2019).

Pour le clonage de molécules d'ADN dans la levure, plusieurs paramètres intrinsèques à la molécule d'ADN à cloner sont importants à prendre en compte :

- La taille du génome ou de la molécule d'ADN. Le plus grand chromosome de la levure a une taille de 1.531 Mb et, à l'heure actuelle, les plus grands génomes bactériens clonés chez la levure ont une taille de 1.8 Mb (Karas *et al.*, 2013). Nous ne savons pas à l'heure actuelle si la machinerie cellulaire de la levure est apte à propager (répliquer) une molécule de plus grande taille. Cependant, il est probable qu'il y ait une limite dans la taille d'une molécule qu'il est possible de cloner dans la levure.



- La présence d'ARS. Ce sont des origines de réplication chez la levure, elles ont une taille d'une centaine de nucléotides et sont composées de séquences riches en AT avec un motif consensus de 11 nucléotides (5'-WTTTATRTTT-3') appelé ACS (ARS Consensus Sequence) (Deshpande and Newlon, 1992). Elles sont considérées comme des séquences permettant de stabiliser une molécule d'ADN ou un génome bactérien dans la levure et ne doivent pas être séparées de plus de 300 kb (Dershowitz and Newlon, 1993). Des génomes ne contenant pas assez d'ARS ne peuvent pas être clonés dans la levure, par exemple, le clonage des génomes de *C. crescentus* et d'*E. coli* a nécessité l'ajout d'ARS supplémentaires (Venetz *et al.*, 2019; Zhou *et al.*, 2016).
- Le pourcentage en G+C. Les génomes avec un faible pourcentage en G+C sont plus faciles à cloner dans la levure. En effet, les ARS étant des séquences riches en AT, il est plus probable de trouver ces séquences dans des génomes avec un faible pourcentage en G+C. C'est pourquoi l'ajout d'ARS est nécessaire pour des génomes bactériens avec un pourcentage en G+C supérieur à 40%, mais n'est pas obligatoire pour des génomes avec un pourcentage en G+C inférieur à 40% (Vashee *et al.*, 2020).
- L'expression de gènes potentiellement toxiques pour la levure. Bien qu'un génome ou une molécule d'ADN bactérien (procaryote) ne soit pas censé être exprimé par la levure (eucaryote), il est tout de même possible que certains gènes potentiellement toxiques pour cette dernière soient exprimés. Le travail de Karas *et al.*, en 2013 rapporte ce phénomène. En effet, le clonage du génome d'*Acholeplasma laidlawii* dans la levure a nécessité l'élimination d'un gène codant une endonucléase (Karas *et al.*, 2012). Le clonage d'un génome n'utilisant pas le code génétique universel semble donc plus simple. C'est notamment le cas des mycoplasmes qui possèdent un code génétique alternatif pour lequel le codon UGA code un tryptophane, au lieu d'un codon stop dans le code génétique universel (Ohama *et al.*, 2008).

Enfin, le clonage et le maintien d'une molécule d'ADN dans la levure requièrent l'ajout de plusieurs séquences appelées « éléments levure ». (i) Un centromère (CEN) nécessaire à la ségrégation du génome bactérien ou de la molécule d'ADN lors de la division cellulaire, (ii) un marqueur de sélection (marqueur d'auxotrophie) et, si nécessaire, (iii) une origine de réplication levure (ARS) pour la réplication du génome bactérien ou de la molécule d'ADN

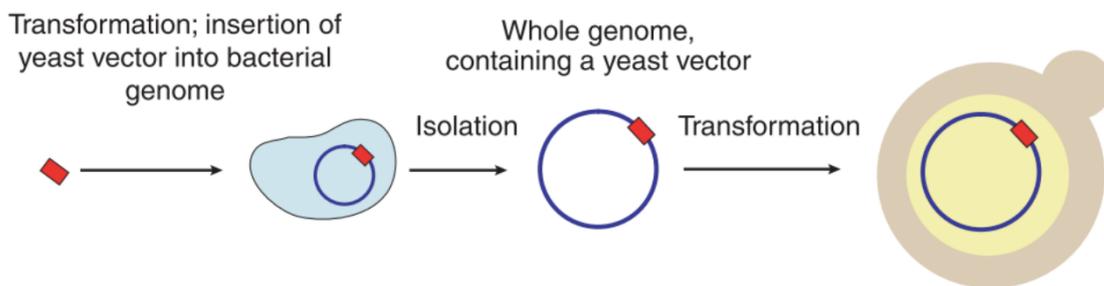


Figure 15 : Schéma de la méthode « classique » pour le clonage de génomes bactérien dans la levure d'après Benders *et al.*, 2010.

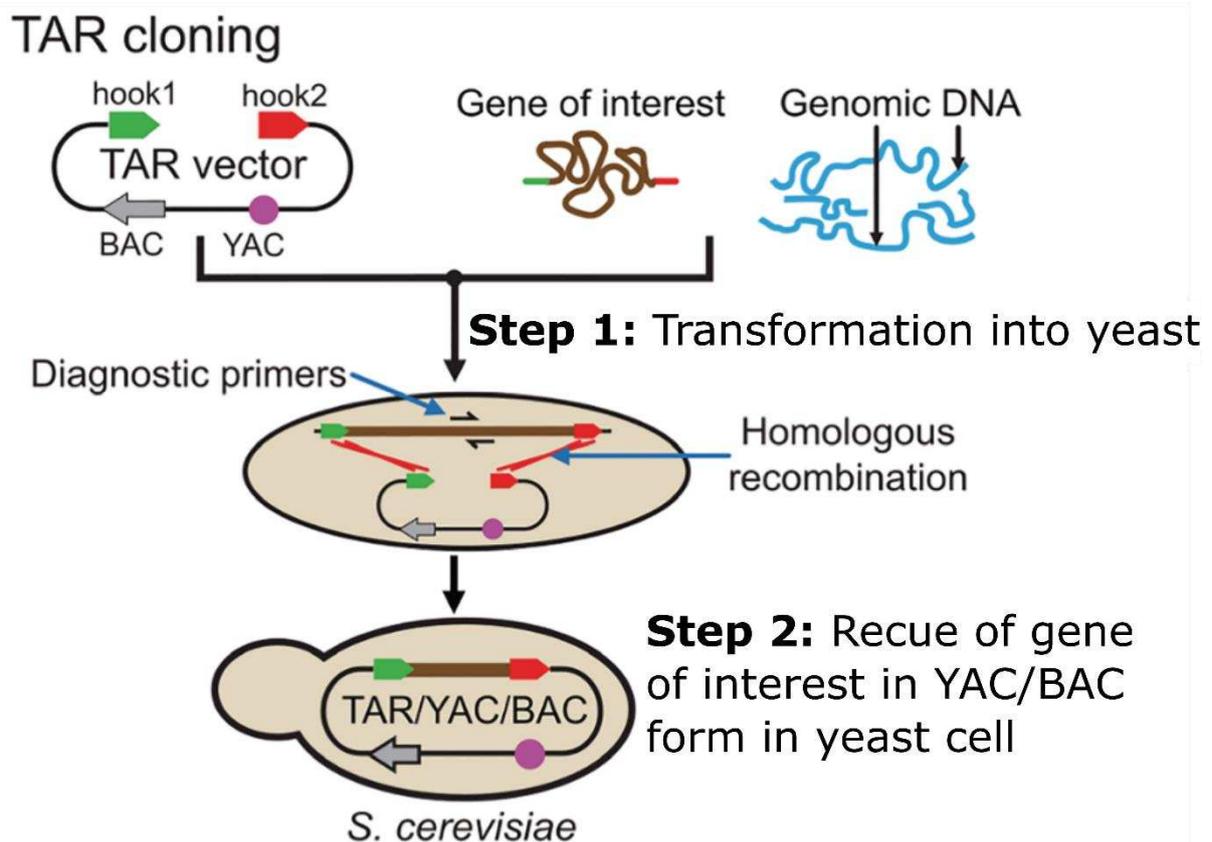


Figure 16 : Schéma de la méthode de TAR-cloning pour le clonage dans la levure d'un fragment de génome, adapté de Kouprina and Larionov, 2023. Cette méthode peut aussi être utilisée pour la capture de génome entier.

(génomés riches en GC). Des ARS supplémentaires peuvent être ajoutées afin de stabiliser la molécule d'ADN dans la levure (Dershowitz and Newlon, 1993).

Actuellement, plusieurs méthodes ont été décrites dans la littérature pour le clonage de fragments de génome ou de génomes entiers dans la levure. Elles sont décrites ci-après :

#### 2.4.1.1. Méthode « classique »

La méthode dite « classique » a été initialement développée pour le clonage des génomes de *M. genitalium*, *M. pneumoniae* et *M. mycoides* subsp. *capri* sous forme de YCps (Lartigue *et al.*, 2009, Benders *et al.*, 2010). Elle consiste à (i) ajouter les éléments levure au génome à capturer par transformation bactérienne, (ii) isoler le génome bactérien marqué avec les éléments levure en bloc d'agarose pour le protéger au maximum des cassures mécaniques, et enfin (iii) le transformer dans des sphéroplastes de levure, par une méthode au PEG décrite en 2008 (Kouprina and Larionov, 2008) (figure 15). Une des limites de cette méthode est qu'elle requiert l'ajout au préalable des éléments levure dans le génome bactérien à cloner. Cette étape peut être limitante pour certaines espèces de bactéries difficilement transformables.

#### 2.4.1.2. Méthode de TAR-cloning

La méthode de TAR-cloning permet de capturer des fragments de génome ou même des génomes entiers sous forme de YAC/BAC (Yeast Artificial Chromosome/Bacterial Artificial Chromosome) circulaires pouvant être propagés dans la levure et dans des cellules bactériennes (Kouprina and Larionov, 2023). Elle est principalement basée sur la capacité de recombinaison homologue de la levure. Elle consiste à linéariser la molécule d'ADN à capturer par digestion enzymatique puis à la co-transformer avec un vecteur TAR qui va permettre de la re-circulariser dans la levure (figure 16). Le vecteur TAR est composé des éléments levures, ainsi que d'une cassette BAC (avec une origine de réplication bactérienne et un marqueur de sélection), le tout bordé par deux séquences d'homologie avec les extrémités de la molécule d'ADN à capturer. Ces régions d'homologie doivent avoir une taille minimale de 60 pb (Noskov *et al.*, 2001). Dans le cadre du clonage de génomes bactérien, uniquement dans la levure, la



cassette BAC n'est pas obligatoire. En effet des génomes entiers peuvent être directement transplantés dans la cellule receveuse voulue, sans passer par *E. coli*. En revanche, la cassette BAC est obligatoire lorsqu'une étape chez *E. coli* est nécessaire, ce qui est souvent le cas lors de la capture de fragments de génomes bactériens. En fonction des caractéristiques intrinsèques de la molécule d'ADN à capturer, la présence d'une ARS dans le vecteur TAR est nécessaire (G+C% > 40%) ou non. Dans le cas où une ARS n'est pas présente dans le vecteur TAR, une ARS doit être présente dans la molécule d'ADN à capturer.

Cette méthode a été initialement développée pour capturer des fragments de génome, notamment humain (Kouprina *et al.*, 1997). Mais elle a aussi été utilisée pour capturer plusieurs génomes bactériens entiers comme ceux de *M. pneumoniae* (0.8 Mb) (Benders *et al.*, 2010), *A. laidlawii* (1.5 Mb) (Karas *et al.*, 2012), *Prochlorococcus marinus MED4* (1.6 Mb) (Tagwerker *et al.*, 2012), *Mycoplasma hominis* (Rideau *et al.*, 2017). Par ailleurs, elle peut être utilisée pour l'assemblage de multiples fragments naturels ou synthétiques en un génome complet, en les co-transformant avec un seul vecteur TAR (Gibson *et al.*, 2008b).

Pour augmenter l'efficacité de la méthode, des voies d'amélioration ont été proposées :

- L'utilisation du système CRISPR-Cas9 (à la place des enzymes de restriction) pour traiter l'ADN génomique et créer des cassures double brin de part et d'autre de la région cible, rendant ses extrémités propices à la recombinaison homologe chez la levure. Ce type de traitement a permis d'augmenter le nombre de clones positifs de 1% à 35% (Kouprina *et al.*, 2021).
- L'utilisation d'un système de contre sélection afin d'éviter un trop grand nombre de faux positifs. En effet, le vecteur TAR « classique » possède toutes les séquences permettant le maintien et la propagation d'une molécule d'ADN dans la levure, il peut ainsi, dans certain cas, se circulariser seul, entraînant l'apparition de faux positifs. Afin de limiter ce phénomène, un vecteur TAR avec un système de contre sélection a été développé. Ce système est basé sur le marqueur URA3 sous le contrôle du promoteur ADH1 de *Schizosaccharomyces pombe*, combiné avec l'utilisation de l'acide 5-fluoroorotique (5-FOA). Les régions d'homologie avec la molécule d'ADN à capturer sont placées entre le promoteur ADH1 et le gène *ura3*. Ainsi, lors de la capture de la séquence cible, le promoteur ADH1 et le gène *ura3* sont éloignés, ce qui empêche

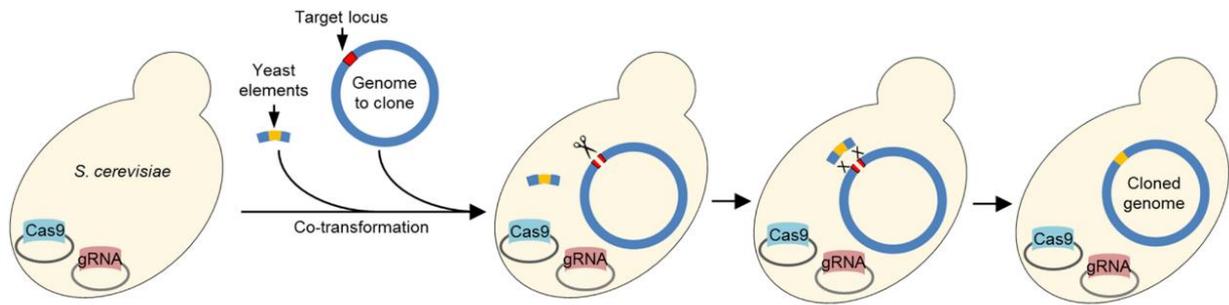


Figure 17 : Schéma de la méthode de CReasPy-Cloning, d'après Ruiz *et al.*, 2019.

l'expression du gène *ura3*, conférant la résistance au 5-FOA. Au contraire, si le vecteur se re-circularise seul, le promoteur ADH1 est réassocié au gène *ura3*, permettant son expression et donc la production d'enzymes qui convertissent le 5-FOA en 5-fluorouracil, toxique pour la levure (Noskov *et al.*, 2003).

#### 2.4.1.3. Méthode CReasPy-Cloning

La méthode de CReasPy-Cloning développée en 2019 combine la capacité de Cas9 à couper l'ADN à un locus spécifié par l'utilisateur et la recombinaison homologue très efficace de la levure pour cloner et modifier simultanément un chromosome bactérien dans la levure (Ruiz *et al.*, 2019). Cette méthode a été mise au point pour cloner et éditer simultanément des génomes de mycoplasmes. Elle consiste à pré-transformer la levure avec le système CRISPR-Cas9, c'est-à-dire un plasmide pour l'expression de la Cas9 et un plasmide pour l'expression d'un ARNg. Ensuite, les levures sont co-transformées avec le génome bactérien à capturer, préalablement isolé en bloc d'agarose, et un pansement de recombinaison. Ce pansement est composé des éléments levure et d'un marqueur de sélection, le tout bordé par deux régions d'homologie avec les extrémités de la molécule d'ADN à capturer. Une fois que le génome bactérien entre dans la levure, le système CRISPR-Cas9 va cibler spécifiquement le locus voulu et produire une cassure double brin ce qui aboutit à la linéarisation du génome bactérien. Le mécanisme de recombinaison homologue de la levure, stimulé par cette cassure double brin va la réparer en utilisant comme matrice le pansement de recombinaison co-transformé, générant ainsi une molécule d'ADN circulaire (figure 17).

L'avantage de cette méthode est qu'elle ne requiert pas l'insertion des éléments levure dans le génome bactérien en amont de la capture. Elle n'est donc pas limitée par une étape de transformation bactérienne. La méthode de CReasPy-Cloning a été utilisée pour la capture dans la levure des génomes de *M. pneumoniae* (0.8 Mb), *Mycoplasma leachii* (1 Mb), *Mycoplasma mycoides* subsp. *mycoides* (1.2 Mb), et *Mycoplasma feriruminatoris* (Ruiz *et al.*, 2019; Talenton *et al.*, 2022). Elle a aussi été utilisée pour supprimer simultanément trois *loci* du génome de *M. pneumoniae* au cours de son clonage (Ruiz *et al.*, 2019).

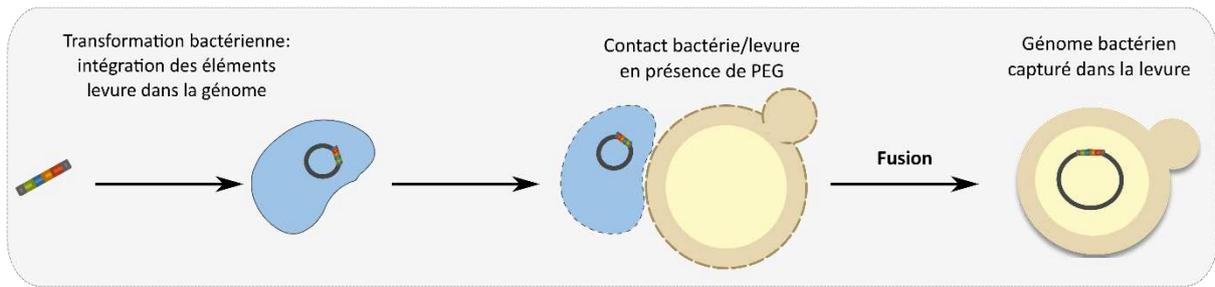


Figure 18 : Schéma de la méthode de Fusion pour le clonage de génomes bactériens dans la levure.

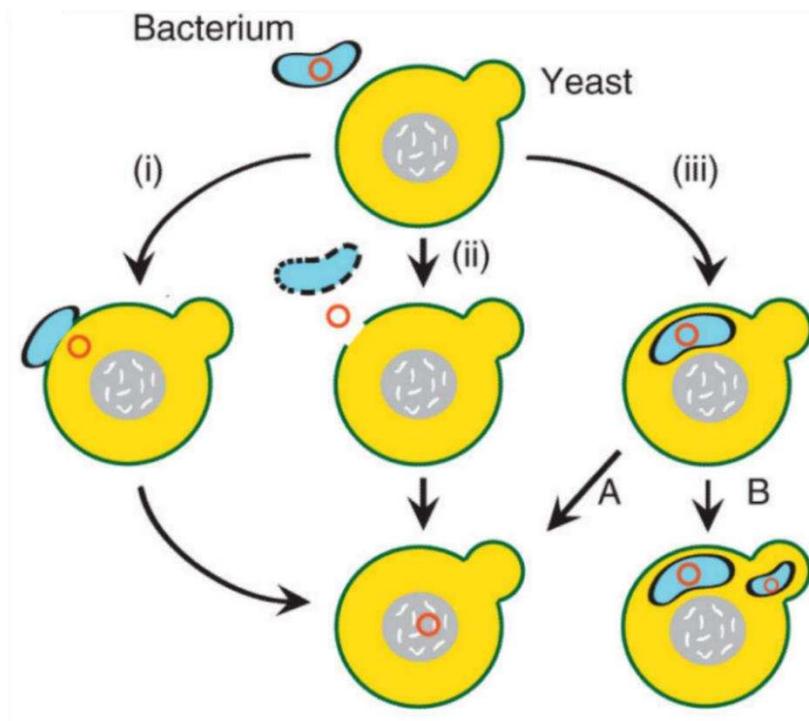


Figure 19 : Représentation schématique des trois scénarios possibles lors du clonage de génomes bactériens dans la levure par la méthode de « fusion » d'après Karas *et al.*, 2013. (i) Scénario de fusion membranaire, (ii) Scénario de lyse de la bactérie et transformation de la levure par le génome bactérien libre dans le milieu, (iii) Scénario d'endocytose de la bactérie par la levure. Bacterium = bactérie. Yeast = Levure

#### 2.4.1.4. Méthode de Fusion

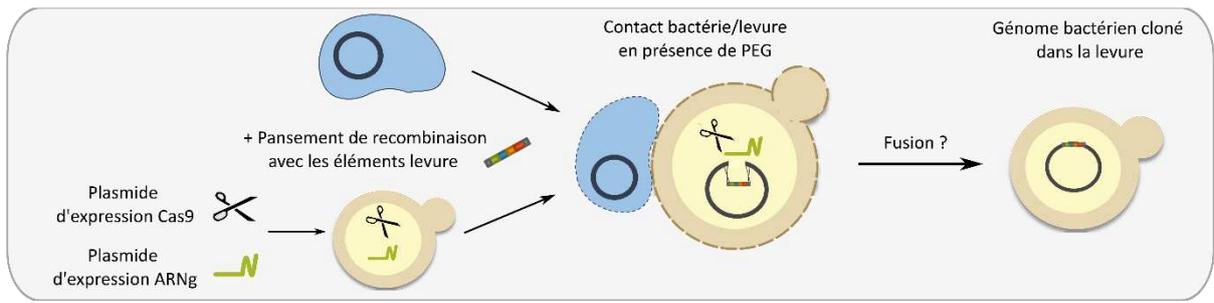
La méthode dite de « Fusion », développée en 2013 (Karas *et al.*, 2013), est basée sur un transfert d'ADN direct par contact cellules/cellules entre bactérie et levure. Cette méthode a été utilisée pour le clonage des génomes de *M. mycoides*, *Mycoplasma capricolum* et *H. influenzae*. Elle consiste à marquer le génome bactérien à capturer avec les éléments levure par transformation bactérienne, puis à mettre les cellules bactériennes sont directement mises en contact avec des sphéroplastes de levure avec un protocole au PEG (figure 18). Pour le cas de *H. influenzae*, une bactérie Gram -, la paroi bactérienne est digérée au préalable par du lysozyme pour produire des protoplastes.

Les auteurs décrivent que le transfert d'ADN peut se faire selon plusieurs scénarios : (i) par fusion membranaire et déversement du contenu de la bactérie dans la levure, (ii) par lyse de la bactérie et relargage de son génome dans le milieu puis transformation de la levure par le génome libre dans le milieu, ou (iii) par endocytose de la bactérie dans la levure (figure 19). Quelques années plus tard, des observations en microscopie électronique en transmission lors de l'étape de contact entre les bactéries et les levures ont permis d'observer ce qui semblait être le scénario d'endocytose (Karas *et al.*, 2019)

L'avantage de cette méthode est qu'elle ne nécessite pas d'isoler le génome bactérien, ce qui permet d'éviter les cassures du génome. Cependant, elle requiert d'ajouter au préalable des éléments levure dans le génome bactérien. Elle est donc dépendante de la capacité à transformer la bactérie avant le clonage de son génome dans la levure.

#### 2.4.1.5. Méthode de CReasPy-Fusion

La méthode de CReasPy-Fusion mise au point en 2023, combine les apports du CReasPy-Cloning et ceux de la méthode de clonage par « fusion » (Guesdon *et al.*, 2023). En effet, elle est basée sur l'utilisation du système CRISPR-Cas9 (comme le CReasPy-Cloning) et le transfert d'ADN par contact cellules/cellules (comme pour la « fusion »). Elle est divisée en trois étapes : (1) le système CRISPR-Cas9 est pré chargé dans la levure par la transformation d'un plasmide d'expression de la Cas9 et un plasmide d'expression d'ARNg spécifique au génome bactérien à cloner, (2) ces levures sont mises en contact avec des cellules bactériennes en présence de PEG. Au cours de cette étape de contact cellules/cellules, un pansement de recombinaison est



**Figure 20 : Schéma de clonage de génome bactérien dans la levure par CReasPy-Fusion.**

co-transformé dans la levure. Il contient les éléments levure ainsi qu'un marqueur d'auxotrophie, le tout bordé de deux régions d'homologie avec les extrémités du *locus* cible, (3) une fois que le génome bactérien est entré dans la levure, il est clivé en un point précis par le système CRISPR-Cas9 puis réparé par la machinerie de recombinaison homologue de la levure qui utilise le pansement de recombinaison co-transformé. Cette réparation aboutit à l'incorporation des éléments levure et à la re-circularisation du génome bactérien (figure 20). Comme pour la méthode de « fusion », plusieurs scénarios peuvent être responsables du transfert d'ADN entre la bactérie et la levure (figure 19).

Cette méthode a été utilisée pour le clonage des génomes de six espèces de mycoplasmes (*Mycoplasma capricolum* subsp. *capricolum*, *M. mycoides* subsp. *capri*, *M. mycoides* subsp. *mycoides*, *Mycoplasma capricolum* subsp. *capripneumoniae*, *Mycoplasma gallisepticum* et *Mycoplasma agalactiae*) ainsi que de grands fragments de génome. Elle permet donc la capture de génomes entiers ainsi que la capture de fragments de génome de façon spécifique en choisissant les sites de coupure avec le système CRISPR-Cas9. Cette méthode n'a pas encore été tentée pour capturer des génomes provenant de bactéries dotées d'une paroi cellulaire.

Les deux avantages de cette méthode sont qu'elle ne requiert ni l'isolement du génome bactérien, ni la pré-incorporation des éléments levure.

#### 2.4.2. Outils génétiques disponibles dans la levure

Une fois la molécule d'intérêt clonée dans la levure (fragment de génome ou génome entier), des outils génétiques puissants sont disponibles pour son ingénierie. Certains de ces outils ont été développés spécifiquement pour la modification de génomes bactériens clonés dans la levure, comme les méthodes TREC (Noskov *et al.*, 2010) et TREC-IN (Chandran *et al.*, 2014). Mais d'autres plus versatiles peuvent aussi être utilisés comme les outils basés sur le système d'édition CRISPR-Cas9 (Tsarmopoulos *et al.*, 2016).



### 2.4.2.1. CRISPR : du système de défense aux outils d'édition du génome

A l'origine, les systèmes CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) sont des systèmes de défense bactériens contre l'invasion par de l'ADN exogène. En 2012, le système naturel CRISPR-Cas9 a ensuite été détourné pour en faire un outil d'ingénierie de génomes permettant de couper l'ADN en un point précis (Jinek *et al.*, 2012).

Le locus CRISPR naturel est composé de séquences répétées de 21 à 48 pb entrecoupées par des séquences uniques appelées « spacers » de 26 à 72 pb, et des gènes *cas* (CRISPR-associated) nécessaires aux systèmes (Deveau *et al.*, 2010). Ce type d'organisation en séquences répétées, séparées par des séquences uniques a été observé pour la première fois chez *E. coli* en 1987. A cette époque, la fonction de ces séquences n'était pas encore comprise (Ishino *et al.*, 1987). Ce n'est qu'en 2002 que l'acronyme CRISPR est proposé pour qualifier ce type de locus (Jansen *et al.*, 2002). Par la suite, l'étude de ces systèmes a permis de les identifier comme étant des systèmes adaptatifs de défense contre de l'ADN exogène, utilisant un mode d'action en trois étapes (Barrangou and Horvath, 2017):

(1) L'adaptation par l'acquisition de nouveaux « spacers » après infection par des phages. Cette étape est médiée par des protéines Cas dont le rôle est spécifique à cet effet. Le nouveau « spacer » étant intégré au génome de la bactérie, la résistance au phage dont il provient est transmise aux cellules filles.

(2) L'expression du système, consiste à exprimer les protéines Cas et à maturer le pre-crRNA qui correspond à un long transcrit primaire produit à partir du *locus* CRISPR de la bactérie, en un crRNA (ARN CRISPR) mature. Ce dernier s'associe à des effecteurs pour former une particule ribonucléique active.

(3) L'interférence, correspond au clivage de l'ADN exogène. La particule ribonucléique est guidée à sa cible présente sur l'ADN exogène par le crRNA. La cible appelée « protospacer » est reconnue par complémentarité de base avec le crRNA. Cette reconnaissance aboutit à la formation d'une R-loop puis au clivage de l'ADN en un point spécifique. Il est important de noter que la reconnaissance de l'ADN exogène se fait par la complémentarité de base mais également par l'identification d'un motif nucléotidique adjacent aux « protospacers » appelée PAM (Protospacer Adjacent Motif) (Sapranauskas *et al.*, 2011). Ce motif est présent sur les



cibles (ADN exogène) mais absent du *locus* CRISPR de la bactérie, permettant au système de ne pas cibler son propre génome, ce qui lui serait létal.

De nos jours, l'ensemble des systèmes CRISPR sont répartis en 2 classes, 6 types et 33 sous types (Makarova *et al.*, 2020). Les systèmes CRISPR de classe 1 incluent les types I, III et IV et 16 sous-types (Makarova *et al.*, 2020). Ils sont définis par des effecteurs multi-protéiques. La mise en place du système nécessite une réaction dite « en cascade » dans laquelle plusieurs protéines Cas entrent en jeu pour la maturation du crRNA, la reconnaissance de la cible, et enfin, l'étape d'interférence (Bernheim, 2017).

Les systèmes de classe 2 regroupent les types II, V et VI et 17 sous types (Makarova *et al.*, 2020). Ils sont caractérisés par un effecteur constitué d'une seule protéine multi-domaine, capable de réaliser l'ensemble de l'activité d'interférence. Les outils CRISPR-Cas dérivent tous de systèmes de cette classe (Bernheim, 2017). Le type II dont est issu l'outil CRISPR-Cas9 est défini par la protéine Cas9 capable de cliver l'ADN exogène ainsi que la présence d'un tracrRNA (trans-activating crRNA) qui se fixe au crRNA (Jinek *et al.*, 2012). La maturation du crRNA est ensuite effectuée par la RNase III. C'est donc tout le complexe tracrRNA/crRNA mature plus communément appelé ARNg, qui sert de guide pour l'effecteur Cas9.

Le système CRISPR-Cas9 a été le premier à être adapté en outil d'ingénierie de génome pour le clivage double brin de l'ADN en un point précis (Jinek *et al.*, 2012). Cette cassure double brin stimule alors les systèmes de réparation de l'ADN par NHEJ (Non-homologous end joining) ou par recombinaison homologue, permettant ainsi l'édition de génome. Emmanuelle Charpentier et Jennifer Doudna ont été récompensées d'un prix Nobel de Chimie en 2020 suite à leurs travaux sur cette technologie. Ce système a ensuite été utilisé chez de nombreux organismes comme des bactéries (Jiang *et al.*, 2013) et des levures (DiCarlo *et al.*, 2013) mais aussi chez des eucaryotes complexes comme par exemple la souris (Shen *et al.*, 2013) ou le lapin (Li *et al.*, 2021). Sa versatilité et son efficacité ont fait sa renommée et ont encouragé la communauté scientifique à l'utiliser.

L'efficacité du système CRISPR-Cas9 en tant qu'outil pour l'ingénierie de génome dépend de la présence de systèmes de réparation de l'ADN chez l'hôte dans lequel il est utilisé. Or, chez de nombreuses bactéries, les systèmes de réparation de l'ADN sont trop peu efficaces ou même absents, empêchant son utilisation. Ce problème a pu être contourné en utilisant la



levure *S. cerevisiae*, connue pour posséder un système de réparation de l'ADN extrêmement efficace, comme plateforme d'ingénierie de génomes bactériens.

Après avoir démontré la puissance de la technologie CRISPR-Cas9 chez la levure (DiCarlo *et al.*, 2013), l'outil a été testé pour modifier des génomes bactériens dans la levure en 2016 et s'est révélé très efficace (Tsarmpopoulos *et al.*, 2016). Au cours de ces travaux, le gène *glpO* a été supprimé du génome de *M. mycoides* subsp. *capri*. Ces expériences ont démontré que le système CRISPR-Cas9 pouvait être utilisé pour éditer un génome bactérien cloné au préalable dans la levure et il est depuis très régulièrement utilisé à cette fin (Nottelet *et al.*, 2021). De plus, il peut aussi être utilisé au cours de l'étape de clonage/modification d'une molécule d'ADN dans la levure comme décrit dans la méthode de CReasPy-cloning (Ruiz *et al.*, 2019) (cf partie 2.4.1.3 de l'introduction) ou la méthode de CReasPy-Fusion (Guesdon *et al.*, 2023) (cf partie 2.4.1.5 de l'introduction).

### 2.4.3. Méthodes pour « sortir » de la levure

La dernière étape du cycle utilisant la levure comme plateforme d'ingénierie génomique consiste à réintroduire la molécule d'ADN, clonée et modifiée dans la levure, dans un compartiment qui lui permettra d'être exprimée. Pour les molécules d'ADN issues de bactéries, cette étape peut être complexe et est souvent considérée comme l'étape limitante du cycle d'ingénierie. Cependant, de nos jours, des méthodes ont été mises au point pour le transfert de génomes entiers ou de fragments de génomes en fonction du type de molécule d'ADN clonée dans la levure.

#### 2.4.3.1. Transplantation de génome

La transplantation de génome (TG) est une méthode qui permet le transfert d'un génome bactérien complet dit « génome donneur » dans une cellule dite « receveuse ». Une fois transféré dans la cellule receveuse, le génome donneur lui confère un nouveau génotype et un nouveau phénotype qu'il est possible de sélectionner afin d'isoler les cellules génétiquement modifiées contrôlées par le nouveau génome. Ces cellules sont appelées « transplants ».

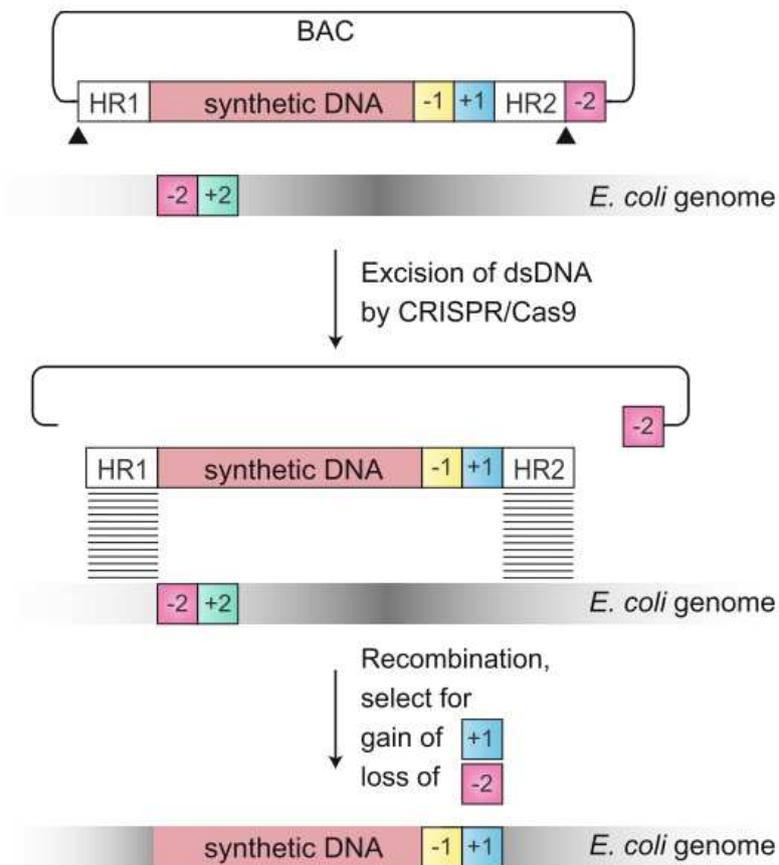


Cette méthode a été mise au point en plusieurs étapes en utilisant des génomes et des cellules receveuses de mycoplasmes. Tout d'abord, en 2007, les premiers essais de TG ont été menés en transférant le génome de *M. mycoides* subsp. *capri* (souche GM12) directement dans une cellule receveuse de *M. capricolum* subsp. *capricolum* sans passer par la levure (Lartigue *et al.*, 2007). Ce travail a démontré pour la première fois qu'il était possible de transplanter un génome bactérien donneur dans une cellule receveuse pour obtenir des transplants. Par la suite, en 2009, le génome de *M. mycoides* subsp. *capri* cloné puis modifié dans la levure a été transplanté dans la cellule receveuse de *M. capricolum* subsp. *capricolum*, montrant qu'il est possible de transférer un génome bactérien complet depuis la levure vers une bactérie receveuse (Lartigue *et al.*, 2009).

Cette méthode de TG a ensuite été utilisée pour la construction de la première cellule contrôlée par un génome synthétique, JCVI Syn1.0. En effet, après avoir assemblé dans la levure une version synthétique du génome de *M. mycoides* subsp. *capri*, les auteurs l'ont ensuite transplanté dans une cellule receveuse de *M. capricolum* subsp. *capricolum* produisant des transplants contrôlés par le génome synthétique (Gibson *et al.*, 2010). En 2016, c'est aussi cette stratégie qui a permis d'obtenir la cellule synthétique minimale, JCVI Syn3.0 (Hutchison *et al.*, 2016).

Lors de la mise au point de la TG, quelques facteurs ont été identifiés comme étant limitants ou ayant un impact fort :

- **La présence des systèmes de restriction modification de la cellule receveuse.** En effet, lors de la mise au point de la méthode pour le transfert de génome depuis la levure, les systèmes de restriction modification de la cellule receveuse ont dû être inactivés afin de protéger le génome donneur de l'action d'enzymes de restriction (Lartigue *et al.*, 2009). Les auteurs ont montré qu'il était également possible de méthyler le génome donneur avec des extraits cellulaires bruts en amont de la réaction afin qu'il ne soit pas reconnu par les systèmes de restriction modification de la cellule receveuse (Lartigue *et al.*, 2009).
- **La distance phylogénétique entre le génome donneur et la cellule receveuse.** Plus le génome donneur est proche phylogénétiquement de la cellule receveuse plus la transplantation est efficace. En effet, l'expression du génome donneur dans la cellule receveuse est un des points clés de la méthode. Dans un premier temps, la cellule



**Figure 21 : Echange de fragments de génome par REXER, d'après Fredens *et al.*, 2019.** Les triangles noirs représentent les sites de coupure utilisés par le systèmes CRISPR-Cas9 pour isoler le fragment d'ADN synthétique du BAC. Cette cassette est bordée de deux régions d'homologie (HR1 et HR2) avec l'ADN génomique à échanger. L'échange de fragment est réalisé par le système de recombinaison lambda-red. Cet évènement est sélectionné par l'intégration de la cassette de double sélection (-1 et +1) portée par le fragment synthétique et par la perte de la cassette de double sélection (-2 et +2) portée par la région génomique échangée.

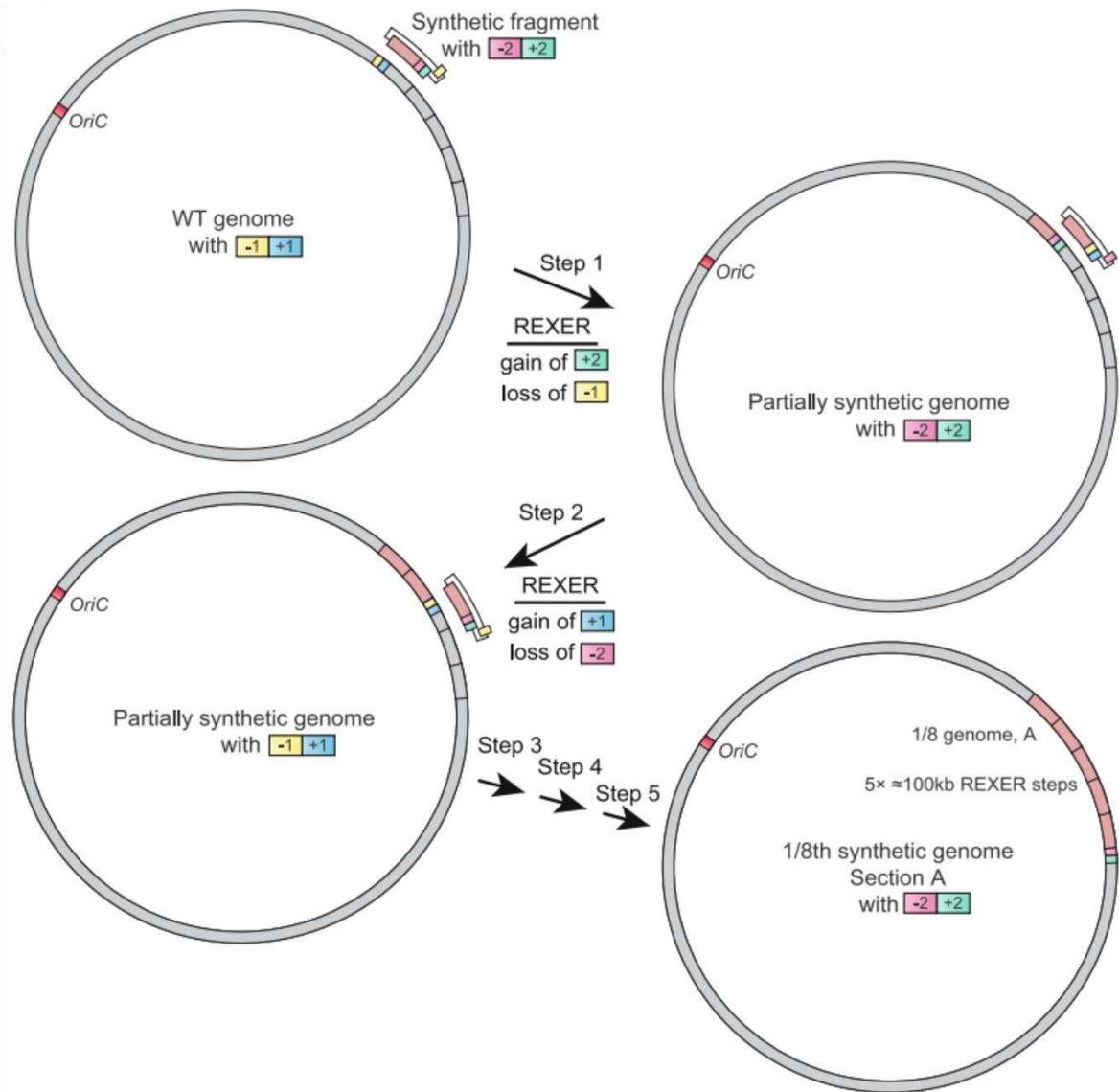
receveuse doit être apte à initier l'expression du génome donneur avec sa propre machinerie cellulaire. Ce processus dépend donc de la proximité phylogénétique reliant l'espèce receveuse à l'espèce donneuse du génome (Labroussaa *et al.*, 2016).

Bien que très puissante, pour le moment, la méthode de TG n'a été mise au point et utilisée que pour des espèces de mycoplasmes (Labroussaa *et al.*, 2019; Gourgues *et al.*, 2024; Talenton *et al.*, 2022). Le point critique de cette méthode repose dans le fait que la cellule receveuse doit être capable d'exprimer le génome donneur dans son entièreté (Labroussaa *et al.*, 2019). Cependant, les paramètres permettant l'expression du génome donneur ainsi que l'élimination du génome natif de la cellule receveuse ne sont pas encore bien compris. Une compréhension fine des mécanismes moléculaires menant à l'obtention de transplants pourrait ouvrir la voie à l'adaptation de la méthode chez d'autres espèces bactériennes.

#### 2.4.3.2. Echange de fragment de génomes

Chez les bactéries pour lesquelles la transplantation de génome n'est pas encore disponible, des méthodes alternatives ont été développées. Elles sont toutes basées sur le transfert et l'échange de fragments de génome. Dans la suite de cette partie, deux d'entre elles seront décrites (REXER/GENESIS et RAGE).

La méthode REXER (Replicon Excision Enhanced Recombination), basée sur l'utilisation du système CRISPR-Cas9 et du système de recombinaison lambda-red, permet chez *E. coli* le remplacement d'un locus d'ADN génomique par une version synthétique portée par un vecteur de type BAC. Cette méthode est découpée en deux étapes : (1) le BAC contenant le fragment synthétique (donneur) d'intérêt est transformé dans une cellule bactérienne receveuse. Le système CRISPR-Cas9 porté par le BAC est utilisé pour isoler le fragment d'ADN synthétique bordé par deux régions d'homologie avec les extrémités du locus génomique à remplacer, (2) l'échange de fragments de génome est effectué via le système de recombinaison lambda-red. La sélection de l'échange est réalisée à la fois avec une sélection positive par l'intégration d'un marqueur porté par le fragment donneur et avec une sélection négative par la perte d'un marqueur initialement porté par le locus génomique échangé (Wang *et al.*, 2016) (figure 21).



**Figure 22 : Schéma illustratif de GENESIS pour le remplacement itératif d'ADN génomique par de l'ADN synthétique, d'après Fredens *et al.*, 2019.** Les cycles itératifs de REXER avec une alternance entre les marqueurs de sélection et de contre sélection ont abouti à la mise en œuvre de la méthode GENESIS. Le premier fragment synthétique inséré par REXER porte la cassette de double sélection (-1 et +1) qui sert de marqueur de contre sélection lors de l'intégration du second fragment synthétique qui porte la cassette (-2 et +2). L'échange de ce second fragment est sélectionné par l'insertion de la cassette (-2 et +2) et la perte de la cassette (-1 et +1).

Des cycles itératifs de REXER, alternant les marqueurs de sélection et contre sélection, permettent le remplacement progressif de fragments de génome adjacents, pouvant même aboutir au remplacement de l'intégralité du génome. Cette stratégie appelée GENESIS (Genome Steppwise Interchange Synthesis) (figure 22) a notamment été utilisée pour la construction d'une souche d'*E. coli* recodée, qui n'utilise que 61 codons au lieu des 64 codons standards. Lors de ce travail, des fragments d'ADN synthétiques de 100 kb ont été assemblés dans la levure puis intégrés de façon itérative dans le génome d'une souche receveuse d'*E. coli* par l'intermédiaire de la méthode REXER (Fredens *et al.*, 2019).

La méthode RAGE (Recombinase-Assisted Genome Engineering) a été développée en 2020 (Garcia-Morales *et al.*, 2020). Pour l'étape d'échange de fragments de génome dans la cellule receveuse, cette méthode se base sur le système de recombinaison site-spécifique cre/lox. Ce système de recombinaison sera détaillé ultérieurement dans cette étude bibliographique (cf partie 3.2.2 de l'introduction).

Cette stratégie se décline en 4 étapes : (1) un fragment de génome d'intérêt est capturé dans la levure, dans un vecteur navette de type YAC/BAC (cf partie 2.4.1.2 de l'introduction). Ce vecteur permet la propagation du fragment de génome bactérien à la fois chez la levure et chez *E. coli*. De plus, il est composé de deux sites lox incompatibles entre eux (le site lox66 et le site loxm2/71), ainsi que du gène *cre*, nécessaire au fonctionnement de système cre/lox. Il est construit de sorte à ce que le fragment de génome bactérien se retrouve capturé entre les deux sites lox, (2) la molécule d'ADN ainsi capturée peut être modifiée à souhait grâce aux outils génétiques disponibles chez la levure (cf partie 2.4.2 de l'introduction), (3) le fragment de génome (modifié ou non) est ensuite transféré chez *E. coli* pour une étape d'amplification de son nombre de copies. Cette étape chez *E. coli* permet ainsi de purifier facilement une grande quantité du fragment d'ADN d'intérêt, (4) ce dernier, qualifié à cette étape de molécule donneuse, est enfin transformé dans une cellule receveuse dans laquelle deux « landing pads » ont préalablement été insérés de part et d'autre du locus génomique à échanger. Ces « landing pads » correspondent à deux sites lox (le site lox71 et le site loxm2/66) incompatibles entre eux mais compatibles avec ceux portés par la molécule donneuse.

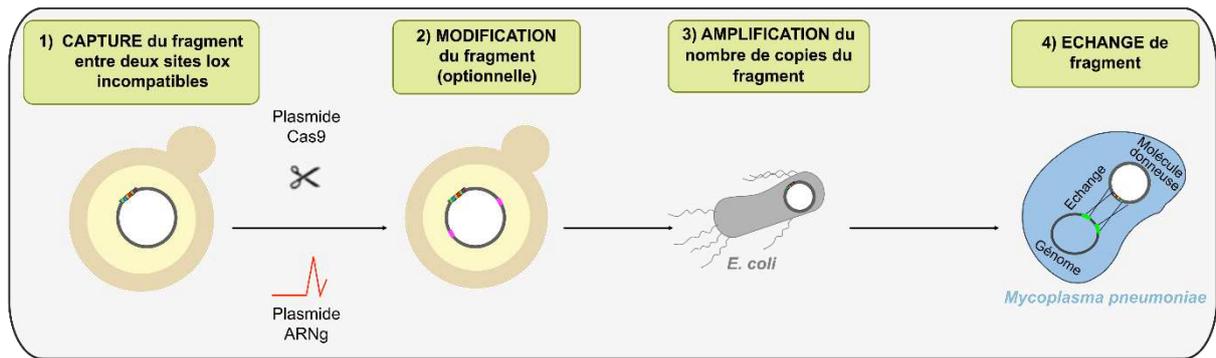


Figure 23 : Schéma de la méthode RAGE développée chez *M. pneumoniae*.

L'échange de fragments de génome est ainsi réalisé par le système cre/lox, aboutissant à une cellule modifiée (figure 23).

Cette stratégie a été développée chez *M. pneumoniae* et a permis de remplacer un locus de 38 kb par une version modifiée ou une version synthétique réduite, générant une souche de *M. pneumoniae* dans laquelle 13 gènes ont été éliminés, soit une délétion de 15.4 kb en une seule fois (Garcia-Morales *et al.*, 2020).

Contrairement à la TG, pour laquelle le mécanisme de remplacement du génome natif par le génome donneur n'est pas bien compris, ce type de stratégie d'échange de fragments de génome repose sur des mécanismes de recombinaison bien connus et bien caractérisés. Par rapport à la TG, ce type d'approche pourrait donc être plus facilement adapté à d'autres espèces de bactéries.

#### 4. Les systèmes de recombinaison génétique

La recombinaison génétique correspond à un échange d'information génétique entre deux segments d'ADN. Ce phénomène peut se produire entre deux molécules d'ADN différentes ou entre deux portions d'une même molécule. C'est un processus fondamental dans l'évolution, la diversité génétique et le maintien de l'intégrité des génomes. La recombinaison génétique est maintenant aussi utilisée en tant qu'outil en biologie moléculaire et ingénierie génétique pour engendrer des modifications ciblées des génomes et pour construire de nouvelles molécules d'ADN. Les différents systèmes de recombinaison génétique sont classés en 4 groupes (Rossignol, 1990) :

- 1) La recombinaison homologue, qui correspond à un événement de recombinaison entre deux régions possédant une similitude de séquence importante. Elle peut être effectuée entre deux molécules d'ADN indépendantes ou sur la même molécule. Enfin, elle peut se produire n'importe où sur le génome tant que des régions homologues ou similaires sont présentes.
- 2) La recombinaison site-spécifique, qui nécessite la présence de séquences nucléotidiques spécifiques, reconnues par une enzyme (recombinase) entraînant la recombinaison. Elle correspond à un événement de recombinaison programmé entre deux sites à des positions précises sur les molécules d'ADN.



- 3) La transposition, qui requiert des séquences spécifiques capables de se déplacer d'une position du génome à une autre. Ces éléments sont appelés transposons et contiennent toutes les séquences nécessaires à leur déplacement.
- 4) La recombinaison illégitime, qui correspond à un évènement de recombinaison entre des séquences d'ADN ne présentant pas de similitude. Elle regroupe tous les types d'évènements de recombinaison qui n'entrent pas dans les 3 premiers groupes.

Dans la partie qui suit ne sont présentés que les systèmes de recombinaison homologue et de recombinaison site-spécifique. En effet, mon projet de thèse ne faisant appel ni à la transposition ni à la recombinaison illégitime, nous avons choisi de ne pas présenter ces thématiques plus en détail.

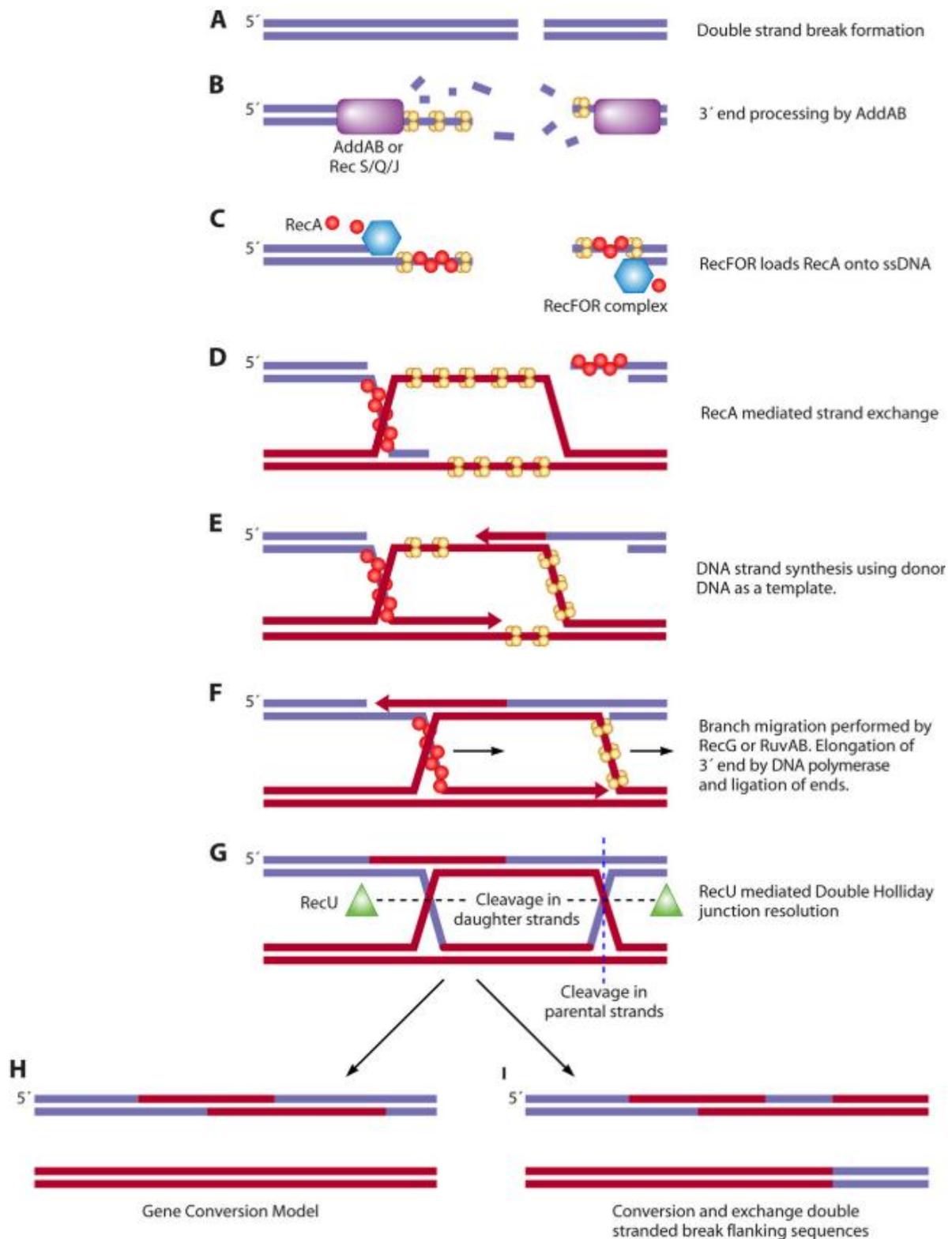
## 4.1. La recombinaison homologue

### 4.1.1. Description générale

La recombinaison homologue correspond à un échange d'information génétique entre deux molécules d'ADN possédant un certain degré d'identité de séquence. C'est un processus qui met en jeu un grand nombre de protéines, dont des recombinases (comme par exemple RecA chez *E. coli* et *B. subtilis* ou Rad51 chez *S. cerevisiae*), une activité de dégradation et de synthèse d'ADN, et qui requiert une source d'énergie, l'ATP (Lenhart *et al.*, 2012).

La recombinaison homologue joue un rôle important dans la diversité des génomes et le maintien de leur intégrité. En effet, c'est un processus qui permet l'acquisition de nouveaux traits suite à l'intégration de fragments d'ADN. Il permet aussi de réparer l'ADN de lésions délétères comme les cassures double brin. De plus, c'est un système central pour les processus de réplication et de maintien des télomères (Li and Heyer, 2008).

Enfin, en biologie moléculaire, la recombinaison homologue est utilisée comme un outil d'ingénierie permettant notamment de réaliser de la mutagenèse dirigée.



**Figure 24 : Représentation schématique de la réparation d'un cassure double brin par le système de recombinaison chez *B subtilis* d'après Lenhart *et al.*, 2012. Le brin cassé est représenté en bleu et le brin donneur intact est représenté en rouge.**

#### 4.1.2. La recombinaison homologue chez *B. subtilis*

*B. subtilis* possède une machinerie de recombinaison homologue qui intervient à la fois dans la réparation de l'ADN et dans l'insertion de fragments d'ADN dans le génome après leur entrée dans la cellule, notamment par compétence naturelle (cf partie 1.4 de l'introduction).

**Réparation de l'ADN.** Chez *B. subtilis*, la réparation de l'ADN par recombinaison homologue nécessite 5 étapes : (i) la reconnaissance d'une extrémité double brin, (ii) le recrutement de la recombinase (RecA chez *B. subtilis*) et la formation d'un nucléofilament, (iii) l'envahissement de brin d'un segment d'ADN similaire intact qui aboutit à la formation d'une boucle D (D loop), (iv) la synthèse d'ADN à partir de l'extrémité 3' OH libre du brin envahisseur en utilisant le brin envahi comme matrice, et enfin (v) la résolution de la jonction de Holliday par coupure enzymatique (Lenhart *et al.*, 2012) (figure 24).

De façon plus détaillée, après une cassure double brin (CDB) (figure 24A), le complexe AddAB (analogue du complexe RecBCD de *E. coli*) se fixe sur l'extrémité double brin (figure 24B). Ce complexe, qui possède des activités hélicase et nucléase, dégrade les deux brins jusqu'à rencontrer une séquence Chi (Crossover Hot spot Instigator) (5'- AGCGG - 3') (Lenhart *et al.*, 2012). Lorsqu'il rencontre cette séquence, son activité nucléasique est modifiée pour ne dégrader que le brin 5'. L'arrêt de la dégradation du brin 3' entraîne l'apparition d'une extrémité 3' libre propice au recrutement de RecA (figure 24C). Les protéines SsbA et SsbB se fixent sur l'ADNsb de l'extrémité 3' libre afin de le protéger de la dégradation, puis les protéines RecA sont recrutées. Mais étant incapables d'interagir directement avec les complexes SsbA/ADNsb et SsbB/ADNsb, elles se fixent sur l'extrémité 3' libre par l'intermédiaire des protéines DprA et RecO qui agissent comme médiateurs. A cette étape, d'autres protéines peuvent également intervenir, comme la protéine RecF qui agit en tant que modulateur positif en améliorant la fixation de RecA, et les protéines RecX et RecU qui agissent comme modulateurs négatifs en inhibant la fixation de RecA sur l'ADNsb (Serrano *et al.*, 2018). Une fois recrutées sur l'ADNsb, les protéines RecA « scannent » le génome à la recherche d'une séquence d'homologie intacte et catalysent un échange de brins (ou envahissement de brins). Cet événement aboutit à la formation d'un hétéroduplex dans lequel le brin envahisseur et le brin complémentaire sont mis de pair, formant une D loop (figure 24D). A ce stade démarre une synthèse d'ADN à partir de l'extrémité 3' libre en utilisant le brin dit



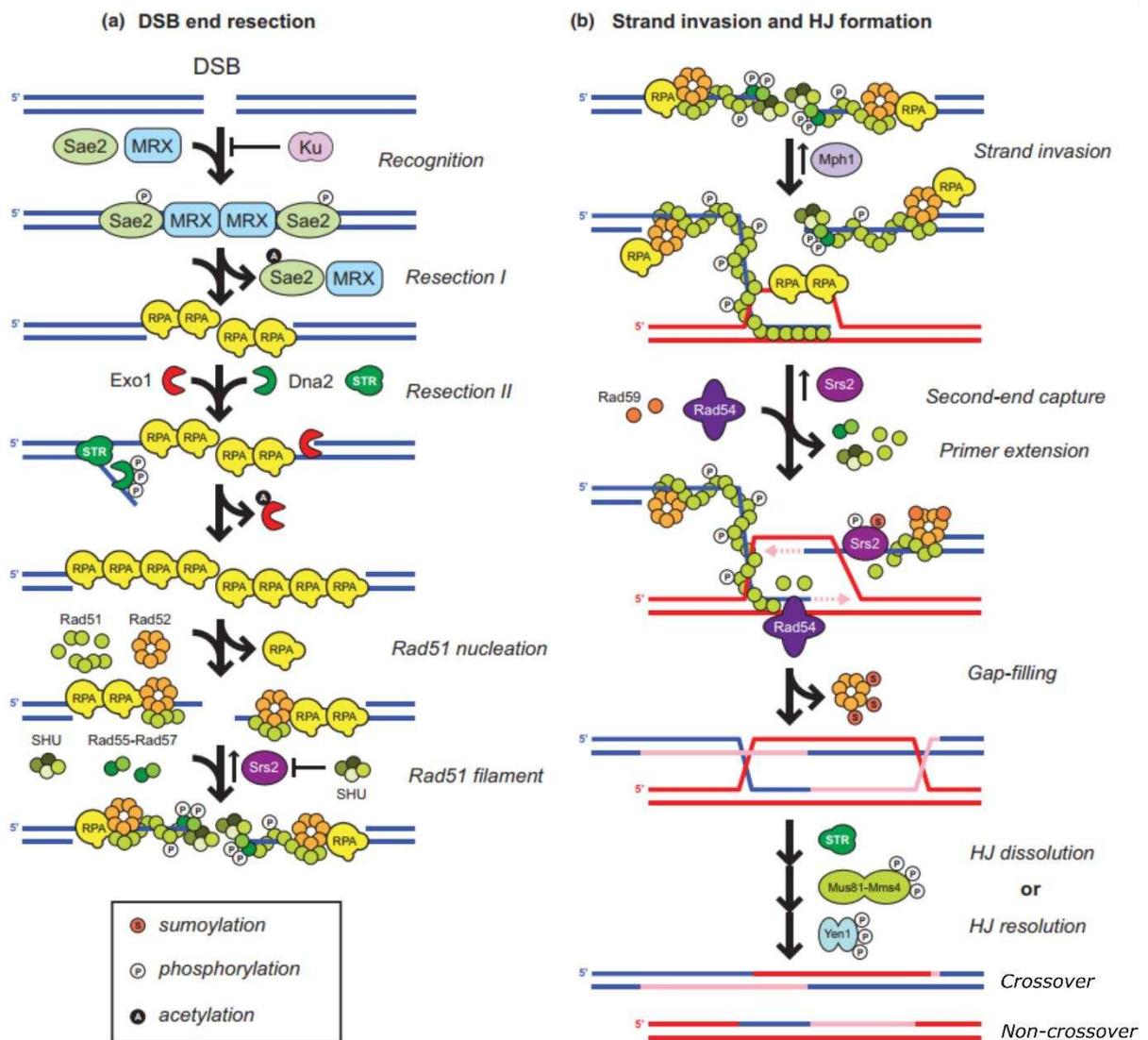
donneur (brin envahi) comme matrice (figure 24E). En parallèle, la D loop est étendue par l'action de différentes translocases : RecD2 (Serrano *et al.*, 2020), RuvAB, RecG et RadA/Sms (Torres *et al.*, 2019) (figure 24F). Concernant le complexe RuvAB, la protéine RuvA se fixe la première sur la D loop puis l'hélicase RuvB qui permet l'expansion est recrutée. Concernant RadA, il est intéressant de noter que son absence réduit de 140 fois l'efficacité de la transformation chromosomique mais pas de la transformation plasmidique (Torres *et al.*, 2019). Après migration de la D loop et synthèse d'ADN en utilisant le brin envahi, la jonction de Holliday (formée des 4 brins : 2 brins envahisseurs et 2 brins envahis) est résolue par une nucléase (figure 24G) dont l'identité n'est pas claire encore aujourd'hui (Serrano *et al.*, 2021). Plusieurs scénarios de résolution sont possibles en fonction de la coupure de la jonction de Holliday (figure 24H). Ce système permet finalement de réparer la CDB sans erreur.

**Intégration de fragments d'ADN dans le génome.** Chez *B. subtilis*, les fragments d'ADN qui entrent dans la cellule par compétence naturelle sont sous forme simple brin (cf partie 1.4 de l'introduction). Ils sont donc directement pris en charge par les protéines SsbA et SsbB qui les protègent de la dégradation par des exonucléases, et le complexe AddAB n'intervient pas. Les étapes qui suivent sont identiques à la section précédente excepté pour l'étape de la résolution de la D loop qui aboutit forcément à l'intégration du fragment d'ADN transformé dans le génome. Des travaux menés en 2000 ont montré qu'au cours du développement de la compétence, la protéine RecA était fortement induite (Fernández *et al.*, 2000).

La faculté de *B. subtilis* à faire de la recombinaison homologe a été mise à profit pour des expériences de mutagénèse dirigée, par exemple lors d'expériences d'insertion ou de délétion de gènes à des *loci* précis ou des substitutions ciblées. C'est donc devenu un outil puissant pour modifier le génome de *B. subtilis*.

#### 4.1.3. La recombinaison homologe chez la levure *S. cerevisiae*

Chez la levure *S. cerevisiae*, la recombinaison homologe est un système qui contribue notamment au maintien de l'intégrité du génome et qui agit comme système de réparation de l'ADN lors des CDB. Pour la réparation de ce type de lésion, deux systèmes sont en compétition, la NHEJ et la recombinaison homologe. Le choix de l'un ou l'autre de ces systèmes est dicté par la phase du cycle cellulaire. En phase G1, la NHEJ est préférentiellement

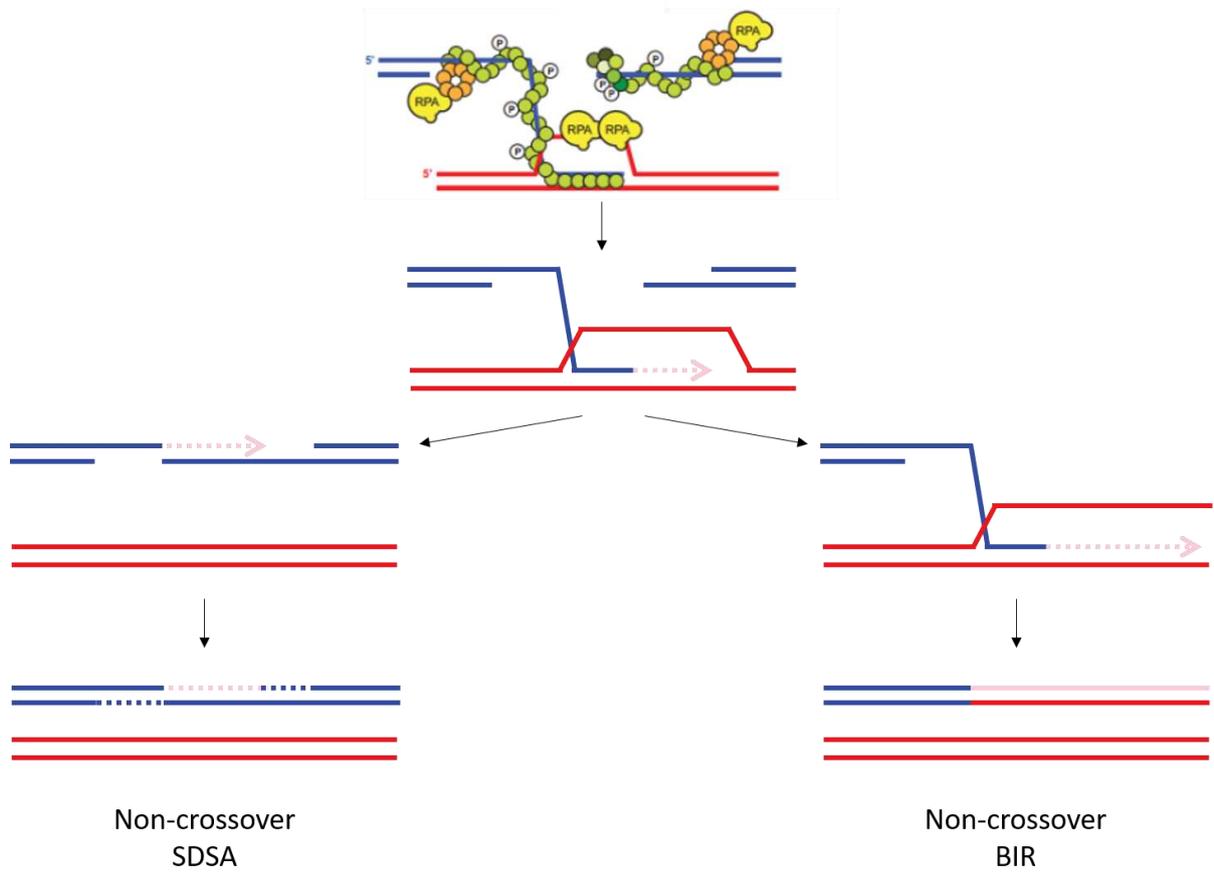


**Figure 25 : Représentation schématique de la réparation d'une cassure double brin par le système de recombinaison chez la levure d'après Mathiasen and Lisby, 2014. (a) Digestion de brin et formation de l'extrémité 3' libre. (b) Envahissement de brin et formation de la jonction de Holliday.**

utilisée alors qu'en phases S et G2, c'est la recombinaison homologue qui est utilisée (Mathiasen and Lisby, 2014). Cela vient du fait qu'il existe chez la levure un ensemble de gènes dont l'expression est dépendante du cycle de croissance. C'est notamment le cas du facteur  $\gamma$ Ku70, qui est fortement exprimé au début de la phase G1 et faiblement exprimé en phase S. Ce facteur fait partie du complexe Ku ( $\gamma$ Ku70 et  $\gamma$ Ku80) qui, en se fixant aux CDB, bloque la recombinaison homologue et laisse une fenêtre de temps pour l'action de la NHEJ.

En absence de Ku, le complexe MRX et l'endonucléase Sae2 reconnaissent la CDB et initient la dégradation de l'ADN de 5' en 3', ce qui correspond à la première étape du processus de recombinaison homologue (figure 25). Après la digestion des 50 à 200 premiers nucléotides, une seconde dégradation pouvant atteindre plusieurs kb est catalysée par Exo1 ou le complexe STR/Dna2, aboutissant à une extrémité 3' OH libre (figure 25). Cette extrémité simple brin est reconnue par les protéines RPA qui viennent s'y fixer pour la protéger. Cependant, ces protéines bloquent la fixation de la recombinase Rad51 sur l'ADNsb. C'est pourquoi le médiateur Rad52 est recruté par les protéines RPA elles-mêmes. Rad52 est capable de déstabiliser les protéines RPA, ce qui permet la fixation de Rad51 sur l'ADNsb et la formation du filament nucléoprotéique (figure 25). De plus, les protéines Rad55 et Rad57 stabilisent le filament, et le complexe SHU le protège de l'hélicase Srs2 (figure 25). Il est important de noter que les protéines RPA se fixent sur l'ADNsb quelle que soit la phase de croissance, mais elles ne recrutent le médiateur Rad52 que lors des phases S et G2. Le filament nucléoprotéique de Rad51 va ensuite « scanner » le génome à la recherche de séquences d'homologie intactes. Une fois la région d'homologie trouvée, Rad51 catalyse la formation d'une D loop par un envahissement de brin. Cette boucle est stabilisée par la fixation de protéines RPA sur le brin désapparié (figure 25). Ensuite, débute une synthèse d'ADN à partir de l'extrémité 3' libre du brin envahisseur en utilisant le brin envahi comme matrice (Mathiasen and Lisby, 2014) (figure 25). Après cette extension de l'extrémité 3', trois voies peuvent être empruntées :

- 1) La voie de réparation « classique ». Dans cette voie, les protéines Rad52 et Rad59 entraînent la capture du second brin dans la D loop, ce qui forme une double jonction de Holliday (dJH). Dans cette voie, les deux extrémités de la CDB sont engagées dans la boucle. Ensuite, la jonction peut être résolue par la résolvasse Mms4-Mus81, aboutissant soit à un produit de « crossover » soit à un produit de « non-crossover ».



**Figure 26 : Représentation schématique des voies SDSA et BIR pour la résolution de la fourche de Holliday lors de la réparation d'une cassure double brin chez la levure.**

Elle peut aussi être désassemblée par le complexe STR, aboutissant à un produit de « non-crossover » (Mathiasen and Lisby, 2014) (figure 25).

- 2) La voie de liaison des brins dépendant de la synthèse (SDSA : Synthesis-Dependent Strand Annealing). Après synthèse, le brin envahisseur est désengagé et apparié avec la seconde extrémité de la CDB, entraînant un produit « non-crossover » (Li and Heyer, 2008) (figure 26).
- 3) La voie de réplication induite par la rupture (BIR : Break-Induced Replication). Dans cette voie, la D loop correspond à une fourche de réplication complète, ce qui entraîne une réplication de tout le reste du chromosome, aboutissant à une perte d'hétérozygotie (Li and Heyer, 2008) (figure 26).

Cette capacité, initialement de réparation de l'ADN est maintenant utilisée pour de l'ingénierie de génomes par transformation de fragments d'ADN. Ces derniers sont traités par le système de recombinaison homologue pour être insérés au génome. La recombinaison homologue chez la levure permet donc de réaliser des modifications génétiques ciblées en utilisant des matrices de recombinaison possédant des séquences ayant un haut degré d'identité avec la zone cible.

## 4.2. Les systèmes de recombinaison site-spécifique

### 4.2.1. Généralités

La recombinaison site-spécifique correspond à un processus d'échange d'ADN programmé entre deux séquences cibles précises, médiée par une famille de protéines appelées recombinases site-spécifiques. La réaction met en jeu deux sites de recombinaison et une recombinase spécifique, capable de reconnaître les sites pour casser puis relier les brins d'ADN. C'est un processus qui ne requiert ni dégradation et synthèse d'ADN, ni cofacteur énergétique. La recombinaison site-spécifique correspond à un besoin de séparer ou joindre des segments d'ADN. Elle joue donc un rôle important dans de nombreuses fonctions comme la régulation de l'expression de gènes, la génération de diversité génétique par l'intégration ou l'excision d'éléments génétiques, ainsi que la réduction de dimères de chromosome en monomères, notamment lors de la réplication (Grindley *et al.*, 2006).

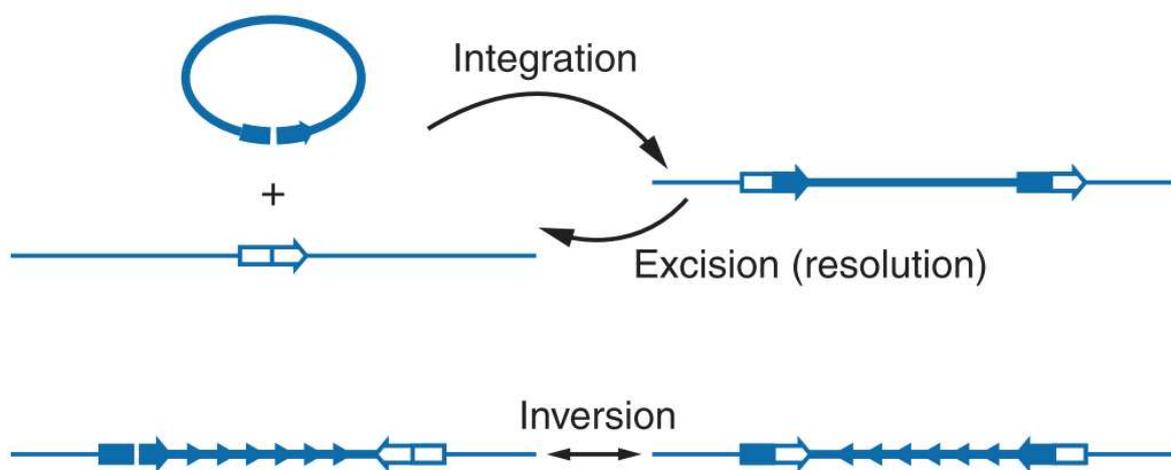


Figure 27 : Schéma des différents scénarios de recombinaison via un système de recombinaison site spécifique, en fonction de l'orientation des sites d'après Grindley *et al.*, 2006.

Quel que soit le type de recombinaison site-spécifique, le mécanisme général reste le même et peut être découpé en quatre étapes : (i) La recombinase se fixe aux deux sites de recombinaison, (ii) un complexe synaptique est formé en associant les deux sites liés aux recombinases, (iii) les recombinases catalysent la cassure, l'échange de brins et le ré-attachement de l'ADN au sein du complexe synaptique, et (iv) le complexe se délisse, libérant les produits de recombinaison (Grindley *et al.*, 2006).

En fonction de l'orientation des sites, trois produits de recombinaison sont possibles : l'intégration, l'excision ou l'inversion. L'intégration résulte de la recombinaison entre deux sites situés sur deux molécules d'ADN différentes dont au moins l'une d'elles est circulaire. L'excision et l'inversion surviennent lorsque les deux sites sont sur la même molécule. Le résultat de la réaction dépend de l'orientation des sites. L'excision provient d'une recombinaison entre deux sites en orientation directe (tête-à-queue) aboutissant à la circularisation de la région présente entre les deux sites. Enfin, l'inversion est le résultat de la recombinaison entre deux sites en orientation inverse (tête-à-tête) (figure 27).

Bien qu'un grand nombre de systèmes de recombinaison site-spécifique ait été identifié, ils peuvent tous être classés en deux familles : les recombinases à sérine et les recombinases à tyrosine. Cette classification est basée sur le mécanisme de recombinaison et sur l'acide aminé porté par la recombinase impliquée dans la formation du complexe recombinase/site de recombinaison.

- La famille des recombinases à sérine, aussi appelée famille « resolvase/invertase », est composée de protéines très hétérologues avec des tailles allant de 180 à 800 acides aminés (AA) avec une organisation des domaines très variables (Grindley *et al.*, 2006). Malgré cette variabilité (notamment au niveau de l'extrémité C-terminale), tous les membres de cette famille possèdent en N-terminal un domaine catalytique caractéristique d'une taille constante d'environ 150 AA. Dans ce domaine, des résidus identifiés comme contribuant à la structure du site actif de recombinaison sont très conservés (Stark, 2015). Les variations au niveau de l'extrémité C-terminale permettent de différencier deux sous-classes de recombinases à sérine :
  - o Les petites recombinases à sérine caractérisées par une extrémité C-terminale de petite taille constituée d'un petit domaine de liaison à l'ADN de type hélice-tour-hélice.



- Les grandes recombinases à sérine qui possèdent des extrémités C-terminales de grandes tailles composées de plusieurs domaines variables de liaison à l'ADN et de régulation.

Quatre systèmes de recombinaison médiés par les recombinases à sérine ont été particulièrement étudiés : (i-ii) les resolvases  $\gamma\delta$  et Tn3 nécessaires à la résolution des intermédiaires de transposition  $\gamma\delta$  et Tn3. Les resolvases catalysent des recombinaisons entre deux sites en orientation directe (tête-à-queue) pour éliminer l'ADN entre ces sites. (iii-iv) Les invertases qui contrôlent l'expression génique des antigènes des flagelles de *Salmonella* (invertase Hin) et des fibres de la queue dans le phage Mu (invertase Gin) (Brown *et al.*, 2011). Les invertases recombinent les sites en orientation inversée (tête-à-tête) pour provoquer une inversion de l'ADN entre les sites (Smith and Thorpe, 2002).

- Les recombinases à tyrosine sont retrouvées principalement chez les procaryotes et leurs virus (bactériophages), même si ces protéines sont aussi présentes chez les archées (Badel *et al.*, 2021) et des eucaryotes (Jayaram *et al.*, 2015). L'intégrase du phage  $\lambda$  est le premier membre de cette famille à avoir été identifié. En effet, en 1974, à partir d'*E. coli* infectées par le phage  $\lambda$ , Ausubel purifie cette protéine qu'il nomme intégrase (AUSUBEL, 1974). Par la suite, en 1975, Nash démontre, *in vitro*, son activité de recombinaison. Ce n'est qu'en 1988 que Pargellis démontre que, dans cette protéine, un résidu tyrosine est responsable de la formation de liaisons covalentes avec l'ADN (Pargellis *et al.*, 1988). C'est à cause de ces travaux que la famille des tyrosine recombinases avait été initialement appelée famille  $\lambda$  Intégrase (Grindley *et al.*, 2006). Elle est composée d'un grand nombre de recombinases, environ 1300 ont été identifiées et classées en familles et sous-familles (Van Houdt *et al.*, 2012). D'un point de vue structural, les recombinases à tyrosine sont assez hétérogènes. Elles sont composées d'un domaine catalytique en C-terminal qui possède, en plus de la tyrosine catalytique, six résidus hautement conservés (Arg, Lys, His, Arg, His/Trp et Asp/Glu) (Jayaram *et al.*, 2015). Certains membres ne contiennent que ce domaine. Cependant dans la plupart des cas, le domaine catalytique est associé à un domaine très variable en N-terminal, permettant des interactions avec d'autres protéines accessoires et qui peuvent aider à la liaison à l'ADN (Grindley *et al.*, 2006). On peut donc différencier les recombinases à tyrosine dites « simples » des recombinases à tyrosine plus complexes.



Les recombinaisons à tyrosine dites « simples » (comme par exemple Cre et Flp) ne nécessitent pas de topologie particulière de l'ADN, ni une orientation spécifique du site cible, ni de protéines accessoires. Elles peuvent catalyser des réactions intra ou intermoléculaires, sur des ADN super enroulés, circulaires ou linéaires. Les recombinaisons à tyrosine plus complexes (comme par exemple la  $\lambda$  Intégrase) ont besoin d'un contexte réactionnel précis qui peut nécessiter une topologie particulière de l'ADN ainsi qu'une interaction avec des protéines accessoires (Jayaram *et al.*, 2015). Chaque recombinaison possède ses propres sites de reconnaissance sur l'ADN. Au minimum, ces sites de recombinaison sont composés de deux séquences d'ADN inversées, permettant la fixation de la recombinaison, séparées par un « spacer » de 6 à 8 pb. Le clivage et l'échange de brins ont lieu à la jonction entre les sites de liaisons des recombinaisons et l'extrémité 5' du « spacer ». Certaines recombinaisons complexes ont aussi besoin de sites accessoires pour la fixation de protéines régulatrices.

#### 4.2.2. Le système cre/lox

Le système cre/lox a été identifié en 1981 (Sternberg and Hamilton, 1981). Il provient du bactériophage P1 qui peut infecter *E. coli* et d'autres bactéries. Au lieu de s'insérer dans le génome de son hôte, ce bactériophage est maintenu sous forme de plasmide circulaire à faible nombre de copies. La circularisation de son ADN est due à la recombinaison site-spécifique Cre (Causes recombination) (Hochman *et al.*, 1983). Cette recombinaison fait partie de la famille des recombinaisons à tyrosine dites « simples ». Elle est donc capable de provoquer un événement de recombinaison quel que soit l'état de la molécule d'ADN (super enroulée, circulaire ou linéaire) sans nécessiter l'intervention de protéines accessoires. Elle se fixe sur des séquences cibles de 34 pb appelées loxP (locus of crossing (x) over of P1). Ces séquences sont composées de deux régions inversées de 13 pb (LE : élément gauche (left element) ; RE : élément droit (right element)) pour la liaison de la recombinaison, séparées par une région « spacer » de 8 pb. C'est cette partie centrale (spacer) qui donne l'orientation du site loxP. Tout comme les autres systèmes de recombinaison site-spécifique, c'est l'orientation des sites de recombinaison (ici les sites loxP) qui dicte le résultat de la réaction (ici catalysée par la Cre recombinaison). La bonne annotation des sites loxP est donc primordiale (Shaw *et al.*, 2021).

**Tableau 3: Séquences de différents sites lox et leurs orientations.** Les mutations sont représentées en rouge.

Nom du site lox	Séquence	Orientation
LoxP « classique »	ATAACTTCGTATA ATGTATGC TATACGAAGTTAT	→
Lox66	TACCGTTCGTATA ATGTATGC TATACGAAGTTAT	→
Lox71	ATAACTTCGTATA ATGTATGC TATACGAA <b>CGGTA</b>	→
Lox72	TACCGTTCGTATA ATGTATGC TATACGAA <b>CGGTA</b>	→
Loxm2/66	TACCGTTCGTATA <b>AGAAACCA</b> TATACGAAGTTAT	→
Loxm2/71	ATAACTTCGTATA <b>AGAAACCA</b> TATACGAAC <b>CGGTA</b>	→
Loxm2/72	TACCGTTCGTATA <b>AGAAACCA</b> TATACGAAC <b>CGGTA</b>	→

**Tableau 4 : Comparaison entre les différents sites mutants au niveau du spacer adapté de Langer et al., 2002.**

Nom du site lox	Séquence du spacer	Orientation	Incompatible avec
LoxP « classique »	ATGTATGC	→	m2, m3, m7, m11, 511
Lox511	ATGTAT <b>AC</b>	→	LoxP et non testé avec les mutants
Loxm2	<b>AGAAACCA</b>	→	LoxP
Loxm3	<b>TAATACCA</b>	→	LoxP, m7
Loxm7	<b>AGATAGAA</b>	→	LoxP, m3
Loxm11	<b>CGATACCA</b>	→	LoxP

En 1995, les deux sites lox « mutants » les plus utilisés, le lox66 et le lox71, ont été décrits (Albert *et al.*, 1995). Le lox66 a ses cinq premières paires de bases différentes du site loxP « classique », alors que le site lox71 a ses cinq dernières paires de bases différentes. Chez ces mutants, la partie « spacer » reste inchangée par rapport au loxP « classique ». Les séquences de ces sites sont répertoriées dans le tableau 3. La recombinaise Cre est capable de reconnaître ces sites mutants avec la même affinité que pour le site loxP « classique ». Après recombinaison de deux loxP par l'action de la Cre recombinaise, le produit résultant correspond à une combinaison des deux sites initiaux. Lorsque deux sites loxP « classiques » sont recombines entre eux, cela aboutit à un site loxP fonctionnel. Au contraire, lorsque deux sites mutants sont utilisés et recombines ensemble, la réaction résulte forcément en la formation d'un lox doublement muté (Albert *et al.*, 1995). La recombinaison entre les sites lox66 et lox71 aboutit donc à la formation d'un site double mutant, le lox72. Ce dernier est dit « inactif » car la Cre recombinaise ne possède aucune affinité avec ce site, elle n'est donc pas capable de s'y fixer (Albert *et al.*, 1995). L'utilisation de sites mutants qui aboutissent, après recombinaison, à un site inactif, permet de réaliser plusieurs réactions dans le même génome sans se soucier de potentielles réactions croisées.

De nos jours, il existe plusieurs autres sites lox « mutants » notamment au niveau du spacer, comme par exemple le site lox511 (Langer *et al.*, 2002). Ce type de mutants possède une efficacité de recombinaison réduite avec un site loxP dont le spacer n'est pas muté, mais conserve une capacité de recombinaison avec des sites qui ont les mêmes mutations au niveau du spacer. Cependant, tous les sites lox mutés au niveau de leur spacer, ne sont pas forcément incompatibles entre eux (tableau 4). En effet, seuls les sites loxm3 et loxm7 possèdent une incompatibilité réciproque (Langer *et al.*, 2002).

On retrouve aussi des mutants qui combinent les deux types de mutation (au niveau des éléments LE ou RE et au niveau du spacer). On peut citer par exemple le site loxm2/66 qui est muté au niveau de son spacer et qui possède les mêmes mutations que le lox66 au niveau de ses cinq premières paires de base (tableau 4). On peut aussi citer le site loxm2/71 qui correspond au lox71 au niveau de ses cinq dernières paires de bases et qui possède les mêmes mutations que le loxm2/66 au niveau de son spacer (tableau 4). Les sites loxm2/66 et loxm2/71 sont donc compatibles entre eux, mais incompatibles avec un site loxP « classique » (Langer *et al.*, 2002).

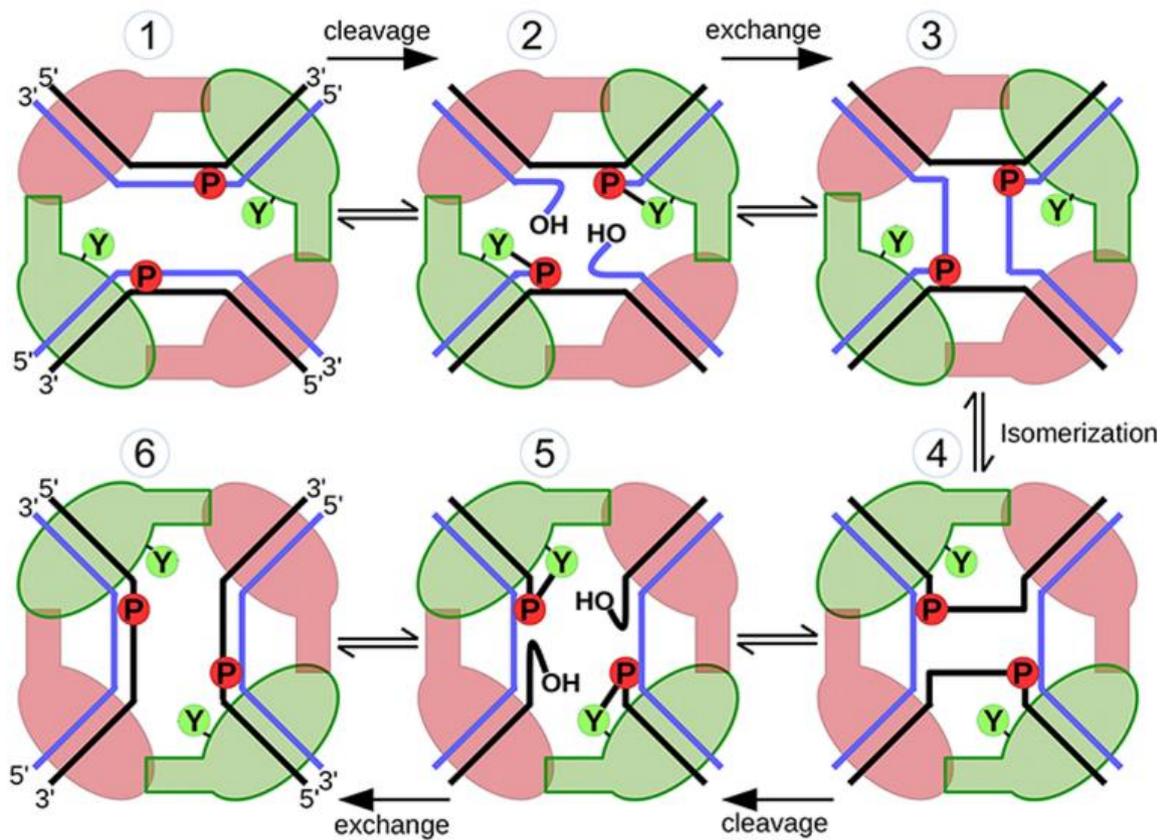
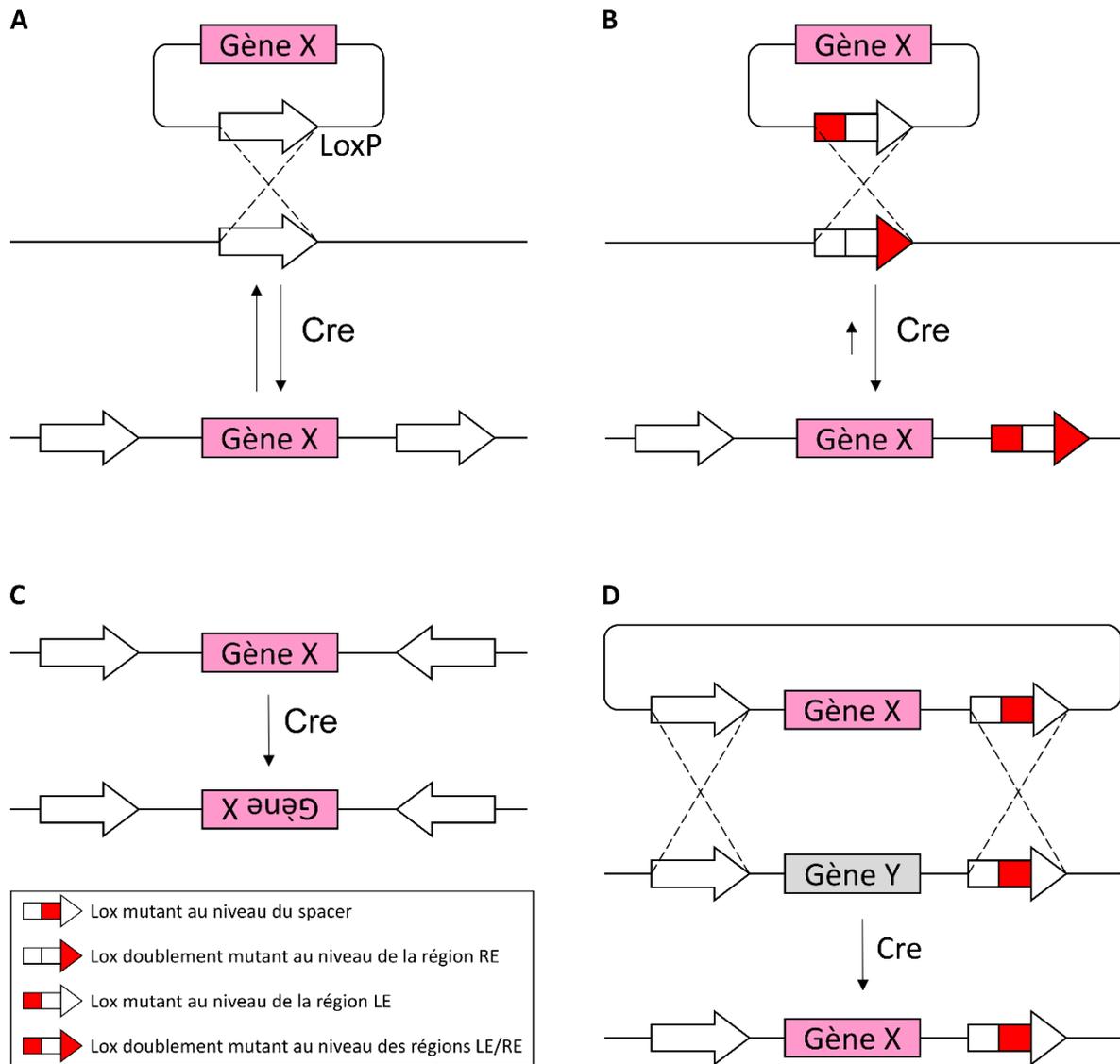


Figure 28 : Représentation schématique du mécanisme de recombinaison site-spécifique via le système cre/lox d'après Meinke *et al.*, 2016. Les recombinases actives sont représentées en vert, les recombinases inactives sont représentées en rouge. Y : tyrosine catalytique. P : phosphate

La compatibilité et incompatibilité entre sites loxP permettent de faciliter l'utilisation du système cre/lox pour provoquer certains événements de modification génétique spécifique. En effet, récemment, un set de 16 variantes de sites lox a été établi. Il a été montré qu'il est possible de les utiliser pour de l'ingénierie de génome multiple, à la fois chez des procaryotes et chez des eucaryotes (Cautereels *et al.*, 2024).

La recombinaison via le système cre/lox est un processus en 6 étapes, qui fait intervenir quatre protéines Cre pour la coupure et la jonction de quatre sites loxP (figure 28). La première étape correspond à la formation du complexe synaptique, composé de quatre recombinases Cre en interaction avec deux sites loxP, formant ainsi un complexe tétramérique (figure 28). Dans ce complexe, deux recombinases Cre opposées sont inactives alors que les deux autres sont actives et peuvent couper un des deux brins d'ADN de chacun des sites loxP. Cette coupure correspond à la deuxième étape et est réalisée par le résidu tyrosine conservé de la Cre recombinase qui va former un intermédiaire 3' phospho-tyrosine, libérant ainsi une extrémité 5'-OH libre (figure 28). La troisième étape consiste en un échange de brins par jonction du 5'-OH libre avec le 3' phospho-tyrosine voisin. Cet échange de brins aboutit à la formation d'une jonction de Holliday (figure 28). Lors de la quatrième étape, cette jonction de Holliday s'isomérisé, ce qui va modifier la structure des promoteurs des recombinases et aboutir à l'inactivation des recombinases précédemment actives et à l'activation de celles précédemment inactives (figure 28). Ensuite les étapes 5 et 6 correspondent à une répétition des étapes de coupure et de jonction cette fois avec le deuxième brin, produisant *in fine* une recombinaison d'ADN (Meinke *et al.*, 2016) (figure 28).

Le système cre/lox, maintenant utilisé pour de l'ingénierie de génome, est capable de réaliser des délétions, des insertions, des inversions ou de l'échange d'ADN (figure 29). L'utilisation du système cre/lox pour engendrer une délétion peut notamment être utilisée pour la production de mutations sans marqueur (cf partie 1.7.2 de l'introduction). Pour de l'insertion, l'utilisation de sites mutants au niveau des régions LE/RE comme les sites lox66 et lox71 qui aboutissent à un site lox72 (inactif), permet de contrôler la réaction de recombinaison dans le sens de l'insertion (figure 29). Enfin, une stratégie appelée RMCE (Recombinase-Mediated Cassette Exchange) consiste à provoquer un échange d'ADN par une double réaction de recombinaison (figure 29D) (Turan *et al.*, 2011). Dans cette stratégie, deux régions d'ADN sur deux molécules différentes sont bordées d'un site loxP « classique » (sans



**Figure 29 : Schéma des différentes réactions de recombinaison possibles via le système cre/lox. [A]** Intégration et délétion d'une cassette d'ADN. **[B]** Utilisation de sites mutés au niveau des régions LE et RE pour diriger la réaction de recombinaison dans le sens de l'insertion. La recombinaison de ces deux sites mutants aboutit à la formation d'un site sauvage et d'un site doublement muté dit « inactif » limitant la réaction de délétion. **[C]** Inversion. **[D]** Echange d'ADN par la stratégie RMCE (Recombinase-Mediated Cassette Exchange).

mutation au niveau du spacer), et d'un site mutant (avec une mutation au niveau du spacer). La recombinase provoque un échange entre ces deux régions d'ADN par une double réaction de recombinaison, entre les sites loxP « classiques » entre eux et les sites mutants entre eux (figure 29).



# Matériels et Méthodes



## Matériels et Méthodes

### 1. Souches microbiennes et conditions de cultures

#### 1.1. *Escherichia coli*

Plusieurs souches commerciales d'*E. coli* ont été utilisées pour propager et amplifier les plasmides :

- ***E. coli* DH5  $\alpha$**  [*fhuA2* $\Delta$ (*argF-lacZ*) *U169 phoA glnV44  $\Phi$ 80* $\Delta$ (*lacZ*)*M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*] commercialisée par New England Biolab (NEB).
- ***E. coli* DH10  $\beta$**  [ $\Delta$ (*ara-leu*) *7697 araD139 fhuA  $\Delta$ lacX74 galK16 galE15 e14-  $\phi$ 80dlacZ* $\Delta$ *M15 recA1 relA1 endA1 nupG rpsL (Str<sup>R</sup>) rph spoT1  $\Delta$ (mrr-hsdRMS-mcrBC)*] commercialisée par New England Biolabs (NEB).
- ***E. coli* ElectroMAX™ Stbl4™** [*mcrA  $\Delta$ (mcrBC-hsdRMS-mrr) recA1 endA1 gyrA96 gal thi-1 supE44  $\lambda$ -relA1  $\Delta$ (lac-proAB)/F' proAB<sup>+</sup>lacI<sup>q</sup>* $\Delta$ *M15 Tn10 (Tet<sup>R</sup>)*] commercialisée par Invitrogen. Cette souche a été utilisée pour le clonage d'inserts de grande taille et/ou instables dans les souches précédemment citées.
- ***E. coli* EPI 300 TransfoMax** [*F- mcrA  $\Delta$ (mrr-hsdRMS-mcrBC)  $\phi$ 80dlacZ* $\Delta$ *M15  $\Delta$ lacX74 recA1 endA1 araD139  $\Delta$ (ara, leu)7697 galU galK  $\lambda$ rpsL nupG trfA dhfr*] commercialisée par Lucigen. Cette souche contient un système permettant de contrôler le nombre de copie du plasmide pCC1FOS (ou dérivés) par cellule.

Ces souches ont été cultivées à 37°C ou 30°C pour la souche ElectroMAX™ Stbl4™, en milieu LB, solide ou liquide (200 rpm), additionné d'ampicilline (100  $\mu$ g/ml), spectinomycine (50  $\mu$ g/ml) ou d'érythromycine (200  $\mu$ g/ml) en fonction du gène de résistance porté par le plasmide.

#### 1.2. *Bacillus subtilis*

Plusieurs souches de *B. subtilis* ont été utilisées au cours de ces travaux. Elles nous été fournies par nos collaborateurs de l'équipe Syber (unité MICALIS, INRAE de Jouy-en-Josas).

- ***B. subtilis* 168** est la souche de référence. Elle a un génome d'une taille de 4.215 Mb. Elle a été identifiée en 1872 par Cohn (Cohn, 1872).

**Tableau 5: Liste des souches construites à partir de la souche AG686, ainsi que les ARS ajoutées dans chacune d'elles.**

<b>Souches construites</b>	<b>ARS insérées</b>
AG 686	-
JB 686	ARSH4
AG 694	ARSH4, ARSH1
AG 703	ARSH4, ARSH1, ARS 1332, ARS1
EDJ 2039	ARSH4, ARSH1, ARS 1332, ARS1, ARS 1018, ARS 416
EDJ 2052	ARSH4, ARSH1, ARS 1332, ARS1, ARS 1018, ARS 416, ARS 727, ARS 516
EDJ 2052 ARS404	ARSH4, ARSH1, ARS 1332, ARS1, ARS 1018, ARS 416, ARS 727, ARS 516, ARS 404
AG 709	ARSH4, ARSH1, ARS 1332, ARS1, ARS 1018, ARS 416, ARS 727, ARS 516, ARS 603, ARS 1120, ARS 914, ARS 1213, ARS 209, ARS 1113, ARSMax2
AG 709 ARS404	ARSH4, ARSH1, ARS 1332, ARS1, ARS 1018, ARS 416, ARS 727, ARS 516, ARS 603, ARS 1120, ARS 914, ARS 1213, ARS 209, ARS 1113, ARSMax2, ARS 404

- ***B. subtilis* Master Strain (MS)** [*trpC2*  $\Delta$ *ICEBs1*,  $\Delta$ *PBSX*,  $\Delta$ *SPB*,  $\Delta$ *skin*,  $\Delta$ (*gtaB yvzH ggaB ggaA yvzI yvzE*), *Δupp::*[ $\lambda$ pR-neo] dérive de la souche de référence *B. subtilis* 168 par réduction itérative de son génome (Dervyn *et al.*, 2023). Elle a un génome d'une taille de 3,981 Mb.
- ***B. subtilis* FE220** [*trpC2*  $\Delta$ *ICEBs1*,  $\Delta$ *PBSX*,  $\Delta$ *SPB*,  $\Delta$ *skin*,  $\Delta$ (*gtaByvzH,ggaB,ggaA,yvzI,yvzE*), *Δupp::*[ $\lambda$ Pr-neo],  $\Delta$ (*sboA-albG*),  $\Delta$ (*ydzT-ydjC*)::*clupp*-Phleo]] dérive de la MS par délétion de la région du prophage 3 (entre les gènes *ydzT* et *ydjC*) qui contient, entre autres, les gènes codant pour le système de restriction modification BsuM. Elle a un génome d'une taille de 3,964 Mb.
- ***B. subtilis* MGP192** (Minimal Genome Project) dérive de la 168 par réduction itérative de son génome (Dervyn *et al.*, 2023). Elle a un génome d'une taille de 2,868 Mb
- ***B. subtilis* AG686** dérive de la MS par réduction itérative de son génome. Elle a un génome moins réduit que la souche MGP192 d'une taille de 2,944 Mb. Elle est modifiée pour être plus compétente grâce à la surexpression des gènes *comSK*. C'est de cette souche que vont dériver toutes les souches contenant les ARS supplémentaires ajoutées tout le long du génome (tableau 5). Elle est résistante à la tétracycline et au chloramphénicol.
- ***B. subtilis* CW22** [BSB1 *amyE*::{*P<sub>veg</sub>* rfp specR}] exprime la RFP de manière constitutive, la protéine se retrouve libre dans le cytoplasme. Cette souche est résistante à la spectinomycine. Elle a été utilisée pour les observations en microscopie à fluorescence. Elle a été construite par l'équipe SyBER (Institut MICALIS UMR 1319 INRAE / AgroParisTech, Jouy-en-Josas).

Ces souches ont été cultivées à 37°C, en milieu LB, solide ou liquide (200 rpm), additionné de spectinomycine (100 µg/ml), chloramphénicol (4 µg/ml), tétracycline (7,5 µg/ml), kanamycine (7,5 µg/ml) ou érythromycine 1 µg/ml puis 5 µg/ml, selon le gène de résistance hébergé.

### 1.3. *Saccharomyces cerevisiae*

Plusieurs souches présentes au laboratoire ont été ramenées du J. Craig Venter Institute (JCVI) par le Dr. Carole Lartigue. Elles ont été utilisées pour le clonage des fragments de génome de *B. subtilis*.

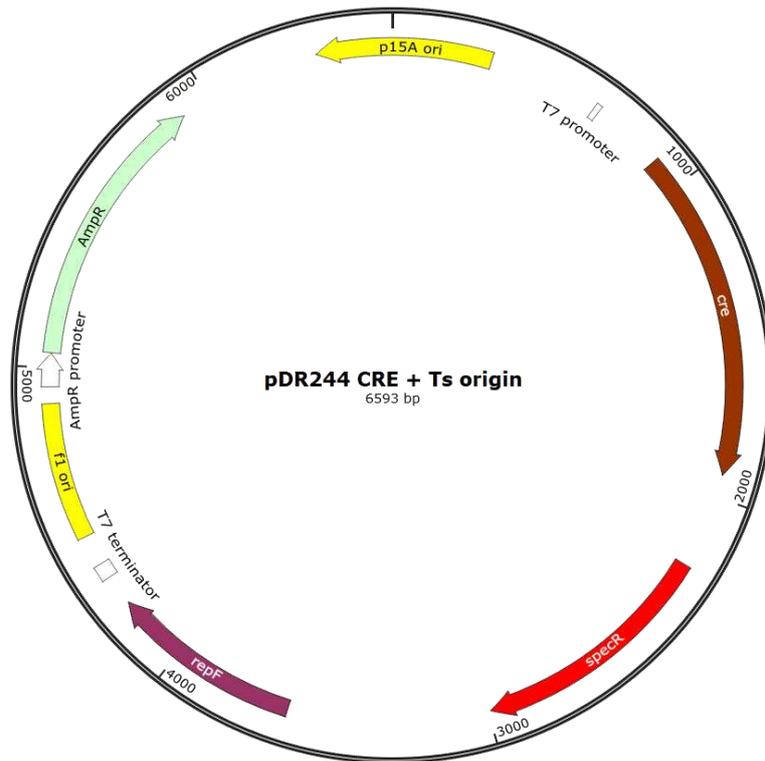


Figure 30: Carte du plasmide pDR244 CRE + Ts origin

- **VL6-48** (ATCC MYA-3666) [*MAT $\alpha$* , *his3- $\Delta$ 200*, *trp1- $\Delta$ 1*, *ura3-52*, *lys2*, *ade2-101*, *met14*, *psi+cir<sup>o</sup>*] a été construite par V. Larionov et N. Kouprina (Larionov *et al.*, 1994).
- **VL6-48N** [*MAT $\alpha$* , *his3- $\Delta$ 200*, *trp1- $\Delta$ 1*, *ura3- $\Delta$ 1*, *lys2*, *ade2-101*, *met14*, *psi+cir<sup>o</sup>*] dérive de la souche VL6-48 par substitution du gène *ura3-52* par une cassette KanMX (Larionov *et al.*, 1997).
- **W303a** [*MAT $\alpha$*  {*leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*} [*phi<sup>+</sup>*]] a été construite par Rodney Rothstein (Rothstein, 1983).

Ces souches ont été cultivées à 30°C en milieu YPDA sous agitation (200 rpm). Les transformants levure ont été cultivés en milieu synthétique défini minimum (Minimal SD Base) solide ou liquide (200 rpm), dépourvu d'histidine (HIS), d'adénine (ADE), d'uracile (URA), de tryptophane (TRP), de leucine (LEU), de lysine (LYS) seul ou en combinaison, en fonction du ou des gènes d'auxotrophie portés par les plasmides ou fragments de génome clonés.

## 2. Techniques de biologie moléculaire

### 2.1. Oligonucléotides

Au cours de ces travaux, des oligonucléotides ont été utilisés à la fois pour réaliser des amplifications PCR nécessaires à (i) la construction des plasmides, (ii) la production de pansements ou (iii) l'analyse de transformants. Ils ont tous été synthétisés par la société prestataire Eurogentec et sont listés dans le tableau S1.

### 2.2. Plasmides

#### 2.2.1. Plasmides commerciaux ou précédemment construits

**Le plasmide pDR244 CRE + Ts origin:** Ce plasmide contient la séquence codant le gène *cre* ainsi qu'une origine de répllication sensible à la température et le marqueur de résistance *specR* conférant la résistance à la spectinomycine (figure 30).



Figure 31: Carte du plasmide pJIM

**Le plasmide pJIM :** Ce plasmide contient une origine de réplication *E. coli* (*ColE1*), ainsi qu'un marqueur de résistance *ermB* conférant la résistance à l'érythromycine. C'est un plasmide répliatif chez *B. subtilis* grâce aux trois gènes *copF*, *repD* et *repE* (figure 31).

**Le plasmide pShuttle *Bsu*-levure :** C'est un plasmide navette répliatif à la fois chez *B. subtilis* et la levure. Il a été construit à partir du pJIM donc il contient comme lui une origine de réplication *E. coli* (*ColE1*), le marqueur de résistance *ermB*. Il est répliatif chez *B. subtilis* grâce aux trois gènes *copF*, *repD* et *repE*. Il contient aussi un second marqueur de sélection, *specR* ainsi que les éléments levure ARSH4, CEN6, et que le marqueur d'auxotrophie pour l'histidine (HIS3) permettant son maintien dans la levure.

**Le plasmide pDR111 :** Ce plasmide contient une origine de réplication *E. coli* (*ColE1*) ainsi que les marqueurs de sélection *specR* et *ampR* conférant respectivement la résistance à la spectinomycine et à l'ampicilline. C'est un plasmide intégratif chez *B. subtilis*. Deux cassettes de recombinaison correspondant au début et à la fin du gène *amyE* permettent l'intégration du marqueur de résistance *specR* dans le génome de *B. subtilis* par double recombinaison.

**Le plasmide pICE ADE2 :** Ce plasmide a été développé par Flagg *et al.* (Flagg *et al.*, 2019) et est commercialisé par Addgene (#122876). Il contient les séquences dites « éléments levure », c'est-à-dire un centromère (CEN6), une origine de réplication levure (ARSH4) et un marqueur d'auxotrophie pour l'adénine (ADE2). Il contient également le marqueur de sélection *ampR*.

**Le plasmide pICE LYS2 :** Ce plasmide a été développé par Flagg *et al.* (Flagg *et al.*, 2019) et est commercialisé par Addgene (#122879). Il contient les éléments levure ARSH4, CEN6 avec un marqueur d'auxotrophie pour la lysine (LYS2). Il contient également le marqueur de sélection *ampR*.

**Le plasmide pCC1 YTRP :** Ce plasmide est un vecteur navette de type BAC que l'on peut cloner à la fois dans la levure et dans *E. coli* (Garcia-Morales *et al.*, 2020). Il contient les éléments levure ARSH4, CEN6 et un marqueur d'auxotrophie pour le tryptophane (TRP1), symbolisés par la lettre « Y » (pour Yeast) dans le nom du plasmide. L'utilisation de la souche *E. coli* EPI300 permet de contrôler le nombre de copies de ce plasmide.

**Le plasmide pTwist CmR *loxm2/66* :** Ce plasmide a été synthétisé par TWIST Bioscience®. Il contient une origine de réplication *E. coli* (*ColE1*) ainsi qu'un marqueur de sélection *ampR*.

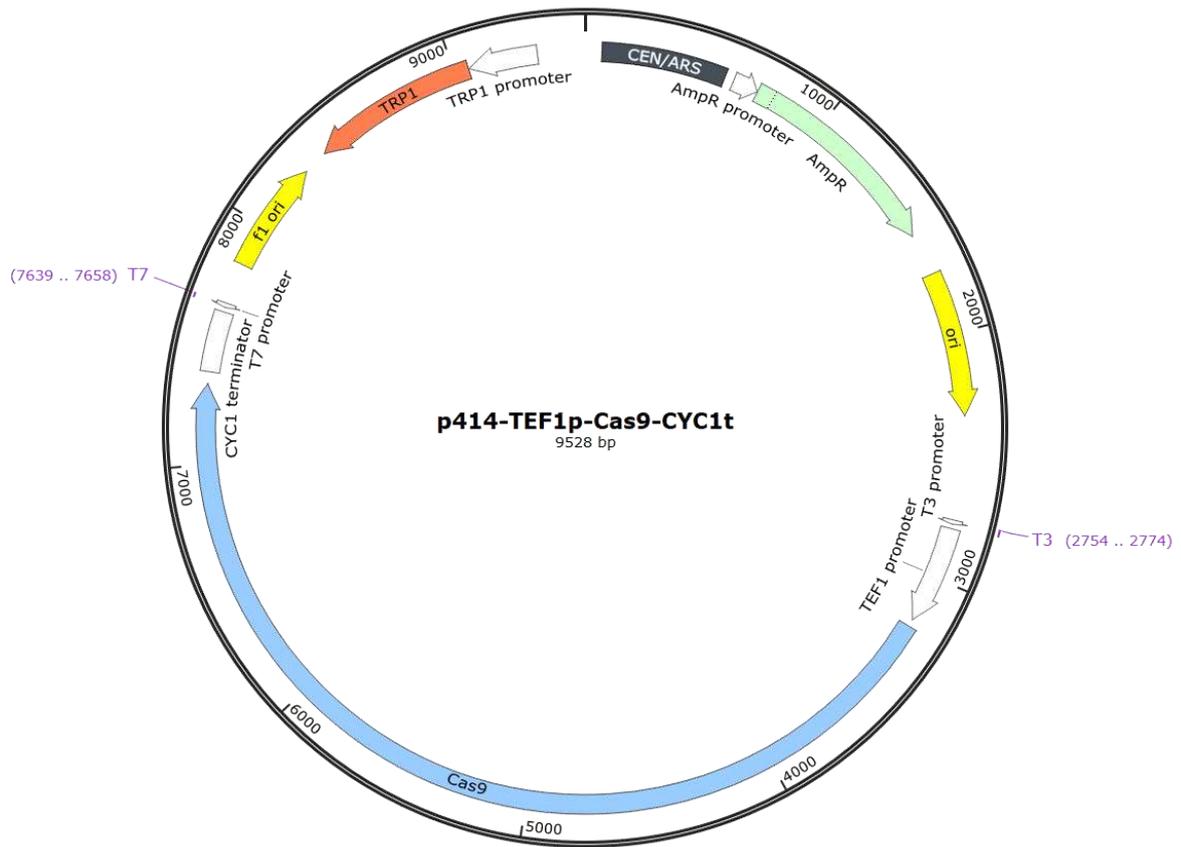


Figure 32: Carte du plasmide p414-TEF1p-Cas9-CYC1t

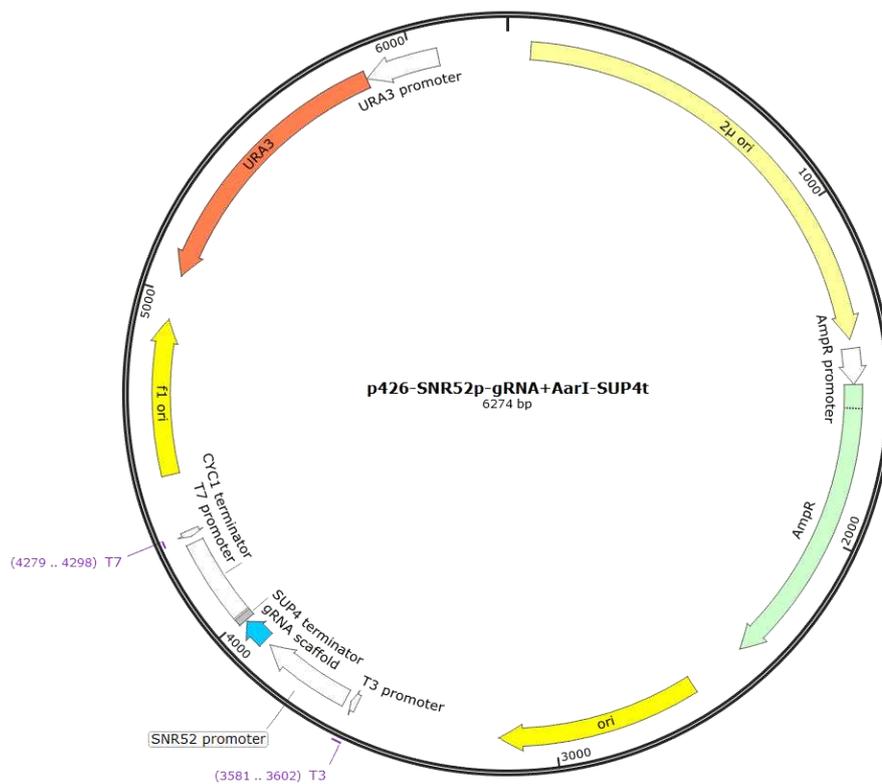


Figure 33: Carte du plasmide p426-SNR52p-gRNA.Aar1.Y-SUP4t

C'est un plasmide intégratif chez *B. subtilis*. Deux cassettes de recombinaison correspondant aux gènes *trmE* et *trmF* permettent l'intégration du marqueur de résistance *cat* conférant la résistance au chloramphénicol ainsi que le site de recombinaison *loxm2/66* (tableau 3).

**Le plasmide pTwist TetR lox71** : Ce plasmide a été synthétisé par TWIST Bioscience®. Il contient une origine de répllication *E. coli* (*ColE1*) ainsi qu'un marqueur de sélection *ampR*. C'est un plasmide intégratif chez *B. subtilis*. Deux cassettes de recombinaison correspondant aux gènes *yxep* et *yxq* permettent l'intégration du marqueur de résistance *tetR* conférant la résistance à la tétracycline, ainsi que le site de recombinaison *lox71* (tableau 3).

**Le plasmide p414-TEF1p-Cas9-CYC1t** : Ce plasmide centromérique permet l'expression constitutive chez la levure, d'une version codon optimisée de la protéine Cas9 de *Streptococcus pyogenes*. Ce plasmide contient les éléments levure ARSH4, CEN6 et le marqueur TRP1. L'expression constitutive de la Cas9 est contrôlée par le promoteur constitutif TEF1P, et le tag SV40 en C-terminal lui confère une localisation nucléaire. Il possède aussi une origine de répllication *E. coli* (*ColE1*) ainsi qu'un marqueur de sélection *ampR*. Le plasmide p414-TEF1p-Cas9-CYC1t a été développé en 2013 (DiCarlo *et al.*, 2013) et est commercialisé par Addgene (#43802) (figure 32).

**Le plasmide p426-SNR52p-gRNA.Aar1.Y-SUP4t** : Ce plasmide permet l'expression chez la levure d'un ARNg, composé d'une séquence de 20 bases spécifiques du gène cible « spacer » qui est fusionné avec un composant structural (tracrARN). L'expression de l'ARNg est sous le contrôle du promoteur constitutif SNR52. C'est un plasmide à grand nombre de copie disposant d'une origine de répllication  $2\mu$  et d'un marqueur d'auxotrophie pour l'uracile (URA3). Il a été optimisé par Tsarmopoulos *et al.* (Tsarmopoulos *et al.*, 2016) à partir du vecteur initialement développé par DiCarlo *et al.* (DiCarlo *et al.*, 2013) et disponible chez Addgene (#43803). Le « spacer » original ciblant le gène de levure CAN1.Y a été remplacé par un site de clonage comprenant deux sites de restriction « tête-bêche » de l'enzyme AarI. Cette région peut être remplacée par le « spacer » de notre choix, par restriction et ligature d'un duplex d'oligonucléotides (figure 33).

**Le plasmide p426-SNR52p-gRNA.Aar1.Y-SUP4t\_ADE2** : Ce plasmide dérive du p426-SNR52p-gRNA.Aar1.Y-SUP4t. Le marqueur URA3 a été remplacé par le marqueur ADE2.

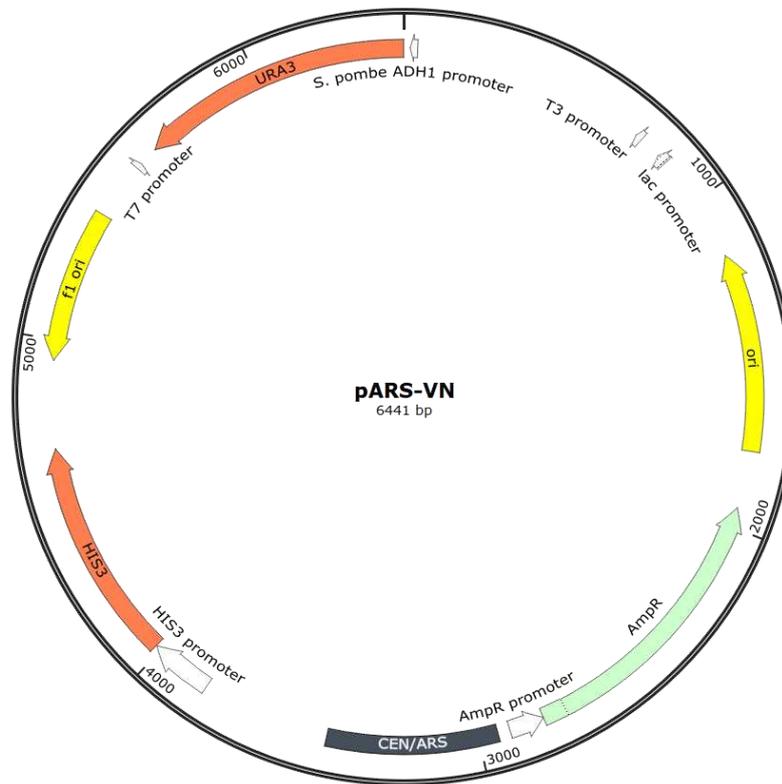


Figure 34: Carte du plasmide pARS-VN

Tableau 6: Liste des plasmides d'expression d'ARNg produits

Utilisation	plasmide	Séquence « spacer »
CReasPy-Fusion	p426-SNR52p-gRNA 1 frgt 2	CGGTTTGCAAGACAGATGCCG
	p426-SNR52p-gRNA 2 frgt 2	TGAATACGCCAATTCTACCG
	p426-SNR52p-gRNA 1 frgt 3	ATATTAAAAACCCTCTACCG
	p426-SNR52p-gRNA 2 frgt 3	ATAACTGTAAAGGATACCCG
	p426-SNR52p-gRNA1 Fgt4-2,5	GTTTTGATAGATTCTGACAG
	p426-SNR52p-gRNA 2 Fgt4-2,5	TGCTGAGATCGAAAAAGCCG
	p426-SNR52p-gRNA 1 Fgt4-2,64	ACTCAGGCCAAAACCTGAACCG
	p426-SNR52p-gRNA 2 Fgt4-2,64	GCAATGTTTCTACTACGCCG
	p426-SNR52p-gRNA2 Frgt5-2,78	AGACGGACTTTAATGCCATG
	p426-SNR52p-gRNA 1 demi genome	ACCATCGTGATAAATAACAG
p426-SNR52p-gRNA2 demi genome	GCTTCATTACGGAGAAACCG	
RAGE	p426 gRNA 1-frgtsBsu-spec-ACH	ATCCGACTATTTGGACGACG
	p426 gRNA 2-frgtsBsu-spec-ACH	TTTCCCCGAAAAGTGCCACC
	p426 gRNA 3-frgtsBsu-spec-ACH	GCTCATGAGACAATAACCCT
	p426 gRNA 1 deletion frgt 7.2	TTAGCGACCCAGATGAATGG
	p426 gRNA 2 deletion frgt 7.2	AAGAAATTCAGCATAAAGCG
	p426 gRNA gRNA 1 marqueur frgt 7.2	GCTGTATCAACCAGACGAAG
	p426 gRNA 2 marqueur frgt 7.2	TGGCTTTGCCTCATGAACG

**Le plasmide pMT85-pRS313-pMiniTetM-pSPEC sans lacZ-sbo** : Ce plasmide contient une origine de réplication *E. coli* et un marqueur de sélection *ampR*. C'est un plasmide intégratif chez *B. subtilis*. Deux cassettes de recombinaison correspondant à des *locus* du génome de *B. subtilis* encadrant l'opéron *sboAX-albAG* permettent l'intégration des éléments levures ARSH4, CEN6, ainsi que le marqueur d'auxotrophie pour l'histidine (HIS3) et du marqueur de sélection *specR*.

**Le plasmide pBlueScript KS(-) (pBSK-)** : Ce plasmide, commercialisé par Stratagene (Genbank #X52329), est un vecteur de clonage qui contient un multi-site de clonage (MCS) avec 21 sites uniques de reconnaissance pour des enzymes de restriction. Ce MCS est flanqué de part et d'autre par les promoteurs T3 et T7. Il contient également le marqueur de sélection *ampR*.

**Le plasmide pARS-VN** : Ce plasmide contient une origine de réplication *E. coli* (*ColE1*), ainsi qu'un marqueur de résistance *ampR*. Il contient les éléments levures ARSH4, CEN6 et le marqueur HIS3. Il possède aussi, sous le contrôle du promoteur pADH1 de *S. pombe*, le marqueur URA3 (figure 34). Il a été développé en 2003 par Noskov *et al.* (Noskov *et al.*, 2003)

**Le plasmide pUG36** : Ce plasmide contient une origine de réplication *E. coli* (*ColE1*), le marqueur de résistance *ampR* ainsi que les éléments levures ARSH4, CEN6, et le marqueur URA3. Il possède le gène *yeGFP* (version codon optimisée de la Green Fluorescent Protein (GFP) pour expression chez *S. cerevisiae*) sous le contrôle du promoteur MET17, un promoteur sensible à la méthionine (Cormack *et al.*, 1997).

**Le plasmide pBSB9** : C'est un plasmide d'expression chez *B. subtilis*, il contient une origine de réplication *E. coli* (*ColE1*), les marqueurs de résistance *ampR* et *cat* ainsi que le gène d'expression de la GFP.

### 2.2.2. Plasmides construits au cours du projet

**Les plasmides d'expression d'ARNg « simple target »** : Plusieurs plasmides permettant l'expression d'un ARNg spécifique ont été produits au cours de ces travaux. Ils sont listés dans le tableau 6.

**Tableau 7: Liste des plasmides ARNg double target produits**

Fragment capturé	ARNg double target	Simple target utilisés
Capture fragment 5	p426-gRNA-Ade2-double-target_Fgt5 G	p426-SNR52p-gRNA2.Frgt5-2,78 et p426-SNR52p-gRNA 1 Fgt4-2,64
	p426-gRNA-Ade2-double-target_Fgt5 H	p426-SNR52p-gRNA2.Frgt5-2,78 et p426-SNR52p-gRNA 2 Fgt4-2,64
	p426-gRNA-Ura3-double-target_Fgt5 C	p426-SNR52p-gRNA2.Frgt5-2,78 et p426-SNR52p-gRNA 1 Fgt4-2,64
	p426-gRNA-Ura3-double-target_Fgt5 D	p426-SNR52p-gRNA2.Frgt5-2,78 et p426-SNR52p-gRNA 2 Fgt4-2,64
Capture fragment 4+5	p426-gRNA-URA-pdt-fgt4+5-A	p426-SNR52p-gRNA2.Frgt5-2,78 et p426-SNR52p-gRNA1_Fgt4-2,5
	p426-gRNA-URA-pdt-fgt4+5-B	p426-SNR52p-gRNA2.Frgt5-2,78 et p426-SNR52p-gRNA 2 Fgt4-2,5
Capture fragment 3+4+5	p426-gRNA-URA-pdt-fgt 3-5_1	p426-SNR52p-gRNA2.Frgt5-2,78 et p426-SNR52p-gRNA 1 frgt 3
	p426-gRNA-URA-pdt-fgt 3-5_2	p426-SNR52p-gRNA2.Frgt5-2,78 et p426-SNR52p-gRNA 2 frgt 3
Capture fragment 2+3+4+5	p426-gRNA-URA-pdt-fgt 2-5_1	p426-SNR52p-gRNA2.Frgt5-2,78 et p426-SNR52p-gRNA 1 frgt 2
	p426-gRNA-URA-pdt-fgt 2-5_2	p426-SNR52p-gRNA2.Frgt5-2,78 et p426-SNR52p-gRNA 2 frgt 2
Capture d'un demi génome	p426-gRNA-URA-pdt-demi genome_1	p426-SNR52p-gRNA2.Frgt5-2,78 et p426-SNR52p-gRNA1 demi genome
	p426-gRNA-URA-pdt-demi genome_2	p426-SNR52p-gRNA2.Frgt5-2,78 et p426-SNR52p-gRNA2 demi genome

Pour construire ces plasmides, le vecteur p426-SNR52p-gRNA.Aarl.Y-SUP4t (Tsarmpopoulos *et al.*, 2016) a été digéré par l'enzyme de restriction Aarl (NEB) en suivant les instructions du fournisseur, puis des inserts de 20 pb nécessaires à la construction des ARNg ont été intégrés. Pour cela, des oligonucléotides complémentaires et spécifiques de la séquence cible ont été mélangés, phosphorylés 30 minutes à 37°C en présence de la T4 polynucleotide kinase (NEB) (PNK4) et hybridés grâce à un gradient de température (0.1°C par seconde jusqu'à 16°C) après dénaturation 5 min à 95°C. La réaction de ligature entre les duplex d'oligonucléotides (possédant après hybridation des extrémités compatibles « Aarl ») et le plasmide linéarisé a été réalisée à 16°C en présence de T4 DNA ligase (Promega). Le produit de ligature a été transformé dans *E. coli* DH5α par chimiocompétence. Les plasmides recombinants ont été extraits par mini-préparation, dosés, puis vérifiés par digestion avec l'enzyme de restriction Aarl, et par séquençage avec les amorces T3 et T7. Ils ont été stockés à -20°C. Les oligonucléotides clonés dans le plasmide p426-SNR52p-gRNA.Aarl.Y-SUP4t ont été sélectionnés en utilisant Benchling (<https://www.benchling.com/>). Cet outil permet de lister au sein d'un locus donné, toutes les séquences de 20 nucléotides présentent en amont d'une séquence PAM (NGG) et de les classer grâce à un score. Le score reflète l'efficacité supposée de l'ARNg lors d'expériences utilisant l'outil CRISPR-Cas9. Les séquences ayant des scores « ON target » et « OFF target » proches de 100, sont prédites comme étant les plus efficaces.

**Les plasmides d'expression d'ARNg « double target » :** Plusieurs plasmides permettant l'expression simultanée de deux ARNg spécifiques ciblant deux séquences différentes ont été produits au cours de ces travaux et sont listés dans le tableau 7. Les différentes versions des **p426-gRNA double target** ont toutes été construites de la même façon :

- La cassette « insert » contient un gène codant une ARNg chimérique, composé d'une séquence de 20 bases, spécifique à la région cible (crARN ou « spacer ») fusionnée avec un composant structural (tracrARN). L'expression de l'ARNg est constitutive sous le contrôle du promoteur SNR52, suivi de deux terminateur le SUP4 terminator et le CYC1 terminator. Elle a été produite par PCR à partir d'un plasmide matrice p426-SNR52p-gRNA1.Y-SUP4t avec les amorces p426\_insertGibson\_F et p426\_insertGibson\_R (tableau S1).
- La cassette « vecteur » correspond à l'ensemble d'un plasmide p426-SNR52p-gRNA.Y-SUP4t. Elle a été produite par PCR avec les amorces Div\_p426\_vecteurGibson\_F et

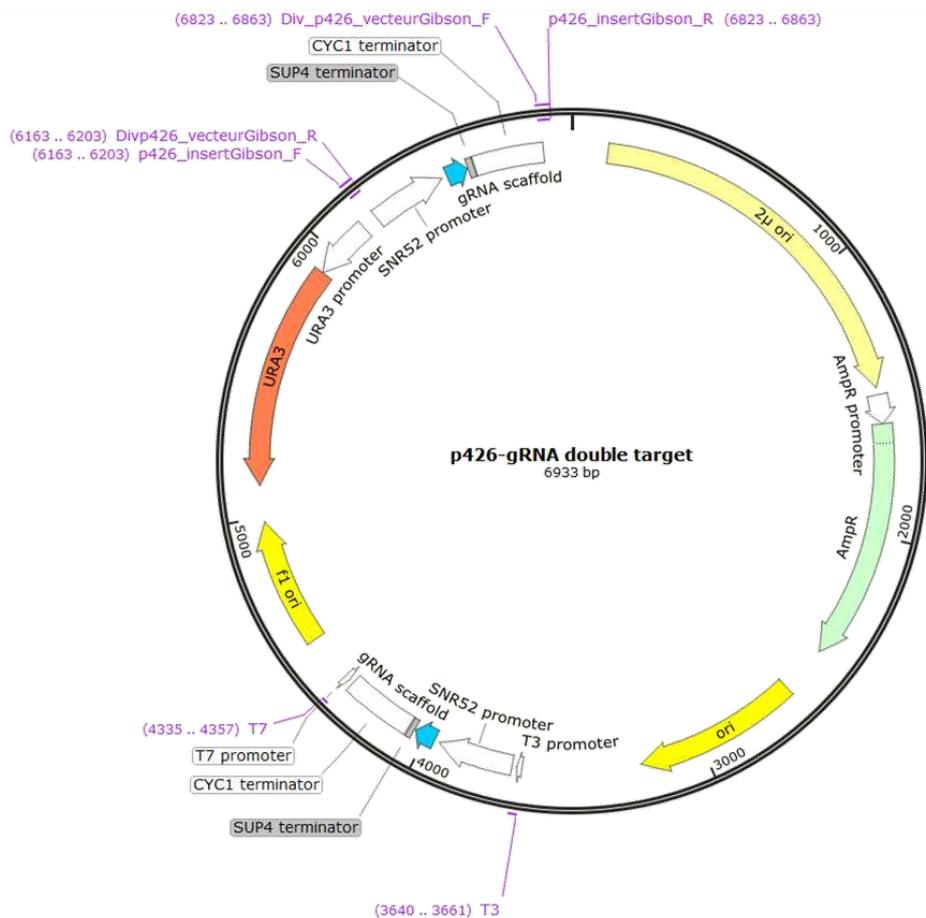


Figure 35: Carte du plasmide p426-gRNA « double target »

Divp426\_vecteurGibson\_R (tableau S1) à partir d'un plasmide matrice p426-SNR52p-gRNA2.Y-SUP4t permettant l'expression d'un second ARNg différent du premier.

Les deux cassettes d'ADN ont été assemblées par Gibson Assembly. Le produit réactionnel a ensuite été transformé dans les cellules *E. coli* DH5  $\alpha$ . L'ADN plasmidique a été extrait des transformants par mini-préparation (kit Macherey-Nagel, NucleoSpin® Plasmid EasyPure) puis vérifié par digestion enzymatique puis par PCR et séquençage au niveau des séquences nécessaires à l'expression des ARNg (figure 35).

**Le plasmide pBSK-pCRE-Bsu** : Le plasmide pBSK-pCRE-Bsu a été construit par assemblage dans *E. coli* du plasmide pBSK- linéarisé et d'une cassette contenant le gène *cre* obtenu à partir du plasmide pDR244 CRE + Ts origin. Les plasmides pBSK- et pDR244 ont été digérés par l'enzyme de restriction XbaI (NEB) en suivant les instructions du fournisseur. Ainsi le vecteur a été linéarisé et l'insert (le gène *cre*) a été isolé. La réaction de ligature entre l'insert et le plasmide linéarisé (possédant après digestion des extrémités compatibles « XbaI ») a été réalisée à 16°C en présence de T4 DNA ligase (Promega). Le produit de ligature a été transformé dans *E. coli* DH5 $\alpha$  par choc thermique. Les plasmides recombinants ont été extraits par mini-préparation puis vérifiés par digestion avec les enzymes de restriction XbaI, EcoRI et PstI. Ce plasmide a été utilisé comme matrice pour amplifier le gène *cre* utilisé pour la construction du plasmide pCC1-YLYS-CRE-SpecR.

**Le plasmide pCC1-YLYS-CRE-SpecR** : Ce plasmide a été construit par Gibson Assembly en assemblant 4 cassettes : pCC1, LYS, CRE et SpecR.

- La cassette pCC1 porte les éléments nécessaires au maintien et à la réplication chez *E. coli*, mais également les éléments ARSH4/CEN6 pour le maintien et la réplication dans la levure. Elle a été produite par PCR à partir du plasmide matrice pCC1 YTRP avec les amorces Gib\_loxm2/71\_pCC1\_F et Gib\_Ade2\_pCC1\_R. L'amorce sens Gib\_loxm2/71\_pCC1\_F porte une queue flottante permettant d'introduire le site loxm2/71 (34 b) (tableau S1).
- La cassette LYS porte le marqueur LYS2. Elle a été produite par PCR en utilisant comme matrice le plasmide pICE LYS2 avec les amorces Gib\_pCC1\_Ade2\_F et Gib\_CreT7\_Ade2\_R\_V2 (tableau S1).

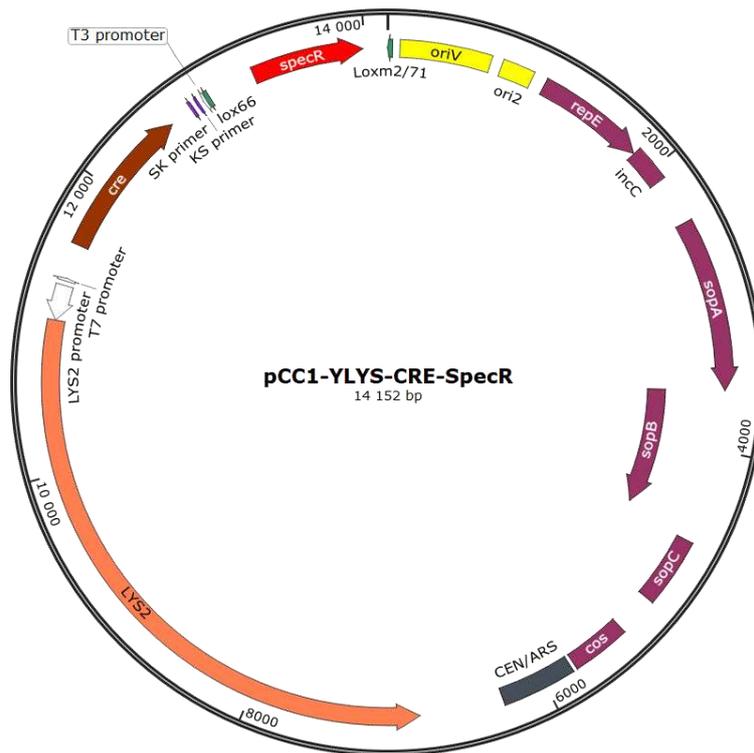


Figure 36: Carte du plasmide pCC1-YLYS-CRE-SpecR

- La cassette CRE porte le gène *cre* sous le contrôle d'un promoteur spécifique à *B. subtilis*. Elle a été produite par PCR en utilisant comme matrice le plasmide pBSK-pCRE-Bsu avec les amorces Gib\_Ade2\_CreT7\_F\_V2 et Gib\_lox66-CreT3\_R (tableau S1).
- La cassette SpecR contient le gène *specR*. Elle a été produite par PCR à partir du plasmide matrice pDR111 et des amorces Gib\_lox66\_Spec\_F et Gib\_loxm2/71\_Spec\_R portant des queues flottantes permettant d'introduire respectivement les sites lox66 et loxm2/71 (tableau S1).

Le kit Q5<sup>®</sup> High-Fidelity DNA Polymerase (NEB) a été utilisé pour l'amplification PCR des 4 cassettes, en suivant les recommandations du fournisseur. Elles ont été assemblées par Gibson Assembly. Le produit réactionnel a ensuite été transformé dans les cellules électro-compétentes *E. coli* EPI 300 Transfomax. L'ADN plasmidique a été extrait des transformants par mini-préparation puis vérifié par PCR (figure 36).

**Le plasmide pTwist CmR loxm2/66 V2** : Ce plasmide est intégratif chez *B. subtilis*, il a été construit par Gibson Assembly en assemblant 2 cassettes d'ADN :

- La cassette pCmR contient le gène *cat* ainsi que son promoteur conférant la résistance au chloramphénicol. Elle a été produite à partir du pBSB9 avec les amorces Gib pCmR-pTwist et Gib pCmR (fin CmR)-pTwist-N (tableau S1).
- La cassette ptwist contient deux zone de recombinaison d'une taille de 1 kb correspondant aux gènes *trmE* et *trmF* ainsi qu'une origine de répllication *E. coli* (*ColE1*) et un marqueur de sélection *ampR*. Elle a été produite à partir du pTwist CmR loxm2/66 avec les amorces Gib pTwist-pCmR et Gib pTwist-pCmR (fin CmR)-N (tableau S1).

Le kit Q5<sup>®</sup> High-Fidelity DNA Polymerase (NEB) a été utilisé pour l'amplification PCR des 2 cassettes, en suivant les recommandations du fournisseur. Elles ont été assemblées par Gibson Assembly. Le produit réactionnel a ensuite été transformé dans les cellules *E. coli* DH5  $\alpha$ . L'ADN plasmidique a été extrait des transformants par mini-préparation puis vérifié par digestion enzymatique.

**Le plasmide pTwist CmR loxm2/66 V3** : Ce plasmide est intégratif chez *B. subtilis*, il a été construit par Gibson Assembly en assemblant 3 cassettes d'ADN:



- La cassette pCmR est identique à celle utilisée pour la construction du plasmide **pTwist CmR loxm2/66 V2**. Cependant elle a été obtenue avec des amorces différentes, Gib pCmR-pTwist et Gib pCmR-promoteur TrmF-N (tableau S1).
- La cassette pTrmF contient le promoteur du gène *trmE*. Elle a été produite à partir de l'ADN génomique (ADNg) de la souche MS avec les amorces Gib promoteur TrmF-pCmR-N et Gib promoteur TrmF-pTwist (tableau S1).
- La cassette ptwist est identique à celle utilisée pour la construction du plasmide **pTwist CmR loxm2/66 V2**. Cependant elle a été obtenue avec des amorces différentes GibpTwist-promoteur TrmF et Gib pTwist-pCmR (tableau S1).

Le kit Q5® High-Fidelity DNA Polymerase (NEB) a été utilisé pour l'amplification PCR des 3 cassettes, en suivant les recommandations du fournisseur. Elles ont été assemblées par Gibson Assembly. Le produit réactionnel a ensuite été transformé dans les cellules *E. coli* DH5  $\alpha$ . L'ADN plasmidique a été extrait des transformants par mini-préparation puis vérifié par digestion enzymatique.

**Le plasmide ptwist loxm2/66 TetR lox 71** : Ce plasmide correspond au ptwist TetR lox71 avec un site de recombinaison supplémentaire, le site loxm2/66. Il est intégratif chez *B. subtilis*, et permet d'insérer le gène de résistance *tetR* bordé de deux sites de recombinaison, loxm2/66 et lox71. Il a été construit à partir du plasmide matrice ptwist TetR lox 71 par KLD (Kinase, Ligase, DpnI) en utilisant le kit Q5 Site-Directed Mutagenesis (NEB). Les amorces KLD loxm2/66 tetR lox71 F et KLD loxm2/66 TetR lox71 R (tableau S1) ont été utilisées pour amplifier par PCR tout le plasmide matrice, de plus, elles contiennent des queues flottantes correspondant au site loxm2/66 permettant de le rajouter au produit de PCR. Ce dernier a ensuite subi un traitement par les enzymes KLD (Kinase, Ligase, DpnI). Le produit réactionnel a été transformé dans les cellules *E. coli* DH5  $\alpha$ . L'ADN plasmidique a été extrait des transformants par mini-préparation, puis vérifié par digestion enzymatique puis par séquençage au niveau du site de recombinaison loxm2/66.

**Le plasmide pDR244  $\Delta$ CRE** : Ce plasmide correspond au plasmide pDR244 CRE + Ts origin sans le gène *cre*. Il a été construit à partir du plasmide matrice pDR244 *cre* + Ts origin par KLD en utilisant le kit Q5 Site-Directed Mutagenesis (NEB). Pour ce faire, les amorces KLD cre-R et KLD cre-F (tableau S1) ont été utilisées et pour amplifier tout le plasmide matrice à l'exception du gène *cre*. Le produit PCR a ensuite subi un traitement par les enzymes KLD (Kinase, Ligase,

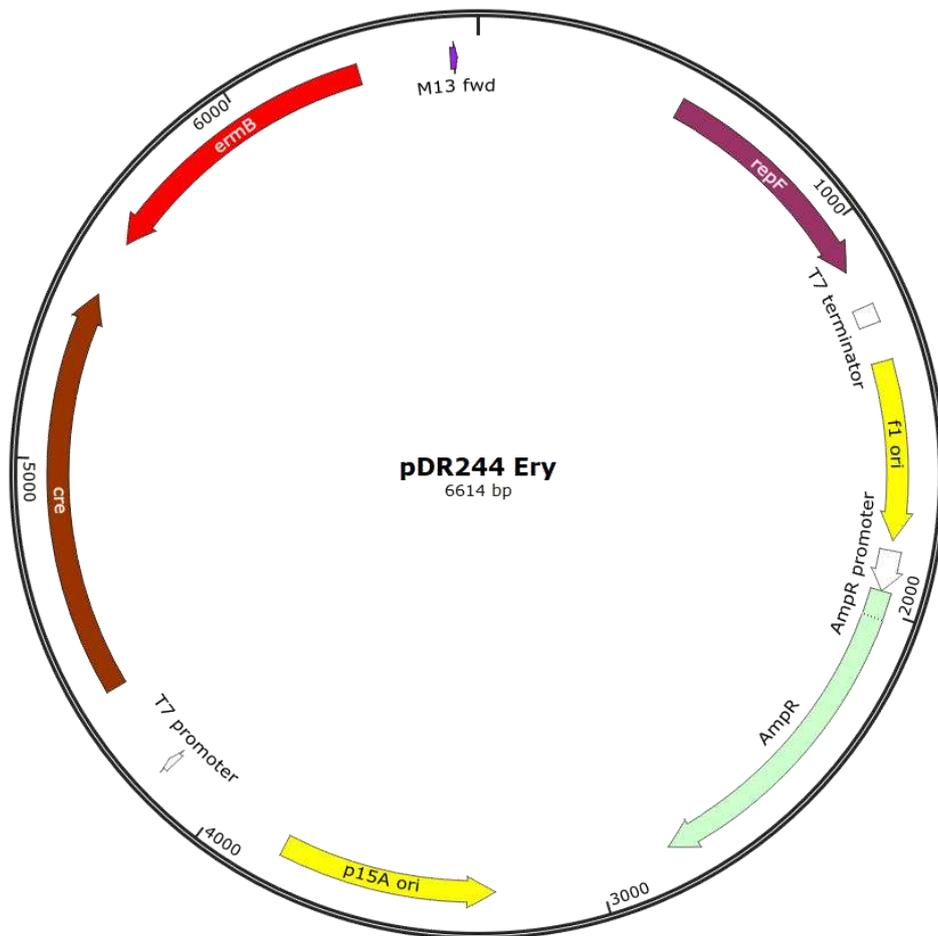


Figure 37: Carte du plasmide pDR244 Ery

DpnI). Le produit réactionnel a été transformé dans les cellules *E. coli* DH5  $\alpha$ . L'ADN plasmidique a été extrait des transformants par mini-préparation, puis vérifié par digestion enzymatique puis par séquençage complet du plasmide.

**Le plasmide pDR244 Ery** : Ce plasmide correspond au plasmide pDR244 CRE + Ts origin avec comme marqueur de résistance le gène *ermB* conférant la résistance à l'érythromycine à la place du marqueur de résistance *specR* conférant la résistance à la spectinomycine. Il a été construit par Gibson Assembly en assemblant deux cassettes :

- La cassette pDR244 contient toutes les séquences du pDR244 CRE + Ts origin excepté le marqueur de résistance *specR*. Elle a été produite par PCR à partir du plasmide matrice pDR244 CRE + Ts origin avec les amorces pDR244 R / Ery\_F2 et pDR244 F / Ery\_R2 (tableau S1).
- La cassette Ery contient le marqueur de résistance *ermB*. Elle a été produite par PCR à partir du plasmide matrice pJIM avec les amorces Ery\_R2 / pDR244 F et Ery\_F2 / pDR244 R (tableau S1).

Le kit Q5<sup>®</sup> High-Fidelity DNA Polymerase (NEB) a été utilisé pour l'amplification PCR en suivant les recommandations du fournisseur. Les deux cassettes ont été assemblées par Gibson Assembly. Le produit réactionnel a ensuite été transformé dans les cellules *E. coli* DH5  $\alpha$ . L'ADN plasmidique a été extrait des transformants par mini-préparation, puis vérifié par digestion enzymatique puis par PCR et séquençage au niveau du gène *cre* et du gène *ermB* (figure 37).

**Le plasmide pARS404 kana**: Ce plasmide est intégratif chez *B. subtilis*, il a été construit par Gibson Assembly en assemblant 4 cassettes : KanaR, ORI, FrIB et YurQ.

- La cassette KanaR contient le marqueur de résistance *kanaR* conférant la résistance à la kanamycine. Elle a été produite par PCR à partir de l'ADNg de la souche *B. subtilis* BKK 16940 (Koo *et al.*, 2017) avec les amorces ARS404 Kana/FrIB et kana/YurQ (tableau S1).
- La cassette ORI contient une origine de répllication *E. coli* (*ColE1*) et un marqueur de sélection *ampR*. Elle a été produite à partir du pDR111 avec les amorces pDR111/YurQ et ORI/FrIB (tableau S1).

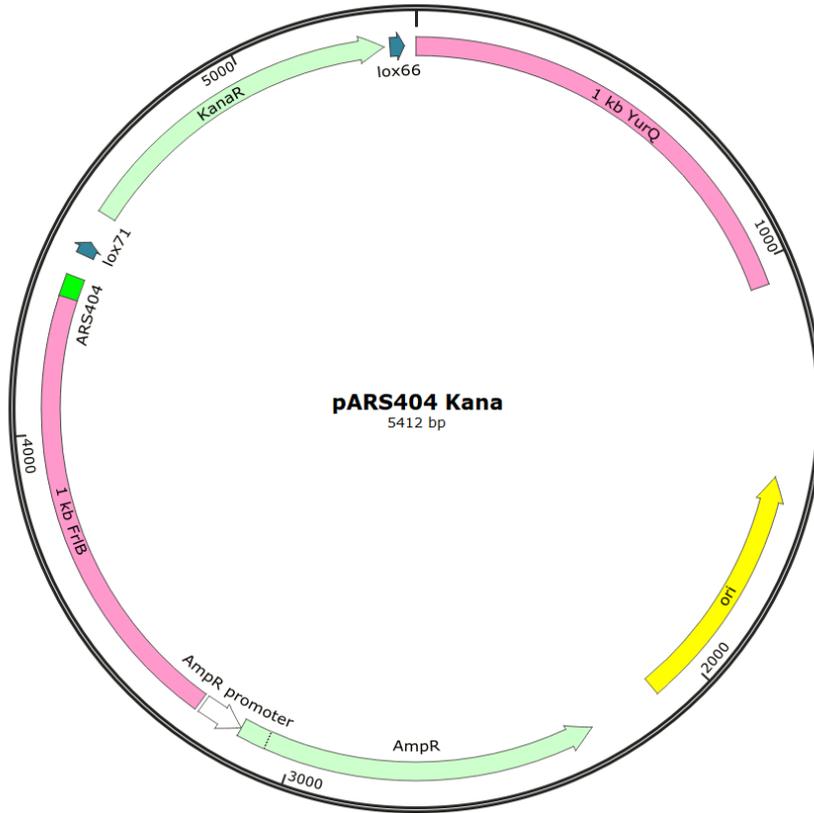


Figure 38: Carte du plasmide pARS404 Kana

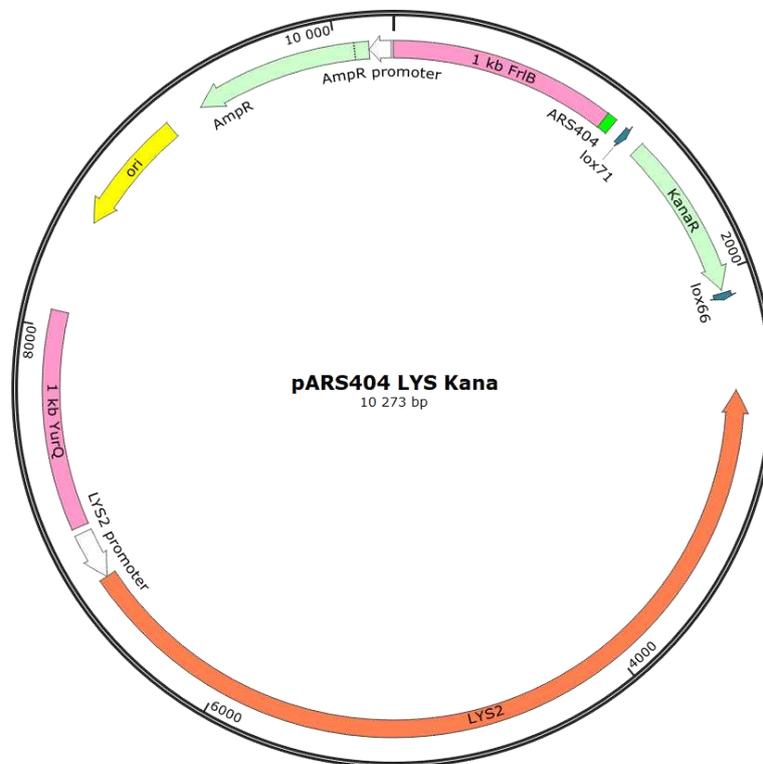


Figure 39: Carte du plasmide pARS404 LYS Kana

- La cassette FrlB correspond à une cassette de recombinaison d'une taille de 1 kb correspondant au gène *frlB*. Elle a été produite à partir de l'ADNg de la souche de *B. subtilis* MS avec les amorces FrlB/ORI et FrlB/ARS404 TetR (tableau S1).
- La cassette YurQ correspond à la deuxième cassette de recombinaison d'une taille de 1 kb correspondant au gène *yurQ*. Elle a été produite à partir de l'ADNg de la souche de *B. subtilis* MS avec les amorces YurQ/pDR111 et YurQ/Kana.

Les quatre cassettes ont été assemblées par Gibson Assembly. Le produit réactionnel a ensuite été transformé dans les cellules *E. coli* ElectroMAX™ Stbl4™. L'ADN plasmidique a été extrait des transformants par mini-préparation, puis vérifié par digestion enzymatique et par séquençage au niveau de l'ARS404 (figure 38).

**Le plasmide pARS404 LYS kana** : Ce plasmide est intégratif chez *B. subtilis*, il est presque identique au pARS404 kana. Il contient une cassette supplémentaire avec le marqueur LYS2. Il a été construit par Gibson Assembly en assemblant 5 cassettes : KanaR, ORI, FrlB, YurQ et LYS.

- Les cassettes ORI et FrlB sont identiques aux cassettes ORI et FrlB utilisées pour construire le plasmide pARS404 kana
- Les cassettes KanaR et YurQ sont également identiques aux cassettes KanaR et YurQ utilisées pour construire le plasmide pARS404 kana mais elles ont été produites avec des amorces différentes. La cassette KanaR a été produite avec les amorces ARS404 Kana/FrlB et Kana/LYS (tableau S1). La cassette YurQ a été produite avec les amorces YurQ/pDR111 et YurQ/LYS (tableau S1).
- La cassette LYS contient le marqueur LYS2 ainsi que son promoteur. Elle a été produite à partir du plasmide pICE LYS2 avec les amorces LYS/YurQ et LYS/Kana (tableau S1).

Les cinq cassettes ont été assemblées par Gibson Assembly. Le produit réactionnel a ensuite été transformé dans les cellules *E. coli* ElectroMAX™ Stbl4™. L'ADN plasmidique a été extrait des transformants par mini-préparation, puis vérifié par digestion enzymatique et par séquençage complet (figure 39).

**Le plasmide p414-TEF1-yeGFP** : Ce plasmide est centromérique chez la levure *S. cerevisiae*, il permet l'expression constitutive chez la levure, d'une version codon optimisée

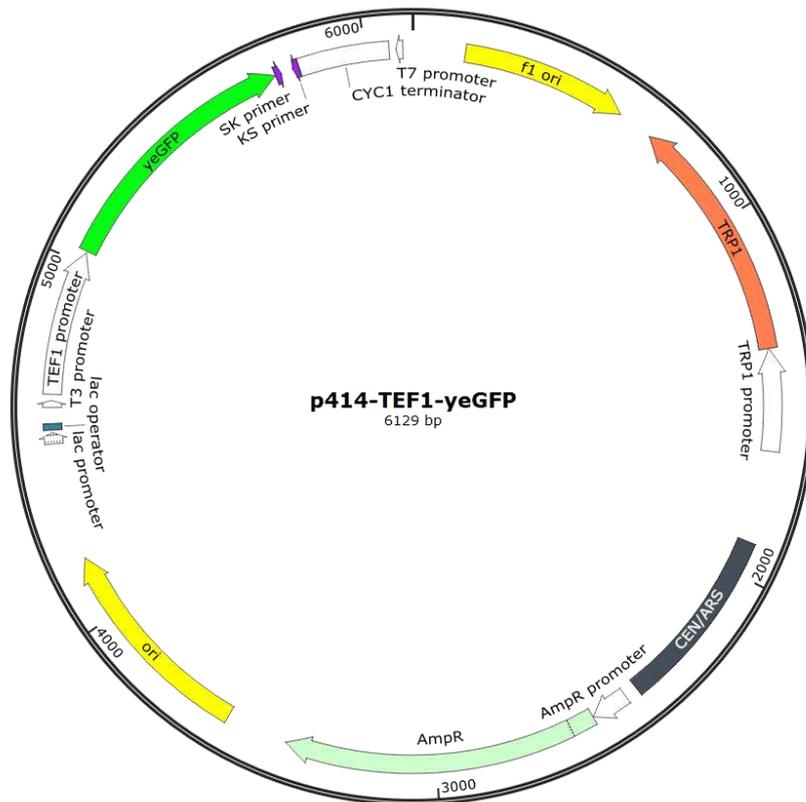


Figure 40: Carte du plasmide p414-TEF1-yeGFP

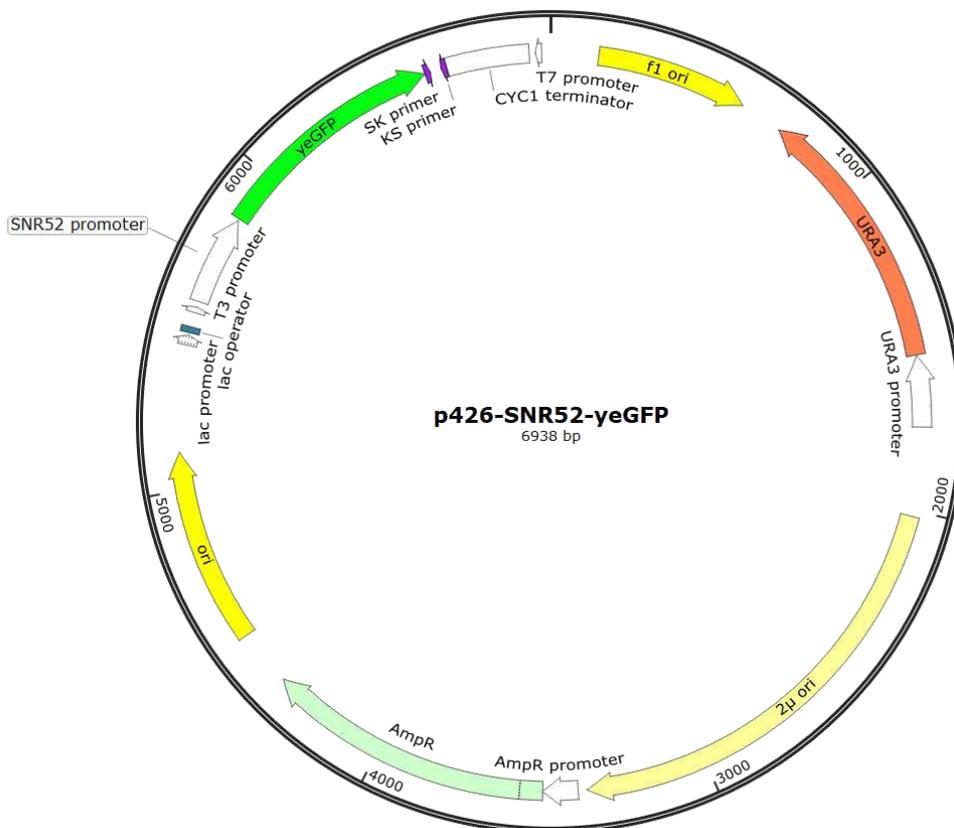


Figure 41: Carte du plasmide p426-SNR52-yeGFP

de la GFP sous le contrôle du promoteur constitutif TEF1P. Il dérive du p414-TEF1p-Cas9-CYC1t et a été construit par Gibson Assembly en assemblant 2 cassettes : yeGFP et p414-TEF1

- La cassette yeGFP contient une version codon optimisée de la GFP (la yeGFP) pour l'expression chez *S. cerevisiae*. Elle a été produite par PCR à partir du pUG36 avec les amorces GFP/vecteur et GFP/TEF1 (tableau S1).
- La cassette p414-TEF1 contient les éléments levure ARSH4, CEN6, le marqueur TRP1 et le promoteur constitutif TEF1P, ainsi qu'une origine de répllication *E. coli* (*ColE1*) et un marqueur de sélection *ampR*. Elle a été produite par PCR à partir du p414-TEF1p-Cas9-CYC1t avec les amorces vecteur/GFP et TEF1/GFP (tableau S1).

Les deux cassettes ont été assemblées par Gibson Assembly. Le produit réactionnel a ensuite été transformé dans les cellules *E. coli* DH5  $\alpha$ . L'ADN plasmidique a été extrait des transformants par mini-préparation puis vérifié par digestion enzymatique puis par séquençage au niveau du promoteur TEF1P et de la yeGFP (figure 40).

**Le plasmide p426-SNR52-yeGFP** : Ce plasmide permet l'expression constitutive chez la levure, d'une version codon optimisée de la GFP sous le contrôle du promoteur constitutif SNR52. Il dérive du p426-SNR52p-gRNA.Aar1.Y-SUP4t et a été construit par Gibson Assembly en assemblant 2 cassettes : yeGFP et p426-SNR52

- La cassette yeGFP est identique à celle décrite pour le p414-TEF1-yeGFP mais elle a été produite avec les amorces GFP/vecteur et GFP/SNR52 (tableau S1).
- La cassette p426-SNR52 contient une origine de répllication à grand nombre de copies 2 $\mu$ , le marqueur URA3 et le promoteur constitutif SNR52, ainsi qu'une origine de répllication *E. coli* (*ColE1*) et un marqueur de sélection *ampR*. Elle a été produite par PCR à partir du p426-SNR52p-gRNA.Aar1.Y-SUP4t avec les amorces SNR52/GFP et vecteur/GFP (tableau S1).

Les deux cassettes ont été assemblées par Gibson Assembly. Le produit réactionnel a ensuite été transformé dans les cellules *E. coli* DH5 $\alpha$ . L'ADN plasmidique a été extrait des transformants par mini-préparation, puis vérifié par digestion enzymatique et par séquençage au niveau du promoteur SNR52 et de la yeGFP (figure 41).



**Le plasmide pCRE Ery** : Ce plasmide permet d'intégrer le gène *cre* dans le génome de *B. subtilis* entre les gènes *friB* et *yurQ*. Il a été construit par Gibson Assembly en assemblant 2 cassettes : CRE/Ery et FriB/YurQ.

- La cassette CRE/Ery contient la séquence codant le gène *cre* ainsi que le marqueur de résistance *ermB*. Elle a été produite par PCR à partir du pDR244 Ery avec les amorces Ery/yurQ et cre/FriB (tableau S1).
- La cassette FriB/YurQ qu'une origine de répllication *E. coli* (*ColE1*) et un marqueur de sélection *ampR* ainsi que deux cassettes de recombinaison d'une taille de 1 kb chacune correspondant aux gènes *friB* et *yurQ*. Elle a été produite par PCR à partir du pARS404 LYS kana avec les amorces YurQ/Ery et FriB/cre (tableau S1).

Le kit Q5® High-Fidelity DNA Polymerase (NEB) a été utilisé pour l'amplification PCR en suivant les recommandations du fournisseur. Les deux cassettes ont été assemblées par Gibson Assembly. Le produit réactionnel a ensuite été transformé dans les cellules *E. coli* DH5  $\alpha$ . L'ADN plasmidique a été extrait des transformants par mini-préparation, puis vérifié par digestion enzymatique puis par séquençage au niveau des gènes *cre* et *ermB*.

**Le pJIM CRE** : Ce plasmide est répllicatif chez *B. subtilis* et permet l'expression de la CRE. Il a été construit par Gibson Assembly en assemblant 2 cassettes : jim et cre

- La cassette jim correspond au pJIM complet il contient donc une origine de répllication *E. coli* (*ColE1*), un marqueur de résistance *ermB* ainsi que les trois gènes *copF*, *repD* et *repE* qui permettent la répllication chez *B. subtilis*. Elle a été produite par PCR à partir du pJIM avec les amorces jim/cre et Ery\_F2 / pDR244 R (tableau S1).
- La cassette cre contient la séquence codant le gène *cre*. Elle a été produite par PCR à partir du pDR244 Ery avec les amorces cre/jim et pDR244 R / Ery\_F2 (tableau S1).

Le kit Q5® High-Fidelity DNA Polymerase (NEB) a été utilisé pour l'amplification PCR en suivant les recommandations du fournisseur. Les deux cassettes ont été assemblées par Gibson Assembly. Le produit réactionnel a ensuite été transformé dans les cellules *E. coli* DH5  $\alpha$ . L'ADN plasmidique a été extrait des transformants par mini-préparation, puis vérifié par digestion enzymatique puis par séquençage au niveau des gènes *cre*.



### 2.3. Gibson Assembly

Plusieurs plasmides ont été assemblés en utilisant le kit NEBuilder® Hifi DNA Assembly Master Mix de chez NEB. Ce kit est basé sur la méthode de Gibson Assembly (Gibson *et al.*, 2009; Gibson, 2011) qui permet d'assembler *in vitro* des fragments d'ADN chevauchants (sur ~40 pb) dans un ordre désiré en utilisant l'activité synchrone de trois enzymes : exonuclease 5'-3', ADN polymérase et ADN ligase.

- 1) L'exonucléase digère les séquences terminales 5' exposant ainsi des extrémités 3' libres.
- 2) Les fragments dont les extrémités sont complémentaires vont ensuite s'hybrider entre eux et la polymérase pourra synthétiser de l'ADN dans le sens 5' en 3' afin de combler les interstices des régions monocaténares.
- 3) Enfin, l'ADN ligase liera les fragments d'ADN ensemble ce qui permettra d'obtenir une seule molécule d'ADN double brin.

### 2.4. KLD

Plusieurs plasmides ont été modifiés à l'aide du kit Q5 Site-Directed Mutagenesis de chez NEB. Ce kit est efficace pour l'insertion, la délétion et la substitution de gène(s). La première étape est une amplification de fragment PCR avec le kit Q5® High-Fidelity DNA Polymerase. La seconde étape est basée sur l'utilisation d'un mélange d'enzymes Kinase, Ligase et Dnpi (KLD) qui permet une phosphorylation, une circularisation du produit PCR et l'élimination du plasmide matrice en une seule étape.

### 2.5. Mini-préparation

La mini-préparation de plasmides permet d'extraire une petite quantité (100 ng à 5 µg) d'ADN plasmidique provenant de bactéries, le plus souvent *E. coli*. Cette extraction est réalisée en utilisant le kit « NucleoSpin® Plasmid EasyPure » de chez Macherey-Nagel, en suivant les instructions du fournisseur. Pour chaque minipréparation, une colonie bactérienne est ensemencée dans 3 ml de milieu LB contenant l'antibiotique approprié. La culture est incubée toute la nuit à 37°C, sous agitation (200 rpm). Le lendemain, 1 ml de culture est prélevé afin



d'en extraire l'ADN plasmidique. L'ADN total est libéré par lyse alcaline des cellules, puis l'ADNg et les plasmides sont précipités différenciellement par neutralisation rapide du pH et séparés par centrifugation. Les plasmides restés dans la phase soluble sont adsorbés sur une membrane de silice, lavés avec une solution alcoolique, puis élués. Les plasmides purifiés sont dosés et conservés à -20°C.

## 2.6. Midi-préparation

La midi-préparation de plasmides permet d'extraire une plus grande quantité d'ADN plasmidique. Une colonie bactérienne estensemencée dans 100 ml de milieu LB contenant l'antibiotique approprié. La culture est incubée toute la nuit à 37°C, sous agitation (200 rpm). L'ADN plasmidique est extrait de la totalité de la culture, en utilisant le kit « NucleoBond® Xtra Midi » de chez Macherey-Nagel en suivant les instructions du fournisseur. Le principe de la purification est le même que celui du kit de mini-préparation décrit ci-dessus. Les plasmides purifiés sont dosés et conservés à -20°C.

## 2.7. Séquençage

### 2.7.1. Séquençage des plasmides

#### 2.7.1.1. Séquençage Sanger

Le séquençage Sanger a été effectué par la société prestataire GENEWIZ™ afin de contrôler les régions d'intérêt des plasmides produits et de valider qu'elles étaient dépourvues de mutations. Les résultats de séquençage ont été étudiés manuellement généralement par alignement avec une séquence théorique avec le logiciel Ape. La fiabilité des séquençages a été évaluée par l'analyse des chromatogrammes.

#### 2.7.1.2. Séquençage des plasmides complets

Le séquençage de plasmide complet a été réalisé par deux sociétés prestataires différentes GENEWIZ™ et PLASMIDSAURUS™ afin de s'assurer de l'intégrité et de la qualité des plasmides. Ces services de séquençage de plasmide complet sont basés sur des

**Tableau 8 : Liste des plasmides entièrement séquencés au cours de ces travaux.**

Plasmides entièrement séquencés
pDR244 CRE + Ts origin
pJIM
pMT85-pRS313-pMiniTetM-pSPEC sans lacZ-sbo
pBSB9
pDR244 $\Delta$ CRE
pARS404 LYS kana
pJIM CRE

technologies de séquençage long reads d'Oxford Nanopore et nécessitent des échantillons dilués à 50 ng/μl. Les plasmides séquencés entièrement au cours de ce travail sont listés dans le tableau 8.

### 2.7.2. Séquençage de génomes bactériens complets

Le séquençage de génomes bactériens complets a été réalisé par la société prestataire PLASMIDSAURUS™, afin de contrôler l'intégrité du génome bactérien de certaines souches et de s'assurer de l'absence de grands remaniements. Ce service est basé sur des technologies de séquençage long reads d'Oxford Nanopore et demande des échantillons dilués à 50 ng/μl. Afin de s'assurer de l'absence de Single Nucleotide Polymorphisms (SNP), du séquençage short reads basé sur des technologies d'Illumina a aussi pu être réalisé sur certaines souches.

## 2.8. Extraction d'ADN génomique de *B. subtilis*

L'ADNg des cellules de *B. subtilis* a été extrait selon trois protocoles différents en fonction des besoins.

- Extraction en suivant le protocole du kit Wizard® Genomic DNA Purification Kit de chez Promega®. A partir d'une culture en milieu liquide ensemencée la veille, 1 ml est prélevé et les cellules sont sédimentées par centrifugation pendant 2 min à 16000 g. Elles sont remises en suspension dans 480 μl d'EDTA 50 mM additionné de lysozyme à une concentration finale de 2,5 mg/ml puis incubées 1 h à 37°C. Les cellules sont ensuite centrifugées, remises en suspension dans 600 μl de « Nuclei Lysis Solution » puis placées au moins 5 min à 80°C pour permettre leur lyse. Après un traitement d'une heure à 37°C avec de la RNase à une concentration finale de 0,02 mg/ml, 200 μl de la « Protein Precipitation Solution » sont ajoutés et le mélange est placé sur glace pendant 5 min ce qui permet de précipiter les protéines. Les débris cellulaires et les protéines sont sédimentés par centrifugation pendant 3 min à 16000 g. L'ADNg toujours présent dans le surnageant après centrifugation, est prélevé puis lavé et précipité par ajout d'isopropanol. L'ADN est récupéré par centrifugation pendant 2 min à 16000 g et l'isopropanol est éliminé. Le culot subit un dernier lavage à l'éthanol 70%. Après séchage pour enlever complètement l'éthanol, le culot d'ADN est remis en

**Tableau 9: Liste et composition des solutions utilisées lors du protocole d'extraction d'ADN total de *S. cerevisiae***

Zymolyase/ME buffer (pour 1 extraction)	H <sub>2</sub> O	100 µl
	β-Mercaptoethanol (12.5 M SIGMA-98%- M3148)	1 µl
	Zymolyase buffer (solution stock : 10 mg/ml)	10 µl
Zymolyase buffer (10 mg/ml) (pour un volume final de 20 ml)	Tris-HCl (solution stock : 1 M) pH 7.5	1 ml
	Glycérol (solution stock : 50 %)	10 ml
	Zymolyase 100T	200 mg
	Qsp H <sub>2</sub> O	20 ml
	Conserver à -20°C	
Acétate de potassium 5 M	Acétate de potassium	29.4 g
	Acide acetic glacial dissous	11.5 ml
	Qsp H <sub>2</sub> O	75 ml

suspension dans 100  $\mu$ l de « DNA Rehydratation Solution » puis dosé par spectrophotométrie au Nanodrop. Il est finalement conservé à 4°C ou à -20°C.

- Extraction en suivant un protocole « rapide » (Brno and Mendel, 2020). A partir d'une culture en milieu liquide ensemencée la veille, 1 ml est prélevé en tube Eppendorf et les cellules sont collectées par centrifugation. Elles sont remises en suspension dans de l'eau stérile puis placées 10 min à 100°C. Cette étape permet de lyser les cellules. Les tubes sont refroidis 15 min à 4°C. Les débris cellulaires sont sédimentés par centrifugation. L'ADNg toujours présent dans le surnageant après centrifugation, est prélevé puis dosé par spectrophotométrie au Nanodrop. Il est finalement conservé à 4°C ou à -20°C. Ce protocole permet une extraction simple d'un grand nombre de transformants pour une analyse rapide.
- Extraction en suivant le protocole décrit dans le kit NucleoBond® AXG Columns de chez Macherey Nagel. Ce protocole est basé sur des colonnes échangeuses d'anions qui permettent d'isoler de l'ADNg pure et peu fragmenté comparativement à d'autres protocoles. Il peut être utilisé pour des analyses par séquençage NGS notamment pour du séquençage de génome complet. L'extraction est réalisée à partir de 1,5 ml de culture diluée à  $DO_{600nm} = 1$  en suivant les recommandations du fournisseur. L'ADNg obtenu est dosé au Qubit puis stocké à 4°C.

## 2.9. Extraction d'ADN total de *S. cerevisiae*

L'ADN total de levure est extrait en suivant le protocole décrit par K. Kouprina et V. Larionov (Kouprina and Larionov, 2008) . Une colonie isolée est piquée et striée sur milieux sélectif et incubée 48h à 30°C. A partir d'un sous clone issu de la strie, un « patch » est réalisé sur milieu sélectif et incubé 48h à 30°C. Les levures prélevées à partir du patch sont mises en suspension pendant 1h à 37°C dans 90  $\mu$ l de zymolyase/ME buffer (tableau 9). 10  $\mu$ l de SDS 2% sont ajoutés à la suspension pour une incubation de 15 min à 70°C. Puis 11  $\mu$ l d'une solution d'acétate de potassium 5 M (tableau 9) sont ajoutés et le mélange est placé 15 min dans la glace. Les débris cellulaires et les protéines sont éliminés grâce à une centrifugation de 10 min à 14000 g à 4°C. L'ADN total toujours présent dans le surnageant après centrifugation, est prélevé et précipité par l'ajout de 90  $\mu$ l d'isopropanol puis centrifugé dans les conditions décrites précédemment. Après séchage pour éliminer toute trace



d'isopropanol, le culot d'ADN est remis en suspension dans 50 µl de TAE 1X (Tris-Acétate 40 mM ; EDTA 1 mM ; pH 8.0), puis dosé par spectrophotométrie au Nanodrop et conservé à 4°C ou à -20°C.

### 2.10. Polymerase Chain Reaction (PCR) simplex

La PCR simplex permet d'amplifier un fragment d'ADN à partir d'un couple d'amorces spécifiques. Elle a été utilisée soit pour vérifier la présence d'un locus donné dans un échantillon, soit pour produire de grande quantité d'un fragment d'ADN. Différentes ADN polymérasés et kit correspondants ont été utilisés en fonction des besoins. Quel que soit le kit utilisé, les recommandations du fournisseur ont toujours été suivies.

- Le kit Q5<sup>®</sup> High-Fidelity PCR (NEB) est basé sur l'ADN polymérase Q5 processive, rapide et environ 280 fois plus fidèle que la Taq polymérase classique entraînant un taux d'erreurs très faible ( $5,3 \times 10^{-7}$ ) (Potapov and Ong, 2017). Ce kit a préférentiellement été utilisé pour de la production de fragments d'ADN ainsi que pour de la détection lors de l'analyse des transformants *B. subtilis*.
- Le kit Advantage 2 PCR (Takara) est basé sur le mélange de polymérasés Advantage<sup>®</sup> 2. Ce kit permet d'amplifier une large gamme d'ADN matrice y compris de l'ADNg complexe issu de transformants levures qui hébergent un fragment de génome de *B. subtilis*. Il a été utilisé préférentiellement pour de la détection lors de l'analyse de transformants levure.

### 2.11. PCR multiplex

La PCR multiplex permet d'amplifier simultanément plusieurs fragments d'ADN en utilisant un cocktail de plusieurs couples d'amorces généralement distribués de manière homogène tout au long de la molécule d'ADN à analyser. Elle a été utilisée pour vérifier l'intégrité de fragments de génome bactérien clonés dans la levure ou transformés dans *E. coli*. Le kit Qiagen<sup>®</sup> Multiplex PCR a été utilisé en suivant les recommandations du fournisseur.



## 2.12. Production des pansements de recombinaison

Les expériences de clonage de fragments de génome dans la levure par CReasPy-Fusion sont basées sur l'utilisation du système CRISPR-Cas9 et la capacité de recombinaison homologue de la levure. Afin de faciliter la réparation de la cassure double brin causée par le système CRISPR-Cas9 et de permettre la re-circularisation de la molécule d'ADN à cloner dans la levure, des matrices de recombinaison sont utilisées. Elles correspondent à des fragments d'ADN double brin contenant des séquences d'homologie avec la région cible et sont appelées pansements de recombinaisons. Les expériences d'édition à un *locus* spécifique de la molécule d'ADN contenue dans la levure nécessitent aussi l'utilisation de tels pansements.

### 2.12.1. Pansements de recombinaison pour l'édition chez la levure

Ces pansements contiennent des séquences spécifiques en fonction de la modification souhaitée. Ils sont produits par PCR à partir de différents plasmides matrices. Ils sont bordés de séquences identiques aux régions à éditer. Ces régions d'homologie sont ajoutées au pansement lors de la PCR en utilisant des amorces avec des queues flottantes. Le produit de PCR est contrôlé par électrophorèse, purifié, dosé puis stocké à -20°C.

### 2.12.2. Pansements de recombinaison aux jonctions de 90 pb

Les pansements de recombinaison de 90 pb ne contiennent aucun élément spécifique. Ils correspondent à deux séquences de 45 pb identiques aux extrémités de la molécule à capturer. Ils permettent la re-circularisation aux jonctions. Ils sont produits par hybridation de deux oligonucléotides complémentaires de 90 bases. Pour ce faire, 10 µl de chaque oligonucléotide à 100 µM sont mis en contact avec 2,5 µl de tampon 10X Advantage 2 PCR. Le mélange est placé à 95°C pendant 5 min, puis la diminution de la température de 0,1°C par sec jusqu'à 16°C permet l'hybridation des oligonucléotides. Le pansement de recombinaison de 90 pb ainsi produit peut être conservé à -20°C.



### 2.12.3. Pansement de recombinaison aux jonctions composé des éléments levure ARSH4 CEN6 HIS3

Le pansement contenant les éléments levures ARSH4 CEN6 HIS3 est produit par PCR à partir du plasmide matrice pARS-VN (figure 34). Il est bordé de séquences identiques aux extrémités de la région à capturer. Ces régions d'homologie sont ajoutées au pansement lors de la PCR en utilisant des amorces avec des queues flottantes. Le produit de PCR est contrôlé par électrophorèse, purifié, dosé puis stocké à -20°C.

### 2.13. Purification des ADN

Les ADN produits par réaction enzymatique, sont purifiés en utilisant différents kits en fonction des besoins. Ces kits fonctionnent de manière très similaire et permettent de purifier des fragments d'ADN allant de 25 pb à 20 kb. Ils sont basés sur l'utilisation d'une colonne avec une membrane de silice exclusive.

- Le kit de purification GeneJET PCR (ThermoScientific) a été principalement utilisé pour la purification des ADNs issus d'amplification PCR.
- Le kit GFX™ PCR DNA, and Gel Band Purification (CYTIVA) a été principalement utilisé pour isoler spécifiquement un fragment d'ADN d'une taille donnée après électrophorèse sur gel d'agarose.

### 2.14. Quantification des ADN

La concentration d'ADN est mesurée en utilisant deux systèmes différents en fonction des besoins.

- Le système Epoch associé à la plaque microvolume Take3 de chez Biotek est basé sur la mesure d'absorbance à une densité optique de 260 nm. La concentration d'ADN est calculée par l'interface logicielle Gen5 à partir de la densité optique mesurée sur 2 µl d'échantillon. Ce système a été utilisé en routine pour la mesure d'un grand nombre d'échantillon.
- Le système Qubit 4 Fluorometer est basé sur l'utilisation de colorants sélectifs émettant de la fluorescence uniquement lorsqu'ils sont liés à l'ADN. Ce système est



beaucoup plus précis que la mesure de l'absorbance UV classique qui peut surestimer la concentration d'ADN dans l'échantillon dû à des contaminants comme notamment les protéines. Il a été utilisé pour des mesures précises sur un petit nombre d'échantillons notamment pour la préparation d'échantillons pour du séquençage long reads.

## 2.15. Digestion des ADN par restriction enzymatique

Les digestions enzymatiques ont été réalisées en suivant les recommandations des fournisseurs. Brièvement, les réactions ont été effectuées dans un volume final de 20 µl à 50 µl, contenant le tampon approprié à 1X, 1 à 5 unités d'enzyme et 0,3 à 1 µg d'ADN à cliver. Le mélange est incubé au moins une heure à la température appropriée à l'enzyme de restriction choisie.

## 2.16. Électrophorèse sur gel d'agarose

Les différents ADN ont été analysés par électrophorèse horizontale sur gel d'agarose (0,8 à 2,5%) dans du tampon TAE 1X. Cinq volumes de solution d'ADN ont été mélangés à un volume de tampon de charge 6X (Gel Loading Dye Purple de chez NEB) et déposés dans chaque puit. Pour définir la taille des ADN analysés, des marqueurs de taille ont été utilisés : (i) l'échelle d'ADN Invitrogen 1 Kb Plus et (ii) l'échelle d'ADN 100 pb de chez Promega ou NEB. L'électrophorèse a été réalisée à 100 volts. Après migration, les acides nucléiques ont été colorés par incubation des gels dans un bain de bromure d'éthidium (BET) à une concentration de 2 µg/ml. Après coloration, les fragments d'ADN ont été visualisés par fluorescence, avec le système d'imagerie Vilbert Lourmat E-BOX VX2.

## 2.17. Génomes en bloc d'agarose

### 2.17.1. Préparation des génomes de *B. subtilis* en bloc d'agarose

L'extraction des génomes de *B. subtilis* en bloc d'agarose a été réalisée en utilisant un protocole développé précédemment au laboratoire, basé sur l'utilisation du kit CHEF Bacterial

**Tableau 10: Liste et composition des solutions utilisées pour la préparation des génomes en blocs d'agarose**

Cell Suspension Buffer (Kit)			
Lyticase Buffer (pour un volume final de 500 ml)	Tris-HCl (solution stock : 1M) pH 7.5	5 ml	
	EDTA (solution stock : 500 mM) pH 8.0	50 ml	
	Qsp H <sub>2</sub> O	500 ml	
Solution de lysozyme (Kit)	Lysozyme Buffer (kit)		
	Lysozyme	1 mg/ml	
Solution de protéinase K (Kit ou maison)	Protéinase K Buffer (Kit ou maison)	EDTA, pH 8.0	100 mM
		Sodium deoxycholate	0.2%
		Sodium lauryl sarcosine pH 8.0	1%
	Protéinase K	1 mg/ml	
Tampon de lavage 1X (pour un volume final de 1 l)	Tris (solution stock : 1 M) pH 7.5	20 ml	
	EDTA (solution stock : 0,5 M) pH 8.0	100 ml	
	Qsp H <sub>2</sub> O	1 l	

Genomic DNA Plug de chez Bio-Rad. Brièvement, les cellules de *B. subtilis* sont mises en culture à 37°C dans 50 ml de LB complétement de streptomycine à 0,75 µg/ml finale. Lorsque la culture atteint une DO<sub>600nm</sub> comprise entre 1,5 et 1,6, du chloramphénicol est ajouté à une concentration finale de 180 µg/ml puis la culture est incubée 55 min à 37°C. Deux ou cinq millilitres de culture sont prélevés dans le but de préparer deux séries de 5 à 6 blocs d'agarose possédant des concentrations en cellules (et donc en génomes) différentes. Après centrifugation, les cellules sont remises en suspension dans 300 µl de Cell Suspension Buffer (tableau 10) puis placées 10 min à 50°C. La suspension est ensuite mélangée à 300 µl de Low Melting Agarose 2% en TAE 1X maintenu en surfusion à 50°C (volume total ~600µl), et coulée dans les puits de la barrette fournie spécialement par le kit Biorad (100µl/puit). Les blocs sont laissés 20 min à température ambiante pour la solidification. Ils sont ensuite démoulés et incubés 2 h à 37°C dans une solution de lysozyme (Lysozyme Buffer + lysozyme à une concentration finale de 1 mg/ml) (tableau 10). Les blocs sont lavés dans 25 ml d'eau stérile 2 fois pendant 5 min sous agitation douce (35 rpm). Après lavages, ils sont traités dans une solution contenant des détergents et de la protéinase K (tableau 10) (kit CHEF Bacterial Genomic DNA Plug de chez Bio-Rad) pendant 48h à 50°C : la solution est changée après 24h. Enfin les blocs sont lavés dans 10 ml de tampon de lavage 1X, 4 fois pendant 1 h sous agitation douce (35 rpm), puis stockés à 4°C.

### 2.17.2. Préparation des génomes de *S. cerevisiae* en bloc d'agarose

Les blocs d'agarose des levures qui hébergent un fragment de génome de *B. subtilis* ont été préparés en utilisant le kit CHEF Mammalian Genomic DNA Plug, de chez Bio-Rad.

En fonction du nombre de blocs à préparer pour une culture de levure, les volumes peuvent être ajustés, ici est décrit le protocole pour la préparation de 3 blocs. Les cellules de *S. cerevisiae* sont mises en culture à 30°C dans 50 ml de milieu minimal SD complétement du « DO supplement » nécessaire. Lorsque la culture atteint une DO<sub>600nm</sub> comprise entre 3 et 4, les cellules sont sédimentées par centrifugation et remises en suspension dans 5 ml d'EDTA 50mM, pH 8,0 froid. Un volume correspondant à  $6 \times 10^9$  cellules est prélevé puis centrifugé 10 min à 5000 rpm et à 4°C. Afin de définir le volume de culture à prélever la concentration cellulaire est définie par dénombrement sur cellule de Malassez. La mesure de DO<sub>600nm</sub> initiale



est aussi prise en compte. Le culot cellulaire est repris dans 150  $\mu$ l de Cell Suspension Buffer (tableau 10). Cette suspension est incubée à 50°C pour 10 min puis 1,5 mg de zymolyase 100T y sont ajoutés. La suspension est ensuite mélangée volume à volume à du Low Melting Agarose 2% en TAE 1X maintenu en surfusion à 50°C, et coulée dans les puits de la barrette fournie spécialement par le kit Biorad (100 $\mu$ l/puit). Trois blocs peuvent être coulés avec les conditions décrites. Les blocs sont laissés 20 min à 4°C pour la solidification. Ils sont ensuite démoulés et incubés 2 h à 37°C dans du Lyticase Buffer pH 7,5 (tableau 10) contenant de la zymolyase 100T à une concentration finale de 5 mg/ml. Les blocs sont lavés dans 25 ml d'eau stérile 2 fois pendant 5 min sous agitation douce (35 rpm). Après lavages, ils sont traités dans une solution contenant des détergents et de la protéinase K (tableau 10) (kit CHEF Mammalian Genomic DNA Plug, de chez Bio-Rad) pendant 48h à 50°C. La solution de protéinase K est changée après 24h. Enfin les blocs sont lavés dans 10 ml de tampon de lavage 1X, 4 fois pendant 1 h sous agitation douce (35 rpm), puis stockés à 4°C.

### 2.17.3. Digestion des blocs d'agarose par restriction enzymatique

L'ADN contenu dans les blocs d'agarose peut être digéré par des enzymes de restriction. Cette digestion se fait sur la moitié d'un bloc (50  $\mu$ l). Chaque demi bloc est lavé pendant 1 h sous agitation douce dans 1 ml de tampon de lavage 1X contenant du fluorure de phénylméthylsulfonyl (PMSF) à une concentration finale de 1 mM. Un deuxième lavage avec le tampon lavage 1X seul est réalisé dans les mêmes conditions. Les demi blocs sont ensuite lavés deux fois 1 h dans 1 ml de tampon de lavage 0,1X (Tris-HCl 2 mM pH 7,5 ; EDTA 5 mM pH 8,0) sous agitation douce. Ils sont ensuite incubés 1 h dans 0,5 ml du tampon adéquat pour l'enzyme de restriction utilisée. Enfin ils sont digérés toute la nuit à la température optimale de l'enzyme choisie (pour un demi bloc : 30-50 unités d'enzyme, 50  $\mu$ l du tampon (1X) approprié et de l'eau en quantité suffisante pour 0,5ml total).

### 2.18. Electrophorèse en champs pulsé (PFGE)

L'électrophorèse en champ pulsé est une méthode qui permet de faire migrer sur un gel d'agarose des molécules d'ADN de grande taille (> 50 kb) en appliquant un champ électrique homogène dont l'orientation alterne (Gilbert Chu, Douglas Vollrath, 1985). Elle a été utilisée



pour estimer la taille des fragments de génome de *B. subtilis* clonés dans la levure *S. cerevisiae* afin de s'assurer de leur intégrité. Elle a aussi été utilisée pour analyser la taille des génomes de *B. subtilis* modifiés à grande échelle.

#### 2.18.1. Analyse des transformants de *S. cerevisiae* qui propagent des fragments de génome de *B. subtilis*

L'analyse par électrophorèse en champ pulsé des demi blocs d'agarose contenant l'ADN des levures transformées avec des fragments de génome de *B. subtilis* est réalisée en suivant un protocole sur 6 jours nécessitant trois migrations afin de pouvoir visualiser proprement les fragments de génome clonés. Il est découpé en deux grandes étapes, la première consiste à éliminer les chromosomes de la levure et la deuxième consiste à digérer les fragments de génome de *B. subtilis* et à les faire migrer.

- Les demi blocs sont lavés dans 1 ml de tampon de lavage. Puis ils sont chargés sur un gel d'agarose Pulsed Field Certified 1 % de chez Bio-Rad, en TAE 1X. La migration est réalisée avec le système d'électrophorèse en champ pulsé CHEF-DR III de chez Bio-Rad pendant 20 h dans 1,9 l de TAE 1X à 14°C. Le temps d'impulsion est augmenté de 50 s à 90 s, avec un angle de 120° et un voltage de 6 V/cm. Pour définir la taille des ADNs, le marqueur de taille utilisé est le CHEF DNA Size Marker *S. cerevisiae* Ladder de chez Bio-Rad. Après migration, les blocs sont retirés du gel et les acides nucléiques sont colorés par incubation du gel dans un bain de SYBR<sup>TM</sup> Gold pendant 40 min. Après coloration, les fragments d'ADN sont visualisés par fluorescence, avec le système d'imagerie Vilbert Lourmat E-BOX VX2. Les demi blocs sont ensuite digérés avec 50 unités de la DNase ATP-dépendante Plasmid-Safe<sup>TM</sup> de chez Lucigen. Après digestion, ils sont lavés 3 fois dans 1 ml de tampon de lavage 1X pendant 30 min sous agitation douce. Puis ils subissent une deuxième migration dans les mêmes conditions que décrit précédemment. Après migration, les blocs sont retirés. A cette étape les chromosomes de la levure ont été éliminés.
- Les demi blocs sont digérés par l'enzyme de restriction souhaitée en fonction de la taille du fragment de *B. subtilis* cloné ainsi que du profil de migration voulu. Après digestion ils sont lavés 3 fois dans 1 ml de tampon de lavage 1X pendant 30 min sous

**Tableau 11: Liste des fragments de génome capturés dans la levure et paramètres pour l'analyse en électrophorèse en champs pulsé (PFGE)**

Fragment de génome	Taille	Enzyme utilisée	Profil de digestion	Paramètres	Marqueur de taille
Fragment 5 de la souche MGP192	125 kb	AvrII	2 bandes : 80 kb et 45 kb	20 h, switch time 1-25, angle 120° et voltage 6 V/cm en TBE 0,5X	<i>S. cerevisiae</i> Ladder de chez bio rad
Fragment 5 de la souche MS	145 kb		2 bandes : 100 kb et 45 kb		
Fragment 5 de la souche FE220	145 kb		2 bandes : 100 kb et 45 kb		
Fragment 4+5 de la souche MGP192	268 kb	FseI	Linéarisation : 268 kb		
Fragment 3+4+5 de la souche AG686	536 kb		3 bandes : 246 kb, 204 kb et 85 kb		
Fragment 2+3+4+5 de la souche AG686	925 kb	SgrDI	4 bandes : 536 kb, 173 kb, 141 kb, 73 kb		
Demi génome de la souche AG686	1,434 Mb		5 bandes : 655 kb, 390 kb, 173 kb, 141 kb, 73 kb		

agitation douce (35 rpm). Puis ils subissent une dernière migration dont les paramètres et le choix du marqueur de taille sont ajustés en fonction du fragment de *B. subtilis* cloné (tableau 11).

### 2.18.2. Analyse des transformants de *B. subtilis*

L'analyse par électrophorèse en champ pulsé des transformants de *B. subtilis* est réalisée sur des demi blocs de *B. subtilis*. Ils sont digérés par l'enzyme de restriction souhaitée en fonction du nombre de site de restriction sur le génome de la souche de *B. subtilis* présente dans le bloc ainsi que du profil de migration voulu. Après digestion, ils sont lavés dans 1 ml de tampon de lavage 1X pendant 1 h sous agitation douce (35 rpm). Puis ils sont chargés sur un gel d'agarose Pulsed Field Certified 1 % de chez Bio-Rad, en TAE 1X. La migration est réalisée dans les mêmes conditions que décrit partie 2.16.1. Après migration les acides nucléiques sont colorés par incubation du gel dans un bain de SYBR<sup>TM</sup> Gold pendant 40 min. Enfin les fragments d'ADN sont visualisés par fluorescence, avec le système d'imagerie Vilbert Lourmat E-BOX VX2.

## 3. Transformation

### 3.1. Transformation d'*E. coli*

Les transformations des souches commerciales ont été réalisées en suivant les recommandations des fournisseurs.

#### 3.1.1. Transformation par choc thermique

Les cellules d'*E. coli* stockées à -80°C sont décongelées sur glace. 20 µl de cellules sont mis en contact avec l'ADN à transformer. Le mélange est incubé 30 min sur glace puis un choc thermique est réalisé au bain marie à 42°C pendant 30 sec. Les cellules sont reprises dans 500 ml de milieu SOC (NEB) puis transférées dans un tube de 13 ml et incubées sous agitation (200 rpm) 1 h à 37°C. Un volume de 250 µl (non dilué ou dilué à 10<sup>-1</sup>) des bactéries est étalé sur milieu LB solide sélectif, en fonction de l'ADN transformé, les transformants sont sélectionnés comme décrit partie 1.1. Enfin les boîtes sont placées à 37°C pendant une nuit.

**Tableau 12 : Liste et composition des solutions utilisées lors de de la transformation de *B. subtilis* avec le protocole MG1/MG2**

MM 10X (pour un volume final de 100 ml)	Potassium Phosphate Dibasique (K <sub>2</sub> HPO <sub>4</sub> )	10,7 g	
	Potassium Phosphate Monobasique (KH <sub>2</sub> PO <sub>4</sub> )	6 g	
	Sulfate d'ammonium ((NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> )	2 g	
	Tri Sodium Citrate dihydraté Na <sub>3</sub> C <sub>6</sub> H <sub>5</sub> O <sub>7</sub>	1 g	
	Qsp H <sub>2</sub> O	100 ml	
	Stérilisé par autoclavage		
MG0 (pour un volume final de 100 ml)	MM 10X	10 ml	
	Glucose (solution stock : 20%)	2,5 ml	
	Sulfate de Magnésium (MgSO <sub>4</sub> ) (solution stock : 1 M)	160 µl	
	Qsp H <sub>2</sub> O	100 ml	
MG1	MG0	50 ml	
	Hydrolysate de caséine (solution stock : 5%)	250 µl	
	Extrait de levure (solution stock : 5%)	1 ml	
	Tryptophane (solution stock : 0,5%) (si besoin en fonction de la souche transformée)	2 ml	
	Stérilisé par filtration à 0,22 µm et conservé à -20°C		
MG2	MG0	50 ml	
	Hydrolysate de caséine (solution stock : 5%)	125 µl	
	Extrait de levure (solution stock : 5%)	250 µl	
	Sulfate de Magnésium (MgSO <sub>4</sub> ) (solution stock : 1 M)	125 µl	
	Stérilisé par filtration à 0,22 µm et conserver à -20°C		
	Nitrate de calcium (Ca(NO <sub>3</sub> ) <sub>2</sub> ) (solution stock : 0,5M) filtré 0,22 µm	250 µl	

### 3.1.2. Transformation par électroporation

Les cellules d'*E. coli* stockées à -80°C sont décongelées sur glace. 20 µl de cellules sont mis en contact avec l'ADN à transformer dans des tubes froids. Le mélange est transféré dans une cuvette d'électroporation froide de 1 mm ou 2 mm. La cuvette est placée dans une électroporateur (Gene Pulser Xcell de chez Bio-Rad) puis les cellules subissent une impulsion électrique. Les paramètres de l'électroporateur sont ajustés en fonction de la souche utilisée en suivant les recommandations des fournisseurs. Les cellules sont reprises dans 500 ml de milieu SOC (NEB) puis transférées dans un tube de 13 ml et incubées sous agitation (200 rpm) 1 h à 37°C ou 30°C pour la souche ElectroMAX™ Stbl4™. Un volume de 250 µl (non dilué ou dilué à 10<sup>-1</sup>) des bactéries est étalé sur milieu LB solide sélectif, en fonction de l'ADN transformé, les transformants sont sélectionnés comme décrit partie 1.1. Enfin les boites sont placées à 37°C (ou 30°C pour la souche ElectroMAX™ Stbl4™) pendant une nuit.

## 3.2. Transformation de *Bacillus subtilis*

### 3.2.1. Protocole MG1/MG2

Ce protocole de transformation de *B. subtilis* est basé sur sa capacité de compétence naturelle. Il a été développé en 1961 par Anagnostopoulos et Spizizen (Anagnostopoulos and Spizizen, 1961). Il consiste à rendre les cellules compétentes par la culture successive en deux milieux, le premier, minimum enrichi appelé MG1 puis le second carencé appelé MG2. Ces milieux sont confectionnés à partir d'une solution MG0 elle-même préparée à partir d'une solution mère MM 10X. Les milieux sont conservés à -20°C et sont décongelés à 4°C (tableau 12). La veille de la transformation, à partir du stock à -80°C les cellules de *B. subtilis* sont mises en culture sur milieu LB solide avec sélection si besoin et placées à 37°C pour la nuit. Une colonie fraîche isolée issue de cette culture estensemencée dans 2 ml de MG1, le tout est incubé 4h30 à 37°C sous agitation. 200 µl de la culture sont dilués dans le milieu MG2 pour un volume final de 2 ml (dilution à 1/10) puis incubés 1h30 dans les mêmes conditions que précédemment. Pour une condition de transformation, 100 µl de cellules issu de la culture MG2 sont mis en contact avec l'ADN à transformer. Le mélange est incubés 20 min à 37°C au bain marie. Puis la totalité est étalée avec des billes sur milieu LB solide sélectif, en fonction

**Tableau 13: Liste et composition des solutions utilisées lors de de la transformation de *B. subtilis* avec le protocole MC**

MC 10X (pour un volume final de 50 ml)	Potassium Phosphate Dibasique (K <sub>2</sub> HPO <sub>4</sub> )	5,35 g
	Potassium Phosphate Monobasique (KH <sub>2</sub> PO <sub>4</sub> )	2,6195 g
	Glucose	10 g
	Citrate de Sodium Trisodique (solution stock : 300 mM)	5 ml
	Citrate d'ammonium Ferrique (solution stock : 2,2%)	0,5 ml
	Hydrolysate de Caséine	0,5 g
	Glutamate de Potassium	1,1 g
	Qsp H <sub>2</sub> O	50 ml
	Stérilisé par filtration à 0,22 µm et conserver à -20°C	
MC 1X (pour un volume final de 20 ml)	MC 10X	2 ml
	H <sub>2</sub> O	18 ml
	Sulfate de Magnésium (MgSO <sub>4</sub> ) (solution stock : 1 M)	67 µl
	Stérilisé par filtration à 0,22 µm et conserver à -20°C	

Solutions mères		
Citrate de Sodium Trisodique (300 mM)	Tri-NA Citrate	8,823g
	Qsp H <sub>2</sub> O	100 ml
	Conservé à température ambiante	
Citrate d'ammonium Ferrique (2,2%)	Citrate d'ammonium Ferrique	2,2g
	Qsp H <sub>2</sub> O	100 ml
	Conservé à 4°C enveloppé dans de l'aluminium	

de l'ADN transformer, les transformants sont sélectionnés comme décrit partie 1.2. Enfin les boîtes sont placées à 37°C pendant au moins une nuit jusqu'à apparition des colonies.

### 3.2.2. Protocole « simple » (MC)

Ce protocole, assez simple, est basé sur la capacité de recombinaison homologe de *B. subtilis* et ne requiert qu'un seul milieu de culture, le milieu « Modified competence » (MC) (tableau 13) (Konkol *et al.*, 2013). Ce milieu est préparé à partir d'une solution mère MC 10X (tableau 13) diluée à 1X dans de l'eau stérile et complétée avec du MgSO<sub>4</sub> à 1 M. La solution mère est conservée à -20°C et est décongelée à 4°C.

La veille de la transformation, à partir du stock à -80°C les cellules de *B. subtilis* sont mises en culture sur milieu LB solide avec sélection si besoin et placées à 37°C pour la nuit. Une colonie fraîche isolée issue de cette culture estensemencée dans 3 ml de milieu MC et est cultivée 6 h à 37°C sous agitation (200 rpm). 200 ml de cette culture sont mis en contact avec 1,5 µg de l'ADN à transformer. Le mélange est incubé 2 h à 37°C au bain marie. Puis la totalité est étalée avec des billes sur milieu LB solide sélectif, en fonction de l'ADN transformer, les transformants sont sélectionnés comme décrit partie 1.2. Enfin les boîtes sont placées à 37°C pendant au moins une nuit.

## 3.3. Transformation de *S. cerevisiae*

### 3.3.1. Transformation à l'acétate de lithium

Ce protocole de transformation de *S. cerevisiae* est adapté de celui décrit en 1995 (Gietz *et al.*, 1995). Il est basé sur l'utilisation du PEG un agent fusogène qui va favoriser l'entrée de l'ADN dans les cellules et de l'acétate de lithium qui va fragiliser les parois des levures. Les levures sont mises en culture en milieu liquide YPDA ou SD additionné du « DO supplement » requis et sont placées à 30°C sous agitation (200 rpm). Lorsque la DO<sub>600nm</sub> est environ égale à 1, les cultures sont diluées au quart dans un volume final de 20 ml de YPDA et remises en culture dans les mêmes conditions que décrites précédemment jusqu'à obtenir de nouveau une DO<sub>600nm</sub> environ égale à 1. Pour une condition de transformation, 2,5 ml de culture sont prélevés et centrifugés pendant 2 min à 6000 g. Le culot cellulaire est lavé deux fois dans 1 ml

**Tableau 14: Liste et composition des solutions utilisées lors du protocole de transformation de *S. cerevisiae* à l'acétate de lithium**

0.1 M LiAc/1X TE (pour un volume final de 100 ml)	Tris-HCl (solution stock : 1 M) pH 7.5	1 ml
	EDTA (solution stock : 0.5 M) pH 7.5	20 µl
	Lithium acetate dehydrate	1.02 g
PEG 4000 40%/0.1 M LiAc/1X TE (pour un volume final de 50 ml)	PEG 4000	20 g
	Qsp 0.1 M LiAc/1X TE	50 ml

**Tableau 15: Liste et composition des solutions utilisées lors du protocole de CReasPy-Fusion**

Sorbitol 1M (pour un volume final de 1 l)	Sorbitol	182 g
	Qsp H <sub>2</sub> O	1 l
	Stérilisé par filtration à 0,22 µm et conservé à température ambiante	
SPEM (pour un volume final de 500 ml)	Na <sub>2</sub> HPO <sub>4</sub> x 2H <sub>2</sub> O	0,69 g
	NaH <sub>2</sub> PO <sub>4</sub> x 2H <sub>2</sub> O	0,18 g
	EDTA (solution stock 0,5 M) pH8	10 ml
	Sorbitol	91 g
	Qsp H <sub>2</sub> O	500 ml
	Stérilisé par filtration à 0,22 µm et conservé à température ambiante	
STC (pour un volume final de 100 ml)	Sorbitol	18,2 g
	Tris-HCl (solution stock : 1 M) pH 7,5	1 ml
	CaCl <sub>2</sub> (solution stock : 1 M)	1 ml
	MgCl <sub>2</sub> (solution stock : 1 M)	225µl
	Qsp H <sub>2</sub> O	100 ml
	Stérilisé par filtration à 0,22 µm et conservé à température ambiante	
SMM 1X (pour un volume final de 50 ml)	saccharose	8,55 g
	Maleate Buffer (solution stock : 0,2 M)	5 ml
	MgCl <sub>2</sub> (solution stock : 1 M) pH = 7,5	1 ml
	Qsp H <sub>2</sub> O	50 ml
	pH ajusté à 6,5	
	Stérilisé par filtration à 0,22 µm et conservé à température ambiante	
PEG 20% (pour un volume final de 30 ml)	PEG 8000 (BP233-1 Fisher Scientific)	6 g
	Tris (solution stock : 1 M) pH 8	300 µl
	CaCl <sub>2</sub> (solution stock : 1 M)	300 µl
	MgCl <sub>2</sub> (solution stock : 1 M)	75 µl
	Qsp H <sub>2</sub> O	30 ml
	Stérilisé par filtration à 0,22 µm et conservé à température ambiante	
SOS (pour un volume final de 50 ml)	Sorbitol	9,1 g
	Extrait de levure	0,125 g
	Bacto peptone	0,25 g
	CaCl <sub>2</sub> (solution stock : 1 M)	300 µl
	Qsp H <sub>2</sub> O	50 ml
	Stérilisé par filtration à 0,22 µm et conservé à température ambiante	

de solution de l'acétate de lithium/TE (0,1 M LiAc/1X TE) (tableau 14). Les cellules sont ensuite incubées 30 min à 30°C, puis sont collectées par centrifugation pendant 3 min à 5000 g et remises en suspension dans 50 µl de 0,1 M LiAc/1X TE. Elles sont mises en contact avec l'ADN à transformer (500 ng) et 50 µg d'ADN carrier (Solution d'ADN de sperme de saumon UltraPure™ de chez Invitrogen™) dénaturé au préalable 10 min à 100°C et refroidi sur glace. 500 µl de PEG 4000 40%/0,1 M LiAc/1X TE (tableau 14) ainsi que 57 µl diméthylsulfoxyde (DMSO) (10% du volume final) sont ajoutés au mélange ADN/levure. Le mélange est ensuite incubé 30 min à 30°C sans agitation, puis un choc thermique à 42°C pendant 25 min est réalisé. Enfin les cellules sont collectées par centrifugation 3 min à 5000 g et remises en suspension dans 200 µl d'eau stérile. Le tout est étalé au râteau sur milieu solide SD complétement avec le « DO supplement » requis. Les boîtes sont placées pendant au moins 2 jours à 30°C.

### 3.3.2. Transformation des sphéroplastes au PEG : méthode de CReasPy-Fusion (fusion avec *B. subtilis*)

Ce protocole de transformation dérive de celui décrit en 2013 par Karas *et al.*, (Karas *et al.*, 2013). Il est basé sur l'utilisation du PEG et met en jeu deux types de cellules, des sphéroplastes de *S. cerevisiae* et des protoplastes de *B. subtilis*. Les sphéroplastes de levures sont obtenus par traitement à la zymolyase et les protoplastes de *B. subtilis* sont obtenus par action du lysozyme. L'ADN à transformer dans la levure est apporté par la cellule bactérienne « fusionnée » avec la levure. Ce protocole est utilisé pour la transformation de molécules d'ADN de grande taille, ici des fragments de génomes de *B. subtilis* de l'ordre de plusieurs centaines de kb voire du Mb. La préparation des deux types de cellules se fait de manière simultanée.

#### - Préparation des sphéroplastes de *S. cerevisiae* :

En fonction de la souche à transformer les cellules de *S. cerevisiae* sont mises en culture dans 50 ml de milieu liquide YPDA ou SD/DO supplement et sont placées à 30°C sous agitation (200 rpm). Lorsque la DO<sub>600nm</sub> est comprise entre 2,5 et 3 elles sont collectées par centrifugation pendant 5 min à 2500 g à 10°C. Le culot est lavé dans 50 ml d'eau stérile et de nouveau centrifugé 5 min à 2500 g à 10°C, puis il est remis en suspension dans 50 ml de sorbitol 1 M (tableau 15) et centrifugé une dernière fois dans les mêmes conditions. Le



surnageant est éliminé et le culot est conservé à 4°C pendant une nuit. Le jour de la transformation le culot est repris dans 20 ml de SPEM (tableau 15). Les levures sont traitées pour obtenir des sphéroplastes en ajoutant 40 µl de zymolyase 20T et 30 µl de β-Mercaptoéthanol. Le tout est placé à 30°C au bain-marie et des mesures de DO sont réalisées régulièrement en diluant les cellules au 1/10<sup>ème</sup> dans du sorbitol 1 M ou de l'eau. Lorsque le rapport de DO sorbitol/H<sub>2</sub>O est compris entre 1,8 et 2 la réaction est immédiatement arrêtée par l'ajout de 30 ml de sorbitol 1 M. Les sphéroplastes produits sont centrifugés à 1500 g pendant 5 min à 4°C, puis le culot est repris délicatement dans 50 ml de sorbitol 1 M. Enfin, ils sont de nouveau centrifugés dans les mêmes conditions et remis en suspension dans 2 ml de STC (tableau 15) par agitation douce du tube (à la main). Les sphéroplastes de levure ainsi obtenus sont laissés à température ambiante au moins 20 min jusqu'à utilisation pour la transformation.

- Préparation des protoplastes de *B. subtilis* :

En fonction de la souche à transformer les cellules de *B. subtilis* sont mises en culture dans 50 ml de LB complété ou non d'un antibiotique et sont placées à 37°C sous agitation (200 rpm). Lorsque la DO<sub>600nm</sub> est comprise entre 0,4 et 1, du chloramphénicol à une concentration finale de 180 µg/ml est ajouté, puis la culture est placée dans les mêmes conditions que précédemment pour 55 min. Elle est ensuite centrifugée à 7000 g pendant 5 min à 10°C, et le culot est repris dans 5 ml de SMM 1X (tableau 15). Pour produire les protoplastes, du lysozyme à une concentration finale de 2 mg/ml est ajouté à la suspension qui est incubée à 37°C sous agitation douce (30 rpm) pour 1 h à 2 h. Ensuite, les protoplastes sont collectés par centrifugation à 2600 g pendant 15 min à 10°C puis sont repris dans 5 ml de SMM 1X. Ils sont de nouveau centrifugés dans les mêmes conditions puis mis en suspension dans 1 ml de SMM 1X. A ce stade, les protoplastes de *B. subtilis* sont prêts à être utilisés.

- Mise en contact des cellules : transformation de la levure

Pour une condition de transformation 200 µl de sphéroplastes de levures sont mis en contact avec 50 µl ou 100 µl de protoplastes de *B. subtilis*. En fonction des conditions 11 µl de pansement de 90 pb ou 1 µg de pansement contenant les éléments levures sont ajoutés. Le tout est laissé 5 min à température ambiante puis additionné de 1 ml de PEG 20 % (tableau 15) puis de nouveau laissé à température ambiante pour 20 min. La suspension cellulaire est



ensuite centrifugée à 1500 g pendant 5 min, le culot est remis en suspension dans 800 µl de SOS (tableau 15) puis placé 2 h à 30°C. Pour l'étalement sur boîte, les cellules sont mélangées à 12 ml de TOP agar (milieu sélectif SD avec 1 M de sorbitol et 3 % d'agar maintenu en surfusion à 55°C au bain marie) et directement étalées sur milieu sélectif solide SD contenant 1 M de sorbitol. Le TOP agar et le milieu sélectif solide SD sont toujours complétés par du chloramphénicol et de l'ampicilline à une concentration finale respectivement de 5 µg/ml et 100 µg/ml. Dans les conditions de contre sélection au 5'-FOA, ces milieux sont en plus additionnés d'acide 5-fluoro-orotique à une concentration finale de 1 g/l. Finalement les boîtes sont placées pendant au moins 3 jours à 30°C.

## 4. Observation en microscopie

### 4.1. Microscopie à fond noir

Les images de microscopie à fond noir ont été réalisées sur un microscope photonique droit Eclipse Ni de chez Nikon équipé d'un condenseur à fond noir (Nikon). L'objectif à immersion dans l'huile utilisé était le Nikon 60X avec une ouverture numérique (NA) de 0,80. Les photographies ont été faites avec la caméra photometrics iris 9 Scientific CMOS (2960 X 2960 pixels) et acquises avec le logiciel NIS-Elements D.

### 4.2. Microscopie à fluorescence

#### 4.2.1. Microscope à Epifluorescence

Les échantillons ont été préparés à partir de cultures de *B. subtilis* et de *S. cerevisiae* à une  $DO_{600nm}$  environ égale à 1 dont 5 µl ont été placés entre lame et lamelle.

Les images de microscopie à épifluorescence ont été réalisées sur un AxioImager de chez Zeiss. Différents objectifs à ouverture numérique (NA) ont été utilisés, du 5X au 20X sans immersion et du 63X au 100X avec immersion dans l'huile. Pour les observations en lumière transmise la caméra CCD cool Snap HQ2 a été utilisée. Pour celles en fluorescence, le filtre GFP-BP a été utilisé pour visualiser les mutants exprimant la GFP, et le filtre mRFP-LP pour ceux exprimant la RFP ou Kate2. Le signal a été collecté par une caméra AxioCam 208 color



(Zeiss). Les images ont été acquises avec le logiciel MicroManager. Elles ont ensuite été traitées par Fiji/ImageJ.

#### 4.2.2. Microscope Confocal

Les échantillons ont été préparés à partir de cultures de *B. subtilis* et de *S. cerevisiae* à une  $DO_{600nm}$  environ égale à 1 dont 5  $\mu$ l ont été placés entre lame et lamelle.

Les images de microscopie confocale ont été réalisées sur un Microscope Confocal Zeiss LSM 880 de chez Zeiss. L'objectif à ouverture numérique (NA) 63X avec immersion dans l'huile a été utilisé. Les lasers d'excitation Argon pour la longueur d'onde 488 nm et Helium neon pour 561 nm ont été utilisés respectivement pour les observations de la GFP et de la RFP. Différents détecteurs ont été employés, le PMT (photomultiplicateur) avec des plages de détection de 597 à 677 nm pour la RFP et 496 à 554 nm pour la GFP, le PMT trans pour les observations en lumière transmise ainsi que le Module Airyscan avec des longueurs d'ondes de détection à 579 nm pour la RFP et 515 nm pour la GFP. Les images ont été acquises et traitées avec les logiciels Zenblack et Zenblue.

La microscopie confocale se caractérise par l'utilisation de LASER (Light Amplification by Stimulated Emission of Radiation) ou de diode qui sont des sources de lumière ponctuelle avec des longueurs d'ondes très précises. Le faisceau d'excitation est balayé grâce à un scanner et est condensé par l'objectif sur le plan focal. Ce système d'illumination par balayage permet de limiter les interférences lumineuses en XY et participe à l'amélioration de la résolution par rapport à un microscope plein champs. De plus la présence d'un diaphragme (appelé « pinhole ») devant le détecteur permet de transmettre jusqu'au détecteur uniquement les photons émis par les molécules se trouvant sur le plan focal. Ainsi, les photons venant des autres plans de l'échantillon sont éliminés. En modifiant la position en z (de haut en bas) de l'échantillon, des sections optiques pourront être acquises et permettront l'acquisition d'images en 3 dimensions. La résolution optimale sera de 200 nm en xy et de 400 nm en z. Cela permet alors de réaliser des images sur des plans situés à différentes profondeurs et de réaliser la reconstruction d'une image à trois dimensions avec une résolution accrue. De plus, l'utilisation du mode de détection spectrale permet d'acquérir les spectres d'émission des échantillons et de valider ou non la spécificité des molécules fluorescentes observées. Cette



technique permet de démontrer que la fluorescence observée ne provient pas d'une fluorescence aspécifique ou une auto fluorescence. Afin d'améliorer encore la résolution (d'un facteur 1,4), nous avons aussi utilisé un détecteur AiryScan.



# Chapitre 1 : Adaptation de la méthode de CReasPy-Fusion à *B. subtilis*



## Chapitre 1 : Adaptation de la méthode de CReasPy-Fusion à *B. subtilis*

### 1. Introduction

Le projet Bacillus 2.0 vise à utiliser la levure *S. cerevisiae* comme plateforme d'assemblage et d'ingénierie de génomes semi-synthétiques de *B. subtilis* pour la construction rapide de châssis bactériens « nouvelle génération » et modulables pour les biotechnologies (= organismes « sur mesure »).

Un prérequis essentiel à ce travail dont le coût est estimé à environ 10 000 euros pour un génome de 2 Mb est de démontrer qu'un génome de *B. subtilis* ne contenant que les gènes essentiels et quasi-essentiels peut être cloné dans la levure et maintenu stable chez cet hôte pendant plusieurs générations.

Pour réaliser cette tâche, l'équipe a sélectionné, après mûres réflexions, la souche *B. subtilis* MGP192. Ce choix s'est fait sur les critères suivants :

- (i) *Caractéristiques liées au génome.* Le génome de cette souche est l'un des plus réduits (2,86 Mb) (Dervyn *et al.*, 2023) de la collection « Minimal Genome Project » de l'équipe SyBER (Institut MICALIS, UMR 1319 INRAE / AgroParisTech, Jouy-en-Josas) avec laquelle le projet a été conçu. Le pourcentage en G+C de ce génome reste identique à celui des souches parentales *B. subtilis* 168 et MS (~44 %), et, proche du seuil critique des 40 % pour la levure (cf partie 2.4.1 de l'introduction). De même, le nombre d'ARS putatives (recherchées via le motif ACS, 5'-WTTTATRTTT-3') (van Houten and Newlon, 1990; Deshpande and Newlon, 1992) reste en proportion équivalent au nombre identifié sur le génome de *B. subtilis* 168 (145 contre 221), et leur distribution relativement homogène le long du génome.
- (ii) *Caractéristiques liées à la souche.* Bien que la taille de son génome ait été fortement réduite (suppression d'1,34 Mb), la souche MGP192 conserve un temps de génération correct en milieu standard comme le milieu LB, ce qui permet de réaliser des expériences sur des échelles de temps comparables à celles de *B. subtilis* 168 ou MS. De plus, les gènes liés à la compétence ayant été préservés, *B. subtilis* MGP192 peut importer de l'ADN exogène depuis l'environnement, signifiant que des expériences de transformation peuvent être réalisées.



Une fois la souche sélectionnée, l'équipe a testé trois méthodes éprouvées de clonage de génomes bactériens dans la levure (cf parties 2.4.1.1, 2.4.1.3, 2.4.1.4 et figures 15, 16, et 18 de l'introduction). Ces expériences, réalisées avant mon arrivée au laboratoire, ont donné les résultats suivants :

- Méthode classique (Benders *et al.*, 2010)(figure 15): aucune colonie levure n'est apparue sur milieu de sélection après transformation indiquant que le génome de *B. subtilis* pré-marqué avec les éléments levure et extrait en bloc d'agarose n'entre probablement pas dans la levure.
- Méthode CReasPy-Cloning (Ruiz *et al.*, 2019) (figure 16) : des colonies levure sont apparues sur milieu de sélection après transformation mais après criblage, il s'est avéré que le génome de *B. subtilis* MGP192 n'était présent dans aucun des transformants analysés. Comme pour la méthode classique, le génome de *B. subtilis* (ici, non marqué) extrait en bloc d'agarose n'entre probablement pas dans la levure ; seul le pansement (petit fragment PCR linéaire) composé des éléments levure ARSH4-CEN6-HIS3 y pénètre, générant des faux positifs sur milieu de sélection.
- Méthode de clonage par fusion (Karas *et al.*, 2013) (figure 18): des colonies levure sont apparues sur milieu de sélection après transformation, le génome de *B. subtilis* MGP192 a été détecté mais il n'était entier que lors des passages précoces. Aux passages plus tardifs, le génome présent était fortement tronqué, laissant place à des molécules d'environ 300 kb couvrant seulement la région du génome de *B. subtilis* dans laquelle était inséré les éléments levure ARSH4-CEN6-HIS3. Le chromosome semblait donc pénétrer et s'installer entier dans le noyau de la levure mais ne s'y maintenait pas intact au cours de la propagation des cellules.

Ces expériences, reproduites plusieurs fois et produisant invariablement les mêmes résultats, nous ont amenés à conclure que la méthode de clonage par fusion était la seule méthode permettant de détecter de l'ADN de *B. subtilis* dans la levure.

L'équipe en a ressorti trois enseignements majeurs : (i) les méthodes utilisant des génomes entiers *B. subtilis* MGP192 « nus », c'est-à-dire extraits en bloc d'agarose et traités à la protéinase K (et donc dans un état probablement « relâché »), ne fonctionnent pas, semblant indiquer que ces méthodes ne sont pas adaptées à des génomes de grande taille comme celui de *B. subtilis* MGP192; (ii) la méthode par fusion, basée sur un contact direct bactérie/levure (et qui ne passe pas par une étape d'extraction de génomes entiers), semble,



à ce jour, la seule méthode permettant l'entrée du génome de *B. subtilis* MGP192 dans la levure. Il est très probable qu'au cours de ce processus le génome soit délivré dans la cellule hôte sous forme de nucléoïde (et donc dans un état très fortement condensé), cette méthode serait donc plus appropriée que les autres pour les génomes de grande taille ; et (iii) après son entrée dans la levure, le génome de *B. subtilis* MGP192 n'est pas stable, il subit de nombreux réarrangements avec certaines régions du génome systématiquement amputées.

Pour tenter de comprendre les bases de l'instabilité du génome de *B. subtilis* MGP192 dans la levure, l'équipe a poursuivi avec la mise en place d'expériences de TAR-cloning (Transformation-Associated Recombination) (Natalay Kouprina and Vladimir Larionov, 2016). Le clonage par recombinaison associée à la transformation (TAR) permet d'isoler de manière sélective, précise et efficace un fragment génomique donné dans un mélange plus ou moins complexe. Plus précisément, au cours de la transformation de sphéroplast, le TAR-cloning exploite le système de recombinaison homologue de la levure, dans le but d'assembler le fragment d'ADN cible avec un vecteur TAR (fragment PCR), qui porte à chaque extrémité des séquences d'environ 60 pb, appelées crochets, identiques aux extrémités du fragment cible. Ceci implique que l'ADN à capturer doit être fragmenté au préalable, soit par cassures mécaniques soit par digestion avec des enzymes de restriction.

L'utilisation du TAR-cloning a permis de cloner la totalité du génome de *B. subtilis* MGP192 en 21 morceaux de tailles variables (~6 kb à ~515 kb). Mais il a surtout dévoilé des zones du génome beaucoup plus instables que d'autres et mis en exergue un ensemble de facteurs certainement responsables de l'instabilité du génome MGP192 dans la levure. Parmi eux, on peut citer : (i) le manque d'ARS fonctionnelles dans certaines régions du chromosome, (ii) la présence de séquences répétées (*e.g.* les nombreuses copies d'opérons ribosomiques concentrées en un endroit du génome pouvant favoriser la survenue d'évènements de recombinaison) et/ou (iii) la présence de gènes dont l'expression simultanée a un effet toxique pour la levure.

Fort de ces connaissances, l'équipe a souhaité aller plus loin en capturant et en amenant des mesures correctives à certaines régions du génome MGP192 mais a rapidement été limitée par les méthodes et outils à disposition qui ne répondaient pas à ses besoins. Elle a alors imaginé une nouvelle stratégie de clonage, nommée CReasPy-Fusion (cf partie 2.4.1.5 de l'introduction et figure 17), qui serait suffisamment flexible pour capturer au choix n'importe



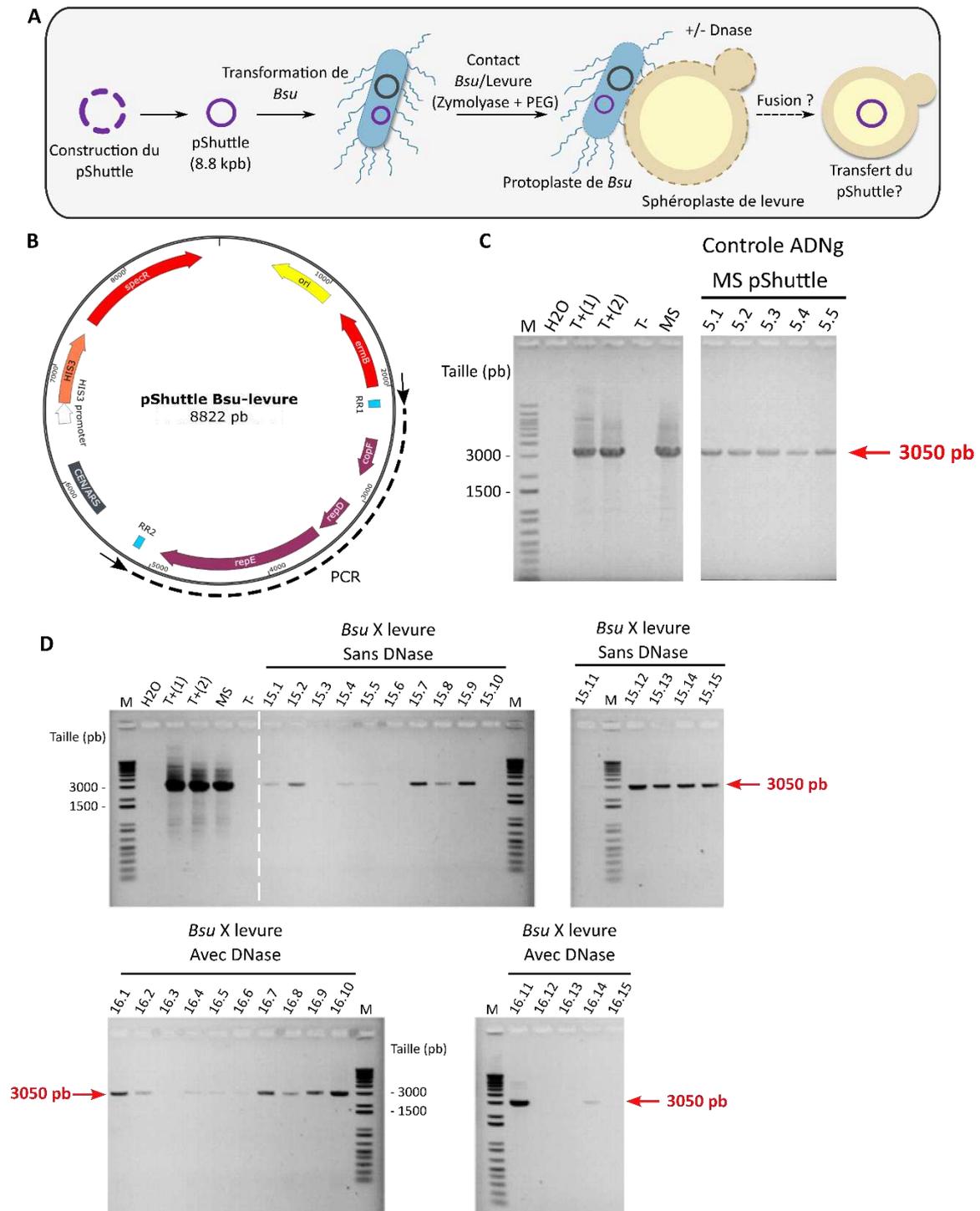
quel fragment de génome (voire un génome entier) et y introduire simultanément des modifications (*e.g.* ajout d'ARS, suppression de *loci* non désirés, etc) selon les besoins.

Au vu des nombreuses difficultés rencontrées avec le génome de *B. subtilis* MGP192, la mise au point de cette nouvelle méthode a été réalisée avec des espèces de mycoplasmes dont les génomes avaient déjà été clonés dans la levure en utilisant d'autres méthodes. Au total, les génomes de six espèces distinctes de mycoplasmes, provenant de trois groupes phylogénétiques distincts, ont été clonés/édités dans la levure. Au cours de ce travail, de grands fragments de génome ont pu être également capturés avec une grande efficacité, démontrant la puissance de la méthode (Guesdon *et al.*, 2023). La preuve de concept étant faite, il ne restait plus qu'à appliquer la méthode à *B. subtilis* MGP192.

Ce premier chapitre décrit l'adaptation de la méthode de CReasPy-Fusion à *B. subtilis*. Pour atteindre cet objectif, un programme de travail en cinq étapes a été établi :

- Etape A : Démontrer le transfert d'un plasmide navette initialement porté par des protoplastes de *B. subtilis* à des sphéroplastes de levures après contact cellules/cellules.
- Etape B : Démontrer le transfert et la modification d'un plasmide navette initialement porté par des protoplastes de *B. subtilis* à des sphéroplastes de levures hébergeant le système CRISPR-Cas9 après contact cellules/cellules.
- Etape C : Démontrer le transfert d'un génome entier marqué de *B. subtilis* à des sphéroplastes de levures portant le système CRISPR-Cas9 après contact cellules/cellules en capturant de manière précise un fragment de génome de *B. subtilis*.
- Etape D : Tenter de simplifier la méthode en travaillant avec des génomes de *B. subtilis* non marqués.
- Etape E : Documenter l'étape de contact « protoplastes de *B. subtilis*/sphéroplastes de levures » grâce à des observations en microscopie à fond noir, microscopie à épifluorescence et microscopie confocale afin de mieux comprendre le processus de transfert d'ADN.

Les étapes A et B ont été réalisées par le Dr. Guesdon avant mon arrivée au laboratoire. Mes travaux de thèse ont donc démarré avec la mise en place des étapes C, D et E. Bien que



**Figure 42 : Expérience de transfert du pShuttle entre les protoplastes de *B. subtilis* (*Bsu*) et les sphéroplastes de levure par contact cellules/cellules. [A] Schéma de la procédure expérimentale. [B] Carte du plasmide navette pShuttle. La PCR d'analyse avec les amorces Verif\_pShuttle\_F2 et Verif\_pShuttle\_R est représentée. [C] Analyse des transformants levure hébergeant le pShuttle après transformation avec l'ADNg de la souche *B. subtilis* hébergeant le pShuttle. La présence du plasmide est contrôlée par PCR avec un couple d'amorces spécifiques produisant un amplicon à 3050 pb. [D] Analyse des transformants levure hébergeant le pShuttle après contact cellules/cellules. M : DNA Ladder 1 kb+ Invitrogen (100-12000 pb). H2O : contrôle négatif, sans ADN. T+(1) : témoin positif, plasmide pJIM. T+(2) : témoin positif plasmide pShuttle. T- : témoin négatif, ADNg de *S. cerevisiae* VL6-48N. MS : témoin positif, ADNg souche MS+pShuttle**

je ne sois pas intervenu sur les deux premières étapes, nous avons choisi de présenter ci-après l'ensemble du programme (*i.e.* étapes A à E) afin de permettre aux lecteurs de bien comprendre la logique derrière l'approche mise en place et l'articulation des différentes étapes.

## 2. Travaux réalisés avant mon arrivée au laboratoire

Avant mon arrivée au laboratoire, les étapes A et B nécessaires à la mise en place de la CREasPy-Fusion chez *B. subtilis* avaient été validées. Les principaux résultats sont résumés ci-dessous

### 2.1. Etape A: Transfert d'un plasmide navette par contact *S. cerevisiae/B. subtilis*

Le groupe de B. Karas avait démontré en 2013 qu'il était possible de transférer chez la levure *S. cerevisiae* des génomes de mycoplasmes (petites bactéries sans paroi) ainsi que ceux de bactéries Gram- (*H. influenzae*) simplement en mettant ces cellules (débarassées de leurs paroi) en contact en présence d'agent fusogène, le PEG (Karas *et al.*, 2013; Karas *et al.*, 2014). Suite à ces résultats, notre groupe avait alors tenté d'étendre ces travaux à la bactérie Gram+ *B. subtilis* (thèse du Dr. Guesdon) mais les données expérimentales amassées indiquaient que le génome, bien qu'apparemment transféré, ne semblait pas se maintenir chez cet hôte.

Dans le but de démontrer que de l'ADN issu de bactéries Gram+ pouvait être transféré chez la levure, nous avons décidé de simplifier l'approche expérimentale en visant à montrer dans un premier temps qu'un plasmide navette (appelé pShuttle), pouvant se propager à la fois chez *B. subtilis* (grâce aux gènes *copF*, *repD* et *repE*) et chez *S. cerevisiae* (grâce aux éléments levure ARSH4, CEN6, HIS3) pouvait être transféré de *B. subtilis* à la levure. Le pShuttle a donc été construit en assemblant par la méthode Gibson tous les éléments nécessaires à sa propagation chez les 2 hôtes, puis il a été transformé dans la souche MS de manière à produire des souches de *B. subtilis* MS hébergeant le plasmide navette. Après transformation et criblage par PCR, un clone a été sélectionné et utilisé pour des expériences de transfert d'ADN plasmidique entre protoplastes de *B. subtilis* et sphéroplastes de levure (souche VL6-48N) après contact cellules/cellules en présence de PEG (figure 42A). Ces expériences ont été menées en absence puis en présence de DNase I afin d'écartier toute

**Tableau 16 : Résultats de l'expérience de transfert du pShuttle par contact cellules/cellules entre les protoplastes de *B. subtilis* et les sphéroplastes de levure en absence ou en présence de DNase I.**

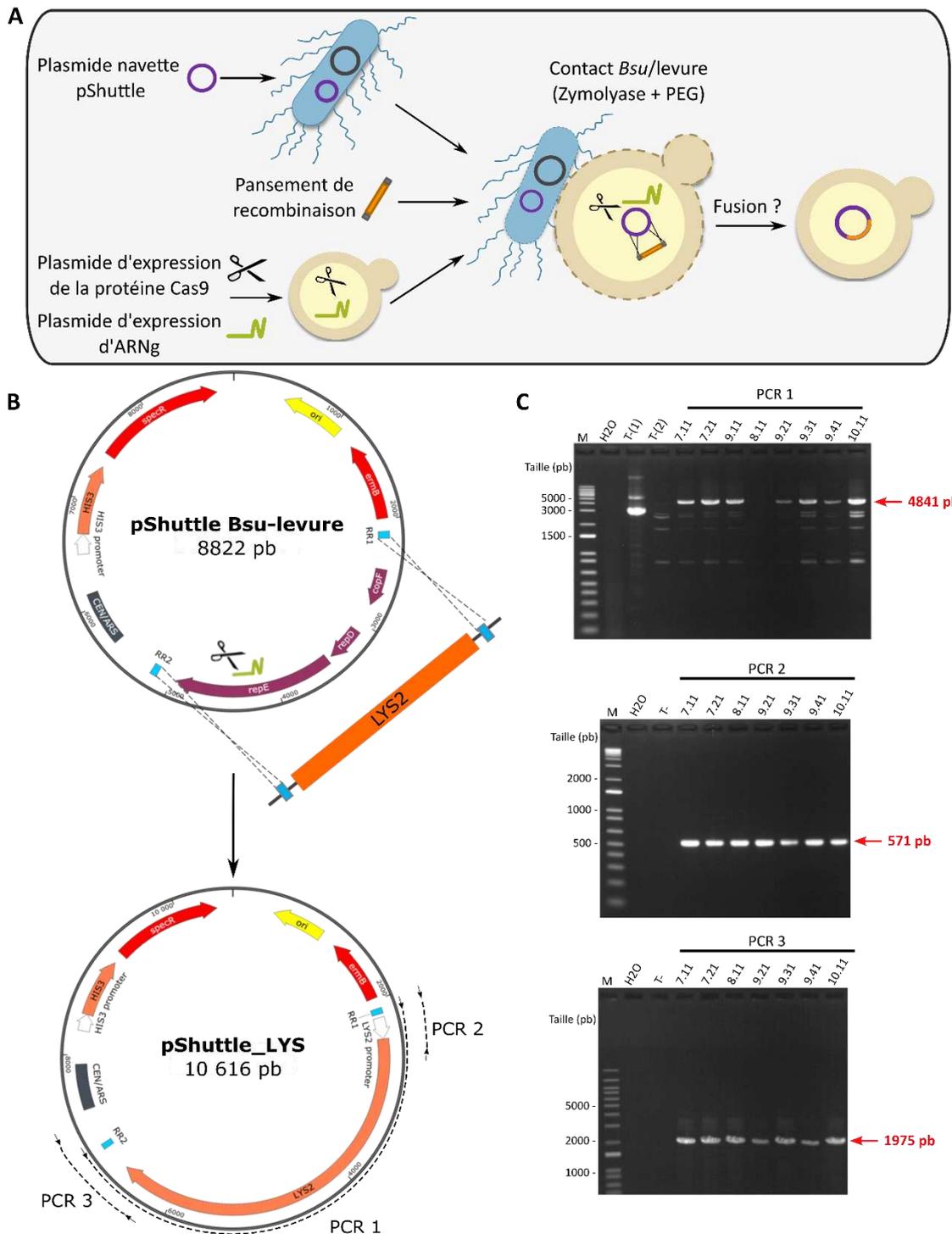
	Conditions expérimentales	Présence de DNase I	CFU	Clones positifs / Clones analysés
				PCR simplex
<b>Contrôle</b>	(ADNg <i>Bsu</i> MS + pShuttle) x levure	Sans DNase I	>300	5/5
	(ADNg <i>Bsu</i> MS + pShuttle) x levure	DNase I	0	-
<b>Fusion <i>Bsu</i> MS/levure</b>	( <i>Bsu</i> MS + pShuttle) x levure	Sans DNase I	40	11/15
	( <i>Bsu</i> MS + pShuttle) x levure	DNase I	52	11/15

\* : Tableau dérivé de la thèse du Dr. Gabrielle Guesdon

possibilité d'entrée de l'ADN plasmidique chez la levure par transformation suite à de la lyse bactérienne. Les résultats ont été les suivants :

- Pour les conditions contrôles, réalisées avec de l'ADN total de la souche *B. subtilis* MS/pShuttle isolé en bloc d'agarose, nous avons obtenu plusieurs centaines de colonies sur milieu sélectif en absence de DNase I alors qu'aucune colonie n'est apparue sur ces mêmes milieux en présence de DNase I (tableau 16). Cinq colonies ont été repiquées à partir de la boîte contrôle sans DNase I puis analysées par PCR à l'aide du couple d'amorces Verif\_pShuttle\_F2 et Verif\_pShuttle\_R ciblant spécifiquement le pShuttle (figure 42B). Pour ces 5 colonies, un amplicon à la taille attendue de 3050 pb a été obtenu (figure 42C). Ces résultats ont permis de valider, d'une part, que le pShuttle (présent dans l'ADN total) pouvait être transféré par transformation et se propager dans les levures au cours des générations et, d'autre part, que la DNase I utilisée était efficace pour dégrader l'ADN plasmidique libre dans le milieu.
- Pour les conditions mettant en jeu les protoplastes de *B. subtilis* MS hébergeant le pShuttle, le nombre de colonies obtenues sur boîte, que ce soit en présence (52 colonies) ou en absence (40 colonies) de DNase I dans le milieu, était sensiblement le même. Pour chacune de ces conditions, 15 clones ont été analysés par PCR en utilisant les mêmes amorces que pour les conditions contrôles (tableau 16). Dans chacun des cas, 11 clones sur 15 ont montré un amplicon à la taille attendue de 3050 pb (figure 42D). Ces résultats indiquaient qu'il était possible d'isoler des levures propageant le pShuttle transféré depuis *B. subtilis* même en présence de DNase I dans le milieu.

La mise en œuvre de cette première expérience nous a permis de confirmer la possibilité de transférer de l'ADN plasmidique entre une bactérie Gram+ (*B. subtilis*) et la levure *S. cerevisiae* après contact cellules/cellules. En effet, l'utilisation de DNase I au cours de l'expérience a permis de démontrer qu'il ne s'agissait pas d'évènements de transformation et que les mécanismes mis en jeu reposaient donc sur un contact direct entre les cellules. Cependant, à ce stade, il n'était pas possible de dire si le transfert d'ADN se faisait : (i) par fusion membranaire entre la bactérie et la levure, entraînant le relargage du plasmide dans le cytoplasme de la levure ou bien (ii) par endocytose de la bactérie portant le plasmide dans la levure.



**Figure 43 : Expérience de transfert et modification du pShuttle depuis *B. subtilis* vers la levure par CRasPy-Fusion. [A]** Schéma de la procédure expérimentale. **[B]** Modification du plasmide pShuttle. Par le système CRISPR-Cas9 porté par la levure : les gènes *repD*, *repE* et *copF* sont remplacés par le marqueur d'auxotrophie *LYS2*. **[C]** Analyse des transformants levure hébergeant le pShuttle modifié. La présence du plasmide modifié est contrôlée par PCR avec trois couples d'amorces. La PCR 1 avec le couple d'amorces Verif\_pShuttle\_F2 et Verif\_pShuttle\_R produit un amplicon de 4481 pb. Les deux autres PCR, spécifiques des jonctions de la zone éditée, produisent des amplicons de 571 pb pour la PCR 2 et 1975 pb pour la PCR 3. M : DNA Ladder 1 kb+ Invitrogen (100-12000 pb). H2O : contrôle négatif sans ADN. T-(1) témoin négatif, plasmide pShuttle. T-(2) : témoin négatif, ADNg de *S. cerevisiae* VL6-48N.

## 2.2. Etape B : Transfert et modification d'un plasmide navette depuis *B. subtilis* vers la levure par CREasPy-Fusion

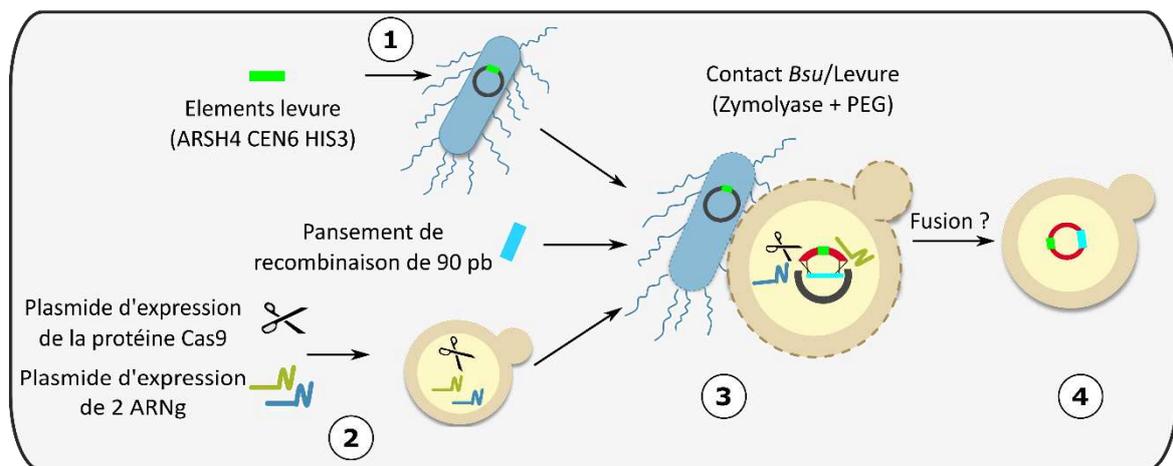
Dans la section précédente nous avons montré qu'il était possible de transférer une molécule d'ADN directement depuis des protoplastes de *B. subtilis* vers des sphéroplastes de levure par contact cellules/cellules. Afin d'adapter la méthode de CREasPy-Fusion à *B. subtilis*, l'étape suivante consistait à utiliser des levures hébergeant le système CRISPR-Cas9 et à démontrer que le pShuttle pouvait non seulement être cloné dans ces nouvelles cellules hôtes mais également être édité au niveau d'un locus choisi par l'expérimentateur (figure 43A). La modification souhaitée (illustrée figure 43B) visait à insérer en lieu et place des gènes *copF*, *repD* et *repE* portés par le pShuttle un second marqueur d'auxotrophie (LYS2) en plus du marqueur déjà présent pour l'histidine (HIS3).

Pour ce faire, la souche *S. cerevisiae* VL6-48N, (utilisée à l'étape A) a été transformée avec un plasmide permettant l'expression de la protéine Cas9 ainsi qu'un plasmide d'expression d'un ARNg ciblant spécifiquement le pShuttle au niveau du gène *repE*. Elle a ensuite été mise en contact (sous forme de sphéroplastes) avec des protoplastes *B. subtilis* MS/pShuttle en présence de PEG. Lors de cette expérience, un contrôle mettant en jeu une préparation de pShuttle en solution a également été utilisée afin de s'assurer que le système CRISPR-Cas9 de la levure était bien efficace pour permettre la modification désirée (contrôle positif). Les résultats ont été les suivants :

- Pour la condition contrôle (pShuttle purifié), nous avons observé plus d'un millier de colonies sur milieu sélectif. Dix d'entre elles ont été repiquées puis analysées par PCR (tableau 17), en utilisant un couple d'amorces Verif\_pShuttle\_F2 et Verif\_pShuttle\_R ciblant le pShuttle spécifiquement au niveau de la région éditée (PCR 1, figure 43B). Sur les 10 testés, 8 clones ont montré un amplicon à la taille attendue (4841 pb) et deux clones ont montré un profil mixte avec deux amplicons de tailles différentes, un à la taille de 4841 pb correspondant au plasmide édité et un à la taille de 3050 pb correspondant au plasmide initial (résultats non montrés). Nous avons conclu que le système CRISPR-Cas9 semblait efficace et permettait la modification du pShuttle dans la levure dans environ 80% des cas.

**Tableau 17 : Résultats de l'expérience de transfert et modification du pShuttle par CReasPy-Fusion.**

	Conditions expérimentales	CFU	Clones positifs / Clones analysés <sup>a</sup>
			PCR simplex
<b>Contrôle</b>	pShuttle seul x levure CRISPR	>1000	8/10
<b>Fusion <i>Bsu</i> MS/levure</b>	( <i>Bsu</i> MS + pShuttle) x levure CRISPR	9	7/8



**Figure 44 : Schéma de la capture, par la méthode de CReasPy-Fusion, d'un fragment de génome de *B. subtilis* pré-marqué avec les éléments levure. [1] Les éléments levure sont intégrés au génome de *B. subtilis* par transformation bactérienne ; [2] Le système CRISPR-Cas9 est pré-chargé dans la levure ; [3] L'ADN de la bactérie est transféré dans la levure par contact cellules/cellules. Le fragment à capturer est isolé du reste du génome par le système CRISPR-Cas9 ; [4] Le fragment est re-circularisé dans la levure par recombinaison homologue avec le pansement de réparation de 90 pb fourni.**

- Pour la condition de contact entre les protoplastes de la souche *B. subtilis* MS propageant le pShuttle et les sphéroplastes de la souche levure VL6-48N pré-chargée avec le système CRISPR-Cas9, 9 colonies ont été obtenues sur milieu sélectif. Huit clones ont été analysés dans un premier temps en utilisant les mêmes amorces que pour la condition contrôle et 7 ont montré un amplicon à la taille attendue (PCR1 : 4841 pb) (figure 43C, gel du haut). Pour le dernier clone (cl 8.11), aucun amplicon n'a été obtenu. Concernant les 7 clones positifs pour la PCR 1, l'intégration de LYS2 a ensuite été confirmée par les PCR 2 et 3 grâce à des couples d'amorces spécifiques couvrant les jonctions de la zone éditée (figure 43B). Les 7 clones ont montré des amplicons aux tailles attendues pour les deux PCR, respectivement 571 pb et 1975 pb (figure 43C), validant ainsi la modification du pShuttle.

Les résultats obtenus au cours de cette seconde expérience nous ont permis de montrer que le système CRISPR-Cas9 présent dans la levure était fonctionnel et permettait la modification du pShuttle (tel que décidé par l'expérimentateur) après son transfert depuis des protoplastes de *B. subtilis* vers des sphéroplastes de levure. L'étape B était ainsi validée.

### 3. Etape C : Capture d'un fragment de génome de *B. subtilis* marqué avec les éléments levures ARSH4 CEN6 HIS3

Les expériences décrites dans la partie précédente ont montré qu'il est possible de (i) transférer un plasmide directement depuis des protoplastes de *B. subtilis* vers des sphéroplastes de levure et de (ii) modifier la molécule transférée grâce au système CRISPR-Cas9 porté par les levures. Ces validations nous ont conduits à passer à l'étape C de notre programme de travail : utiliser la méthode de CReasPy-Fusion pour capturer de manière spécifique, non plus une molécule plasmidique, mais un fragment de génome de *B. subtilis* dans la levure (figure 44).

C'est ici qu'a réellement commencé mon travail de thèse.

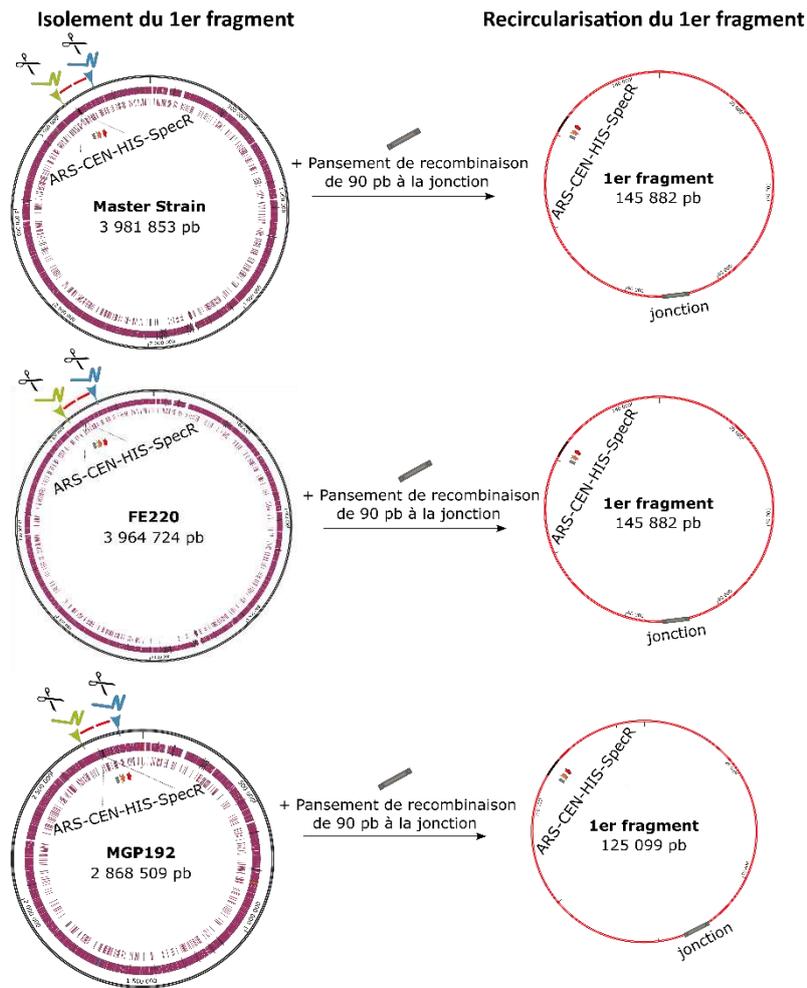


Figure 45: Zone cible pour la capture du premier fragment de génome correspondant à la région entre les gènes *atpA* et *IrgB* chez les souches Master Strain (MS), FE220 et MGP192.

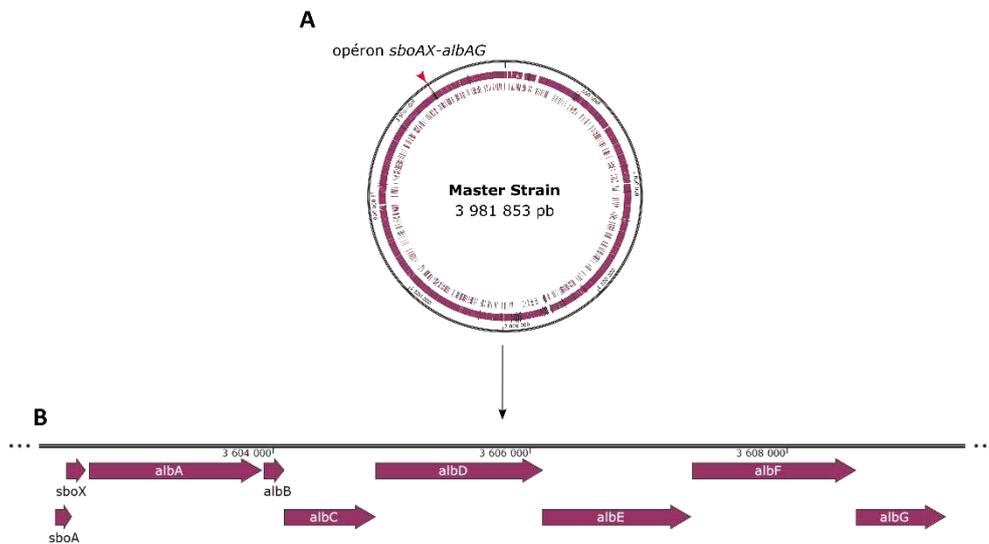


Figure 46: Représentation schématique de l'opéron ciblé par le système CRISPR-Cas9 pour l'insertion des éléments levure. [A] Localisation de l'opéron cible sur la carte du génome de la souche MS. [B] Opéron cible *sboAX-albAG*.

### 3.1. Construction des souches *B. subtilis* MS, FE220 et MGP192 marquées avec les éléments levures ARSH4 CEN6 HIS3

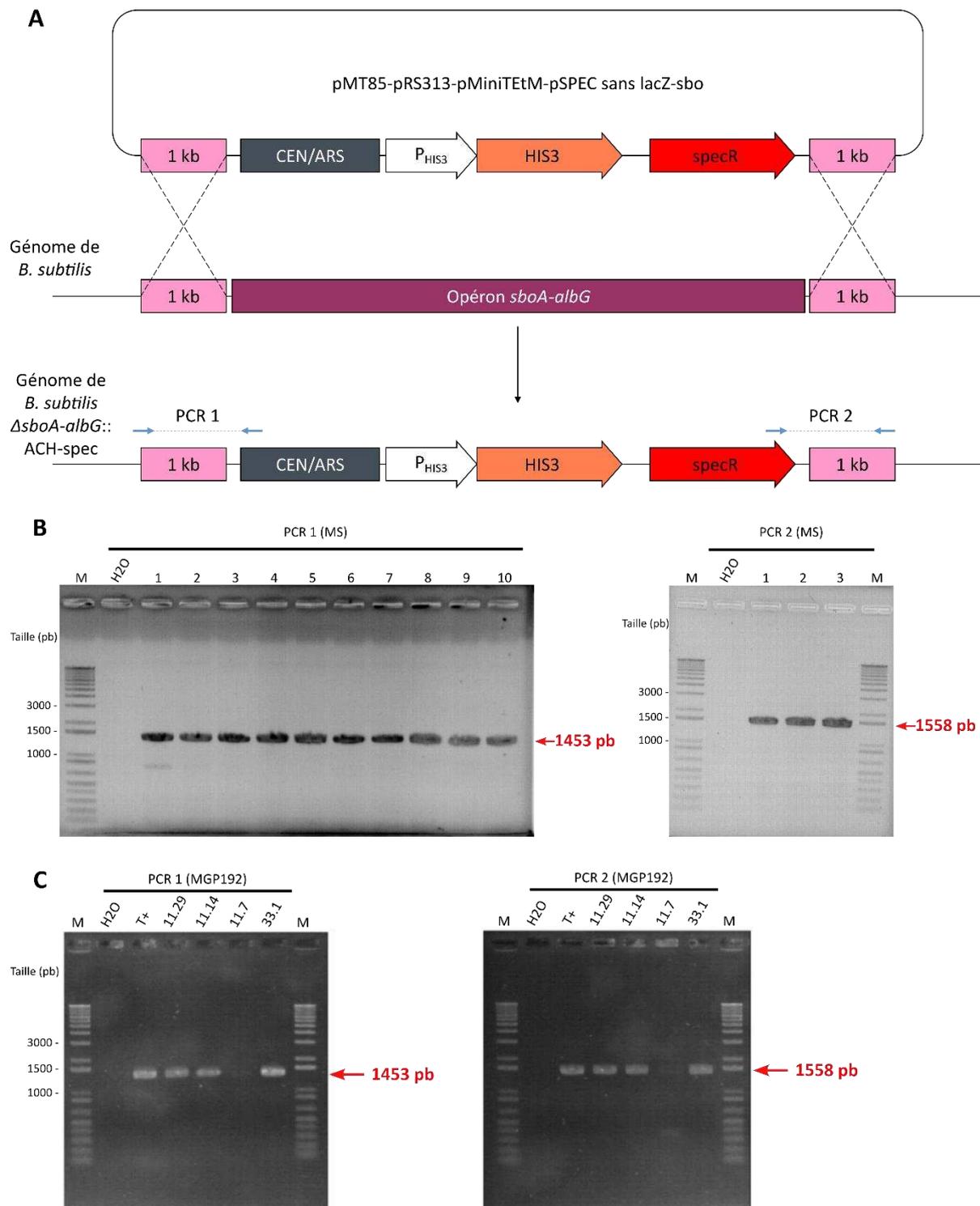
Une molécule d'ADN peut se maintenir dans la levure si et seulement si elle possède un centromère (CEN), une origine de réplication autonome (ARS), et un marqueur d'auxotrophie. Ces séquences, appelées éléments levure, devaient donc être intégrées au génome de *B. subtilis* avant ou pendant son transfert. Nous avons deux choix : (i) soit pré-marquer le génome avec ces éléments par transformation bactérienne soit (ii) amener les éléments levure au cours de la capture, sur le pansement de re-circularisation du fragment.

*B. subtilis* étant facilement transformable, nous avons choisi dans un premier temps, d'insérer au préalable ces éléments à son génome (figure 44, étape 1).

Trois souches de *B. subtilis* ont été sélectionnées pour ce travail :

- La souche MS (3,98 Mb) considérée pour nos études comme la souche de référence.
- La souche FE220 (3,96 Mb) équivalente à la souche MS mais dépourvue du système de restriction modification BsuM. Ce choix a été guidé par un ensemble de travaux du laboratoire de B. Karas montrant que la présence de systèmes de restriction modification chez les bactéries (et plus spécifiquement d'endonucléases de restriction) pouvait réduire voire inhiber complètement le transfert d'ADN après contact bactéries/levures (Karas *et al.*, 2013; Karas *et al.*, 2014). Cette souche a donc été retenue pour évaluer l'impact de la présence du système de restriction modification BsuM sur la méthode de CREasPy-Fusion.
- La souche MGP192 (2,86 Mb) qui possède un génome fortement réduit par rapport aux deux autres souches et qui permet donc d'évaluer l'impact de la taille du génome bactérien tout comme celui de la taille du fragment à cloner sur la CREasPy-Fusion.

Le premier fragment que nous souhaitions capturer correspondait à la région située entre les gènes *atpA* et *irgB* d'une taille de 145 kb pour les souches MS et FE220 et d'une taille de 125 kb pour la souche MGP192 (figure 45). Il nous fallait donc marquer cette section du génome avec les « éléments levure ». Nous avons choisi de les insérer à la place d'une région non essentielle présente dans les 3 génomes : l'opéron *sboAX-albAG* (figure 46). Cet opéron regroupe l'ensemble des gènes nécessaires à la synthèse de la subtilosine A, un lantibiotique



**Figure 47 : Insertion des éléments levure dans le génome de *B. subtilis*.** [A] Schéma de l'insertion des éléments levure à partir du plasmide pMT85-pRS313-pMiniTetM-pSPEC sans lacZ-sbo. [B] Analyse des transformants de la souche *B. subtilis* MS issus de la première expérience. L'intégration des éléments levure est validée par la PCR 1 (avec les amorces Verif\_sbo\_F et Verif\_sbo\_R2) et la PCR 2 (avec les amorces Verif\_sbo\_SPEC\_F2 et Verif\_sbo\_R), au niveau des jonctions produisant des amplicons de 1453 pb pour la PCR 1 et 1558 pb pour la PCR 2. [C] Analyse des transformants de la souche *B. subtilis* MGP 192 issus de la deuxième expérience. L'intégration des éléments levure est validée de la même manière que pour la souche MS. M: DNA Ladder 1 kb+ Invitrogen (100-12000 pb). H2O: contrôle négatif, sans ADN. T+: témoin positif, ADNg de la souche *MS* $\Delta sboA-alb::ACH-spec$ .

produit par *B. subtilis* (Babasaki *et al.*, 1985) actif contre des pathogènes parmi lesquels *Listeria monocytogenes*, *Gardnerella vaginalis*, *Streptococcus agalactiae* et *Micrococcus luteus* (Sutyak *et al.*, 2008).

Pour ce faire, nous avons utilisé le plasmide pMT85-pRS313-pMiniTetM-pSPEC sans lacZ-sbo qui porte les éléments levure ainsi qu'un marqueur de sélection *specR* conférant la résistance à la spectinomycine, le tout bordé par deux cassettes de recombinaison correspondant aux régions du génome de *B. subtilis* encadrant l'opéron *sboAX-albAG*. L'insertion des éléments levure dans le génome de *B. subtilis* est schématisée figure 47A.

Les trois souches ont été construites au cours de 3 expériences de transformations différentes. Lors du premier essai, nous avons transformé le plasmide dans les souches MS et MGP192 avec le protocole dit « simple » basé sur le milieu « Modified Competence » (MC). Les transformants ont été sélectionnés sur milieu LB solide contenant 100 µg/ml de spectinomycine. Après 24 h d'incubation à 37°C, aucune colonie n'est apparue pour la souche MGP192 alors que 34 colonies étaient visibles pour la souche MS. Dix d'entre elles ont été analysées par PCR dans le but de vérifier l'intégration des éléments levure au site désiré. Plus précisément, deux PCR ont été réalisées, une à chaque jonction (PCR 1 et PCR 2, figure 47A). Pour les 10 clones, une bande unique à la taille attendue de 1453 pb pour la PCR 1 est apparue (figure 47B). Trois clones (cl. 1, 2 et 3) ont été sélectionnés et analysés avec la PCR 2. Ils ont tous montré une bande unique à la taille de 1558 pb (figure 47B), suggérant que les éléments levure avaient bien été insérés à la place de l'opéron *sboAX-albAG* dans la souche *B. subtilis* MS.

Après cette première expérience, nous avons donc obtenu la souche  $MS\Delta sbo-alb::ACH-spec$ . Cependant, l'essai réalisé avec la souche MGP192 avait échoué. Suite à ces résultats, nous avons entamé des discussions avec nos collaborateurs qui nous ont transmis plusieurs informations importantes : (i) les souches à génomes réduits (comme MGP192) sont difficilement transformables et il est généralement plus efficace de transformer une souche de référence (comme les souches 168 ou MS) avec le plasmide d'intérêt puis d'utiliser l'ADNg de la souche de référence modifiée pour transformer les souches de *B. subtilis* à génomes réduits ; (ii) un second protocole (protocole MG1/MG2), toujours basé sur la compétence naturelle de *B. subtilis* et utilisant deux milieux très distincts, peut être appliqué aux souches à génomes réduits.

**Tableau 18 : Résultats de la transformation de *B. subtilis* MGP192 avec le protocole MG1/MG2 pour l'insertion des éléments levure.**

Conditions	Temps d'incubation en MG2 <sup>a</sup>	ADN transformé <sup>b</sup>	Nombre de CFU (après transformation)	Clones positifs / Clones analysés
				PCR simplex
1	30 min	ADNg <i>MSΔsbo-alb::ACH-spec</i>	42	2/30 <sup>c</sup> + 0/1
2		pMT85-pRS313-pMiniTetM-pSPEC sans lacZ-sbo	0	-
3		pMT85-pRS313-pMiniTetM-pSPEC sans lacZ-sbo digéré SmaI	0	-
4	1h00	ADNg <i>MSΔsbo-alb::ACH-spec</i>	1	0/1
5		pMT85-pRS313-pMiniTetM-pSPEC sans lacZ-sbo	0	-
6		pMT85-pRS313-pMiniTetM-pSPEC sans lacZ-sbo digéré SmaI	0	-
7	1h30	ADNg <i>MSΔsbo-alb::ACH-spec</i>	0	-
8		pMT85-pRS313-pMiniTetM-pSPEC sans lacZ-sbo	0	-
9		pMT85-pRS313-pMiniTetM-pSPEC sans lacZ-sbo digéré SmaI	0	-
10	2h00	ADNg <i>MSΔsbo-alb::ACH-spec</i>	0	-
11		pMT85-pRS313-pMiniTetM-pSPEC sans lacZ-sbo	0	-
12		pMT85-pRS313-pMiniTetM-pSPEC sans lacZ-sbo digéré SmaI	1	1/1
13	2h30	ADNg <i>MSΔsbo-alb::ACH-spec</i>	220	0/30 <sup>c</sup>
14		pMT85-pRS313-pMiniTetM-pSPEC sans lacZ-sbo	43	0/30 <sup>c</sup>
15		pMT85-pRS313-pMiniTetM-pSPEC sans lacZ-sbo digéré SmaI	108	0/30 <sup>c</sup>
16	3h00	ADNg <i>MSΔsbo-alb::ACH-spec</i>	0	-
17		pMT85-pRS313-pMiniTetM-pSPEC sans lacZ-sbo	60	0/30 <sup>c</sup>
18		pMT85-pRS313-pMiniTetM-pSPEC sans lacZ-sbo digéré SmaI	1	0/1

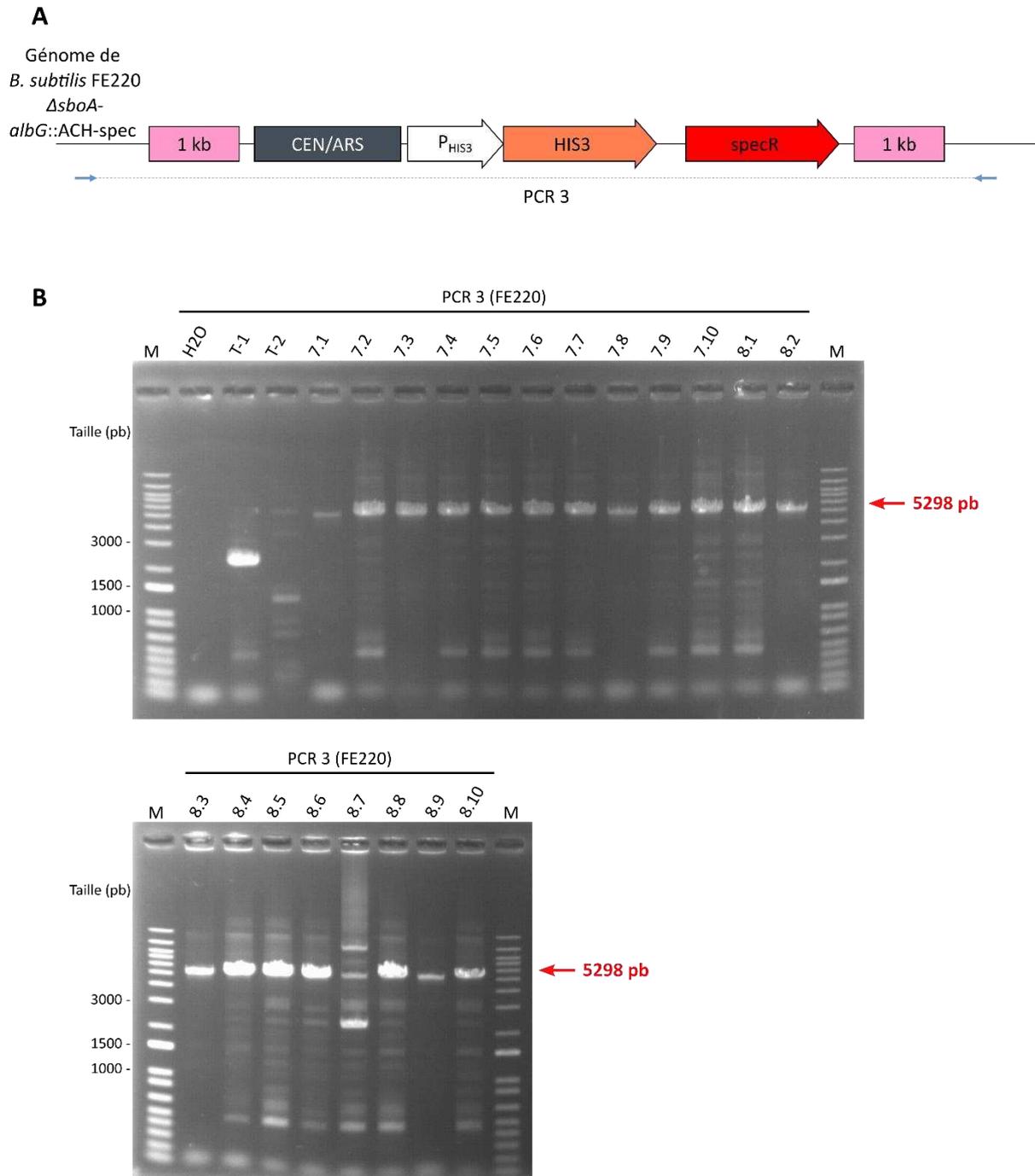
<sup>a</sup> : MG2 (milieu carencé). Composition indiquée tableau 12 du matériels et méthodes.

<sup>b</sup> : Le plasmide pMT85-pRS313-pMiniTetM-pSPEC sans lacZ-sbo a été introduit par transformation soit sous forme circulaire soit sous forme linéaire après digestion par l'enzyme SmaI. L'ADNg de la souche *MSΔsbo-alb::ACH-spec* a aussi été utilisée

<sup>c</sup> : clones analysés en pool

Une seconde expérience de transformation mettant en jeu la souche MGP192, le protocole MG1/MG2 et de l'ADNg de la souche *MSΔsbo-alb::ACH-spec* (en parallèle du plasmide seul) a donc été tentée. L'étape critique de ce protocole de transformation est le temps d'incubation dans le milieu MG2. N'ayant pas d'idées préconçues sur le temps à utiliser pour la souche MGP192, les cellules ont été mises en culture en milieu MG1 pendant un temps fixe (4h30). Puis les cellules issues de la culture MG1 ont été diluées toutes les 30 min pendant 3 h en milieu MG2 (les temps d'incubation se répartissaient ainsi de 30 min à 3 h en milieu MG2). Après cette étape, les cellules ont été mises au contact (i) d'ADNg de la souche *MSΔsbo-alb::ACH-spec*, (ii) de plasmide circulaire ou (iii) de plasmide digéré par l'enzyme de restriction *SmaI* (permettant d'isoler la cassette d'intérêt) puis étalées sur milieu LB gélosé contenant 100 µg/ml de spectinomycine.

Après 48 h à 37°C, un total de 476 colonies a été observé sur boîte (tableau 18). Dans le but de cribler un grand nombre de clones, dix pools de 3 clones ont été réalisés pour les conditions montrant un grand nombre de colonies sur milieu de sélection (*i.e.* conditions 1, 13, 14, 15 et 17). Les colonies obtenues pour les conditions 4, 12 et 18 ainsi qu'une colonie isolée de la condition 1 ont, quant à elles, été analysées individuellement (tableau 18). Au total, 154 clones ont donc été analysés mais la plupart se sont avérés être de faux positifs. En effet, sur les 50 pools de trois coloniesensemencés, seuls deux pools ont montré une croissance bactérienne (pools 11.29 et 11.14 issus de la condition 1) et sur les 4 clones individuels, seuls deux clones (cl 11.7 et 33.1 issus respectivement des conditions 1 et 12) ont montré une croissance bactérienne. Ces 4 cultures ont été analysées. Comme pour la souche *MSΔsbo-alb::ACH-spec* précédemment, l'intégration des éléments levure a été vérifiée grâce aux PCR 1 et PCR 2. Les pools 11.29 et 11.14 ainsi que le clone 33.1 ont montré le profil attendu pour les deux PCR (figure 47C). Etant donné que le clone 33.1 était un clone individuel, nous avons choisi de le sélectionner et de ne pas poursuivre avec les « pools ». A noter, que ce clone 33.1 provenait d'une condition de transformation avec le plasmide et non pas avec l'ADNg de la souche *MSΔsbo-alb::ACH-spec*. Lors de cette seconde expérience, nous avons donc enfin réussi à insérer les éléments levure dans le génome de la souche *B. subtilis* MGP192 à la place de l'opéron *sboAX-albAG* nous permettant enfin d'obtenir la souche *MGP192Δsbo-alb::ACH-spec*.



**Figure 48 : Insertion des éléments levure dans le génome de *B. subtilis* FE220. [A]** Schéma des éléments levure insérés dans le génome bactérien. **[B]** Analyse des transformants de la souche *B. subtilis* FE220. L'intégration des éléments levure est validée par une PCR utilisant les amorces screening\_ARS\_CEN\_HIS\_F et screening\_ARS\_CEN\_HIS\_R, en dehors de bras de recombinaison du plasmide (et produisant un amplicon de 5298 pb. Si les éléments levure ne sont pas insérés la PCR produit un amplicon 2229 pb. « M » : DNA Ladder 1 kb+ Invitrogen (100-12000 pb). « H2O » : contrôle négatif sans ADN. « T-1 » : témoin négatif ADNg de la souche FE220 non modifiée. « T-2 » : témoin négatif, plasmide pMT85-pRS313-pMiniTetM-pSPEC sans lacZ-sbo.

Pour la troisième et dernière expérience, la souche FE220 a été transformée avec le plasmide pMT85-pRS313-pMiniTEtM-pSPEC sans lacZ-sbo de la même manière que pour la souche MS. Après 24 h d'incubation à 37°C, plusieurs centaines de colonies sont apparues sur milieu de sélection. Vingt clones ont été analysés par PCR afin de vérifier l'intégration des éléments levure à l'endroit souhaité dans le génome. Cette fois, une seule PCR a été réalisée avec les amorces screening\_ARS\_CEN\_HIS\_F et screening\_ARS\_CEN\_HIS\_R (PCR 3) permettant de mettre en évidence l'intégration complète de la cassette en lieu et place de l'opéron *sboAX-albAG* (figure 48A). Sur les 20 clones analysés, cinq clones (cl 7.3, 7.8, 8.2, 8.3 et 8.9) ont montré un seul amplicon à la taille attendue (5298 pb, figure 48B), attestant l'obtention de la souche FE220 $\Delta$ *sbo-alb*::ACH-spec.

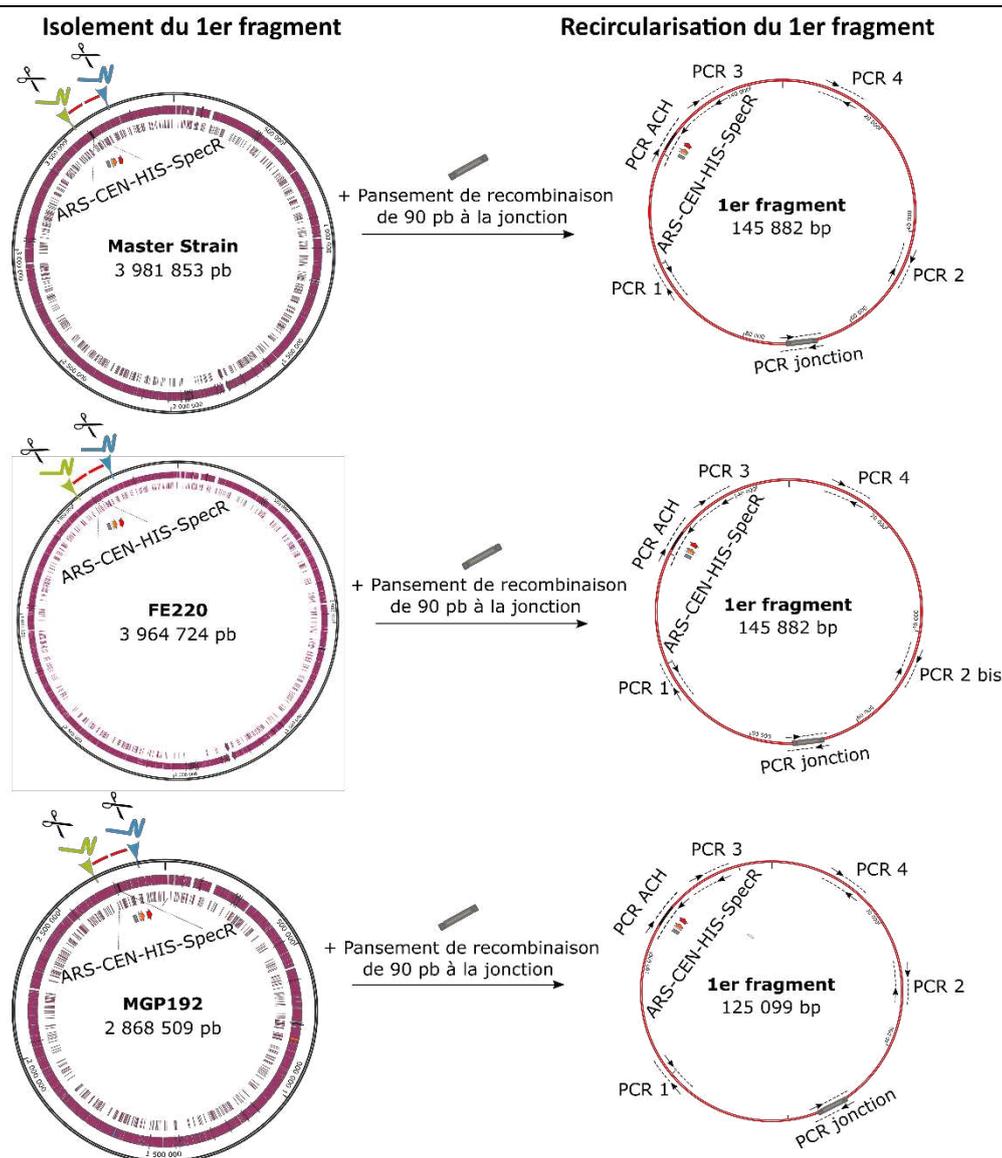
Les génomes des trois souches obtenues MS $\Delta$ *sbo-alb*::ACH-spec, MGP192 $\Delta$ *sbo-alb*::ACH-spec et FE220 $\Delta$ *sbo-alb*::ACH-spec ont tous été séquencés entièrement par la société prestataire PLASMIDSAURUS™. Nous avons donc pu valider l'intégration des « éléments levure » et du marqueur *specR* ainsi que l'absence de remaniements sur le reste des génomes.

### 3.2. Préparation des levures avec le système CRISPR-Cas9

Pour capturer un fragment de génome de *B. subtilis*, les levures doivent héberger et exprimer un système CRISPR-Cas9 fonctionnel et spécifique du fragment à cloner (figure 44, étape 2). Pour ce travail, nous avons sélectionné deux souches levures (VL6-48N et W303a) et les avons toutes deux pré-chargées par transformation (protocole à l'acétate de lithium) avec le système CRISPR-Cas9. L'introduction de ce système s'est faite en deux temps. Le plasmide p414-TEF1p-Cas9-CYC1t (DiCarlo *et al.*, 2013) qui permet l'expression de la Cas9 et qui ne varie pas au cours des expériences a d'abord été introduit dans ces souches de façon à sélectionner un clone VL6-48N/Cas9 et un clone W303/Cas9 corrects qui ont été stockés au congélateur à -80°C. En fonction du fragment à capturer, ces clones ont ensuite été transformés avec des plasmides p426-gRNA « double target » qui permettent d'exprimer simultanément deux ARNg différents, chacun d'entre eux ciblant une région spécifique du génome bactérien. Les transformants issus de cette deuxième transformation ont toujours été analysés de la même manière, *i.e.* par PCR en utilisant le couple d'amorces dites universelles T3 et T7. Ce couple d'amorces s'hybridant à la fois sur le plasmide p414-TEF1p-Cas9-CYC1t et p426-gRNA double

**Tableau 19 : Résumé des expériences de capture, par CReasPy-Fusion, de fragments de génome de *B. subtilis* pré-marqués avec les éléments levure.**

Fragment de génome	Souches <i>Bacillus subtilis</i>	CFU			Clones positifs / Clones analysés	
		1 <sup>er</sup> expérience	2 <sup>nd</sup> expérience	3 <sup>eme</sup> expérience	PCR simplex	PFGE
1 <sup>er</sup> fragment	MGP192Δ <i>sbo-alb</i> ::ACH-Spec	78	98	49	28/78 (Exp 1)	2/2
	MSΔ <i>sbo-alb</i> ::ACH-Spec	2	0	-	2/2	2/2
	FE220Δ <i>sbo-alb</i> ::ACH-Spec	-	0	2	2/2	2/2
2 <sup>nd</sup> fragment	MGP192Δ <i>sbo-alb</i> ::ACH-Spec	38	119	-	5/38 (Exp 1)	4/4
	MSΔ <i>sbo-alb</i> ::ACH-Spec	0	0	-	-	-
	FE220Δ <i>sbo-alb</i> ::ACH-Spec	-	0	-	-	-



**Figure 49 : Capture par CReasPy-Fusion, du premier fragment de génome de *B. subtilis* marqué avec les éléments levure. Isolement du fragment par deux duplex ARNg-Cas9. Re-circularisation du fragment par le système de recombinaison homologue de la levure. Le pansement de recombinaison de 90 pb, co-transformé, est utilisé par la levure pour réparer les cassures double brin et former un fragment circulaire. ACH : ARS CEN HIS**

target permet en effet de vérifier la présence simultanée des deux plasmides. Les clones avec le profil attendu en PCR, c'est-à-dire montrant des bandes sur gel d'agarose à 4902 pb (plasmide d'expression de la Cas9) et à 713 pb (plasmide d'expression des ARNg), ont été conservés et analysés par séquençage au niveau des régions des ARNg.

Au cours de ce travail, plusieurs clones levure ont été construits et obtenus. Les clones se distinguaient les uns des autres par les ARNg qu'ils portaient et exprimaient, spécifiques de la région du génome de *B. subtilis* à capturer.

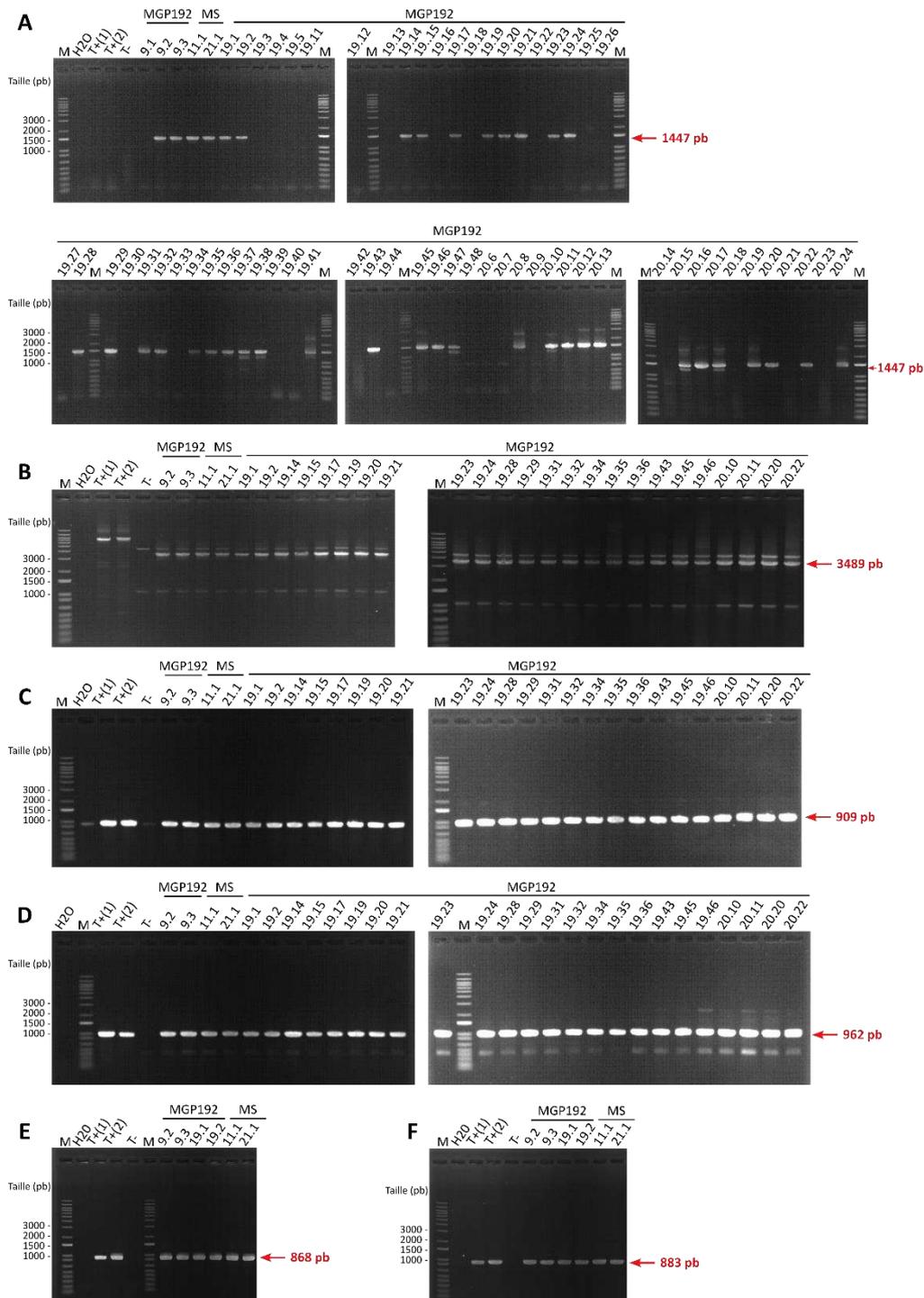
### 3.3. Preuve de concept : Capture d'un premier fragment de génome d'une taille de 125 kb à 150 kb pour 3 souches de *B. subtilis* différentes

Pour démontrer que la CReasPy-Fusion était une méthode adaptée à la capture spécifique de fragments de génome de *B. subtilis*, une preuve de concept a été réalisée avec l'isolement d'un premier fragment de génome provenant de trois souches différentes. Pour les premières expériences, nous avons choisi de capturer la région du génome entre les gènes *atpA* et *IrgB* (figure 45). Elle a une taille de 145 kb pour les souches *MSΔsbo-alb::ACH-spec* et *FE220Δsbo-alb::ACH-spec* et 125 kb pour la souche minimale *MGP192Δsbo-alb::ACH-spec*.

Pour y parvenir, nous avons donc commencé par pré-charger la souche levure VL6-48N/Cas9 avec le plasmide p426-gRNA « double target » permettant l'expression de deux ARNg ciblant spécifiquement les gènes *atpA* et *ywbE*, de part et d'autre de la région à capturer. Parallèlement à ce travail, nous avons produit, par hybridation de deux grands oligonucléotides, un pansement de recombinaison de 90 pb correspondant aux deux bornes du fragment à capturer (figure 44, rectangle bleu). Ce pansement de recombinaison est ajouté lors du contact sphéroplast/protoplastes, juste avant addition de PEG (figure 44, étape 3).

Pour capturer le premier fragment de génome de *B. subtilis* à partir de 3 souches différentes, 3 expériences indépendantes ont été menées et sont résumées dans le tableau 19.

Lors de la première expérience, nous avons mis en contact des protoplastes des souches de *B. subtilis* *MSΔsbo-alb::ACH-spec* ou *MGP192Δsbo-alb::ACH-spec* avec des sphéroplast



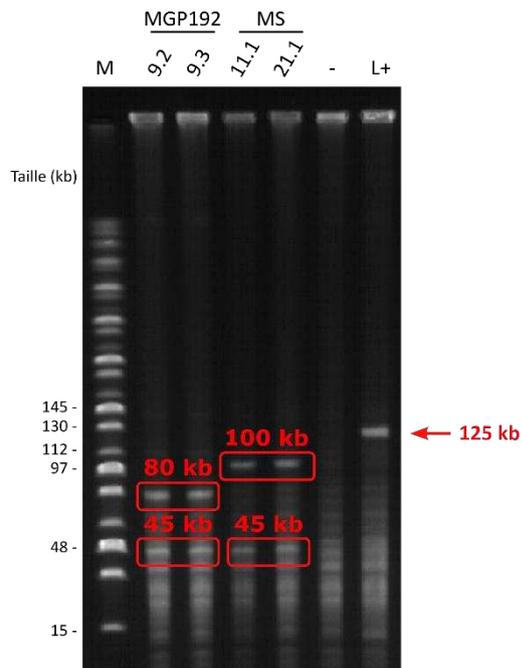
**Figure 50 : Analyse des transformants levure propageant le premier fragment de génome des souches *B. subtilis* MGP192 (125 kb) et MS (145 kb).** La présence du fragment est contrôlée avec 6 couples d'amorces distribués de façon homogène le long de ce dernier. **[A]** La re-circularisation du fragment est confirmée avec les amorces Verif\_Frgt5-jonction\_F/Verif\_Frgt5-jonction\_R produisant un amplicon de 1447 pb uniquement possible en cas de jonction des extrémités. **[B]** La présence des éléments levure est confirmée avec le couple d'amorces ScreenPST\_Bsu\_sbo\_F/ScreenPST\_Bsu\_sbo\_R qui produit un amplicon de 3489 pb. L'intégrité du fragment est évaluée avec 4 couples d'amorces produisant des amplicons aux tailles de **[C]** 909 pb (PCR 1) **[D]** 962 pb (PCR 2) **[E]** 868 pb (PCR 3) **[F]** 883 pb (PCR 4). M: DNA Ladder 1 kb+ Invitrogen (100-12000 pb). H2O: contrôle négatif, sans ADN. T+(1): témoin positif, avec l'ADNg de la souche MGP192. T+(2): témoin positif, avec l'ADNg de la souche MS. T- : témoin négatif avec l'ADNg de *S. cerevisiae* VL6-48N.

VL6-48N pré-chargée avec le système CRISPR-Cas9. Les transformants ont été sélectionnés sur milieu gélosé SD-HIS.

- Pour la souche MGP192, un total de 78 transformants a été obtenu, ils ont tous été analysés. La présence du fragment a été recherchée à l'aide de 6 PCR indépendantes mettant en jeu chacune une paire d'amorces spécifiques. La 1<sup>ère</sup> PCR, appelée PCR jonction, réalisée avec les amorces Verif\_Frgt5-jonction\_F/Verif\_Frgt5-jonction\_R, permet de générer un amplicon de 1447 pb si le fragment a été correctement circularisé dans la levure (figure 49). La 2<sup>ème</sup> PCR, appelée PCR ARS-CEN-HIS, réalisée avec les amorces ScreenPST\_Bsu\_sbo\_F/ScreenPST\_Bsu\_sbo\_R permet de générer un amplicon de 3489 pb si les éléments levure et le marqueur spectinomycine pré-intégrés au fragment cible (cf paragraphe 3.1 du chapitre 1) sont toujours présents. Enfin, les PCR 1 à 4, réalisées à l'aide de couples d'amorces distribués tout le long du fragment cible (figure 49), permettent de générer des amplicons de 909 pb (PCR 1), 962 pb (PCR 2), 868 pb (PCR 3) et 883 pb (PCR 4) respectivement si le fragment est intègre, c'est-à-dire sans réarrangement majeur. Sur les 78 clones analysés, 26 ont montré le profil attendu pour les PCR au niveau de la jonction, des éléments levure et les PCR 1 et 2. Puis seulement 4 d'entre eux ont été analysés avec les PCR 3 et 4 et ils ont tous montré les profils attendus pour ces amplifications (figure 50).

Pour la souche MS, seulement 2 colonies sont apparues sur boîte. Les 2 clones ont été analysés par PCR comme décrit précédemment pour la souche MGP192, et ont présenté tous deux le profil attendu pour les 6 PCR (figure 50). Afin de s'assurer de la présence du fragment dans son intégralité et de l'absence de grands remaniements, une seconde analyse a été faite par électrophorèse en champ pulsé (PFGE). Pour cela, deux transformants levure (pour chaque souche de *B. subtilis*) ont été choisis et la taille de la molécule d'ADN clonée a été évaluée. Le criblage s'est fait en utilisant l'enzyme de restriction AvrII dont le site est présent une fois sur le fragment cible et une fois sur le pansement de recombinaison (au niveau de HIS3). L'hydrolyse du fragment par AvrII génère donc un profil à 2 bandes d'environ 80 kb et 45 kb pour le fragment issu de la souche MGP192 et d'environ 100 kb et 45 kb pour celui issu de la souche MS. Tous les clones de levure analysés ont présenté le profil à deux bandes avec les tailles attendues (figure 51).

Cette première expérience, nous a permis de valider, pour la première fois, notre capacité à capturer de façon spécifique un fragment de génome de *B. subtilis* par CReasPy-Fusion



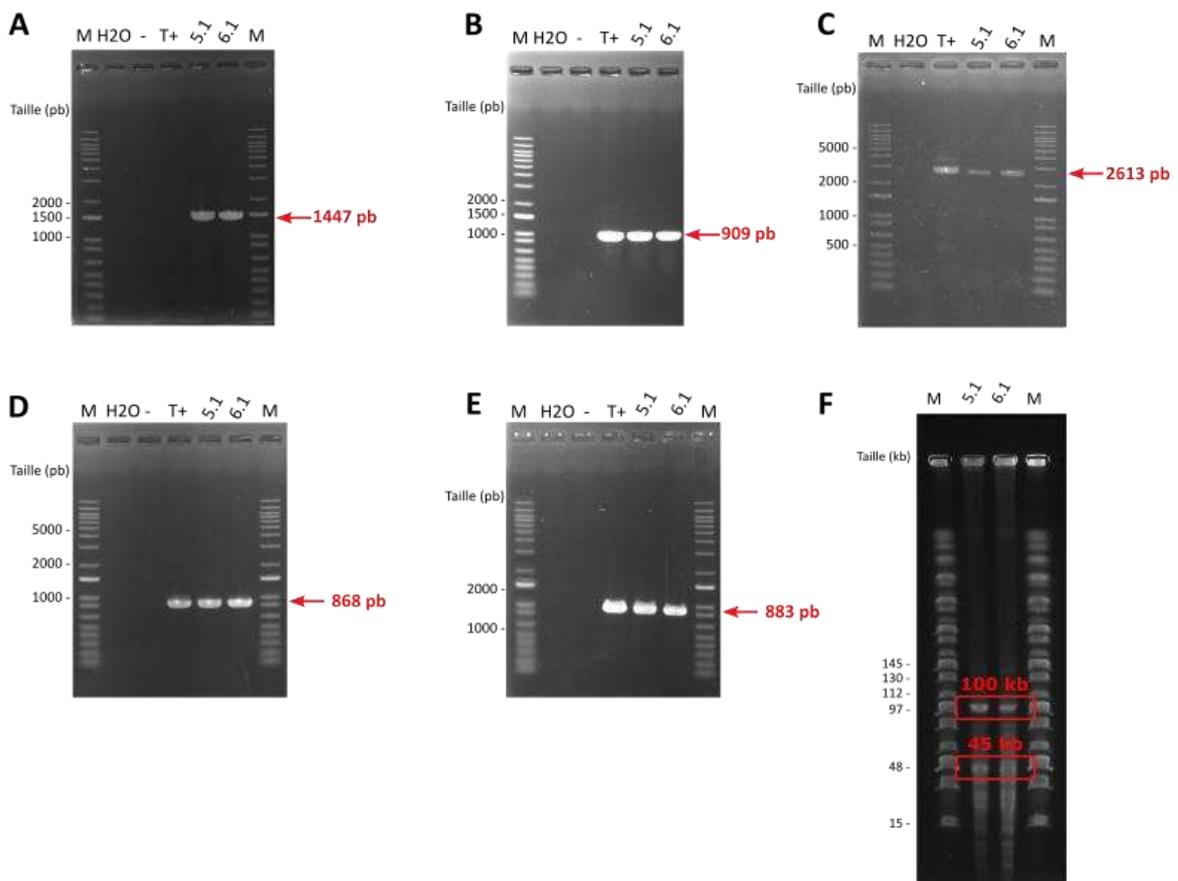
**Figure 51 : Analyse de l'intégrité et la taille du premier fragment de génome capturé à partir des souches *B. subtilis* MGP192 (125 kb) et MS (145 kb) par restriction enzymatique et PFGE.** La digestion par l'enzyme *AvrII* linéarise le fragment en une molécule de 125 kb pour la souche MGP192 et 145 kb pour la souche MS. La présence d'un site de restriction additionnel au niveau des éléments levure pré-intégrés induit un profil à 2 bandes d'environ 80 kb et 45 kb pour le fragment issu de la souche MGP192 et 100 kb et 45 kb pour celui issu des souches MS. M: Mid range PFG ladder NEB (15.0 - 242.5 kbp). - : témoin négatif, ADNg de *S. cerevisiae* VL6-48N. L+ : témoin positif, ADNg de transformant levure hébergeant le fragment capturé par TAR-cloning.

prouvant ainsi que nous étions capables (i) de transférer par contact cellules/cellules le génome entier de *B. subtilis* des souches MGP192 (2,868 Mb) et MS (3,981 Mb) et (ii) d'en isoler spécifiquement un fragment. Il est important de noter, cependant, que la capture semble plus efficace pour la souche MGP192 dont le génome et le fragment capturé (125 kb) sont de plus petites tailles que pour la souche MS, suggérant un impact de la taille du génome bactérien initial ou une limite de la taille du fragment qu'il est possible de cloner.

Afin d'évaluer l'impact de la présence du système de restriction modification BsuM, nous avons réalisé une seconde expérience. Cette fois, nous avons mis en contact des protoplastes des souches de *B. subtilis*  $MS\Delta sbo-alb::ACH-spec$ ,  $MGP192\Delta sbo-alb::ACH-spec$  et  $FE220\Delta sbo-alb::ACH-spec$  avec les sphéroplastes de VL6-48N pré-chargée avec les systèmes CRISPR-Cas9 adéquats ; la souche *B. subtilis* MGP192 servant ici de contrôle positif pour la capture d'un fragment de génome.

Lors de cette expérience, aucune colonie n'est apparue sur les conditions avec les souches *B. subtilis* FE220 et MS. En revanche, 98 colonies ont été obtenues pour les conditions avec la souche *B. subtilis* MGP192. Vingt clones ont été analysés par PCR avec le couple d'amorces correspondant à la PCR 2. Sur les 20 clones analysés, 9 ont montré le profil attendu avec un amplicon à 962 pb. L'objectif de cette expérience était de questionner l'impact de la présence du système de restriction modification BsuM sur la capture d'un fragment de génome par CReasPy-Fusion en comparant l'efficacité de capture entre la souche MS et la souche FE220. Comme aucune colonie n'est apparue sur ces conditions, l'analyse des transformants n'a pas été poursuivie. Ces résultats ont montré une nouvelle fois que la capture de fragment de génome est plus probable avec la souche MGP192 (2,86 Mb) dont le génome est le plus petit par rapport aux souches FE220 (3,96 Mb) et MS (3,98 Mb). Ces observations soulignent ici encore l'impact de la taille du génome initial et/ou la taille de la molécule clonée, sur l'efficacité de capture. Ces premiers résultats ont suggéré aussi que la présence ou l'absence du système BsuM ne semblait pas impacter la capture d'un fragment de génome dans la levure par la méthode de CReasPy-Fusion.

Lors de cette deuxième expérience, aucun transformant n'a pu être obtenu avec les souches FE220 et MS. Or, lors de la première expérience, 2 transformants hébergeant un fragment de génome issu de la souche MS avaient pu être obtenus. La souche FE220 étant similaire à la souche MS, il semblait donc possible d'en capturer un fragment de génome. Nous



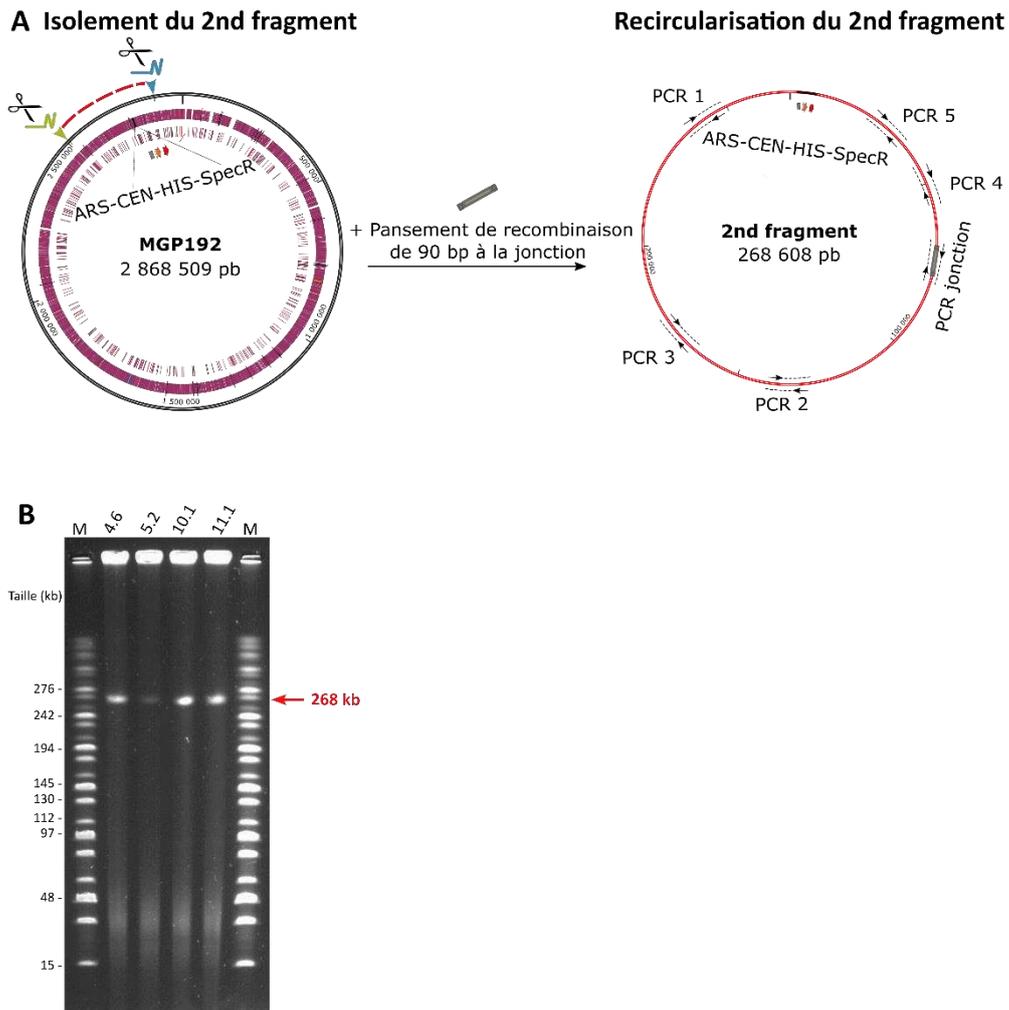
**Figure 52 : Analyse des transformants levure propageant le premier fragment de génome de la souche *B. subtilis* FE220 (145 kb).** La présence du fragment est contrôlée avec 6 couples d'amorces distribués de façon homogène le long de ce dernier. **[A]** La re-circularisation du fragment est confirmée avec les amorces Verif\_Frgt5-jonction\_F/Verif\_Frgt5-jonction\_R produisant un amplicon de 1447 pb uniquement possible en cas jonction des extrémités. L'intégrité du fragment est évaluée avec 4 couples d'amorces produisant des amplicons au tailles de **[B]** 909 pb (PCR 1) **[C]** 2613 pb (PCR 2bis) **[D]** 868 pb (PCR 3) **[E]** 883 pb (PCR 4). M: DNA Ladder 1 kb+ Invitrogen (100-12000 pb). H2O: contrôle négatif, sans ADN. - : témoin négatif avec l'ADNg de *S. cerevisiae* VL6-48N. T+ : témoin positif, avec l'ADNg de la souche FE220. **[F]** L'intégrité et la taille du fragment sont validés par restriction enzymatique et analyse en PFGE. La digestion par l'enzyme AvrII linéarise le fragment en une molécule de 145 kb. La présence d'un site de restriction additionnel au niveau des éléments levure pré-intégrés induit un profil à 2 bandes d'environ 100 kb et 45 kb. « M »: Mid range PFG ladder NEB (15.0 - 242.5 kbp).

avons donc réalisé une troisième expérience pour tenter de capturer un fragment de génome issu de la souche FE220.

Pour cette tentative, nous avons mis en contact des protoplastes des souches de *B. subtilis* FE220 $\Delta$ sbo-alb::ACH-spec et MGP192 $\Delta$ sbo-alb::ACH-spec (contrôle positif) avec des sphéroplastes VL6-48N pré-chargée avec les systèmes CRISPR-Cas9. Deux colonies sont apparues sur boîte de sélection SD-HIS pour la souche FE220 (49 pour la souche MGP192, tableau 19). Les 2 clones ont été analysés par PCR en utilisant 5 des 6 couples d'amorces conçus pour MGP192 (PCR 1, PCR 2, PCR 3, PCR 4 et PCR jonction, figure 49). Toutes les PCR se sont avérées positives (figure 52), exceptée la PCR 2, alors que le contrôle positif avec l'ADNg de la souche FE220 montrait bien un amplicon à la taille attendue pour cette PCR (962 pb). Afin de contrôler cette région du fragment, nous avons conçu un autre couple d'amorce (PCR 2bis) qui permettait de produire un amplicon de 2 613 pb englobant la région correspondant à la PCR 2 (962 pb). Étonnamment, la PCR 2bis s'est révélée positive pour les deux clones, indiquant la présence probable du fragment FE220 de 145 kb dans la levure (figure 52C). Ces résultats ont été confirmés lors de l'analyse des clones en PFGE. Les 2 clones présentaient en effet un profil à deux bandes (~100 kb et ~45 kb, figure 52F) correspondant au profil attendu lors d'une digestion du fragment par AvrII. Avec cette expérience, nous avons enfin réussi à capturer par CREasPy-Fusion un fragment de génome d'environ 145 kb issu de la souche FE220.

A ce stade, nous n'avons pas de réelles hypothèses pour expliquer le résultat de la PCR 2 et il est possible que la séquence du fragment d'environ 145 kb cloné dans la levure ne soit pas identique à la séquence du génome de la souche FE220 au niveau de ce locus. Il s'agit peut-être d'une région variable du génome sur laquelle une des deux amorces de la PCR 2 ne s'hybride pas correctement. Pour tenter de comprendre ce résultat, l'amplicon issu de la PCR 2bis qui englobe la région correspondant à la PCR 2 (962 pb) sera envoyé au séquençage prochainement.

L'ensemble de ces résultats a confirmé que (i) nous étions bien capables de transférer le génome de *B. subtilis* dans la levure et d'en capturer un fragment spécifique à partir de 3 souches différentes, (ii) l'efficacité de capture est bien supérieure pour la souche MGP192 (2,86 Mb) possédant le plus petit génome, et (iii) l'absence du système de restriction modification BsuM (souche FE220) n'augmente pas l'efficacité de capture de fragment.



**Figure 53 : Capture par CReasPy-Fusion, du second fragment de génome de *B. subtilis* marqué avec les éléments levure. [A] Isolement du fragment par deux duplex ARNg-Cas9. Re-circularisation du fragment par le système de recombinaison homologue de la levure. [B] L'intégrité et la taille du fragment sont validés par restriction enzymatique et analyse en PFGE. La digestion par l'enzyme FseI linéarise le fragment en une molécule de 268 kb. M: Mid range PFG ladder NEB (15.0 - 242.5 kbp).**

### 3.4. Capture d'un deuxième fragment de génome d'une taille supérieure à 250 kb

Dans la section précédente, nous avons réussi à capturer dans la levure, de façon spécifique, un premier fragment de génome de *B. subtilis* d'une taille de 145 kb pour les souches MS et FE220 et 125 kb pour la souche MGP192.

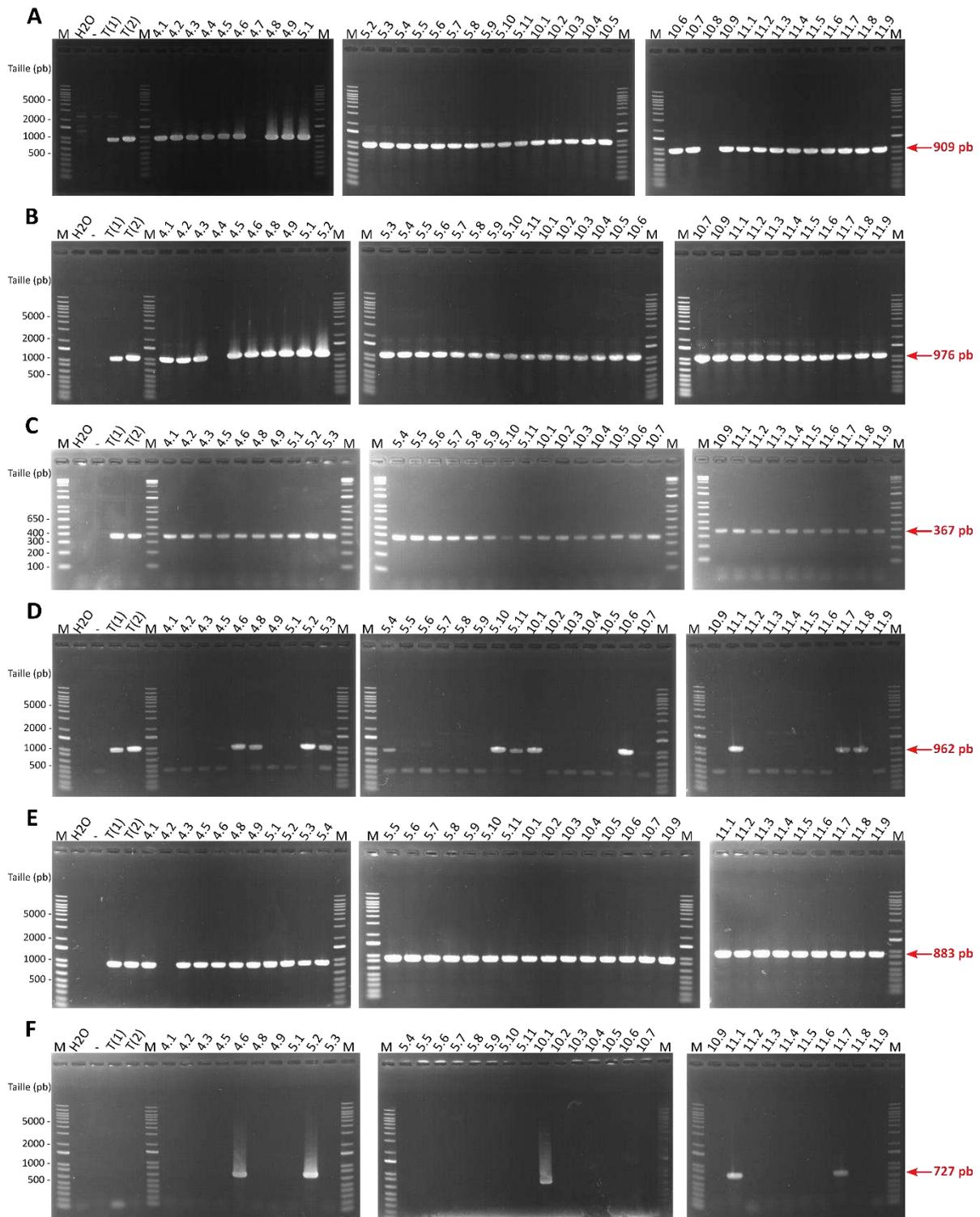
Afin de montrer la robustesse et la répétabilité de la méthode de CReasPy-Fusion, nous avons choisi de tenter la capture d'un second fragment de génome de plus grande taille, correspondant à la région entre les gènes *yvcB* et *irgB* (figure 53A). Ce fragment a une taille de 350 kb pour les souches *MSΔsbo-alb::ACH-spec* et *FE220Δsbo-alb::ACH-spec* et 268 kb pour la souche *MGP192Δsbo-alb::ACH-spec*. Il est important de noter que l'entièreté du premier fragment (de 125 kb ou 145 kb) fait partie de ce deuxième fragment.

La zone à capturer étant différente de celle de la première expérience, nous avons dû pré-charger la souche levure VL6-48N/Cas9 avec le système CRISPR-Cas9 ciblant spécifiquement les gènes *yvcB* et *ywbE* de part et d'autre de la nouvelle région à cloner. Nous avons également dû produire un nouveau pansement de recombinaison de 90 pb correspondant aux deux bornes de ce second fragment. Deux expériences indépendantes ont été menées, elles sont résumées dans le tableau 19.

Lors de la première expérience, des protoplastes des souches de *B. subtilis* *MSΔsbo-alb::ACH-spec* et *MGP192Δsbo-alb::ACH-spec* ont été mis au contact de sphéroplastes VL6-48N levure portant le système CRISPR-Cas9 spécifique du second fragment puis les transformants ont été sélectionnés sur milieu solide SD-HIS.

Concernant la souche MS, aucune colonie n'est apparue sur milieu de sélection SD-HIS. En revanche, pour la souche MGP192, un total de 38 transformants a été obtenu. Ces clones ont tous été analysés à l'aide d'une série de 6 PCR indépendantes mettant en jeu chacune une paire d'amorces spécifiques.

Les PCR 1 à 5 réalisées à l'aide de couples d'amorces distribués de façon homogène le long du fragment cible (figure 53A) permettent de générer des amplicons de 909 pb (PCR 1), 976 pb (PCR 2), 367 pb (PCR 3), 962 pb (PCR 4) et 883 pb (PCR 5) respectivement si le fragment est présent et intègre, c'est-à-dire sans réarrangement majeur. La « PCR jonction » réalisée avec les amorces *Verif\_Frgt5-jonction\_F/Verif\_Fgt4\_jonction\_R* permet de générer un



**Figure 54 : Analyse des transformants levure propageant le deuxième fragment de génome de la souche *B. subtilis* MGP192 (268 kb).** La présence du fragment est contrôlée avec 6 couples d'amorces distribués de façon homogène le long de ce dernier. L'intégrité du fragment est évaluée avec 5 couples produisant des amplicons au tailles de **[A] 909 pb (PCR 1) [B] 976 pb (PCR 2) [C] 367 pb (PCR 3) [D] 962 pb (PCR 4) [E] 883 pb (PCR 5)**. La re-circularisation du fragment est confirmée avec une paire d'amorces produisant un amplicon de 727 pb uniquement possible en cas jonction des extrémités. M: DNA Ladder 1 kb+ Invitrogen (100-12000 pb). H2O : contrôle négatif sans ADN. - : témoin négatif avec l'ADNg de *S. cerevisiae* VL6-48N. T(1) : témoin positif avec l'ADNg de la souche MGP192. T(2) : témoin positif avec l'ADNg de la souche MS.

amplicon de 727 pb si le fragment a été correctement circularisé dans la levure (figure 53A). Sur les 38 clones levure analysés, 5 ont montré le profil attendu pour toutes les PCR (figure 54 A, B, C, D, E, F). Quatre d'entre eux (cl. 4.6, 5.2, 10.1 et 11.1) ont été sélectionnés afin d'évaluer la taille du fragment cloné par PFGE. Les 4 clones analysés ont tous présenté un profil identique au profil théorique attendu après hydrolyse du fragment par FseI avec une bande à 268 kb (figure 53B).

Ces résultats ont montré une nouvelle fois que nous étions capables de transférer par contact cellules/cellules le génome de la souche MGP192 (2,86 Mb) dans la levure et d'en isoler spécifiquement un fragment, cette fois de plus grande taille (268 kb, soit environ deux fois la taille du précédent). Ils ont mis également à nouveau en avant nos difficultés à travailler avec des souches possédant des génomes plus grands.

Afin de s'assurer qu'aucune erreur technique n'avait été commise avec la souche MS, cette expérience a été répétée une deuxième fois. Des protoplastes des souches de *B. subtilis* MS $\Delta$ sbo-*alb*::ACH-spec, mais également FE220 $\Delta$ sbo-*alb*::ACH-spec et MGP192 $\Delta$ sbo-*alb*::ACH-spec (contrôle positif) ont donc à nouveau été mis en contact avec des sphéroplastes VL6-48N pré-chargées avec le système CRISPR-Cas9 spécifique au second fragment. Comme attendu, aucune colonie n'est apparue sur boîte de sélection SD-HIS pour les conditions mettant en jeu « *B. subtilis* MS » (même résultat pour « *B. subtilis* FE220 ») confirmant les résultats obtenus lors de la première expérience. Pour la souche MGP192, un total de 119 transformants a été obtenu, dont 20 ont été analysés par PCR à l'aide des amorces de la PCR 4. Sur les 20 clones analysés, 7 ont montré le profil attendu. L'objectif de cette expérience était de capturer un second fragment de génome provenant des souches MS (ou FE220). Comme aucun clone n'a pu être obtenu dans ces conditions, l'analyse des transformants issus des conditions avec la MGP192 n'a pas été poursuivie.

Ces résultats démontraient clairement que nous n'étions pas capables de capturer un fragment de génome de plus grande taille à partir des souches MS et FE220, alors que nous n'avions rencontré aucune difficulté avec la souche MGP192. Cela nous a amené à reformuler des hypothèses précédemment émises. En effet, après la capture du premier fragment de génome (cf. paragraphe 3.3), nous avons émis une première hypothèse selon laquelle la taille du fragment à capturer était limitante car peu de transformants hébergeant un fragment



d'une taille de 150 kb avaient été obtenus. Or ici nous avons réussi à capturer un fragment d'une taille de 268 kb à partir de la souche MGP192, indiquant que des fragments plus grands de *B. subtilis* peuvent être capturés dans la levure par CREasPy-Fusion. Nous émettons désormais comme hypothèse que plus que la taille du fragment à capturer, c'est surtout la taille initiale du génome bactérien et sa composition en gènes et/ou séquences répétées qui constituent une ou des barrière(s) à la capture de fragment dans la levure *S. cerevisiae*.

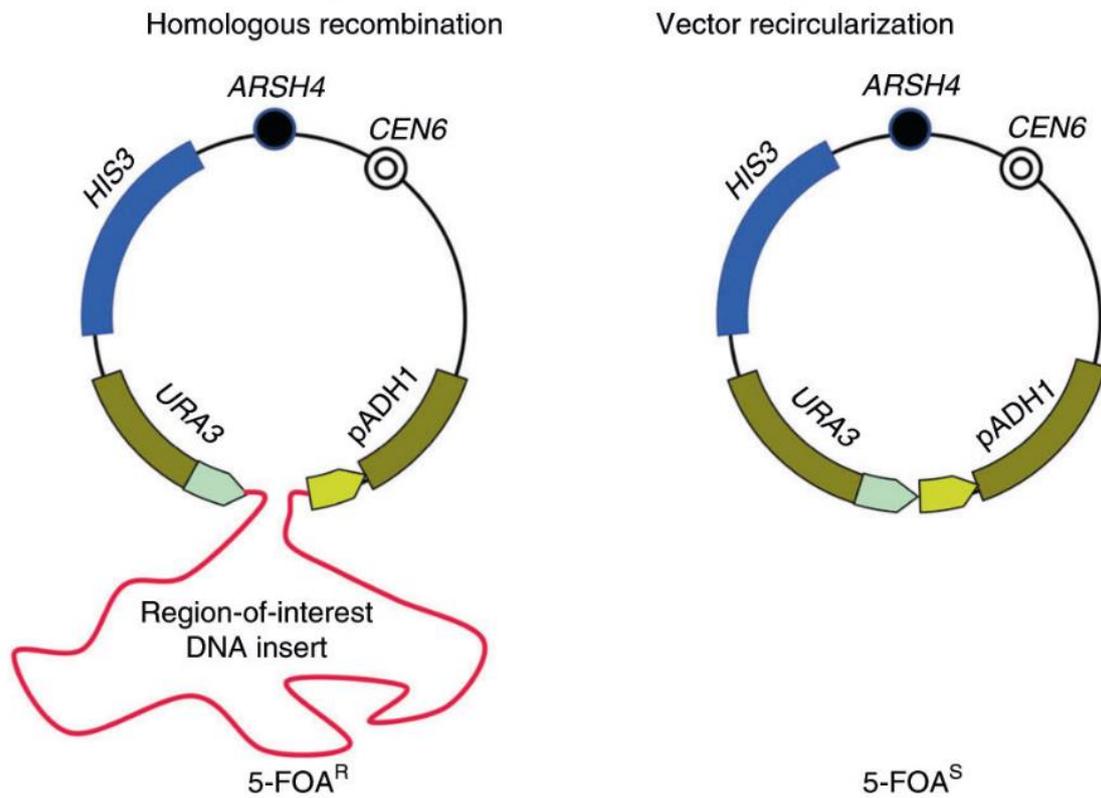
#### 4. Etape D : Capture d'un fragment de génome de *B. subtilis* non marqué par contre sélection au 5'FOA

Lors des expériences précédentes, nous avons réussi à capturer par CREasPy-Fusion plusieurs fragments de génome à partir de souches de *B. subtilis* pré-marquées avec les éléments levure. Cependant, dans cette approche, nous nous sommes rendu compte que l'étape d'insertion des éléments levure dans le génome bactérien par transformation pouvait être critique. En effet, certaines espèces/souches bactériennes sont peu transformables, comme c'est le cas par exemple de la souche minimale MGP192 par exemple, voire non transformables. Nous avons alors réfléchi à une seconde version du protocole qui permettrait de contourner l'étape de transformation et présenterait les avantages suivants : (i) rendre la méthodologie accessible à un plus grand nombre d'espèces et (ii) gagner du temps en réduisant le nombre d'étapes.

Nous avons entrepris de mettre en place cette nouvelle version de la CREasPy-Fusion en utilisant une souche de *B. subtilis* non marquée avec les éléments levure.

##### 4.1. Préparation du pansement de recombinaison contenant les éléments levure

Pour qu'un fragment de génome puisse se maintenir dans *S. cerevisiae*, les éléments levure (ARS, CEN et marqueur d'auxotrophie) doivent être présents sur ce fragment. S'ils ne sont pas introduits dans le génome bactérien par transformation avant le contact cellules/cellules, alors ils doivent être apportés et insérés au fragment à capturer pendant le contact cellules/cellules. La seule possibilité est donc qu'ils figurent sur le pansement de recombinaison qui permet de circulariser le fragment cible. Comme des travaux ont montré qu'une telle matrice de réparation pouvait parfois se circulariser sur elle-même (Noskov *et al.*,



**Figure 55 : Schéma du principe de la contre sélection par le 5-FOA.** Ce schéma provient de la publication de Kouprina et Larionov de 2020 (Kouprina *et al.*, 2020). Il illustre le principe de la contre sélection au 5'-FOA. Le fragment d'intérêt à capturer est représenté en rouge. Lorsque le vecteur pARS-VN entraîne la re-circularisation du fragment à capturer par recombinaison homologue (homologous recombination), les clones levures propageant cette molécule sont résistants au 5-FAO (5-FOA<sup>R</sup>). Si le vecteur se re-circularise seul (vector recircularisation), le gène URA3 est exprimé et les clones le propageant sont sensibles au 5-FOA (5-FOA<sup>S</sup>) et meurent.

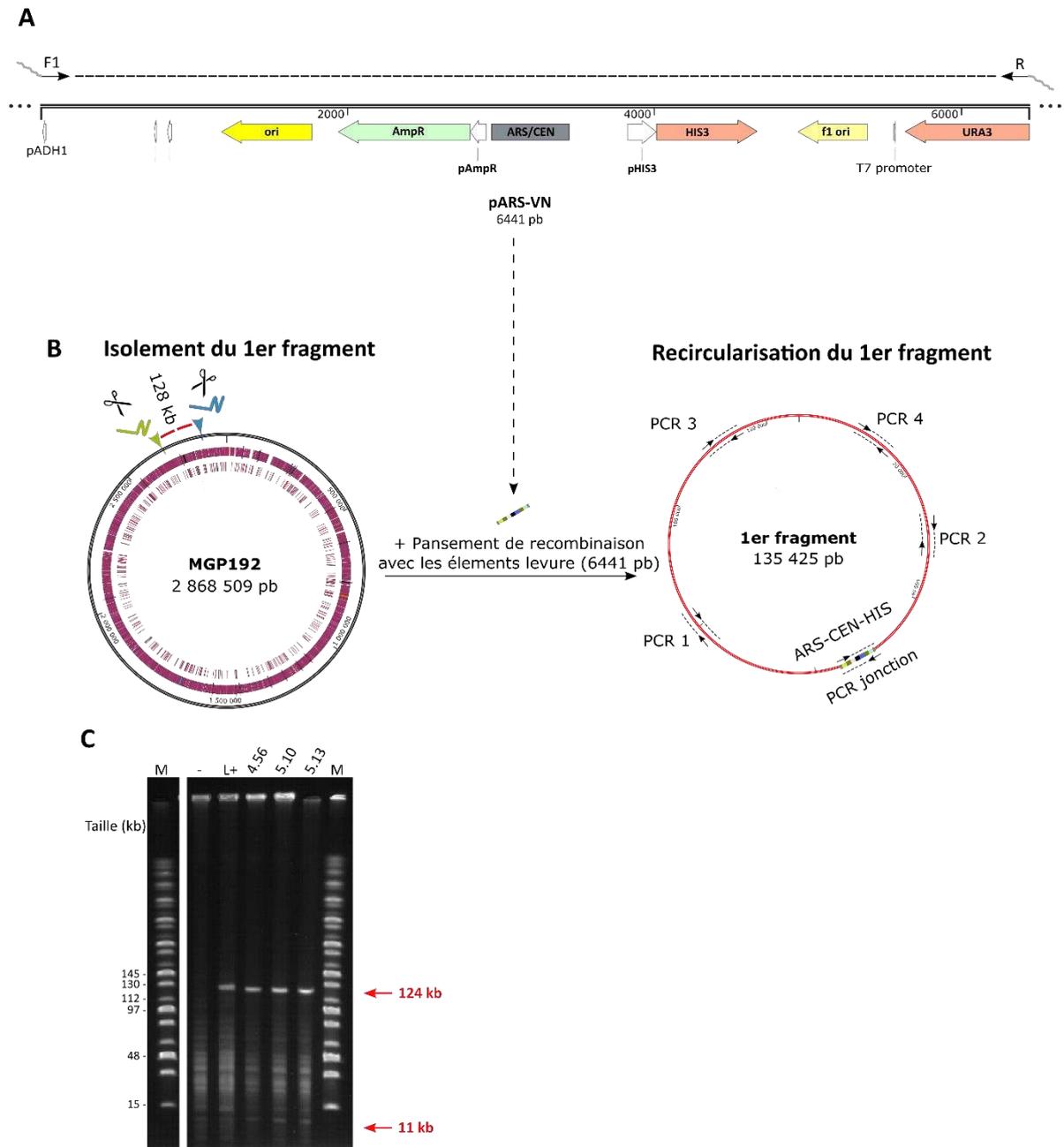
2003) et générer un grand nombre de faux positifs, nous avons ajouté à ce « design » les éléments nécessaires à une contre sélection basée sur le marqueur URA3 et le 5'-FOA (acide 5-fluoroorotique) (Sherman *et al.*, 1987). Le principe de cette contre-sélection est illustrée figure 55. Brièvement, les régions d'homologie avec le fragment cible à capturer (rectangles bleu clair et vert clair sur le schéma) sont placées entre le promoteur pADH1 et le gène URA3. Ainsi, si le fragment cible (= région d'intérêt en rouge) s'assemble au pansement de recombinaison par recombinaison homologue, alors le gène URA3 éloigné de son promoteur n'est pas exprimé, et les levures résistent au 5-FOA présent dans le milieu. Au contraire, si le pansement se referme sur lui-même, le gène URA3 réassocié au promoteur constitutif pADH1 est exprimé et produit les enzymes convertissant le 5-FOA en 5-fluorouracil, toxique pour la levure. En théorie, grâce à ce système, seuls les clones levures hébergeant le fragment d'intérêt sont capables de pousser sur le milieu de contre-sélection SD-HIS + 5FOA.

Les pansements à produire pour la version 2 du protocole ont donc des tailles bien supérieures (~6,4 kb) à ceux nécessaires pour la version 1 du protocole (90 pb). Ils contiennent : (i) les éléments levure ARSH4, CEN6 et HIS3, (ii) le marqueur URA3 nécessaire au système de contre sélection au 5'-FOA séparé de son promoteur pADH1, et (iii) à chaque extrémité, des séquences d'environ 60 pb identiques aux extrémités du fragment de génome à cloner.

Tous les pansements ont été produits par PCR en utilisant comme matrice le plasmide pARS-VN (Noskov *et al.*, 2003) et des amorces de 80 bases incluant des queues flottantes de 60 bases spécifiques du fragment à capturer (figure 56A).

#### 4.2. Capture d'un premier fragment de génome par contre sélection au 5'FOA

Afin de démontrer que nous étions capables de capturer un fragment de génome à partir d'une souche bactérienne non modifiée, nous avons choisi de capturer le fragment correspondant à la région entre les gènes *atpA* et *IrgB* des souches MGP192 (~ 128 kb) (figure 56B) et MS (~ 149 kb) originales.



**Figure 56 : Capture par CREasPy-Fusion, d'un premier fragment de génome de *B. subtilis* non marqué avec les éléments levure, par contre sélection au 5'-FOA. [A]** Utilisation du pARS-VN pour la production du pansement de recombinaison contenant les éléments levure. Les régions d'homologie avec le fragment à capturer sont ajoutées lors de la PCR. Pour ce faire, des longues amorces F1 et R (80 bases total) incluant des queues flottantes de 60 bases spécifiques du fragment à capturer ont été utilisées. **[B]** Isolement du fragment par deux duplex ARNg-Cas9. Re-circularisation du fragment par le système de recombinaison homologue de la levure. Le pansement de recombinaison produit à partir du pARS-VN est utilisé par la levure pour réparer les cassures double brin et former un fragment circulaire. Les éléments levure sont ajoutés sur le fragment lors de cette étape. **[C]** L'intégrité et la taille du fragment sont validés par restriction enzymatique et analyse en PFGE. La digestion par l'enzyme AvrII linéarise le fragment. La présence d'un site de restriction additionnel au niveau des éléments levure apportés sur le pansement induit un profil à 2 bandes d'environ 124 kb et 11 kb pour le fragment issu de la souche MGP192. M : Mid range PFG ladder NEB (15.0 - 242.5 kbp). - : témoin négatif, ADNg de VL6-48N. L+ : témoin positif, ADNg de transformant levure hébergeant le fragment capturé par TAR-cloning.

Pour cela, nous avons eu besoin d'une souche levure pré-chargée avec le système CRISPR-Cas9 et un plasmide p426-gRNA « double target » permettant l'expression de deux ARNg ciblant les gènes *atpA* et *ywbE*. Un clone levure VL6-48N hébergeant ces plasmides avait été produit lors de la capture du premier fragment pré-marqué (cf. partie 3.3 du chapitre 1) mais n'a pas pu être réutilisé ici pour deux raisons simples. La première est que le plasmide p426-gRNA « double target » présent dans ce clone portait le marqueur URA3, dont nous avons besoin pour créer le pansement d'environ 6.4 kb. Le plasmide p426-gRNA « double target » ciblant les gènes *atpA* et *ywbE* a donc été reconstruit avec le marqueur ADE2 à la place d'URA3. La deuxième raison est que la souche levure VL6-48N n'est pas utilisable avec le marqueur ADE2. Il a donc fallu identifier une autre souche avec laquelle travailler et nous avons choisi la souche W303a. La construction de cette souche a été réalisée comme décrit dans la section 3.1.

Pour mettre en place cette expérience, nous avons mis en contact des protoplastes des souches de *B. subtilis* MS et MGP192 avec des sphéroplastes W303a pré-chargée avec le système CRISPR-Cas9 spécifique au fragment cible puis les transformants ont été sélectionnés sur milieu solide SD-HIS complémenté avec du 5'-FOA.

Un total de 92 et 60 transformants levure ont été obtenus respectivement pour la souche MGP192 et MS. Ils ont tous été analysés, premièrement par PCR pour évaluer la présence du fragment puis comme précédemment par PFGE. Le nombre de transformants étant élevé, ils ont été rassemblés en pools de 5 pour extraire l'ADNg et effectuer la première PCR (PCR 1). Trois pools ont montré le profil attendu avec un amplicon à 909 pb. L'ADNg des clones issus de ces pools a donc été extrait individuellement et analysé par PCR. L'intégrité du fragment cloné a été évaluée par PCR à l'aide de 4 couples d'amorces placées tout le long du fragment, donnant des bandes de 909 pb (PCR 1), 962 pb (PCR 2), 868 pb (PCR 3) et 883 pb (PCR 4). Puis la circularisation du fragment a été confirmée par une PCR produisant un amplicon de 7 888 pb. Trois clones (cl. 4.56, 5.10 et 5.13) ont montré le profil attendu pour toutes les PCR, ils étaient tous issus des conditions mettant en jeu la souche *B. subtilis* MGP192. La taille des fragments clonés a ensuite été évaluée par restriction enzymatique (AvrII) et analyse par PFGE. Les 3 clones analysés présentaient tous le profil attendu à deux bandes (124 kb + 11 kb) correspondant au profil théorique après digestion du fragment cloné par AvrII (figure 56C).



Cette expérience montre qu'il est possible de capturer dans la levure *S. cerevisiae*, par CREasPy-Fusion, un fragment de génome de *B. subtilis* n'ayant pas été au préalable marqué avec les éléments levure. Comme précédemment, il semblerait que la capture soit nettement plus efficace pour la souche MGP192 que la souche MS (voire possible exclusivement avec MGP192).

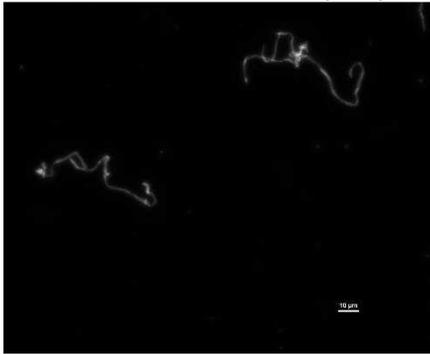
A ce stade, nous avons donc développé deux versions du protocole de CREasPy-Fusion. La première version utilise comme matériel de départ des souches bactériennes génétiquement modifiées et requiert une étape supplémentaire de transformation. La seconde version utilise, quant à elle, des souches bactériennes non modifiées génétiquement et est donc plus rapide. Malgré sa simplicité, cette seconde version semble cependant moins efficace. En effet, en se référant aux résultats obtenus avec la souche *B. subtilis* MGP192 pour la capture du premier fragment de génome : 26 clones ont été testés positifs sur 78 clones analysés pour la version 1 du protocole et 3 clones ont été testés positifs sur 92 clones analysés pour la version 2 du protocole.

## 5. Etape E : Observation des évènements de fusion cellulaire entre levure et *B. subtilis*

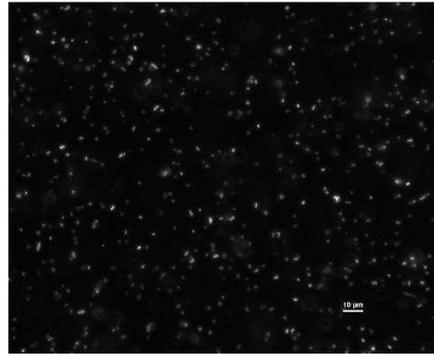
Dans les parties précédentes, nous avons montré que la méthode de CREasPy-Fusion pouvait être adaptée à *B. subtilis*, en démontrant de manière séquentielle (1) le transfert d'un plasmide entre des protoplastes de *B. subtilis* et des sphéroplastes de levure par contact cellules/cellules puis (2) le transfert et la modification de ce plasmide, (3) la capture de deux fragments de génome pré-marqués avec les éléments levure, à partir de 3 souches de *B. subtilis* et enfin (4) la capture d'un fragment de génome non pré-marqué, grâce à la contre sélection au 5'-FOA.

Lors de la mise au point de la méthode de clonage de génome entier par fusion par l'équipe de B. Karas (cf partie 2.4.1.4 de l'introduction) sur laquelle est basée la CREasPy-Fusion, les auteurs ont proposé que le transfert d'ADN pouvait se produire selon 3 scénarios : (i) une fusion membranaire entre la bactérie et la levure, (ii) une lyse bactérienne suivie d'une transformation de la levure avec l'ADN bactérien libéré, ou encore (iii) une endocytose de la bactérie par la levure (Karas *et al.*, 2013) (figure 19). Nos expériences de transfert du plasmide

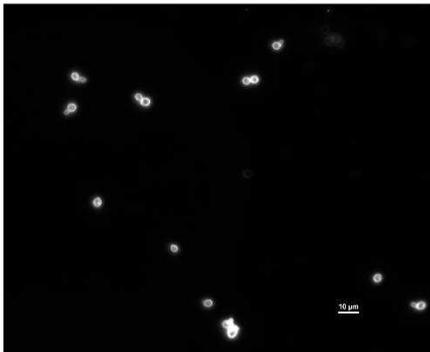
**A** *B. subtilis* MGP192 avant lysozyme



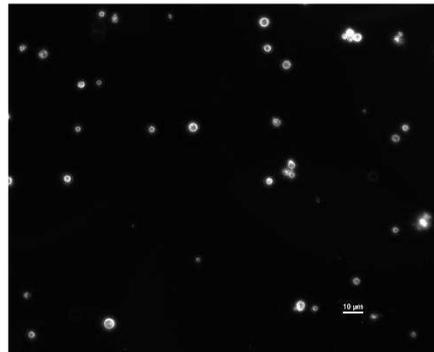
**B** *B. subtilis* MGP192 après lysozyme



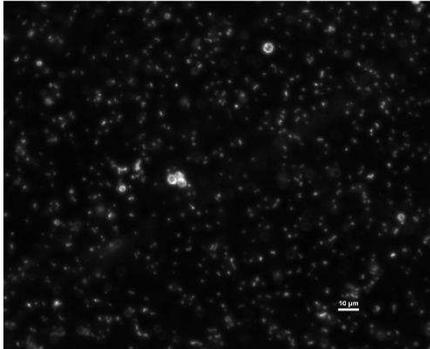
**C** *S. cerevisiae* avant zymolyase



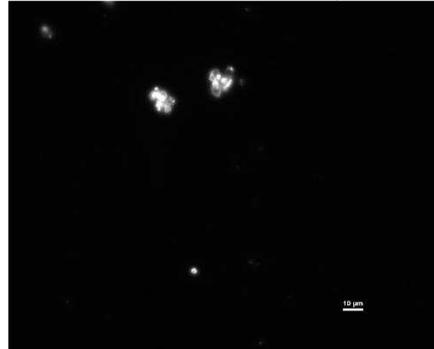
**D** *S. cerevisiae* après zymolyase



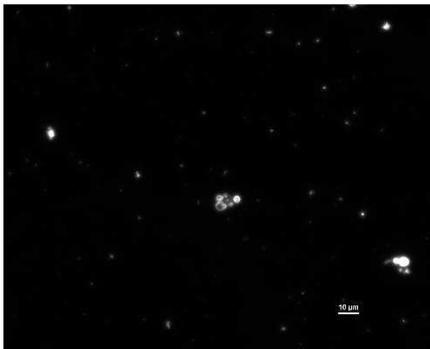
**E** *B. subtilis* + *S. cerevisiae* avant PEG



**F** *B. subtilis* + *S. cerevisiae* après PEG



**G** *B. subtilis* + *S. cerevisiae* en SOS



**Figure 57 : Observation en microscopie à fond noir d'une expérience de CReasPy-Fusion. [A]** *B. subtilis* MGP192 avant traitement au lysozyme. **[B]** Protoplastes de *B. subtilis*. **[C]** Levure avant le traitement à la zymolyase. **[D]** Sphéroplastes de levure. **[E]** Protoplastes de *B. subtilis* et sphéroplastes de levure mis en contact en absence de PEG. **[F]** Puis en présence de PEG. **[G]** Enfin en milieu SOS. Echelle : 10  $\mu$ m. Objectif : 60x

navette en présence de Dnase I (cf partie 2.1, chapitre 1) ont permis d'invalider le scénario de la lyse bactérienne. En revanche, les 2 autres scénarios demeurent encore possibles.

Afin de mieux comprendre les mécanismes qui sous-tendent le transfert d'ADN entre protoplastes bactériens et sphéroplastes de levure, nous avons mis en place des expériences préliminaires de microscopie à fond noir, microscopie à épifluorescence et microscopie confocale.

### 5.1. Observation des événements de fusion en microscopie à fond noir

Ne sachant pas si les conditions dans lesquelles étaient menées la CReasPy-Fusion étaient adaptées à des observations en microscopie, notre premier essai a consisté à mimer une expérience de CReasPy-Fusion, et à effectuer des prélèvements à des moments clés de l'expérience afin de les observer au microscope à fond noir.

Pour y parvenir, l'expérience décrite paragraphe 3.3 qui mettait en jeu des protoplastes de la souche *B. subtilis* MGP192 $\Delta$ sbo-*alb*::ACH-spec et des sphéroplastes VL6-48N hébergeant le système CRISPR-Cas9, a été reproduite (excepté que les cellules n'ont pas été étalées sur milieu sélectif à la fin du protocole). Au cours de cet essai, des échantillons de 5  $\mu$ L ont été prélevés, placés entre lame et lamelle et observés au microscope à fond noir (objectif 60x).

Les premières observations ont été faites sur la culture de *B. subtilis* MGP192 à une  $DO_{600nm}=0,4$  avant et après traitement au lysozyme (figures 57A et 57B). Avant traitement, toutes les cellules présentaient une forme très allongée (avec des tailles très variables pouvant aller de  $\sim 30 \mu m$  à  $\sim 110 \mu m$  et en moyenne  $\sim 60 \mu m$ ), un phénotype caractéristique de la souche à génome réduit MGP192. Cette souche est en effet affectée au niveau du cycle cellulaire et de la division cellulaire : ses cellules ne sont pas sous forme de bacilles (2 à 4  $\mu m$ ) à bout arrondis mais se présentent sous forme de très longs filaments (Dervyn *et al.*, 2023). Après dégradation de la paroi des cellules par l'action du lysozyme, les cellules avaient, en revanche, une forme bien arrondie spécifique de protoplastes ( $\sim 1,3 \mu m$  de diamètre) (figure 57B).

Nous avons poursuivi en observant la culture de *S. cerevisiae* avant et après traitement à la zymolyase (figures 57C et 57D). Avant traitement, les cellules présentaient une forme ovoïde ( $\sim 4,5 \mu m$  de diamètre) avec parfois des bourgeons (= futures cellules filles). Après



traitement à la zymolyase et dégradation de la paroi, les sphéroplastes de levure présentait une forme bien arrondie (~3,7  $\mu\text{m}$ ).

Pour terminer, après avoir observé les protoplastes et sphéroplastes séparément, nous avons souhaité les observer lors de leur mise en contact avant ajout du PEG (figure 57E), puis après ajout du PEG (figure 57F), et enfin dans le milieu de régénération SOS (cf partie 3.3.2 matériels et méthodes, pour le protocole) (figure 57G). Avant ajout du PEG, les deux types cellulaires étaient distribués de façon homogène sur la lame et les protoplastes semblaient plus nombreux que les sphéroplastes (pas de co-localisation). En contraste, après ajout du PEG, de gros agrégats, parsemés de façon hétérogène sur la lame, s'étaient formés. Les protoplastes de *B. subtilis*, tout comme les sphéroplastes de levures, n'étaient plus discernables, et avaient laissé place à de gros amas cellulaires. De façon plus surprenante, les agrégats étaient encore bien visibles en milieu SOS, mais comme pour l'étape du PEG, il n'était pas possible de définir leur composition (figure 57G).

La mise en œuvre de cette première expérience nous a permis de valider que dans les conditions classiques de CReasPy-Fusion nous étions capable d'observer (i) les cellules et les protoplastes de *B. subtilis* (ii) les cellules et les sphéroplastes de levure, mais aussi (iii) la mise en contact protoplastes/sphéroplastes en absence ou présence de PEG et plus tard en milieu SOS. Les concentrations cellulaires par champ étaient correctes pour chacune des étapes et aucun ajustement n'était donc nécessaire pour passer aux étapes suivantes qui visaient à tenter de mieux documenter le contact sphéroplastes/protoplastes.

## 5.2. Observation des évènements de fusion en microscopie à épifluorescence.

Pour mieux visualiser les évènements ayant lieu au moment du contact « protoplastes/sphéroplastes » ainsi que ceux ayant lieu après ce contact, nous avons pensé que des expériences de microscopie à fluorescence pourraient être plus appropriées. Mais ce type d'expériences nécessite d'avoir à disposition des souches exprimant différentes protéines fluorescentes, un matériel que nous n'avions pas au laboratoire et qu'il fallait donc acquérir ou construire.



Concernant *B. subtilis*, nos collaborateurs de l'équipe SyBER possédaient dans leur souchothèque plusieurs souches de *B. subtilis* fluorescentes. Parmi elles, nous avons sélectionné la souche de *B. subtilis* BSBI exprimant la mRFP de manière constitutive dans le cytosol (appelée CW22, cf partie 1.2 du matériels et méthodes). La mRFP (pour RFP monomérique) (Campbell *et al.*, 2002) dérive de la RFP (red fluorescent protein) découverte en 1999 (Matz *et al.*, 1999) et issue du champignon corallien *Discosoma sp.* Cette protéine émet une fluorescence dans le rouge à une longueur d'onde optimale de 583 nm.

Concernant *S. cerevisiae*, n'identifiant aucun laboratoire possédant une souche de *S. cerevisiae* exprimant une protéine fluorescente de manière constitutive dans le cytosol, nous avons décidé de construire nous-même une souche exprimant la GFP. La GFP (green fluorescent protein) est issue de la méduse *Aequorea victoria* et a été découverte en 1962 par Shimomura (SHIMOMURA *et al.*, 1962). Parmi les nombreux variants développés (Cormack *et al.*, 1996), nous avons sélectionné la yeGFP (yeast-enhanced GFP), une version « codon optimisée » pour l'expression chez *S. cerevisiae* (Cormack *et al.*, 1997). Cette protéine fluorescente émet dans le vert à une longueur d'onde optimale de 510 nm.

La GFP et la RFP ont pour avantage d'être stables et de former des chromophores de façon spontanée, sans nécessiter de co-facteurs.

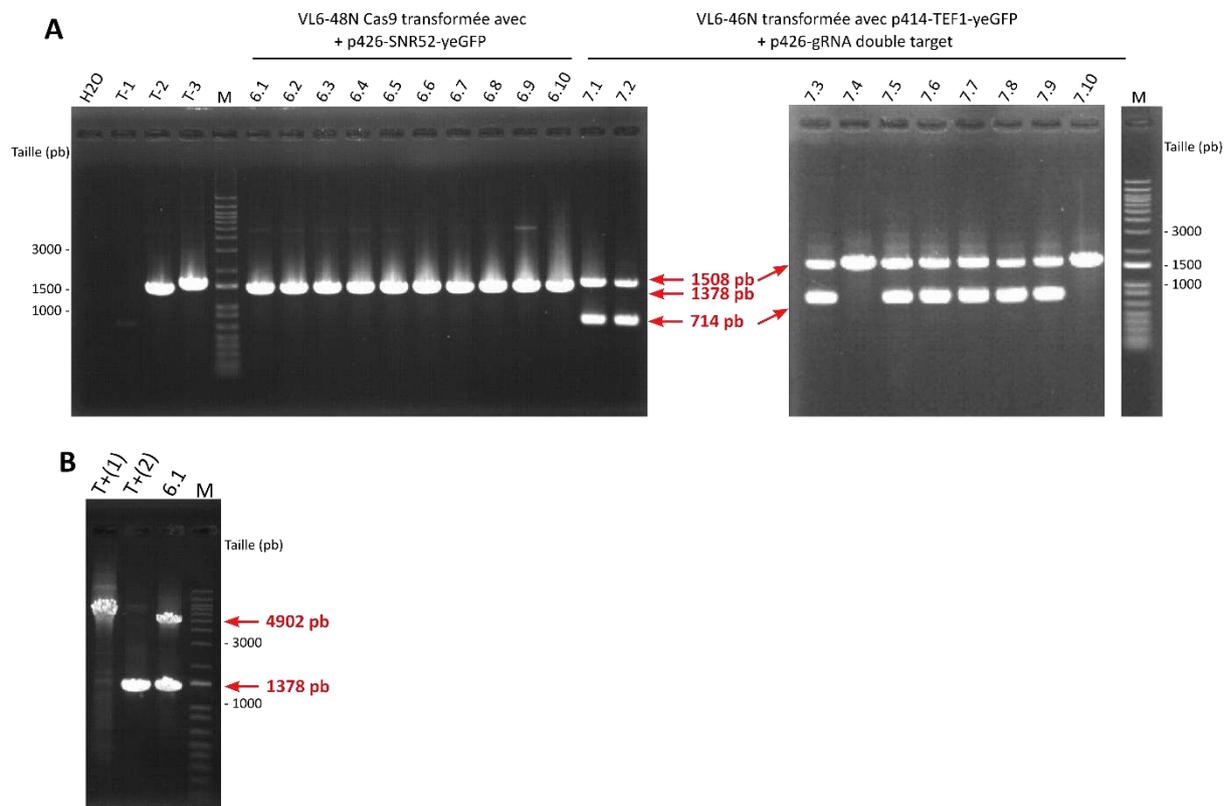
### 5.2.1. Construction de levures fluorescentes

Pour se placer dans les conditions les plus proches possibles des expériences de CReasPy-Fusion, nous avons souhaité que la levure possède des plasmides similaires à ceux utilisés pour l'expression du système CRISPR-Cas9. Nous avons donc construit deux plasmides portant le gène codant la yeGFP :

- Le p414-TEF1-yeGFP qui correspond au plasmide d'expression de la Cas9 (p414-TEF1p-Cas9-CYC1t) dans lequel le gène codant la Cas9 a été remplacé par le gène yeGFP. Ce plasmide dérive d'un plasmide centromérique et est donc mono-copie (une copie par cellule).
- Le p426-SNR52-yeGFP qui correspond au plasmide d'expression d'ARNg (p426-SNR52p-gRNA.Aar1.Y-SUP4t) dans lequel la séquence qui code l'ARNg a été remplacée

**Tableau 20 : Essais de transformation de la souche VL6-48N dans le but de produire des levures exprimant la GFP.**

Souche levure	Plasmide	CFU	Clones positifs / Clones analysés	Clone sélectionné
VL6-48N	Sans ADN	0	-	-
	p414-TEF1-yeGFP + p426-gRNA double target	~100	8/10	7.5
VL6-48N Cas9 (p414-TEF1p-Cas9-CYC1t)	Sans ADN	0	-	-
	p426-SNR52-yeGFP	>300	10/10	6.1



**Figure 58 : Analyse des transformants levures issus de la transformation des souches VL6-48N et VL6-48N/Cas9 dans le but de produire des levures exprimant la GFP. [A]** La présence des plasmides p414-TEF1-yeGFP et p426-gRNA double target a été contrôlée chez les clones 7.1 à 7.10 avec le couple d'amorce T3/T7 produisant des amplicons de 1508 pb pour le p414-TEF1-yeGFP et 714 pb pour le p426-gRNA double target. La présence du p426-SNR52-yeGFP dans la souche VL6-48N/Cas9 (cl 6.1 à 6.10) a été contrôlée elle aussi avec le couple d'amorce T3/T7 produisant un amplicon de 1378 pb. **[B]** La présence des plasmides p426-SNR52-yeGFP et p414-TEF1p-Cas9-CYC1 a été contrôlée chez le clone 6.1 avec le couple d'amorce T3/T7 et un temps d'élongation plus long permettant de produire un amplicon de 4902 pb pour le plasmide p414-TEF1p-Cas9-CYC1 et 1378 pb pour le plasmide p426-SNR52-yeGFP. M: DNA Ladder 1 kb+ Invitrogen (100-12000 pb). H2O: contrôle négatif, sans ADN. T-1: témoin négatif, avec l'ADNg de *S. cerevisiae* VL6-48N. T-2: témoin positif, plasmide p426-SNR52-yeGFP. T-3: témoin positif, plasmide p414-TEF1-yeGFP. T+(1): témoin positif, plasmide p414-TEF1p-Cas9-CYC1. T+(2) témoin positif, plasmide p426-SNR52-yeGFP.

par le gène yeGFP. Ce plasmide dérive d'un plasmide 2 $\mu$  et est donc multi-copie (plusieurs copies par cellule).

Après validation des constructions, le plasmide p414-TEF1-yeGFP a été co-transformé dans la souche VL6-48N avec le p426-gRNA « double target » utilisé pour la capture du premier fragment de génome et le p426-SNR52-yeGFP a été transformé dans la souche VL6-48N hébergeant déjà le p414-TEF1p-Cas9-CYC1t (tableau 20). Les transformations de ces deux souches levure ont été réalisées en utilisant le protocole à l'acétate de lithium.

Dans chaque cas, plusieurs centaines de colonies sont apparues sur milieu de sélection SD-TRP-URA. Dix colonies de chaque condition ont été analysées par PCR simplex avec le couple d'amorces universelles T3/T7 permettant de détecter à la fois la présence des plasmides basés sur le p414 et ceux basés sur le p426.

Concernant la souche VL6-48N transformée avec les 2 plasmides, 8 clones sur 10 ont présenté un profil à deux bandes avec la première à une taille de 1508 pb correspondant à la présence du p414-TEF1-yeGFP et l'autre à une taille de 714 pb correspondant à la présence du p426-gRNA double target (figure 58A clone 7.1 à 7.10). A l'issue de ces analyses, le clone 7.5 a été conservé (tableau 20).

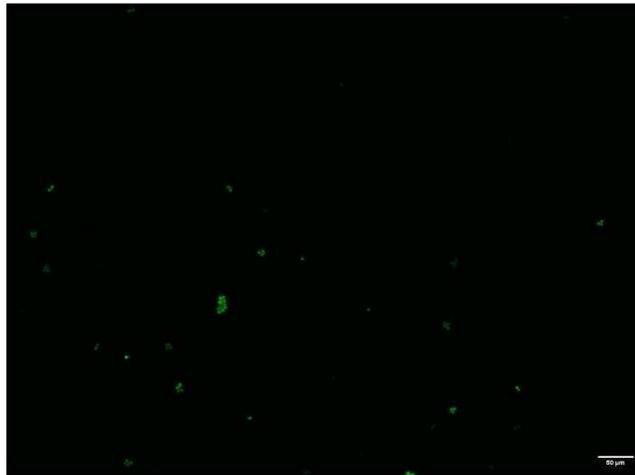
Concernant la transformation de la souche VL6-48N/Cas9 avec le p426-SNR52-yeGFP, nous avons commencé par vérifier la présence du nouveau plasmide (p426-SNR52-yeGFP) dans les clones repiqués. Comme le montre la figure 58A, tous les clones testés (cl 6.1 à 6.10) présentaient la bande attendue à 1378 pb. Le clone 6.1 a été sélectionné (tableau 20) et nous avons vérifié dans ce clone la présence des deux plasmides en effectuant la même PCR mais avec un temps d'élongation plus long. Un profil à deux bandes, la première à une taille de 4902 pb correspondant à la présence du p414-TEF1p-Cas9-CYC1 et l'autre à une taille de 1378 pb correspondant à la présence du p426-SNR52-yeGFP, était bien visible sur gel d'agarose (figure 58B).

Pour les deux clones sélectionnés (cl 7.5 et cl 6.1, tableau 20), la séquence du gène yeGFP a été vérifiée par séquençage Sanger. La séquence étant correcte pour chacun de ces clones, plusieurs tubes ont été stockés à -80°C.

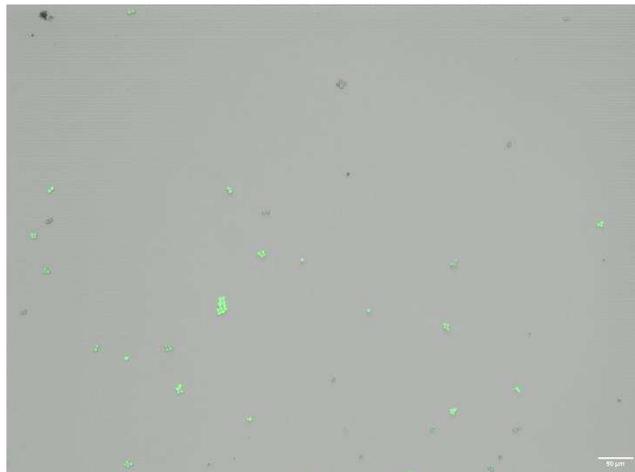
**A** Levure/GFP cl 7.5 observation en lumière transmise



**B** Levure/GFP cl 7.5 observation avec filtre GFP



**C** Levure/GFP cl 7.5 superposition des deux premières images



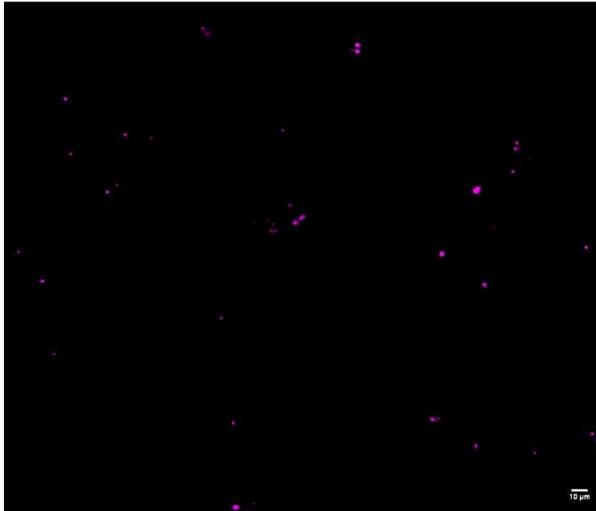
**Figure 59 : Observation en microscopie en épifluorescence : Analyse phénotypique de la souche levure VL6-48N p414-TEF1-yeGFP + p426-gRNA double target. [A] Observation en lumière transmise. [B] Observation avec un filtre GFP. [C] Superposition des images en lumière transmise et avec le filtre GFP. Echelle : 50 µm. Objectif : 10x**

### 5.2.2. Observation du contact *B. subtilis*/levure en microscopie à épifluorescence

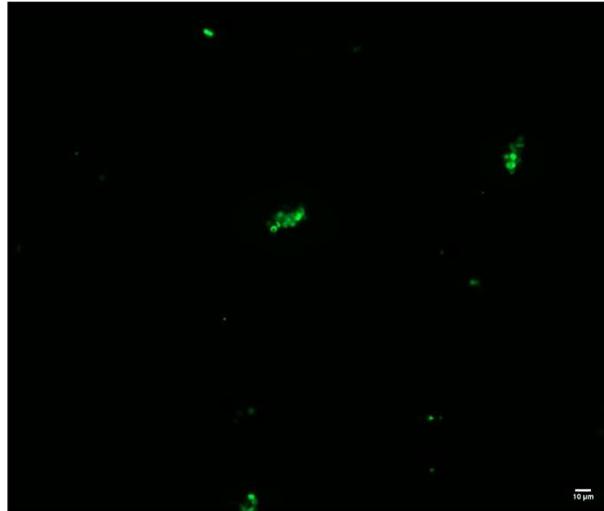
Afin de s'assurer que le clone de *B. subtilis*/RFP CW22 fourni par nos collaborateurs et que les clones levure/yeGFP construits dans notre équipe étaient bien fluorescents, nous les avons mis en culture et observés en microscopie à épifluorescence. Cette méthode est basée sur l'utilisation d'un microscope optique combiné avec des filtres permettant d'observer la fluorescence émise par un objet. Ici les protéines fluorescentes GFP et RFP produites respectivement par la levure et *B. subtilis* permettent de visualiser directement la fluorescence dans les cellules vivantes.

Pour cela, 5  $\mu$ L d'une culture *B. subtilis*/RFP ( $DO_{600nm}=1$ ) et 5 $\mu$ L des deux cultures levures/yeGFP ( $DO_{600nm}=1$ ) ont été prélevés et chaque échantillon a été déposé entre lame et lamelle. Lorsque les échantillons ont été observés en lumière transmise (grossissement 20x), des cellules isolées non colorées en forme de bâtonnet étaient visibles pour *B. subtilis*/RFP et des cellules isolées non colorées ovoïdes bourgeonnantes étaient visibles pour les deux clones levures/yeGFP (cl 7.5 et cl 6.1 tableau 20) (figure 59A). Ces mêmes échantillons observés après excitation à des longueurs d'onde de 561 nm pour la détection de la mRFP et à 488 nm pour la détection de la yeGFP, ont laissé apparaître des cellules qui émettaient une fluorescence unique dans le rouge pour le clone *B. subtilis* et une fluorescence unique dans le vert pour les levures mais seulement pour un des deux clones levures testés (cl 7.5) et non les deux, comme attendu (figure 59B). De plus, alors que la population de *B. subtilis*/RFP était homogène et que toutes les cellules observées par champ étaient positives à la RFP, le clone levure/yeGFP cl 7.5 présentait environ 60% de cellules positives à la yeGFP (figure 59C), un phénomène décrit par ailleurs dans la littérature (Kaishima *et al.*, 2016). Cependant, les cellules positives émettant de forte intensité de fluorescence dans le vert, nous avons décidé de conserver ce clone pour la suite des expériences. Ces premières observations nous ont permis de mettre en lumière plusieurs points essentiels : (i) le clone *B. subtilis*/RFP CW22 et le clone levure/yeGFP cl 7.5 exprimaient respectivement des protéines mRFP et yeGFP fonctionnelles et en quantité suffisante pour être détectées en microscopie à fluorescence, (ii) toutes les cellules du clone levure/yeGFP cl 7.5 n'exprimaient pas la yeGFP, (iii) le clone levure/yeGFP cl 6.1 dans lequel

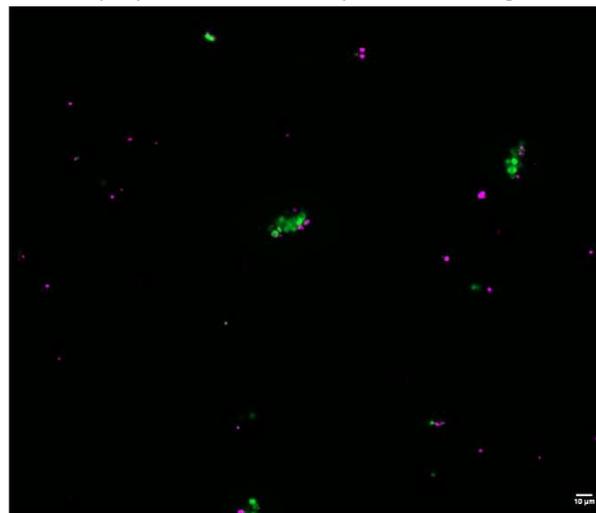
**A** Contact *B. subtilis* RFP/ *S. cerevisiae* GFP  
observation avec filtre RFP



**B** Contact *B. subtilis* RFP/ *S. cerevisiae* GFP  
observation avec filtre GFP



**C** Contact *B. subtilis* RFP/ *S. cerevisiae* GFP  
superposition des deux premières images



**Figure 60 : Observation en microscopie en épifluorescence : Etape de contact, en présence de PEG, entre les protoplastes de *B. subtilis* CW22 et les sphéroplastes de levure VL6-48N p414-TEF1-yeGFP + p426-gRNA double target. [A] Observation avec un filtre RFP. Afin de pouvoir observer les cellules, le contraste de cette image a été modifié sur ImageJ. [B] Observation avec un filtre GFP. [C] Superposition des images obtenues avec le filtre RFP et le filtre GFP. Echelle : 10 μm. Objectif : 20x**

le gène yeGFP était porté par le plasmide multi-copie p426-SNR52-yeGFP n'émettait pas de fluorescence dans le vert. Nous avons émis deux hypothèses pour expliquer ce résultat :

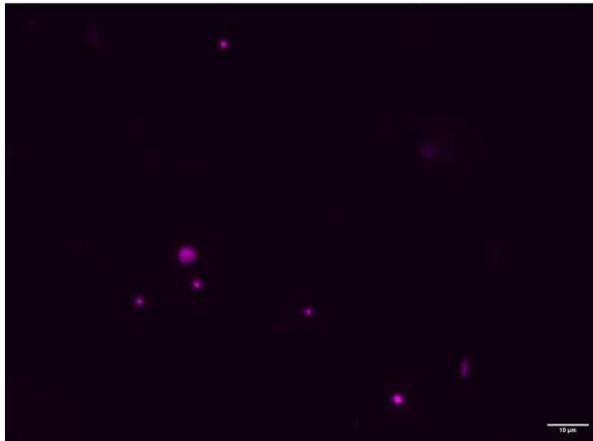
- La protéine yeGFP exprimée à partir d'un plasmide multi-copie est présente en quantité trop élevée dans la cellule, affectant l'émission de fluorescence.
- Une/des mutation(s) présente(s) ailleurs dans le génome interfère(nt) avec l'expression du gène yeGFP ou l'émission de fluorescence dans ce clone particulier.

Il aurait été intéressant de tester d'autres clones levure/p426-SNR52-yeGFP pour conforter l'une ou l'autre des hypothèses, mais par manque de temps, cela n'a pas été fait.

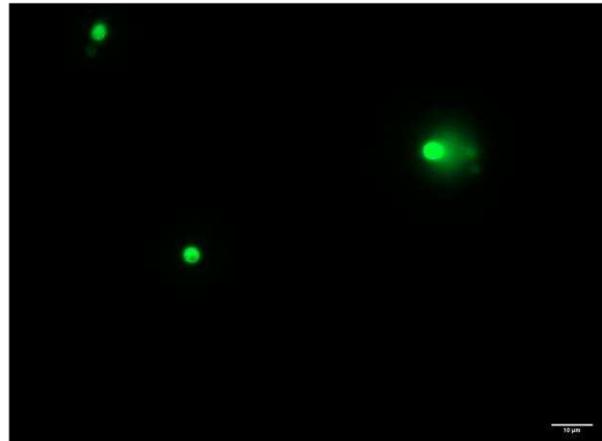
Suite à cette première expérience, nous avons à disposition un clone *B. subtilis* fluorescent en mRFP (CW22) et un clone levure fluorescent en yeGFP (cl 7.5). Une deuxième expérience a donc été mise en place dans le but d'observer cette fois les interactions protoplastes/sphéroplastes au cours de la CREasPy-Fusion. Pour ce faire, nous avons produits des protoplastes à partir du clone *B. subtilis*/RFP CW22, et des sphéroplastes du clone levure/yeGFP cl 7.5, puis nous les avons mis en contact en présence de PEG. Lors de cette expérience, les protoplastes de *B. subtilis* et les sphéroplastes de levure ont d'abord été observés séparément dans le but de contrôler leur fluorescence après l'action respectivement du lysozyme et de la zymolyase. Ils ont ensuite été observés ensemble en absence puis en présence de PEG. Et enfin, ils ont été observés dans le milieu de régénération des levures (le milieu SOS) soit directement après ajout (T0) soit après 1 heure d'incubation (T1h). Pour toutes les étapes, nous avons contrôlé la spécificité de fluorescence en observant les échantillons avec des filtres DAPI, RFP et GFP.

Les résultats ont été les suivants. Nous avons observé qu'après traitement au lysozyme, les protoplastes *B. subtilis*/RFP CW22 étaient fluorescents et qu'ils émettaient une fluorescence uniquement dans le rouge. De même, après traitement à la zymolyase, les sphéroplastes de levure étaient également toujours fluorescents et émettaient une fluorescence uniquement dans le vert. Après mise en contact des deux échantillons (sans PEG), la fluorescence respective de chaque échantillon était conservée dans le mélange, offrant des images sur lesquelles des levures vertes coexistaient à côté de bactéries rouges plus petites. Lorsque le PEG a été ajouté, des images radicalement différentes ont été capturées. En effet, les nombreuses cellules isolées avaient laissé place à de nombreux agrégats constitués de levures émettant dans le vert et de bactéries émettant dans le rouge,

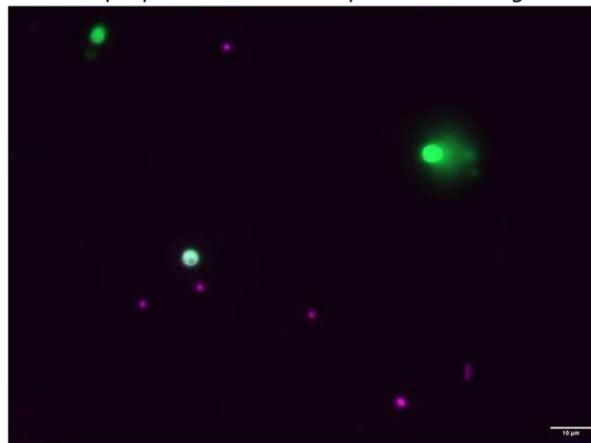
**A** *B. subtilis* RFP/*S. cerevisiae* GFP en SOS  
observation avec filtre RFP



**B** *B. subtilis* RFP/*S. cerevisiae* GFP en SOS  
observation avec filtre GFP



**C** *B. subtilis* RFP/*S. cerevisiae* GFP en SOS  
superposition des deux premières images



**Figure 61 : Observation en microscopie en épifluorescence : En milieu SOS. [A]** Observation avec un filtre RFP. **[B]** Observation avec un filtre GFP. **[C]** Superposition des images obtenues avec le filtre RFP et le filtre GFP. Présence d'une cellule levure fluorescente en RFP et en GFP. Echelle : 10  $\mu\text{m}$ . Objectif : 63x

montrant clairement une co-localisation des deux types cellulaires à cette étape (figure 60) et un possible contact physique. Plus tard, en milieu SOS, nos observations se sont affinées et notre réflexion s'est éclaircie lorsque nous avons mis en évidence pour la première fois quelques levures présentant une fluorescence à la fois en GFP (comme attendu) mais également en RFP. La fluorescence visualisée en GFP et RFP émanait de l'ensemble de la cellule (excepté au niveau de ce qui semblait être une vacuole) et non d'un point précis de la cellule, suggérant un transfert de la totalité du contenu de la bactérie dans la levure *S. cerevisiae* (figure 61). Ces résultats inattendus nous ont encouragés à poursuivre, c'est pourquoi des expériences en microscopie confocale ont ensuite été menées.

### 5.3. Observation des évènements de fusion en microscopie confocale

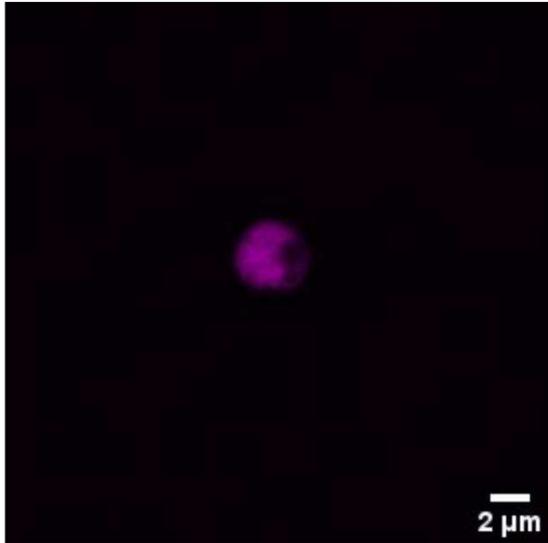
Les observations en microscopie à épifluorescence nous ont permis de montrer que suite à l'ajout de PEG, les bactéries et levures étaient pour la plupart regroupées sous forme de gros amas cellulaires, amas qui persistaient ensuite dans le milieu de régénération SOS. Le résultat le plus frappant de cette série d'expériences a été la mise en évidence en milieu SOS de levures positives à la fois en GFP et en RFP, soulignant le fait qu'après ajout de PEG, les bactéries et levures non seulement co-localisent mais interagissent.

Dans le but de documenter au mieux ce contact physique et les étapes qui s'ensuivent, nous avons souhaité observer nos échantillons en microscopie confocale. Le microscope confocal est un microscope à épi-fluorescence qui permet d'imager uniquement les molécules fluorescentes qui sont localisées sur le plan focal de l'objectif en place.

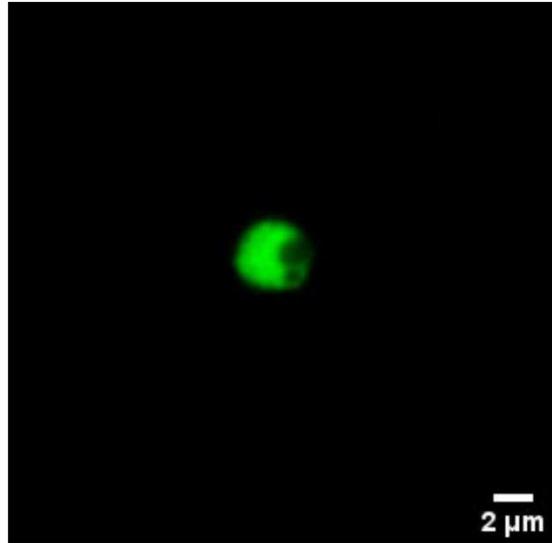
Pour obtenir ces données, une nouvelle expérience de CREasPy-Fusion mettant en jeu le clone *B. subtilis* CW22 fluorescent en RFP et le clone levure fluorescent en GFP (cl 7.5) a été menée. Comme pour les expériences précédentes, l'émission de fluorescence a été vérifiée à chaque étape du protocole mais seulement documentée par la prise d'images au moment du contact bactéries/levures/PEG puis au moment de l'ajout du milieu de régénération (T0) et après 30 minutes d'incubation dans ce milieu (T30).

Plusieurs éléments intéressants sont ressortis de cette expérience. Tout d'abord, les observations réalisées en milieu SOS (T0 et T30) nous ont permis de confirmer les résultats de microscopie à épifluorescence. En effet, des levures présentant une fluorescence à la fois en GFP et en RFP ont été observées aux deux temps (figure 62). Les spectres d'émission de ces

**A** Observation en SOS avec filtre RFP après contact *B. subtilis* RFP / *S. cerevisiae* GFP



**B** Observation en SOS avec filtre GFP après contact *B. subtilis* RFP / *S. cerevisiae* GFP



**C** Superposition des deux premières images



**Figure 62 : Observation en microscopie confocale : En milieu SOS une cellule de levure fluorescente en RFP et GFP. [A] Observation en RFP. [B] Observation en GFP. [C] Superposition des images obtenues en RFP et en GFP. Echelle : 2  $\mu\text{m}$ . Objectif : 63x**

cellules ont été récupérés confirmant la présence de levures doublement fluorescentes. Dans le but de s'assurer que l'apparition de levures fluorescente en RFP était réellement due à une interaction avec *B. subtilis* CW22, nous avons réalisé un échantillon contrôle dans lequel 50  $\mu\text{L}$  de milieu SMM1X (sans protoplastes de *B. subtilis* CW22) ont été ajoutés à 200  $\mu\text{L}$  de sphéroplastes de levures (cl 7.5). Ce tube a ensuite été traité comme les autres tubes. Sur ce contrôle, aucune levure n'est apparue doublement fluorescente. Cela nous a permis de valider le fait que l'apparition de levures fluorescentes en RFP était bel et bien due à la présence de la bactérie *B. subtilis* CW22, confortant le scénario suivant : contact membrane/membrane puis déversement du contenu bactérien dans le cytosol de la levure.

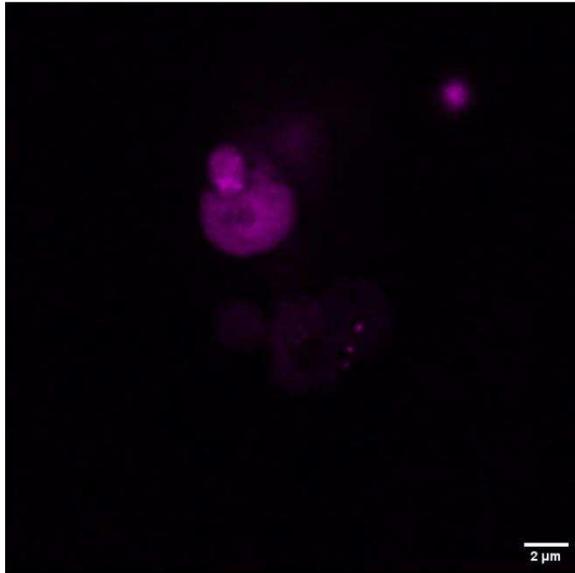
Pour mieux comprendre cet évènement de transfert entre bactéries et levures, des observations ont été réalisées lors du contact bactéries/levures/PEG (étape qui précède l'étape du SOS). A notre grande surprise, deux phénomènes différents ont pu être visualisés à cette étape :

- Le premier phénomène identique à celui observé précédemment en milieu SOS (que ce soit en microscopie à épifluorescence et en microscopie confocale) est appelé désormais : « transfert d'ADN par déversement du contenu bactérien » (figure 63). Les images capturées lors de cette séance semblent montrer des levures déformées capable « d'engloutir » des protoplastes, lesquels déversent ensuite leur contenu cellulaire dans le cytoplasme de l'hôte, le rendant fluorescent en RFP (figure 63).
- Le second phénomène, très différent du premier, montre très vraisemblablement une cellule entière de *B. subtilis* à l'intérieur d'une levure (figure 64). Les images acquises ici ne nous permettent pas de décrire le processus aboutissant à l'installation de *B. subtilis* au sein de la levure. Dès lors, plusieurs scénarios peuvent être envisagés. Nous émettons comme hypothèse que cette inclusion intracellulaire est le résultat d'une invagination de la membrane cytoplasmique suivie de la formation d'une vésicule d'endocytose. C'est pourquoi nous avons qualifié ce type d'évènement : transfert d'ADN par « endocytose ».

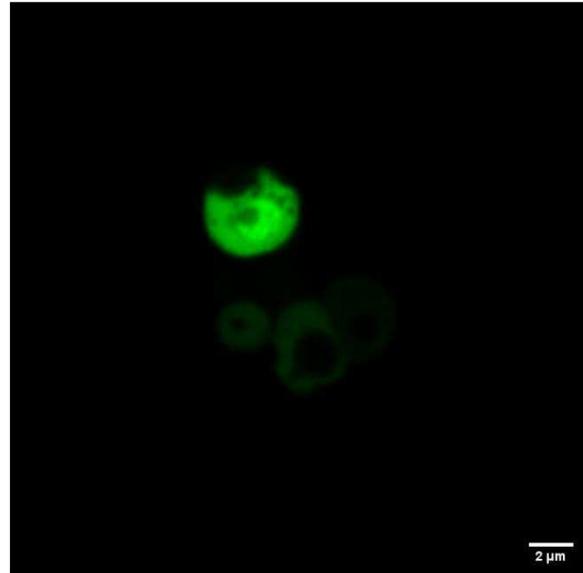
Cette dernière expérience a montré qu'il y avait potentiellement plusieurs phénomènes entraînant l'apparition de levures fluorescentes à la fois en GFP et en RFP.

Cette expérience de microscopie nous a permis de confirmer que c'est au moment de l'ajout du PEG que les interactions protoplastes/sphéroplastes ont lieu. Les images acquises

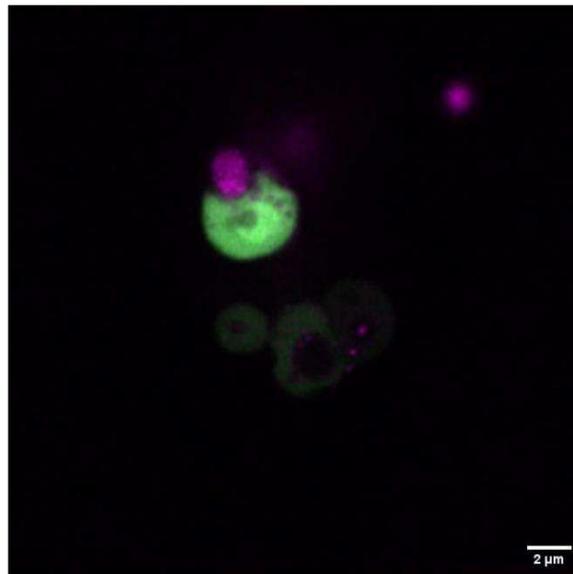
**A** *B. subtilis* RFP/*S. cerevisiae* GFP "deversement"  
en PEG observation avec filtre RFP



**B** *B. subtilis* RFP/*S. cerevisiae* GFP "deversement"  
en PEG observation avec filtre GFP



**C** *B. subtilis* RFP/*S. cerevisiae* GFP "deversement"  
PEG superposition des deux premières images



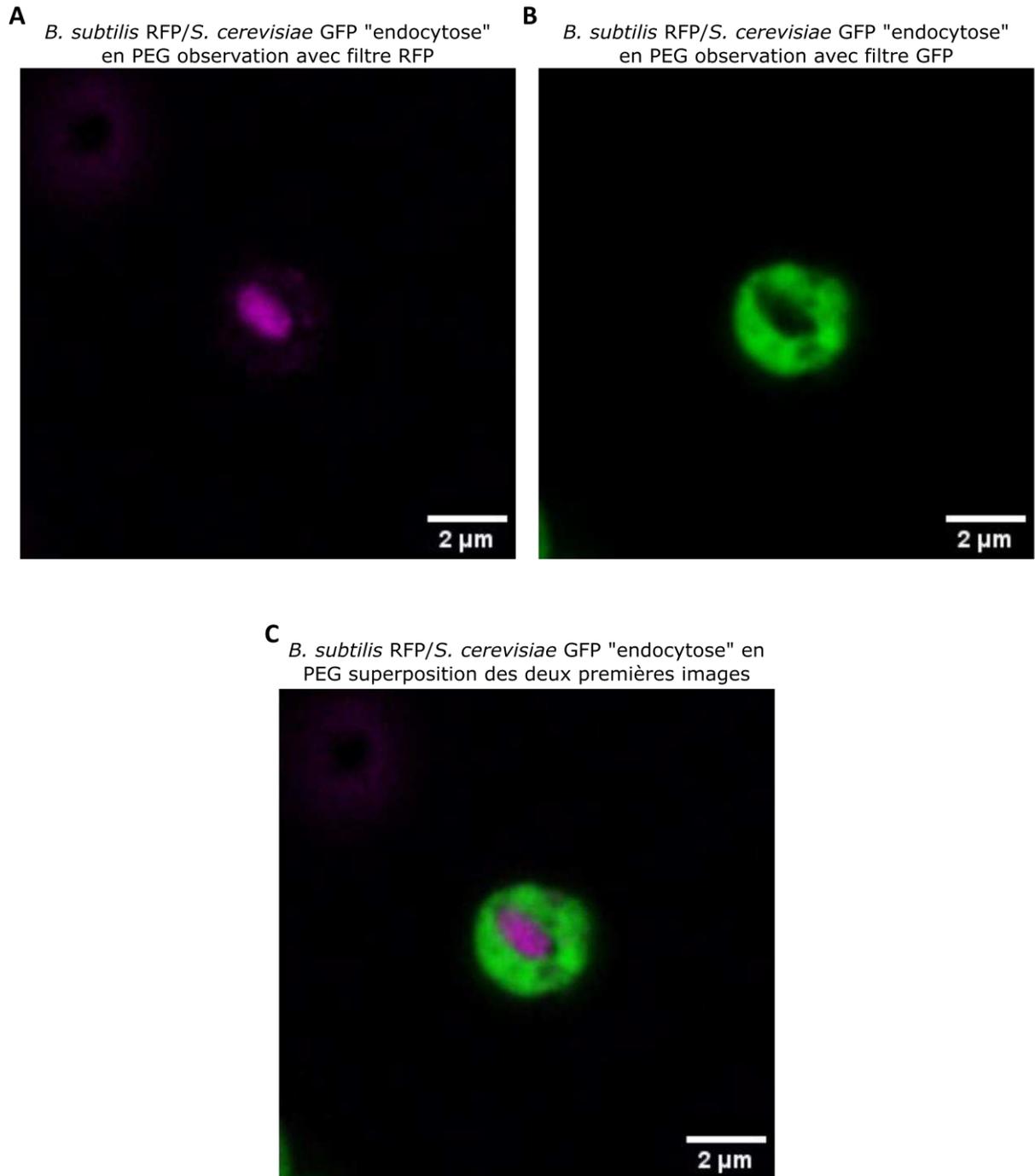
**Figure 63 : Observation en microscopie confocale : Événement de « déversement » lors de l'étape de contact entre les protoplastes de *B. subtilis* et les sphéroplastes de levure, en présence de PEG. [A] Observation en RFP. [B] Observation en GFP. [C] Superposition des images obtenues en RFP et en GFP. Echelle : 2 μm. Objectif : 63x**

semblent montrer deux types d'évènements : (i) contact protoplastes/sphéroplastes puis déversement du contenu bactérien dans la levure et (ii) contact protoplastes/sphéroplastes puis entrée de la bactérie par un phénomène apparenté à de l'endocytose. Ce second scénario n'a été observé qu'à l'étape du PEG. Il est possible qu'il soit rare et que des observations supplémentaires soient nécessaires pour l'observer en milieu SOS. Une autre possibilité est qu'il soit de courte durée et qu'il ne soit plus visible en milieu SOS. Les levures ayant « endocyté » des bactéries ne sont peut-être pas viables et meurent rapidement. Il est possible également que les bactéries « endocytées » soient désagrégées entraînant soit la mort des levures, soit le relargage de leur contenu cellulaire dans le cytoplasme de la levure la rendant RFP positive en milieu SOS.

Quel que soit le processus mis en jeu, à ce jour, il nous est impossible de déterminer quel scénario est propice à du transfert/capture d'ADN et donc à l'obtention de clones levure hébergeant un fragment du génome de *B. subtilis* lors des expériences de capture par CReasPy-Fusion : le déversement ? l'endocytose ? les deux ? aucun des deux signifiant qu'un autre mécanisme (non identifié) existe ?

## 6. Conclusion

Au cours de ce chapitre, nous avons décrit l'adaptation de la méthode de CReasPy-Fusion, développée à l'origine chez les mycoplasmes (Guesdon *et al.*, 2023), à *B. subtilis*. Passer directement des mycoplasmes à *B. subtilis* était un choix ambitieux puisque ces microorganismes possèdent des caractéristiques bien distinctes. Les mycoplasmes sont de petites bactéries sans paroi, qui abritent des génomes de petite taille (<1.5Mb), à faible pourcentage G+C ( $\leq 40\%$ ), qui utilisent un code génétique non standard, alors que *B. subtilis* est une bactérie à paroi Gram+ possédant un grand génome (2,86 Mb pour la souche à génome réduit MGP192 et 3,98 Mb pour la souche MS), à fort pourcentage en G+C (44 %) et utilisant un code génétique universel. Cependant, comme des travaux antérieurs suggéraient que le génome de *B. subtilis* MGP192 pouvait entrer dans la levure par la méthode de fusion et s'y maintenir entier le temps de quelques divisions cellulaires, nous pensions possible le transfert de la méthode à cet organisme pour capturer et modifier des fragments de son génome sans contrainte.



**Figure 64 : Observation en microscopie confocale : Événement « d'endocytose » lors de l'étape de contact entre les protoplastes de *B. subtilis* et les sphéroplastes de levure, en présence de PEG. [A] Observation en RFP. [B] Observation en GFP. [C] Superposition des images obtenues en RFP et en GFP. Echelle : 2 µm. Objectif : 63x**

Une démarche méthodique et ordonnée a été pensée et appliquée. Elle comprenait trois grands volets :

- **Le premier volet (étapes A et B)** a consisté à revenir vers des expériences simples mettant en jeu un plasmide navette pouvant se propager à la fois chez *B. subtilis* et chez la levure, avec pour premier objectif de confirmer que des molécules d'ADN pouvaient être transférées de *B. subtilis* à la levure *S. cerevisiae* après contact protoplastes/sphéroplastes/PEG (étape A) et pour second objectif de montrer que les molécules d'ADN transférées pouvaient être subséquemment modifiées par le système d'édition CRISPR-Cas9 portés par les sphéroplastes de levure (étape B). Aucune difficulté n'a été rencontrée lors de ces expériences, prouvant que dans des conditions adéquates la bactérie Gram+ *B. subtilis* et la cellule eucaryote *S. cerevisiae* peuvent échanger de l'ADN après contact membrane/membrane. L'utilisation de DNase I au cours de ces expériences élimine totalement la possibilité d'un transfert de plasmide suite à une transformation de la levure avec des molécules d'ADN libres dans l'environnement. L'expérience réalisée au cours de l'étape B constitue la première CREasPy-Fusion réalisée avec succès chez *B. subtilis*, mais appliquée à une molécule plasmidique.
- **Le second volet (étapes C et D)** a consisté à tenter de capturer de façon spécifique non plus des plasmides mais des fragments de génomes issus de 3 souches différentes de *B. subtilis*. Dans un premier temps, la région du génome située entre les gènes *atpA* et *irgB*, pré-marquée avec les éléments levure provenant des souches MS (145 kb), FE220 (145 kb) et MGP192 (125 kb), a été capturée avec succès dans la levure (étape C). Encouragée par ce résultat, la capture d'un fragment marqué de plus grande taille (268 kb) compris entre les gènes *yvcB* et *irgB* a été tentée. Ce fragment a également été cloné avec succès dans la levure mais seulement lorsqu'il provenait de la souche MGP192. L'hypothèse la plus probable expliquant ce résultat est que le fragment équivalent provenant des souches MS et FE220, fragments automatiquement plus grands (350 kb), contiendrait des gènes toxiques pour la levure ou/et des éléments générant de l'instabilité (séquences répétées par exemple). Après avoir montré que la CREasPy-Fusion fonctionnait chez *B. subtilis*, nous avons souhaité, dans un second temps, proposer une version allégée du protocole qui ne nécessiterait pas de pré-



marquer le génome (étape D). Nous avons alors tenté de capturer la première région, celle située entre les gènes *atpA* et *irgB*, lorsqu'elle n'était pas pré-marquée avec les éléments levure, en les ajoutant au cours du clonage. Nous avons réussi à isoler des clones levure hébergeant le fragment cible mais avons noté des différences d'efficacité significative entre les deux expériences. En effet, la capture du fragment de génome *atpA /irgB* non marqué a généré environ 3% de clones positifs alors que la capture du même fragment pré-marqué a généré environ 33% de clones positifs. Ce résultat indique qu'il est plus difficile de capturer un fragment d'un mélange complexe lorsque ce dernier n'est pas marqué, même avec l'aide d'un système de contre-sélection (5'FOA dans notre cas). Une marge de progression est peut-être possible en optimisant certaines étapes du protocole et/ou en testant d'autres systèmes de contre-sélection.

L'ensemble des expériences de capture de fragment de génome de *B. subtilis* dans la levure par CREasPy-Fusion nous a montré que nous étions capables (i) de transférer par contact cellules/cellules le génome entier de *B. subtilis* des souches MS (3.98 Mb), FE220 (3.96 Mb), et MGP192 (2.86 Mb) dans la levure, (ii) d'en isoler spécifiquement un fragment ; dans nos exemples, des fragments d'une taille de 145 kb pour les souches FE220 et MS et d'une taille de 268 kb pour la souche MGP192, puis (iii) si nécessaire, d'ajouter les éléments levure au cours de la capture.

La méthode de CREasPy-Fusion peut donc être appliquée à la bactérie Gram+ *B. subtilis*. C'est un résultat majeur pour le projet Bacillus 2.0. En effet, les travaux de thèse du Dr. Guesdon avaient montré que nous ne pouvions pas cloner le génome de *B. subtilis* MGP192 dans la levure dans son entièreté. Suite à ce résultat, des expériences de TAR-cloning avaient été réalisées dans le but de capturer ce génome en morceaux et tenter de localiser des régions/*loci* problématiques. Une première banque de levure (d'une vingtaine de clones) propageant différents fragments ou sous-fragments du génome (de 6 kb à 515 kb) avait ainsi été produite et nous avait permis de commencer à poser des hypothèses sur les raisons de nos échecs répétés à cloner ce génome. Mais l'extension de la banque a rapidement été limitée par la technique du TAR-cloning elle-même, qui est soumise à la présence des sites de restriction et ne permettait pas de choisir strictement la portion chromosomique que nous souhaitions capturer. La méthode CREasPy-Fusion, développée ici, permet de contourner ce problème puisqu'il est désormais possible d'isoler un fragment d'intérêt en choisissant



précisément la localisation du ou des sites de coupure. Nous pouvons désormais imaginer utiliser cette méthode pour cloner n'importe quelle zone du génome de *B. subtilis* MGP192 et possiblement de n'importe quelle taille, même si ce paramètre reste encore à étudier. Chez les mycoplasmes, le plus grand génome cloné en CREasPy-Fusion a une taille d'1.2 Mb (Guesdon *et al.*, 2023). Ce résultat laisse penser que des quarts de génomes *B. subtilis* MGP192 (0.715 Mb) voire des demi-génomes *B. subtilis* MGP192 (1.43 Mb) pourraient être capturés dans la levure si ces derniers ne contiennent pas de zones conflictuelles. Dans le cas contraire, des mesures correctives pourraient être effectuées sur le génome *B. subtilis* MGP192 soit en amont de la CREasPy-Fusion par transformation de la bactérie, soit pendant la CREasPy-Fusion grâce au système d'édition CRISPR-Cas9 présent dans les levures. Ainsi nous pouvons imaginer ajouter des ARS le long du génome si ces dernières font défaut, ou éliminer les *loci* bloquant le clonage de certaines portions du génome (*e.g.* opérons ribosomiques).

La méthode CREasPy-Fusion présente de nombreux avantages dans le cadre de ce projet. Elle a aussi ses limites, puisque des facteurs réduisant son efficacité ont été identifiés. On peut citer par exemple la présence de gènes codant certaines nucléases chez certaines espèces bactériennes (Guesdon *et al.*, 2023). Nos résultats montrent que le seul système de restriction modification identifié dans le génome de *B. subtilis* ne semble pas être problématique dans nos expériences (souche FE220 vs souche MS). Il est important de noter que ce système est absent de la souche *B. subtilis* MGP192.

- **Le troisième et dernier volet (étape E)** de cette étude a consisté à réaliser des observations en microscopie à fluorescence afin de documenter les étapes permettant de passer d'un contact protoplastes/sphéroplastes/PEG à de l'ADN bactérien cloné dans la levure. Ce travail n'est pas terminé mais les expériences réalisées au cours de ma thèse, nous ont déjà permis d'acquérir un ensemble d'images qui pourraient refléter, à ce stade, la mise en jeu de deux types d'interactions protoplastes/sphéroplastes:
  - Le premier au cours duquel la cellule bactérienne semble déverser son contenu directement dans le cytoplasme de la levure et que nous appelons « transfert d'ADN par déversement ». Un tel mécanisme impliquerait un contact protoplastes/sphéroplastes suivi d'une fusion des membranes phospholipidiques.



- Le second au cours duquel la bactérie entière est « endocytée » par la levure formant une vésicule interne emprisonnant la bactérie et que nous appelons « transfert d'ADN par endocytose ». Un tel mécanisme impliquerait une invagination de la membrane cytoplasmique de la levure, suivie par la formation d'une vésicule interne qui se détacherait de la membrane pour rejoindre le cytosol. A ce stade, le devenir de la vésicule n'est pas connu. Il est possible que la vésicule migre vers le noyau pour y déposer le nucléoïde bactérien. Il est également possible que la vésicule se désagrège dans le cytoplasme de la levure, aboutissant alors, comme dans le premier cas de figure, à un déversement du contenu bactérien dans le cytosol de la levure. Il est à noter que, dans ce scénario, le nucléoïde bactérien doit encore rejoindre le noyau de la levure. Par quels mécanismes ? Tout reste à élucider.

Ces résultats viennent étayer les premières observations réalisées par B. Karas *et al.*, en microscopie électronique lors du contact entre mycoplasmes et levures (Karas *et al.*, 2019). Il est évident que des études supplémentaires sont nécessaires pour identifier les mécanismes qui sous-tendent le processus de CREasPy-Fusion. Pour aller plus loin, nous envisageons d'une part de marquer, parallèlement aux bactéries et levures, l'ADN bactérien au DAPI afin de suivre son devenir au cours d'une CREasPy-Fusion en microscopie à fluorescence. D'autre part, des expériences de microscopie corrélative (CLEM), qui allie microscopie confocale et microscopie électronique, sont également en train d'être discutées. Cette technique permet de repérer un événement d'intérêt en microscopie confocale puis de le caractériser au niveau ultra-structural en microscopie électronique. Cependant, sa mise en place peut prendre du temps si l'on veut obtenir des résultats convenables. Il faut, entre autres, définir la méthode de fixation pour que l'échantillon reste viable et que les événements d'intérêt soient toujours observables.

Ce type d'observation pourrait nous permettre de définir s'il y a une réelle fusion entre les protoplastes de *B. subtilis* et les sphéroplastes de levure. Nous pourrions observer en microscopie électronique soit (i) une fusion membranaire entre la bactérie et la levure soit (ii) la présence de la bactérie dans la levure lors de l'événement « d'endocytose ». Par la suite, cette méthode pourrait nous permettre de comprendre le devenir de la bactérie après



endocytose : la vésicule contenant les protoplastes de *B. subtilis* (ou les protoplastes de *B. subtilis* eux-mêmes) fusionne-t-elle avec la membrane nucléaire de la levure pour y délivrer l'ADN bactérien ? Est-ce que l'ADN bactérien est libéré dans le cytoplasme de la levure et migre ensuite vers le noyau de la levure ? Est-ce que ces structures « Euca/proca-ryotiques » ne sont pas viables et meurent rapidement?

Une publication rassemblant l'ensemble de ces travaux est en cours d'écriture. Elle devrait être soumise dans une revue scientifique évaluée par les pairs d'ici le mois de décembre 2024.

**Titre possible:**

**Cloning of *Bacillus subtilis* large genome fragments in yeast using CReasPy-Fusion**



## Chapitre 2 : Utilisation de la méthode de CReasPy-Fusion pour le clonage de fragments de génome de *B. subtilis* de grande taille dans la levure



## Chapitre 2 : Utilisation de la méthode de CReasPy-Fusion pour le clonage de fragments de génome de *B. subtilis* de grande taille dans la levure

### 1. Introduction

Comme expliqué dans le chapitre 1, un des principaux objectifs du projet « Bacillus 2.0 » était désormais de réussir à cloner un génome réduit de *B. subtilis* dans la levure, et que ce génome se maintienne de façon stable chez cet hôte au cours du temps. Mais suite à nos multiples échecs, nous pouvions nous questionner sur la faisabilité d'une telle étape. Était-ce réellement possible de cloner un tel génome dans la levure ? Comment y parvenir ? Ces deux questions méritent d'être posées et réfléchies pour les raisons suivantes :

- La levure *S. cerevisiae* est un organisme unicellulaire eucaryote qui possède un noyau abritant 16 chromosomes (12 Mb au total) dont le plus grand a une taille proche d'1,531 Mb dans la souche de référence S288C (Engel *et al.*, 2014). Cet organisme a depuis longtemps été utilisé comme hôte pour propager et modifier le matériel génétique d'autres organismes, que ce soit sous forme de chromosomes artificiels de levure (YAC) (par exemple de l'ADN humain (Larionov *et al.*, 1996)) ou de plasmides centromériques de levure (YCps) (par exemple des génomes de mycoplasmes (Benders *et al.*, 2010) ou synthétiques (Gibson *et al.*, 2010)) (Vashee *et al.*, 2020). A ce jour, les plus grandes molécules exogènes clonées dans cet organisme ont des tailles de 2,3 Mb pour un YAC (molécule linéaire) (Marschall *et al.*, 1999) et de 1.8 Mb pour un YCp (molécule circulaire) (Karas *et al.*, 2013; Labroussaa *et al.*, 2016). On peut donc légitimement se demander si la levure peut héberger une molécule d'ADN surnuméraire d'environ 2.8 Mb, soit environ 1.3 Mb plus grand que son plus grand chromosome.

- Si maintenant on décide de s'affranchir de cette barrière de taille et de considérer qu'une molécule de 2.8 Mb peut être introduite et maintenue dans la levure sans répercussion sur la viabilité des cellules, alors comment expliquer que tous les essais réalisés pour tenter de cloner le génome entier de *B. subtilis* MGP192 dans la levure ont échoués jusqu'à présent ? Les expériences de TAR-cloning menées pendant la thèse du Dr. Guesdon ont permis d'apporter quelques éléments de réponse et de poser des hypothèses quant aux facteurs



responsables de nos échecs (des hypothèses qui pour certaines étaient fortement appuyées par des données de la littérature). Ainsi, la taille mise à part, nos résultats suggéraient :

- Premièrement, que plusieurs portions du génome de *B. subtilis* MGP192 ne possédaient pas suffisamment d'ARS fonctionnelles pour supporter la réplication correcte du génome complet dans la levure. Pour rappel, le nombre d'ARS (recherché via le motif ACS pour ARS Consensus Sequence, 5'-WTTTATRTTTW-3') dans le génome de MGP192, dont le GC% est proche de 44%, est estimée à environ 36 (soit en moyenne une ARS tous les 80 kb), mais il est très probable que les éléments ARS-like identifiés dans le génome de *B. subtilis* ne soient pas tous actifs dans la levure et donc que la distance séparant deux ARS fonctionnelles soit supérieure à 300 kb, un point critique chez la levure (Dershowitz and Newlon, 1993). Les fragments 1 et 7 semblent particulièrement concernés par ce problème mais le génome dans son ensemble doit lui aussi pâtir de ce déficit en ARS (figure 2, de l'avant-propos). De façon intéressante, il est décrit dans la littérature que les génomes artificiels et riches en bases G+C de *C. crescentus* (0,79 Mb et 57 % G+C) (Venetz *et al.*, 2019) et *E. coli* (1.03 Mb et 51 % G+C) (Zhou *et al.*, 2016), qui sont environ 3 fois moins grands que celui de MGP192, ont nécessité l'ajout d'une dizaine d'ARS pour être clonés et maintenus intacts dans la levure.
- Deuxièmement, que des *loci* fortement répétés dans le génome de MGP192 servaient de sites pour des événements de recombinaison homologue entraînant la perte de certaines régions du génome bactérien (exemple des opérons ribosomiques sur le fragment 7).
- Troisièmement, que des gènes dont l'expression est délétère pour la levure étaient présents dans le génome de MGP192. Les génomes bactériens ne sont pas censés s'exprimer chez la levure, mais le génome de *B. subtilis* MGP192 utilise le code génétique universel tout comme *S. cerevisiae*, et ceci peut favoriser la lecture de certaines séquences dont le produit d'expression peut être nocif pour la levure. Un tel phénomène a notamment été décrit lors du clonage du génome d'*A. laidlawii* dans *S. cerevisiae* (Karas *et al.*, 2012).

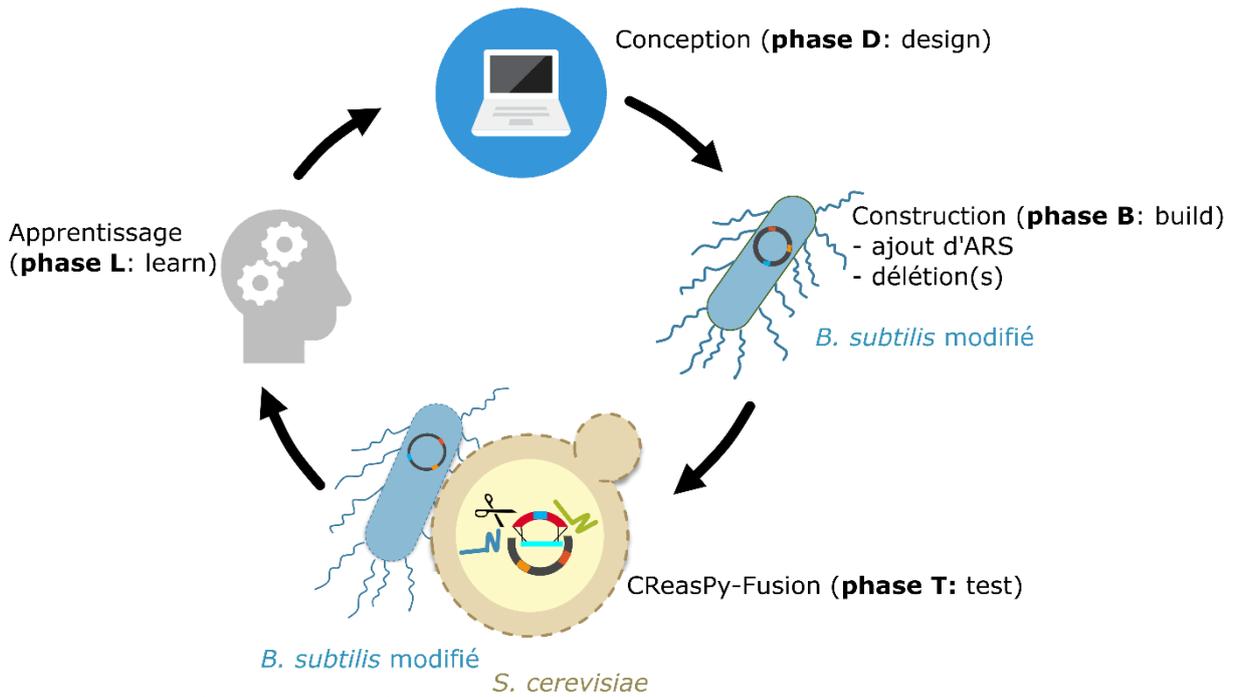


Figure 65 : Schéma de la stratégie mise en place pour construire un génome de *B. subtilis* qui puisse être cloné et maintenu stable dans la levure.

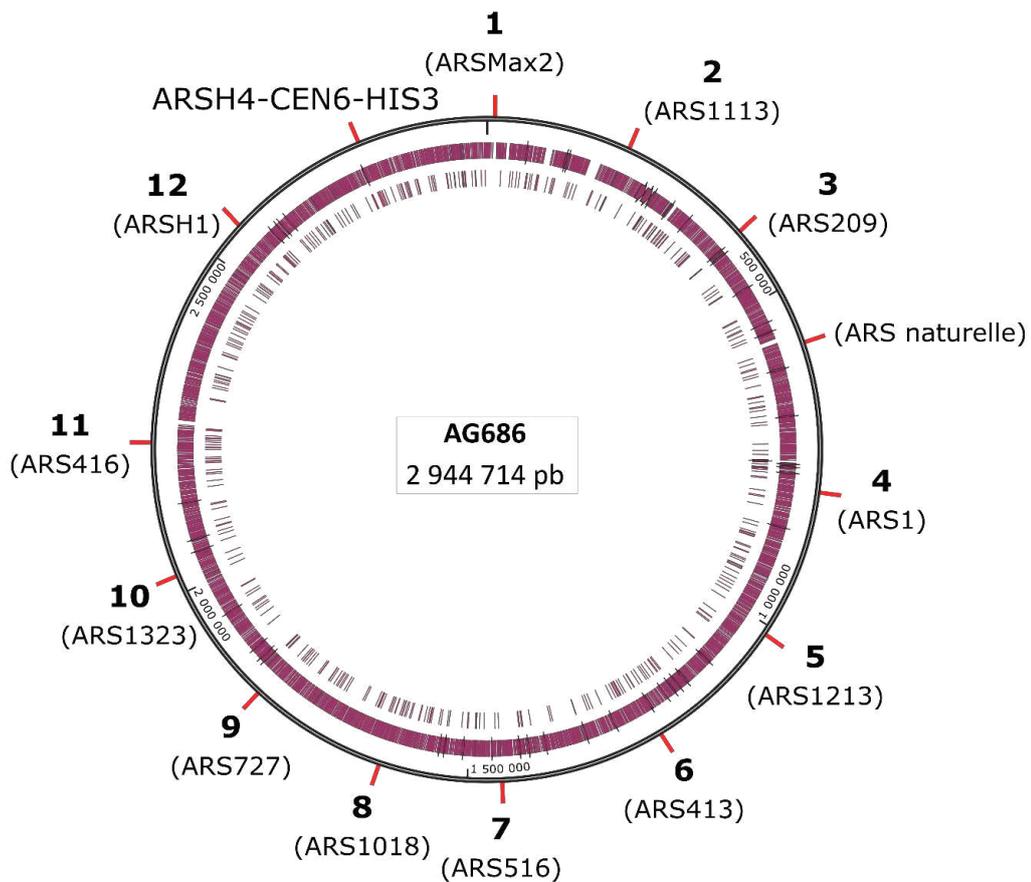


Figure 66 : Carte théorique du génome de la souche *B. subtilis* à construire d'après le premier design (design n°1). Les ARS supplémentaires sont indiquées par les numéros qui correspondent aux constructions répertoriées dans le tableau 21. La position des éléments levure ARSH4-CEN6-HIS3 tel que décrit chapitre 1 (cf partie 3.1 du chapitre 1) ainsi que celle de l'ARS naturelle sont aussi indiquées.

Pour tenter de répondre à ces questions essentielles (Est-ce possible ? et comment y parvenir ?), nous avons proposé, comme suite au travail amorcé, d'introduire des modifications dans le génome de *B. subtilis* puis d'utiliser la méthode de CREasPy-Fusion pour tenter de cloner « pas à pas », dans la levure, des fragments de plus en plus grands de ce génome. La méthode de CREasPy-Fusion semblait ici la plus pertinente puisque, comme démontré dans le chapitre 1, le point fort de cette méthode est de laisser le libre choix à l'expérimentateur de sélectionner les sites de coupure et de capturer à sa convenance n'importe quel fragment de génome.

Pour atteindre cet objectif, nous avons mis en place une stratégie basée sur le cycle DBTL (Design, Built, Test, Learn, figure 65) très utilisé dans le domaine de la BS. Cette stratégie en 4 étapes se décompose comme suit : (i) Conception *in silico* des modifications souhaitées (phase D), (ii) Insertion des modifications dans le génome de *B. subtilis* par transformation (phase B), (iii) Expériences de capture de fragments de génome modifiés dans la levure par CREasPy-Fusion (phase T), (iv) Apprentissage (phase L) : à cette étape, si la région du génome modifiée est clonée et maintenue stable dans la levure, alors la phase suivante consiste à capturer un fragment modifié de plus grande taille. En revanche, si la région modifiée n'est pas maintenue dans la levure, la phase suivante consiste à proposer un nouveau « design » (ex : ajout d'ARS supplémentaires, et/ou la délétion de certains gènes ou *loci* répétés) et répéter le cycle.

Une telle stratégie devrait nous permettre d'identifier progressivement les régions du génome qui posent problème, de proposer et tester des solutions et de valider celles qui fonctionnent, notre objectif étant de cloner pas à pas des fragments de plus en plus grands pour aller jusqu'à la capture du génome complet de *B. subtilis* dans la levure.

Le chapitre 2 décrit les efforts en cours pour déployer cette stratégie et atteindre nos objectifs. Comme planifié, le travail a démarré avec la conception sur ordinateur d'un premier génome de *B. subtilis* modifié. Ce premier design, issu de nos réflexions autour des expériences de TAR-cloning mais aussi des données de la littérature, proposait l'insertion de 13 ARS le long du génome de *B. subtilis* (figure 66) et d'un marqueur de sélection. Nous avons donc entrepris la construction de la souche *B. subtilis* conçue sur ordinateur (travail toujours en cours) et la capture, au fur et mesure des avancées, des fragments modifiés par CREasPy-Fusion. Les données acquises lors de nos différentes tentatives de capture de fragments de



génomique contenant des ARS supplémentaires ont conduit à réviser le premier design et à en proposer un second que l'on espère plus adapté.

Le travail présenté dans le chapitre 2 est réalisé en étroite collaboration avec l'équipe SyBER (Institut MICALIS, UMR 1319 INRAE / AgroParisTech, Jouy-en-Josas). Sur le plan théorique, toutes les décisions sont prises de manière collégiale. Sur le plan expérimental, le travail est partagé. Nos collaborateurs sont en charge de l'ajout des ARS (ou toutes autres modifications décidées en cours de route) dans le génome de *B. subtilis*. Leur connaissance de l'organisme et de son génome leur permet, en particulier, d'insérer les modifications souhaitées de façon précise dans le génome de *B. subtilis*, et à ne pas interférer avec les unités de transcription. De notre côté, nous sommes en charge des expériences de capture des fragments modifiés par CREasPy-Fusion et du processus de criblage. Tous les résultats sont partagés et discutés.

## 2. Design d'une souche *B. subtilis* possédant 13 ARS dans son génome

### 2.1. Choix de la souche de départ

Dans le chapitre précédent, nous avons montré que la CREasPy-Fusion permettait de capturer efficacement des fragments de génomes à partir de la souche minimale *B. subtilis* MGP192 (ce qui n'était pas le cas des souches parentales MS et FE220). Cette souche, dont le génome est très fortement réduit, a cependant un inconvénient majeur : elle est difficilement transformable. Elle semblait donc peu indiquée pour le travail à effectuer dans cette seconde partie, qui nécessitait l'insertion en cascade de nombreuses ARS (et marqueurs) dans le génome de *B. subtilis* (et sur du plus long terme, possiblement la délétion de plusieurs régions).

La stratégie globale reposant entièrement sur notre capacité à produire ces souches modifiées, il nous fallait donc identifier une souche de départ possédant les caractéristiques suivantes : un génome le plus réduit possible (pour les étapes de CREasPy-Fusion) et une compétence naturelle optimale (pour les étapes de transformation). Cela n'est pas simple car les souches à génome réduit sont généralement peu compétentes.

Nous avons opté pour la souche MGP144 de la collection « Minimal Genome Project » dont le génome restait de petite taille de (2,944 Mb) et qui était plus facilement

**Tableau 21 : Liste des ARS supplémentaires à ajouter au génome de *B. subtilis* d'après le design n°1.**

Numéro de la construction	Gène proche	Nom de l'ARS	Séquence de l'ARS	Références
1	<i>gyrA</i>	ARSMaX2	TTCAAATTTTATGTTTTGTTTCGAGATCCACAGTTTTTCGGAGCATAG GAAGCTCGTATATGATAACAAAAAATAATATATAAAAAATAAAAA ATAATGA	Venez et al., 2019
2	<i>570</i>	ARS1113	CAATCTTCATTTGTTTACGTAATAATGAAAACGTTTTGTGTTAGAATT GACAGATTGAAGTTAATTTTAGGTGTCGGGAGAAACATTTAATCC CTTTTTATGAATATAATTTAAGTACTAAAAAATCACCATAATTATAC AAAGTAACTGCAATTTAAAAGTACGTTTATAATTAATTCACATACAC ATGTTATTGTATCTAACTAAATAGTCTCGAAATAGTTCAATTGA	Venez et al., 2019
3	<i>yerl</i>	ARS209	ACCTTTGTTTTATTTTTTTTCTTCATTCCGTAACCTTCTACCTTCTTTA TTTACTTTTCTAAAATCCAAATACAAAACATAAAAAAT	Venez et al., 2019
4	<i>oppA</i>	ARS1	AGCTTACATTTTATGTTAGCTGGTGGACTGACGCCAGAAAATGTTG GTGATGCGCTTAGATTAATGGCGTTATTGGTGTGATGTAAGCGG AGGTGTGGAGACAAATGGTGTAAGAACTCTAACAAAATAGCAAA TTTCGTCAAAAATGCTAAGAAATAGGTTATTACTGAGTAGTATTTAT TTAAGTATTGTTGTGCACTTGCCTGCAAGCCTTTTGAAGCAAGC ATAAAAGATCTAACATAAAATCTGTAATAACAAGATGTAAGA TAATGCTAAATCATTGGCTTTTATTGATTGTACAGGAAAATATA CATCGCAGGGGGTTGACTTTTACCATTTCACCGCAATGGAATCAAAC TTGTTGAAGAGAATGTTACAGGCGCATACGCTACAATGACCCGAT TCTTGCTAGCCTTTTCTCGGTCTTGCAAAACAACCGCCGCGCAGCTTAG TATATAAATACACATGTACATACCTCTCTCCGTATCCTCGTAATCATT TTCTTGATTTATCGTCTTTTCTGCTGTAATAACTTTATCACACTTATCT CAAATACACTTATTAACCGCTTTTACTATTATCTTCTACGCTGACAGT AATATCAAACAGTGACACATATTAACACAGTGGTTTTCTTGCATAA ACACCATCAGCCTCAAGTCGTCAAGTAAAGATTTCTGTTCATGCGAG ATAGATAACAATCTATATGTTGATAATTAGCGTTGCCATCAATGC GAGATCCGTTTAAACCGACCTAGTGCCTTACCCACGTTCCGGTCC ACTGTGTGCCGAACATGCTCCTTCACTATTTTAAACATGTGG	Liachko et al., 2013
5	<i>ykwD</i>	ARS1213	TGTTTACATTTCAAGTCTAATTCTGTGCTTTTACCGAAGAGGAATTTT CATAAATACGGAGAAAATATAAAAAAAGT	Venez et al., 2019
6	<i>ylqB</i>	ARS413	GAATAATAGAATGAGTTACATACGTAACGTTAGTATTATTATTAT TATTCTATTTCTTTCCGTTTTTATTATTTATATATATATATATATAT ATATATTTGGGTGTTTATTTAGGTAGGGCATAAGGATTTACTGTC GGCATGAACCTACTTGAACAAAAAATAAAAAATCAAAGAAAAATA ATGTAGGCTTTTGCACATCGACTATATATAAATATATATATACACAC TTCTGTGCAGAA	Liachko et al., 2013
7	<i>yocI</i>	ARS516	TTTTTCATAATATTTACTTTTTGGTGCAGAAATCTGTAATATTTTAAA ACAAAATCTCA	Liachko et al., 2013
8	<i>ypuC/2</i>	ARS1018	TCAGAGCAATACGTGTTTATTTTTCAACAATTCGCATGGCATTATGT ATCATTGCAGCGTAAATTTTACATTTAGTAATATCTAATAACATATA CACAAAAAAGGGACAGATTTTTGAATTAATAAACCAAGCGTAAAT CAGTATTGCTTTACCGCAGTAGGGGAGTTATCTACTCCTGCTTAA CGTAGAGAAACAATAGTAAAAGTTATGCTTGTGTTGGAATAAAAAT	Venez et al., 2019
9	<i>lepA</i>	ARS727	TTTTTTATTGTTTCTAATAACTTCATTTGTTACTTTTTGCCTTTTGT TATTATGAAATAATGCAGGTTAACATAAGTAAACATAAAAATATG	Venez et al., 2019
10	<i>infC</i>	ARS1323	CAATTAACCAAGTCCTTTTATTTTTAATGGTTTTATTATGAATTGTT TATGTTTTATGTTCTTTATACTCAAAATCAATAAGTGCTTATTTAGCG AATTTTCTTCTACGAATATATATAAGTCATATCTAAGACTTCCATCCG TGCTTTATTACGTTTGTATGATGCATACAATGATGCCAAATCAGCTT GATGCTGTTGGTATTCACTCGTC	Venez et al., 2019
11	<i>menC</i>	ARS416	GGCCTTTTGAAGCAAGCATAAAAAGTCTAAACATAAAAATCTGTA AAATAACAAG	Venez et al., 2019
12	<i>cwI/O</i>	ARSHI	TTACAGATTTTATGTTTAGATCTTTTATGCTTGCTTTTCAAAGGCCT GCGGGCAAGTGACAAACAATACTTAAATATATACTACTTAGTAATA ACCTA	Venez et al., 2019

transformable que la souche MGP192, mais avons choisi de la modifier, préalablement à notre travail, de façon à amplifier sa compétence naturelle et à sécuriser les premières étapes de transformation. A partir de la souche MGP144, nos collaborateurs ont donc construit la souche AG686 (pop088 *ykoS::comSK-tet*) qui sur-exprime deux gènes impliqués dans la compétence naturelle chez *B. subtilis*, les gènes *comSK* (Maier, 2020).

L'ensemble des travaux présentés ci-après, y compris le design, ont été réalisés à partir de la souche parentale AG686 et de son génome.

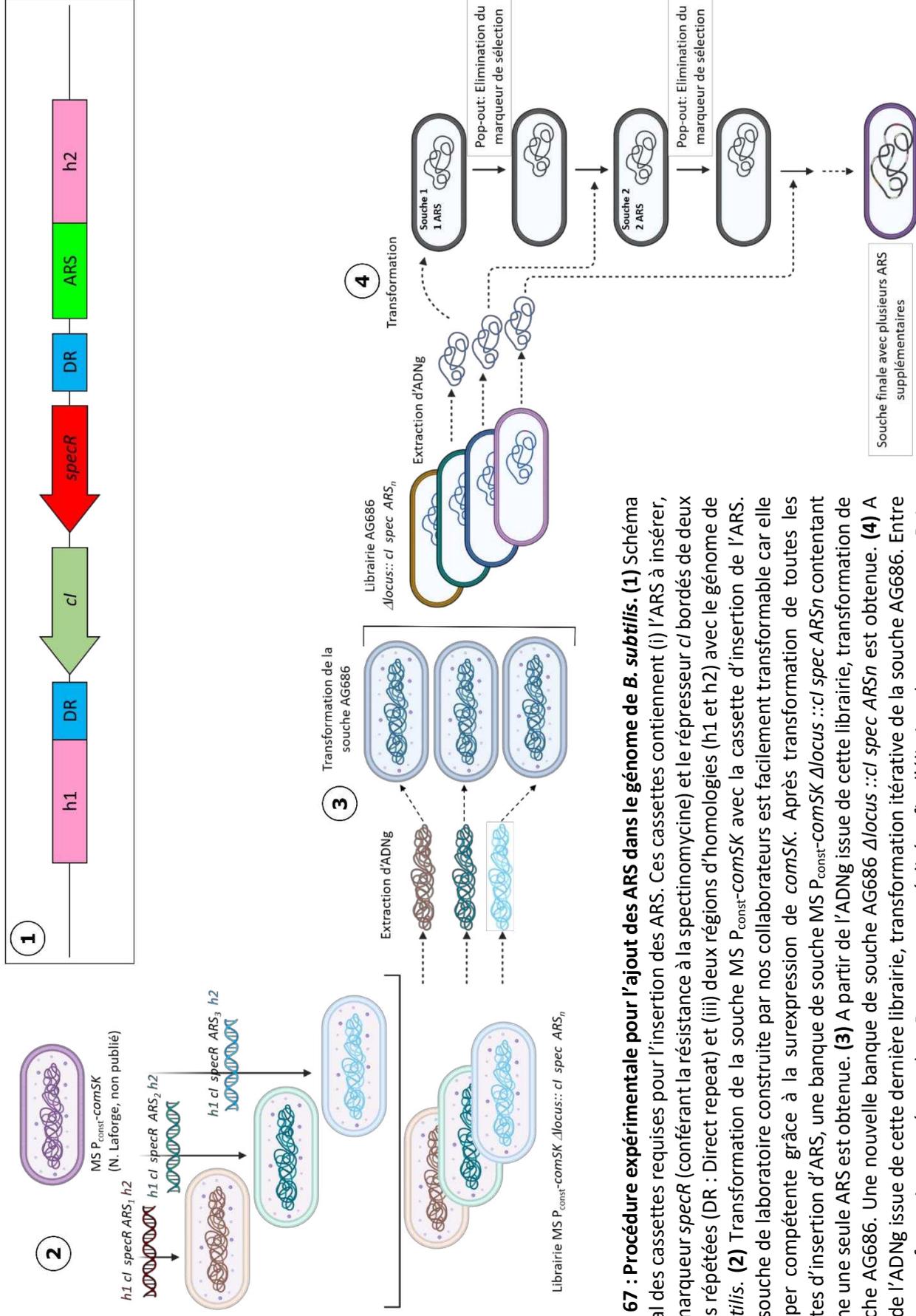
## 2.2. Conception sur ordinateur d'un premier génome modifié à partir du génome de la souche AG686

Une fois le génome de base identifié, le consortium a réfléchi aux modifications à y introduire pour réussir à le cloner dans la levure *S. cerevisiae*, avec deux questions constamment à l'esprit : le type de modifications à faire et leur emplacement sur le génome. Pour répondre à ces questions, le consortium s'est appuyé sur un ensemble de ressources, parmi elles :

(i) les expériences de TAR-cloning qui avaient permis d'identifier et lister les régions du génome les plus instables et qui nécessitaient d'être revues. Nous pouvons citer en particulier le fragment 7 (482 kb) et le fragment 1 (1500 kb) (figure 3 et partie 2 de l'avant-propos).

(ii) les analyses bio-informatiques réalisées sur ces régions instables (et le reste du génome) qui permettaient de faire des hypothèses sur les corrections à apporter (insertion d'ARS, insertion de marqueurs de sélection, délétions de séquences répétées et/ou possiblement toxiques).

(iii) les données de la littérature notamment celles issues du clonage du génome de *C. crescentus* (Venetz *et al.*, 2019). Le clonage de ce génome synthétique d'une taille de 0.79 Mb avait en effet nécessité l'ajout de 10 ARS et 5 marqueurs d'auxotrophie. L'ajout de ces ARS (et marqueurs), dont les séquences étaient disponibles dans la publication, avaient permis de stabiliser ce génome, elles étaient donc fonctionnelles. Ces séquences ont été récupérées afin d'être réutilisées (tableau 21). Par ailleurs, rapporté à la taille du génome de *B. subtilis* (2.944 Mb), la théorie voudrait qu'environ 36 ARS et 18 marqueurs d'auxotrophie soient ajoutés au génome. Cependant, ce n'est pas tant la taille du génome, mais le pourcentage en G+C qui est



**Figure 67 : Procédure expérimentale pour l'ajout des ARS dans le génome de *B. subtilis*. (1)** Schéma général des cassettes requises pour l'insertion des ARS. Ces cassettes contiennent (i) l'ARS à insérer, (ii) le marqueur *specR* (conférant la résistance à la spectinomycine) et le répresseur *cl* bordés de deux régions répétées (DR : Direct repeat) et (iii) deux régions d'homologies ( $h1$  et  $h2$ ) avec le génome de *B. subtilis*. **(2)** Transformation de la souche MS  $P_{const}$ -*comSK* avec la cassette d'insertion de l'ARS. Cette souche de laboratoire construite par nos collaborateurs est facilement transformable car elle est super compétente grâce à la surexpression de *comSK*. Après transformation de toutes les cassettes d'insertion d'ARS, une banque de souche MS  $P_{const}$ -*comSK*  $\Delta$ locus::*cl spec ARSn* contenant chacune une seule ARS est obtenue. **(3)** A partir de l'ADNg issue de cette librairie, transformation de la souche AG686. Une nouvelle banque de souche AG686  $\Delta$ locus::*cl spec ARSn* est obtenue. **(4)** A partir de l'ADNg issue de cette dernière librairie, transformation itérative de la souche AG686. Entre chaque transformation une étape de « Pop-out » est réalisée afin d'éliminer le marqueur *specR*. La souche contenant plusieurs ARS est obtenue à cette étape. Finalement cette méthode permet d'ajouter plusieurs ARS sans marqueur de sélection.

à prendre en compte ici. En effet, la probabilité de trouver des sites riches en AT susceptibles de fonctionner comme des ARS dans les génomes bactériens à faible teneur en G+C est plus élevée que dans les génomes bactériens à forte teneur en G+C. Le génome de *C. crescentus* possédant un pourcentage en G+C de 57% (contre 44% pour *B. subtilis*), il contient donc moins d'ARS putatives potentiellement actives que le génome de *B. subtilis*.

(iv) les données de la littérature issues des travaux de Dershowitz *et al.*, en 1993 (Dershowitz and Newlon, 1993), selon lesquels les ARS prédites comme les plus fortes sont capables de stabiliser des séquences d'ADN de 100 à 300 kb dans la levure, et ceux de Liachko *et al.*, en 2013 qui ont conçu de courtes séquences d'ADN synthétique conservant la fonction d'ARS (Liachko *et al.*, 2013).

L'ensemble de ces données a amené le consortium à proposer un génome de *B. subtilis* MGP192 intégrant les modifications suivantes : en complément de l'ARS de 17 nucléotides naturellement présente dans le génome de *B. subtilis*, ajout de 13 ARS supplémentaires réparties de manière homogène tout le long du génome avec une distance moyenne entre chaque ARS d'environ 200 kb. La plus grande distance séparant deux ARS (entre l'ARSH1 (construction 12) et l'ARS416 (construction 11)) était de 338 kb (tableau 21, figure 66). Cette distance était certes supérieure à 300 kb mais plusieurs ARS putatives avaient été identifiées dans cette région, nous faisons donc le pari que certaines d'entre elles étaient fonctionnelles.

### 2.3. Construction des souches de *B. subtilis* AG686 modifiées

Toutes les séquences ARS (et marqueurs) ont été insérées en suivant la procédure expérimentale en 4 étapes illustrée figure 67. Cette procédure permet d'introduire des modifications au génome de *B. subtilis* sans laisser de trace ou marqueur au site d'insertion (figure 67). Ce travail a été réalisé par nos collaborateurs de l'équipe SyBER (Institut MICALIS, UMR 1319 INRAE / AgroParisTech, Jouy-en-Josas). Il est présenté brièvement ci-dessous :

**Etape 1** : Construction de 13 cassettes d'ADN linéaires portant chacune une ARS spécifique. Plus précisément, ces cassettes sont composées (i) de l'ARS à insérer, (ii) du marqueur de résistance *specR* et du répresseur *cl* bordés de deux régions répétées de 100 pb (appelées Direct Repeat ou DR, nécessaires à l'élimination du marqueur de résistance lors de la dernière



étape), puis (iii) à chaque extrémité, deux séquences d'environ 1 kb (h1 et h2) identiques aux *loci* ciblés sur le génome de *B. subtilis* (figure 67).

**Etape 2 :** Construction d'une première banque de 13 clones contenant chacun une ARS spécifique à partir de la souche sauvage MS P<sub>const</sub>-comSK. Cette dernière correspond à la souche MS (3,98 Mb) qui sur-exprime les gènes *comSK*, sous le contrôle d'un promoteur constitutif. Elle possède un taux de transformation élevé avec des plasmides circulaires mais aussi des constructions d'ADN linéaires (figure 67).

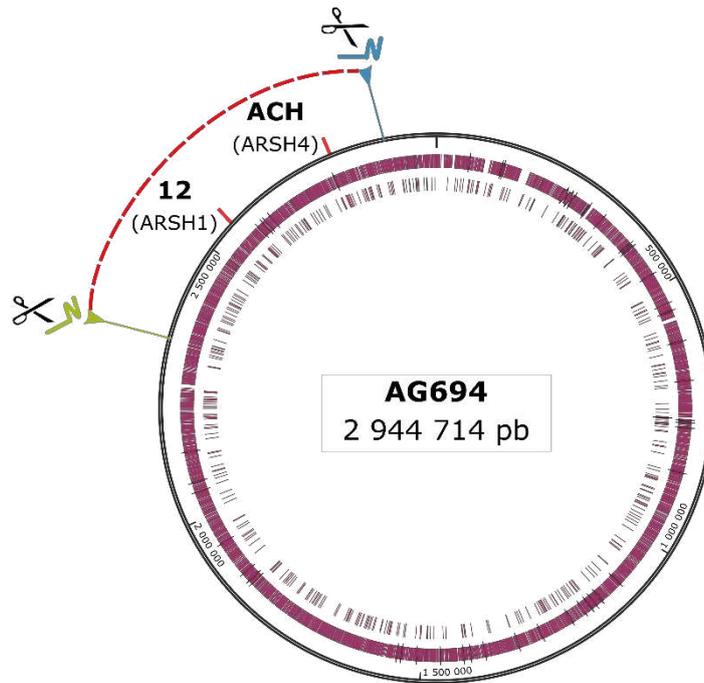
**Etape 3 :** Construction d'une seconde banque de 13 clones contenant chacun une ARS spécifique à partir de la souche AG686. Cette souche est difficilement transformable avec des plasmides circulaires ou des cassettes d'ADN linéaires. Pour obtenir les mutants d'intérêt, il est nécessaire de transformer cette souche avec l'ADNg modifié de la souche MS provenant de l'étape 2 (figure 67).

**Etape 4 :** Intégration itérative des ARS dans la souche à génome réduit à partir de l'ADNg issu de la seconde banque. A cette étape, l'intégration se fait grâce au marqueur de résistance *specR* (étape dite de Pop-in). Dans un second temps, le marqueur de résistance est éliminé par contre-sélection basée sur le système opérateur-répresseur utilisant le répresseur *cl* (étape dites de Pop-out, cf partie 1.7.2 de l'introduction). Une fois le marqueur éliminé, une autre ARS peut être insérée (figure 67). Il est ainsi possible d'insérer plusieurs ARS sans ajouter de marqueur de résistance. Cette méthode de Pop-in/Pop-out a été publiée en 2023 par Dervyn *et al.*, (Dervyn *et al.*, 2023).

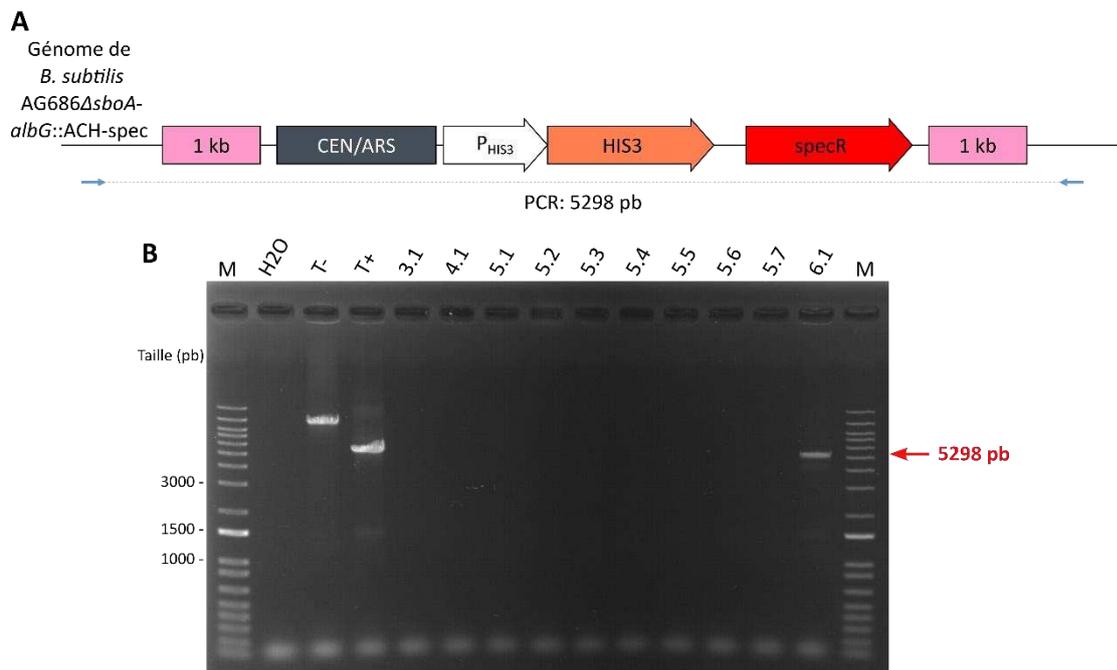
Pour toutes les souches possédant des ARS additionnelles, la dernière cassette insérée a toujours été la cassette portant les éléments levure ARSH4, CEN6 et HIS3 et le marqueur *specR* (cf partie 3.1 chapitre 1) de façon à générer une souche résistante à la spectinomycine.

### 3. Premier essai de capture d'un fragment de génome de *B. subtilis* de 536 kb modifié par l'ajout d'ARS

Dans le chapitre 1, nous avons montré la capture dans la levure d'un fragment de génome de 268 kb non modifié à partir de la souche *B. subtilis* MGP192. Pour progresser dans notre quête, nous avons souhaité capturer un fragment de 536 kb correspondant à la région entre les gènes *IrgB* et *yufO*.



**Figure 68 : Carte théorique du génome de la souche *B. subtilis* AG694 avec l'ARSH4 et l'ARSH1, basé sur le design n°1. La région qui correspond au fragment de 536 kb entre les gènes *irgB* et *yufO* est indiquée. ACH : ARSH4-CEN6-HIS3.**



**Figure 69 : Insertion des éléments levure dans le génome de la souche AG686. [A] Schéma des éléments levure insérés dans le génome de *B. subtilis* AG686. La PCR pour l'analyse des transformants est représentée. [B] Analyse des transformants de la souche *B. subtilis* AG686. L'intégration des éléments levure est validée par la PCR utilisant les amorces `screening_ARS_CEN_HIS_F` et `screening_ARS_CEN_HIS_R` en dehors de bras de recombinaison et produisant une bande à 5298 pb si les éléments levure sont bien insérés. « M » : DNA Ladder 1 kb+ Invitrogen (100-12000 pb). « H2O » : contrôle négatif, sans ADN. « T- » : témoin négatif, ADNg de la souche AG686 non modifiée. « T+ » : témoin positif, ADNg de la souche AG694 (qui contient déjà les éléments levure).**

Pour ce faire, nos collaborateurs ont construit une souche appelée AG694 (dérivant de AG686) contenant deux ARS localisées au niveau du fragment de 536 kb à capturer : l'ARSH1 (construction 12) et l'ARSH4 amenée avec les éléments levure CEN6 et HIS3 (figure 68, tableau 21). Parallèlement à ce travail, notre équipe a construit une souche ne contenant qu'une seule ARS pour servir de témoin à l'expérience, sa construction est décrite ci-après.

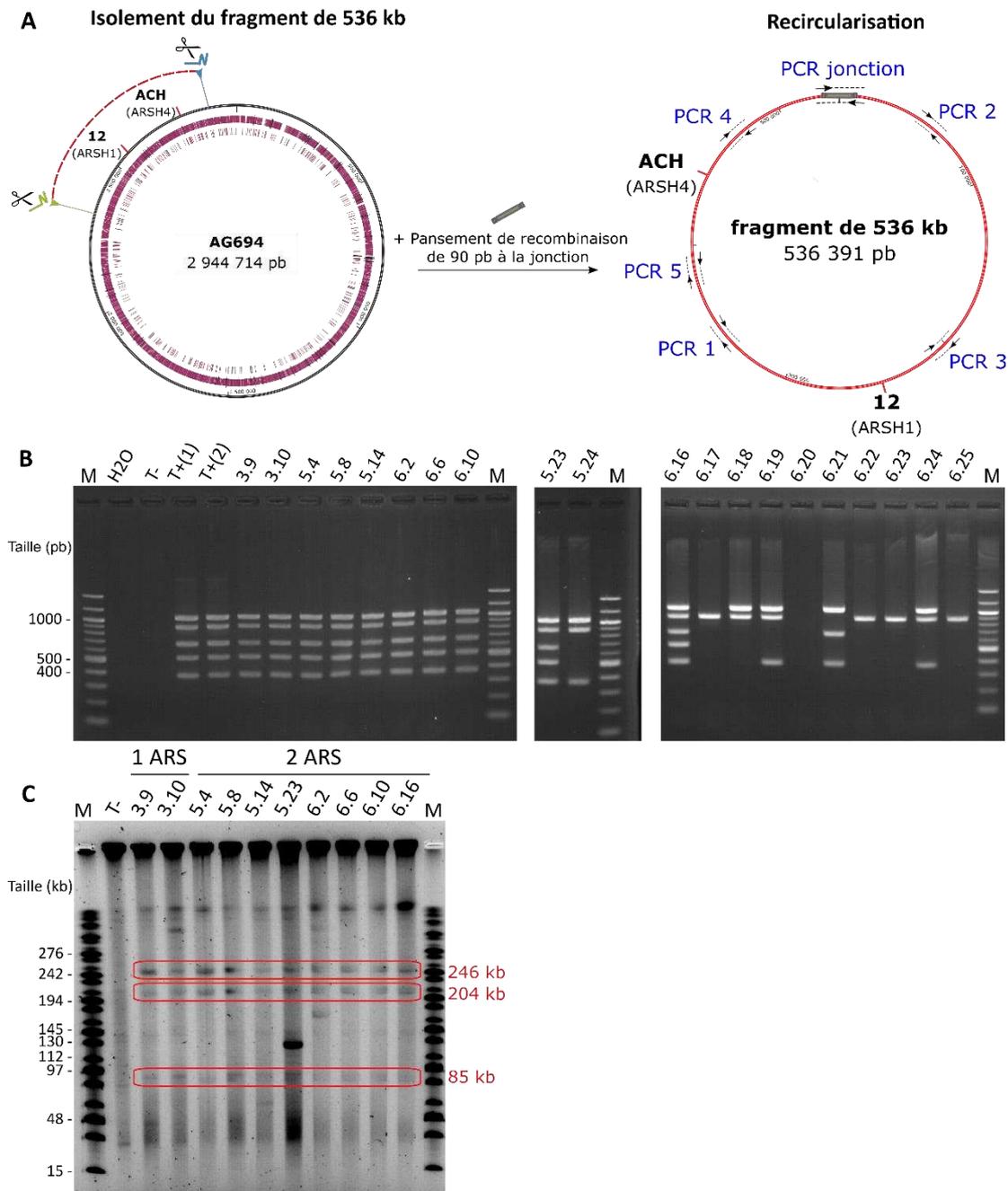
### 3.1. Construction d'une souche témoin avec une seule ARS

Afin de prouver que l'ajout d'ARS avait un réel impact sur l'efficacité de capture, une souche « contrôle » contenant seulement l'ARSH4 et les éléments levure CEN6 et HIS3 a été construite. Pour que cette souche contrôle soit la plus pertinente possible, la cassette d'ADN a été insérée dans le génome de la souche parentale AG686 au même locus que la souche AG694.

Pour y parvenir, la souche AG686 a été transformée avec le plasmide pMT85-pRS313-pMiniTetM-pSPEC sans lacZ-sbo ou avec de l'ADNg de la souche MGP192 $\Delta$ sbo-*alb*::ACH-spec obtenue dans le chapitre 1 (protocole MC). Les transformants ont été sélectionnés sur du milieu LB solide contenant 100  $\mu$ g/ml de spectinomycine. Après 24 h d'incubation à 37°C, 8 colonies sont apparues pour la condition de transformation avec l'ADNg de la souche MGP192 $\Delta$ sbo-*alb*::ACH-spec et 2 pour la condition de transformation avec le plasmide. Afin de vérifier l'intégration des éléments levure, ces 10 clones ont été analysés par PCR avec les amorces screening\_ARS\_CEN\_HIS\_F et screening\_ARS\_CEN\_HIS\_R permettant de mettre en évidence l'intégration complète de la cassette. Un clone (cl 6.1) a montré le profil attendu en PCR avec une bande unique à 5298 pb (figure 69A et B). Ce locus a été séquencé par la méthode de SANGER. L'analyse des séquences n'a montré aucune mutation sur la séquence analysée. La souche contrôle « AG686 $\Delta$ sbo-*alb*::ACH-spec » ne contenant qu'une seule ARS supplémentaire était donc prête pour les expériences de capture de fragments de génome par CREasPy-Fusion.

### 3.2. Capture du fragment de 536 kb par CREasPy-Fusion

Pour effectuer cette expérience, nous avons au préalable préparé tout le matériel biologique nécessaire à sa mise en place, à savoir :



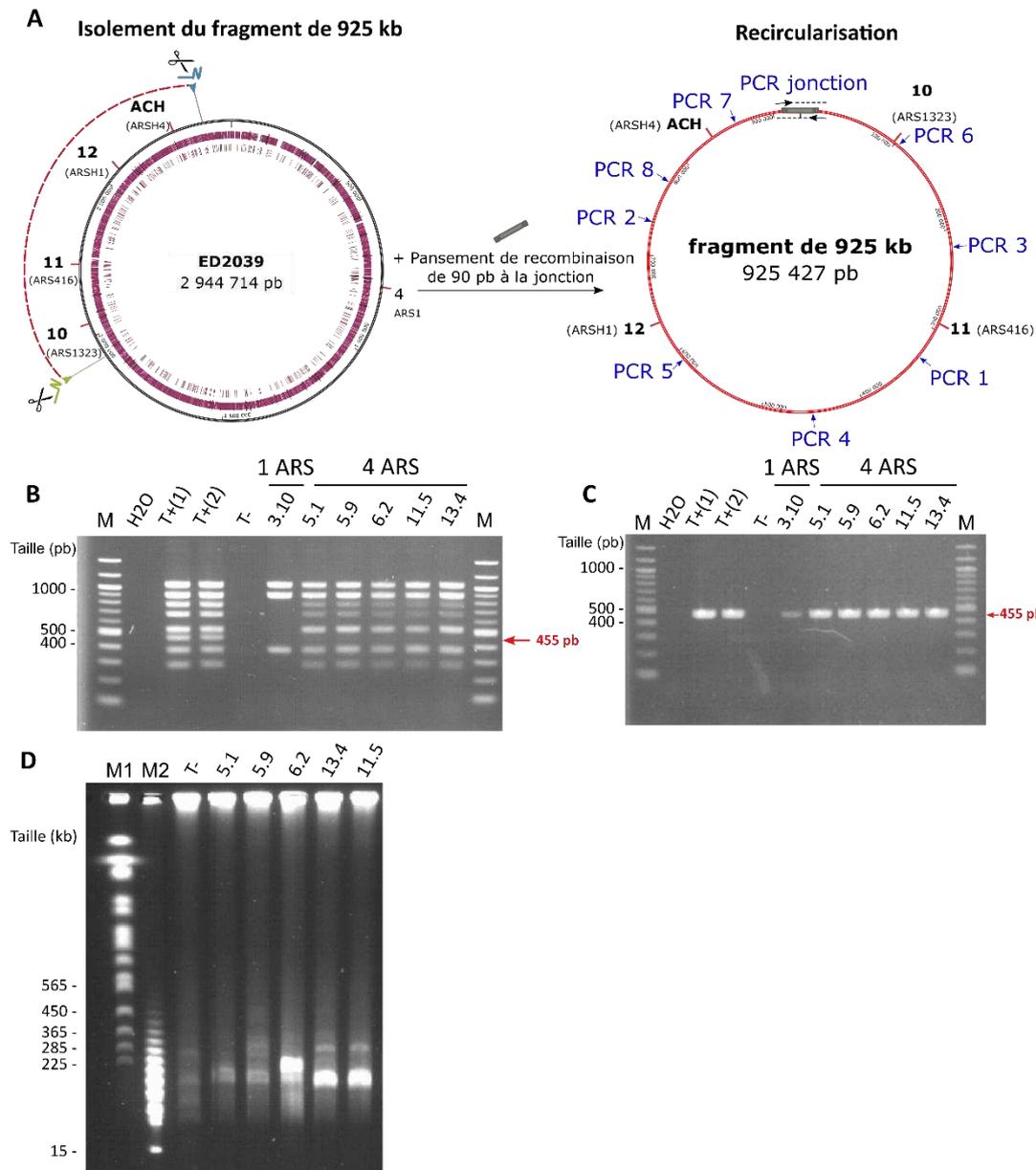
**Figure 70 : Capture par CReasPy-Fusion, d'un fragment du génome *B. subtilis* d'une taille de 536 kb.**

**[A]** Isolement du fragment par deux duplex ARNg-Cas9. Re-circularisation du fragment par le système de recombinaison homologue de la levure. Le pansement de recombinaison de 90 pb, co-transformé, est utilisé par la levure pour réparer les cassures double brin et former un fragment circulaire. Les positions des ARS supplémentaires sont indiquées. **[B]** Première analyse de l'intégrité du fragment par PCR multiplex avec 5 couples d'amorce distribués de façon homogène le long du fragment cible, permettant de générer simultanément 5 amplicons de 367 pb (PCR 1), 516 pb (PCR 2), 664 pb (PCR 3), 883 pb (PCR 4) et 1020 pb (PCR 5). « M » : DNA Ladder 100bp NEB (100-1517pb). « H2O » : contrôle négatif sans ADN. « T- » : témoin négatif, ADNg de *S. cerevisiae* VL6-48N. « T+(1) » : témoin positif, ADNg de la souche AG686Δ*sbo-alb*::ACH-spec. « T+(2) » : témoin positif, ADNg de la souche AG694. **[C]** L'intégrité et la taille du fragment sont finalement validés par restriction enzymatique et analyse en PFGE. La digestion du fragment par l'enzyme FseI produit 3 fragments de 246 kb, 204 kb et 85 kb. « M » : Mid range PFG ladder NEB (15.0 - 242.5 kbp). « T- » : témoin négatif, ADNg de *S. cerevisiae* VL6-48N. ACH : ARSH4-CEN6-HIS3

- Des protoplastes de la souche AG694 contenant deux ARS et des protoplastes de la souche AG686 $\Delta$ *sbo-alb*::ACH-spec contenant une seule ARS.
- Des sphéroplastes de levure VL6-48N exprimant le système CRISPR-Cas9 ciblant spécifiquement les gènes *ywbE* et *yufO* de part et d'autre de la région de 536 kb à cloner (figure 70A).
- Un pansement de recombinaison de 90 pb correspondant aux deux bornes du fragment à capturer.

Une fois prêts, les protoplastes (AG694 ou AG686 $\Delta$ *sbo-alb*::ACH-spec), les sphéroplastes VL6-48N et le pansement ont été mélangés et du PEG a été ajouté. Après 20 minutes de contact et régénération en milieu SOS, les transformants ont été étalés sur milieu solide SD-HIS et incubés pendant 48 h à 30°C. Pour la souche AG694, 344 colonies sont apparues sur les boîtes contre 239 pour la souche AG686 $\Delta$ *sbo-alb*::ACH-spec. Pour chacune de ces deux conditions, 56 colonies ont été repiquées en milieu liquide SD-HIS. Après 48 h à 30°C sous agitation, nous avons observé une croissance dans 49 tubes pour la souche AG694 et 48 pour la souche AG686 $\Delta$ *sbo-alb*::ACH-spec. L'ADNg de ces clones a été extrait puis analysé par PCR, d'abord à l'aide du couple d'amorces Verif\_Fgt5-jonction\_F et Verif\_jonction frgt 3-5 situés au niveau de la jonction de circularisation du fragment (taille attendue 644 pb). L'intégrité du fragment a ensuite été évaluée par PCR multiplex à l'aide de 5 couples d'amorces distribués de façon homogène le long du fragment cible, et permettant de générer simultanément cinq amplicons de 367 pb (PCR 1), 516 pb (PCR 2), 664 pb (PCR 3), 883 pb (PCR 4) et 1020 pb (PCR 5) (figure 70B). Par souci de clarté, seulement 20 clones sont présentés figure 70B. Au bilan, 8 clones pour la souche AG694 (cl 5.4, 5.8, 5.14, 5.23, 6.2, 6.6, 6.10 et 6.16) et 2 clones pour la souche AG686 $\Delta$ *sbo-alb*::ACH-spec (cl 3.9 et 3.10) ont montré les profils attendus pour toutes les PCR.

Suite à ce criblage PCR, la taille du fragment a été évaluée par migration par PFGE après restriction enzymatique avec l'enzyme FseI. La présence de 3 sites de restriction sur le fragment capturé devait aboutir à l'hydrolyse de la molécule en 3 fragments linéaires de 246 kb, 204 kb et 85 kb. Les 10 clones analysés ont tous présenté le profil à 3 bandes attendu (figure 70C). Cependant, 3 d'entre eux (cl 3.10, 5.23 et 6.2) ont aussi montré des bandes supplémentaires à des tailles non attendues (figure 70C).



**Figure 71 : Capture par CREasPy-Fusion, d'un fragment du génome *B. subtilis* d'une taille de 925 kb.**

**[A]** Isolement du fragment par deux duplex ARNg-Cas9. Re-circularisation du fragment par le système de recombinaison homologue de la levure. Les positions des ARS supplémentaires sont indiquées. **[B]** Première analyse de l'intégrité du fragment par PCR multiplex avec 8 couples d'amorces produisant simultanément 8 amplicons de 269 pb (PCR 1), 367 pb (PCR 2), 455 pb (PCR 3), 516 pb (PCR 4), 664 pb (PCR 5), 770 pb (PCR 6), 883 pb (PCR 7) et 1020 pb (PCR 8). **[C]** PCR simplex avec le couple d'amorce Bsu\_Frgt2.2\_R et Bsu\_Frgt2.2\_F (PCR 3). « M » : DNA Ladder 100bp NEB (100-1517pb). « H2O » : contrôle négatif sans ADN. « T+(1) » : témoin positif, ADNg de la souche ED 2039. « T+(2) » : témoin positif, ADNg de la souche AG686 $\Delta$ sbo-alb::ACH-spec. « T- » : témoin négatif, ADNg de *S. cerevisiae* VL6-48N. **[D]** L'intégrité et la taille du fragment sont contrôlés par restriction enzymatique et analyse en PFGE. La digestion du fragment par l'enzyme SgrDI produit 4 fragments de 73 kb, 173 kb, 141 kb et 536 kb. « M 1 » *S. cerevisiae* chromosomal DNA ladder Bio-Rad (225-2200 kb). « M 2 » : Mid range PFG ladder NEB (15.0 - 242.5 kbp). « T- » : témoin négatif, ADNg de *S. cerevisiae* VL6-48N. ACH : ARSH4-CEN6-HIS3.

Ces résultats ont clairement montré qu'il était possible de capturer un fragment d'environ 536 kb possédant une ou 2 ARS supplémentaires. La capture semble plus efficace lorsque deux ARS sont présentes sur le fragment, puisque 6 clones sur les 49 analysés ont été validés pour la capture du fragment issu de la souche AG694 possédant 2 ARS contre 1 clone sur 48 pour la souche AG686 $\Delta$ *sbo-alb*::ACH-spec possédant une seule ARS. Cette première expérience nous a confortés dans l'idée que l'ajout d'ARS tout le long du génome était nécessaire pour son maintien dans la levure. Cependant, elle sera reproduite afin de démontrer que la présence de 2 ARS augmente bien l'efficacité de capture et de stabilité de la molécule dans la levure.

#### 4. Essais de capture d'un fragment de génome de *B. subtilis* de 925 kb par CREasPy-Fusion

##### 4.1. Premier essai de capture du fragment de 925 kb

Après avoir réussi à capturer un fragment de 536 kb possédant 2 ARS (ARSH1 et ARSH4), nous avons décidé de tenter la capture d'un fragment de 925 kb correspondant à la région entre les gènes *IrgB* et *obg* (figure 71A).

Pour cette tentative, nos collaborateurs ont construit la souche ED2039 contenant un total de 5 ARS dont 4 dans la région d'intérêt, correspondant à l'ARSH4 avec les éléments levure, l'ARSH1 (construction 12), l'ARS416 (construction 11) et l'ARS1323 (construction 10) (figure 71A, tableau 21). De notre côté, nous avons préparé la souche levure VL6-48N exprimant le système CRISPR-Cas9 ciblant spécifiquement les gènes *ywbE* et *obg* de part et d'autre de la région de 925 kb à cloner et nous avons produit un nouveau pansement de recombinaison de 90 pb correspondant aux deux bornes de ce fragment.

La capture a été réalisée en mettant des protoplastes de la souche de *B. subtilis* ED2039 (4 ARS dans la région cible) ou bien des protoplastes de la souche AG686 $\Delta$ *sbo-alb*::ACH-spec (une seule ARS) en contact avec des sphéroplastes de la souche de levure pré-chargée avec le système CRISPR-Cas9 spécifique au fragment de 925 kb et du pansement. A la fin de l'expérience, les suspensions cellulaires ont été étalées sur milieu solide SD-HIS.

Après 48h d'incubation à 30°C, plusieurs centaines de colonies sont apparues sur boîte de sélection que ce soit pour les conditions avec la souche ED2039 ou celles avec la souche AG686 $\Delta$ *sbo-alb*::ACH-spec. Un total de 120 clones a été repiqué en milieu liquide (60 pour

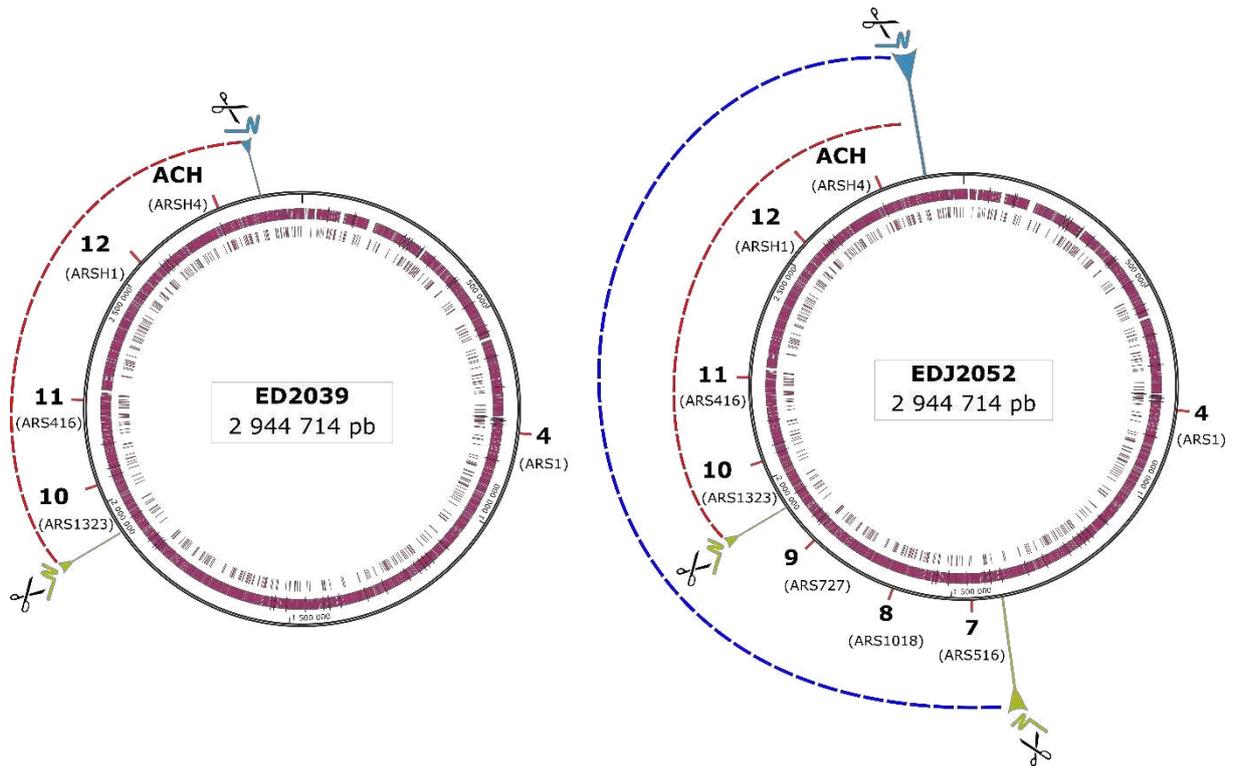


chaque condition). Quarante-six clones issus des conditions avec la souche ED2039 se sont multipliés en milieu sélectif contre 40 pour celles issues de la souche AG686 $\Delta$ *sbo-*alb*::ACH-spec*. Comme précédemment, l'analyse des clones a été faite par PCR simplex, PCR multiplex puis PFGE.

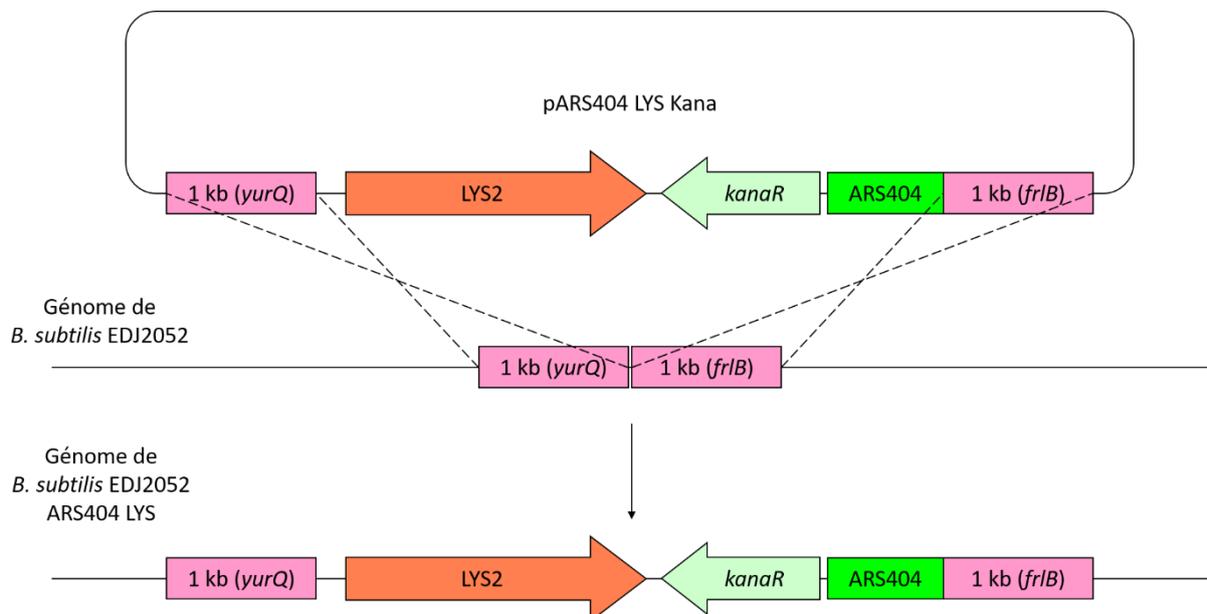
La circularisation du fragment a été confirmée grâce au couple d'amorces Verif\_Fgt5-jonction\_F et Verif\_jonction frgt 2-5\_R produisant un amplicon de 1077 pb seulement si le fragment est présent dans la levure sous forme circulaire. Six clones ont montré la présence de cet amplicon, 5 issus des conditions de capture du fragment avec 4 ARS et 1 issu des conditions de capture du fragment avec 1 ARS. L'intégrité du fragment a ensuite été évaluée par PCR multiplex à l'aide de 8 couples d'amorces produisant simultanément 8 amplicons spécifiques de 269 pb (PCR 1), 367 pb (PCR 2), 455 pb (PCR 3), 516 pb (PCR 4), 664 pb (PCR 5), 770 pb (PCR 6), 883 pb (PCR 7) et 1020 pb (PCR 8). Le clone issu de la capture du fragment avec une seule ARS n'a montré que 3 bandes sur les 8 bandes attendues. Tous les autres ont montré un profil à 7 bandes, puisque la bande correspondant à la PCR 3 (455 pb) (couple d'amorces Bsu\_Frgt2.2\_R et Bsu\_Frgt2.2\_F) était très peu visible sur gel (figure 71B). Une PCR simplex a donc été mise en place avec ce couple d'amorces spécifiquement. Les 6 clones analysés ont tous montré la bande attendue à 455 pb indiquant (i) que le locus était possiblement présent dans les clones sélectionnés et (ii) que vraisemblablement le couple d'amorces (Bsu\_Frgt2.2\_R et Bsu\_Frgt2.2\_F) ne fonctionnait pas bien dans le cadre d'une PCR multiplex dont les conditions sont parfois difficiles à optimiser. Pour le clone issu de la capture du fragment avec une seule ARS (cl 3.10), la bande était également présente mais avec une intensité plus faible que pour les autres clones (figure 71C).

Les 5 clones issus de la capture du fragment avec 4 ARS montrant les amplicons attendus pour l'ensemble des couples d'amorces en PCR simplex/multiplex ont été sélectionnés puis analysés en PFGE après hydrolyse de leurs génomes par SgrDI (profil théorique attendu : 4 bandes de 73 kb, 173 kb, 141 kb et 536 kb). Aucun des 5 clones analysés n'a montré le profil attendu à 4 bandes (figure 71D), malgré deux répétitions du PFGE.

Ces derniers résultats indiquaient que le fragment de 925 kb n'avait pas été capturé dans la levure, et suggéraient que la proposition initiale (addition de 4 ARS dans cette région de 925 kb) n'était pas optimale et devait être revue. Malgré cette déconvenue, les résultats obtenus en PCR multiplex, avec des profils quasi-complets pour les fragments « 4 ARS » contre un profil très incomplet pour le fragment « 1 ARS » étaient encourageants et semblaient indiquer que



**Figure 72 : Comparaison des cartes théoriques des génomes des souches *B. subtilis* ED2039 (5 ARS) et EDJ2052 (8 ARS) basées sur le design n°1. Les ARS supplémentaires sont indiqués par les numéros qui correspondent aux constructions répertoriées dans le tableau 21. ACH : ARSH4-CEN6-HIS3. La zone cible de la capture du fragment de 925 kb est représentée par les pointillés rouges, celle de la capture du demi génome est représentée par les pointillés bleus.**



**Figure 73 : Schéma de l'insertion de l'ARS404 LYS Kana dans le génome de la souche *B. subtilis* EDJ2052 à partir du plasmide pARS404 LYS Kana. L'insertion de la cassette ARS404 kana seule se fait de la même manière. La différence réside dans l'absence du marqueur d'auxotrophie LYS2.**

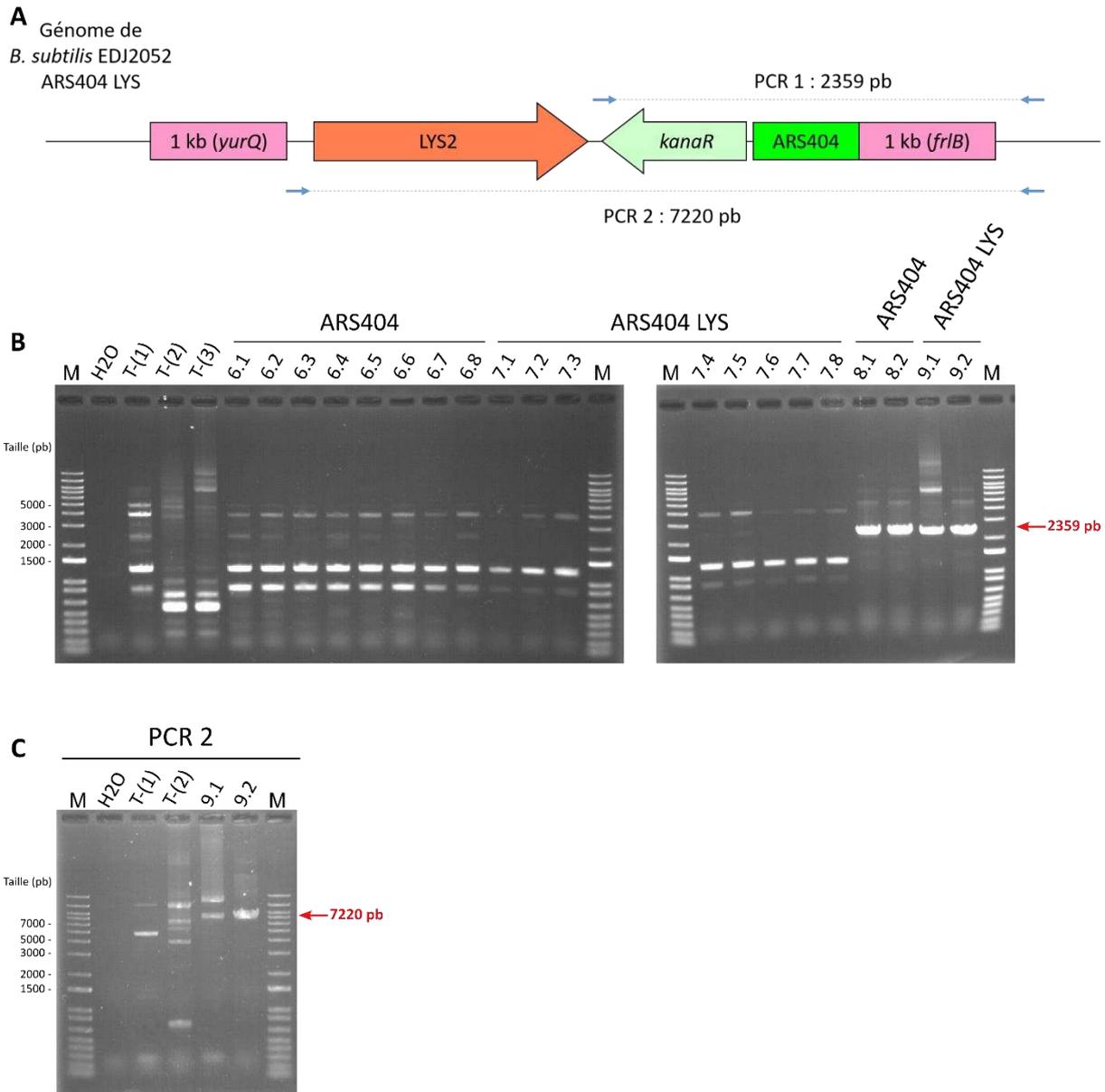
nous étions, malgré tout, sur la bonne voie. Ces résultats nous ont montré que le fragment avec une seule ARS supplémentaire, issu de la souche AG686 $\Delta$ *sbo-alb::ACH-spec*, était loin d'être complet. Nous en avons conclu que l'ajout d'une seule ARS sur un fragment de 925 kb n'était pas suffisant pour le stabiliser.

#### 4.2. Ajout de l'ARS 404 avec ou sans le marqueur d'auxotrophie pour la lysine

Dans la section précédente, nous n'avons pas réussi à cloner un fragment de 925 kb contenant 4 ARS supplémentaires. Les premières expériences de capture des fragments de 536 kb et de 925 kb semblaient pourtant indiquer que l'ajout d'ARS pouvait être efficace pour augmenter la stabilité du fragment dans la levure. Nous avons donc décidé d'ajouter une cinquième ARS sur le fragment de 925 kb. Pour choisir où insérer cette nouvelle ARS, nous avons scrupuleusement analysé ce fragment au niveau de sa séquence et avons noté que nous avons laissé une distance d'environ 338 kb entre les ARSH1 et ARS416, ce qui pouvait constituer un point de faiblesse. Il est en effet indiqué dans la littérature que les ARS prédites comme fortes sont capables de stabiliser des séquences d'ADN de 100 à 300 kb dans la levure (Dershowitz and Newlon, 1993). Nous avons donc décidé de placer une nouvelle ARS à équidistance de ces deux dernières, plus précisément entre les gènes non essentiels *frlB* et *yurQ*. Pour ce travail, nous avons sélectionné l'ARS404 d'une taille de 53 pb identifiée comme forte par Liachko *et al.* (Liachko *et al.*, 2013).

Des données de la littérature indiquaient aussi que la présence de plusieurs marqueurs d'auxotrophie pouvait être bénéfique pour le clonage de grandes molécules d'ADN dans la levure (Venetz *et al.*, 2019; Zhou *et al.*, 2016). Nous avons donc fait le choix d'ajouter soit l'ARS404 seule, soit associée avec le marqueur LYS2, les marqueurs HIS3, URA3, TRP1 étant déjà utilisés respectivement dans la cassette des éléments levure (ARSH4-CEN6-HIS3), dans le plasmide d'expression des ARNg et dans le plasmide d'expression de la Cas9.

Afin de pouvoir tester ce nouveau design le plus rapidement possible, notre équipe a proposé d'ajouter elle-même ces éléments dans la souche ED2039 (4 ARS) sans passer par la procédure décrite partie 2.3 (figure 67) qui permet, entre autres, de recycler le marqueur de résistance et qui est donc longue à mettre en place. Cette proposition a été validée au moment où nos collaborateurs venaient de terminer la construction d'une nouvelle souche, la souche EDJ2052 contenant 8 ARS (figure 72). Après plusieurs discussions, le consortium a finalement



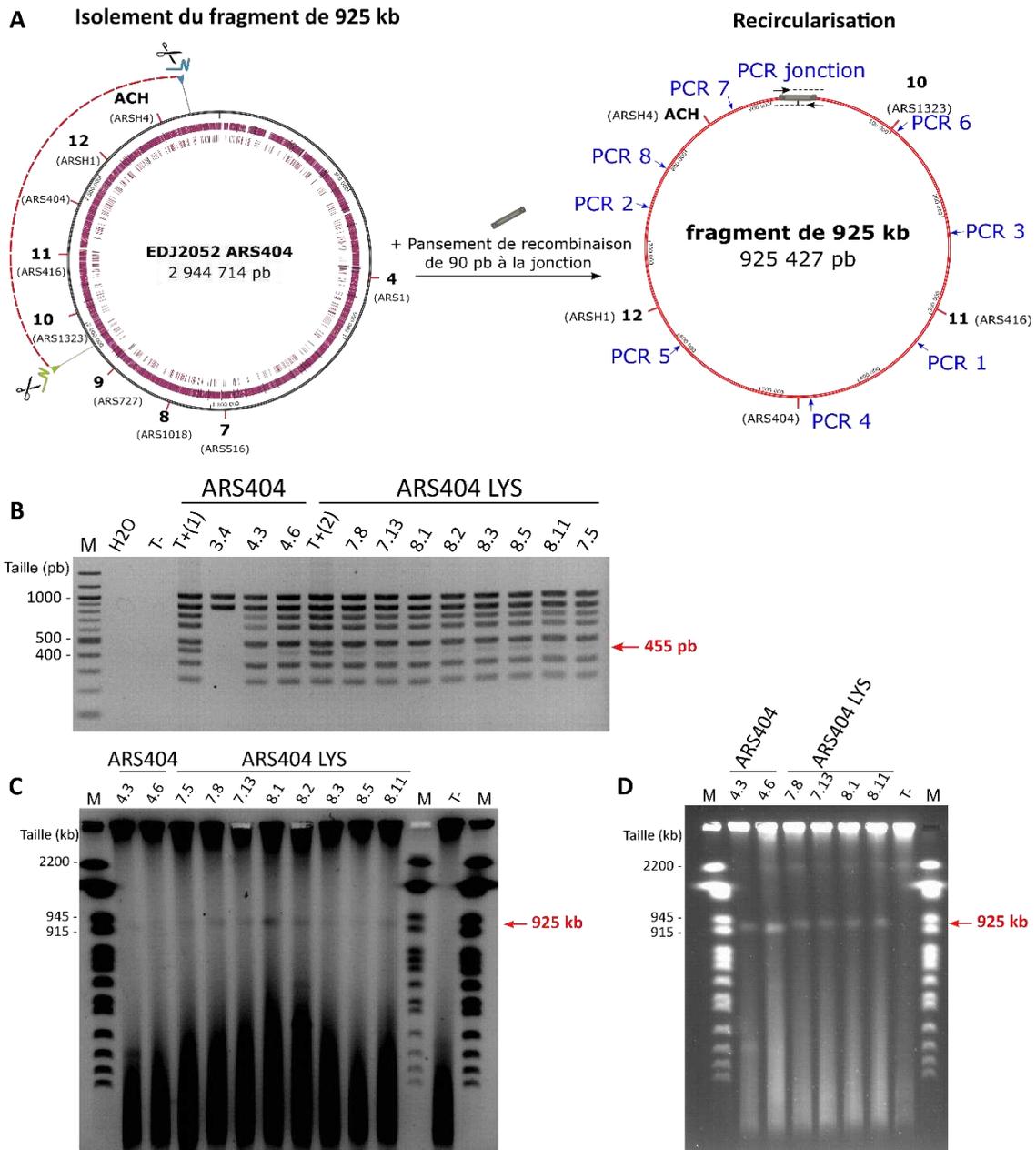
**Figure 74 : Construction des souches EDJ2052 ARS404 et EDJ2052 ARS404 LYS.** **[A]** Schéma de la cassette ARS404 LYS Kana insérée dans le génome de *B. subtilis* EDJ2052. L'insertion de la cassette ARS404 Kana est identique, la seule différence est l'absence du marqueur LYS2. **[B]** Analyse des transformants par PCR avec les amorces screening ARS404 F et Kana/YurQ (PCR 1). L'intégration des cassettes ARS404 LYS Kana et ARS404 Kana est validée par cette PCR produisant une bande à 2359 pb. **[C]** Chez les transformants issus de la transformation du pARS404 LYS Kana, contrôle de la présence de LYS2 par PCR avec les amorces screening ARS404 F et LYS/YurQ (PCR 2) qui produit un amplicon de 7220 pb. « M » : DNA Ladder 1 kb+ Invitrogen (100-12000 pb). « H2O » : contrôle négatif, sans ADN. « T-(1) » témoin négatif, ADNg de la souche EDJ2059. « T-(2) » : témoin négatif, PCR 1 : plasmide pARS404 Kana ; PCR 2 : plasmide pARS404 LYS Kana « T-(3) » : témoin négatif, plasmide pARS404 LYS Kana.

fait le choix d'ajouter l'ARS404 +/- LYS2 au génome de la souche *B. subtilis* EDJ2052 par transformation classique (protocole MC). Deux raisons ont motivé ce choix : (i) le fragment de 925 kb possédant 4 ARS était identique entre la souche ED2039 utilisée lors du premier essai de capture et la souche EDJ2052 (figure 72), ajouter la cassette ARS404 (+/-LYS2) à l'une ou l'autre souche pour capturer le fragment de 925 kb ne changeait donc rien a priori et (ii) l'insertion de ARS404 (+/-LYS2) dans le génome de la souche EDJ2052 permettait de tester non seulement la capture du fragment de 925 kb, mais également la capture d'un demi-génome (~1.466 Mb) (figure 72) dans une même expérience.

Pour obtenir les souches EDJ2052 ARS404 (+/-LYS2), les plasmides pARS404-Kana et pARS404-LYS-Kana ont été construits. Ces plasmides sont des plasmides intégratifs dans lesquels le gène de résistance à la kanamycine (*kanaR*) et l'ARS404 (+/-LYS2) sont bordés de deux cassettes de recombinaison (~1 kb) correspondant en amont au gène *frlB* et en aval au gène *yurQ* (figure 73). Après validation, ces plasmides ont été transformés dans la souche EDJ2052 en utilisant le protocole simple MC. Cette souche étant résistante à la spectinomycine, les transformants ont été sélectionnés soit sur milieu LB solide contenant de la kanamycine (7,5 µg/ml) et de la spectinomycine (100 µg/ml), soit sur milieu LB solide contenant seulement de la kanamycine (7,5 µg/ml).

Après une nuit d'incubation à 37°C, 32 colonies ont été obtenues pour les conditions de transformation du pARS404-kana et 24 pour le pARS404-LYS-kana. Dix clones de chaque ont été analysés au niveau génotypique (figure 74A). Pour le plasmide pARS404-kana, la PCR 1 a permis de confirmer l'insertion de l'ARS404 et du gène *kanaR* au locus désiré pour 2 clones sur 10 (cl 8.1 et 8.2) (bande à 2359 pb, figure 74B). Pour le plasmide pARS404-LYS-kana, la PCR 1 a également confirmé l'intégration de la cassette pour 2 clones sur 10 (cl 9.1 et 9.2). Le reste des clones montraient une bande à 1273 pb correspondant à la taille du locus non modifié. Pour les 2 clones issus de la transformation du pARS404-LYS-Kana, la présence du marqueur LYS2 au locus a été contrôlée avec la PCR 2 qui produit un amplicon de 7220 pb (figure 74A). Un seul des 2 clones (cl 9.2) a montré le profil attendu, avec une bande unique à 7220 pb (figure 74C). Pour terminer cette analyse, nous avons contrôlé l'absence de mutation au niveau de l'ARS404 et du marqueur LYS2 par séquençage Sanger. Un clone a pu être validé pour chaque condition (cl 8.1 et cl 9.2).

Suite à cette expérience, les deux souches EDJ2052 ARS404 et EDJ2052 ARS404 LYS contenant respectivement l'ARS404 seule ou l'ARS404 avec le marqueur LYS2 ont été



**Figure 75 : Capture par CReasPy-Fusion, d'un fragment du génome *B. subtilis* d'une taille de 925 kb à partir des souche EDJ2052 ARS404 et EDJ2052 ARS404 LYS. [A] Isolement du fragment par deux duplex ARNg-Cas9. Re-circularisation du fragment par le système de recombinaison homologue de la levure. Seule la souche EDJ2052 ARS404 est présentée mais le principe est le même pour la souche EDJ2052 ARS404 LYS. [B] Première analyse de l'intégrité du fragment par PCR multiplex avec 8 couples d'amorces produisant simultanément 8 amplicons de 269 pb (PCR 1), 367 pb (PCR 2), 455 pb (PCR 3), 516 pb (PCR 4), 664 pb (PCR 5), 770 pb (PCR 6), 883 pb (PCR 7) et 1020 pb (PCR 8) « M » : DNA Ladder 100bp NEB (100-1517pb). « H2O » : contrôle négatif sans ADN. « T- » : témoin négatif, ADNg de *S. cerevisiae* VL6-48N. « T+(1) » : témoin positif, ADNg de la souche EDJ2052 ARS404. « T+(2) » : témoin positif, ADNg de la souche EDJ2052 ARS404 LYS. [C] L'intégrité et la taille du fragment sont validés par restriction enzymatique et analyse en PFGE. La digestion du fragment par l'enzyme *AbsI* doit en théorie le linéariser en une molécule de 925 kb. « M » *S. cerevisiae* chromosomal DNA ladder Bio-Rad (225-2200 kb). « T- » : témoin négatif, ADNg de *S. cerevisiae* VL6-48N. ACH : ARSH4-CEN6-HIS3.**

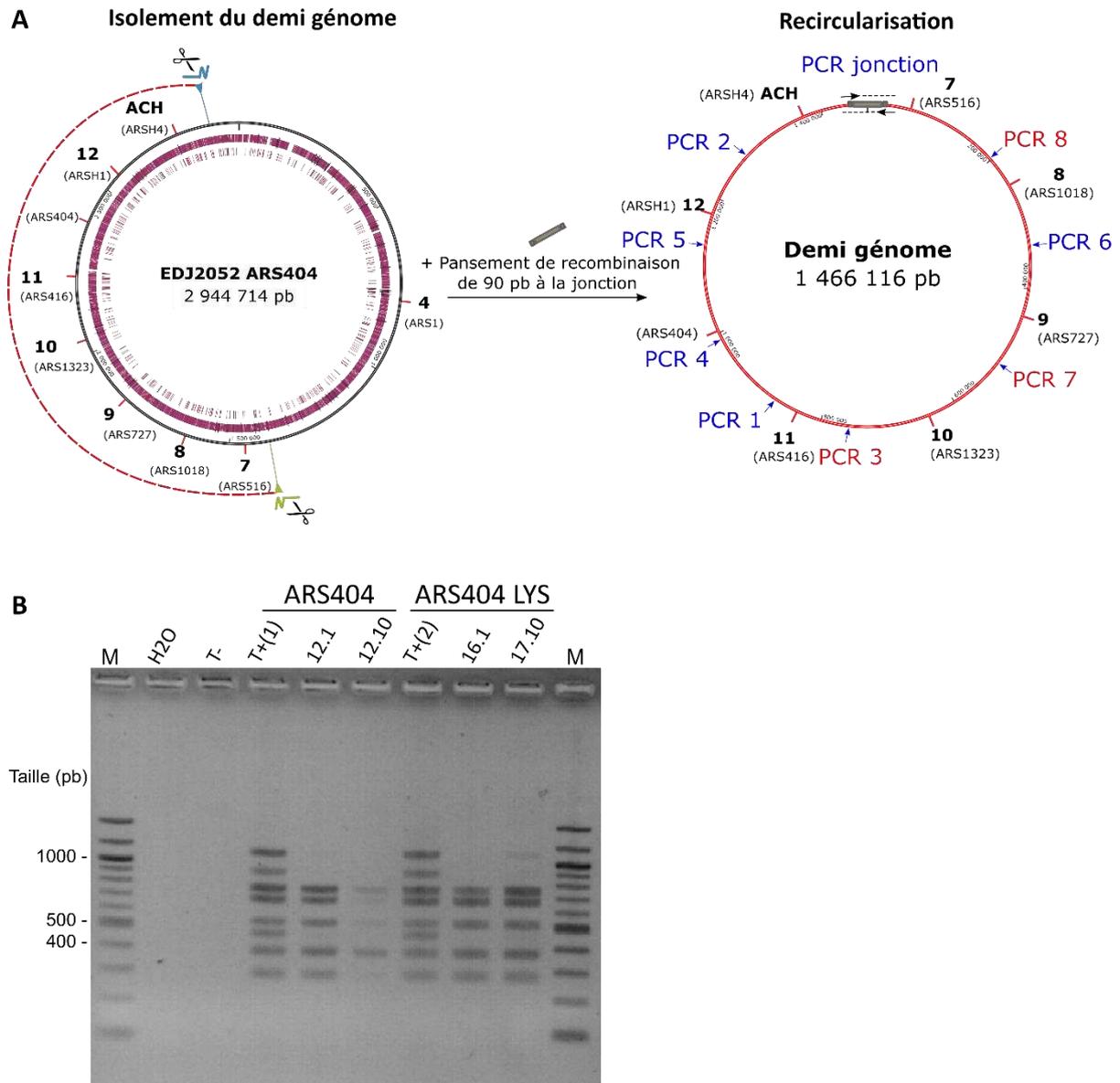
obtenues. Au niveau du fragment de 925 kb, ces deux souches contenaient 5 ARS, la souche EDJ2052 ARS404 LYS possédant en plus un deuxième marqueur d'auxotrophie LYS2. Au niveau du demi-génome de 1,46 Mb, ces deux souches contenaient 8 ARS, la souche EDJ2052 ARS404 LYS possédant en plus le deuxième marqueur d'auxotrophie LYS2.

### 4.3. Deuxième essai de capture du fragment de 925 kb par CReasPy-Fusion

Pour ce deuxième essai de capture du fragment de 925 kb, nous avons utilisé des protoplastes des souches de *B. subtilis* EDJ2052 ARS404 (5 ARS) et EDJ2052 ARS404 LYS (5 ARS et marqueur LYS2), des sphéroplastes de la levure VL6-48N pré-chargée avec le système CRISPR-Cas9 spécifique au fragment de 925 kb et le pansement de réparation adéquat. Les transformants ont été sélectionnés sur milieu solide SD-HIS pour les conditions avec la souche EDJ2052 ARS404 et sur SD-HIS-LYS pour celles avec la souche EDJ2052 ARS404 LYS.

Après 5 jours d'incubation à 30°C, une colonie est apparue sur le contrôle négatif « levure seule » (sans pansement et sans avoir été mis en contact avec *B. subtilis*). Pour les conditions utilisant la souche EDJ2052 ARS404, 198 colonies ont été obtenues. Pour celles avec la souche EDJ2052 ARS404 LYS, 87 transformants ont été observés sur milieu gélosé. Soixante clones au total (30 pour chaque condition) ont été analysés par PCR simplex, PCR multiplex, puis PFGE. Comme lors de la première tentative, la première PCR simplex a permis de contrôler la circularisation du fragment avec un amplicon à la taille de 1077 pb possible uniquement si le fragment s'est circularisé. Sur les 60 clones analysés, 10 ont montré un profil à une bande à la taille attendue. Deux étaient issus des conditions de capture du fragment sans le marqueur LYS2, et les 8 autres provenaient de la capture du fragment avec le marqueur LYS2. L'intégrité du fragment a ensuite été évaluée par PCR multiplex en utilisant le jeu d'amorces décrit précédemment (cf partie 4.1 du chapitre 2) (figure 75A). Cette fois, tous les clones ont présenté le profil à 8 bandes attendu, avec la bande à 455 pb correspondant à la PCR 3 visible mais cependant faible par rapport aux témoins positifs (figure 75B). Le clone 3.4 (négatif en PCR simplex) mais conservé pour l'analyse en PCR multiplex, présentait quant à lui un profil à 2 bandes (figure 75B).

Pour terminer cette analyse, la taille des fragments clonés a été contrôlée par restriction enzymatique suivi d'un PFGE en utilisant l'enzyme de restriction AbsI qui permet de linéariser le fragment en une molécule unique de 925 kb. Après migration et révélation, le gel n'était



**Figure 76 : Capture par CReasPy-Fusion, d'un demi génome de *B. subtilis* d'une taille de 1.466 Mb à partir des souche EDJ2052 ARS404 et EDJ2052 ARS404 LYS. [A]** Isolement du demi génome par deux duplex ARNg-Cas9. Re-circularisation du demi génome par le système de recombinaison homologue de la levure. Seule la souche EDJ2052 ARS404 est présentée mais le principe est le même pour la souche EDJ2052 ARS404 LYS. **[B]** Analyse de l'intégrité du fragment par PCR multiplex avec 8 couples d'amorces, permettant de générer simultanément 8 amplicons de 269 pb (PCR 1), 367 pb (PCR 2), 455 pb (PCR 3), 516 pb (PCR 4), 664 pb (PCR 5), 736 pb (PCR 6), 888 pb (PCR 7) et 1084 pb (PCR 8). « M » : DNA Ladder 100bp NEB (100-1517pb). « H2O » : contrôle négatif, sans ADN. « T- » : témoin négatif, ADNg de *S. cerevisiae* VL6-48N. « T+(1) » : témoin positif, ADNg de la souche EDJ2052 ARS404. « T+(2) » : témoin positif, ADNg de la souche EDJ2052 ARS404 LYS.

pas très net, mais nous avons tout de même observé une bande (faible) à la taille attendue pour tous les clones analysés alors qu'aucune bande n'était visible sur le contrôle négatif correspondant aux levures VL6-48N non transformées (piste T-, figure 75C). Pour confirmer ce résultat, l'hydrolyse par *AbsI* et le PFGE ont été refaits sur les clones les plus prometteurs (cl 4.3, 4.6, 7.8, 7.13, 8.1 et 8.11). La bande à 925 kb était cette fois bien visible, indiquant que le fragment d'intérêt était bien présent dans les clones levure analysés (figure 75D).

Ces résultats montrent pour la première fois la capture dans la levure d'un fragment de génome de *B. subtilis* proche d'1 Mb. Pour y arriver, il a fallu l'agrémenter non pas de 4 ARS mais de 5 ARS. En effet, nos résultats semblent bien indiquer que c'est l'ajout de l'ARS404 sur le fragment de 925 kb (en plus des 4 autres) qui a permis de cloner ce fragment dans la levure. Par ailleurs, nos résultats semblent également montrer que la présence du marqueur LYS2 augmente l'efficacité de capture. En effet, 8 clones issus de la capture du fragment avec le marqueur LYS2 (sur 30 analysés) semblent corrects contre 2 clones sur 30 analysés pour la capture du fragment sans le marqueur LYS2

## 5. Capture d'un demi-génome de *B. subtilis*

Parallèlement à la capture du fragment de 925 kb modifié, nous avons tenté la capture de la région comprise entre les gènes *IrgB* et *parE*, d'une taille de 1.466 Mb correspondant à un demi-génome (figure 76A). Pour ce faire, nous avons pré-chargé la souche levure VL6-48N avec le système CRISPR-Cas9 ciblant spécifiquement les gènes *ywbE* et *plsY* de part et d'autre de la région à cloner et avons produit le nouveau pansement de recombinaison de 90 pb correspondant aux deux bornes du demi-génome ciblé. Ces éléments ont été mis au contact des protoplastes des souches de *B. subtilis* EDJ2052 ARS404 et EDJ2052 ARS404 LYS et de PEG avant d'être étalés sur milieu sélectif SD-HIS et SD-HIS-LYS respectivement.

Après 48h à 30°C, 150 colonies sont apparues sur boîte pour la souche EDJ2052 ARS404 contre 148 colonies pour la souche EDJ2052 ARS404 LYS. Soixante clones (30 pour chaque condition) ont été analysés par PCR simplex au niveau de la jonction de re-circularisation en utilisant le couple d'amorces *Verif\_demi-genome-jonction\_R2* et *Verif\_Fgt5-jonction\_F* qui permettait de générer un amplicon de 945 pb si le fragment était correctement circularisé dans la levure. Quatre clones (2 par condition) ont montré l'amplicon attendu. Ces 4 clones



ont ensuite été analysés par PCR multiplex à l'aide de 8 couples d'amorces distribués de façon homogène le long du demi-génome cible, et permettant de générer 8 amplicons spécifiques de 269 pb (PCR 1), 367 pb (PCR 2), 455 pb (PCR 3), 516 pb (PCR 4), 664 pb (PCR 5), 736 pb (PCR 6), 888 pb (PCR 7) et 1084 pb (PCR 8) si le demi-génome est présent et intègre (figure 76A). Les clones analysés n'ont pas présenté le profil à 8 bandes attendu mais un profil à 5 bandes identique entre tous les clones (figure 76B). De façon intéressante, les PCR positives (*i.e.* les PCR 1, PCR 2, PCR 4, PCR 5 et la PCR jonction) se situent au niveau de *loci* contigus sur le fragment circularisé et les PCR négatives (= bandes absentes (PCR 7) ou de très faibles intensités (PCR 8 pour le clone 17.10) et PCR 3) se concentrent vers une autre partie du fragment (figure 76A et B). La seule PCR qui enfreint cette règle est la PCR 6 (736 pb). Ces résultats semblent indiquer que les régions adjacentes aux PCR 7, PCR 8 et possiblement PCR 3 ne sont pas stables en l'état dans la levure (figure 76A et B).

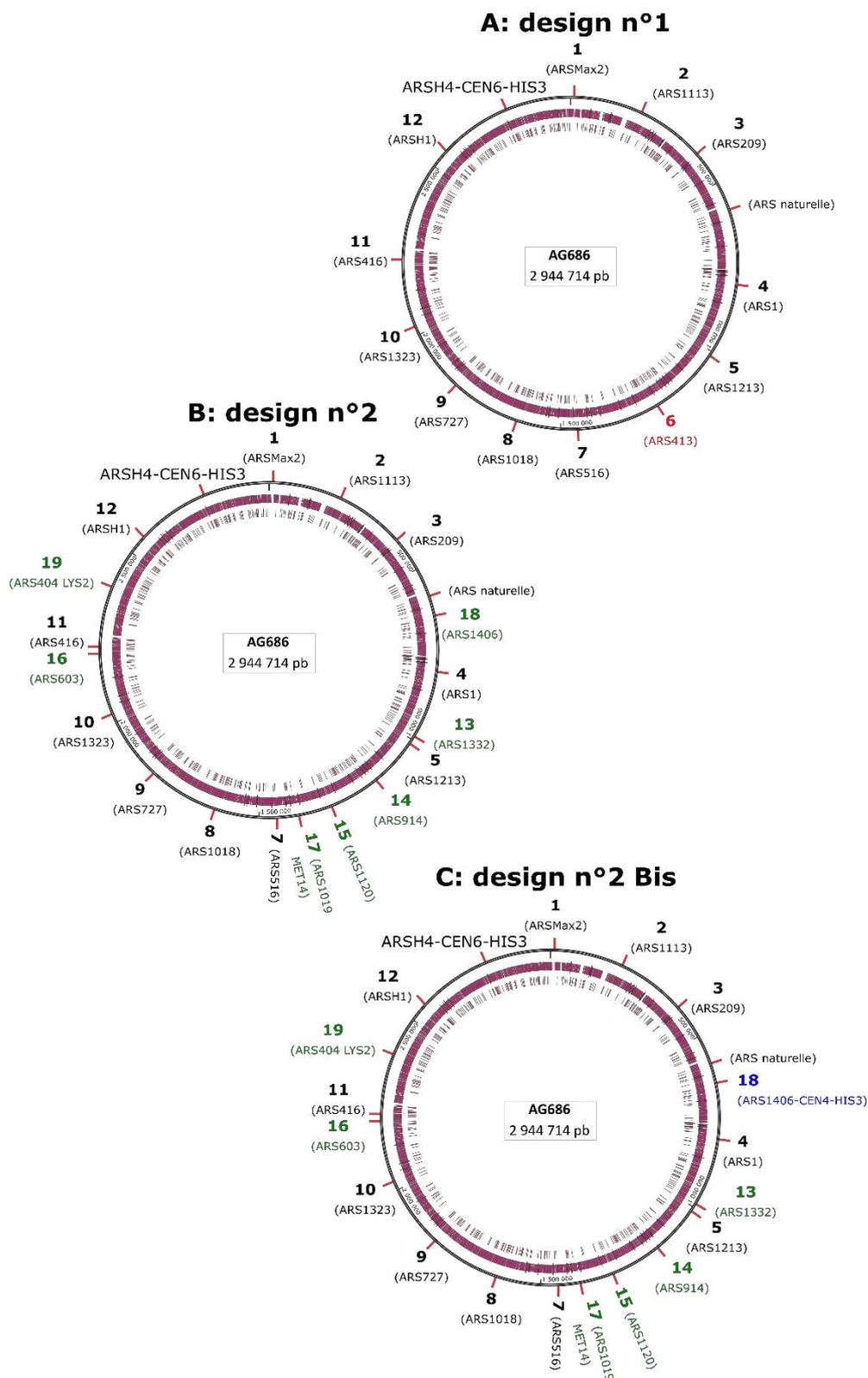
Ces résultats indiquaient que le demi-génome de *B. subtilis* EDJ2052 ARS404 LYS ne pouvait pas être cloné en l'état dans la levure VL6-48N et qu'il était certainement nécessaire de le modifier davantage pour y arriver.

## 6. Conception sur ordinateur d'un génome modifié de *B. subtilis* contenant 19 ARS et 3 marqueurs d'auxotrophie

Une série de quatre CREasPy-Fusion a été menée dans le cadre de ce chapitre dans le but de cloner dans la levure des fragments du génome de *B. subtilis* de plus en plus grands. Après avoir analysé les résultats de ces premières expériences et les avoir confrontés les uns aux autres (= phase « L » du cycle DBTL, figure 65), le consortium a décidé de faire une pause et revoir la proposition initiale (13 ARS, HIS3 et CEN6) avant de poursuivre les expériences de capture (phase « D » du cycle DBTL, figure 65).

Un ensemble de modifications a été proposé et validé par l'ensemble des participants. Ces modifications sont exposées ci-dessous sans ordre particulier :

- 1) Lors des travaux de capture du fragment de 925 kb, nous avons montré que l'ajout de l'ARS404 et du marqueur LYS2 semblait augmenter la stabilité du fragment dans la levure. Ces conclusions confortées par des données de la littérature (Venetz *et al.*, 2019; Zhou *et al.*, 2016) ont été prises en compte et l'insertion de l'ARS404 avec le



**Figure 77 : Comparaison des cartes théoriques des génomes des souches de *B. subtilis* à construire selon les différents design.** Les ARS supplémentaires sont indiqués par les numéros qui correspondent aux constructions répertoriées dans le tableau 22. La position des éléments levure ARSH4-CEN6-HIS3 décrit chapitre 1 (cf partie 3.1 du chapitre 1) ainsi que celle de l'ARS naturelle sont aussi indiquées. **[A]** Design n°1. La construction 6 (ARS413), éliminée dans le design n°2, est représentée en rouge. **[B]** Design n°2. Les constructions ajoutées sont représentées en vert. **[C]** Design n°2 Bis. La construction 18, différente du design n°2, est représentée en bleu.

marqueur LYS2 entre les gènes *yurQ* et *frlB* (figure 77) a donc été validée pour ce second design.

- 2) Suite à cette décision, nous avons observé la carte du génome et avons noté que les marqueurs HIS3 et LYS2 étaient proches, laissant une grande partie du génome sans marqueur d'auxotrophie. Nous avons donc décidé d'ajouter le marqueur d'auxotrophie MET14 à la construction 17 (*i.e.* associé à l'ARS1019) dans le but de stabiliser cette section du génome qui était éloignée des marqueurs HIS3 et LYS2.
- 3) Parallèlement à nos expériences de CReasPy-Fusion, nos collaborateurs avaient poursuivi la construction de la souche initiale contenant 13 ARS (premier design). Au cours de ces travaux, ils se sont rendus compte que la construction 6 (=ARS413), (figure 77, tableau 22) placée entre les gènes *ylqB* et *ylxM* ne pouvait pas être intégrée à ce locus en raison d'un problème technique. Cette ARS a donc été éliminée du second design et remplacée par les constructions 14, 15 et 17 (ARS914, ARS1120, ARS1019) afin de stabiliser cette zone (figure 77, tableau 22).
- 4) La capacité de l'ARS naturelle à stabiliser une molécule d'ADN dans la levure n'a pas encore été évaluée par manque de temps. Pour pallier une éventuelle faible efficacité de cette ARS, il a été décidé d'ajouter la construction 18 (=ARS1406) à 63 kb de l'ARS naturelle (figure 77, tableau 22).
- 5) En relisant la littérature, nous nous sommes aperçus que les ARS416 (construction 11) et ARS1213 (construction 5) étaient potentiellement peu ou pas efficaces pour stabiliser une molécule d'ADN dans la levure (Venetz *et al.*, 2019). Pour pallier ce manque d'efficacité, nous avons décidé de « doubler » ces ARS en plaçant des ARS additionnelles au niveau de ces *loci*. Le second design intègre donc la construction 16 (=ARS603) placée à 20 kb de l'ARS416, entre les gènes *ytoA* et *ytnA*, et la construction 13 (=ARS1332) placée à 15 kb de l'ARS1213 entre les gènes *stoA* et *zosA* (figure 77).
- 6) Concernant les captures de fragments de grande taille (demi-génome (>1.466 Mb) et plus grand), un problème concernant la taille de la molécule à cloner pourrait survenir. Il se peut donc que malgré la présence d'un nombre d'ARS fonctionnelles suffisantes (ainsi que de marqueurs d'auxotrophie), la levure ne soit pas apte à accueillir et propager le génome complet de *B. subtilis* de 2.944 Mb pour la souche utilisée ici. Il nous est donc apparu judicieux de produire une souche pour laquelle il serait possible de capturer le génome complet sous forme de deux « demi-génomes » qui

**Tableau 22 : Liste des ARS supplémentaires à ajouter au génome de *B. subtilis* d'après les design n°2 et n°2 Bis. La construction 6 (grisée) est éliminée de ces design.**

Numéro de la construction	Gène proche	Nom de l'ARS	Séquence de l'ARS	Marqueur d'auxotrophie	Référence
1	<i>gyrA</i>	ARSMaX2	TTCAAAATTTTATGTTTTGTTTCGAGATCCACAGTTTTTCGGAGCATAGGAAGCTCGTAT ATGATAACAAAAATAAATATATAAAAAATAAAAAATAATGA	-	Venez et al., 2019
2	<i>S70</i>	ARS1113	CAATCTCATTGTTTACGTA AAAATGAAAACGTTTTGTGTTAGAATTGACAGATTGAA GTTTAAATTTAGGTGTCGGGAGAAACATTTTAAATCCCTTTTATGAATATAATTTAAGT ACTAAAAAATCACCATAATTATACAAAGTAAGTCAATTTAAAAAGTACGTTTTATAAT TAATTCACATACACATGTTATTGTATCTAACTAAATAGTCTCGAAATAGTTC AATTGA	-	Venez et al., 2019
3	<i>yerl</i>	ARS209	ACCTTTGTTATTTTTTTTCTTCATTCCGTAACCTCTTACCTCTTTATTTACTTTCTAA AATCCAAATACAAAACATAAAAAAT	-	Venez et al., 2019
4	<i>oppA</i>	ARS1	AGCTTACATTTTATGTTAGCTGGTGGACTGACGCCAGAAAATGTTGGTGATGCGCTTA GATTAATGGCGTTATTGGTGTGATGTAAGCGGAGGTGGAGACAAATGGTGTA AAGACTTAACAAAATAGCAAATTCGTA AAAATGCTAAGAAATAGGTTATTACTGA GTAGTATTTTAAAGTATTGTTGTGCACTGCCTGCAAGCCTTTTGAAAAGCAAGCA TAAAAGATCTAAACATAAAATCTGTA AAAATAACAAGATGTAAGATAATGCTAAATCA TTTGGCTTTGATTGATTGTACAGGAAAATATACATCGCAGGGGGTTCCTTTTACCAT TTCACCGCAATGGAATCAAATCTGTTGAAGAGAATGTTACAGGCGCATACGCTACAA TGACCCGATTCTTGTAGCCTTTCTCGGTCTTGCAAAACACCGCCGCGCAGCTTAGTAT ATAAATACACATGTACATACTCTCTCCGTATCCTCGTAATCATTTCCTGTTATTATCGT CTTTTCGCTGTA AAAACTTTATCACACTTATCTCAAATACACTTATAACCGCTTTACTA TTACTCTACGCTGACAGTAATATCAAACAGTGACACATATTAACACAGTGGTTTTCT TGATAAACACCATCAGCCTCAAGTCTCAAGTAAAGATTTCTGTTTCATGCAGATAGA TAACAATCTATGTTGATAATTAGCGTTGCCTCATCAATGCGAGATCCGTTTAAACGG ACCCTAGTGCACTTACCCACGTTCCGTCCTGTGTGCCGAACATGCTCTTCACTATT TTACATGTGG	-	Liachko et al., 2013
5	<i>ykwD</i>	ARS1213	TGTTTACATTTCAAGTCTAATCTGTGCTTTTACCGAAGAGGAATTTTCATAAATACGG AGAAAATATAAAAAAAAAGT	-	Venez et al., 2019
6	<i>ylqB</i>	ARS413	GAATAATAGAATGAGTTACATACGTAACGTTAGTATTATTATTATTCTATTTC TTTCGGTTTTTATTATTATATATATATATATATATATATTGGGGTTTATATTTA GGTAGGGCATAAGGATTTACTGTGCGCATGAACCTACTTGAACAAAAAATAAAAAATC AAAAAGAAAAATAATGTAGGCTTTTGACATCGACTATATAAATATATATATACACAC TTCTGTGCAGAA	-	Liachko et al., 2013
7	<i>yocI</i>	ARS516	TTTTTCATAATTTTACTTTTTGGTGCAGAAATCTGTAATATTTTAAAAACAAATCTCA	-	Liachko et al., 2013
8	<i>ypuC/2</i>	ARS1018	TCAGAGCAATACGTTTATTTTTTCAACTTCGCATGGCATTATGTATCATTGCAG CGTAAATTTTTACATTTAGTAATATCTAATAACATATACACAAAAAGGGACAGATTT TTGAATTAATAAACCAAGCGTAAATCAGTATTGCTCTTACCGCAGTAGGGGGAGTT ATCTACTCTGCTTAACGTAGAGAAACAATAGTAAAAGTTATGCTTGTGTTGGAATA AAAAAT	-	Venez et al., 2019
9	<i>lepA</i>	ARS727	TTTTTTTATTGTTCTAATAACTTCAATTTGTTACTTTTTGCTTTTGTATTATGGA AATAATGCAGGTTAACATAAGTAAACATAAAATATG	-	Venez et al., 2019
10	<i>infC</i>	ARS1323	CAATTAACCAAGTCTTTTATTTTTTAAATGGTTTATTTATGAATGTTTATGTTTATG TTCTTTACTCAAAAATCAATAAGTGCTTTTTTACGGAAATTTCTCTACGAATATAT ATAAGTCATATCTAAGACTTCCATCCGTCCTTTTACGTTTGTATGATGCATACAAT GATGCCAAATCAGCTTGATGCTGTTGGTATTCAGTCGTC	-	Venez et al., 2019
11	<i>menC</i>	ARS416	GGCCTTTGAAAAGCAAGCATAAAAGATCTAAACATAAAATCTGTA AAAATAACAAG	-	Venez et al., 2019
12	<i>cwI</i>	ARSHI	TTACAGATTTTATGTTTATGATCTTTTATGCTTCTTTTCAAAGGCTCGGGCAAGT GCACAAACAATACTTAAATATATACTACTTAGTAATAACCTA	-	Venez et al., 2019
13	<i>zosA</i>	ARS1332	TTCTGCTATTGATTGGTAGTGTTCGGGTTAATTCATTTTTTGTTCATAACAAGAA CCAAACATAAAAAAATACTTTAAGAATAAACAAGACTGATGAAGCATAGAT	-	Liachko et al., 2013
14	<i>ylyA</i>	ARS914	GTATTGTCGTGATAGTAGCATGAATCTAGCGTGGAAATCTTTTATTTAAGGTCGCG TTGTGACGACAAAACATTA AAAAACACATGTGTTGATTGTAGACTTGATACACTC	-	Liachko et al., 2013
15	<i>ymfF</i>	ARS1120	CTGCGAAGAAAGAAAAAGAAATATTCACTCTAGGAGCCAAAAACATAAAAAA AAATAATTA AAAAATACTTTTATTGAATGCTTAATTTAGTGACGCGATTACCTAT ATACCAGGTTAT	-	Liachko et al., 2013
16	<i>ytnA</i>	ARS603	TCATAGGTTAAATAGGACGTTCTCCAAAATATGTAACAATACTATATAGAATAATAT ATGTGTTTTATTCTGTTTTAATATTTCCAGATTTTAAAGTTTTGATACAAAATATGA AATTAGGGACTTCAAAAATCTAATTTA	-	Liachko et al., 2013
17	<i>xynC</i>	ARS1019	TGTATATAAAGAGGTTGTA AAAATGACGAATAAATAGTATTTAGGCAACTAAAG ATAAAAAAATATTATTATTTTTTTCGCGGTGCGTGGCAGATTTTTTTGACATGC GGAA	MET14	Liachko et al., 2013
18	<i>yhdF</i>	ARS1406	TTGTA AAATTTACTACCCTTATCTTCAATGGTGTTTTTAATGTTTTGTAGAGG AAAACATAAAGTTGCTATTTTTTATTCAGATCGCAATCCATGTTTGACACATTCTA	-	Liachko et al., 2013
18'				CEN4-HIS3	
19	<i>frlB</i>	ARS404	AGAAAAAGTCAAAAATGGTTTTTTTCATCCAAAATATTAATTTTACTTTTA	LYS2	Liachko et al., 2013

' : design alternatif au niveau de la construction

potentiellement pourraient ensuite être assemblés à travers des expériences de mating. Deux versions ont donc été proposées pour ce second design. Les deux versions seront identiques l'une à l'autre, excepté au niveau de la construction 18 (ARS1406). En effet, dans la seconde version, l'ARS1406 sera associée au centromère CEN4 et au marqueur HIS3 qui ne sera donc plus présent au côté de l'ARSH4 (tableau 22).

Pour résumer les modifications réalisées au niveau de ce second design : (i) la construction 19 correspondant à l'ajout de l'ARS404 et du marqueur LYS2 a été validée ; (ii) les constructions 13 (ARS1332), 14 (ARS914), 15 (ARS1120), 16 (ARS603), 17 (ARS1019), et 18 (ARS1406) ont été ajoutées et (iii) la construction 6 a été éliminée (figure 77) (tableau 22).

En se basant sur ce second design, la souche *B. subtilis* qu'il faut désormais construire contient 19 ARS réparties de manière homogène le long du génome. Il est important de noter qu'avec ce design, la distance maximale séparant 2 ARS est de 201 kb entre les constructions 4 et 13. Un design alternatif avec un second centromère (CEN4) au niveau de la construction 18 a aussi été proposé.

## 7. Conclusion

Jusqu'à ce jour, le génome de *B. subtilis* n'a jamais pu être cloné entier dans la levure. Or, cloner ce génome dans la levure serait une avancée majeure pour le projet « Bacillus 2.0 ». D'une part, nous aurions la démonstration que la levure est un hôte pertinent pour assembler et construire des génomes semi-synthétiques de *B. subtilis*, et d'autre part, nous pourrions amorcer les expériences de transplantation de génomes sans lesquelles les souches de *B. subtilis* modifiées ne peuvent pas être produites (figure 1, étape 5).

Au cours de ce chapitre, nous décrivons un travail en cours, basé sur un cycle DBTL (Design, Build, test, Learn) qui vise à produire une version du génome de *B. subtilis* qui puisse être clonée dans la levure et maintenue stable pendant plusieurs générations. Cette stratégie, mise en place en collaboration étroite avec l'équipe SyBER (Institut MICALIS, UMR 1319 INRAE / AgroParisTech, Jouy-en-Josas), repose en grande partie sur la méthode de CREasPy-Fusion initialement développée chez les mycoplasmes (Guesdon *et al.*, 2023) et adaptée à *B. subtilis* tel que décrit dans le chapitre 1.



Pour atteindre cet objectif, la première étape a consisté à concevoir sur ordinateur un génome de *B. subtilis* modifié de façon à contenir des modules qui le stabilisent et à être dépourvu de *loci* qui le fragilisent. Dans un premier temps, nous nous sommes focalisés sur les modules à ajouter et avons proposé un premier design qui incluait 13 ARS et un marqueur d'auxotrophie. Ce design émanait (i) des résultats obtenus au cours des expériences de TAR-cloning qui avaient, entre autres, permis d'identifier certains facteurs empêchant le clonage du génome complet dans la levure comme le manque d'ARS fonctionnelles et la présence de régions répétées et/ou de gènes dont l'expression peut être toxique pour la levure, mais aussi (ii) d'analyses bioinformatiques et (iii) de données de la littérature.

La construction de la souche *ad hoc* a été faite par nos collaborateurs de l'équipe SyBER (Institut MICALIS, UMR 1319 INRAE / AgroParisTech, Jouy-en-Josas) en suivant un processus de production basé sur la compétence naturelle de *B. subtilis* (figure 67) et permettant de générer, sans trace ni marqueur, aussi bien des souches *B. subtilis* portant des ARS uniques (total 13 souches) que des souches dans lesquelles les ARS s'accumulaient dans un ordre bien défini.

Après vérification, ces souches étaient ensuite envoyées, les unes après les autres, à Bordeaux pour des expériences de CReasPy-Fusion qui permettaient de valider ou invalider expérimentalement, l'apport des ARS, sur la stabilité des fragments du génome de *B. subtilis* dans la levure. Les données acquises lors de ces expériences ont montré que le plan de construction initial n'était pas optimal et qu'il était nécessaire de le réviser (ajout d'ARS et de marqueurs d'auxotrophie supplémentaires), si nous voulions cloner de plus grands fragments de *B. subtilis* dans la levure *S. cerevisiae*.

La première expérience de CReasPy-Fusion menée dans le cadre de ce projet a montré que (i) nous étions capables de capturer un fragment de 536 kb contenant une seule ou deux ARS supplémentaires et (ii) que la présence de deux ARS, comparée à une seule ARS, semblait augmenter l'efficacité de capture. Avec cette expérience, la stabilité de la région du génome de *B. subtilis* comprise entre les gènes *IrgB* et *yufO* avec des ARS supplémentaires a pu être validée.

Les expériences de CReasPy-Fusion suivantes ont été menées dans l'objectif de capturer la région du génome de *B. subtilis* comprise entre les gènes *IrgB* et *obg*, d'une taille proche de



1 Mb, précisément de 925 kb. La capture de cette région semblait assez ambitieuse car elle était d'une taille quasiment deux fois supérieure à celle du dernier fragment capturé (536 kb). Lors de la première tentative, le fragment de 925 kb n'a pas pu être capturé que ce soit avec une seule ou 4 ARS supplémentaires. Ce n'est que lors d'une seconde tentative, après l'ajout d'une cinquième ARS seule ou accompagnée d'un second marqueur d'auxotrophie (LYS2), que le fragment a pu être cloné entièrement dans la levure. L'ensemble de ces deux expériences a montré (i) que nous étions capables de capturer un fragment de 925 kb avec 5 ARS supplémentaires que ce soit avec un seul ou 2 marqueurs d'auxotrophie, et que (ii) la présence d'une seule ARS supplémentaire n'était définitivement pas suffisante pour stabiliser un fragment de cette taille dans la levure. Avec cette expérience, la stabilité de la région du génome de *B. subtilis* comprise entre les gènes *IrgB* et *obg* avec 5 ARS supplémentaires a pu être validée que ce soit avec ou sans le marqueur LYS2. Il est important de noter que lors de ce travail, tous les fragments de génome, de tailles croissantes, englobaient le fragment (de plus petite taille) capturé lors de l'étape précédente.

Enfin, une dernière expérience de CReasPy-Fusion a été réalisée dans le but de capturer un demi-génome. Cette expérience a montré que malgré l'ajout de 8 ARS supplémentaires ainsi qu'un ou deux marqueurs d'auxotrophie, la capture de la région comprise entre les gènes *IrgB* et *parE* de 1.466 Mb n'est pas encore possible. A ce stade plusieurs facteurs, seul ou en combinaison, peuvent expliquer cet échec :

- **Le design.** Les essais de capture du demi-génome ont été réalisés avec les souches de *B. subtilis* EDJ2052 ARS404 et EDJ2052 ARS404 LYS2 qui possédaient 8 ARS en plus du marqueur HIS3 et LYS2 (EDJ2052 ARS404 LYS2). Ce design n'est peut-être pas suffisant ou optimal pour cloner ce fragment dans la levure.
- **La taille du fragment.** La CReasPy-Fusion est une méthode relativement récente (Guesdon *et al.*, 2023). Elle a été développée pour des espèces du genre *Mycoplasma* qui possèdent de petits génomes. A ce jour, aucune équipe n'a encore tenté de cloner un génome de taille supérieure à 1.2 Mb dans la levure par CReasPy-Fusion. Nous ne savons donc pas si cette technique permet de capturer des fragments avec des tailles supérieures et ne connaissons pas la taille limite des fragments à cloner. Il pourrait être intéressant, parallèlement à nos expériences, de réaliser des expériences « contrôles » mettant en jeu des génomes de mycoplasmes (ou espèces apparentées de la classe *Mollicutes*) possédant des tailles supérieures à 1.2 Mb afin de valider ces



points. Les génomes d'*A. laidlawii* (1.5 Mb) (Karas *et al.*, 2012) ou *S. citri* (1.8 Mb) (Labroussaa *et al.*, 2016) ont par exemple déjà été clonés dans la levure par d'autres techniques et pourraient être utilisés au cours d'expériences de CREasPy-Fusion pour répondre à ces questions.

- **La présence d'une ou plusieurs régions « problématiques » pour la levure.** Outre le design et/ou la taille, un troisième facteur peut potentiellement à lui seul (ou en combinaison) abroger le clonage du demi-génome dans la levure. Il s'agit du contenu en gène(s) du fragment ou de sa séquence. En effet, des données de la littérature montrent que des gènes émanant de séquences procaryotes peuvent parfois être exprimés dans un contexte eucaryote. Si le produit d'expression est toxique ou mal toléré par son hôte, des remaniements excluant ces zones peuvent être sélectionnés au cours de la propagation des cellules et générer des clones levures hébergeant des fragments tronqués. De même, la présence de nombreuses séquences répétées sur le fragment cible peut induire des événements de crossing-over par recombinaison homologue générant des fragments dont la taille et la composition ne seront pas celles recherchées.

Dans le cas du demi-génome analysé ici, on peut noter par exemple que lors de la PCR multiplex, les couples d'amorces correspondant aux PCR 7, PCR 8 (et possiblement PCR 3) ne produisent pas d'amplicon. Il est probable que ces régions contiennent des gènes ou séquences critiques pour l'hôte levure. Il pourrait être intéressant d'identifier ces zones et tenter de les supprimer pour confirmer nos hypothèses.

L'ensemble des expériences de CREasPy-Fusion menées jusqu'à présent ont montré que nous étions capables de stabiliser un fragment de 925 kb avec 5 ARS, mais que si nous voulions capturer un fragment de plus grande taille alors les régions adjacentes aux PCR 7 et 8 réalisées lors de l'analyse des transformants de la tentative de capture du demi-génome (figure 76) devaient être modifiées pour être clonées.

En accord avec nos collaborateurs, un nouveau design a été proposé, avec cette fois l'ajout de 19 ARS supplémentaires le long du génome de *B. subtilis* (figure 77). En addition de ces 19 séquences, 2 marqueurs d'auxotrophie ont aussi été ajoutés. Un design alternatif a



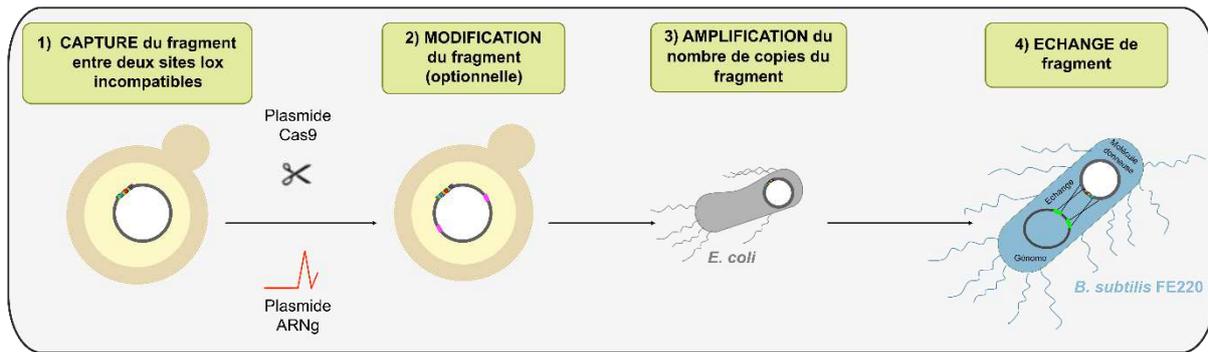
également été proposé avec l'ajout d'un deuxième CEN qui permettrait de tenter de cloner le génome bactérien en deux moitiés distinctes.

Lorsque les souches correspondantes seront prêtes, la capture du demi-génome sera de nouveau tentée. Une fois cette première moitié stabilisée dans la levure, la deuxième moitié du génome pourra aussi être clonée grâce à l'ajout du second centromère dans le design secondaire. Il est fort possible que la capture de cette seconde moitié de génome ne soit pas triviale. En effet, lors des expériences de TAR-cloning, la région comprise entre les gènes *sacX* et *swrC* a été classée comme une région instable, probablement à cause de la présence de régions répétées correspondant aux nombreuses copies d'opérons ribosomiques (cf partie 2 de l'avant-propos). La capture de cette deuxième moitié de génome pourrait donc nécessiter deux étapes. La première consisterait à capturer ce demi-génome sans cette région afin de démontrer qu'elle pose bien problème. La deuxième étape consisterait à tenter de stabiliser cette région en supprimant certaines répétitions, puis à tenter de cloner cette version modifiée dans la levure. Si cela donne les résultats escomptés, des tentatives de clonage du génome complet sans cette région (ou avec la région modifiée) pourront être envisagées ; cela correspondrait à la capture d'une molécule de 2,5 Mb dans la levure, ce qui n'a jamais été réalisé.

Pour conclure, la stratégie que nous avons mise en place et décrite au cours de ce chapitre pourrait *in fine* permettre de développer une version du génome de *B. subtilis* qui serait complètement stable dans la levure. L'utilisation de la méthode de CReasPy-Fusion dans cette stratégie montre sa flexibilité, mais on peut aussi imaginer cloner dans la levure la version retravaillée du génome de *B. subtilis* avec d'autres méthodes de clonage comme la méthode « classique » (Benders *et al.*, 2010), la méthode de CReasPy-Cloning (Ruiz *et al.*, 2019) ou la méthode de clonage par fusion (Karas *et al.*, 2013).



## Chapitre 3 : Adaptation de la méthode RAGE à *B. subtilis*



**Figure 78 : Schéma bilan de la méthode RAGE.** La méthode RAGE se divise en 4 étapes: **1)** Capture dans la levure d'un fragment de génome bactérien d'intérêt entre deux sites lox incompatibles. **2)** Modification du fragment (délétion, insertion, remplacement) en utilisant les outils génétiques disponibles chez cet hôte. **3)** Transformation du fragment dans *E. coli*, pour permettre l'amplification du nombre de copie. **4)** Echange de fragments d'ADNg, médié par le système de recombinaison site-spécifique cre/lox, dans une cellule receveuse.

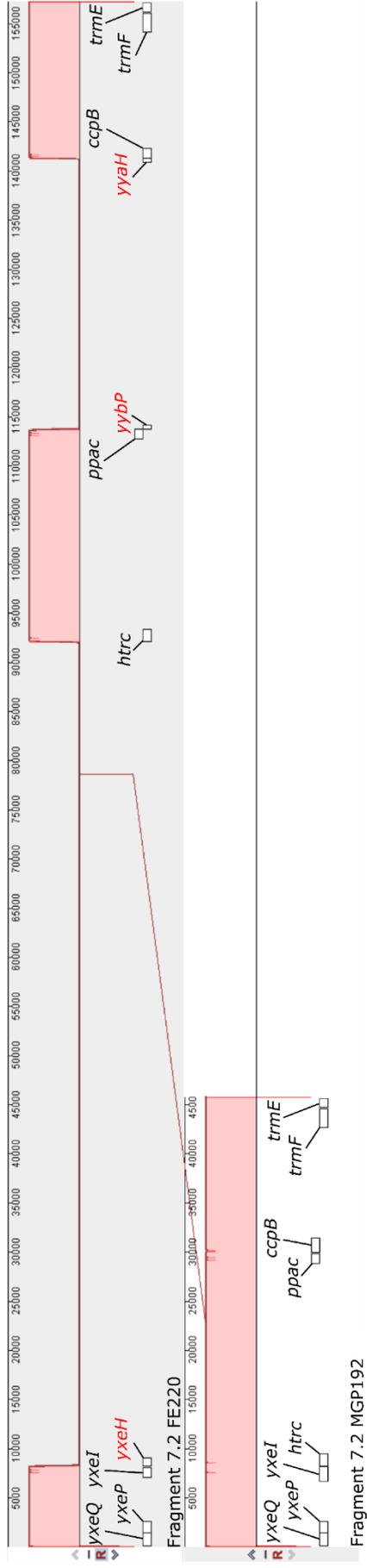
## Chapitre 3 : Adaptation de la méthode RAGE à *B. subtilis*

### 1. Introduction

Le projet Bacillus 2.0 avait deux objectifs majeurs : (1) assembler un génome semi-synthétique dans la levure, puis (2) le transférer dans une bactérie afin d'obtenir une souche modifiée à grande échelle. Pour cette seconde et dernière étape, nous avons planifié d'utiliser une méthode connue sous le terme de « Transplantation de génomes » (TG) qui permet de transférer un génome bactérien entier dans une cellule dite receveuse et de générer des cellules dont le génotype et le phénotype sont dictés uniquement par le nouveau génome entrant (Lartigue *et al.*, 2007; Lartigue *et al.*, 2009). Cette méthode avait notamment été utilisée pour la construction de la première cellule synthétique JCVI Syn1.0 (Gibson *et al.*, 2010) et de la première cellule synthétique (quasi) minimale JCVI Syn3.0 (Hutchinson *et al.*, 2016). Cette méthode, très puissante, permet de travailler à l'échelle de génomes entiers, cependant, elle nécessite de nombreuses mises au point avant d'être fonctionnelle chez une nouvelle espèce. L'adapter à *B. subtilis* constituait donc un véritable défi.

Au démarrage de ma thèse, les membres du consortium étaient de plus en plus conscients que le premier objectif (si tant est qu'il puisse être atteint) allait demander plus de temps et de travail que prévu. Cette prise de conscience les a incités à réexaminer le programme de travail et à suspendre les tâches/sous-tâches en lien avec l'objectif 2 tant que des questions autour de la faisabilité de l'objectif 1 subsistaient.

Au cours des mois qui ont suivi cette décision, le clonage dans la levure, ne serait-ce que d'un génome natif de *B. subtilis*, n'était toujours pas réalisé. Cependant des progrès étaient notés puisque plusieurs fragments (allant d'une dizaine de kilo paires de bases à plusieurs centaines) avaient été clonés avec succès chez cet hôte grâce à la mise en place d'expériences de TAR-cloning puis de CREasPy-Fusion (cf. chapitres 1 et 2). Une fois clonés dans la levure, des expériences d'ingénierie de génomes médiées par le système d'édition CRISPR-Cas9, avaient été également réalisées et montraient qu'il était possible de modifier les fragments de génome de *B. subtilis* maintenus dans la levure. L'ensemble de ces résultats nous ont amenés à réfléchir à une méthode alternative à la TG qui permettrait de modifier *B. subtilis*, et donc de générer des mutants mais sans passer par la manipulation de génomes entiers.



**Figure 79 : Alignement entre le fragment 7.2 de *B. subtilis* FE220 (155 kb) et celui de *B. subtilis* MGP192 (44 kb).** Les régions présentent dans les deux séquences sont représentées en rose. Les cassettes blanches correspondent à des gènes, ceux écrits en rouge sont absents du fragment de la MGP192, ceux écrits en noir sont présents dans les deux versions.

Ainsi, dans le but de valoriser le travail réalisé jusque-là et de montrer que la levure restait une plateforme pertinente et adaptée pour l'ingénierie du génome de *B. subtilis*, nous avons tenté d'adapter la méthode RAGE, développée en 2020 chez les mycoplasmes (Garcia-Morales *et al.*, 2020), à *B. subtilis*. Le principe de cette méthode est de transférer, non pas un génome complet, mais seulement d'une portion de ce génome de la levure vers la bactérie à éditer. Elle se rapproche des méthodes REXER/GENESIS mises au point chez *E. coli* pour construire les premiers génomes réduits et recodés (Fredens *et al.*, 2019).

D'un point de vue pratique, elle se décline en 4 étapes : (1) capture, entre deux sites lox incompatibles, d'un fragment de génome bactérien d'intérêt dans la levure, (2) modification du fragment capturé en utilisant les outils génétiques disponibles chez la levure, (3) transformation du fragment dans *E. coli* pour permettre l'amplification de son nombre de copies, puis (4) échange de fragments dans la cellule receveuse via le système de recombinaison sites-spécifiques cre/lox (figure 78). En somme, cette stratégie est la combinaison de plusieurs méthodes déjà connues, à savoir (i) les méthodes de capture dans la levure des fragments de génomes bactérien comme le TAR-cloning (Natalay Kouprina and Vladimir Larionov, 2016) ou la CReasPy-Fusion (Guesdon *et al.*, 2023), (ii) les méthodes d'ingénierie de génome dans la levure, et enfin (iii) la méthode RMCE (Recombinase-mediated cassette exchange), une technique d'échange de fragments d'ADN basée sur le système cre/lox (Noskov *et al.*, 2015).

Pour adapter la méthode RAGE à *B. subtilis*, nous avons choisi de travailler sur la région située entre les gènes *yxeQ* et *trmE* du génome de *B. subtilis*. Cette région a une taille de 155 kb chez la souche *B. subtilis* FE220 que nous avons sélectionnée comme cellule receveuse, et une taille de 44 kb chez la souche à génome réduit MGP192. Le fragment de 155 kb dans le génome de la souche FE220 correspond au fragment de 44 kb dans le génome de la souche MGP192 mais s'en distingue par la présence de nombreux gènes non essentiels (121 au total), qui ont été éliminés du fragment de 44 kb (figure 79). En d'autres termes, tous les gènes présents sur le fragment de 44 kb sont des gènes essentiels, retrouvés à l'identique sur le fragment de 155 kb. L'inverse n'est pas vrai puisque les gènes non essentiels trouvés sur le fragment de 155 kb sont absents du fragment de 44 kb.

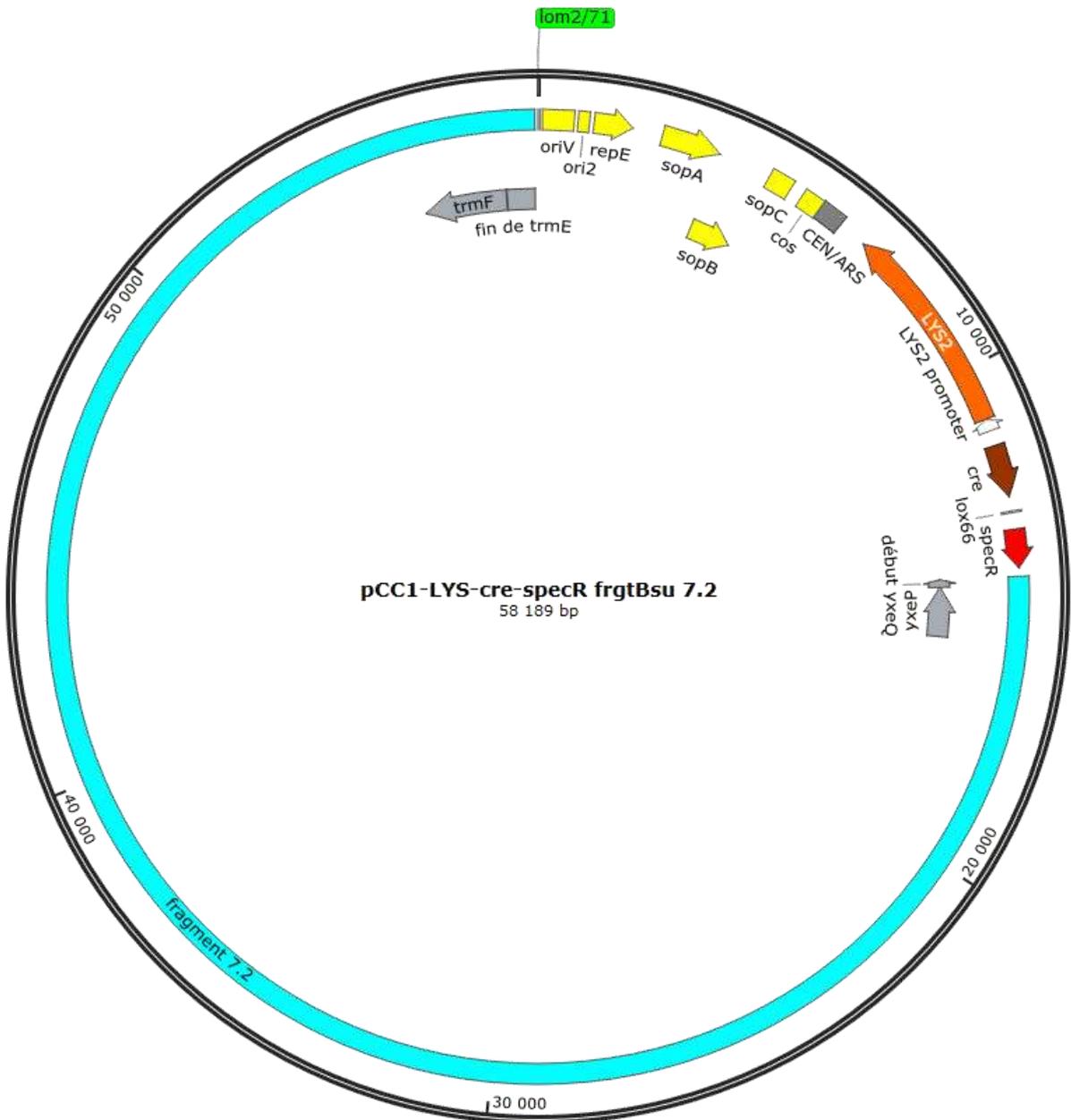


Figure 80 : Carte de la molécule donneuse pCC1-LYS-cre-specR frgtBsu 7.2 d'~58 kb.

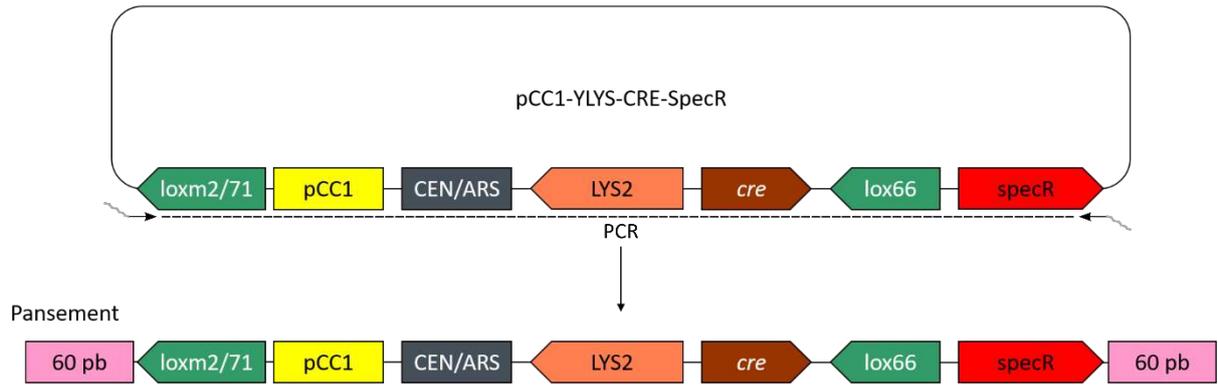
A mon arrivée au laboratoire, la région de 44 kb du génome réduit MGP192 avait déjà été capturée dans la levure par TAR-cloning. Lors de ces expériences, cette région avait été appelée « fragment 7.2 » (figure 3 de l'avant-propos) (une dénomination que nous conserverons tout au long de ce chapitre). L'étape 1 du RAGE qui consiste à capturer un fragment cible dans la levure, était donc partiellement réalisée puisque le fragment d'intérêt était cloné dans la levure, seulement il n'était pas bordé de deux sites lox incompatibles. Nous sommes donc passés à l'étape 2 du RAGE, qui permet, si besoin, de modifier le fragment cloné dans la levure grâce au système d'édition CRISPR-Cas9. Une fois les modifications réalisées, nous avons poursuivi avec l'étape 3, et enfin l'étape 4 qui a nécessité, comme décrit ci-après, de très nombreux ajustements avant de fonctionner.

## 2. Etape 2 du RAGE : Modification du fragment 7.2 cloné dans la levure pour obtenir la molécule donneuse pCC1-LYS-cre-specR frgtBsu 7.2

Le fragment 7.2, qui correspond à une région de 44 kb dans le génome de *B. subtilis* MGP192, avait été cloné précédemment dans la levure et se maintenait stable chez cet hôte grâce à la présence des éléments levure (ARS, CEN, HIS). Bien qu'essentiels, ces éléments n'étaient, cependant, pas suffisant pour mettre en place la méthode RAGE. Le premier travail a donc consisté à modifier la molécule clonée dans la levure, en remplaçant le locus ARS CEN HIS par un « pansement de recombinaison » portant tous les éléments nécessaires à la mise en œuvre de la méthode RAGE, aboutissant à la molécule pCC1-LYS-cre-specR frgtBsu 7.2 d'environ 58 kb (figure 80).

### 2.1. Production du pansement de recombinaison contenant toutes les séquences nécessaires au RAGE

Pour insérer, sur le fragment 7.2 cloné dans la levure, tous les éléments nécessaires à la méthode RAGE, il nous a fallu concevoir puis fabriquer un pansement de recombinaison contenant les éléments nécessaires au maintien de la molécule chez 3 hôtes différents, à savoir la levure *S. cerevisiae*, *E. coli* et *B. subtilis*. Ainsi le pansement contenait : (i) les éléments ARS, CEN et le marqueur d'auxotrophie pour la lysine (LYS2), pour le maintien et la sélection chez la levure, (ii) le plasmide CopyControl pCC1 pour le maintien et le contrôle du nombre de copie chez *E. coli*, et (iii) le gène *cre*, deux sites lox incompatibles entre eux, et le gène de



**Figure 81 : Schéma de la production du pansément de recombinaison qui porte toutes les séquences nécessaires à la mise en place de la méthode RAGE.** Le pansément (~14 kb) est produit par PCR à partir du plasmide pCC1-YLYS-CRE-SpecR avec les amorces PST\_pCC1pYAdCreSpec7.2\_F et PST\_pCC1pYAdCreSpec7.2\_R, ces amorces possèdent des queues flottantes d'environ 60 pb identiques aux extrémités de fragment 7.2 qui permettent d'ajouter ces séquences au pansément lors de la PCR.

résistance à la spectinomycine *specR*, pour l'étape d'échange via le système *cre/lox* dans la cellule receveuse *B. subtilis* (figure 81).

Pour produire ce pansement, nous avons commencé par construire un plasmide appelé pCC1-YLYS-CRE-SpecR, qui regroupait tous les éléments cités précédemment. Puis, à partir de ce plasmide, une PCR a été réalisée en utilisant des amorces à queues flottantes d'environ 60 bases pour ajouter, de part et d'autre du pansement, des régions identiques aux séquences bordant le locus cible (figure 81). Le produit PCR d'une taille d'environ 14 kb a été purifié puis stocké à -20°C jusqu'à utilisation.

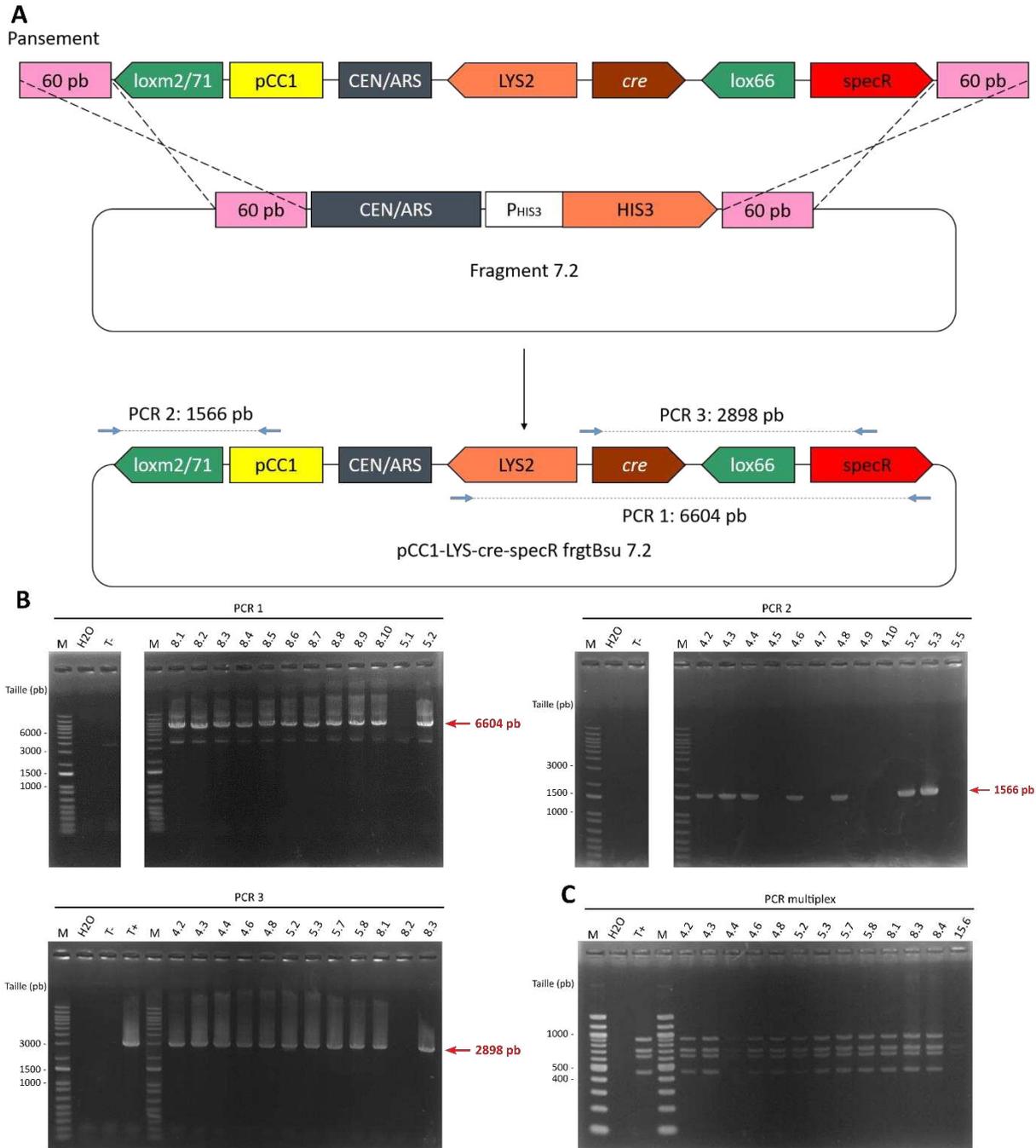
## 2.2. Edition du fragment 7.2 cloné dans la levure

Pour modifier le fragment 7.2 maintenu dans la levure, nous avons choisi d'utiliser le système d'édition CRISPR-Cas9 chez la levure. Pour ce faire, deux clones levures de notre collection TAR, les clones VL6-48[MGP192sous-fgt 7.2] cl **4.1** et **4.3** qui hébergeaient le fragment 7.2, ont été mis en culture puis transformés avec le plasmide p414-TEF1p-Cas9-CYC1t permettant l'expression de la Cas9. Les transformants ont été sélectionnés sur un milieu solide SD-HIS-TRP puis analysés par PCR afin de vérifier la présence du plasmide d'expression de la Cas9. Deux clones ont ainsi été sélectionnés:

- VL6-48[MGP192 sous-fgt 7.2]Cas9 cl **2.1** provenant de VL6-48[MGP192 sous-fgt 7.2] cl 4.1
- VL6-48[MGP192 sous-fgt 7.2]Cas9 cl **6.4** provenant de VL6-48[MGP192 sous-fgt 7.2] cl 4.3

Ces deux clones ont ensuite été transformés avec un plasmide d'expression d'un ARNg ciblant spécifiquement le locus à éditer (ARS CEN HIS) et le pansement de recombinaison fraîchement produit. Après 48 h à 30°C sur milieu solide SD-LYS-TRP-URA, 441 colonies sont apparues pour le clone levure 2.1 et 401 pour le clone levure 6.4. Pour chaque condition, 40 clones ont été piqués puis analysés par PCR.

Plus précisément, l'intégration des éléments nécessaires au RAGE en lieu et place du locus ARS CEN HIS a été validée par 3 PCR indépendantes : les deux premières se situant aux jonctions de la zone éditée (PCR 1 et 2) et la dernière au niveau des gènes *cre/specR* (PCR 3) (figure 82A). Au total, 23 clones ont présenté les amplicons aux tailles attendues pour les trois



**Figure 82 : Modification du fragment 7.2 par insertion du pansement de recombinaison. [A]** Schéma de l'insertion par recombinaison homologue du pansement contenant toutes les séquences nécessaires à la mise en place de la méthode RAGE sur le fragment 7.2. **[B]** L'intégration du pansement est validée par 3 PCR simplex produisant des amplicons de 6604 pb avec les amorces Verif LYS F et Verif 7.2 ADE R pour la PCR 1, 1566 pb avec les amorces verif 7.2 pCC1 F et Verif pCC1 R pour la PCR 2 et 2898 pb avec les amorces Gib ADE2 cre T7 F V2 et Gib loxm2/71 Spec R pour la PCR 3. « M » : DNA Ladder 1 kb+ Invitrogen (100-12000 pb). « H2O » : contrôle négatif sans ADN. T- : témoin négatif ADNg de la souche levure. T+ : contrôle positif avec le pansement de recombinaison. Par souci de clarté seulement 12 clones par PCR sont montrés. **[C]** L'intégrité du fragment 7.2 est validée par une PCR multiplex produisant 4 bandes. « H2O » : contrôle négatif sans ADN. T+ : contrôle positif ADNg de la souche *B. subtilis* MGP192. M : DNA Ladder 100bp NEB (100-1517 bp). Par souci de clarté, seulement 13 clones sur les 23 analysés sont présentés ici.

PCR (tailles respectives: 6604 pb, 1566 pb et 2898 pb ; figure 82B). L'intégrité du reste du fragment a ensuite été contrôlée par PCR multiplex à l'aide de 4 couples d'amorces placés tout le long du fragment. Sur les 23 clones analysés, 19 ont montré le profil attendu à 4 bandes (figure 82C). L'absence de mutation au niveau des sites lox et du gène *cre* a été validée par séquençage Sanger.

Ces résultats ont montré que les éléments nécessaires à la mise en place de la méthode RAGE avaient bien été insérés sur le fragment 7.2 à la place du locus ARS CEN HIS, validant ainsi l'étape 2 du RAGE qui offre la possibilité de modifier la molécule capturée en étape 1, si cela est nécessaire.

Nous n'avons pas entrepris de modifications supplémentaires sur ce fragment et sommes donc passés directement à l'étape 3 après ce travail. En effet, la version du fragment 7.2 maintenue dans la levure provient de la souche *B. subtilis* MGP192 qui possède un génome réduit d'une taille de 2,86 Mb (Dervyn *et al.*, 2023) et de nombreux gènes non-essentiels ont déjà été supprimés de ce génome (y compris au niveau du fragment 7.2). La version du fragment 7.2 provenant de la souche MGP192 (44 kb) est donc suffisamment différente de celle présente dans la cellule receveuse FE220 (155 kb).

Pour la suite des expériences les clones suivants ont été sélectionnés :

- VL6-48 [MGP192 frgt7.2 cl 4.1]Cas9 PST\_pCC1-YLYS-CRE-SpecR cl **4.8** obtenu à partir de VL6-48[MGP192 sous-fgt 7.2]Cas9 cl 2.1
- VL6-48 [MGP192 frgt7.2 cl 4.3]Cas9 PST\_pCC1-YLYS-CRE-SpecR cl **16.2** obtenu à partir de VL6-48[MGP192 sous-fgt 7.2]Cas9 cl 6.4

### 3. Etape 3 du RAGE : Transformation de la molécule donneuse (pCC1-LYS-cre-specR frgtBsu 7.2) dans *E. coli* EPI 300

A ce jour, il n'existe pas de protocole permettant d'extraire de l'ADN plasmidique hautement pur à partir de la levure *S. cerevisiae*. Nous avons donc introduit une étape de propagation de la molécule donneuse chez *E. coli* dans le but de l'amplifier et de récupérer par midi-préparation un matériel pur et en quantité suffisante pour l'étape 4 du RAGE



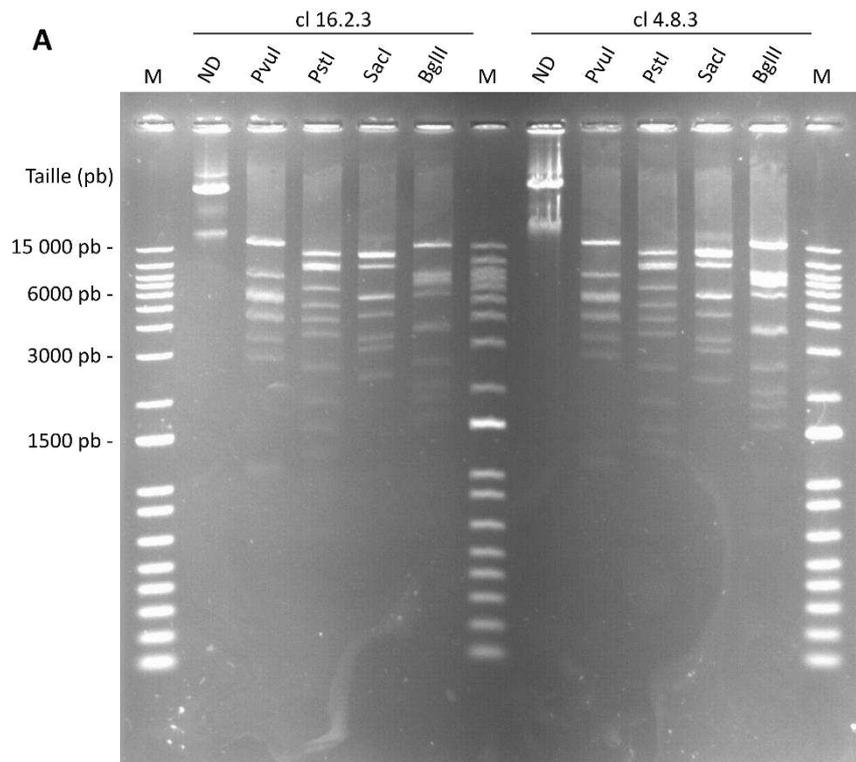
Pour réaliser cette étape, nous avons choisi de travailler avec le plasmide CopyControl pCC1 (LGC, Biosearch Technologies) et sa souche hôte *E. coli* EPI300, qui permettent de propager des inserts de grande taille, et/ou potentiellement instables. Le principe est le suivant :

- Le plasmide pCC1 possède deux origines de réplifications. La première (réplicon du facteur F) permet son maintien en simple copie, et la seconde (oriV) permet sa réplification à un nombre élevé de copies.
- La souche hôte EPI300 contient le gène *trfA* dont l'expression, nécessaire à l'initiation de la réplification à partir de l'origine de réplification oriV du pCC1, est sous le contrôle d'un promoteur inductible.

Au cours de la propagation des cellules en milieu LB additionné d'antibiotique, le plasmide, contrôlé par le réplicon du facteur F, se maintient à une copie par cellule, ce qui garantit sa stabilité. L'ajout d'une solution d'induction, peu de temps avant l'extraction, induit l'expression de *trfA*, qui initie la réplification du plasmide à partir de l'oriV et augmente le nombre de copies du plasmide de 10 à 20 par cellule. L'ADN plasmidique présent en plus grande quantité est extrait par midi-préparation puis vérifié, dosé et stocké.

En prévision de cette étape, les séquences correspondantes au plasmide CopyControl pCC1 avaient été pré-intégrées au pansement de réparation utilisé à l'étape 2 du RAGE pour modifier le fragment 7.2 maintenu sous forme de plasmide centromérique dans la levure (cf partie 2.2 et figure 81). L'ADN des clones levure VL6-48 [MGP192 frgt7.2 cl 4.1]Cas9 PST\_pCC1-YLYS-CRE-SpecR cl **4.8** et VL6-48 [MGP192 frgt7.2 cl 4.3]Cas9 PST\_pCC1-YLYS-CRE-SpecR cl **16.2** pouvait donc être extrait et utilisé pour transformer la souche *E. coli* EPI300 par électroporation. Pour chaque clone levure, une vingtaine de colonies est apparue sur milieu de sélection LB contenant 100 µg/ml de spectinomycine. Dix colonies (par condition) ont été piquées et analysées pour la présence du plasmide pCC1-LYS-cre-specR frgtBsu 7.2 d'environ 58 kb.

A cette étape, les plasmides sont maintenus en simple copie, nous avons donc criblé les transformants *E. coli* par PCR. La première PCR a été réalisée à l'aide d'un couple d'amorces s'hybridant au niveau d'une des jonctions entre le fragment 7.2 et le pansement (= PCR 2 décrit dans la section précédente, taille attendue : 1566 pb). La deuxième PCR a été réalisée à l'aide d'un couple d'amorces s'hybridant au niveau du gène *specR* (taille attendue: 1312 pb).



**B**

PvuI	PstI	SacI	BglII
390	150	8	667
1109	345	410	1593
1166	643	2343	1863
2812	1270	3017	2067
3343	1541	3359	2528
4140	1924	4382	3512
4312	2572	5560	3632
5230	3532	5562	5554
5441	4148	9412	6584
5650	4932	11760	7044
7783	6311	12376	7744
16831	8920		15399
	9534		
	12367		

**Figure 83 : Digestions des molécules donneuses purifiées et issues des deux clones *E. coli* sélectionnés.** Les digestions ont été réalisées avec 4 enzymes de restriction différentes : PvuI, PstI, SacI, BglII. **[A]** Profil des molécules donneuses sur gel d'agarose coloré au BET, après digestion par les enzymes PvuI, PstI, SacI, BglII. ND : Non digéré. « M » : DNA Ladder 1 kb+ Invitrogen (100-12000 pb). **[B]** Tableau récapitulatif des tailles (en pb) des fragments de digestion attendus avec les 4 enzymes utilisées.

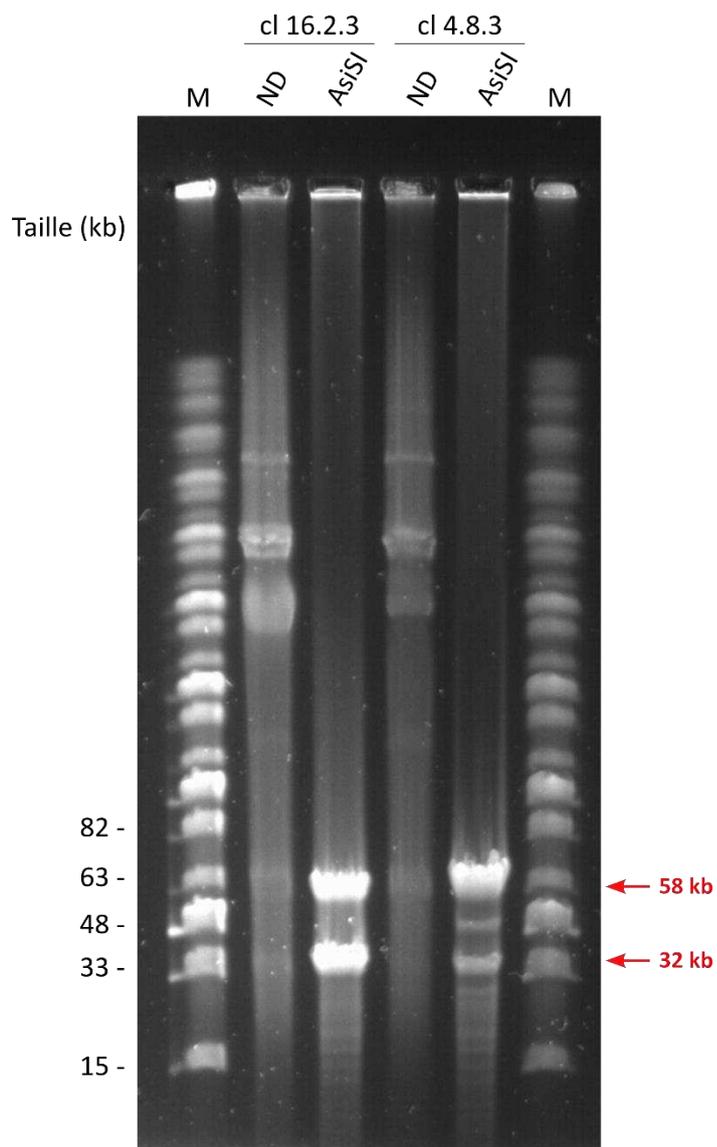
Sur les 20 clones analysés, 15 ont montré des amplicons aux tailles attendues pour les deux PCR. Deux clones ont été sélectionnés, un pour chaque condition :

- *E. coli* EPI300 pCC1-LYS-cre-specR frgtBsu 7.2 cl **4.8.3** provenant de la condition de transformation avec l'ADN de la souche VL6-48 [Bsu MGP192 frgt7.2 cl 4.1]Cas9 PST\_pCC1-YLYS-CRE-SpecR cl 4.8
- *E. coli* EPI300 pCC1-LYS-cre-specR frgtBsu 7.2 cl **16.2.3** provenant de la condition de transformation avec l'ADN de la souche VL6-48 [Bsu MGP192 frgt7.2 cl 4.3]Cas9 PST\_pCC1-YLYS-CRE-SpecR cl 16.2

Après ajout de la solution d'induction et amplification du nombre de copies via le système CopyControl, le fragment a été extrait par midi-préparation à partir des deux transformants sélectionnés. Sa présence et son intégrité ont été contrôlées par restriction avec 4 enzymes différentes (PvuI, PstI, SacI, BglII). Le plasmide pCC1-LYS-cre-specR frgtBsu 7.2 issu des deux clones a présenté le profil attendu pour les 4 digestions (figure 83). A partir de cette étape, le plasmide isolé par midi-préparation d'une taille d'environ 58 kb correspondait à la molécule donneuse pour la mise en place de la méthode RAGE (figure 80).

Pour s'assurer de l'absence de mutation sur la totalité du plasmide, notamment au niveau des sites lox, et des gènes *cre* et *specR*, la molécule donneuse issue du clone *E. coli* EPI300 pCC1-LYS-cre-specR frgtBsu 7.2 cl **4.8.3** a été séquencée complètement par la société prestataire PLASMIDSAURUS. L'analyse des séquences a montré que la molécule donneuse était présente en deux versions. L'une d'entre elles avait la taille attendue de 58 kb et l'autre semblait plus réduite avec une taille de 32 kb.

Pour contrôler la présence de ces deux molécules dans nos tubes stocks, une analyse par PFGE a été effectuée. Les molécules donneuses issues des deux clones *E. coli* ont été digérées avec l'enzyme de restriction AsiSI qui les linéarise. Après migration, elles ont montré toutes les deux, un profil à deux bandes d'environ 32 kb et 58 kb (figure 84). Ces résultats ont montré que les deux versions de la molécule donneuse étaient bien présentes chez les deux clones *E. coli* sélectionnés. Après une analyse des gènes présents et absents sur la molécule de 32 kb, nous nous sommes rendus compte que cette dernière était dépourvue de plusieurs gènes essentiels, comme par exemple *ssbA* et *DnaC* (Kobayashi *et al.*, 2003). Nous avons conclu que la version de 32 kb ne permettrait pas de générer des clones viables de *B. subtilis* en cas



**Figure 84 : Analyse par gel d'électrophorèse en champ pulsé (PFGE) des molécules donneuses purifiées et issues des deux clones *E. coli* sélectionnés.** La digestion par l'enzyme AsiSI linéarise le fragment. La présence des deux versions du fragment sont visibles avec deux bandes aux tailles de 32 kb et 58 kb. « M »: Mid range PFG ladder NEB (15.0 - 242.5 kbp). ND : Contrôle non digéré.

d'échange de fragments à l'étape 4. Etant donné que la molécule donneuse désirée d'une taille d'environ 58 kb était aussi présente et sans mutation, chez les deux clones *E. coli* sélectionnés, nous avons décidé, pour aller plus vite, de les valider et de poursuivre nos expériences avec ces deux préparations.

Dès que nous pourrons, cette étape sera refaite afin de tenter d'obtenir une préparation pure de la molécule donneuse et retenter l'étape 4 du RAGE dans de meilleures conditions.

#### 4. Etape 4 : Echange de fragments d'ADN chez *B. subtilis*

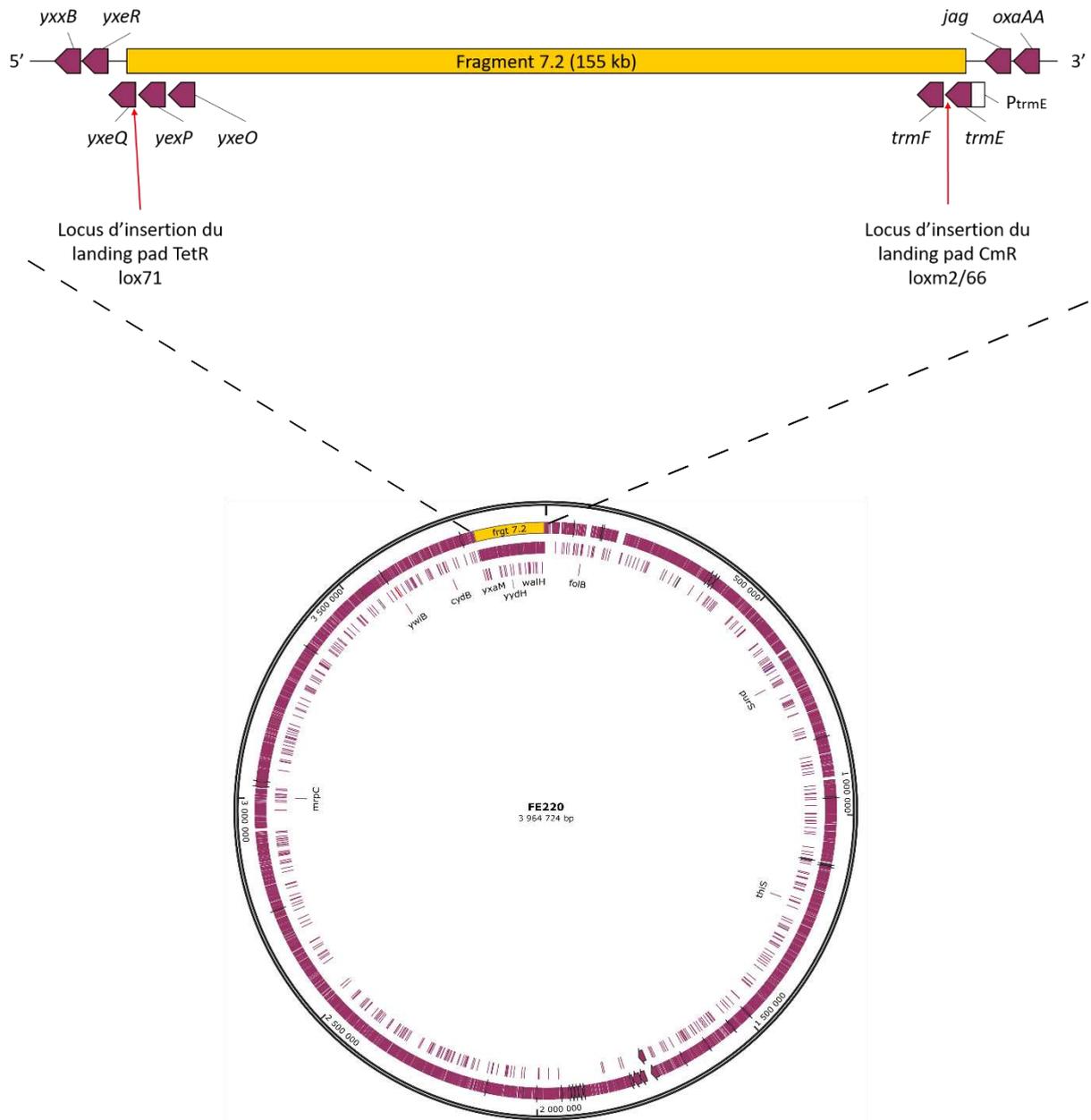
La dernière étape de la méthode RAGE consistait à transférer la molécule donneuse isolée à partir d'*E. coli*, chez *B. subtilis*, et à provoquer un échange de fragments avec la région chromosomique correspondante, grâce au système de recombinaison site-spécifique cre/lox. Pour que la stratégie fonctionne et que l'échange ait lieu, le génome de la cellule receveuse doit contenir (au même titre que la molécule donneuse) des sites lox de part et d'autre de la région à éditer, et devait donc être modifiée.

Notre choix s'est porté sur la souche *B. subtilis* FE220 (3,96 Mb) qui diffère de la souche MS par l'absence du système de restriction-modification BsuM. Les systèmes de restriction-modification sont des systèmes de défense qui peuvent, dans certaines circonstances, dégrader l'ADN exogène après son entrée dans la cellule. Nous avons préféré ne pas prendre ce risque.

##### 4.1. Construction de la cellule receveuse à partir de la souche *B. subtilis* FE220

Dans la molécule donneuse, le fragment 7.2 de 44 kb que l'on souhaite échanger est bordé de deux sites lox incompatibles entre eux : lox66 et loxm2/71. Pour remplacer la région 7.2 du génome de *B. subtilis* FE220 (155 kb) par celui de la molécule donneuse, la région 7.2 du génome de *B. subtilis* FE220 doit également être encadrée par deux sites lox incompatibles entre eux mais compatibles avec ceux de la molécule donneuse. Nous avons sélectionné les sites lox71 (compatible avec lox66) et loxm2/66 (compatible avec loxm2/71). La séquence et la compatibilité de ces sites sont détaillées dans la partie 4.2.2 et tableau 4 de l'introduction.

L'insertion des sites lox, de part et d'autre de la région 7.2 du génome de *B. subtilis* FE220 ne s'est pas faite simultanément. Elle a nécessité deux étapes de transformation et



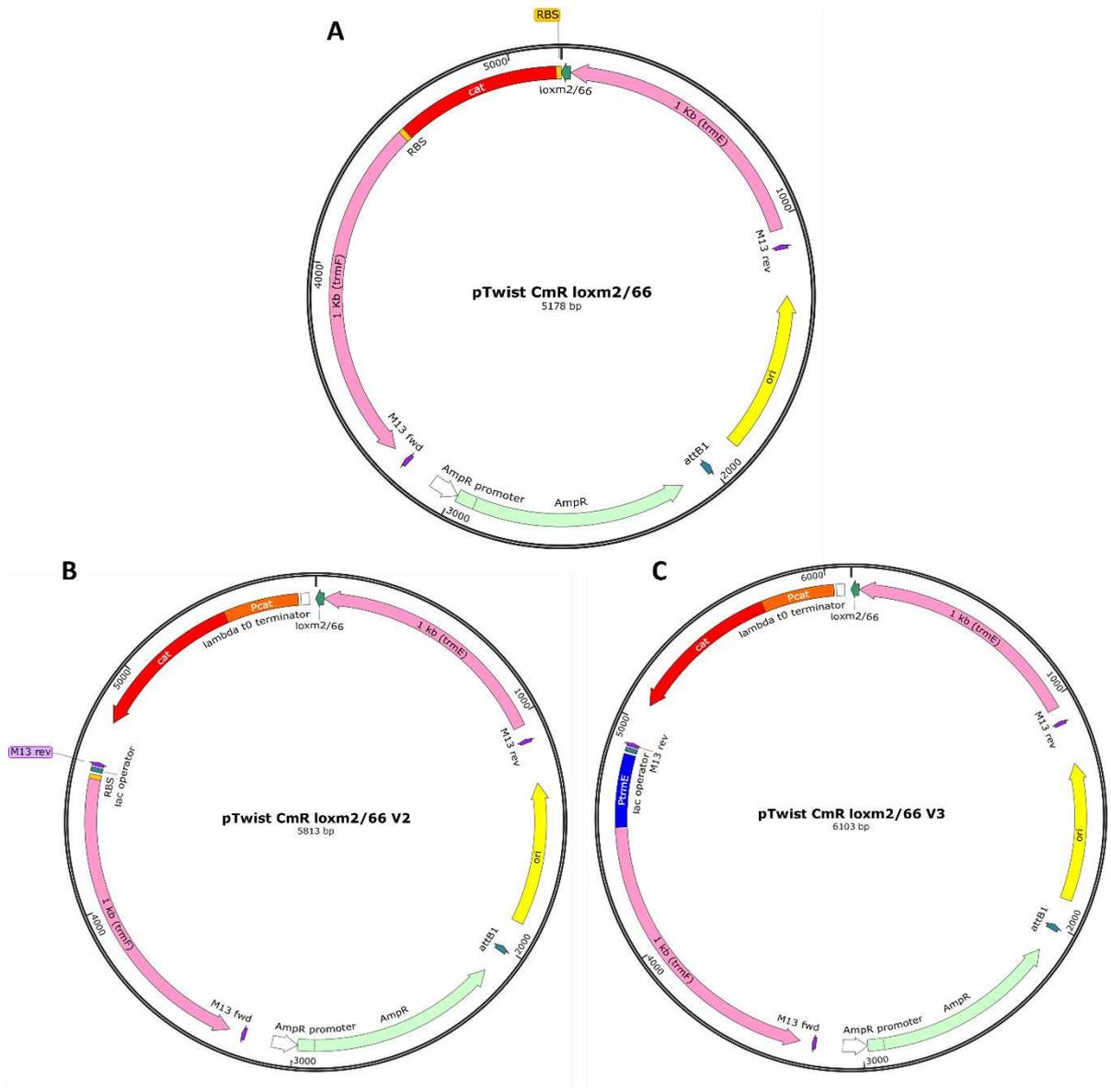
**Figure 85 : Fragment 7.2 de la souche FE220 choisie pour être la cellule receveuse. Les *loci* pour l'insertion des landings pad TetR lox71 et CmR loxm2/66 sont représentés.**

l'introduction concomitante de marqueurs de sélection. Au cours de la première transformation, le site *loxm2/66* a été introduit avec le marqueur de résistance au chloramphénicol (gène *cat*), créant le « landing pad » CmR *loxm2/66* sur la borne 3' du segment de 155 kb. Au cours de la deuxième transformation, le site *lox71* a été introduit avec le marqueur de résistance à la tétracycline (gène *tetR*), créant le « landing pad » TetR *lox71* sur la borne 5' du segment de 155 kb (figure 85). Ces expériences sont détaillées ci-dessous.

#### 4.1.1. Ajout du premier landing pad « CmR *loxm2/66* » dans le génome de *B. subtilis* FE220 (borne 3')

Pour introduire le premier « landing pad » dans le génome de *B. subtilis* FE220, nous avons commandé chez Twist Bioscience le plasmide « pTwist CmR *loxm2/66* » portant le site *loxm2/66* et le gène *cat*, bordés de part et d'autre par des séquences d'environ 1 kb correspondant, en amont, à la fin du gène *trmF* et, en aval, au début du gène *trmE* (figure 86A). Après réception, le plasmide a été amplifié, stocké puis utilisé pour isoler, par digestion enzymatique, la cassette CmR *loxm2/66* et ses bras d'homologie. La souche *B. subtilis* FE220 a ensuite été transformée (protocole MC) avec le plasmide circulaire ou avec la cassette isolée purifiée. Un contrôle positif de transformation a également été réalisé avec le pDR111, un plasmide intégratif portant la résistance à la spectinomycine. Les transformants ont été sélectionnés sur milieu LB solide contenant 5 µg/ml de chloramphénicol pour les conditions d'insertion du « landing pad » ou contenant 100 µg/ml de spectinomycine pour le contrôle positif. Après 24 h d'incubation à 37°C, de nombreuses colonies sont apparues sur milieu de sélection pour le contrôle positif, mais aucune colonie n'était visible pour les conditions de l'insertion de la cassette CmR *loxm2/66* (plasmide ou cassette isolée).

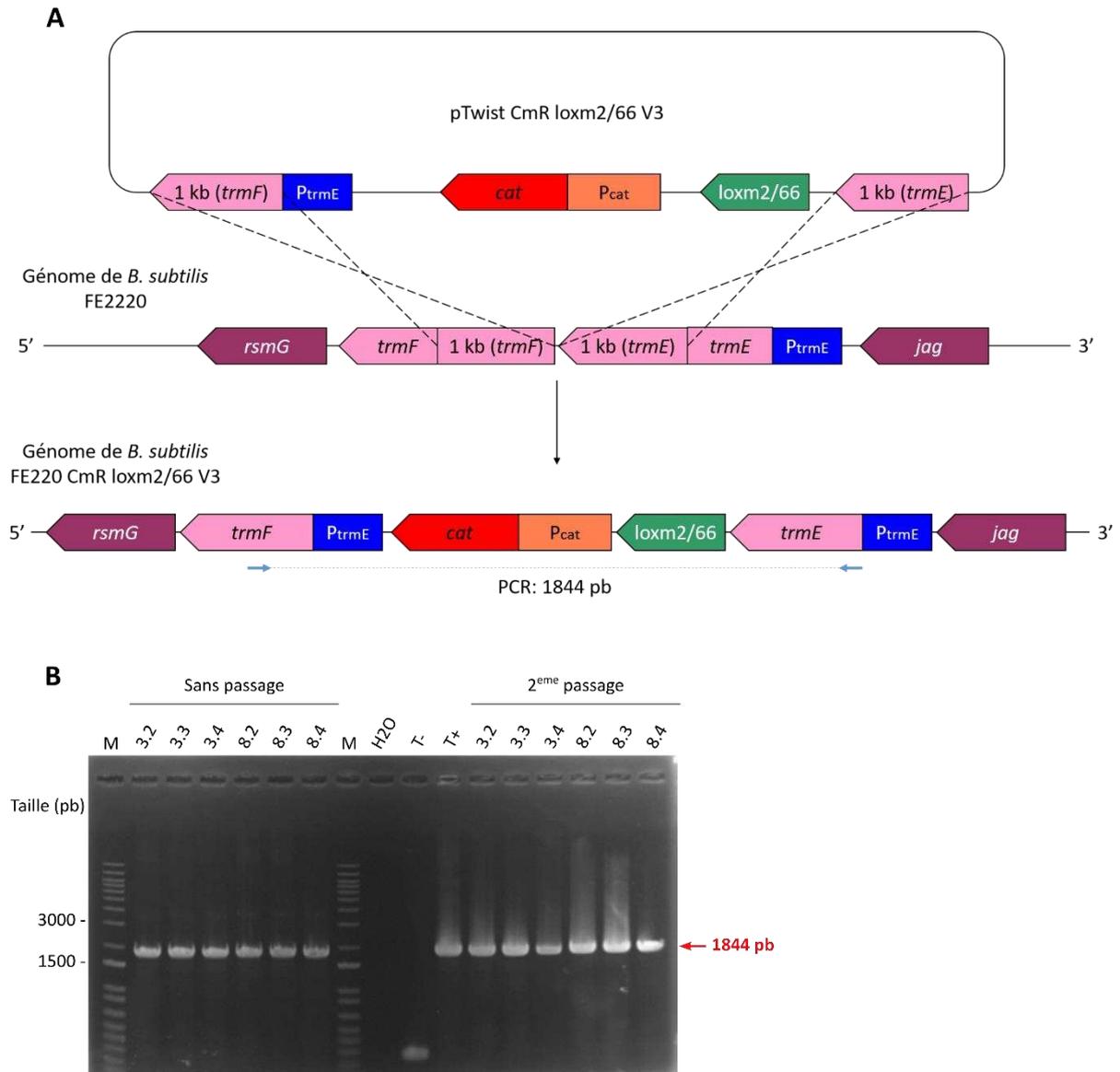
Nous avons réitéré l'expérience quatre fois en changeant certains paramètres : diminution de la concentration en chloramphénicol sur les boîtes à 4 µg/ml, ajout de glucose 20% et/ou chloramphénicol durant l'étape de mise en contact de l'ADN avec les cellules et même utilisation de la souche MS à la place de la FE220. Pour tous ces essais, de nombreuses colonies sont apparues sur milieu de sélection pour le contrôle positif pDR111, indiquant que le protocole de transformation fonctionnait. Malheureusement aucun transformant n'est apparu pour les conditions d'insertion du premier « landing pad ».



**Figure 86 : Carte des différentes version du pTwist CmR loxm2/66. [A] pTwist CmR loxm2/66 V1. [B] pTwist CmR loxm2/66 V2 avec le promoteur Pcat en plus. [C] pTwist CmR loxm2/66 V3 avec les promoteurs Pcat et PtrmE en plus.**

Après ces échecs répétés, nous sommes revenus sur le design du plasmide « pTwist CmR loxm2/66 » qui avait été fait en collaboration avec l'équipe SyBER (Institut MICALIS, UMR 1319 INRAE / AgroParisTech, Jouy-en-Josas). Dans ce 1<sup>er</sup> design, le gène conférant la résistance au chloramphénicol n'était pas sous le contrôle de sa séquence promotrice originale mais seulement d'une séquence de 18 pb contenant un RBS. Cette décision avait été prise pour les raisons suivantes. (i) des données de transcriptomique suggèrent que le promoteur *trmE* contrôle l'expression du gène *trmE* et du gène en aval *trmF* en condition normale de croissance (structure opéronique), (ii) suite à l'intégration de la cassette au site d'insertion, le gène *cat* se retrouve juste en aval du gène *trmE* et du promoteur *trmE*. Nous pensions donc que le gène *cat* pourrait être exprimé (suite à son intégration) à partir du promoteur *trmE*, et que seul un RBS suffirait. Les résultats obtenus semblaient contredire les hypothèses émises. Le gène *cat* n'était apparemment pas (ou trop peu) exprimé à partir de ce promoteur, et ne conférait pas la résistance souhaitée au chloramphénicol.

Nous avons donc décidé de produire deux nouvelles versions du plasmide pTwist CmR loxm2/66 : une version 2 (V2) dans laquelle nous avons ajouté en amont du gène *cat* son promoteur natif, générant le plasmide pTwist CmR loxm2/66 V2 (figure 86B), et une version 3 (V3) dans laquelle nous avons ajouté en amont du gène *cat* mais aussi du gène *trmF* (qui possédait seulement un RBS dans les versions V1 et V2), les promoteurs natifs de ces gènes, générant le plasmide pTwist CmR loxm2/66 V3 (figure 86C). Deux nouvelles expériences de transformation ont été tentées. Pour la première, le plasmide complet circulaire ou la cassette isolée de la version V2 ont été transformés dans la souche *B. subtilis* FE220 avec le protocole simple MC. Les transformants ont été sélectionnés sur milieu LB solide contenant 4 µg/ml de chloramphénicol. Après 24 h d'incubation à 37°C, plusieurs centaines de colonies sont apparues sur les essais aussi bien avec le plasmide complet circulaire qu'avec la cassette isolée. La présence du gène *cat* a été contrôlée par une première PCR produisant un amplicon à la taille de 1295 pb. Sur 20 clones analysés, 13 ont montré le profil attendu. Chez ces 13 clones, la présence du landing pad CmR loxm2/66 V2 a été vérifiée avec une seconde PCR produisant un amplicon à la taille de 1554 pb si la cassette est bien insérée, et 265 pb en absence d'insertion. Tous les clones ont montré un profil à deux bandes (1554 pb et 265 pb) indiquant un mélange de populations. Nous avons donc fait un sous clonage en réalisant des stries sur milieu LB solide contenant 5 µg/ml de chloramphénicol. Une colonie isolée issue de chacune des stries a été analysée avec la seconde PCR. Dix sous clones ont montré le profil



**Figure 87 : Insertion du « landing pad » CmR loxm2/66 V3 dans le génome de la souche FE220. [A]** Schéma de l'insertion de la cassette CmR loxm2/66 dans le génome de la souche FE220 à partir du plasmide pTwist CmR loxm2/66. **[B]** L'insertion de la cassette est validée par une PCR simple avec les amorces Verif receveuse F et Verif receveuse R produisant un amplicon à la taille de 1844 pb. « M » : DNA Ladder 1 kb+ Invitrogen (100-12000 pb). « H<sub>2</sub>O » : contrôle négatif sans ADN. T-: témoin négatif ADNg de la souche *B. subtilis* FE220. T+: contrôle positif plasmide pTwist CmR loxm2/66 V3.

attendu pour l'insertion du landing pad CmR loxm2/66 V2 avec une seule bande à la taille de 1554 pb.

Pour la deuxième expérience, le plasmide complet circulaire ou la cassette isolée de la version V3 ont été transformés dans la souche *B. subtilis* FE220 avec le protocole simple MC. Les transformants ont été sélectionnés de la même manière que pour l'expérience précédente. Après 24 h d'incubation à 37°C, un total de 399 colonies est apparu sur boîte, dont 30 ont été analysées par PCR. La présence du landing pad CmR loxm2/66 V3 a été vérifiée avec une PCR produisant un amplicon à la taille de 1844 pb si la cassette est bien insérée (figure 87A), et 265 pb en absence d'insertion. Sur les 30 clones analysés, 23 ont montré un profil à deux bandes (1844 pb et 265 pb), indiquant un mélange de populations, et 6 ont montré une seule bande à la taille attendue pour l'insertion de la cassette. Sur ces 6 clones, un sous clonage a tout de même été réalisé de manière identique à l'expérience précédente. Les sous clones ont été ré-analysés et montraient à nouveau une seule bande à 1844 pb, correspondant à l'insertion de la cassette (figure 87B).

Ces deux expériences nous ont permis d'isoler des transformants de la souche *B. subtilis* FE220 dans laquelle le premier « landing pad » était intégré, avec deux versions différentes aboutissant à deux souches différentes. Nous les avons appelées : FE220 CmR loxm2/66 **V2** et FE220 CmR loxm2/66 **V3**.

#### 4.1.2. Ajout du second landing pad « TetR lox71 » dans le génome de *B. subtilis* FE220 CmR loxm2/66 et V3 (borne 5')

Pour introduire le second « landing pad » dans le génome de FE220 CmR loxm2/66 **V2** et **V3**, nous avons commandé chez Twist Bioscience le plasmide « pTwist TetR lox71 » portant le site lox71 et le gène *tetR* conférant la résistance à la tétracycline, bordés de part et d'autre par des séquences d'environ 1 kb correspondant, en amont, à la fin du gène *yxeQ* et, en aval, au gène *yxeP* (figure 88).

Après réception, le plasmide a été amplifié, stocké, puis utilisé pour isoler par digestion enzymatique la cassette TetR lox71 et ses bras d'homologie. Les souches FE220 CmR loxm2/66

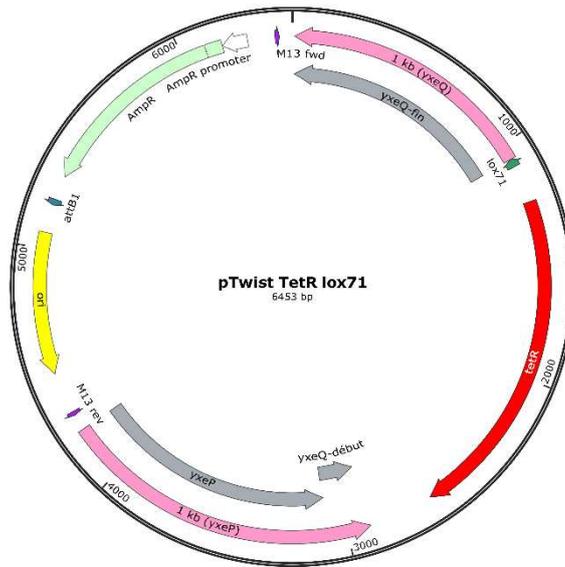


Figure 88 : Carte du plasmide pTwist TetR lox71

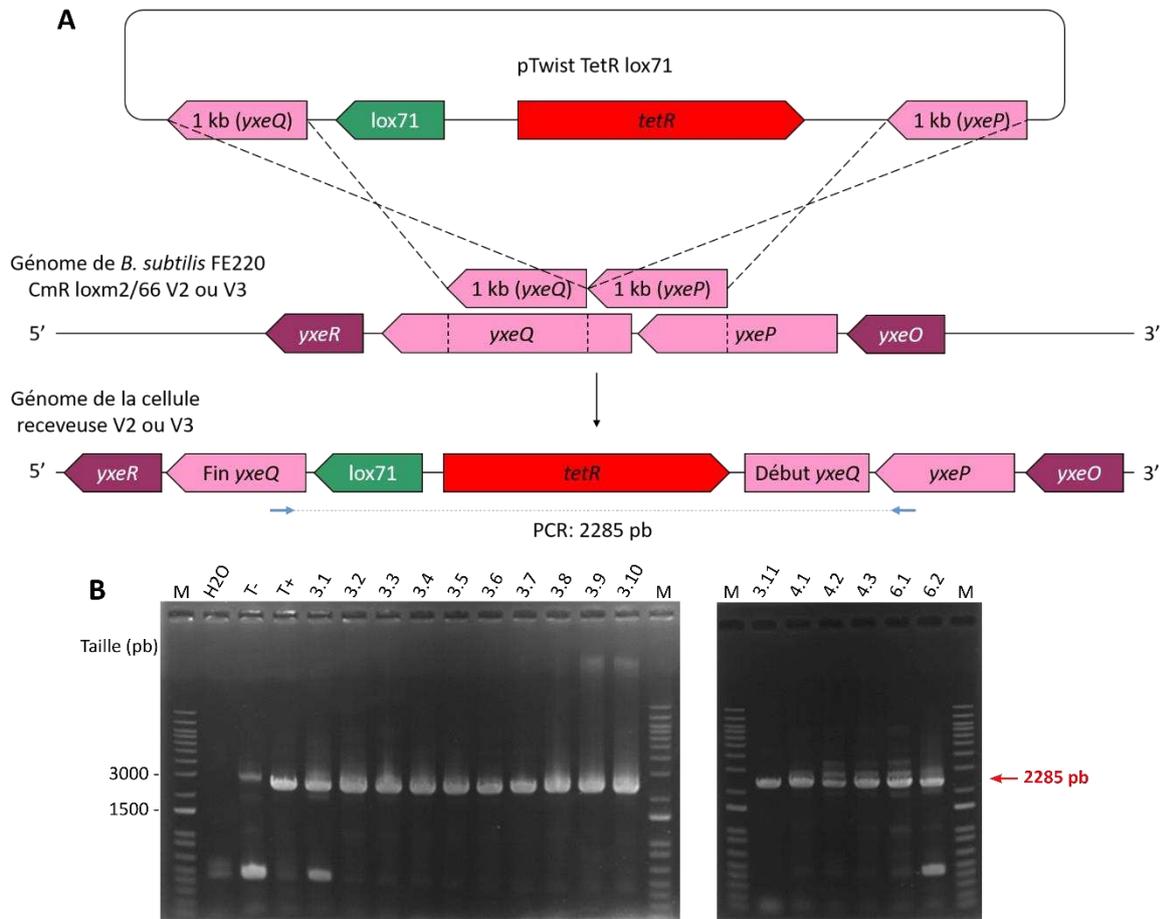


Figure 89 : Insertion du « landing pad » TetR lox71 pour obtenir la cellule receveuse V3. [A] Schéma de l'insertion de la cassette TetR lox71 à partir du plasmide pTwist TetR lox71. [B] L'insertion de la cassette est validée par une PCR simple avec les amorces FE220 receveuse YxeQ F et FE220 receveuse YxeP R produisant un amplicon à la taille de 2285 pb. « M » : DNA Ladder 1 kb+ Invitrogen (100-12000 pb). « H2O » : contrôle négatif sans ADN. T-: témoin négatif ADNg de la souche *B. subtilis* FE220. T+: contrôle positif plasmide pTwist TetR lox71.

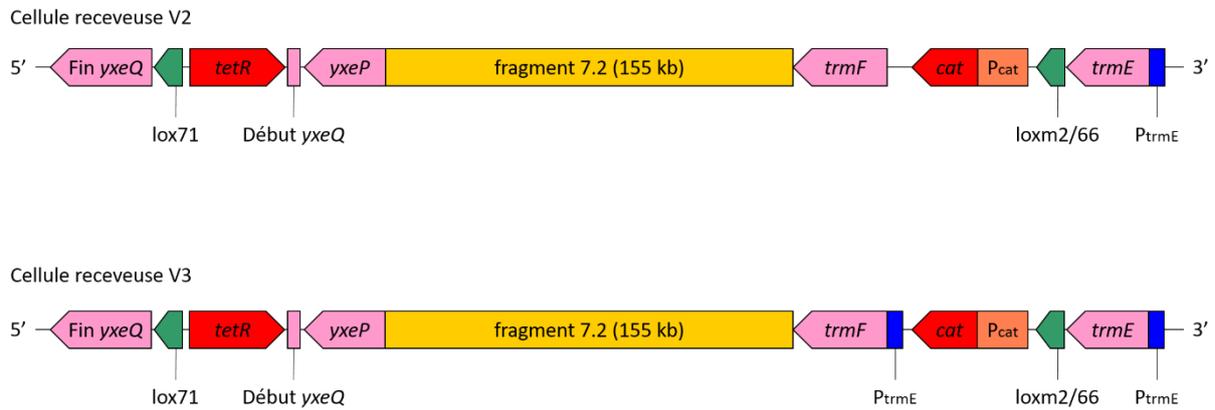
V2 et V3 ont ensuite été transformées (protocole MC) avec le plasmide circulaire ou avec la cassette isolée purifiée (au cours de deux expériences indépendantes).

Lors de la première expérience (réalisée avec la souche FE220 CmR loxm2/66 **V2**), les transformants ont été sélectionnés sur milieu LB solide contenant 7,5 µg/ml de tétracycline. Après 48 h d'incubation à 37°C, 8 colonies sont apparues sur milieu de sélection. La présence du landing pad TetR lox71 a été validée avec une PCR produisant un amplicon à la taille de 2285 pb en cas d'insertion de la cassette, et 451 pb en absence d'insertion. Sur les 8 clones analysés, 2 ont montré un profil à 2 bandes (2285 pb et 451 pb) indiquant un mélange de populations, et 6 autres ont montré un profil à une seule bande à la taille attendue, de 2285 pb, correspondant à l'insertion de la cassette. Chez ces 6 derniers, la présence du premier « landing pad » (CmR loxm2/66 V2) a aussi été contrôlée et validée par PCR. Puis 3 d'entre eux ont été séquencés au niveau des deux « landing pads » afin de s'assurer de l'absence de mutation sur les sites lox. Finalement deux d'entre eux ont été validés.

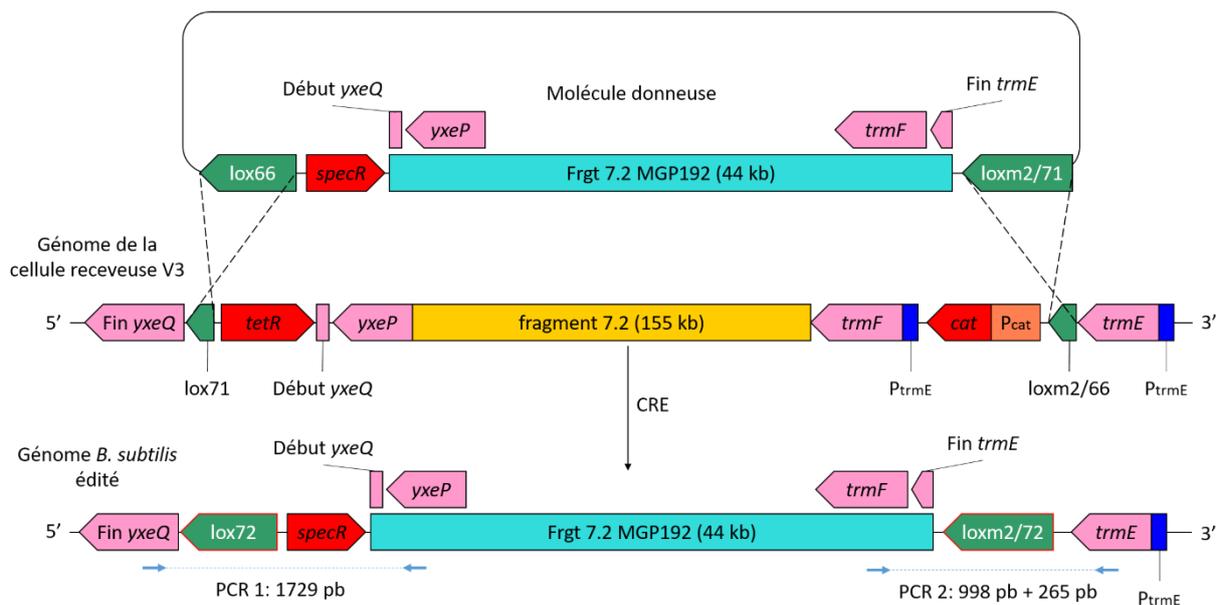
Cette expérience nous a permis d'obtenir une première cellule receveuse FE220 dans laquelle les « landing pads », CmR loxm2/66 **V2** et TetR lox71, étaient insérés de part et d'autre de la région chromosomique 7.2 à éditer.

Lors de la deuxième expérience (réalisée avec la souche FE220 CmR loxm2/66 **V3**), 16 colonies sont apparues sur milieu LB contenant 7,5 µg/ml de tétracycline. La présence du landing pad TetR lox71 a été contrôlée dans les 16 clones comme décrit précédemment. Sur les 16 clones, deux ont montré un profil à 2 bandes (2285 pb et 451 pb) et 14 ont montré un profil à une bande à la taille attendue, de 2285 pb, correspondant à l'insertion de la cassette (figure 89). La présence du premier « landing pad » (CmR loxm2/66 V3) a ensuite été contrôlée. Tous les clones analysés présentaient la bande attendue à 1844 pb, correspondant à la présence du premier « landing pad ». Afin de s'assurer de l'absence de mutation au niveau des deux sites lox, un séquençage Sanger a été réalisé sur 12 de ces clones. Nous avons pu en valider 3.

Cette expérience nous a permis d'obtenir une deuxième cellule receveuse FE220 dans laquelle les « landing pads », CmR loxm2/66 **V3** et TetR lox71, étaient insérés de part et d'autre de la région chromosomique 7.2 à éditer.



**Figure 90 : Représentation schématique de la région chromosomique correspondant au fragment 7.2 chez les deux cellules receveuses obtenues. La version 3 contient en amont du gène *trmF* le promoteur du gène *trmE*.**



**Figure 91 : Schéma de l'évènement d'échange du fragment 7.2 dans la cellule receveuse FE220 V3. L'activité recombinase de CRE provoque un échange entre les régions bordées de sites lox du génome de la cellule receveuse et la molécule donneuse. Cet échange est couplé du côté 5' à l'incorporation du marqueur de résistance à la spectinomycine (*specR*) et à l'élimination du gène *tetR* et côté 3' à l'élimination du gène *cat* et du P<sub>trmE</sub> en amont du *trmF* ainsi que l'incorporation de la fin du gène *trmE*.**

L'ensemble de ces expériences nous a permis d'obtenir deux souches receveuses pour l'étape 4 du RAGE (figure 90). Le gène *trmF* étant finement régulé, nous avons préféré poursuivre notre travail avec la souche V3 dans cette souche, une séquence promotrice a été ajoutée, non seulement en amont du gène *cat*, mais également en amont du gène *trmF*, pour garantir leur expression (figure 90). Cette dernière a été nommée la **cellule receveuse FE220 V3**.

Pour la suite de nos analyses et de nos expériences, nous avons sélectionné le clone 3.4 de cette souche. Le génome de ce clone a été séquencé complètement par la société prestataire PLASMIDSAURUS. L'analyse des séquences a montré que les sites lox étaient bien présents, sans mutation et qu'il n'y avait pas de remaniements dans le reste du génome. Ainsi, la **cellule receveuse FE220 V3 cl 3.4** était validée pour la suite de nos expériences.

#### 4.2. Première tentative d'échange de fragments via le système cre/lox

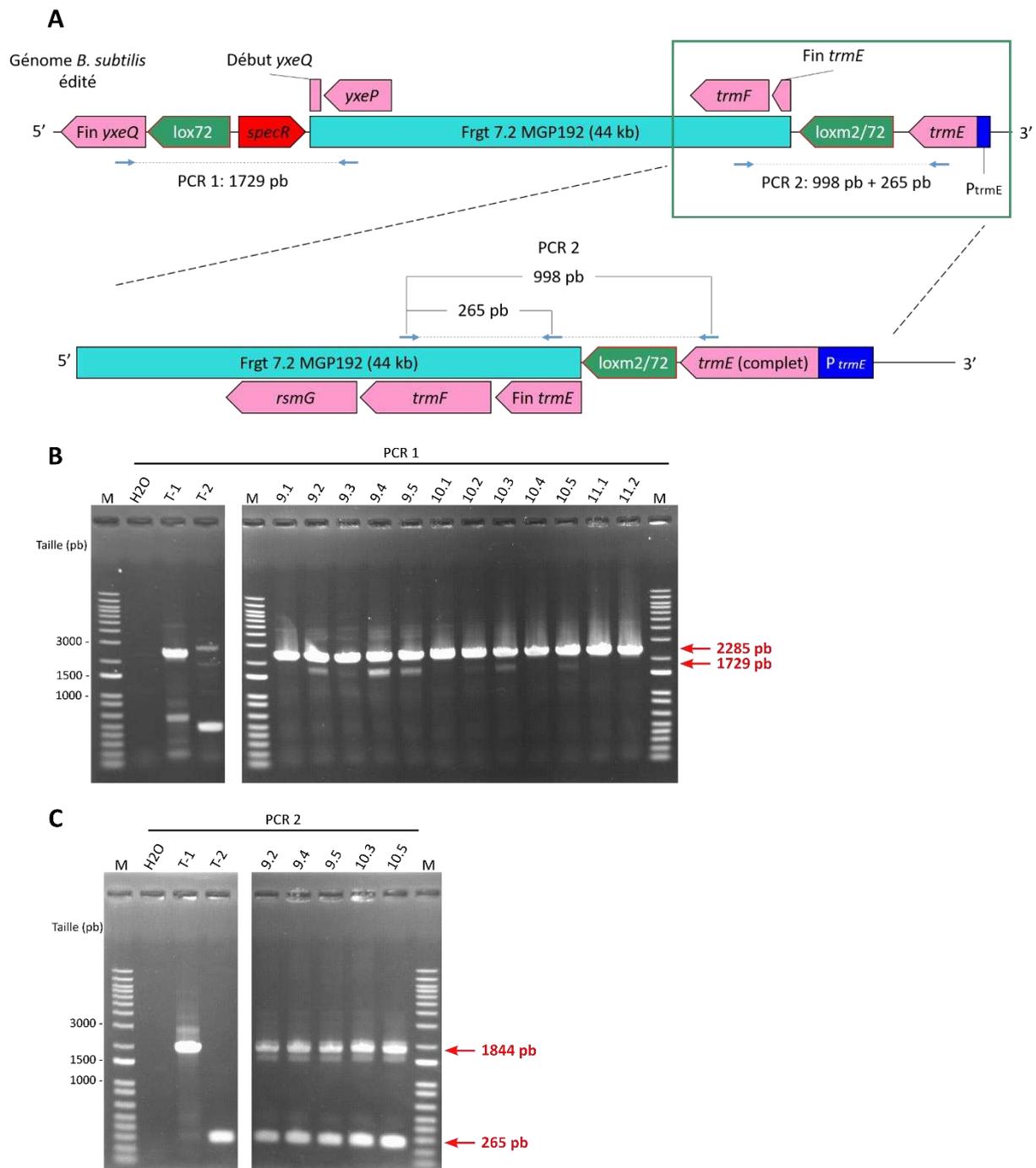
Dans les sections précédentes, nous avons réussi à obtenir deux molécules donneuses issues des clones :

- *E. coli* EPI300 pCC1-LYS-cre-specR frgtBsu 7.2 cl **4.8.3**
- *E. coli* EPI300 pCC1-LYS-cre-specR frgtBsu 7.2 cl **16.2.3**

Nous avons également obtenu la **cellule receveuse FE220 V3 cl 3.4** contenant les sites lox71 et loxm2/66 de part et d'autre de la région chromosomique à échanger.

La dernière étape de la méthode RAGE consiste à introduire la molécule donneuse dans la cellule receveuse et à provoquer un échange de fragments avec la région chromosomique correspondante, via le système de recombinaison site-spécifique cre/lox (figure 91).

A ce stade, notre capacité à faire rentrer une molécule d'ADN de grande taille dans *B. subtilis* pouvait être questionnée : est-il possible d'introduire une molécule d'ADN circulaire de 58 kb dans *B. subtilis* par compétence naturelle ? N'ayant aucune information sur le sujet, nous avons d'abord pensé à recourir à des protocoles de transformation artificielle comme des protocoles de transformation par électroporation, ou bien des protocoles de transformation médiée par le PEG, un agent fusogène connu pour permettre l'entrée de molécule d'ADN de taille supérieure à 1 Mb chez certains organismes (Lartigue *et al.*, 2007;



**Figure 92 : Analyse des transformants issus de la première tentative d'échange de fragment de génome via le système *cre/lox* chez *B. subtilis*.** [A] Schéma détaillé de la région au niveau de la PCR 2 après échange de fragment. [B] L'échange de fragment est contrôlé par deux PCR au niveau des jonctions entre le génome et le fragment issu de la molécule donneuse. En cas d'échange de fragment, la PCR 1 avec les amorces FE220 receveuse YxeQ F et FE220 receveuse YxeP R produit un amplicon de 1729 pb. [C] La PCR 2 avec les amorces Verif receveuse F et Verif receveuse R produit deux amplicons aux tailles de 998 pb et 265 pb. « M » : DNA Ladder 1 kb+ Invitrogen (100-12000 pb). « H2O » : contrôle négatif sans ADN. T-1: témoin négatif plasmide pTwist TetR lox71 pour la PCR 1 et le plasmide pTwist CmR loxm2/66 pour la PCR 2. T-2: contrôle négatif avec l'ADNg de la souche *B. subtilis* FE220.

Lartigue *et al.*, 2009; Labroussaa *et al.*, 2016). Cependant l'utilisation de tels protocoles nécessitait des mises au point et/ou des optimisations que nous estimions chronophages dans la cadre de ma thèse.

Ainsi, dans le but d'entamer le plus rapidement possible les essais d'échange de fragments dans la cellule receveuse, nous avons décidé de tester d'introduire la molécule donneuse de 58 kb dans la cellule receveuse en utilisant la compétence naturelle de *B. subtilis*. Pour ce faire, nous avons utilisé le protocole de transformation simple MC. Lors de ces expériences, deux contrôles ont été systématiquement intégrés, un contrôle négatif sans ADN et un contrôle positif avec le plasmide intégratif pDR111 qui porte un gène de résistance à la spectinomycine. Lors du premier essai, la cellule receveuse FE220 V3 cl 3.4 a été transformée avec les molécules donneuses purifiées issues des deux clones *E. coli* ou directement avec l'ADNg extrait des levures avant l'étape 3 d'amplification. Les transformants ont été sélectionnés sur milieu LB solide contenant 100 µg/ml de spectinomycine. Pour le contrôle négatif, aucune colonie n'est apparue sur boîte, alors que pour le contrôle positif, 171 colonies étaient visibles, validant ainsi l'expérience. Pour les conditions de transformation avec les molécules donneuses issues des clones *E. coli*, plus d'un millier de colonies ont été obtenues et 10 ont été analysées. Pour les conditions de transformation avec l'ADNg des levures, seulement 2 colonies ont été obtenues et analysées.

Afin de confirmer l'échange de fragments, deux PCR simplex ont été réalisées au niveau des bornes du fragment 7.2. La PCR 1 a permis de contrôler l'incorporation du gène de résistance pour la spectinomycine (*specR*) et la perte du gène de résistance à la tétracycline (*tetR*) après recombinaison entre les sites le *lox66* de la molécule donneuse et le *lox71* de la cellule receveuse (borne 5'). Cette PCR peut produire deux amplicons différents, le premier à la taille de 1729 pb si la recombinaison site-spécifique a eu lieu, et le second à la taille de 2285 pb en absence de cet évènement de recombinaison (figure 92A). Sur les 12 clones analysés, 5 ont montré un profil à 2 bandes aux tailles de 2285 pb et 1729 pb, indiquant que les clones n'étaient pas purs. Ces 5 clones provenaient tous des conditions de transformation avec les molécules donneuses issues d'un des deux clones *E. coli*. Les 7 autres n'ont présenté que la bande à 2285 pb (figure 92B). Cette première PCR a montré que la transformation avec l'ADNg extrait des levures hébergeant la molécule donneuse ne permettait pas de sélectionner des clones positifs. Pour ce qui est des autres clones (avec la molécule donneuse issue d'*E. coli*),



les résultats de PCR indiquaient que dans le meilleur des cas (cl 9.2, 9.4, 9.5, 10.3 et 10.5), seule une fraction minime de la population était potentiellement issue d'évènements de recombinaison site-spécifique entre le site lox66 de la molécule donneuse et le lox71 de la cellule receveuse, la bande à 2285 pb étant nettement plus intense que celle à 1729 pb.

Malgré des résultats mitigés, nous avons décidé de poursuivre l'analyse sur les 5 clones présentant une double bande. Sur ces 5 clones, une deuxième PCR a été effectuée, permettant cette fois de contrôler la perte du gène de résistance au chloramphénicol (*cat*), après recombinaison entre les sites loxm2/71 de la molécule donneuse et loxm2/66 de la cellule receveuse. Après recombinaison site-spécifique, le gène *cat* est normalement remplacé par la fin du gène *trmE*, présent sur la molécule donneuse (figure 92A). Cet évènement entraîne une répétition directe de 699 pb, car le gène *trmE* chromosomique est présent au niveau de ce locus. Cette PCR peut donc produire deux profils différents, le premier avec deux amplicons aux tailles de 265 pb et 998 pb si la recombinaison site-spécifique a eu lieu (figure 92A), et le second avec un seul amplicon à la taille de 1844 pb en absence de cet évènement de recombinaison. Les 5 clones analysés ont bien montré un profil à deux bandes mais, présentant des tailles à 1844 pb et 265 pb, et non à 998 pb comme attendu (figure 92C). Ces résultats montraient très clairement que la recombinaison site-spécifique entre le loxm2/71 de la molécule donneuse et le loxm2/66 de la cellule receveuse n'avait pas eu lieu.

Les résultats des PCR 1 et PCR 2 indiquaient donc que le processus d'échange via le système cre/lox ne fonctionnait pas comme anticipé. Sachant que la molécule donneuse de 58 kb n'était pas basée sur un plasmide répliatif, nous ne comprenions pas pourquoi nous récupérions autant de clones résistant à la spectinomycine sur nos boîtes de sélection. Afin de déterminer si cela était dû à la présence de la molécule donneuse dans nos clones, nous avons décidé de réaliser une troisième PCR à l'aide d'un couple d'amorces s'hybridant spécifiquement au niveau du gène *cre* et du marqueur LYS2 de la molécule donneuse (amplifié attendu : 4083pb). La totalité des clones analysés ont présenté la bande attendue à 4083 pb, indiquant que la molécule donneuse était toujours présente dans les transformants. Nous avons donc tenté de l'éliminer en réalisant 4 sous-clonages. Malheureusement l'analyse des sous-clones par PCR a montré que la molécule donneuse était toujours présente.



Les résultats de cette expérience ont laissé apparaître, qu'en l'état, le système de recombinaison site-spécifique cre/lox n'était pas suffisant pour permettre l'échange de fragments dans la cellule receveuse. Nous faisons vraisemblablement face à deux problèmes : (i) les échanges site-spécifiques semblaient rares du côté « lox71-lox66 » (gauche ou 5') et inexistant du côté « loxm2/71-loxm2/66 » (droit ou 3') et (ii) la molécule donneuse, pourtant non répliquative, semblait persister dans les clones, peut-être sous forme intégrée ?

A ce stade, les résultats obtenus étaient peu convaincants et remettaient en cause la faisabilité de la méthode. Un seul élément positif était à retenir de cette expérience : la molécule donneuse de 58 kb semblait bel et bien entrer dans la cellule receveuse via le protocole de transformation MC basée sur la compétence naturelle de *B. subtilis*. A partir de ce moment-là, nous n'avons plus cherché à développer d'autres méthodes de transformation.

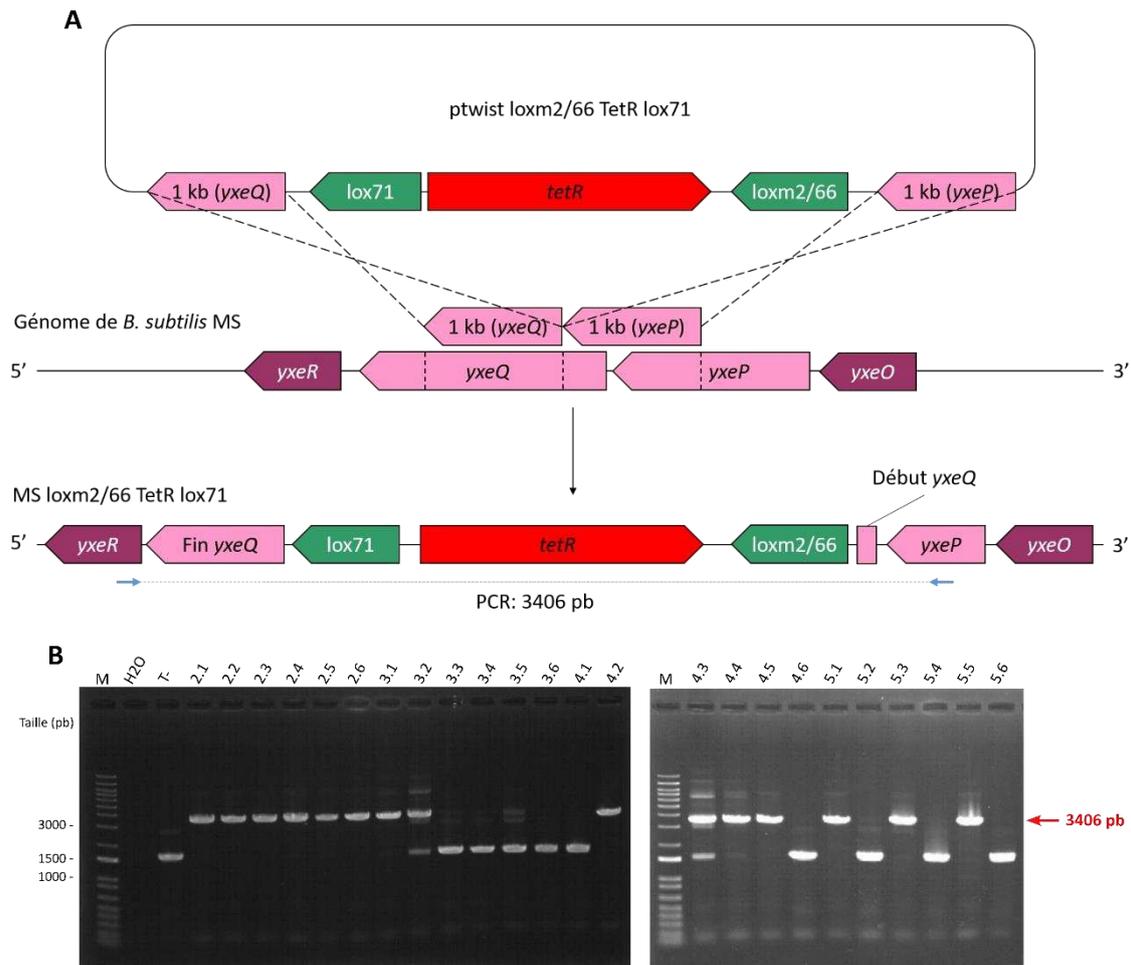
## 5. Validation de la fonctionnalité du système cre/lox chez *B. subtilis*

Après ce premier échec, nous avons souhaité nous rassurer et valider le fait que le système cre/lox était suffisamment efficace chez *B. subtilis* pour permettre l'échange de fragments désiré.

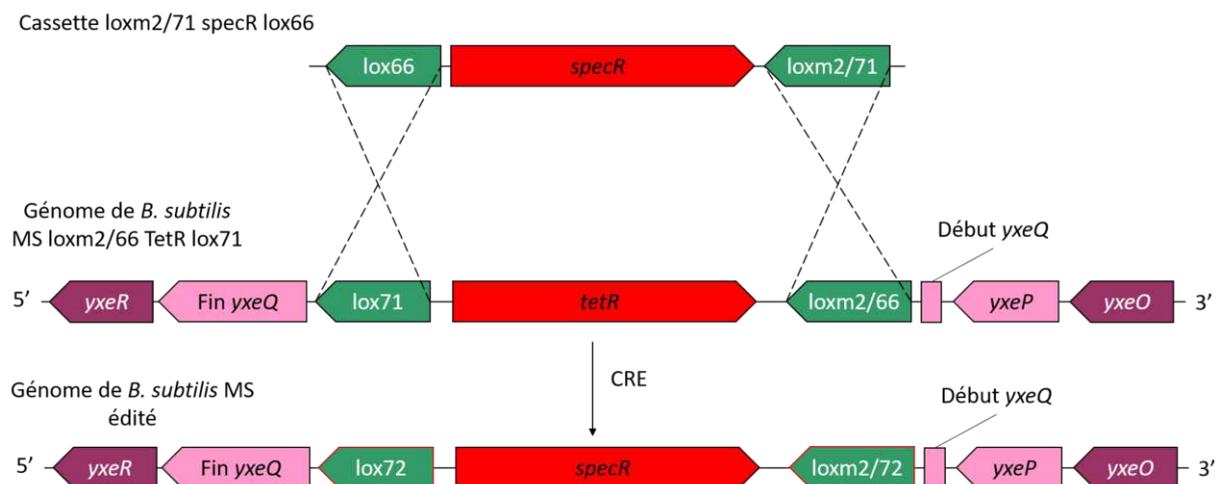
L'utilisation du système cre/lox chez *B. subtilis* avait, en effet, été décrit de nombreuses fois dans la littérature, notamment pour supprimer des marqueurs de sélection devenus inutiles ou non désirés après des travaux de génie génétique (Yan *et al.*, 2008; Koo *et al.*, 2017). Nous avons décidé de revenir vers des expériences simples afin d'évaluer l'efficacité du système cre/lox entre nos mains. Pour ce faire, des expériences d'échange, non pas de grands fragments de génome (155 kb vs 44 kb) mais de simples marqueurs de sélection (*tetR* vs *specR*) ont été entreprises.

### 5.1. Construction d'une cellule receveuse pour un échange de marqueurs

Pour entamer ces expériences, il nous fallait construire une cellule receveuse dans laquelle un marqueur de résistance antibiotique encadré de deux sites lox incompatibles (lox71 et loxm2/66) soit intégré au génome. Pour produire cette cellule receveuse, nous avons utilisé le plasmide ptwist loxm2/66 TetR lox71, dans lequel les sites loxm2/66 et lox71 bordant le



**Figure 93 : Insertion de la cassette loxm2/66 TetR lox71 pour obtenir une cellule receveuse pour un échange de marqueur via le système cre/lox [A] Schéma de l'insertion de la cassette loxm2/66 TetR lox71 dans la souche *B. subtilis* MS. [B] L'insertion est contrôlée par une PCR simple avec les amorces screening landing pad tet F et FE220 receveuse YxeP R produisant un amplicon de 3406 pb. « M » : DNA Ladder 1 kb+ Invitrogen (100-12000 pb). « H2O » : contrôle négatif sans ADN. T-: témoin négatif ADNg de la souche *B. subtilis* MS.**



**Figure 94 : Echange de marqueur via le système cre/lox. L'activité recombinase de CRE provoque l'intégration du gène *specR* à la place du gène *tetR*.**

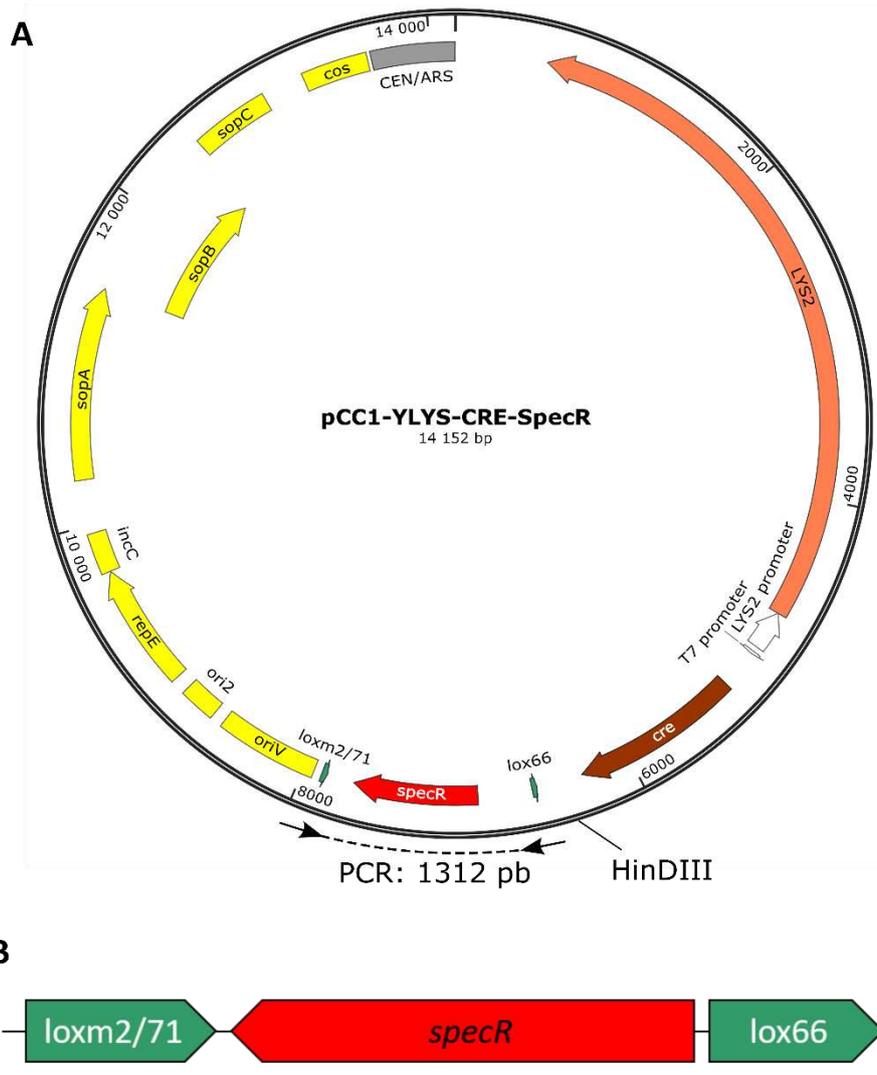
gène de résistance à la tétracycline sont flanqués de deux bras d'homologie d'environ 1 kb correspondant, en amont, à la fin du gène *yxep* et, en aval, au gène *yxq* (figure 93A).

La souche MS a été transformée avec le plasmide *ptwist loxm2/66 TetR lox71* en utilisant le protocole simple MC. Les transformants ont été sélectionnés sur milieu LB solide contenant 7,5 µg/ml. Après 24 h à 37°C, aucune colonie n'est apparue sur le contrôle négatif sans ADN alors que 259 colonies sont apparues sur les conditions de transformation avec le plasmide. Vingt-quatre d'entre elles ont été analysées par PCR avec un couple d'amorces permettant de générer un amplicon de 3406 pb si la cassette *loxm2/66 TetR lox71* a été insérée au locus et de 1538 pb si la région reste inchangée. Sur les 24 analysés, 12 ont montré une bande à 3406 pb (figure 93B). Pour ces 12 clones, l'absence de mutation au niveau de la cassette insérée dans le génome de *B. subtilis* a été confirmée par séquençage Sanger.

La souche MS *loxm2/66 TetR lox71*, dans laquelle le gène de résistance à la tétracycline (*tetR*) était intégré et bordé de part et d'autre avec les sites *loxm2/66* et *lox71*, était donc obtenue.

Nous avons eu l'idée de profiter de la mise en œuvre de cette expérience pour tester différents paramètres, en particulier la nécessité de pré-exprimer la protéine CRE dans la cellule receveuse avant d'introduire la cassette à échanger. Pour cela, il nous fallait construire une deuxième cellule receveuse. La souche MS *loxm2/66 TetR lox71*, nouvellement obtenue, a donc été transformée avec le *pDR244 Ery*, un plasmide répliquatif et sensible à la température, qui contient le gène *cre*. Les transformants ont été sélectionnés sur milieu LB solide contenant soit, 7,5 µg/ml de tétracycline et 1 µg/ml d'érythromycine, soit seulement 1 µg/ml d'érythromycine. Après 72h d'incubation à 37°C, dix transformants ont été obtenus sur boîte. Ils ont tous été analysés par PCR pour contrôler la présence du *pDR244 Ery* ainsi que celle de la cassette *loxm2/66 TetR lox71*. Trois d'entre eux ont montré les profils attendus, validant la présence du *pDR244 Ery* et de la cassette *loxm2/66 TetR lox71* dans ces clones.

A ce stade, les deux souches MS *loxm2/66 TetR lox71* et MS *loxm2/66 TetR lox71 + pDR244 Ery* étaient prêtes et pouvaient donc être utilisées pour les expériences d'échange de marqueurs de résistance via le système *cre/lox* (*i.e.* échange de la cassette (*tet*) *loxm2/66-TetR-lox71* porté par le génome de *B. subtilis* par la cassette (*spec*) *loxm2/71-specR-lox66* porté par un plasmide) (figure 94).



**Figure 95 : Représentation des molécules donneuses utilisées pour les expériences d'échange de marqueurs via le système cre/lox chez *B. subtilis*.** [A] Carte du plasmide pCC1-YLYS-CRE-SpecR. La PCR avec les amorces Gib\_lox66\_Spec\_F et Gib\_loxm2/71\_Spec\_R qui permet de produire la cassette loxm2/71 specR lox66 seule est représentée. Le site de coupure avec l'enzyme HindIII qui permet de linéariser le plasmide est aussi représenté. [B] Schéma de la cassette loxm2/71 specR lox66 seule, produite par PCR.

## 5.2. Echange de marqueurs via le système cre/lox chez *B. subtilis*

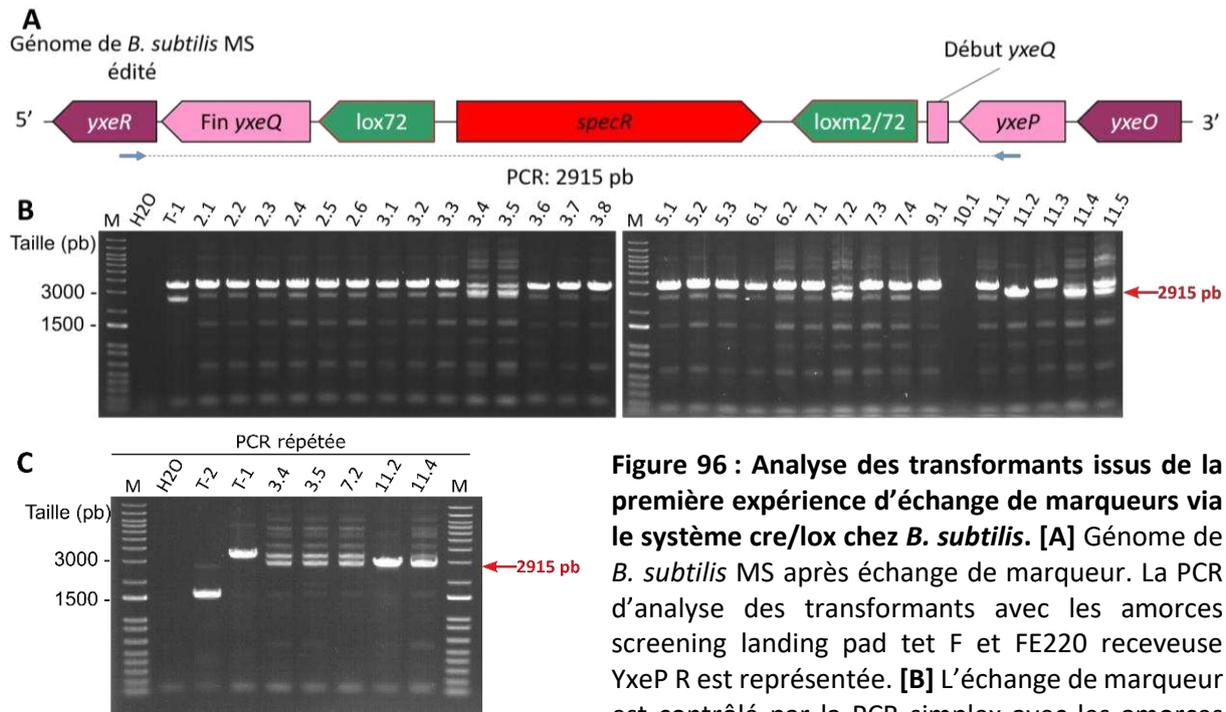
Les objectifs de ces expériences étaient (i) de définir si nous étions capables de faire fonctionner le système cre/lox chez *B. subtilis*, et (ii) de mettre en évidence les conditions dans lesquelles le système était le plus efficace. Diverses conditions expérimentales ont donc été testées en parallèle :

- 1) Pré-expression de la CRE en comparant les souches « MS loxm2/66 TetR lox71 + pDR244 Ery » et « MS loxm2/66 TetR lox71 » à 30°C ou à 37°C. Le plasmide pDR244, qui exprime de manière constitutive la recombinase CRE, contient une origine de réplication sensible à la température. A température permissive (30°C), le plasmide est maintenu et transmis aux cellules filles de génération en génération. A température non permissive (42°C), le plasmide pDR244 est rapidement éliminé des cellules. Ces expériences ont été menées à 30°C ou à 37°C (température optimale de *B. subtilis* mais plus proche de 42°C). Après transformation, les cellules ont été étalées sur milieu de sélection LB spectinomycine 100 µg/ml ou LB spectinomycine 100µg/ml plus erythromycine 1 µg/ml.
- 2) Double expression de la CRE en apportant une molécule d'ADN portant elle aussi le gène *cre*.
- 3) Ajout d'une étape d'expression de la protéine CRE en culture liquide après avoir introduit la molécule d'ADN à échanger.

Deux expériences indépendantes ont été réalisées au cours desquelles les souches MS loxm2/66 TetR lox71 et MS loxm2/66 TetR lox71 + pDR244 Ery ont été transformées à l'aide du protocole simple MC. Pour tester tous les paramètres décrits précédemment, nous avons utilisé trois molécules d'ADN différentes. Elles contenaient toutes la cassette loxm2/71 specR lox66, composée du gène de résistance à la spectinomycine (*specR*) bordé de deux sites lox incompatibles entre eux (loxm2/71 et lox66). La première molécule d'ADN correspondait au plasmide pCC1-YLYS-CRE-SpecR (figure 95A). Pour rappel, ce plasmide d'environ 14 kb avait été utilisé pour ajouter, sur le fragment 7.2, toutes les séquences nécessaires à la mise en place de la méthode RAGE. Il contenait donc la cassette loxm2/71 specR lox66 ainsi que le gène *cre*. Il devait, théoriquement, être suffisant pour provoquer un échange d'ADN via le système cre/lox. La seconde molécule d'ADN correspondait au même plasmide mais sous

**Tableau 23 : Conditions expérimentales de la première expérience d'échange de marqueurs via le système cre/lox chez *B. subtilis*.**

Conditions	Souche receveuse	Température	ADN transformé	Sélection	CFU	Clones positifs / Clones analysés
1	MS loxm2/66 TetR lox71	37°C	Sans ADN	LB spec 100 µg/ml	3	-
2			pCC1-YLYS-CRE-SpecR linéaire		6	0/6
3			pCC1-YLYS-CRE-SpecR circulaire		8	0/8
4	MS loxm2/66 TetR lox71 + pDR244 Ery		Sans ADN	LB spec 100 µg/ml	3	-
5			Cassette loxm2/71 specR lox66		3	0/3
6			pCC1-YLYS-CRE-SpecR linéaire		2	0/2
7			pCC1-YLYS-CRE-SpecR circulaire		>300	0/4
8	MS loxm2/66 TetR lox71 + pDR244 Ery		Sans ADN	LB spec 100 µg/ml + ery 1 µg/ml	0	-
9			Cassette loxm2/71 specR lox66		1	0/1
10			pCC1-YLYS-CRE-SpecR linéaire		1	0/1
11			pCC1-YLYS-CRE-SpecR circulaire		>100	2/5
12			pDR111		>100	-



**Figure 96 : Analyse des transformants issus de la première expérience d'échange de marqueurs via le système cre/lox chez *B. subtilis*.** [A] Génome de *B. subtilis* MS après échange de marqueur. La PCR d'analyse des transformants avec les amorces screening landing pad tet F et FE220 receveuse YxeP R est représentée. [B] L'échange de marqueur est contrôlé par la PCR simplex avec les amorces

screening landing pad tet F et FE220 receveuse YxeP R produisant un amplicon de 2915 pb. [C] Afin de s'assurer des résultats de la PCR, cette dernière a été répétée sur les clones 3.4, 3.5, 7.2, 11.2, 11.4. « M » : DNA Ladder 1 kb+ Invitrogen (100-12000 pb). « H2O » : contrôle négatif sans ADN. T-1: témoin négatif ADNg de la souche *B. subtilis* MS loxm2/66 tetR lox71. T-2: témoin négatif ADNg de la souche *B. subtilis* MS sauvage.

forme linéaire après digestion avec l'enzyme HindIII. Enfin la troisième molécule correspondait à la cassette loxm2/71 specR lox66 seule, produite par PCR à partir de ce plasmide (figure 95B).

Lors de la première expérience, la souche MS loxm2/66 TetR lox71 a été transformée avec le plasmide pCC1-YLYS-CRE-SpecR sous formes circulaire ou linéaire après digestion avec l'enzyme HindIII (expression de la CRE amenée par le plasmide donneur, tableau 23, conditions 2, et 3). La souche MS loxm2/66 TetR lox71 + pDR244 Ery a, elle aussi, été transformée avec le plasmide pCC1-YLYS-CRE-SpecR sous formes circulaire ou linéaire (expression de la CRE amenée par le plasmide donneur et la cellule receveuse, tableau 23, conditions 6, 7, 10 et 11) ainsi qu'avec la cassette loxm2/71 specR lox66 seule (expression de la protéine CRE amené par la cellule receveuse, tableau 23, conditions 5 et 9). Les transformants ont été sélectionnés sur milieu solide LB contenant soit 100 µg/ml de spectinomycine et 1 µg/ml d'érythromycine, soit seulement 100 µg/ml de spectinomycine. Après 72h d'incubation à 37°C, sur les contrôles négatifs sans ADN, 3 colonies sont apparues dans les conditions 1 et 4, et aucune colonie n'est apparue dans la condition 8, suggérant la présence de faux positifs sur les boites contenant seulement 100 µg/ml de spectinomycine. Concernant les essais, alors que les conditions 2-3, 5-6 et 9-10 présentaient peu de colonies, les conditions 7 et 11, qui correspondaient à la souche MS loxm2/66 TetR lox71 + pDR244 Ery transformée avec le plasmide circulaire pCC1-YLYS-CRE-SpecR (donc double expression de la CRE) présentaient plusieurs centaines de colonies sur boite (tableau 23).

Un total de 30 transformants a été analysé. Cette analyse a été réalisée par PCR au niveau du locus échangé, en utilisant un couple d'amorces qui permettait de générer un amplicon de 2915 pb si le marqueur *tetR* avait bel et bien été remplacé par le marqueur *specR* et un amplicon de 3406 pb si le marqueur *tetR* était toujours présent. Seulement 2 transformants (issus uniquement de la condition 11 : cl 11.2 et 11.4), ont présenté une bande unique à 2915 pb (figure 96). Les amplicons ont été purifiés puis envoyés au séquençage pour déterminer la séquence des sites lox. Les analyses de séquences ont permis de valider la recombinaison au niveau des sites lox dans les deux clones.

Malgré un taux de réussite faible, ces résultats semblaient indiquer que, pour réaliser un échange via le système cre/lox, les meilleures conditions correspondaient à :

**Tableau 24 : Conditions expérimentales de la deuxième expérience d'échange de marqueurs via le système cre/lox chez *B. subtilis*.**

Conditions	Souche receveuse	Température	ADN transformé	Etape de culture liquide avant étalement	Sélection	CFU	Clones positifs / Clones analysés
1	MS loxm2/66 TetR lox71 + pDR244 Ery	37°C	Sans ADN	-	LB spec 100 µg/ml	-	-
2				+		Tapis	-
3			pCC1-YLYS-CRE- SpecR circulaire	-		>500	0/10
4				+		Tapis	-
5			Sans ADN	-	LB spec 100 µg/ml + ery 1 µg/ml	-	-
6				+		-	-
7			pCC1-YLYS-CRE- SpecR circulaire	-		100	1/10
8				+		Tapis	-
9			pDR111	-	100	-	
10				+	Tapis	-	
11		30°C	pCC1-YLYS-CRE- SpecR circulaire	-	LB spec 100 µg/ml	100	3/10
12				+		Tapis	-
13				-	LB spec 100 µg/ml + ery 1 µg/ml	100	4/10
14				+		Tapis*	5/10

\*Malgré la présence d'un grand nombre de colonies, quelques une étaient isolées.

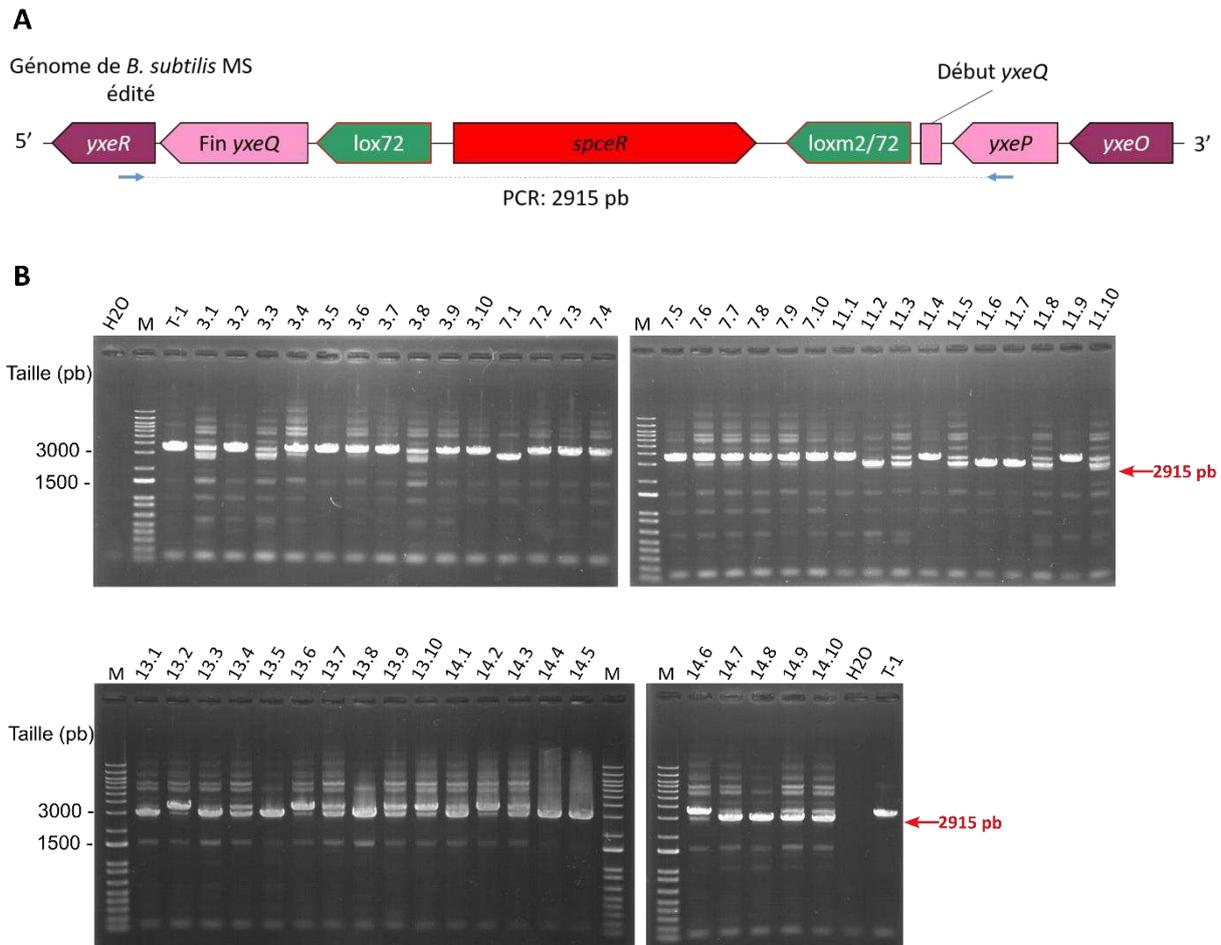
- (i) La transformation d'une souche pré-exprimant la CRE (MS loxm2/66 TetR + pDR244 Ery) avec un plasmide circulaire exprimant également la CRE (pCC1-YLYS-CRE-SpecR). Le système semble donc nécessiter une double expression de la CRE.
- (ii) L'utilisation de milieu de sélection garantissant le maintien du plasmide portant *cre* après transformation (érythromycine pour le plasmide pDR244 Ery).

Afin d'affiner ces résultats, une seconde expérience a été réalisée. Lors de cet essai, la souche MS loxm2/66 TetR lox71 + pDR244 Ery a été transformée avec le plasmide circulaire pCC1-YLYS-CRE-SpecR et l'expérience a été menée soit à 37°C, soit à 30°C. De plus, pour laisser le temps au système *cre/lox* de s'exprimer, une incubation d'une nuit en culture liquide a été ajoutée avant étalement sur milieu de sélection. Comme précédemment, les transformants ont été étalés sur milieu solide LB contenant soit 100 µg/ml de spectinomycine et 1 µg/ml d'érythromycine, soit 100 µg/ml de spectinomycine. Ces antibiotiques étaient présents lors de l'incubation d'une nuit en culture liquide.

Toutes les conditions de transformation sont répertoriées dans le tableau 24. Après 24 h d'incubation à 37°C ou à 30°C, aucune colonie n'est apparue sur les boîtes de contrôles négatifs sans ADN (conditions 1, 5 et 6), excepté pour la condition 2 qui présentait un tapis cellulaire. Cette condition correspondait au contrôle négatif sans ADN en sélection simple (100 µg/ml de spectinomycine) pour laquelle l'incubation d'une nuit en milieu liquide avait été ajoutée.

Pour toutes les conditions de transformation sans incubation d'une nuit en milieu liquide (conditions 3, 7, 11 et 13), plusieurs centaines de colonies ont été obtenues sur boîte. Lorsque l'étape de culture liquide a été ajoutée (conditions 4, 8 et 12), des tapis cellulaires sont apparus, empêchant l'analyse de clones isolés. Seule la condition 14 a permis d'obtenir quelques colonies isolées malgré la présence de très nombreuses colonies sur milieu de sélection (tableau 24).

L'analyse des transformants a été réalisée de manière identique à l'expérience précédente. Sur 50 clones analysés, 13 ont montré une bande unique comme attendue à 2915 pb (figure 97). Sur ces 13 clones, 12 provenaient des conditions 11 (cl 11.2, 11.6 et 11.7), 13 (cl 13.1, 13.3, 13.5 et 13.8) et 14 (cl 14.1, 14.4, 14.5, 14.7 et 14.8) menées à 30°C, suggérant fortement que cette température, propice au maintien du plasmide et donc à l'expression de



**Figure 97 : Analyse des transformants issus de la seconde expérience d'échange de marqueurs via le système cre/lox chez *B. subtilis*. [A] Génome de *B. subtilis* MS après échange de marqueur. La PCR d'analyse des transformants avec les amorces screening landing pad tet F et FE220 receveuse YxeP R est représentée. [B] L'échange de marqueur est contrôlé par la PCR simple avec les amorces screening landing pad tet F et FE220 receveuse YxeP R produisant un amplicon de 2915 pb. « M » : DNA Ladder 1 kb+ Invitrogen (100-12000 pb). « H2O » : contrôle négatif sans ADN. T-1: témoin négatif ADNg de la souche *B. subtilis* MS loxm2/66 tetR lox71. T-2: témoin négatif ADNg de la souche *B. subtilis* MS sauvage.**

*cre*, favorisait les évènements de recombinaison site-spécifique. En revanche, l'incubation d'une nuit en culture liquide avant étalement ne semblait guère augmenter l'efficacité de l'échange.

Avec ces deux expériences nous avons pu (i) prouver que le système *cre/lox* était bien apte à provoquer un échange de marqueurs chez *B. subtilis*, et (ii) mettre en évidence des conditions de transformation dans lesquelles le système semblait plus efficace. Cette dernière consistait à transformer une cellule receveuse qui pré-exprimait la CRE avec une molécule donneuse circulaire exprimant elle aussi la CRE. Cette transformation devait être réalisée à 30°C et le pDR244 Ery devait constamment être maintenu avec 1 µg/ml d'érythromycine dans tous les milieux de sélection.

## 6. Test de la méthode RAGE avec pré-expression de la CRE

Dans la section précédente, nous avons montré que les conditions les plus favorables à un échange de marqueurs via le système *cre/lox* chez *B. subtilis* étaient de pré-exprimer la CRE dans la cellule receveuse et d'apporter une molécule donneuse l'exprimant aussi. Ces nouvelles données nous ont incités à retenter l'étape 4 du RAGE avec une cellule receveuse FE220 V3 pré-exprimant la protéine CRE. Pour cela, nous avons mis en place trois procédures qui sont détaillées ci-dessous :

- Co-transformation de la cellule receveuse FE220 V3 avec le pDR244 Ery et la molécule donneuse pCC1-LYS-cre-specR frgtBsu 7.2 (figure 80)
- Insertion du gène *cre* dans le génome de la cellule receveuse FE220 V3 puis transformation de FE220 V3 + *cre* avec la molécule donneuse pCC1-LYS-cre-specR frgtBsu 7.2
- Transformation successive de la cellule receveuse FE220 V3 avec le pDR244 Ery, puis de FE220 V3 pDR244 Ery avec la molécule donneuse pCC1-LYS-cre-specR frgtBsu 7.2.

### 6.1. Co-transformation de la cellule receveuse FE220 V3 avec le pDR244 Ery et la molécule donneuse (deuxième essai)

Parmi les 3 procédures envisagées, celle basée sur la co-transformation du pDR244 Ery et de la molécule donneuse paraissait la plus critique puisqu'elle ne permettait pas de pré-exprimer le gène *cre* dans la cellule receveuse mais seulement de l'amener en plusieurs copies

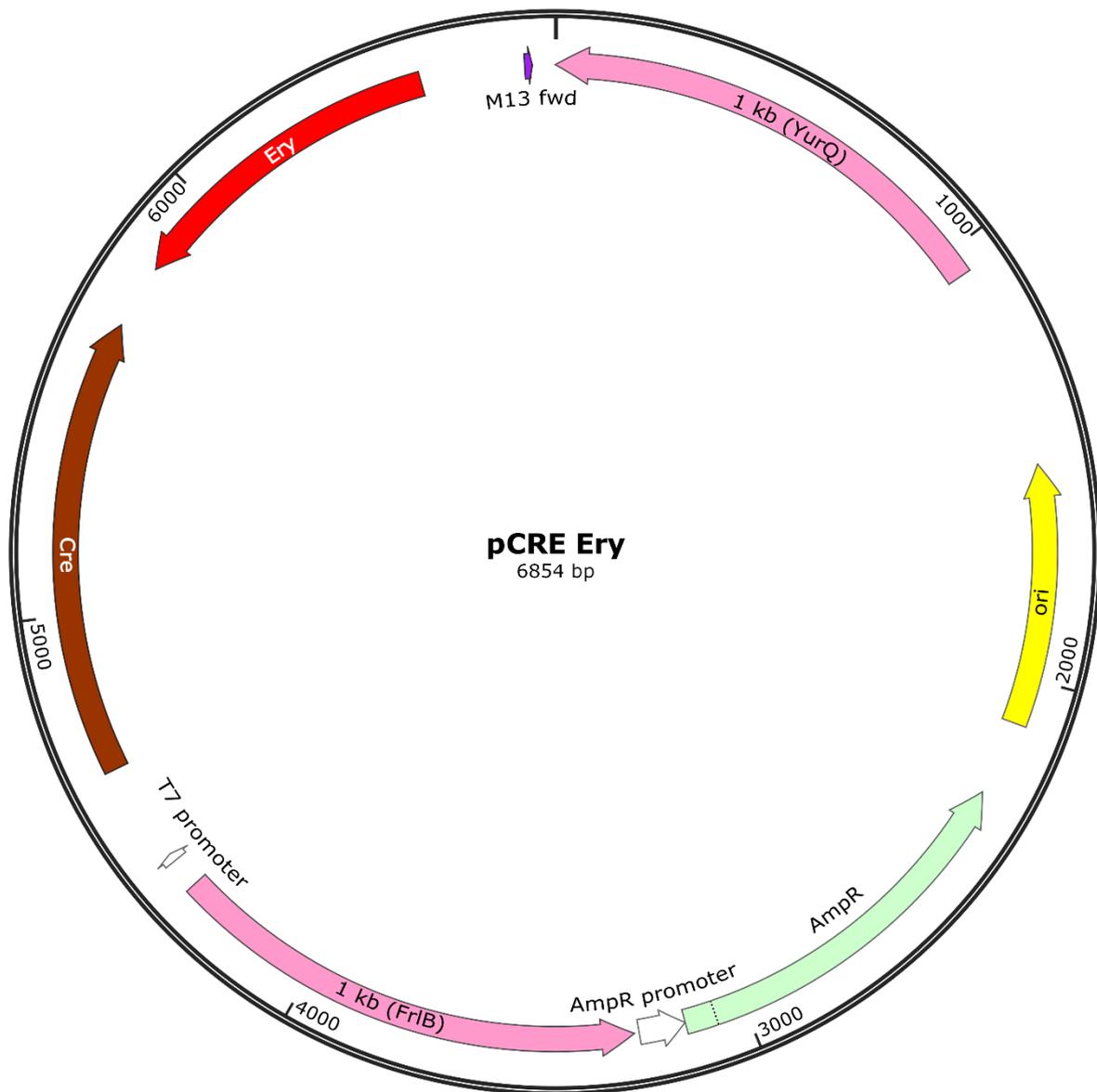


Figure 98 : Carte du plasmide pCRE Ery.

au moment de l'échange. Cependant, étant plus simple et plus rapide que les deux autres méthodes, nous avons souhaité la tester.

La cellule receveuse FE220 V3 a donc été co-transformée avec le plasmide pDR244 Ery et la molécule donneuse pCC1-LYS-cre-specR frgtBsu 7.2 en utilisant le protocole MC. Puis les transformants ont été sélectionnés sur milieu LB solide contenant 100 µg/ml de spectinomycine et 1 µg/ml d'érythromycine. Après 72h d'incubation à 30°C, aucune colonie n'est apparue sur le contrôle négatif alors que deux colonies étaient visibles pour les conditions d'échange de fragments. Cependant, après les avoir repiquées pour les analyser, aucune des deux n'a poussé en milieu liquide. La tentative de co-transformation a donc été un échec. Il est possible que peu de cellules (voire aucune cellule) n'aient été transformées par les deux plasmides, d'autant plus qu'un des deux, a une taille de 58 kb. Il est aussi possible (dans le cas où les deux plasmides seraient entrés dans les cellules) que l'expression des gènes *cre* n'ait pas été suffisante pour permettre l'échange.

## 6.2. Transformation de la molécule donneuse dans une cellule receveuse FE220 V3 portant une copie du gène *cre* inséré à son génome (troisième essai)

Les données que nous commençons à accumuler suggéraient que la dernière étape du RAGE devait être réalisée avec une cellule receveuse pré-exprimant la CRE. Le gène *cre* devait donc être présent dans la cellule receveuse FE220 V3 avant l'introduction de la molécule donneuse. Deux options s'offraient à nous : intégrer une copie du gène *cre* au génome de la cellule receveuse, ou bien transformer FE220 V3 avec le plasmide pDR244 Ery utilisé au cours des expériences d'échange de marqueurs. Nous avons tenté les deux.

Nous avons commencé par l'expérience d'intégration du gène *cre* dans le génome de la cellule receveuse FE220 V3. Pour ce faire, nous avons construit le plasmide pCRE Ery qui porte le gène *cre* (le même que dans pDR244), bordé de deux bras de recombinaison de 1 kb, correspondant aux gènes *yurQ* et *FrlB* permettant son insertion dans le génome de *B. subtilis* (figure 98). La cellule receveuse FE220 V3 a ensuite été transformée avec ce plasmide (protocole de transformation MC) et les transformants ont été sélectionnés sur milieu LB solide contenant soit 1 µg/ml d'érythromycine, 4 µg/ml de chloramphénicol et 7,5 µg/ml de tétracycline, soit seulement 1 µg/ml d'érythromycine. Après 72 h d'incubation à 37°C, aucune

**Tableau 25 : Conditions expérimentales de l'essai d'échange de grands fragments de génome via le système cre/lox lorsque le gène cre est inséré au génome.**

Conditions	Souche receveuse	Température	ADN transformé	Sélection	CFU	Clones positifs / Clones analysés	
						PCR 1	PCR 2
1	cellule receveuse V3 + cre	37°C	Sans ADN	LB spec 100 µg/ml + ery 1 µg/ml	0	-	-
2			pDR111		NC <sup>a</sup>	-	-
3			Molécule donneuse clone 4.8.3		Tapis <sup>b</sup>	0/10	-
4			Molécule donneuse clone 16.2.3		Tapis <sup>b</sup>	0/10	-
5		30°C	Sans ADN		0	-	-
6			pDR111		NC <sup>a</sup>	-	-
7			Molécule donneuse clone 4.8.3		Tapis <sup>b</sup>	0/10	-
8			Molécule donneuse clone 16.2.3		Tapis <sup>b</sup>	1/10	0/1

<sup>a</sup> Non communiqué

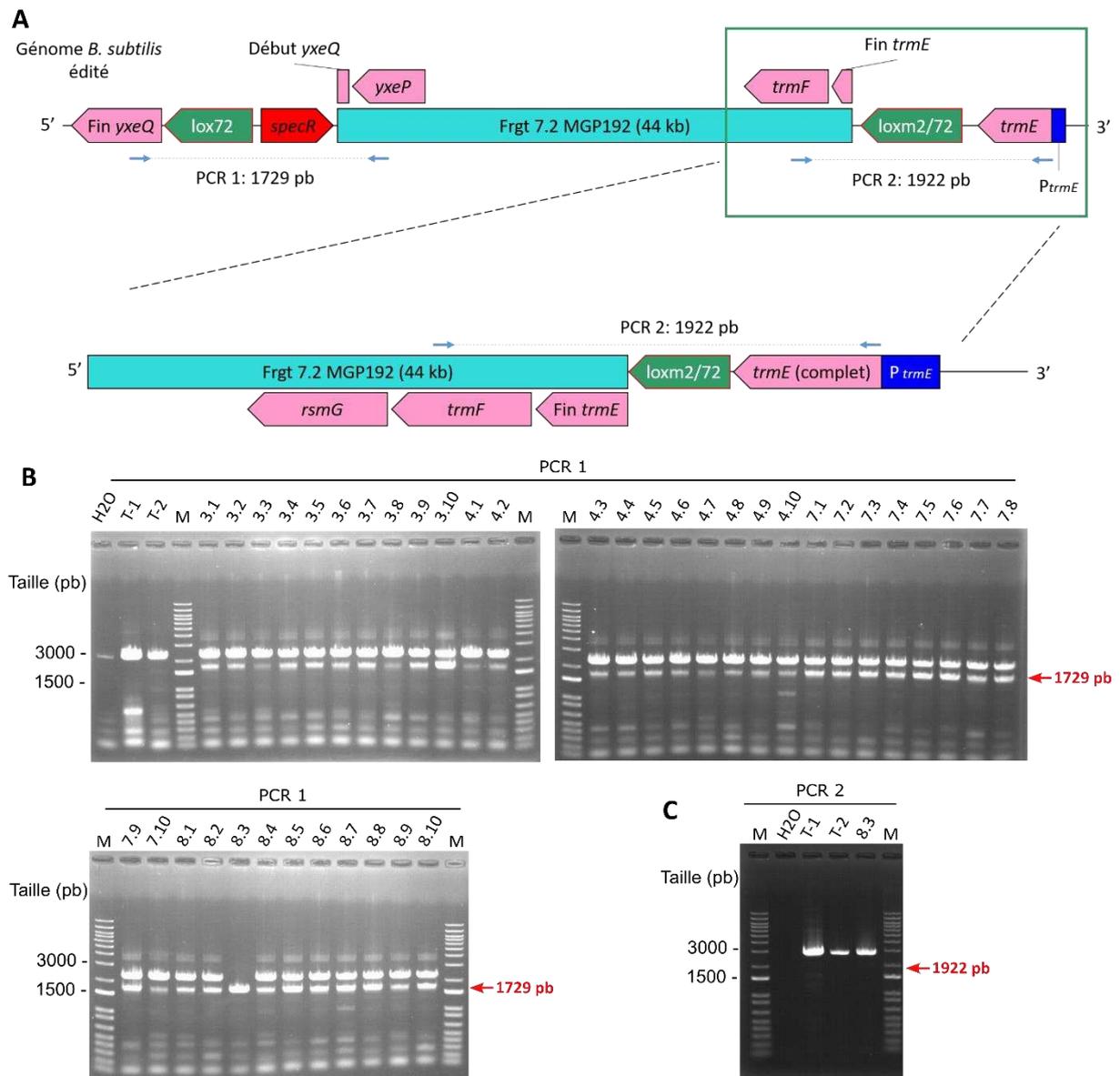
<sup>b</sup> Des colonies isolées n'ont pu être observées que sur les boîtes d'étalement des dilutions à 10<sup>-3</sup>.

colonie n'est apparue sur les contrôles négatifs sans ADN. Pour les conditions de transformation avec le pCRE Ery, 7 colonies ont été obtenues sur les boîtes de sélection (1 pour les conditions de sélection avec les trois antibiotiques, et 6 pour les conditions de sélection avec uniquement l'érythromycine). Elles ont toutes été repiquées, mais seulement 3 d'entre elles ont montré une croissance en milieu liquide additionné d'antibiotique. Les résultats de PCR avec la présence d'un amplicon à 5055 pb (contre 2527 pb pour les clones négatifs) ont permis de confirmer l'insertion du gène *cre* au locus voulu dans un clone issu des conditions de sélection avec seulement l'érythromycine. Chez ce clone, la présence des sites *loxm2/66* et *lox71* a été vérifiée validant l'obtention de la cellule receveuse FE220 V3 + *cre*.

Après obtention de la cellule receveuse V3 + *cre*, cette dernière a été transformée avec les molécules donneuses pCC1-LYS-cre-specR frgtBsu 7.2 issues des deux clones *E. coli*, puis étalées sur milieu LB solide contenant 100 µg/ml de spectinomycine et 1 µg/ml d'érythromycine. Même si dans cette expérience, le gène *cre* n'était pas porté par le plasmide thermosensible pDR244 et ne nécessitait pas d'être particulièrement maintenu à 30°C, nous avons conservé ces deux conditions et avons réalisé l'expérience à 30°C et à 37°C en parallèle (tableau 25).

Après 24 h d'incubation pour les conditions à 37°C et 48 h pour celles à 30°C, aucune colonie n'est apparue sur boîte pour les contrôles négatifs. En revanche, plusieurs centaines de colonies sont apparues pour les conditions de transformation avec les molécules donneuses. Des colonies isolées n'ont pu être obtenues que sur les boîtes de dilution 10<sup>-3</sup>. Quarante clones (20 issus des conditions à 37°C et 20 issus de celles à 30°C) ont été analysés à l'aide de deux PCR. La première utilisait le même couple d'amorces que dans la partie 4.2 du chapitre 3, la seconde, en revanche, utilisait un couple d'amorces différents, produisant un amplicon de 1922 pb si l'échange avait eu lieu et de 2768 pb si la région restait intacte (figure 99A).

Pour la PCR 1, qui permettait de contrôler l'incorporation du gène *specR* et la perte de *tetR*, 1 clone sur 40 (issu d'une condition de transformation à 30°C), a présenté le profil attendu (i.e. une bande à 1729 pb), tandis que les 39 autres ont présenté un profil mixte à 2 bandes (2285 pb + 1729 pb) suggérant que ces clones n'étaient pas purs (figure 99B). Étonnamment, à la différence de la première expérience sans pré-expression de la CRE (cf



**Figure 99 : Analyse des transformants de l'expérience d'échange de grands fragments de génome via le système cre/lox lorsque le gène cre est inséré au génome. [A] Schéma du locus échangé dans le génome de *B. subtilis* édité. Les PCR 1 et 2 sont représentées. La PCR 1 avec les amorces FE220 receveuse YxeQ F et FE220 receveuse YxeP R est identique seulement pour le premier essai (cf partie 4.2). La PCR 2 utilise les amorces Verif receveuse F et Seq échange CmR R. [B] L'échange de fragment est contrôlé par deux PCR au niveau des jonctions entre le génome et le fragment issu de la molécule donneuse (représentées en A). En cas d'échange de fragment, la PCR 1 produit un amplicon de 1729 pb [C] La PCR 2 produit un amplicon de 1922 pb. « M » : DNA Ladder 1 kb+ Invitrogen (100-12000 pb). « H2O » : contrôle négatif sans ADN. T-1: témoin négatif plasmide pTwist TetR lox71 en B et pTwist CmR loxm2/66 V3 en C. T-2: témoin négatif ADN<sub>g</sub> de la cellule receveuse FE220 V3 + cre.**

partie 4.2 du chapitre 3), l'intensité des deux bandes était cette fois quasi identique, en tout cas pour les clones issus des conditions de transformation à 30°C.

Ces résultats étaient intéressants car ils montraient pour la première fois une recombinaison complète entre le site lox71 de la cellule receveuse et le site lox66 de la molécule donneuse. De plus, on notait que chez les clones mixtes, la fraction de la population ayant effectué la recombinaison site-spécifique semblait plus importante que lors du premier essai (cf partie 4.2 du chapitre 3). Enfin, la transformation à 30°C semblait donc, encore une fois, offrir des conditions favorables à l'enclenchement des événements de recombinaison site-spécifiques.

Nous avons poursuivi l'analyse avec la PCR 2, qui permettait de contrôler la perte du gène de résistance au chloramphénicol *cat* sur la borne 3', du clone présentant une bande unique lors de la PCR 1 (cl 8.3). Malheureusement, le clone analysé n'a montré qu'une seule bande à la taille de 2768 pb correspondant au profil de la cellule receveuse (figure 99C).

L'ensemble de ces résultats a montré que l'échange de fragments via le système cre/lox chez *B. subtilis* n'avait, une nouvelle fois, pas fonctionné. Cependant pour la première fois, nous avons observé une recombinaison entre le site lox71 de la cellule receveuse et le site lox66 de la molécule donneuse. Cela nous a permis de confirmer la nécessité (i) d'effectuer la transformation à 30°C et (ii) de pré-exprimer la CRE avant transformation.

Au vu de ces résultats, les niveaux d'expression et de pré-expression de la CRE nous ont semblé être un point critique pour la mise en place de la méthode RAGE. Le gène *cre*, intégré au génome n'était présent qu'en une seule copie, son expression n'était potentiellement pas suffisante pour induire la recombinaison au niveau des deux bornes du fragment. Il nous fallait donc trouver un moyen d'augmenter le niveau d'expression du gène *cre*.

### 6.3. Pré-expression de la CRE dans une cellule receveuse FE220 V3 avec le plasmide répliatif pDR244 Ery

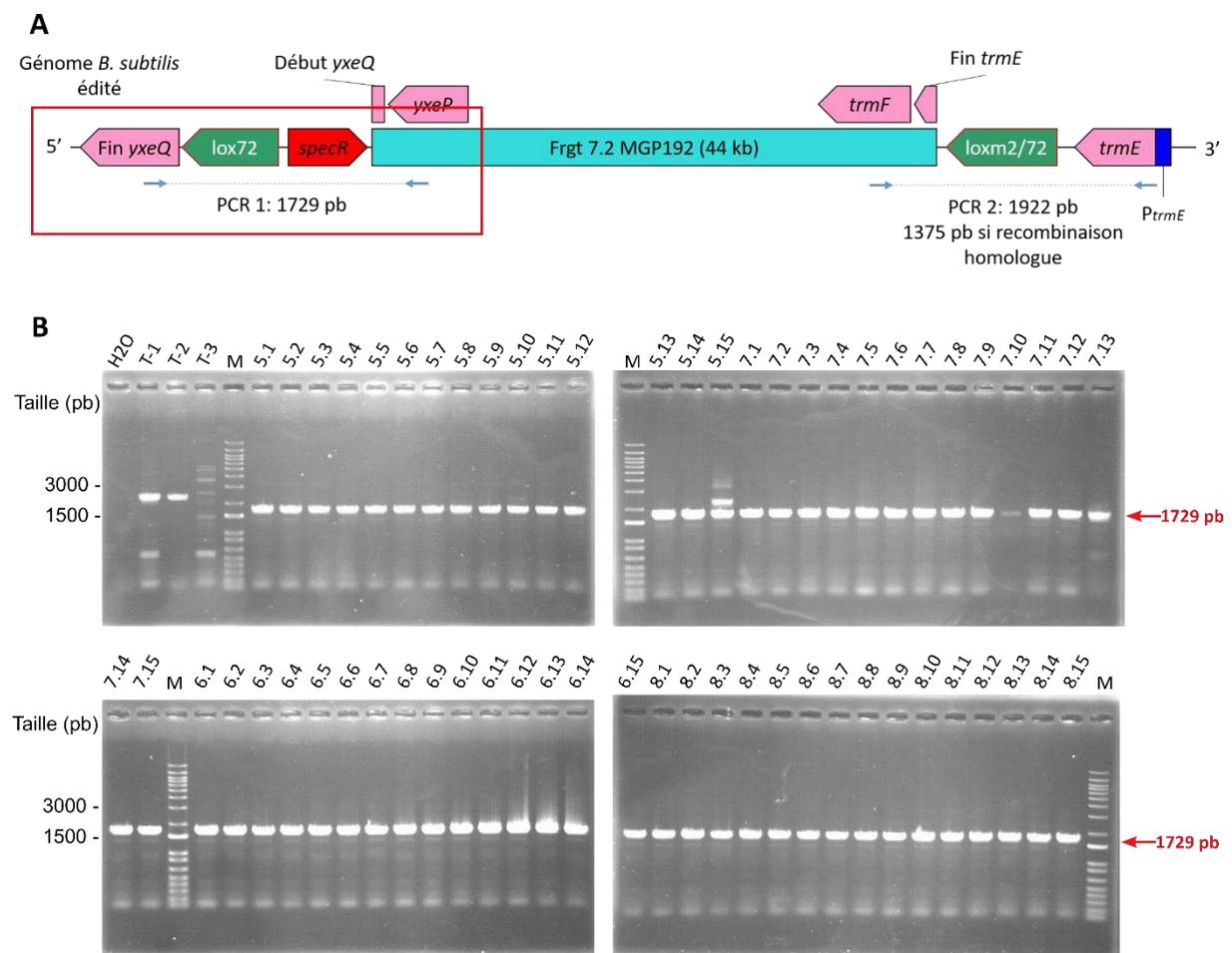
Dans la section précédente, nous avons vu que la pré-expression de CRE avec le gène inséré dans le génome bactérien n'était pas suffisante pour provoquer l'échange de fragments désiré. Afin d'exprimer la CRE plus fortement, nous avons donc choisi d'utiliser un plasmide

**Tableau 26 : Conditions expérimentales de l'essai d'échange de grands fragments de génome via le système cre/lox lorsque la CRE est pré-exprimée par l'intermédiaire du pDR244 Ery.**

Conditions	Souche receveuse	Température	ADN transformé	Sélection	CFU	Clones positifs / Clones analysés			
						PCR 1	PCR 2	PCR longues	PFGE
1	cellule receveuse V3 + pDR244 Ery	30°C	Sans ADN	LB spec 100 µg/ml + ery 1 µg/ml	0	-	-	-	-
2			pDR111		NC <sup>a</sup>	-	-	-	-
3			Molécule donneuse clone 4.8.3		Tapis <sup>b</sup>	29/30	11/29	3/11	3/3
4			Molécule donneuse clone 16.2.3		Tapis <sup>b</sup>	30/30	9/30	1/9	0/1

<sup>a</sup> Non communiqué

<sup>b</sup> Des colonies isolées n'ont pu être observées que sur les boîtes d'étalement des dilutions à 10<sup>-3</sup>.



**Figure 100 : Analyse des transformants de la quatrième expérience d'échange de fragment via le système cre/lox chez *B. subtilis* lorsque la CRE est pré-exprimée par le pDR244 Ery avec la PCR 1. [A] Schéma du résultat de l'échange de fragment. [B] Analyse des transformants par la PCR 1 avec les amorces FE220 receveuse YxeQ F et FE220 receveuse YxeP R. En cas d'échange de fragment elle produit un amplicon de 1729 pb « M » : DNA Ladder 1 kb+ Invitrogen (100-12000 pb). « H2O » : contrôle négatif sans ADN. T-1 : témoin négatif plasmide pTwist TetR lox71. T-2: témoin négatif ADNg de la cellule receveuse + pDR244 Ery. T-3 : témoins négatif molécule donneuse issue du clone *E. coli* 4.8.3.**

réplicatif, le pDR244 Ery. En effet ce dernier avait été utilisé lors des expériences d'échange de marqueurs (cf partie 5.2 du chapitre 3), et nous savions donc qu'il était suffisant pour provoquer un échange d'ADN via le système cre/lox chez *B. subtilis*.

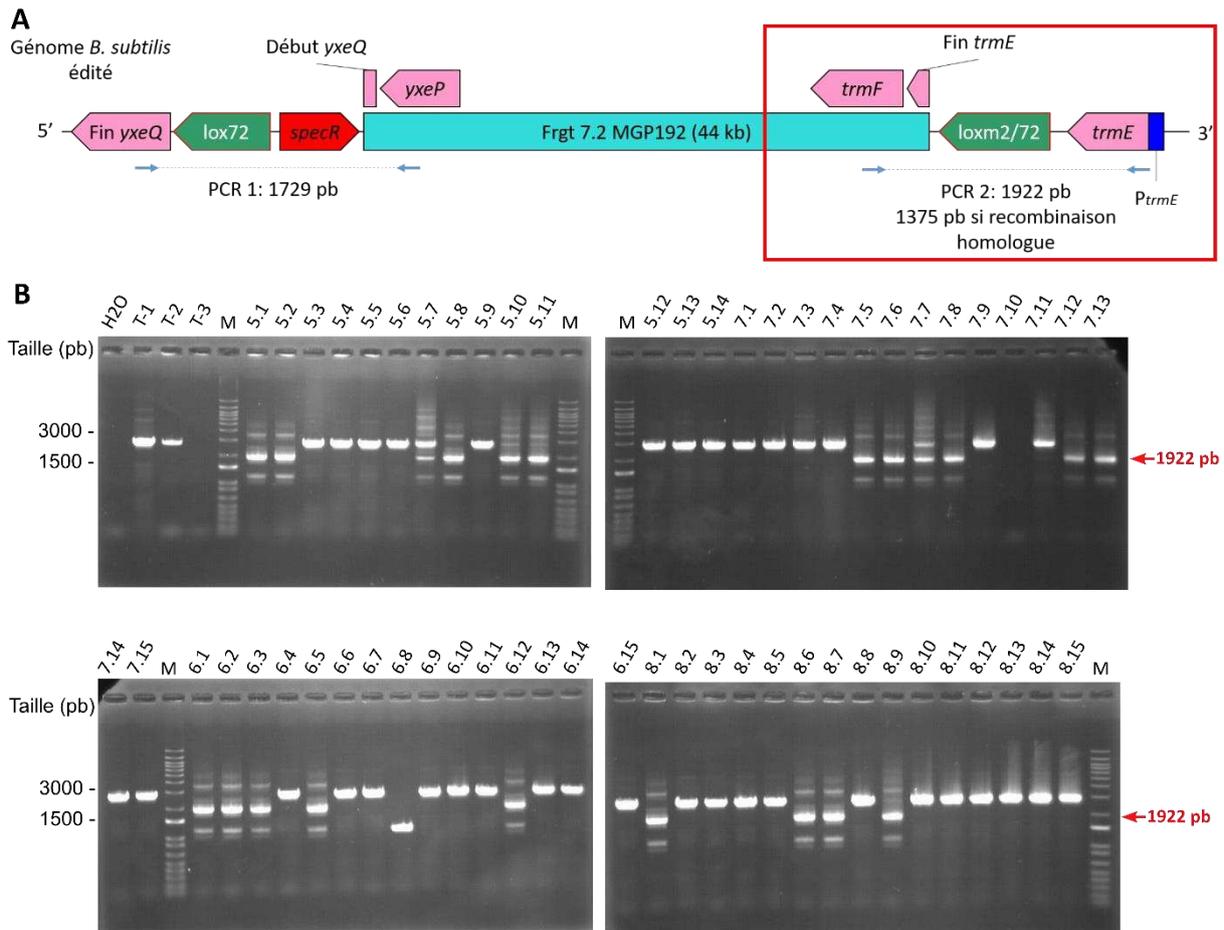
La cellule receveuse FE220 V3 a donc été transformée par le pDR244 Ery avec le protocole simple MC. Les transformants ont été sélectionnés sur milieu LB solide contenant soit 1 µg/ml d'érythromycine, 4 µg/ml de chloramphénicol et 7,5 µg/ml de tétracycline soit seulement 1 µg/ml d'érythromycine. Après 72 h d'incubation à 37°C, 17 clones ont été obtenus. Ils ont tous été repiqués pour être analysés mais seulement 11 ont poussé en milieu liquide : 1 seul issu de la condition de sélection avec les trois antibiotiques (cl 13.1) et 10 issus de la condition de sélection avec l'érythromycine seule (cl 11.1 à 11.10). La présence du pDR244 Ery a été contrôlée par PCR et a pu être validée chez 10 clones. Pour 3 d'entre eux (cl 11.1, 11.13 et 13.1), la présence des deux sites lox de part et d'autre de la région chromosomique, correspondant au fragment 7.2, a également été vérifiée et validée. Une analyse par séquençage Sanger au niveau du gène *cre* du pDR244 Ery hébergé par ces clones a permis de confirmer l'absence de mutation dans ce gène. La cellule receveuse FE220 V3 + pDR244 Ery pré-exprimant la CRE sur un plasmide réplcatif était disponible pour l'étape suivante.

#### 6.3.1. Quatrième tentative d'échange de fragments avec une cellule receveuse pré-exprimant la CRE sur le pDR244 Ery

Pour la quatrième tentative d'échange de fragments, la cellule receveuse FE220 V3 + pDR244 Ery a été transformée avec les molécules donneuses purifiées issues des deux clones *E. coli* (cf partie 3 et figure 3 du chapitre 3). La transformation a été réalisée à 30°C avec le protocole simple MC. Les transformants ont été sélectionnés sur milieu LB solide contenant 100 µg/ml de spectinomycine et 1 µg/ml d'érythromycine (dilution jusqu'à 10<sup>-3</sup>) (tableau 26).

Après 24h d'incubation à 30°C, aucune colonie n'est apparue sur le contrôle négatif alors que plusieurs centaines de colonies étaient visibles pour les conditions de transformation avec les molécules donneuses. Soixante clones isolés ont été piqués puis analysés par PCR simplex à l'aide de 4 couples d'amorces situés aux bornes du fragment 7.2.

La PCR 1, identique à celle décrite paragraphe 4.2, a permis de contrôler l'incorporation du gène *specR* et la perte de *tetR* (figure 100A). A notre grande surprise, 59 clones sur 60 analysés présentèrent une bande unique à 1729 pb (figure 100B) suggérant que la

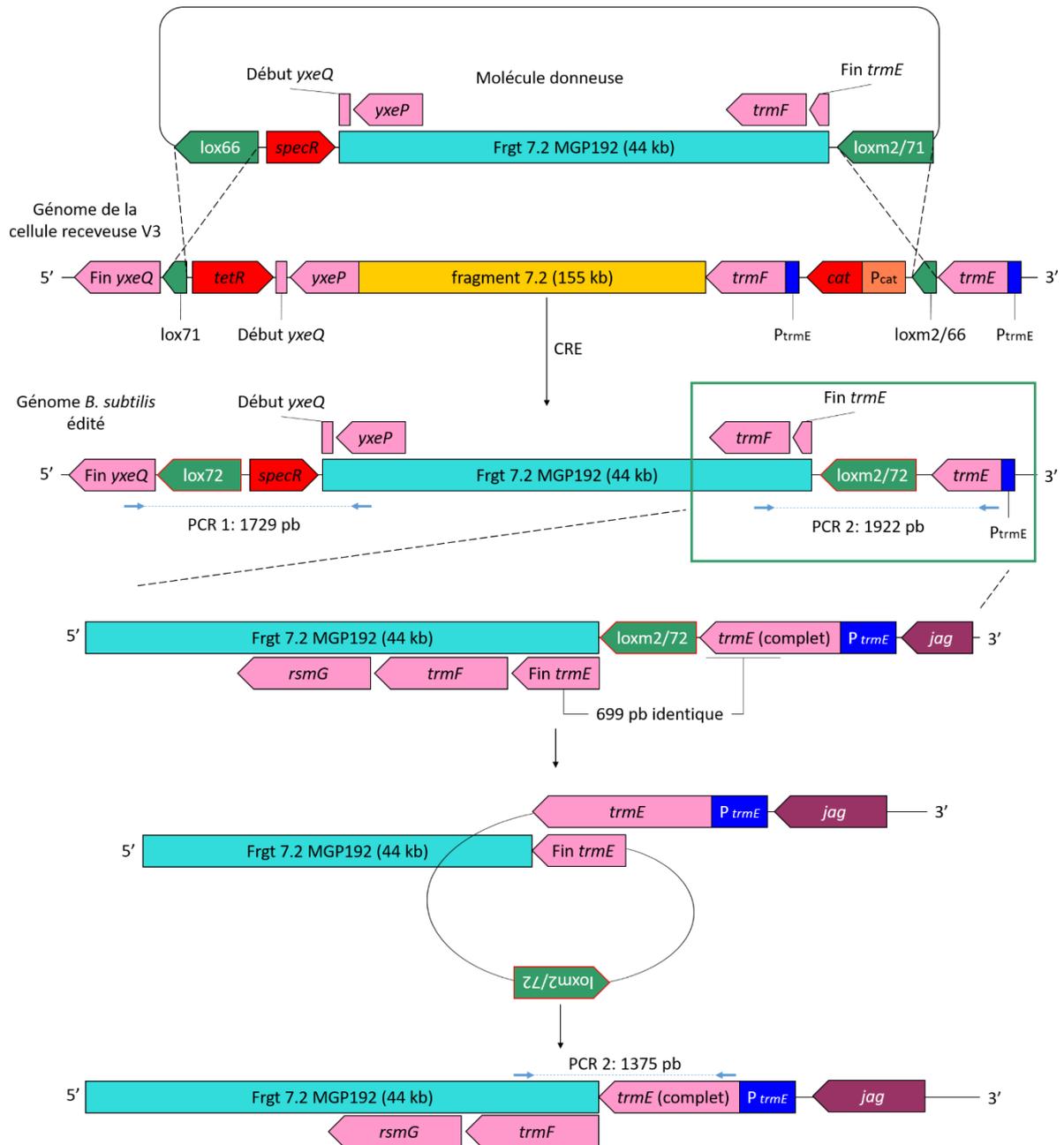


**Figure 101: Analyse des transformants de la quatrième expérience avec la PCR 2. [A]** Schéma du résultat de l'échange de fragment. **[B]** Analyse des transformants par la PCR 2 avec les amorces Verif receveuse F et Seq echange CmR R. En cas d'échange de fragment, la PCR 2 produit un amplicon de 1922 « M » : DNA Ladder 1 kb+ Invitrogen (100-12000 pb). « H2O » : contrôle négatif sans ADN. T-1 : témoin négatif plasmide pTwist pTwist CmR loxm2/66 V3. T-2 : témoin négatif ADNg de la cellule receveuse + pDR244 Ery. T-3 : témoin négatif molécule donneuse issue du clone *E. coli* 4.8.3.

recombinaison site-spécifique entre le site lox71 de la cellule receveuse et le site lox66 de la molécule donneuse avait eu lieu dans la quasi-totalité des clones analysés.

Nous avons donc poursuivi l'analyse des 59 clones avec la PCR 2 permettant de contrôler la perte du gène de résistance au chloramphénicol *cat* (figure 101A). Elle a été réalisée de la même manière que lors du troisième essai (cf partie 6.2 du chapitre 3). Sur les 59 clones analysés, (i) 36 présentaient un amplicon PCR à 2768 pb identique au témoin T-2 (ADNg cellule receveuse FE220 V3) témoignant de l'absence d'évènement de recombinaison site-spécifique à ce locus chez ces clones ; (ii) un clone le 7.10 n'a montré aucune bande ; (iii) 19 montraient un profil à deux bandes, la première à 1922 pb correspondant à la bande attendue après recombinaison entre les sites loxm2/66 et loxm2/71 et la deuxième à 1375 pb correspondant étonnamment à la souche FE220 WT ; (iv) un clone (le clone 6.8) montrait uniquement la bande à 1375 pb correspondant à la souche FE220 WT et enfin (v) deux clones ont présenté trois bandes au taille de 2768 pb, 1922 pb et 1375 pb attestant d'un mélange de population (figure 101B). Après réflexion, nous nous sommes rendus compte que la bande à 1375 pb, pouvait être la résultante de deux phénomènes illustrés figure 102 et 103 et repris ci-après:

- 1) La recombinaison site-spécifique entre les sites lox de la molécule donneuse et de la cellule receveuse aboutit, comme attendu, au remplacement du fragment 7.2 de la cellule receveuse (155 kb) par celui de la molécule donneuse (44 kb). Ce remplacement fait apparaître dans le génome des transformants une répétition directe en tandem de 699 pb dans laquelle est enchâssée un site loxm2/72 non fonctionnel. Cette répétition est due à la présence sur l'extrémité 3' du fragment 7.2 de la molécule donneuse, de la fin du gène *trmE* (699 pb sur 1380 pb total) (figure 102). Nous avons émis l'hypothèse que la présence de cette duplication stimulait la machinerie de recombinaison homologue de la cellule et entraînait la formation d'une boucle qui permettait d'une part, l'éjection du site loxm2/72 et d'autre part, un retour à la séquence sauvage (*i.e.* celle de la souche FE220 WT), expliquant la présence d'un amplicon à 1375 pb seul ou en mélange avec un amplicon à 1922 pb (figure 102).
- 2) La recombinaison site-spécifique entre le site lox71 de la molécule donneuse et le lox66 de la cellule receveuse a lieu au niveau de la borne 5', aboutissant au site lox72 non fonctionnel. En revanche, la recombinaison site-spécifique entre le site loxm2/71 de la molécule donneuse et le loxm2/66 de la cellule receveuse n'a pas lieu, et est remplacée par un évènement de recombinaison homologue entre la fin du gène *trmE* présent sur la



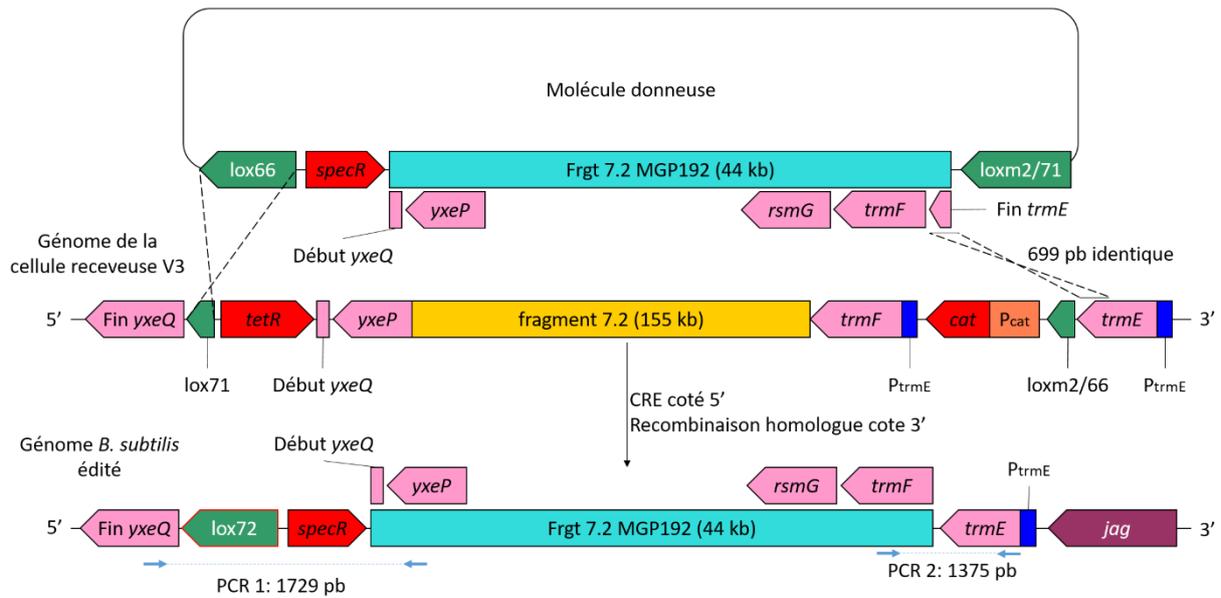
**Figure 102 : Premier scénario de recombinaison homologue après une étape de recombinaison site spécifique cre/lox.** Dans ce scénario, l'étape d'échange de fragment a bien lieu par l'intermédiaire du système cre/lox. Cet échange fait apparaître une répétition directe de 699 pb qui borde le site *loxm2/72* inactif. Cette répétition directe est due à la présence sur l'extrémité 3' du fragment 7.2 de la molécule donneuse de la fin du gène *trmE*. La machinerie de recombinaison homologue stimulée par cette répétition directe, forme une boucle qui permet l'élimination du site *loxm2/72* et le retour à la séquence sauvage.

molécule donneuse et le gène *trmE* présent dans le génome (figure 103). Ce scénario aboutit tout comme le scénario 1) à l'élimination du site *loxm2/72* et un retour à la séquence sauvage (*i.e.* celle de la souche FE220 WT) (figure 103).

Quel que soit le scénario pour la bande à 1375 pb, nos analyses semblaient quand même indiquer que nous détectons pour la première fois des recombinaisons via le système *cre/lox* des deux côtés de la région à échanger. Du côté de la PCR 1, la quasi-totalité des clones testés semblait avoir recombinaison correctement, témoignant d'une très haute fréquence de recombinaison site-spécifique à ce locus. Cette efficacité, quelque peu étonnante, était sans doute à relativiser. En effet, la recombinaison *lox71-lox66* est couplée, à cette borne, à l'intégration d'un marqueur de sélection, le marqueur *specR* (en remplacement du marqueur *tetR*). La présence de spectinomycine 100 µg/ml dans les milieux de culture (solide et liquide) après transformation constitue certainement une telle pression que seuls les clones ayant effectué la recombinaison site-spécifique attendue sont capables de se multiplier. Du côté de la PCR 2, en revanche, aucun marqueur n'est apporté par la molécule donneuse et aucune pression de sélection ne peut donc être appliquée. Malgré cela, 19, voire 20 clones (*i.e.* à minima ceux présentant la bande à 1922 pb) sur 60 analysés ont montré un profil pouvant correspondre à la recombinaison site-spécifique désirée.

Avant de mettre en place un PFGE et estimer la taille des génomes dans les clones d'intérêt, nous avons souhaité affiner ces premiers résultats en réalisant d'autres tests. Le fait que le fragment 7.2 de 44 kb porté par la molécule donneuse contiennent deux grandes délétions par rapport au fragment 7.2 de 155 kb porté par la cellule receveuse FE220 V3 (figure 104A), nous a permis de mettre en place deux autres PCR (PCR 3 et PCR 4) que nous avons appelé « PCR longues » car les amplicons susceptibles d'être produits ont des tailles respectivement de 8.8 kb pour la borne 5' et 17.3 kb pour la borne 3'. Comme illustrés sur la figure 104A, ces amplicons ne peuvent être générés que si l'échange de fragments entre molécule donneuse et cellule receveuse a eu lieu, les amorces sens et anti-sens ayant été placées sur des *loci* trop éloignés dans la séquence originale (*i.e.* dans le fragment 7.2 la cellule receveuse FE220 V3, figure 104A).

Pour la PCR 3, située au niveau de la borne 5', 14 clones sur les 20 testés ont montré la bande attendue à 8,8 kb, venant corroborer les résultats de la PCR 1, et l'hypothèse selon

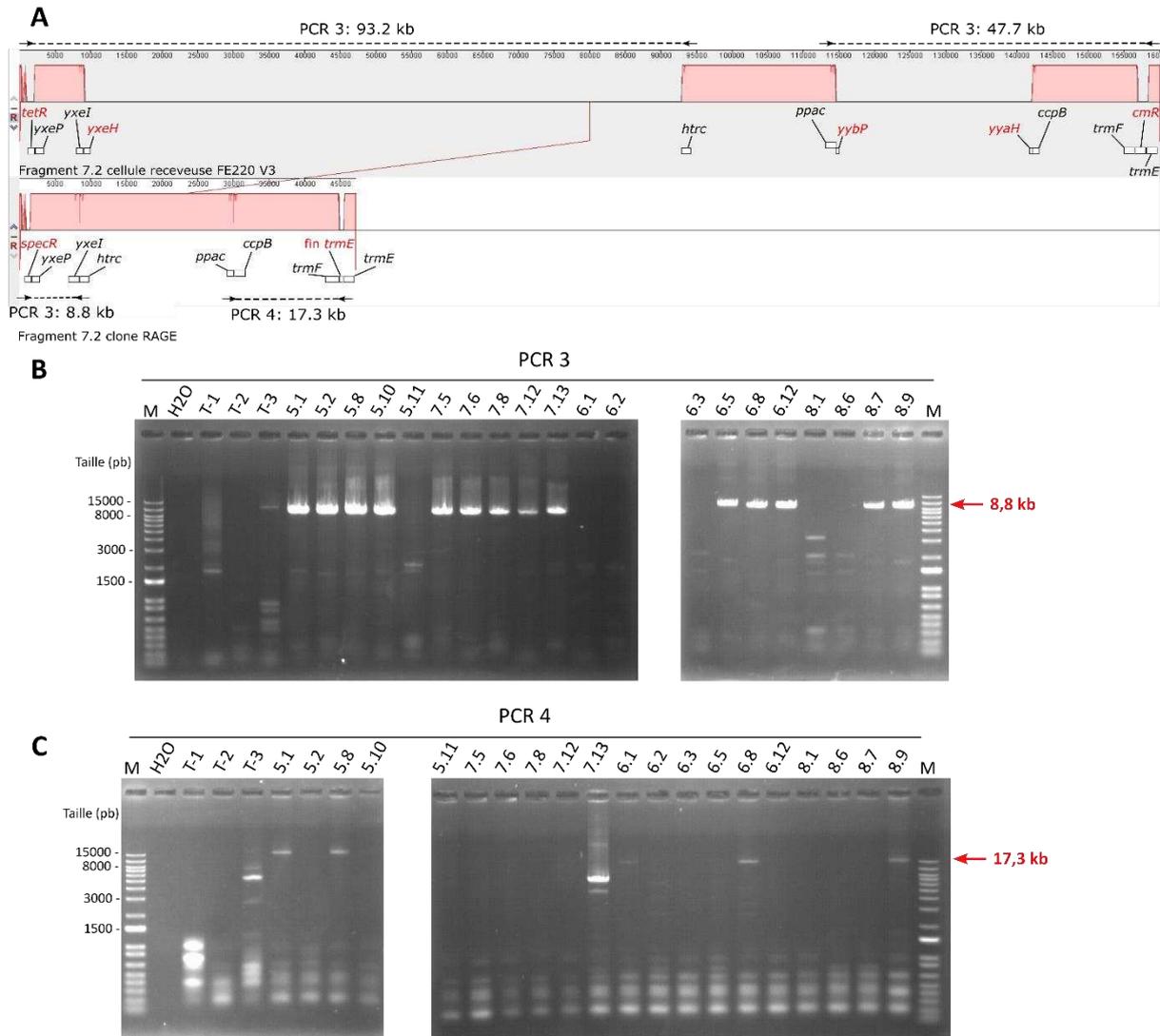


**Figure 103 : Second scénario : recombinaison site spécifique cre/lox du côté 5' et recombinaison homologue du côté 3'.** Dans ce scénario, l'échange de fragment de génome est effectué par un évènement de recombinaison site spécifique cre/lox du côté 5' entre le *lox71* de la molécule donneuse et le *lox66* de la cellule receveuse aboutissant au site inactif *lox72*. Du côté 3', la recombinaison site spécifique cre/lox est remplacée par un évènement de recombinaison homologue entre la fin du gène *trmE* sur la molécule donneuse et le gène *trmE* de la cellule receveuse.

laquelle les sites lox71 et lox66 avaient recombines à ce site (figure 104B). Pour la PCR 4, située au niveau de la borne 3', 5 clones sur les 20 testés ont présenté la bande attendue à 17,3 kb, venant ici supporter les résultats de la PCR 2, et l'hypothèse selon laquelle les sites loxm2/71 et loxm2/66 avaient recombines (figure 104C). L'amplification de 17,3 kb étant techniquement difficile à mettre en place, il est possible que des clones présentant la bonne configuration n'aient pas été détectés.

En croisant l'ensemble des résultats PCR (PCR 1 à 4), 4 clones ont finalement présenté tous les profils attendus en PCR (cl 5.1, 5.8, 6.8 et 8.9). Ces 4 clones ont été analysés en PFGE après hydrolyse enzymatique de leur génome par l'enzyme SgrDI. En théorie, un profil à 10 bandes (de 39 kb à 1,864 Mb) est attendu sur gel d'agarose après digestion des ADN, migration et coloration au Sybr Gold. Ce profil est identique au profil attendu pour les souches contrôles FE220 WT et FE220 V3, excepté pour le fragment SgrDI incluant le fragment 7.2 qui a une taille d'environ 580 kb dans les souches contrôles, et d'environ 469 kb (111 kb de moins) dans les transformants. Sur les 4 clones analysés, 3 (cl 5.1, 5.8 et 6.8) ont montré cette différence de taille au niveau du morceau d'intérêt, et le dernier a présenté le profil correspondant à celui de la cellule receveuse FE220 V3 avec une bande à 580 kb (figure 105). Les résultats de cette expérience, nous ont amené à conclure que l'échange de fragments via le système cre/lox (ou système cre/lox et recombinaison homologue) semblait avoir eu lieu dans au moins 3 clones.

Afin de valider la robustesse de la méthode RAGE pour l'échange de grands fragments de génome chez *B. subtilis*, l'étape 4 du RAGE a été répétée à l'identique. Après transformation, 40 clones ont été analysés et 4 (cl 6.3, 6.7, 7.10 et 8.5) ont montré les profils attendus pour les PCR 1 à 4 (figure 106A). Parmi les 4 clones, le 7.10 présentait un profil à une bande de 1375 pb pour la PCR 2 (équivalent au clone 6.8 présenté figure 101), pouvant être le résultat d'un des deux scénarios décrits précédemment. Ces 4 clones ont été analysés par séquençage Sanger au niveau des sites lox. Du côté 3', trois d'entre eux ont présenté la séquence correspondant au site lox recombine, le loxm2/72. Le dernier (cl 7.10) a lui montré une absence de site lox, ce qui nous paraissait cohérent avec les résultats de la PCR 2 (bande unique à 1375 pb). Du côté 5' deux clones ont présenté les séquences du site lox recombine c'est-à-dire le lox72. Pour les deux autres clones, les séquences étant de mauvaise qualité, les résultats n'ont pas pu être interprétés. L'analyse PFGE a montré que sur ces 4 clones, 3 présentaient le profil à 10 bandes attendu avec la bande à 469 kb (au lieu de 580 kb dans les



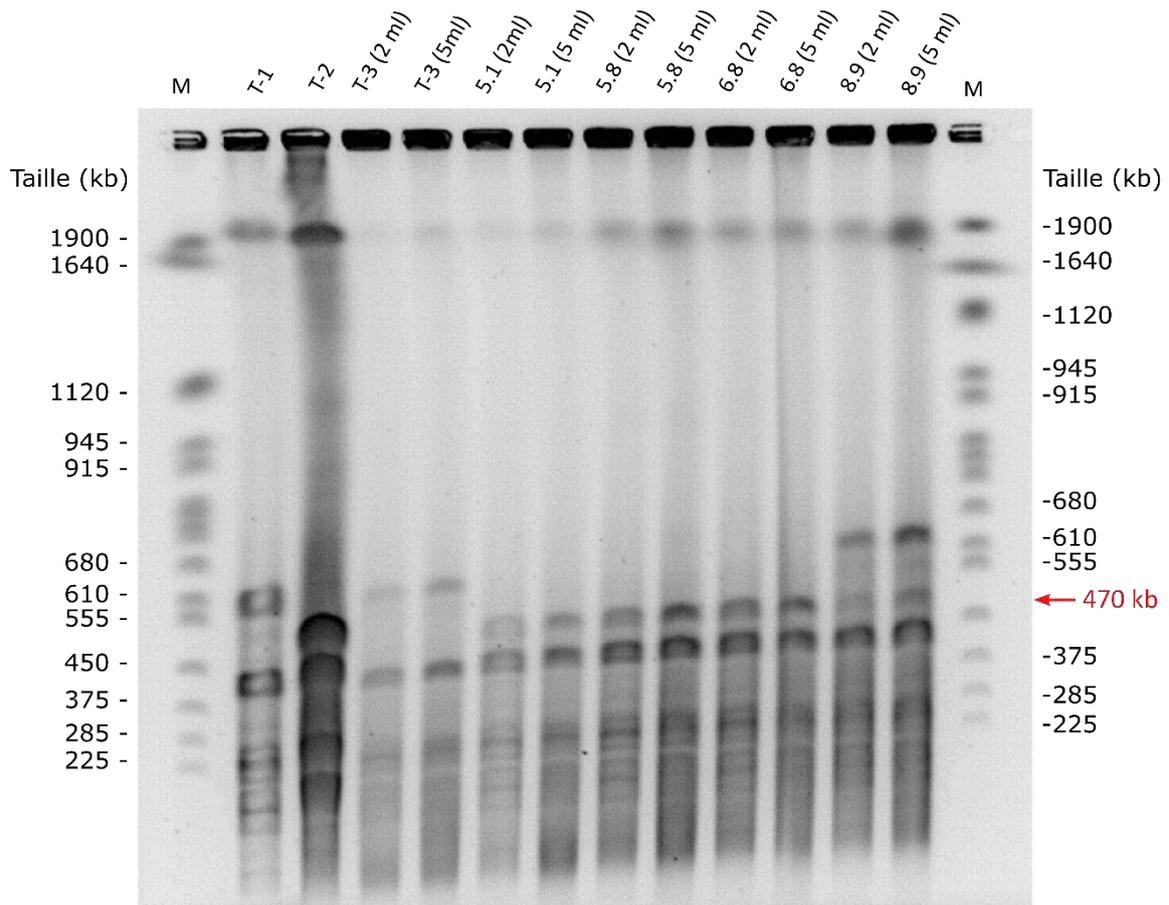
**Figure 104 : Analyse des transformants de la quatrième expérience d'échange de fragment avec les PCR 3 et 4. [A]** Alignement entre le fragment 7.2 de *B. subtilis* FE220 (155 kb) et celui de *B. subtilis* MGP192 (44 kb). Les PCR 3 et 4 sont représentées. **[B]** Après une première analyse avec les PCR 1 et 2, les transformants sont analysés avec les PCR 3 et 4 dites longues, permettant une amplification qu'en cas d'échange de fragment. La PCR 3 avec les amorces FE220 receveuse YxeQ F et Verif\_htrc\_R permet de contrôler l'extrémité 5' du locus cible et produit un amplicon de 8,8 kb. **[C]** La PCR 4 avec les amorces ppaC\_F et trme\_frgt1\_bamhi\_R permet de contrôler l'extrémité 3' du locus cible et produit un amplicon de 17,3 kb pour la PCR 3. « M » : DNA Ladder 1 kb+ Invitrogen (100-12000 pb). « H2O » : contrôle négatif sans ADN. T-1 : témoin négatif ADNg de la souche *B. subtilis* FE220. T-2 : témoin négatif ADNg de la cellule receveuse V3 + pDR244 Ery. T-3 : témoin négatif molécule donneuse issue du clone *E. coli* 4.8.3.

souches contrôles) (figure 106B). Cette deuxième expérience a permis de valider notre capacité à induire un échange de fragments via le système cre/lox chez *B. subtilis*, démontrant la robustesse de la méthode RAGE.

Afin de confirmer les résultats de ces deux expériences, mais aussi pour contrôler l'absence de grands remaniements, d'indels ou encore de SNPs, la dernière analyse a consisté à faire le séquençage complet des génomes des 3 clones obtenus lors de la dernière expérience réalisée (cl 7.10, 6.7 et 8.5). Ces 3 génomes ont donc été séquencés en long read (Oxford Nanopore technologies) par la société prestataire PLASMIDSAURUS, ainsi qu'en short read (Illumina) par nos collaborateurs de l'équipe SyBER (Institut MICALIS, UMR 1319 INRAE / AgroParisTech, Jouy-en-Josas). Les analyses des séquences ont confirmé :

- la présence des sites lox72 (résultant de la recombinaison du lox71 et du lox66) et loxm2/72 (résultant de la recombinaison du loxm2/66 et du loxm2/71) de part et d'autre de la zone à échanger pour les clones cl 6.7 et cl 8.5.
- la présence du site lox72 à la borne 5' de la zone à échanger dans le clone 7.10, ainsi que la disparition du site loxm2/72, et un retour à une séquence de type sauvage à la borne 3' de ce clone ; un résultat attendu et en adéquation avec les résultats de la PCR 2 (bande unique à 1375 pb pour le clone 7.10). Comme mentionné plus haut, deux scénarios peuvent expliquer ce résultat (recombinaison site-spécifique au niveau des deux bornes ou recombinaison site-spécifique à une borne et recombinaison homologue à l'autre borne). Cependant aujourd'hui, nous ne sommes pas capables de dire si la séquence qui a été déterminée pour le clone 7.10 est le résultat du scénario 1 ou 2.
- la présence du fragment 7.2 issu de la souche MGP192 chez les clones 6.7.

L'ensemble de ces résultats ont montré pour la première fois que le fragment 7.2 de la cellule receveuse d'une taille de 155 kb avait été remplacé par le fragment 7.2 de la souche MGP192 d'une taille de 44 kb (apporté dans la cellule par la molécule donneuse). La méthode RAGE nous semblait donc enfin fonctionner chez *B. subtilis* pour remplacer de grands fragments de génome. Cependant, après une analyse plus approfondie des séquences, effectuée avec l'aide de nos collaborateurs, nous nous sommes rendus compte qu'une copie du plasmide pDR244 Ery (d'autres étant libres) semblait être intégrée au génome et que son



**Figure 105 : Analyse en électrophorèse en champ pulsé des clones issus de la quatrième expérience d'échange de fragments via le système cre/lox.** La digestion du génome de la souche *B. subtilis* FE220 par l'enzyme SgrDI produit 10 morceaux. Celui qui contient le fragment 7.2 à une taille initiale de 580 kb. Pour les clones dans lesquels l'échange a eu lieu, ce morceau est réduit de 111 kb et a une taille de 469 kb. La digestion du génome de la souche MGP192 par la même enzyme produit un morceau de 482 kb et a servi de contrôle de taille pour ce PFGE. « M »: *S. cerevisiae* chromosomal DNA ladder Bio-Rad (225-2200 kpb). T-1: témoin négatif digestion de l'ADNg de la souche FE220. T-2: témoin négatif digestion de l'ADNg de la souche MGP192. T-3 : digestion de l'ADNg de la cellule receveuse V3.

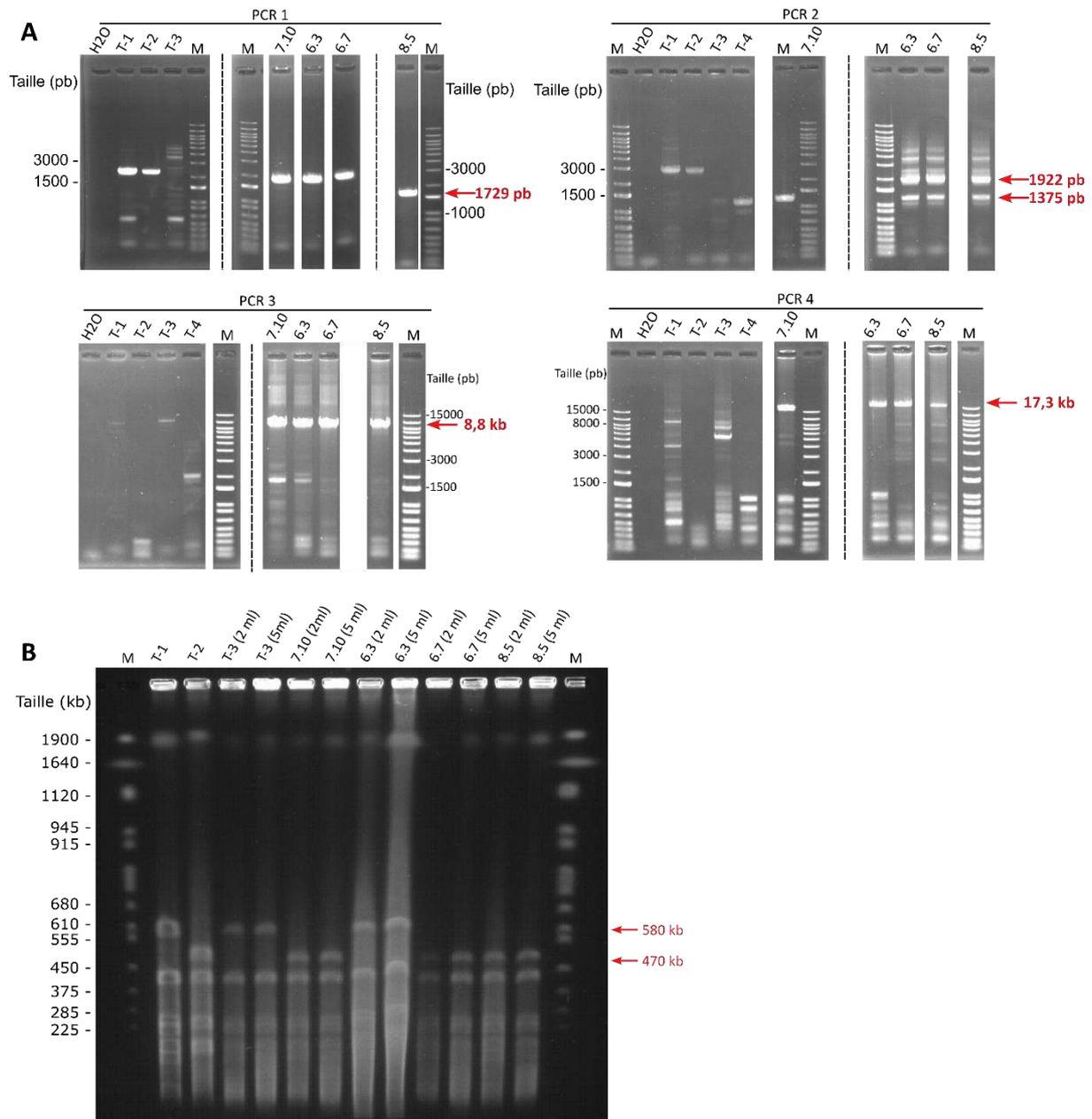
insertion avait entraîné une délétion de 42 kb entre les gènes *ppsA* et *yngD* (le reste du génome semblant intact). Ces observations nous ont surpris car ce plasmide est censé être un plasmide répliatif et donc doit se maintenir libre dans la cellule hôte.

Etonnamment, le lieu d'insertion et la zone délétée (non essentielle (Dervyn *et al.*, 2023)) était identique pour tous les clones séquencés. Nous nous sommes donc demandés si cette insertion/délétion se produisait pendant l'étape 4 du RAGE ou bien en amont. Le génome de la cellule receveuse FE220 V3 ainsi que celui de la cellule receveuse FE220 V3 + pDR244 Ery ont donc été envoyés chez PLASMIDSAURUS pour effectuer le séquençage long-read de leur génome. L'analyse des séquences a montré que la zone entre les gènes *ppsA* et *yngD* était intacte dans la souche FE220 V3, donc avant ajout du pDR244 Ery. En revanche, elle était absente et remplacée par une copie du pDR244 Ery dans la cellule receveuse V3 + pDR244 Ery.

Afin de déterminer si, par malchance, nous avons choisi pour l'étape 4 du RAGE, le seul clone FE220 V3 + pDR244 Ery (cl 13.1) présentant cette insertion, le génome d'un autre clone (cl 11.1) a été séquencé en long-read. Les analyses ont montré que dans ce clone, le pDR244 Ery était également intégré au génome, entraînant cette fois une délétion un peu plus large de 68 kb (entre les gènes *ftsR* et *ynfC*), incluant la zone de 42kb.

Pour définir si le pDR244 Ery ne pouvait s'insérer qu'à ce locus, le génome de la souche MS loxm2/66 TetR lox71 + pDR244 Ery utilisée pour les expériences d'échange de marqueur via le système cre/lox (cf partie 5.1 du chapitre 3) a aussi été séquencé. Dans cette souche, le plasmide semblait aussi être intégré au génome, mais en aval du gène *yxeP*, c'est-à-dire non loin de la cassette loxm2/66 TetR lox71 (insérée entre les gènes *yxeP* et *yxeQ*). Le locus entre les gènes *ppsA* et *yngD* était intact. L'ensemble de ces analyses semblait donc montrer que des copies du plasmide pDR244 Ery (partielles ou complètes) s'inséraient au génome de façon quasi systématique, mais pas toujours au même locus.

Suite à ces analyses, il devenait évident que le pDR244 Ery posait problème et il nous était impossible de valider la méthode RAGE dans de telles conditions.



**Figure 106 : Analyse des quatre clones sélectionnés lors de l'expérience de répétition de la tentative d'échange de fragment de génome via cre/lox chez *B. subtilis* lorsque la CRE est pré-exprimée par le pDR244 Ery. [A]** En cas d'échange de fragments, les PCR 1, 2, 3, et 4 produisent respectivement des bandes de 1729 pb, 1922 pb ou 1375 pb, 8,8 kb et 17,3 kb. « M » : DNA Ladder 1 kb+ Invitrogen (100-12000 pb). « H2O » : contrôle négatif sans ADN. T-1 : témoin négatif plasmide pTwist TetR lox1 (pour les PCR 1 et 3) pTwist CmR loxm2/66 V3 (pour les PCR 2 et 4). T-2 : témoin négatif ADNg de la cellule receveuse V3 + pDR244 Ery. T-3 : témoin négatif molécule donneuse issue du clone *E. coli* 4.8.3. T-4 : témoin négatif ADNg FE220 sauvage. **[B]** Analyse en électrophorèse en champ pulsé. « M » : *S. cerevisiae* chromosomal DNA ladder Bio-Rad (225-2200 kpb). T-1: témoin négatif digestion de l'ADNg de la souche FE220. T-2: témoin négatif digestion de l'ADNg de la souche MGP192. T-3 : digestion de l'ADNg de la cellule receveuse V3.

### 6.3.2. Plasmides répliatifs chez *B. subtilis*

Au vu des expériences précédentes, il semblait évident que la cellule receveuse doit exprimer le gène *cre* à partir d'un plasmide répliatif pour que l'étape 4 du RAGE fonctionne. C'était d'ailleurs ce que nous voulions mettre en place lorsque nous avons commencé à travailler avec le plasmide pDR244 Ery, décrit comme un plasmide répliatif dans la littérature. Seulement, contre toute attente, le plasmide ne s'est pas comporté comme tel au cours de nos expériences, des copies du plasmide s'insérant de façon aléatoire dans le génome de *B. subtilis*.

A partir de là, nous avons cherché dans la littérature et auprès de la communauté d'autres plasmides répliatifs chez *B. subtilis*. Nous avons identifié le plasmide pJIM portant le marqueur de résistance à l'érythromycine (comme le pDR244 Ery) et nous avons commencé, avant toute autre chose, par vérifier si ce plasmide se répliquait de façon stable chez *B. subtilis*.

La cellule receveuse FE220 V3 a donc été transformée par le plasmide pJIM avec le protocole de transformation simple MC. Nous avons profité de cette expérience pour inclure un plasmide ayant une origine commune avec le pDR244 Ery : le plasmide pDR244  $\Delta cre$ . Ce plasmide est, en effet, basé sur le pDR244, il porte donc le gène de résistance à la spectinomycine (au lieu de l'érythromycine chez le pDR244 Ery) et est dépourvu du gène *cre*. Notre idée ici était de s'assurer que le comportement du pDR244 Ery n'était pas dû à la présence du gène *cre* sur le plasmide. Les transformants ont été sélectionnés sur milieu LB solide contenant 1  $\mu\text{g/ml}$  d'érythromycine pour les conditions avec le pJIM et 100  $\mu\text{g/ml}$  de spectinomycine pour celles avec le pDR244  $\Delta cre$ . Après sélection, la présence des plasmides a été contrôlée par des PCR. Deux clones, l'un hébergeant le pJIM et l'autre le pDR244  $\Delta cre$  ont été sélectionnés, puis les ADN ont été séquencés en long read.

L'analyse des séquences a montré pour le clone hébergeant le pJIM, que le génome était intact et que le plasmide était lui aussi présent entier et libre dans la cellule. Concernant le clone hébergeant le pDR244  $\Delta cre$ , le génome était également intact mais le plasmide semblait absent. Aucun « read » ne s'alignait sur le plasmide. Afin de s'assurer de la présence ou de l'absence du plasmide, cette étude sera refaite avec le pDR244  $\Delta cre$ .

Cette expérience nous a permis de vérifier que le pJIM se comportait différemment du pDR244 Ery, et ne s'insérait pas dans le génome de la cellule receveuse FE220 V3. Ce plasmide restait libre dans le cytoplasme et se répliquait de manière stable, il pouvait donc

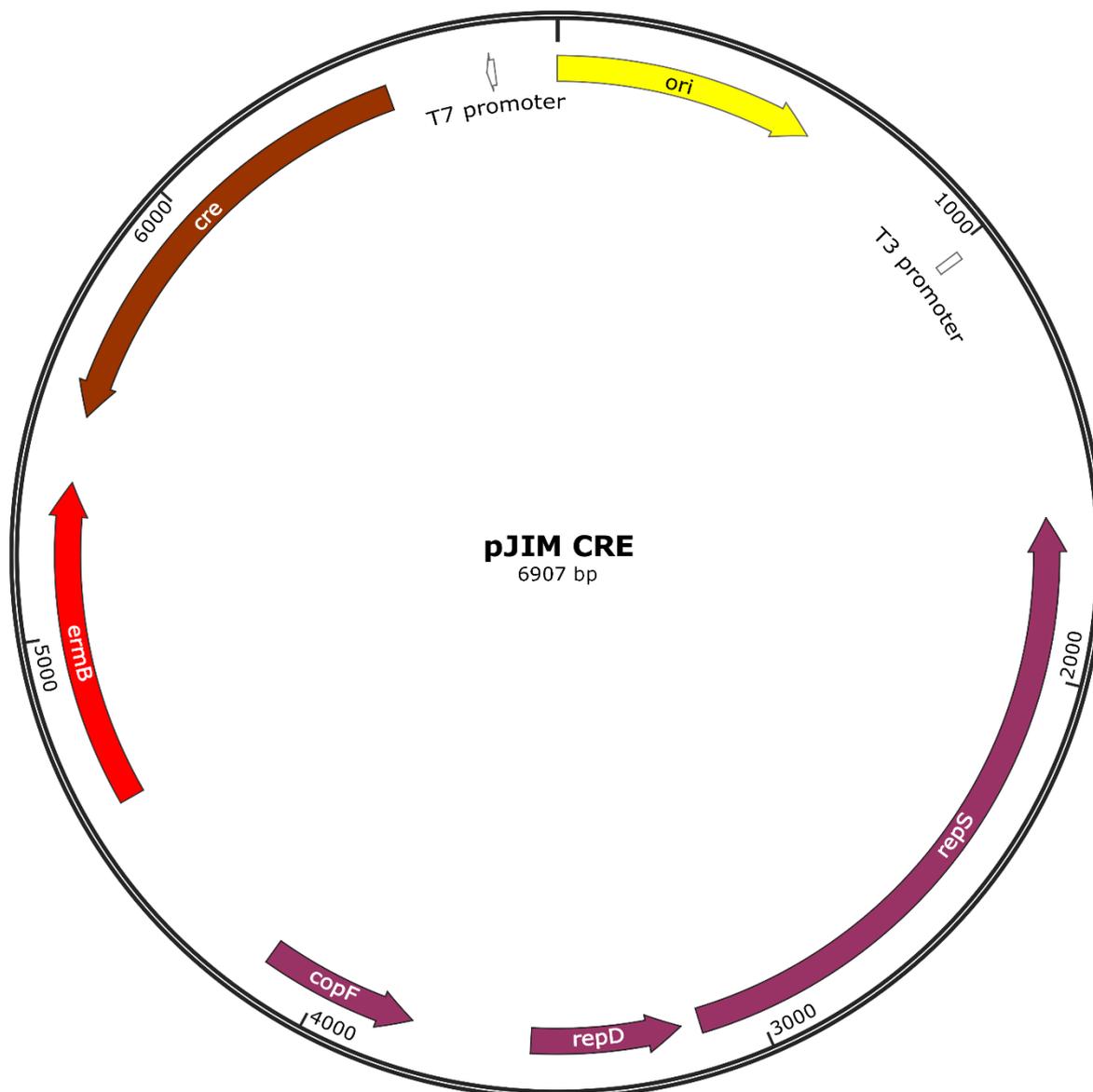


Figure 107 : Carte du plasmide pJIM CRE

être utilisé dans le cadre de la méthode RAGE pour pré-exprimer la CRE dans la cellule receveuse.

#### 6.4. Pré expression de la CRE avec le plasmide répliatif le pJIM CRE

Lors des analyses réalisées dans la section précédente, nous avons conclu que le pJIM était bien répliatif et capable de se maintenir chez *B. subtilis*. Nous avons donc choisi de l'utiliser pour tenter de pré-exprimer la CRE dans la cellule receveuse FE220 V3. Pour ce faire, nous avons transformé le plasmide pJIM CRE qui correspond au plasmide pJIM dans lequel le gène *cre* a été inséré (figure 107), dans la cellule receveuse FE220 V3 avec le protocole simple MC. Les transformants ont été sélectionnés sur milieu LB solide contenant soit 1 µg/ml d'érythromycine, 4 µg/ml de chloramphénicol et 7,5 µg/ml de tétracycline soit seulement 1 µg/ml d'érythromycine. Après 48 h d'incubation à 37°C, 35 colonies ont été obtenues sur milieu de sélection et 15 clones ont été analysés par PCR. La présence du plasmide a d'abord été vérifiée grâce à une PCR utilisant des amorces s'hybridant spécifiquement au pJIM CRE. L'amplicon a pu être détecté dans 13 clones sur 15 analysés. Puis, la présence des landing pads « TetR lox71 » et « CmR loxm2/66 » a été vérifiée sur 4 d'entre eux et validée des deux côtés.

Pour s'assurer que le plasmide ne s'était pas intégré au génome, l'ADN de 3 clones (cl 2.3, cl 3.3 et cl 5.3) a été extrait et séquencé en long-read par la société prestataire PLASMIDSAURUS. L'analyse des séquences a montré que (i) les génomes étaient intacts et que (ii) le pJIM CRE était bien présent sous forme ectopique.

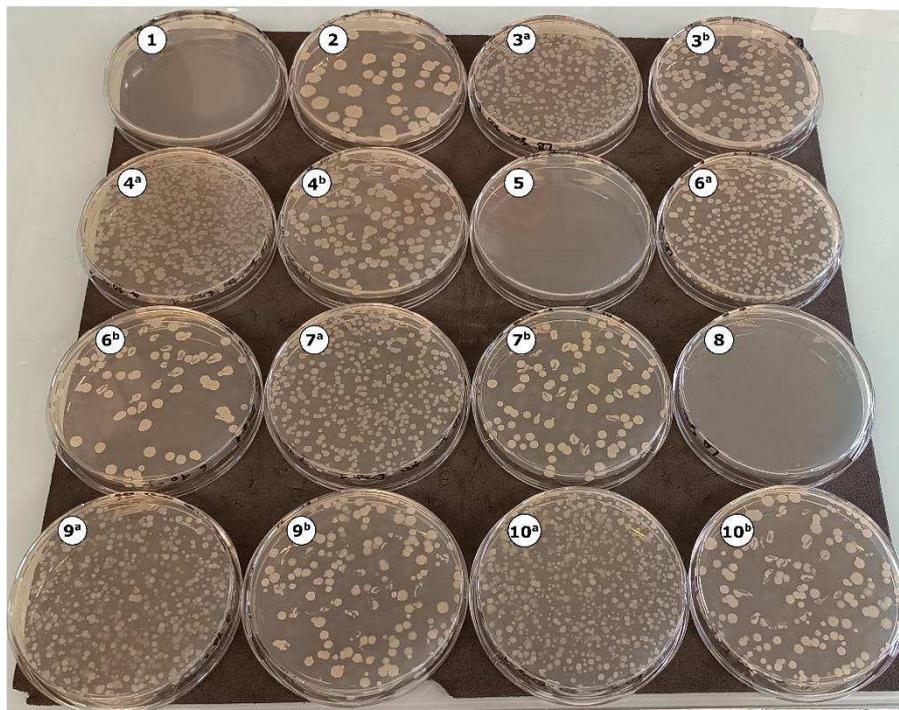
L'ensemble de ces résultats a montré que nous avons trois nouvelles **cellules receveuses FE220 V3 + pJIM CRE** qui hébergeait un plasmide répliatif portant le gène *cre* et dont les génomes étaient intacts. Ces clones pouvaient donc être utilisés pour des essais d'échange de fragments de génome de *B. subtilis* via le système *cre/lox*.

#### 6.5. Cinquième tentative d'échange de fragments avec une cellule receveuse hébergeant le plasmide répliatif pJIM CRE

Dans la section précédente, nous avons obtenu la cellule receveuse FE220 V3 + pJIM CRE (3 clones différents ont été validés : cl 2.3, cl 3.3 et cl 5.3). Pour ce cinquième essai, nous avons décidé de conserver les 3 clones dans le but de maximiser nos chances de réussite. Ils ont donc été transformés par les molécules donneuses pCC1-LYS-cre-specR frgtBsu 7.2 purifiées, issues

**Tableau 27 : Conditions expérimentales du cinquième essai d'échange de fragments de génome via le système cre/lox chez *B. subtilis* lorsque la CRE est pré-exprimée par le pJIM CRE.**

Conditions	Souche receveuse	ADN transformé	Sélection	Clones positifs / Clones analysés		
				PCR	PCR longues	Séquençage
1	cellule receveuse V3 + pJIM CRE cl 2.3	Sans ADN	LB spec 100 µg/ml + ery 1 µg/ml	-	-	-
2		pDR111		-	-	-
3		Molécule donneuse clone 4.8.3		0/20	-	-
4		Molécule donneuse clone 16.2.3		0/20	-	-
5	cellule receveuse V3 + pJIM CRE cl 3.3	Sans ADN		-	-	-
6		Molécule donneuse clone 4.8.3		0/20	-	-
7		Molécule donneuse clone 16.2.3		1/20	0/1	0/1
8		Sans ADN		-	-	-
9	cellule receveuse V3 + pJIM CRE cl 5.3	Molécule donneuse clone 4.8.3		0/20	-	-
10		Molécule donneuse clone 16.2.3		3/20	1/3	1/3

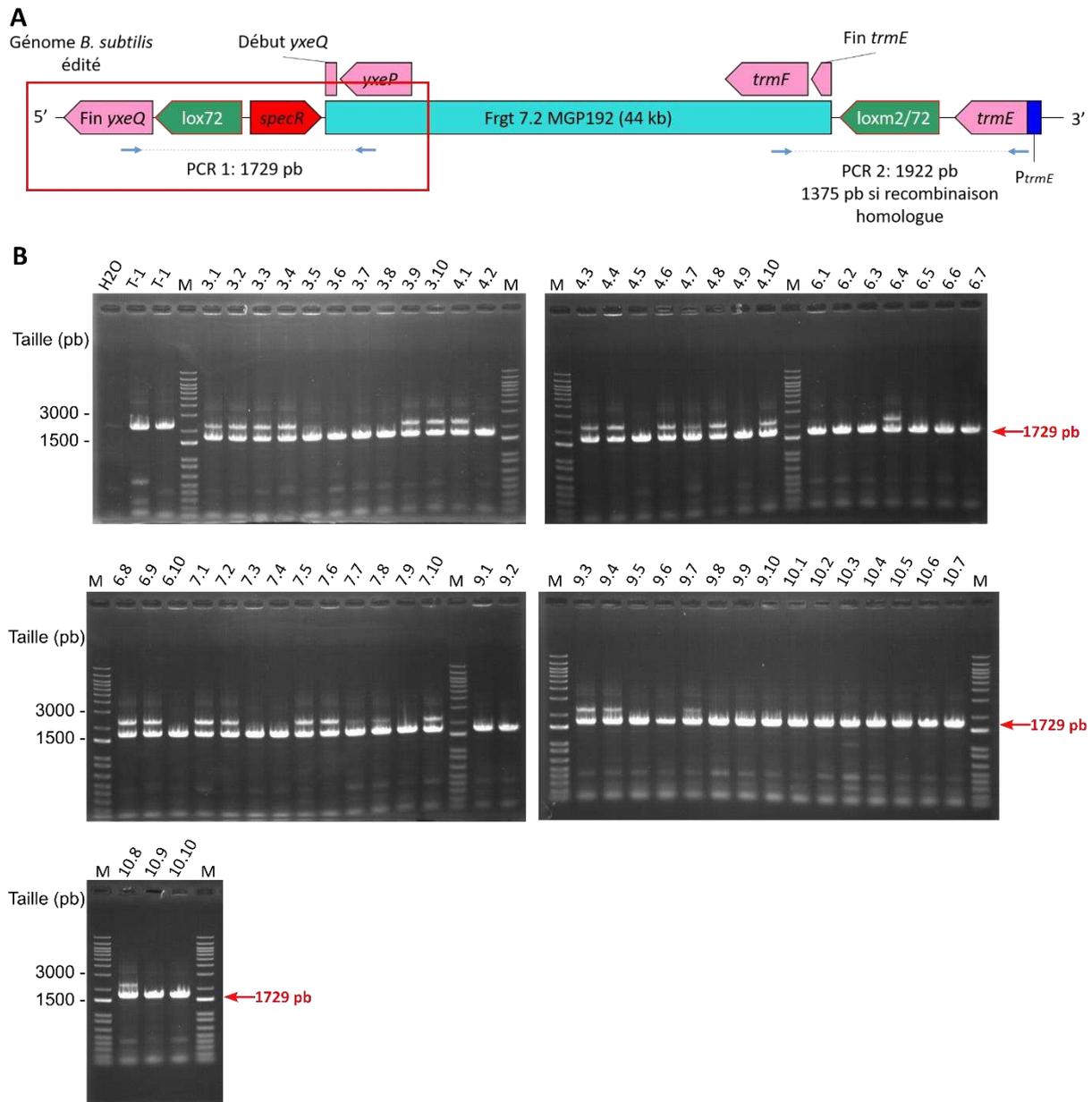


**Figure 108 : Photographie des boîtes de transformation de la cinquième expérience d'échange de fragment via le système cre/lox chez *B. subtilis* lorsque la CRE est pré-exprimée par le pJIM CRE. Les numéros correspondent aux conditions de transformations listées dans le tableau 27. <sup>a</sup> : étalement des dilution 10<sup>-2</sup>. <sup>b</sup> : étalement des dilution 10<sup>-3</sup>.**

des deux clones *E. coli* avec le protocole simple MC. Pour suivre les meilleures conditions définies précédemment, la transformation a été réalisée à 30°C et l'étape de culture liquide avant étalement n'a pas été ajoutée. Puis, comme pour les deux derniers essais, les échantillons ont été dilués jusqu'à  $10^{-3}$ , avant étalement sur milieu gélosé (tableau 27). Les transformants ont été sélectionnés sur milieu LB solide contenant 100 µg/ml de spectinomycine et 1 µg/ml d'érythromycine. Après 48 h d'incubation à 30°C, aucune colonie n'est apparue sur les contrôles négatifs. En revanche, pour les conditions de transformation avec les molécules donneuses, plusieurs centaines de colonies sont apparues sur milieu de sélection (figure 108). Un total de 60 clones, a été repiqué puis analysé de manière identique à la quatrième tentative : PCR 1, PCR 2, PCR 3, PCR 4, et séquençage complet de génome (cf partie 6.3.1 du chapitre 3).

Pour la PCR 1 (figure 109A), 33 clones sur les 60 analysés ont montré une bande à la taille de 1729 pb, correspondant au profil de recombinaison désiré entre le site *lox71* de la cellule receveuse et le site *lox66* de la molécule donneuse. Tous les autres ont montré un profil à 2 bandes aux tailles de 2285 pb et 1729 pb suggérant que ces clones n'étaient pas purs. A la différence des tentatives précédentes, la bande à 1729 pb était plus intense que celle à 2285 pb pour tous les clones (figure 109B). L'analyse a été poursuivie avec la mise en place de la PCR 2 (figure 110A) sur les 33 clones positifs et purs pour la PCR 1. Cette PCR a été réalisée de la même manière que lors des troisième et quatrième essais (cf partie 6.2 du chapitre 3). Sur les 33 clones analysés, 3 d'entre eux (cl 7.3, cl 10.2 et cl 10.4) ont montré un profil à deux bandes, la première à la taille de 1922 pb correspondant à la recombinaison entre les sites *loxm2/66* et *loxm2/71*, et la seconde à la taille de 1375 pb correspondant à un profil FE220 WT. Un quatrième clone, le cl 7.3, a montré un profil à une seule bande de la taille de 1375 pb (figure 110B). Suite à ces deux premières PCR, nous notons une nouvelle fois, que nous sélectionnions plus facilement des clones positifs du côté *lox71-lox66* dont la recombinaison est adjointe à l'intégration du marqueur *specR*, que du côté *loxm2/71-loxm2/66* dont la recombinaison n'est associée à aucun marqueur de sélection.

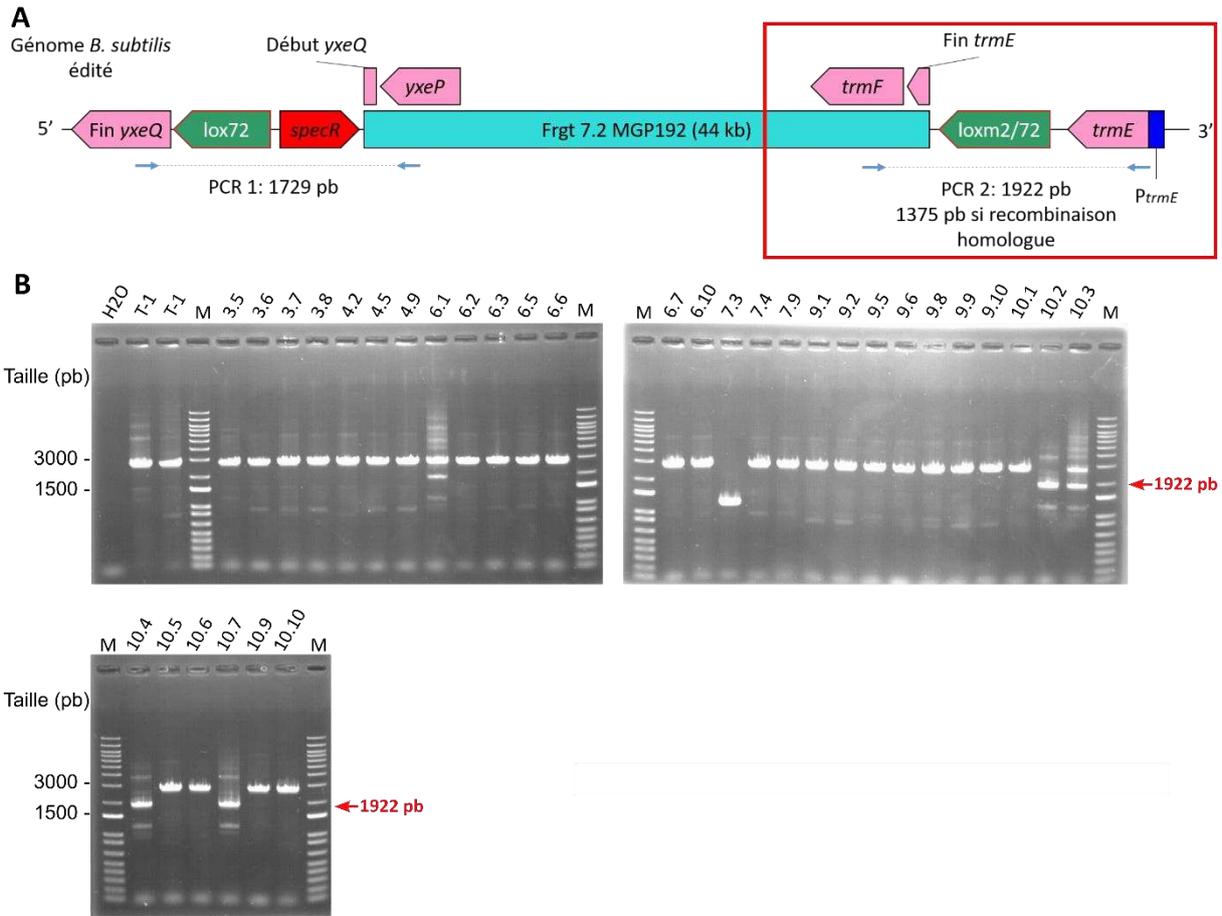
Les 4 clones ont ensuite été analysés avec les PCR dites longues. Ces dernières ont été réalisées comme décrit partie 6.3.1 et figure 111A. Pour la PCR 3, les clones 10.4 et 10.7 ont montré la bande à la taille de 8,8 kb et pour la PCR 4, seul le clone 10.7 a montré une bande à la taille de 17,3 kb (figure 111B). Les amplifications de grande taille notamment celle de 17,3



**Figure 109 : Analyse des transformants de la cinquième expérience d'échange de fragment via le système *cre/lox* chez *B. subtilis* lorsque la CRE est pré-exprimée par le *pJIM CRE* avec la PCR 1. [A] Schéma du résultat de l'échange de fragment. [B] En cas d'échange de fragment, la PCR 1 avec les amorces FE220 receveuse YxeQ F et FE220 receveuse YxeP R produit un amplicon de 1729 pb. « H2O » : contrôle négatif sans ADN. T-1 : témoin négatif plasmide pTwist TetR *lox71*. T-2: témoin négatif ADN de la cellule receveuse FE220 V3 + *pJIM CRE* cl 2.3.**

kb pouvant être techniquement difficile à réaliser, nous avons choisi de poursuivre l'analyse de ces 4 clones. Leurs génomes ont donc été extraits puis séquencés en long-read par la société prestataire PLASMIDSAURUS. Au vu des résultats de PCR, il était fort possible que ces génomes (excepté le cl 10.7) ne soient pas corrects mais récupérer leurs séquences pouvait aussi nous permettre de mieux comprendre notre méthode en prenant connaissance des différents types d'évènements qui pouvaient avoir lieu au sein de la cellule receveuse durant l'étape 4 du RAGE. L'analyse des séquences de ces 4 clones a révélé les faits suivants :

- En dehors de la zone cible les génomes paraissaient intacts, aucun grand remaniement n'ayant pu être détecté.
- Le site lox72 était présent dans les 4 génomes supportant l'idée que placer un marqueur de résistance à proximité du site de recombinaison est bénéfique.
- Le site loxm2/72, était présent dans les génomes des clones 10.2 et 10.7 mais absent des génomes des clones 7.3 et 10.4. Etrangement, pour ces deux derniers, il n'y avait aucune trace de site lox (pas même le loxm2/66 de la cellule receveuse) alors que la séquence s'est avérée correspondre à celle de la cellule receveuse FE220 V3. Les clones 7.3 et 10.4 étaient donc incorrects après séquençage. Il s'agissait vraisemblablement des mélanges de populations (et non de « clones » à proprement parler), expliquant le fait que ces transformants présentaient parfois les bandes attendues en PCR (PCR 1, PCR 2 et PCR 3 (pour le cl 10.4)).
- Les sites recombinés et non fonctionnels lox72 et loxm2/72 étaient bien présents dans les génomes des clones cl 10.2 et cl 10.7, indiquant que le système cre/lox avait été actif dans ces clones. Mais en regardant de plus près la séquence de la région échangée, nous nous sommes aperçus qu'elle avait bien la taille attendue de 44 kb dans le clone 10.7 (signifiant que ce clone présentait donc une séquence nucléotidique qui s'alignait parfaitement sur la séquence attendue), mais une taille de 129 kb dans le cl.10.2. Après une analyse plus approfondie, cette séquence correspondait à une séquence hybride entre la version du fragment 7.2 de la molécule donneuse de 44 kb et celle de la cellule receveuse de 155 kb. Notre interprétation est que ce génome est le résultat de divers évènements de recombinaison mettant en jeu des évènements de recombinaison homologue et site-spécifique. Cela est tout-à-fait possible quand on sait que tous les gènes présents dans le fragment de 44 kb de la molécule donneuse



**Figure 110 : Analyse des transformants de la cinquième expérience d'échange de fragment avec la PCR 2. [A] Schéma du résultat de l'échange de fragment. [B] En cas d'échange de fragment, la PCR 2 avec les amorces Verif receveuse F et Seq échange CmR R produit un amplicon de 1922 pb si un évènement de recombinaison homologue survient, cela produit un amplicon de 1375 pb. « H2O » : contrôle négatif sans ADN. T-1 : témoin négatif plasmide pTwist CmR loxm2/66 V3. T-2 : témoin négatif ADN de la cellule receveuse FE220 V3 + pJIM CRE cl 2.3.**

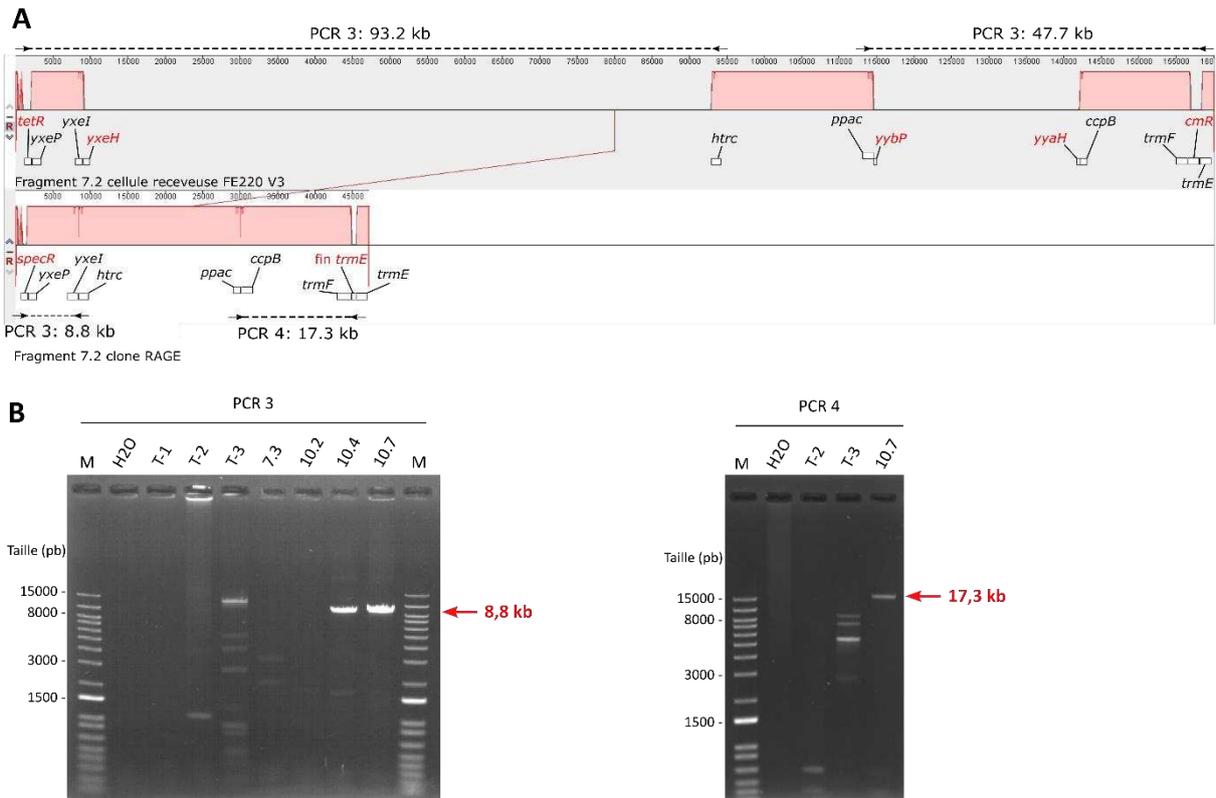
sont présents à l'identique dans le fragment de 155 kb et que *B. subtilis* possède une machinerie de recombinaison homologue très efficace.

Finalement, comme pensé après le criblage PCR, seul le clone 10.7 présentait la séquence attendue. Les deux sites lox recombinaison étaient bien présents de part et d'autre du fragment 7.2 correspondant à la version de 44 kb provenant de la molécule donneuse (figure 112). Grâce à cette expérience, nous avons réussi pour la première fois à obtenir un mutant de *B. subtilis* dans lequel un fragment génomique de 155 kb avait été remplacé par un fragment de 44 kb sans modifier le reste du génome. Cette expérience indique qu'il est possible de faire de l'échange de fragments médié par le système cre/lox chez *B. subtilis*. Dans notre exemple, le fragment de 44 kb est identique au fragment de 155 kb excepté qu'il est dépourvu des gènes considérés comme non-essentiels. Ceci peut ouvrir la voie à la construction de génomes modifiés à grande échelle ou à la construction de génomes réduits ou minimaux.

## 7. Conclusion

En 2020 une stratégie nommée RAGE, a été développée chez les mycoplasmes, dans le but de capturer dans la levure de grands fragments de génome bactérien (plusieurs dizaines de kb), puis d'utiliser cet hôte comme plateforme d'ingénierie génétique pour les modifier, et enfin de les transformer dans la bactérie à éditer (Garcia-Morales *et al.*, 2020). Après transformation, un échange d'ADN est provoqué par le système cre/lox, permettant l'insertion du fragment modifié à la place de la région chromosomique correspondante, générant ainsi des souches bactériennes mutantes. Elle comprend 4 étapes : (1) capture, entre deux sites lox incompatibles, d'un fragment de génome bactérien d'intérêt, (2) modification du fragment grâce aux outils génétiques disponibles chez la levure (optionnelle), (3) amplification du nombre de copies du fragment chez *E. coli* et (4) Echange de fragment dans la cellule receveuse via le système de recombinaison sites-spécifiques cre/lox (figure 78).

Au cours de ce chapitre, nous avons décrit l'adaptation de cette méthode à *B. subtilis*. Pour ce faire, nous avons utilisé comme matériel de départ, deux clones levures hébergeant le fragment 7.2 du génome de la souche MGP192 d'une taille de 44 kb capturé par TAR-cloning. A partir de ces deux clones, nous avons obtenu deux molécules donneuses de la taille de 58



**Figure 111 : Analyse des transformants de la cinquième expérience d'échange de fragment via le système cre/lox chez *B. subtilis*, avec les PCR 3 et 4. [A] Alignement entre le fragment 7.2 de *B. subtilis* FE220 (155 kb) et celui de *B. subtilis* MGP192 (44 kb). Les PCR 3 et 4 sont représentées. [B] Après une première analyse avec les PCR 1 et 2, les transformants sont analysés avec les PCR 3 et 4 dites longues, permettant une amplification qu'en cas d'échange de fragment. La PCR 3 avec les amorces FE220 receveuse YxeQ F et Verif\_htrc\_R produit un amplicon de 8,8 kb et la PCR 4 avec les amorces ppaC\_F et trmE\_frgt1\_bamHI\_R produit un amplicon de 17,3 kb pour la PCR 3. « M » : DNA Ladder 1 kb+ Invitrogen (100-12000 pb). « H2O » : contrôle négatif sans ADN. T-1 : témoin négatif ADNg de la souche *B. subtilis* FE220. T-2 : témoin négatif ADNg de la cellule receveuse V3 + pJIM CRE. T-3 : témoin négatif molécule donneuse issue du clone *E. coli* 4.8.3.**

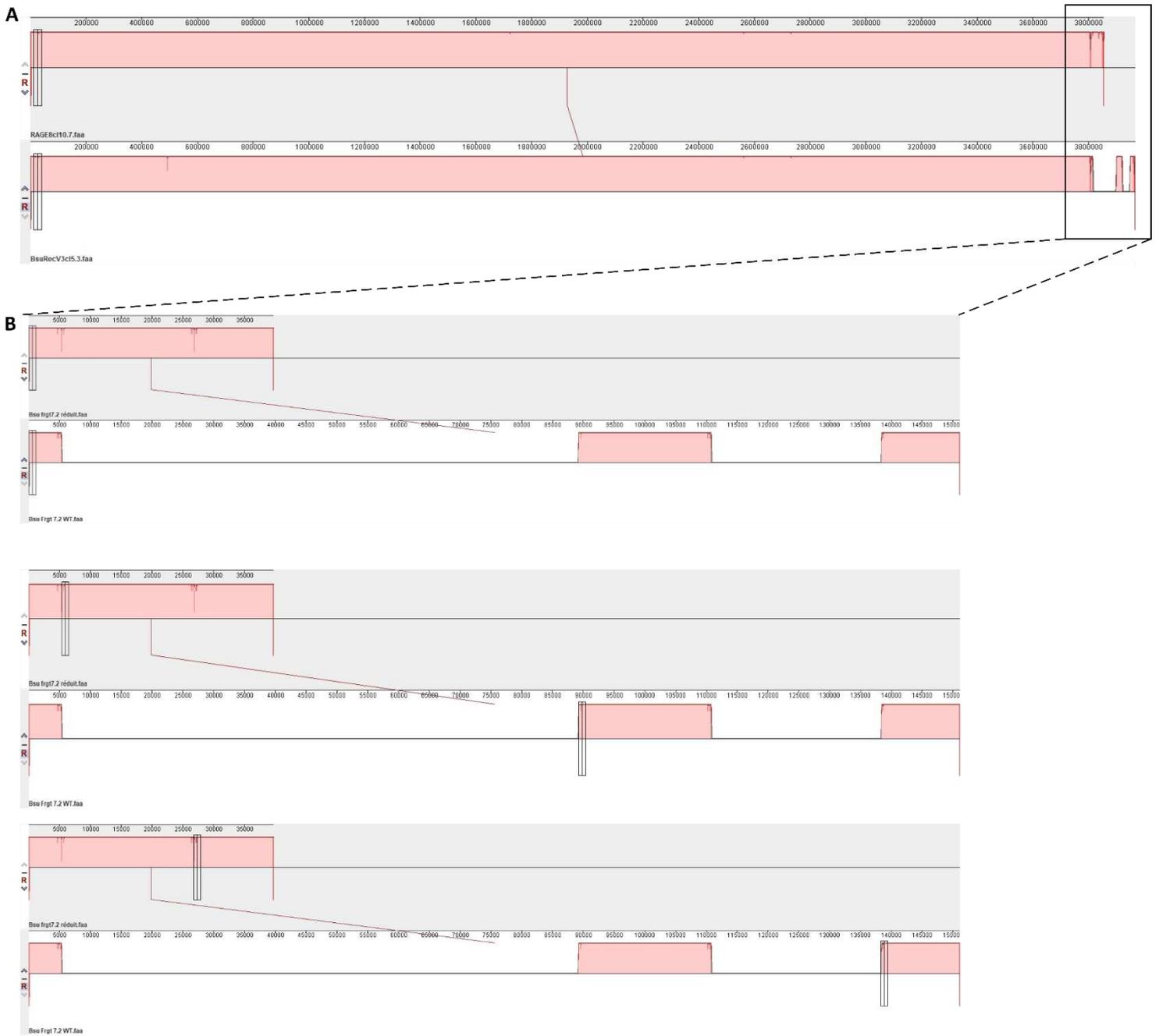
kb correspondant au fragment 7.2 agrémenté des éléments nécessaires à la méthode RAGE (étape 2 d'édition dans la levure). Les molécules modifiées ont ensuite été transformées dans *E. coli* Epi300 (étape 3) afin de produire du matériel de bonne qualité et en bonne quantité. A cette étape, nous nous sommes aperçus qu'une molécule d'une taille de 32 kb contaminait notre préparation dans les deux clones sélectionnés. Cette version était dépourvue de plusieurs gènes essentiels comparée à la version de 58 kb, nous en avons donc déduit que ce fragment ne pourrait pas générer de clones viables en cas d'échange de fragments chez *B. subtilis*. Nous l'avons validé de manière à pouvoir passer rapidement à l'étape 4 du RAGE. Mais nous étions tout-à-fait conscients que ce choix n'était pas optimal et que cela diminuerait forcément l'efficacité de la méthode. C'est pourquoi, nous prévoyons de tenter de la purifier et reproduire l'étape 4 avec une préparation contenant seulement la molécule à 58 kb.

Pour passer à l'étape 4, nous avons commencé par construire, à partir de la souche *B. subtilis* FE220, la cellule receveuse FE220 V3 contenant les sites loxm2/66 et lox71, de part et d'autre de la région chromosomique correspondant au fragment 7.2, d'une taille de 155 kb dans cette souche, puis nous avons réalisé plusieurs essais de transformation avec la molécule donneuse. Ils sont résumés dans le tableau 28.

Les étapes 1 à 3, étant des étapes bien maîtrisées par notre laboratoire, l'étape la plus critique, de notre point de vue, était l'étape 4. D'une part, nous n'étions pas certains que les protocoles de transformations standards de *B. subtilis* (notamment ceux basés sur la compétence naturelle) permettent l'entrée de grandes molécules. D'autre part, bien que l'efficacité du système cre/lox avait été démontrée chez *B. subtilis* pour de l'excision de marqueur de sélection (Cai and Chen, 2021; Koo *et al.*, 2017), son utilisation pour l'échange de grandes portions d'ADN de plusieurs dizaines de kb, n'avait, à notre connaissance, jamais été décrite.

Le premier essai réalisé a montré que, dans l'état, le système cre/lox n'était pas assez efficace pour provoquer l'échange désiré. A ce stade, la molécule donneuse semblait, cependant, bel et bien avoir été transférée dans la cellule receveuse via le protocole de transformation simple MC, basé sur la compétence naturelle, ce qui était déjà en soi un résultat.

Après cet échec, nous avons voulu confirmer que le système cre/lox était bien adéquat pour la mise en place de la méthode RAGE chez *B. subtilis*. Des expériences d'échange, non pas de grands fragments de génome, mais seulement de marqueurs de résistances ont donc



**Figure 112 : Alignement entre la séquence du clone 10.7 du dernier essai d'échange de fragment et la séquence de la cellule receveuse FE220 V3. [A] Alignement des séquences correspondantes aux génomes complets des deux souches. La zone de l'échange de fragment est encadrée. [B] Focus sur la zone de l'échange de fragment. Les 3 alignements sont identiques. Les boites noires représentent les régions identiques sur les deux séquences. Les zones en roses correspondent aux régions présentes sur les deux séquences. Les zones en blanc sur la séquence de la cellule receveuse correspondent aux régions absentes du clone 10.7. Elles représentent les régions éliminées après l'échange de fragment.**

été entreprises. Ces expériences nous ont permis non seulement de confirmer que le système cre/lox était apte à produire un échange d'ADN chez *B. subtilis*, mais aussi de définir un ensemble de paramètres permettant à ce système de fonctionner de manière optimale. Parmi les paramètres testés, une pré-expression de la protéine CRE dans la cellule receveuse par l'intermédiaire d'un plasmide multi-copie et la conduite des expériences de transformation à 30°C semblaient être les plus impactants.

Suite à ces résultats, de nouvelles tentatives d'échange de fragments de génome ont été tentées à 30°C, en utilisant la cellule receveuse FE220 V3 dans laquelle la protéine CRE était pré-exprimée. Nous avons essayé de pré-exprimer cette protéine de 3 manières différentes. La première, qui correspondait à l'essai de co-transformation d'un plasmide d'expression de la CRE en même temps que la molécule donneuse a été un échec. Cet essai, qui n'était pas, à proprement parler une expérience de « pré-expression » de la CRE, mais plutôt un essai de « sur-expression » de la CRE, a confirmé l'idée que les protéines devaient être déjà présentes « prêtes à intervenir » dans la cellule receveuse pour que le RAGE fonctionne. Ce point a été confirmé lors de la deuxième expérience, au cours de laquelle le gène *cre* était inséré dans le génome de la cellule receveuse FE220 V3. C'est en effet, à ce stade, que nous avons observé pour la première fois, l'activité du système cre/lox permettant la recombinaison site-spécifique au niveau d'un des côtés de la région à éditer. Néanmoins il n'était pas encore suffisant pour provoquer une recombinaison des deux côtés et donc le remplacement du fragment 7.2 de 155 kb par celui de 44 kb.

Nous avons poursuivi nos tentatives en pré-exprimant la CRE avec le plasmide réplcatif pDR244 Ery. Lors de cet essai, nous avons observé pour la première fois, l'échange de fragments désiré. En effet, lors de cette expérience, nous avons réussi à échanger, en une seule fois, une région chromosomique de 155 kb par une version réduite de la taille de 44 kb entraînant une réduction du génome de 111 kb. Cependant, après analyse approfondie des séquences, nous nous sommes rendus compte qu'un problème inattendu était survenu. En effet, une copie du plasmide pDR244 Ery permettant l'expression de la CRE semblait s'être insérée dans le génome bactérien, provoquant une délétion de plusieurs dizaines de kb. A cette étape, même si les résultats de l'expérience, nous confortaient sur notre capacité à réaliser des échanges de fragment dans *B. subtilis*, il nous était impossible de valider la méthode en l'état.

**Tableau 28 : Récapitulatif des essais d'échange de fragment via le système cre/lox chez *B. subtilis***

Essai	Conditions	CFU	Clones positifs / Clones analysés				
			PCR 1	PCR 2	Longues PCR	PFGE	Séquençage
1	classique	12	0/12	0/12	-	-	-
2	Co-transformation : pDR244 Ery + molécule donneuse	0	-	-	-	-	-
3	Pré expression de la CRE : gène inséré	40	1/40	0/40	-	-	-
4 + répétition	Pré expression de la CRE : pDR244 Ery	100	99/100	36/99	8/36	6/8	0/3
5	Pré expression de la CRE : pJIM CRE	60	33/60	4/33	1/4	-	1/4

Nous avons donc tenté de trouver un autre moyen de pré-exprimer la protéine CRE. Après quelques recherches, nous avons identifié le plasmide réplcatif pJIM. Après avoir vérifié que ce plasmide se comportait comme prédit et se maintenait sous forme libre au cours de la propagation des cellules, il a été utilisé pour la pré-expression de la CRE dans la cellule receveuse FE220 V3. Lors de ce nouvel essai, nous avons enfin obtenu un clone dans lequel le fragment 7.2 de 155 kb avait été remplacé par son homologue de 44 kb via le système cre/lox (cl 10.7) et le reste du génome était intègre. Cette expérience nous a permis de valider, pour la première fois, la dernière étape de la méthode RAGE chez *B. subtilis*. L'efficacité de l'échange était faible puisqu'un seul clone sur 60 analysés avait réalisé l'évènement voulu. Néanmoins l'évènement attendu était là.

L'ensemble de ces résultats ont démontré que la méthode RAGE pouvait être utilisée chez *B. subtilis*. La mise en place de cette stratégie a montré (i) qu'il était possible d'introduire une molécule d'ADN de grande taille (58 kb) chez *B. subtilis* par compétence naturelle (protocole MC), et (ii) que la pré-expression de la CRE était nécessaire pour que le système cre/lox provoque un échange de grands fragments de génome dans la cellule receveuse. Nous pensons que plusieurs paramètres peuvent encore être optimisés :

- Premièrement, tout au long de nos essais d'échanges de fragment, nous avons utilisé des molécules donneuses qui n'étaient pas pures, dû à la présence de la version à 32 kb. L'efficacité de l'échange pourrait sûrement être augmentée avec une molécule donneuse pure.
- Deuxièmement, le nombre de transformants ayant réalisé la recombinaison site-spécifique du côté où le gène *specR* est inséré à la place de *tetR* est bien plus important que ceux l'ayant effectué du côté où seulement le gène *cat* est éliminé. La présence d'un marqueur de sélection proche du site lox sur la molécule donneuse semble faciliter la sélection des clones dans lesquels la recombinaison site-spécifique a eu lieu. L'ajout d'un marqueur supplémentaire sur la molécule donneuse proche du loxm2/71 permettrait certainement d'augmenter l'efficacité d'échange de fragments.
- Troisièmement, pour prévenir les évènements de recombinaison homologue intempestifs, l'utilisation d'une cellule receveuse dépourvue de cette capacité pourrait être une bonne alternative. Il est cependant important de prendre en compte que la recombinaison homologue est un mécanisme lié à la compétence naturelle, son inactivation pourrait donc entraîner un dysfonctionnement de la compétence



naturelle. Afin de pallier ce problème, plusieurs options sont envisageables : (i) tenter d'inactiver des gènes impliqués dans la recombinaison homologue mais pas dans la compétence naturelle, (ii) tenter de sur-exprimer des facteurs modulant ou inhibant la recombinaison homologue sans affecter la capacité de transformation, comme par exemple la protéine RecU qui module l'activité de RecA (Maier, 2020; Serrano *et al.*, 2018), ou (iii) développer un nouveau protocole de transformation artificielle (Chang and Cohen, 1979), non basée sur la compétence naturelle si cela est possible.

- Enfin, le design du pansement (cf partie 2.1 du chapitre 3) pourrait être encore affiné de manière à incorporer (après échange de fragments dans le génome de *B. subtilis*) des répétitions directes à chaque borne. Cet ajout pourrait en effet donner la possibilité d'éliminer les sites lox et revenir à une séquence sauvage ou encore recycler les sites lox pour recommencer le cycle au locus suivant par exemple.

Dans le cadre du projet Bacillus 2.0, cette méthode peut être proposée comme une vraie alternative à la TG, en offrant notamment, la possibilité de produire des souches de *B. subtilis* mutantes à grande échelle. Pour déterminer les limites de la stratégie, il pourrait être intéressant de réaliser les expériences suivantes :

- (i) Tenter de réduire encore davantage le fragment 7.2 de 44 kb cloné dans la levure : par exemple l'essentialité du locus situé entre les gènes *yycD* et *cotF* (8,5 kb) n'a pas encore été déterminée.
- (ii) Répéter des expériences de RAGE avec des fragments plus grands que le fragment 7.2 afin de déterminer la taille maximale avec laquelle il est possible de travailler.
- (iii) Travailler avec un fragment d'ADN synthétique

Il est aussi important de noter que ce travail fait office de preuve de concept. Pour démontrer la faisabilité de la méthode RAGE, nous avons voulu ré-utiliser un fragment de génome capturé par TAR-cloning, cela a amené un certain nombre de contraintes, en particulier dans le design de la cellule receveuse. Si une telle méthode devait être appliquée à la construction itérative du génome semi-synthétique de *B. subtilis* les fragments de génome seraient capturés dans la levure par la méthode de CReasPy-Fusion qui permet de choisir minutieusement les bornes du fragment à échanger.



Les travaux réalisés au cours de ce chapitre doivent encore être perfectionnés (reproduction, optimisation, analyses expertes de toutes les séquences obtenues...) mais feront vraisemblablement l'objet d'une deuxième publication, soit seuls, soit en lien avec les travaux de TAR-cloning réalisés pendant la thèse du Dr. Guesdon.



## Discussion et conclusion

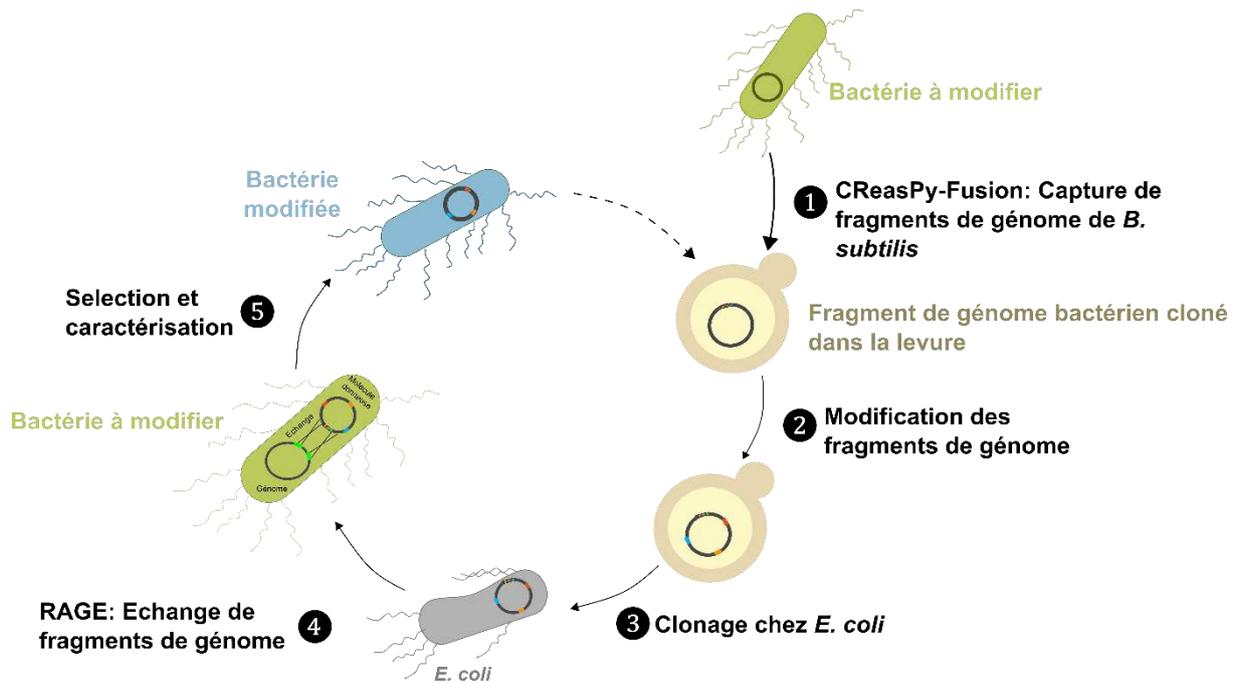


## Discussion et conclusion

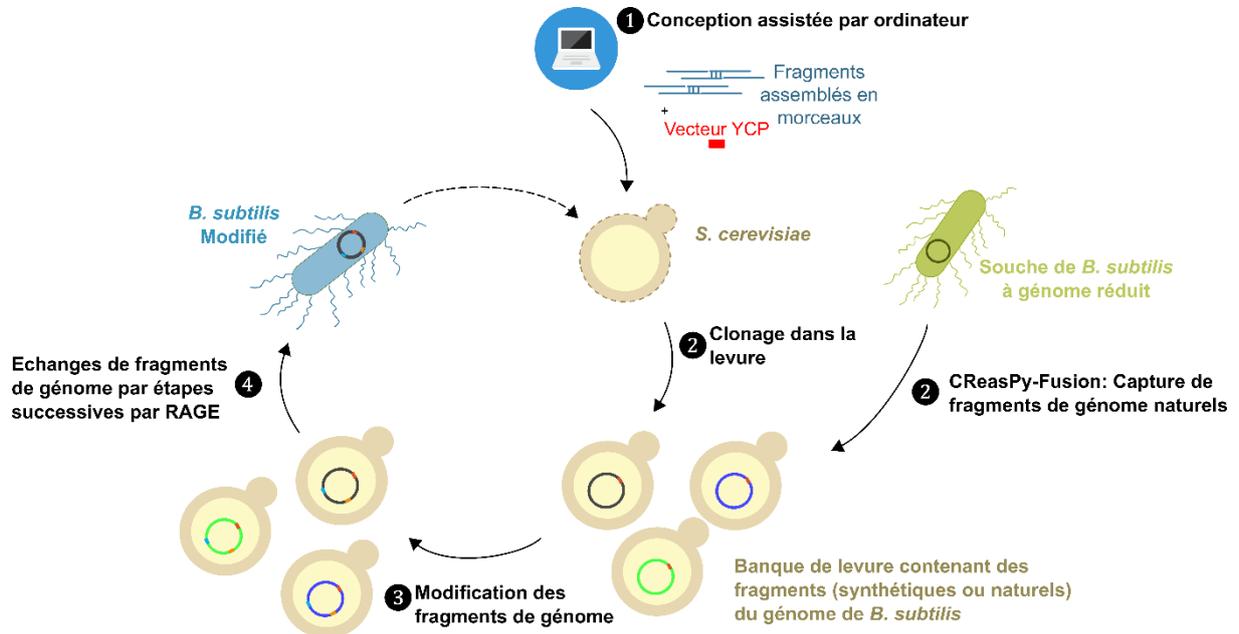
Le projet Bacillus 2.0 avait pour objectif de transposer à *B. subtilis* trois technologies de pointe du domaine de la BS dans le but de munir cet organisme d'un ensemble de méthodes efficaces pour la construction rapide de châssis microbiens utiles à l'industrie. Ces méthodes incluaient (i) la synthèse et l'assemblage de génomes entiers dans la levure, (ii) l'ingénierie de ces génomes grâce au répertoire d'outils génétiques disponibles chez la levure, et enfin (iii) leur transplantation dans une cellule receveuse adéquate de façon à générer les souches mutantes *ad hoc* (figure 1 de l'avant-propos).

Seulement, des difficultés ont été rencontrées dès les premières étapes du projet puisque le consortium n'a pas réussi à valider le premier prérequis au projet qui était de cloner et stabiliser le génome natif d'une souche de *B. subtilis* dans la levure. Face à cette situation, les objectifs de départ ont été revus et le projet a été ré-orienté de façon à répondre aux questions qui prédominaient les discussions du consortium: Est-ce toujours judicieux de vouloir utiliser la levure *S. cerevisiae* comme plateforme d'ingénierie pour générer des mutants de *B. subtilis*? Est-ce que cette stratégie, qui a été opérationnelle chez les mycoplasmes, est adaptée à des bactéries plus conventionnelles comme *B. subtilis*? Faut-il poursuivre les efforts ou bien changer de plan d'action?

Des données obtenues au cours de la thèse du Dr. Guesdon (TAR-cloning, expériences de fusion...) puis d'autres données issues de la littérature sur le clonage des génomes d'*E. coli* (Zhou *et al.*, 2016) et de *C. crescentus* (Venetz *et al.*, 2019) dans la levure ont finalement encouragé le consortium à ne pas totalement abandonner la stratégie initiale et à adopter une démarche plus méthodique visant (i) à cloner dans la levure un premier fragment de génome de *B. subtilis* de petite taille, puis, si les résultats étaient satisfaisants, à en cloner un second un peu plus grand (incluant le premier), puis un troisième encore plus grand (incluant le second), et ainsi de suite... et (ii) si besoin à modifier chacun de ces fragments, en apportant les modifications nécessaires (ajout d'ARS, délétions...), jusqu'à ce que le fragment d'intérêt se maintienne correctement dans la levure. Ce travail étant à la fois long (travail toujours en cours) et incertain (personne ne peut en effet prédire si le génome entier sera un jour cloné dans la levure), la démarche méthodique adoptée intégrait un deuxième aspect : tenter d'éditer *B. subtilis*, non pas en remplaçant son génome complet par un autre génome par TG, mais en modifiant son génome morceaux par morceaux.



**Figure 113 : Combinaison des méthodes de CRasPy-Fusion et RAGE dans un cycle d'ingénierie de génome bactérien à grande échelle. 1)** Clonage de grands fragments de génome dans la levure par CRasPy-Fusion. **2)** Modification des fragments de génome dans la levure (étape optionnelle). **3)** Clonage chez *E. coli* pour amplifier le nombre de copies. **4)** Echange de fragments de génome dans une cellule bactérienne receveuse via le système cre/lox (RAGE) **5)** Sélection et caractérisation des bactéries modifiées.



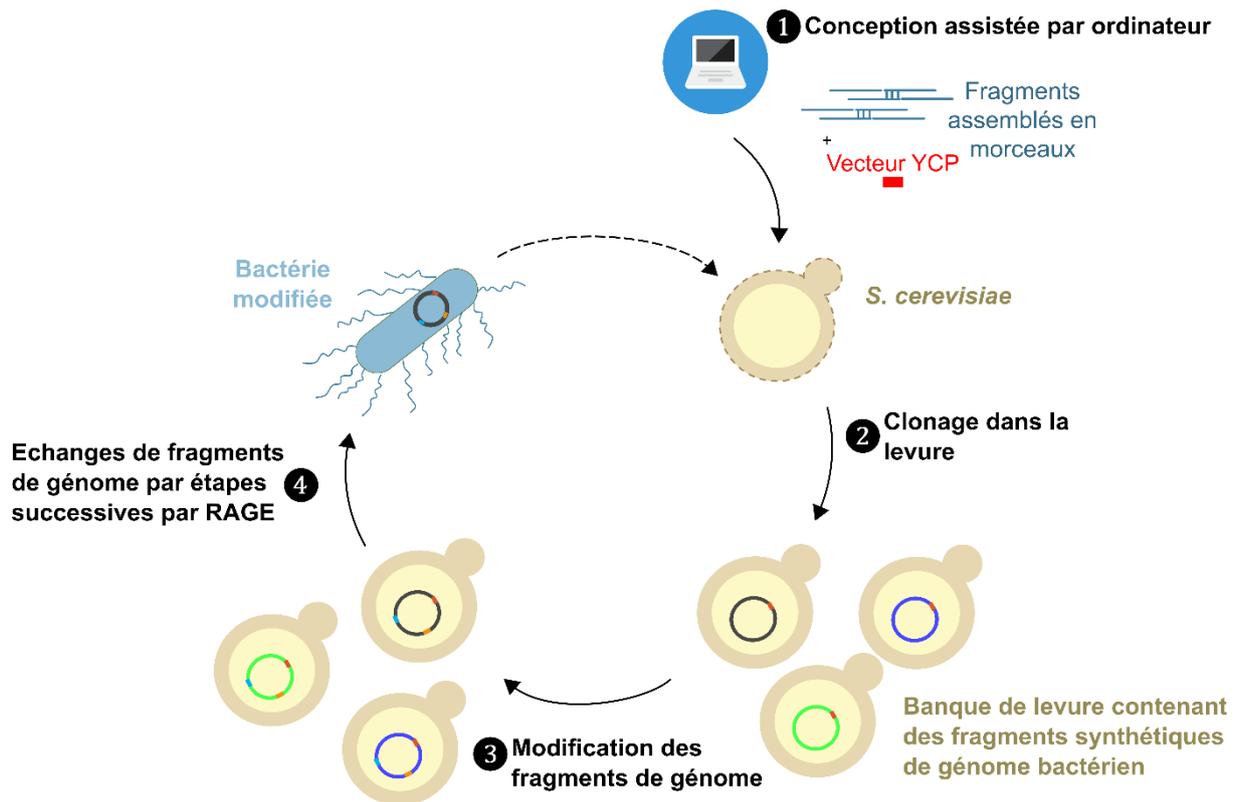
**Figure 114 : Cycle d'ingénierie pour la construction de cellules contrôlées par un génome semi-synthétique appliqué à *B. subtilis*. 1)** Conception assistée par ordinateur d'un génome semi-synthétique de *B. subtilis*. **2)** Assemblage de fragments de génome synthétiques dans la levure, ou clonage de fragments de génome natifs issus de souches réduites de *B. subtilis*. **3)** Modification des fragments de génome dans la levure (étape optionnelle). **4)** Echanges de fragments successifs dans une cellule bactérienne receveuse grâce à la méthode RAGE pour modifier l'ensemble du génome et ainsi obtenir une cellule contrôlée par le génome semi-synthétique conçu à l'étape 1.

C'est dans ce contexte qu'a démarré mon projet de thèse. Il s'est finalement articulé autour du développement de deux méthodes distinctes pour la manipulation et l'ingénierie de grands fragments de génome de *B. subtilis*. Ces méthodes qui font toutes deux appel, à la levure comme hôte intermédiaire de clonage sont les suivantes :

(i) la méthode de CREasPy-Fusion qui permet, par sa flexibilité, de cloner des fragments de génome ou des génomes entiers dans la levure (Guesdon *et al.*, 2023).

(ii) la méthode RAGE (Garcia-Morales *et al.*, 2020; Wang *et al.*, 2019; Santos and Yoshikuni, 2014; Santos *et al.*, 2013) qui permet le transfert de fragments de génome bactérien clonés dans la levure dans une cellule bactérienne receveuse, dans le but de provoquer un échange de fragments d'ADN et générer des mutants. Cette technique s'inspire de la technique RMCE (Roebroek *et al.*, 2011; Schetelig *et al.*, 2019), une procédure de génétique inverse permettant la modification systématique et répétée des génomes eucaryotes supérieurs par intégration ciblée, sous l'action de recombinases site-spécifiques (Turan *et al.*, 2011; Turan *et al.*, 2013).

Les progrès réalisés au cours de la thèse sur ces deux méthodes sont tels que des perspectives d'ingénierie du génome de *B. subtilis* à grande échelle, incluant la construction de génomes semi-artificiels ou artificiels (objectif du projet Bacillus 2.0), commencent à se dessiner. En effet, la combinaison des méthodes de CREasPy-Fusion (pour la capture de morceaux de génome natifs dans la levure) et de RAGE (pour la modification des fragments natifs clonés dans la levure, leur extraction et leur échange dans la cellule receveuse) permet désormais d'envisager des expériences de construction de mutants d'intérêt pour le projet et même au-delà (figure 113). Mais il est possible de voir plus loin puisque l'association de ces méthodes à des expériences d'assemblage de fragments de génome synthétiques permet également d'envisager (i) la production de cellules contrôlées par des génomes semi-artificiels (qui seraient composés de fragments natifs modifiés et échangés par CREasPy-Fusion/RAGE et de fragments synthétiques assemblés dans la levure puis échangés par RAGE), (figure 114) ou (ii) de cellules contrôlées par des génomes entièrement artificiels (figure 115). Nous convenons que ces approches seraient plus longues que l'approche envisagée figure 1 mais si le génome de *B. subtilis* ne peut pas être cloné entièrement dans la levure, alors la combinaison de ces méthodes pourrait constituer une véritable alternative à la stratégie proposée initialement dans le projet Bacillus 2.0.



**Figure 115 : Cycle d'ingénierie de génome pour la production de cellules contrôlées par un génome entièrement synthétique. 1) Conception assistée par ordinateur du génome synthétique. 2) Assemblage de fragments synthétiques en grands fragments dans la levure. 3) Modification des fragments de génome dans la levure (étape optionnelle). 4) Echanges de fragments successifs dans une cellule bactérienne receveuse grâce à la méthode RAGE pour modifier l'ensemble du génome et ainsi obtenir une cellule contrôlée par le génome synthétique conçu à l'étape 1.**

Dans la première partie de cette discussion, les méthodes de CReasPy-Fusion et de RAGE adaptées à *B. subtilis* sont ré-abordées séparément afin de mettre en évidence l'apport de chacune d'elles ainsi que leurs avantages et limites. Dans la seconde partie de cette discussion est abordée la combinaison des deux méthodes (et de la méthode de synthèse d'ADN) dans une stratégie de modification du génome de *B. subtilis* à grande échelle dans le cadre du projet Bacillus 2.0. Enfin, ces méthodes n'étant pas limitées à *B. subtilis*, des applications annexes seront aussi discutées.

## 1. La méthode de CReasPy-Fusion

La première partie de ce travail de thèse a consisté à adapter la méthode de CReasPy-Fusion à *B. subtilis*. Cette méthode avait été imaginée dans le cadre du projet Bacillus 2.0 pour répondre aux contraintes imposées par la méthode du TAR-cloning (cf partie avant-propos) mais au regard des difficultés rencontrées avec le génome de *B. subtilis*, l'équipe a préféré valider la faisabilité de cette approche avec des génomes bactériens qui avait déjà été clonés dans la levure par d'autres approches et de ne l'appliquer à *B. subtilis* qu'une fois son efficacité démontrée. La méthode de CReasPy-Fusion a donc été développée dans un premier temps chez les mycoplasmes, des petites bactéries sans paroi qui avaient également été utilisées pour mettre au point toutes les autres méthodes de clonage de génomes bactériens dans la levure décrites jusqu'à présent dans la littérature (Benders *et al.*, 2010; Natalay Kouprina and Vladimir Larionov, 2016; Ruiz *et al.*, 2019; Karas *et al.*, 2013). Au cours de ce travail, des grands fragments de génome sélectionnés par l'expérimentateur (et non par la présence/absence de sites de restriction) ont été capturés avec une grande efficacité dans la levure. Des génomes entiers ont également été clonés/édités avec succès dans la levure. Un seul génome sur les sept testés a été réfractaire au clonage. Après analyses, il s'agissait du génome qui possédait le plus grand nombre de systèmes de restriction-modification (Guesdon *et al.*, 2023), suggérant que les gènes codant des nucléases bactériennes pouvaient être un obstacle au clonage par CReasPy-Fusion.

Ceci peut se comprendre car, pour rappel, la méthode de CReasPy-Fusion est basée (i) sur la méthode dite de « Fusion » mise au point en 2013 par l'équipe de B. Karas (Karas *et al.*, 2013) qui permet de transférer des molécules d'ADN directement par contact cellules/cellules



(entre bactérie et levure) et (ii) le système CRISPR-Cas9 pour l'isolement spécifique de la molécule d'intérêt ou d'une partie de cette molécule (figure 20). Après contact protoplastes/sphéroplastes, il est fort probable que l'ADN ainsi que l'ensemble du contenu cellulaire de la bactérie soient intégralement libérés dans la levure au cours d'un mécanisme qui n'est pas encore décrit. Ainsi il est vraisemblable que la présence de certaines protéines cytoplasmiques comme des nucléases (endo- ou exo-nucléases) ou d'autres facteurs soit une entrave au clonage. Chez *M. gallisepticum* (*Mgal*) par exemple, il a été montré que l'inactivation des nucléases MnuA et Cas9 augmentait le nombre de transformants levures propageant le génome de *Mgal* (Guesdon *et al.*, 2023). Chez *M. capricolum* subsp. *capricolum*, c'est l'élimination du gène codant une enzyme de restriction (un autre type de nucléases) qui a eu un impact sur l'efficacité de transfert. Chez *M. mycoides* subsp. *capri* (*Mmc*), en plus des gènes codant des enzymes de restriction, l'absence du gène codant la perméase GlpF (protéine facilitant l'absorption du glycérol), a augmenté de manière significative (jusqu'à 21 fois) le transfert du génome de *Mmc* dans la levure (Karas *et al.*, 2019).

Forts de ces connaissances, nous avons procédé par étapes pour adapter la méthode de CReasPy-Fusion à *B. subtilis*, en commençant par des expériences simples et en augmentant progressivement la complexité de la méthode. Ainsi, au cours de la première étape, nous avons simplement cherché à démontrer qu'il était possible de transférer une molécule plasmidique par contact cellules/cellules entre une bactérie à paroi Gram+ et la levure ; un phénomène, qui à notre connaissance, n'avait jamais été décrit auparavant. En effet, Karas *et al.*, avaient bien démontré, au cours d'expériences de Fusion, la possibilité de transférer de l'ADN entre mycoplasmes et levures et entre bactérie Gram- (*H. influenzae*) et levure (Karas *et al.*, 2013), mais n'avaient jamais travaillé avec des bactéries Gram+.

Ensuite, au cours d'une deuxième étape, nous avons introduit le système CRISPR-Cas9 dans la levure, puis nous avons réitéré l'expérience dans le but, cette fois, de cloner et éditer simultanément la molécule plasmidique dans la levure.

Les résultats étant très satisfaisants, nous avons tenté, au cours de la troisième étape, de cloner/éditer non plus une molécule plasmidique, mais des fragments de génome de *B. subtilis* d'une centaine de kilobases environ et cela à partir de trois souches différentes : *B. subtilis* MS (3.98 Mb), *B. subtilis* FE220 (3.96 Mb) et *B. subtilis* MGP192 (2.86 Mb). Ces expériences ont montré d'une part que, contrairement à ce qui a été observé chez *H. influenzae* et les



mycoplasmes (Karas *et al.*, 2013), la présence (souche *B. subtilis* MS) ou l'absence (*B. subtilis* FE220) du système de restriction-modification BsuM ne semblait pas influencer sur l'efficacité de la méthode, puisqu'un fragment d'environ 145 kb a pu être capturé dans les deux cas avec une très faible efficacité. D'autre part, elles ont révélé qu'alors qu'un fragment de 125 kb pour la souche à génome réduit MGP192 et de 145 kb pour les souches à génomes non réduits MS et FE220 pouvaient être capturés dans la levure, un fragment environ deux fois plus grand (268 kb pour MGP192 et 350 kb pour MS et FE220) ne pouvait être capturé qu'à partir de la souche MGP192. L'ensemble de ces résultats ont mis en lumière deux faits notables :

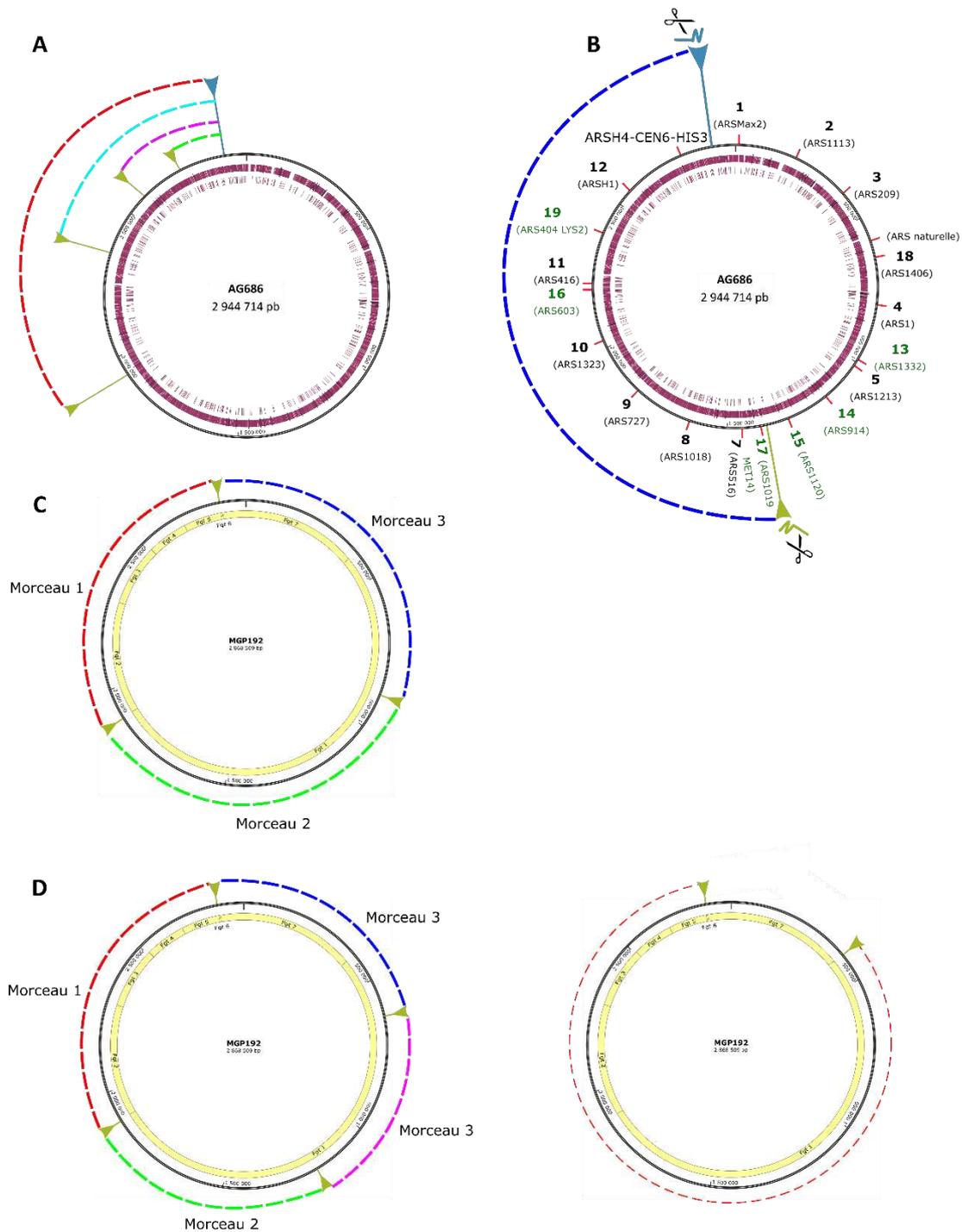
- (1) L'obtention de transformants levures hébergeant un fragment de 145 kb provenant des souches MS (3.98 Mb) et FE220 (3.96 Mb) suggère que des génomes de taille égale ou supérieure à 3.98 Mb peuvent bel et bien entrer et être pris en charge par la levure. Ce résultat, s'il devait se confirmer, serait spectaculaire puisqu'à ce jour, le plus grand génome cloné et maintenu entier dans la levure en CReasPy-Fusion est celui de *M. mycoides* subsp. *mycoides* qui a une taille de 1.18 Mb seulement (Guesdon *et al.*, 2023). Il est fort possible que des génomes plus grands puissent être capturés dans la levure en utilisant cette méthode mais aucune équipe n'a encore tenté de le faire. Au regard de cette situation et des objectifs du projet (cloner une version réduite et modifiée du génome de *B. subtilis* d'environ 2.86 Mb), il pourrait être intéressant de démontrer que des génomes de taille supérieure à 1.18 Mb peuvent véritablement être clonés entiers dans la levure par CReasPy-Fusion. Des expériences mettant en jeu d'autres espèces avec des génomes aux caractéristiques différentes pourraient être réalisées dans le but de démontrer que la taille initiale du génome ne constitue pas à elle seule un paramètre bloquant le clonage d'un génome dans la levure. On peut ainsi proposer des espèces comme : *A. laidlawii* (1,5 Mb, code génétique universel, déjà cloné dans la levure (Karas *et al.*, 2012)), *S. citri* (1,8 Mb, code génétique non standard, déjà cloné dans la levure (Labroussaa *et al.*, 2016)), *H. influenzae* (1,8 Mb, code génétique universel, déjà cloné dans la levure (Karas *et al.*, 2013)), *Spiroplasma ixodetis* (2.2 Mb, code génétique non standard, jamais cloné dans la levure (Tully *et al.*, 1995)), ou encore des souches de *Lactobacillus* (1.8 à 3.3 Mb, 33 à 51% GC, jamais cloné dans la levure (Kant *et al.*, 2011)).
- (2) Sachant que des fragments de taille supérieure à 350 kb peuvent être clonés dans la levure par CReasPy-Fusion, l'absence de transformants levures hébergeant le



fragment de 350 kb à partir des souches MS (3.98 Mb) et FE220 (3.96 Mb) peut suggérer que le contenu en gènes (et l'expression de certains gènes compris dans ces fragments) est délétère pour la levure et constitue une limite au clonage ; un point déjà mis en évidence dans les expériences de TAR-cloning menées par le Dr. Guesdon. Au vu des efficacités de capture (très faibles) obtenues avec le fragment de 145 kb pour ces deux souches, il est également possible que l'efficacité de capture du fragment de 350 kb soit si faible qu'elle ne nous permette pas de détecter de clones positifs.

Au cours de la quatrième étape, nous avons tenté d'optimiser la méthode, les résultats sont prometteurs mais nécessitent encore du travail.

Enfin, la cinquième et dernière étape de ce travail a consisté à comprendre et documenter le phénomène de transfert d'ADN lors du contact cellules/cellules entre la bactérie et la levure, grâce à des observations en microscopie notamment à fluorescence (épifluorescence et confocale). Lors de ces expériences, nous avons pu mettre en évidence deux phénomènes que nous avons appelés « déversement » et « endocytose », et qui corroborent des observations déjà faites par le groupe de B. Karas (microscopie électronique) lors du contact mycoplasmes/levures (Karas *et al.*, 2019). Malgré ces efforts, ce travail reste pour le moment trop exploratoire et notre compréhension des mécanismes de transfert de l'ADN depuis la bactérie vers la levure ainsi que son devenir (entrée dans le noyau) est encore, à ce jour, très limitée. Comme mentionné dans la discussion du chapitre 1, des observations en microscopie corrélative (CLEM), qui allie microscopie à fluorescence et microscopie électronique, pourraient apporter une meilleure vision (d'un point de vue structural) de l'interaction entre la bactérie et la levure lors du contact. Cette technique pourrait également permettre de déceler s'il y a réellement fusion membranaire lors des événements dits de « déversement » et « d'endocytose », et de déterminer si ces deux phénomènes sont deux phénomènes liés ou indépendants. Enfin, en marquant le génome bactérien, il pourrait être possible de le suivre au cours de son transfert, puis d'observer son devenir dans la levure, ainsi que sa localisation finale (voire ses interactions avec les chromosomes levures si ces derniers sont également marqués).



**Figure 116 : Capture de fragments de génome de *B. subtilis* par CREasPy-Fusion. [A]** Récapitulatif des fragments de génome capturés par CREasPy-Fusion à l'heure actuelle. Le plus grand fragment de 925 kb (en rouge) inclut le premier fragment de 125 kb (en vert), le deuxième de 268 kb (en rose) et le troisième de 536 kb (en cyan). **[B]** Demi-génome à tenter de capturer avec 10 ARS additionnelles selon le deuxième design. **[C]** Clonage du génome de *B. subtilis* en 3 morceaux d'environ 1 Mb. Le morceau 1 (en rouge) est déjà cloné dans la levure. Le morceau 3 (en bleu) risque d'être difficile à cloner car il contient le fragment 7 instable dans la levure. Certains gènes ou régions de ce fragment 7 devront être supprimés. **[D]** Autres configurations de capture du génome de *B. subtilis* en différents morceaux. Le morceau 1 (en rouge) étant déjà cloné dans la levure, les efforts peuvent être concentrés sur le reste du génome qui peut être cloné en plusieurs morceaux. Le fragment 7 étant instable, on peut aussi imaginer cloner l'ensemble du génome excepté ce fragment.

Après avoir démontré que la méthode de CReasPy-Fusion pouvait être adaptée à *B. subtilis* pour capturer des fragments de génome d'une centaine de kilobases dans la levure (chapitre 1), nous l'avons utilisée au cours du chapitre 2, pour tenter de cloner chez cet hôte, des fragments du génome de *B. subtilis* de plus en plus grands, qui incluaient à chaque fois le fragment précédemment cloné (figure 116A). Les travaux du Dr. Guesdon avaient montré qu'il serait nécessaire d'ajouter des ARS dans certaines parties du génome de *B. subtilis* et d'en éliminer certains *loci* si nous voulions parvenir à le cloner entier dans la levure. A partir de ces informations, un premier génome modifié a été conçu par ordinateur, et la construction de la souche correspondante a été entamée. A partir de cette souche modifiée, des expériences de CReasPy-Fusion ont été menées. Un fragment de 536 kb a facilement été capturé dans la levure, qu'il contienne 1 ou 2 ARS. En revanche, la capture du fragment suivant de 925 kb contenant 4 ARS a échoué. Ce n'est seulement lorsqu'une cinquième ARS a été ajoutée (soit en moyenne une ARS tous les 185 kb environ) que le fragment a pu être capturé dans la levure, signifiant que notre 1<sup>er</sup> design n'était pas optimal et qu'il devait être modifié urgemment pour poursuivre le projet. Un second génome modifié a donc été conçu par ordinateur (le design est décrit en fin de chapitre 2) ; les souches correspondantes sont en cours de construction et des essais de CReasPy-Fusion seront bientôt réalisés avec un nouvel essai de capture d'un demi-génome contenant 10 ARS puis des trois quarts du génome (figure 116B).

Parallèlement à ces expériences, d'autres essais de capture de fragments d'environ 1 Mb seront tentées sur d'autres régions du génome afin de savoir si le génome de *B. subtilis* réduit peut être cloné en 3 segments d'environ 1 Mb dans la levure comme illustré (figure 116C). Pour le segment de 1 Mb contenant le fragment 7, plusieurs gènes seront supprimés avant l'expérience. Sachant qu'au début de ma thèse le génome de *B. subtilis* MGP192 était cloné dans la levure en 21 morceaux de 6 kb à 515 kb, l'obtenir en 3 fragments d'environ 1 Mb serait une grande avancée pour le projet. Cependant, ce ne serait pas une fin en soi, puisque l'objectif est de cloner ce génome en un morceau dans la levure. De par la flexibilité de la méthode, de nombreuses autres configurations peuvent être imaginées et de nombreuses captures peuvent être réalisées (figure 116D, schéma non exhaustifs). Pour chacune de ces expériences, des mesures correctives pourront être amenées et ceci jusqu'à capturer une version du génome qui puisse être clonée entière dans la levure.

Un tel accomplissement signifierait que le génome semi-synthétique (appelé *SynBsu*) de taille plus petite (~1.89 Mb) que le génome natif/réduit sur lequel nous travaillons



actuellement (2.86 Mb) devrait se maintenir sans trop de difficulté dans la levure. Ce génome semi-synthétique, conçu à partir (i) de données de la littérature, (ii) de données propres à l'équipe SyBER (Institut MICALIS, UMR 1319 INRAE / AgroParisTech, Jouy-en-Josas) et (iii) d'un ensemble de règles définies par le consortium pour convenir des gènes ou *loci* à conserver ou à exclure, est en constante « réactualisation » puisque sa séquence est fonction des résultats des expériences de CREasPy-Fusion (ajout de nouvelles ARS, nouveaux marqueurs, délétion de gènes ou *loci*...). L'alignement actuel des séquences *SynBsu* et du génome natif/réduit montre que de nombreuses zones se recoupent. Cependant, il existe des *loci* présents dans le génome natif/réduit qui sont absents de *SynBsu*, et à l'inverse (et plus problématique) des *loci* présents dans *SynBsu* et absents du génome natif/réduit. Lors de l'assemblage du génome *SynBsu* dans la levure, une attention particulière sera donc portée à ces régions qui n'auront jamais été clonées dans la levure auparavant.

Pour conclure, avec ce travail nous fournissons une nouvelle méthode de clonage de génomes dans la levure qui offre certains avantages par rapport aux méthodes publiées précédemment. En effet, contrairement à la méthode « classique » ou la méthode par « Fusion », la CREasPy-Fusion a l'avantage de pouvoir être utilisée chez des bactéries qui ne se prêtent pas à la transformation ou à la modification génétique. Elle permet aussi de s'affranchir des étapes fastidieuses de purification des chromosomes dans des blocs d'agarose, ce qui réduit à la fois le temps de préparation et les coûts. Malgré ces avantages non négligeables, la CREasPy-Fusion a aussi quelques limites puisqu'il a été montré que des facteurs biologiques (présence de nucléases...) peuvent influencer son efficacité (Guesdon *et al.*, 2023).

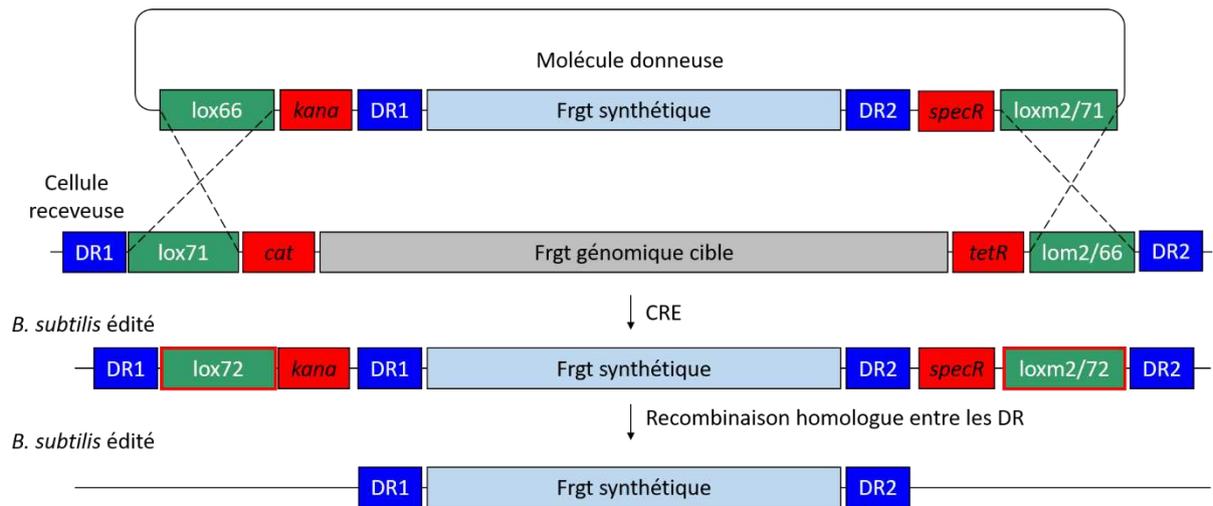
Pour l'heure, la CREasPy-Fusion peut être utilisée chez les mycoplasmes (bactéries sans paroi) ainsi que chez *B. subtilis* (bactérie Gram+). Il est possible qu'elle puisse être adaptée à un grand nombre de bactéries. Les travaux menés pour l'adaptation et l'utilisation de cette méthode chez *B. subtilis* démontrent en effet sa versatilité, tant d'un point de vue du modèle bactérien pour lequel elle peut être utilisée que du point de vue de son application grâce à la possibilité de capturer de façon spécifique n'importe quelle région d'un génome bactérien.



## 2. La méthode RAGE

La seconde partie de ce travail de thèse (chapitre 3) a consisté à adapter la méthode RAGE à *B. subtilis*. Cette méthode d'ingénierie et d'échange de fragments de génome a été développée en 2020 pour une bactérie réputée difficile à modifier sur le plan génétique, *M. pneumoniae* (Garcia-Morales *et al.*, 2020). Elle est basée sur l'utilisation de la levure *S. cerevisiae* comme plateforme d'ingénierie de fragments de génome puis sur un échange de cassettes d'ADN médié par le système cre/lox dans la cellule bactérienne à éditer (appelée aussi la cellule receveuse). Plus précisément, elle se décline en 4 étapes : 1) capture dans la levure d'un fragment de génome bactérien d'intérêt entre deux sites lox incompatibles, 2) modification du fragment en utilisant les outils génétiques disponibles chez cet hôte, 3) transformation du fragment chez *E. coli*, pour permettre l'amplification de son nombre de copies, et 4) échange de fragments d'ADN, médié par le système de recombinaison site-spécifique cre/lox, dans une cellule receveuse (figure 23).

Pour démontrer que cette méthode pouvait être appliquée à *B. subtilis*, nous avons choisi de remplacer un fragment de génome d'une taille de 155 kb par sa version réduite d'une taille de 44 kb, entraînant une réduction de 111 kb en une seule expérience ; un choix fort audacieux puisque nous aurions pu choisir de réaliser cette preuve de concept sur des fragments de petite taille (une dizaine de kb par exemple), ce qui aurait certainement facilité l'étape d'échange de fragments dans la cellule receveuse mais également l'étape d'entrée de la molécule donneuse dans la cellule receveuse. Deux raisons nous ont amenés à faire ce choix. Tout d'abord, nous avons à disposition au laboratoire une collection de clones levures hébergeant déjà des morceaux de génome de *B. subtilis* que nous pouvions mettre à profit pour l'expérience. Nous avons donc sélectionné, dans cette collection, le clone levure hébergeant l'un des plus petits morceaux de génome de *B. subtilis* et avons veillé à ce que le fragment choisi ne contienne pas de gènes ou régions repérés comme problématiques au cours des expériences de TAR-cloning (opéron ribosomique, surfactine...). Ensuite, nous savions qu'accomplir une expérience avec des fragments aussi grands aurait un impact direct sur le projet puisque cela indiquerait qu'à défaut de pouvoir amener le génome semi-synthétique dans la cellule receveuse en une fois (figure 1), nous pourrions potentiellement

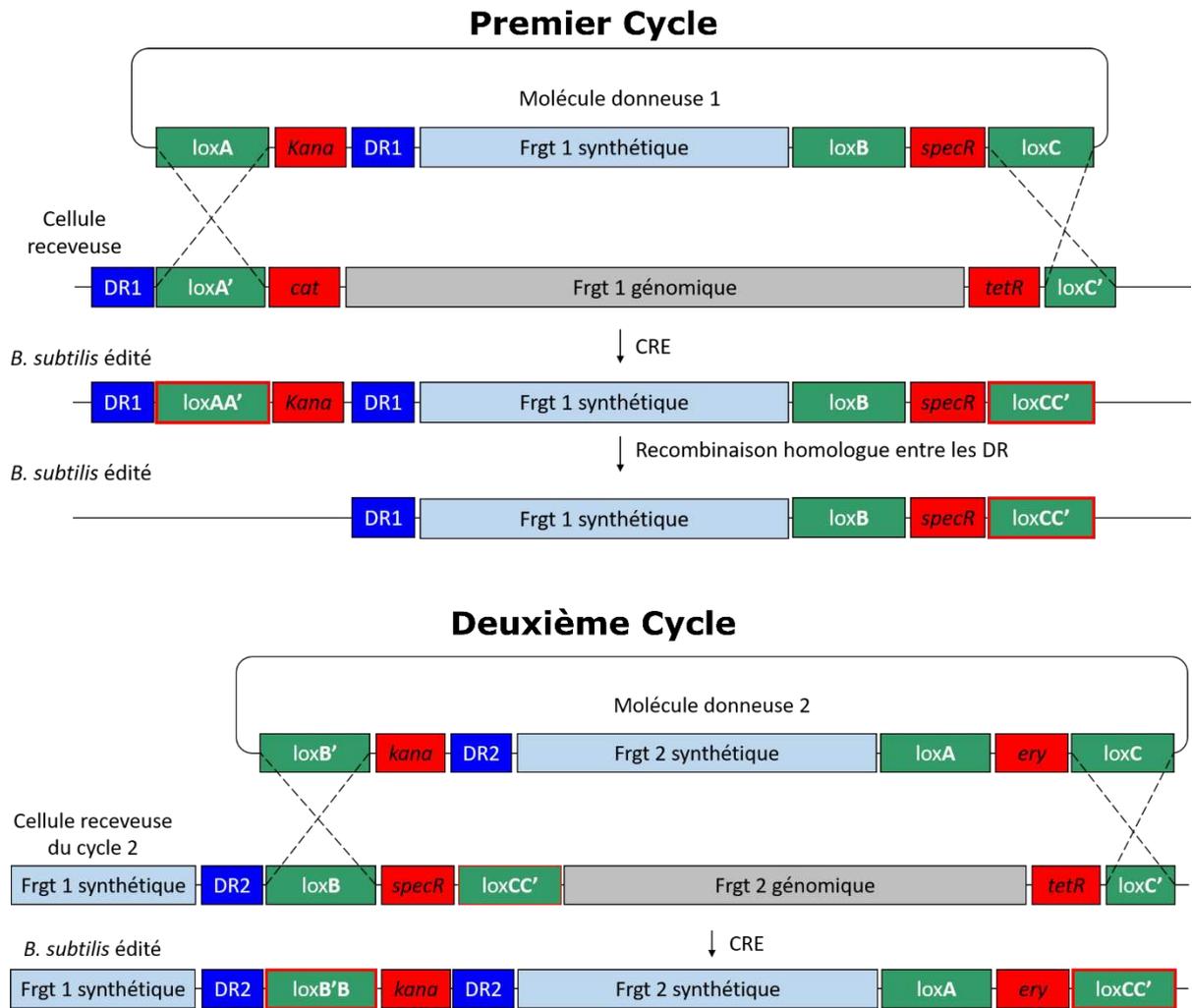


**Figure 117 : Optimisation de la méthode RAGE.** Deux marqueurs sont ajoutés à chaque borne de la molécule donneuse. De plus, des DR (Direct repeat) sont ajoutées afin de pouvoir éliminer les cicatrices laissées après recombinaison des sites lox et recycler les marqueurs de sélection. Cette étape est réalisée par des passages en milieu sans sélection. Durant ces passages, le système de recombinaison homologue de *B. subtilis* induit des événements de recombinaison entre les DR1 (entre eux) et les DR2 (entre eux), permettant l'élimination des deux cicatrices et des deux marqueurs de part et d'autre du fragment échangé.

le générer (après optimisation de la méthode) par étapes successives comme cela a été fait chez *E. coli* avec les méthodes REXER/GENESIS (Fredens *et al.* 2019).

Après avoir essuyé de nombreux échecs qui, comme évoqué dans la discussion du chapitre 3, semblaient plus fortement liés au niveau d'expression de la protéine CRE qu'à un problème de taille de fragments, nous avons finalement réussi à remplacer le fragment 7.2 de 155 kb de la cellule receveuse par le fragment *ad hoc* de 44 kb via le système cre/lox, et ceci sans induire de mutations ailleurs dans le génome. Nous sommes conscients que les efficacités obtenues lors des expériences qui ont permis de sélectionner les clones corrects sont un peu faibles, avec 1 clone positif sur 60 testés pour l'expérience du cinquième essai de RAGE par exemple ; les autres étant des clones hybrides montrant un site lox recombiné au lieu de deux (et parfois aucun) et portant un fragment 7.2 de taille inattendue (généralement entre 44 kb et 155 kb), signifiant que des évènements de recombinaison homologue ont lieu dans ces clones. Mais ces résultats sont extrêmement récents et nous pensons qu'une marge de progression est possible, certaines de nos données suggérant que des améliorations peuvent être apportées au système pour réduire le nombre de faux positifs. Ainsi, que ce soit pour un cycle RAGE unique ou pour des cycles itératifs de RAGE (méthode itRAGE), il serait pertinent :

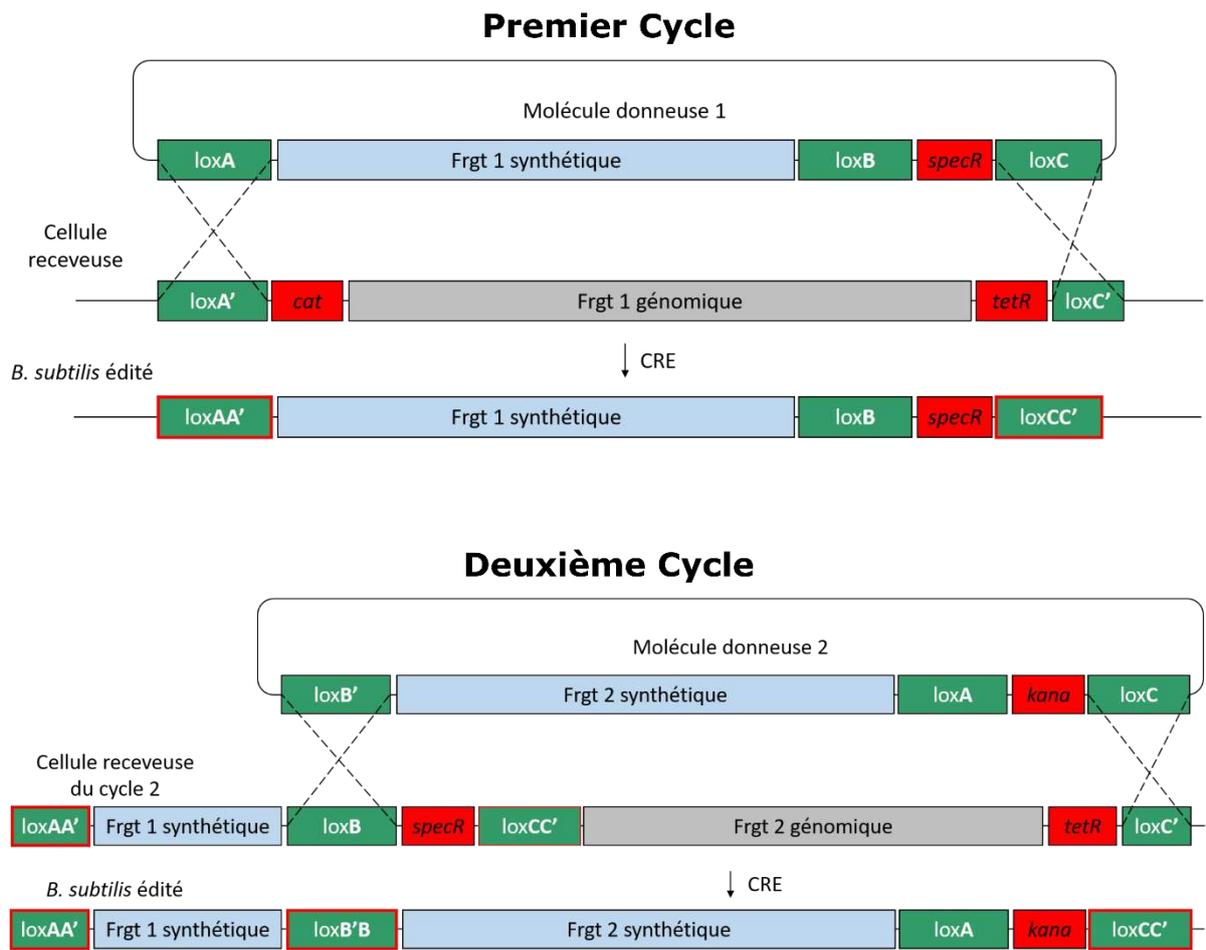
- (1) de revoir le design de la molécule donneuse de manière à y ajouter un marqueur de sélection à chaque borne ainsi que des séquences répétées en tandem pour recycler les marqueurs et les sites lox (figure 117).
- (2) de travailler avec des cellules receveuses dans lesquelles la machinerie de recombinaison homologue fonctionne peu, dans le but de favoriser la recombinaison site-spécifique.
- (3) et pour la méthode itRAGE, d'intégrer dès le premier cycle à la molécule donneuse un « landing pad » qui servirait au cycle suivant de façon à n'avoir à ajouter que le deuxième « landing pad » à chaque cycle. Deux versions du itRAGE peuvent être imaginées : la première est une version complètement optimisée dans laquelle deux marqueurs sont apportés à chaque borne de la molécule donneuse, qui requiert l'incorporation de séquences répétées en tandem afin de recycler un des deux marqueurs (figure 118). Pour pouvoir l'utiliser, l'organisme cible doit tout de même être doté d'une capacité de recombinaison homologue, un minimum efficace, pour recycler les marqueurs de résistance (figure 118). Au contraire, une version plus



**Figure 118 : Schéma de deux cycles de la version optimisée du itRAGE.** Lors du premier cycle, l'échange de fragment permet d'insérer le fragment 1 issu de la molécule donneuse (dans cet exemple, un fragment synthétique) à la place du fragment 1 de la cellule receveuse. Ensuite, la recombinaison homologue entre les deux DR permet d'éliminer le loxAA' et le marqueur *kana*. Cette étape permet notamment de recycler le marqueur *kana*. Le site loxB apporté sur la molécule donneuse sera utilisé comme « landing pad » pour le deuxième cycle. Afin de produire la cellule receveuse pour le deuxième cycle, il suffit d'insérer le « landing pad » (*tetR*, *loxC'*). Le second cycle permet l'insertion du fragment 2 issu de la molécule donneuse à la place du fragment 2 de la cellule receveuse. Les sites loxA, B et C sont incompatibles entre eux. Les sites loxAA', B'B, CC' et encadrés en rouge sont inactifs.

classique du itRAGE qui ne dépend pas de la capacité de recombinaison homologue peut être proposée, dans laquelle un seul marqueur de sélection est amené avec la molécule donneuse (figure 119).

Nous émettons donc l'hypothèse que les faibles efficacités obtenues au cours de nos expériences sont dues à un paramétrage qui n'est pas encore optimal et qui sera testé prochainement. Au-delà du fait que la méthode peut certainement être optimisée, ces faibles efficacités peuvent aussi s'expliquer par le fait que la molécule donneuse avec laquelle nous avons travaillé n'était pas pure (mélange de molécules de 32 kb et 58 kb), et comme mentionné un peu plus haut, que les expériences aient d'emblée été réalisées avec des fragments donneur (molécule donneuse de 58 kb au total) et receveur (155 kb) de grandes tailles. Il est fort possible qu'un échange de fragments de petites tailles eût été bien plus efficace comparé à celui réalisé pendant nos expériences. Néanmoins, la validation de clones corrects lors de nos expériences après criblage d'un nombre de colonies qui restait gérable pour l'expérimentateur (~60 clones) nous interroge désormais sur les tailles maximales des fragments (donneurs et receveurs) avec lesquels il serait possible de travailler. Peut-on imaginer par exemple pouvoir échanger des segments de génomes de 250 kb voire 500 kb par leurs homologues réduits ? Ceci signifierait qu'un génome réduit (ou réduit/recodé) pourrait être obtenu après seulement 6 à 12 cycles, une solution acceptable, à défaut de pouvoir transplanter des génomes entiers. Pour mettre en place une telle stratégie, il serait nécessaire de démontrer non seulement que l'échange médié par le système cre/lox peut se faire, mais aussi que des molécules donneuses de grandes tailles peuvent pénétrer la cellule receveuse. Lors de nos expériences par exemple, la molécule donneuse, d'une taille de 58 kb, a été introduite dans la cellule receveuse en utilisant un protocole basé sur la compétence naturelle de *B. subtilis*. Cependant, il ne nous est pas possible d'affirmer que des molécules de plus grandes tailles pourraient, elles aussi, être introduites dans la cellule receveuse par compétence naturelle. Pour cela, des tests doivent être réalisés. Parallèlement à ces tests, des protocoles de transformation artificiels permettant d'introduire de grandes molécules d'ADN dans une cellule pourraient également être développés. Des protocoles de transformation artificiels utilisant le PEG sont par exemple utilisés chez les mycoplasmes pour transférer des plasmides mais aussi des génomes complets de taille supérieure à 1 Mb (Labroussaa *et al.*, 2019). De façon intéressante, un protocole basé sur le PEG a été décrit en 1979 pour *B. subtilis*



**Figure 119 : Schéma de deux cycles de la version « classique » du itRAGE.** Lors du premier cycle, l'échange de fragment permet d'insérer le fragment 1 issu de la molécule donneuse (dans cet exemple, un fragment synthétique) à la place du fragment 1 de la cellule receveuse. Le site loxB apporté sur la molécule donneuse sera utilisé comme « landing pad » pour le deuxième cycle. Afin de produire la cellule receveuse pour le deuxième cycle, il suffit d'insérer le « landing pad » (*tetR*, loxC'). Le second cycle permet l'insertion du fragment 2 issu de la molécule donneuse à la place du fragment 2 de la cellule receveuse. Les sites loxA, B et C sont incompatibles entre eux. Les sites loxAA', B'B, CC et encadrés en rouge sont inactifs. A noter que cette version du itRAGE n'a pas besoin d'un système de recombinaison homologue efficace pour l'élimination d'un marqueur, mais laisse la cicatrice d'un site lox inactif à chaque cycle.

(Chang and Cohen, 1979). Ce dernier pourrait servir de base au développement d'une nouvelle méthode de transfert d'ADN chez *B. subtilis* permettant de travailler avec des molécules donneuses de grandes tailles comme 250 kb, 500 kb, voire d'environ 1 Mb ou plus.

Parallèlement aux réflexions menées sur l'efficacité de la méthode et son optimisation, nous commençons également à réfléchir à des premières expériences d'échange d'un fragment génomique par un fragment synthétique. En effet, dans la publication originale de Garcia-Morales *et al.*, un fragment de génome synthétique (22 kb) a été assemblé dans la levure et utilisé comme molécule donneuse pour remplacer un fragment de génome de 38 kb chez *M. pneumoniae* (Garcia-Morales *et al.*, 2020). Une telle expérience est désormais envisageable chez *B. subtilis* en travaillant par exemple avec une molécule donneuse de 58 kb qui aurait été entièrement synthétisée et assemblée dans la levure (ou un tout autre fragment mais de taille équivalente, en tous cas pour le moment). Cette réalisation permettrait de faire la preuve de concept qu'un fragment de génome peut être remplacé par son homologue synthétique réduit chez *B. subtilis*, posant ainsi la « première pierre » au projet Bacillus 2.0.

Pour conclure sur cette partie, la méthode RAGE vient offrir un nouvel outil d'ingénierie à grande échelle pour le génome de *B. subtilis* et complète ainsi une boîte à outils déjà bien garnie (cf partie 1.7.2 de l'introduction). Elle nous a permis de montrer une nouvelle application du système *cre/lox* chez *B. subtilis*. En effet, ce système est couramment utilisé pour l'élimination de marqueurs de sélection (Cai and Chen, 2021; Koo *et al.*, 2017) mais n'a, à notre connaissance, jamais été utilisé pour l'échange de grands fragments d'ADN. L'adaptation de cette méthode à *B. subtilis* démontre qu'elle est efficace même chez une bactérie ayant une machinerie de recombinaison homologue performante (tout du moins dans une perspective de réduction du génome en utilisant des fragments de génome donneurs réduits par rapport à ceux de la cellule receveuse). On peut donc maintenant l'utiliser pour de l'échange de fragments de génome dans des cellules capables de recombinaison homologue (*B. subtilis*) ou non (*M. pneumoniae*), tant que des « landing pads » peuvent être introduits de façon ciblée de part et d'autre de la région à échanger. Elle nécessite encore quelques étapes d'optimisation mais si celles-ci s'avèrent possibles, la méthode RAGE pourrait être adaptée à d'autres espèces bactériennes modèles ou non-modèles. Comparée à la méthode d'échange de fragments de génome REXER/GENESIS, qui est basée sur le système de recombinaison



lambda/red et CRISPR-Cas9 (Wang *et al.*, 2016), elle a l'avantage d'exploiter un système de recombinaison (le système cre/lox) fonctionnel dans de nombreuses espèces vivantes, des bactéries aux levures en passant par les plantes et les cellules de mammifères (Yan *et al.*, 2008; Sauer, 1987; Gilbertson, 2003; Sauer and Henderson, 1988) et aussi largement étudié (van Duynes, 2015). Très récemment, par exemple, Cautereels *et al.*, ont établi un ensemble de 16 nouveaux variants symétriques du site loxP utilisables pour l'ingénierie génomique multiplexée chez les procaryotes (*E. coli*) et les eucaryotes (*S. cerevisiae* et *Zea mays*) (Cautereels *et al.*, 2024), laissant entrevoir de nombreuses possibilités pour la construction de génomes largement remaniés ou synthétiques réduits.

### 3. Applications des méthodes de CReasPy-Fusion et RAGE

#### 3.1. Construction de mutants de *B. subtilis* ou autres organismes

Les méthodes de CReasPy-Fusion et de RAGE pourraient être combinées dans un cycle d'ingénierie de génome bactérien à grande échelle pour produire des mutants de *B. subtilis*. Comme illustré figure 113, le cycle serait divisé en 5 étapes : 1) clonage dans la levure par la méthode de CReasPy-Fusion d'un fragment de génome bactérien, de façon spécifique, entre deux sites lox incompatibles, 2) modification du fragment de génome grâce aux outils génétiques disponibles chez la levure, 3) transformation du fragment de génome, éventuellement modifié, dans *E. coli* pour une étape d'amplification de son nombre de copies, 4) échange de fragments de génome dans une cellule bactérienne receveuse via le système de recombinaison site-spécifique cre/lox, et 5) sélection et caractérisation des bactéries modifiées.

A l'heure actuelle, ces deux méthodes fonctionnent séparément chez les mycoplasmes et chez *B. subtilis*. Il est donc fort probable que leur enchaînement (sous forme de cycle unique ou de cycles successifs) soit possible chez ces micro-organismes et peut-être au-delà. Comme démontré au cours de cette thèse, ce cycle pourrait servir à remplacer de grands fragments génomiques par leurs homologues plus ou moins réduits dans le but de générer des mutants d'intérêt. Il pourrait aussi servir à remplacer de grands fragments génomiques par des homologues contenant des mutations ponctuelles ou bien dans lesquels des gènes hétérologues auraient été ajoutés. De même, en modifiant un peu le design du RAGE, on peut aussi envisager faire de l'insertion et ajouter des groupes de gènes permettant



d'implémenter de nouvelles voies de biosynthèse à l'organisme d'intérêt. Comme illustré sur la figure 114, les mutants (de *B. subtilis* ou autres organismes) déficients pour une fonction ou bien réalisant une fonction donnée qui pourraient être issus de ce cycle d'ingénierie sont générés suite à (i) des expériences de transfert de « gènes préexistants » entre organismes (*B. subtilis*/levure, levure/*E.coli*, *E.coli*/*B. subtilis*) et aussi (ii) des expériences de modification de gènes grâce aux outils génétiques disponibles. De par sa vision, ses objectifs, mais aussi ses outils et techniques, le domaine de la BS révolutionne peu à peu ce champ en offrant la possibilité de s'abstraire de la contrainte « d'ADN préexistant » et de passer par la technologie dite de « synthèse de l'ADN ». Bien que les coûts actuels soient encore élevés, de plus en plus de chercheurs commandent des cassettes d'ADN à assembler ou bien « prêtes à l'emploi » auprès d'entreprises spécialisées dans la fabrication de séquences d'ADN (Gene foundries).

Les techniques de synthèse/assemblage d'ADN pourraient être associées aux méthodes de CRISPR-Cas9 et de RAGE et fournir plus de flexibilité pour la construction de mutants mais aussi de cellules contrôlées par des génomes semi-synthétiques voire entièrement synthétiques

### 3.2. Construction de cellules contrôlées par des génomes entièrement synthétiques ou semi-synthétiques

#### 3.2.1. Construction de cellules contrôlées par des génomes entièrement synthétiques

Dans le domaine de la BS, la construction de « châssis microbiens » est l'un des principaux axes de recherche. Un châssis microbien peut être défini comme un micro-organisme modulable dans lequel des circuits génétiques peuvent être introduits, maintenus et exprimés grâce à la machinerie cellulaire de l'hôte. L'un des moyens d'améliorer les châssis microbiens consiste à produire des « châssis simplifiés », contrôlés par des génomes réduits, ou des « châssis minimaux » contrôlés par des génomes minimaux, c'est-à-dire dépourvus de tous les gènes non-essentiels. L'idée derrière cette minimisation du génome est de réduire au maximum les interactions possibles entre les circuits génétiques nouvellement intégrés et le réseau métabolique de l'hôte. De plus, la réduction du génome de l'hôte permettrait d'augmenter la production d'une molécule d'intérêt (Morimoto *et al.*, 2008).



En ce qui concerne la quête d'une cellule minimale, le résultat le plus marquant a été obtenu en 2016 avec la construction de la première cellule minimale contrôlée par un génome synthétique réduit, d'une taille de 531 kb, ce qui en fait l'organisme capable d'auto-réplication contrôlé par le plus petit génome (Hutchison *et al.*, 2016). Ce travail a été réalisé en combinant trois grandes technologies du domaine de la BS : 1) la technologie de synthèse d'ADN, 2) la technologie d'assemblage de génomes synthétiques dans la levure, qui nécessitait donc de pouvoir cloner entièrement le génome bactérien dans la levure et 3) la technologie de TG qui consiste à transférer un génome entier (synthétique ou naturel) dans une cellule bactérienne receveuse afin d'obtenir une cellule modifiée contrôlée par le génome synthétique ou naturel. Adapter de tels outils à des organismes d'intérêt industriel tels que *B. subtilis* constituerait une avancée majeure et ouvrirait la voie à la construction de châssis microbiens synthétiques à visée industrielle, dont l'avantage serait de contrôler l'ensemble de leurs propriétés biologiques. Mais comme nous l'avons vu au cours de ces travaux de thèse, ce n'est pas trivial. A l'heure actuelle, de nombreux génomes viraux, bactériens et eucaryotes ont été clonés dans la levure mais aucun de taille supérieure à 1.8 Mb (Vashee *et al.*, 2020) et la technologie de TG n'est disponible que chez les mycoplasmes (Labroussaa *et al.*, 2016).

Ainsi, développer une approche basée sur le remplacement de fragments de génome et non sur le remplacement d'un génome complet semble être une bonne alternative. Un exemple récent et marquant utilisant ce type d'approche est la construction d'une cellule d'*E. coli* contrôlée par un génome réduit et entièrement recodé (Fredens *et al.*, 2019), un travail réalisé grâce aux méthodes REXER/GENESIS basées sur les systèmes d'édition lambda-red et CRISPR-Cas9. Au regard de tels résultats, envisager de développer une stratégie similaire, basée sur la méthode itRAGE pour la production de cellules contrôlées par un génome synthétique de *B. subtilis* (ou autres organismes, à compter que la méthode itRAGE fonctionne chez ces organismes), semble donc réaliste. Cette stratégie serait découpée en 4 étapes (figure 115) : 1) conception assistée par ordinateur d'un génome bactérien synthétique, 2) assemblage de fragments d'ADN synthétique de 10 kb en grands fragments de 100 kb dans la levure, 3) modification des fragments de génome par les outils génétiques disponibles chez la levure (optionnel), 4) échanges itératifs de fragments de génome grâce à la stratégie itRAGE dans une cellule receveuse afin d'obtenir une cellule contrôlée par le génome synthétique conçu à l'étape 1.



Une approche itérative paraît plus longue que l'assemblage d'un génome entier dans la levure suivi de sa transplantation, mais elle possède l'avantage de pouvoir vérifier après chaque cycle la viabilité des cellules et donc s'assurer que le fragment synthétique nouvellement apporté dans la cellule est correct.

### 3.2.2. Construction de cellules contrôlées par des génomes semi-synthétiques (projet Bacillus 2.0)

Dans le cadre du projet Bacillus 2.0, le consortium a fait le choix de construire un génome semi-synthétique (plutôt qu'un génome entièrement synthétique) pour deux raisons. D'une part, il souhaitait réduire les coûts liés à la synthèse d'ADN. D'autre part, il avait à disposition une collection de souches de *B. subtilis* à génomes réduits construites par réduction itérative à partir du génome de la souche parentale *B. subtilis* 168 (Dervyn *et al.*, 2023; Reuß *et al.*, 2017). Au sein de cette collection, les génomes les plus réduits pouvaient donc servir de matériel de départ pour récupérer des sections de génome de *B. subtilis* ne retenant déjà que (ou quasi que) des gènes essentiels, qu'il n'était pas nécessaire de synthétiser à nouveau.

Le consortium a commencé à récupérer certaines sections, par PCR puis assemblage d'amplicons chevauchants dans la levure. Désormais, il est aussi possible de les capturer directement en utilisant la méthode de CReasPy-Fusion. Une fois dans la levure, ces fragments natifs pourront être modifiés, si besoin, grâce aux outils génétiques disponibles dans la levure, ou utilisés tels quels.

Pour ce projet, si le plan ne change pas, nous devrions donc aboutir à une banque de clones levures hébergeant soit des fragments 100% synthétiques, soit des fragments natifs réduits (figure 114) ; ceci n'est pas problématique et revient au même d'avoir une banque de clones levures hébergeant que des fragments synthétiques ou une banque de clones levures hébergeant que des fragments natifs réduits. En revanche, à partir de là, deux stratégies peuvent être envisagées selon les avancées du projet.

#### 1) Première stratégie

Si la méthode RAGE ne permet pas d'introduire et d'échanger des fragments de taille supérieure à 200 kb dans la cellule à éditer, alors le cycle d'ingénierie sera le suivant : 1) conception d'un génome de *B. subtilis* semi-synthétique minimal sur ordinateur, 2)



assemblage de fragments d'ADN chevauchants d'environ 10 kb en fragments de 100 kb dans la levure, ou capture de fragments d'environ 100 kb à partir de génome issus de souches de *B. subtilis* minimales, 3) si besoin, modification de ces fragments en utilisant les outils génétiques disponibles chez la levure, 4) échanges de fragments de façon itérative grâce à la méthode itRAGE dans une cellule receveuse afin d'obtenir une cellule contrôlée par un génome complètement modifié (figure 114).

## 2) Deuxième stratégie

En revanche, si la méthode RAGE est peu ou pas limitée par la taille des fragments avec lesquels il est possible de travailler, alors la souche châssis (*SynBsu*, ~2Mb) pourrait être construite en quelques cycles et plus exactement :

- 8 cycles si des fragments de 250 kb peuvent être échangés.
- 4 cycles si des fragments de 500 kb peuvent être échangés.
- 2 cycles si des fragments de 1 Mb peuvent être échangés.
- Voire un cycle si des fragments encore plus grands peuvent être échangés. Les progrès faits sur la partie « CreasPy-Fusion » suggèrent qu'une version complète ou quasi-complète (*i.e.* sans le fragment 7 ou avec un fragment 7 modifié) du génome de *B. subtilis* pourrait être clonée dans la levure dans le futur. Si tel est le cas, nous pourrions tenter d'utiliser la méthode RAGE (basée sur la compétence) ou peut-être une variante (basée sur un protocole de transformation artificiel) afin de provoquer un échange de la quasi-totalité du génome (version sans le fragment 7) dans la cellule receveuse. Le génome résultant de cet échange serait hybride car le fragment 7 proviendrait de la cellule receveuse et le reste du génome correspondrait à la version semi-synthétique assemblée dans la levure. Une telle souche pourrait être utilisée telle quelle en tant que châssis.

Il est également possible de modifier la région du fragment 7 en supprimant certains gènes ou bien en les déplaçant ailleurs dans le génome de façon à appauvrir des régions riches en séquences répétées. En effet, lors des expériences de TAR-cloning réalisées par le Dr. Guesdon, la totalité du fragment 7 a été cloné dans la levure en sous-fragments, ce qui semble indiquer que, séparément, les régions ne posent pas de problèmes alors que réunies elles sont délétères pour la levure. On peut aussi envisager d'assembler une version synthétique ou semi-



synthétique réduite de chacun de ces sous-fragments dans la levure. A partir de ces versions, il serait alors possible d'utiliser la méthode itRAGE afin d'échanger de façon itérative l'ensemble de la région du fragment 7 par les versions synthétiques ou semi-synthétiques réduites assemblées dans la levure. Ce travail aboutirait à la construction d'une souche contrôlée par un génome entièrement modifié.

Cependant, une telle stratégie requiert quelques ajustements notamment au niveau de la méthode RAGE :

- Travailler avec une cellule receveuse qui n'est plus capable de faire de la recombinaison homologe. En effet, l'ensemble des gènes présents dans la version semi-synthétique du génome serait aussi présent dans le génome de la cellule receveuse, ce qui pourrait être à l'origine d'évènement de recombinaison homologe intempestifs.
- Utiliser plusieurs marqueurs répartis de façon homogène sur l'ensemble du génome semi-synthétique réduit quasi complet afin de forcer son maintien.
- Développer un protocole de transformation efficace afin de transférer le génome semi-synthétique réduit quasi complet assemblé dans la levure, dans la cellule receveuse de *B. subtilis*

Cette seconde stratégie, beaucoup plus ambitieuse que la première, se rapproche de celle proposée dans le projet initial par la construction et l'utilisation d'un génome quasi-complet. Elle est également plus incertaine, car même si nous arrivons désormais à cloner des fragments du génome de *B. subtilis* d'environ 1 Mb dans la levure (en CREasPy-Fusion), nous n'avons pas encore démontré que de tels fragments de génomes peuvent ensuite être mobilisés pour du RAGE.

#### 4. Conclusion

Pour conclure, ce projet de thèse a consisté à adapter à *B. subtilis* des outils de BS initialement développés chez les mycoplasmes pour l'ingénierie de grands fragments de génome.



Les premiers travaux réalisés au cours de ce projet ont permis d'adapter la méthode de CReasPy-Fusion à *B. subtilis* pour la capture de façon spécifique de fragments de génome bactérien dans la levure. Cette adaptation a été l'occasion de réaliser des observations en microscopie dans le but de comprendre les mécanismes de transfert d'ADN lors d'un contact cellules/cellules. Par la suite l'utilisation de la méthode de CReasPy-Fusion a permis d'initier un travail de capture de fragments de génome de tailles croissantes, qui espérons-le nous mènera vers la capture d'une version complète (modifiée) du génome de *B. subtilis* dans la levure. Les derniers travaux réalisés au cours de la thèse ont, quant à eux, permis d'adapter la méthode RAGE à *B. subtilis* et démontrer que le génome de *B. subtilis* pouvait être modifié par section.

En combinant ces deux méthodes, il semblerait désormais qu'un cycle d'ingénierie du génome de *B. subtilis* à grande échelle, utilisant la levure *S. cerevisiae*, puisse être mis en place, offrant une voie alternative au projet Bacillus 2.0. Au regard des difficultés rencontrées, cette solution est acceptable à ce jour. Mais, il faut avouer que certains progrès réalisés au cours de la mise en place des méthodes CReasPy-Fusion et RAGE, et en particulier de la méthode CReasPy-Fusion avec le clonage d'un fragment de génome de *B. subtilis* d'environ 1 Mb (soit 1/3 de son génome) nous laisse penser que le clonage d'un génome réduit/modifié de *B. subtilis* est redevenu un objectif réalisable chez la levure. Si la méthode de TG peut ensuite être mise en place chez *B. subtilis* alors le cycle présenté en figure 1 pourrait pour la première fois devenir opérationnel chez une bactérie Gram+.



# Annexes



## Annexes

### 1. Liste des abréviations

**ACS** : ARS Consensus Sequence

**A, C, T, G, U** : Adénine, Cytosine, Thymine, Guanine, Uracile

**ADE** : Adénine

**AND** : Acide DésoxyRiboNucleique

**ADNdb** : AND double brin

**ADNg** : AND génomique

**ADNr** : ADN ribosomique

**ADNsb** : AND simple brin

**ARNg** : ARN guide

**ARNt** : ARN de transfert

**ARS** : Autonomously Replicative Sequence

**ATP** : Adénosine Tri-Phosphate

**BAC** : Bacterial Artificial Chromosome

**BET** : Bromure d'Ethidium

**BS** : Biologie de Synthèse

**Cas** : CRISPR associated protein

**CDB** : cassure double brin

**CDS** : Coding DNA Sequence

**CEN** : centromère

**CRISPR** : clustered regularly interspaced short palindromic repeat

**crRNA** : ARN CRISPR

**DBTL** : Design Built Test Learn

**DMSO** : Diméthylsulfoxyde

**DO** : Densité Optique

**DO supplement** : Drop Out supplement

**DR** : Direct repeat (séquences répétées directes)

**GFP** : Green Fluorescent Protein

**GRAS** : Generally Recognized as Safe



**HIS** : Histidine  
**JCVI** : John Craig Venter Institute  
**LB** : Lysogeny Broth  
**LYS** : Lysine  
**LEU** : Leucine  
**MC** : Modified Competence  
**NHEJ** : Non-homologous end joining  
**NEB** : New England Biolabs  
**NGS** : Séquençage Nouvelle Génération  
**PAM** : Protospacer Adjacent Motif  
**PCR** : Polymerase Chain Reaction  
**PEG** : PolyEthylene Glycol  
**PFGE** : Pulsed-Field Gel Electrophoresis  
**PMSF** : fluorure de phénylméthylsulfonyle  
**Qsp** : Quantité Suffisante Pour  
**RAGE** : Recombinase-Assisted Genome Engineering  
**RBS** : Ribosome Binding Site  
**RFP** : Red Fluorescent Protein  
**RMCE** : Recombinase Mediated Cassette Exchange  
**Rpm**: Rotation Par Minute  
**SNP** : single nucleotide polymorphism  
**TAR** : Transformation-Associated Recombinaison  
**TEL** : sequence télomérique  
**TG** : transplantation de génome  
**tracrRNA** : trans-activating crRNA  
**TRP** : tryptophane  
**URA** : Uracile  
**UV** : Ultra Violet  
**YAC** : Yeast Artificial Chromosome  
**YCp** : plasmide centromérique de levure  
**YPDA** : Yeast Peptone Dextrose Adenine



## 2. Liste des amorces utilisées au cours de ces travaux

Tableau S1 : Listes des amorces utilisées au cours de cette thèse

Nom de l'amorce	Séquence
p426_insertGibson_F	GTACAATCTGCTCTGATGCCGGCTGGAGCTTCTTTGAAAAG
p426_insertGibson_R	GTCGGGGCTGGCTTAACTATGGTACCGGCCGCAAATTAAG
Div_p426_vecteurGibson_F	CTTTAATTTGCGGCCGGTACCATAGTTAAGCCAGCCCCGAC
Divp426_vecteurGibson_R	CTTTTCAAAGAAGCTCCAGCCGGCATCAGAGCAGATTGTAC
Gib_loxm2/71_pCC1_F	TACCGTTCGTATATGGTTTCTTATACGAAGTTATTACAGCCAGTAGTGCTC
Gib_Ade2_pCC1_R	GCACTCTCAGTACAATCTGCGCATCCGCTTACAGACAAGC
Gib_pCC1_Ade2_F	GCTTGTCTGTAAGCGGATGCGCAGATTGACTGAGAGTGC
Gib_CreT7_Ade2_R_V2	CCTATAGTGAGTCGTATTAGCATCTGTGCGGTATTTTAC
Gib_Ade2_CreT7_F_V2	GTGAAATACCGCACAGATGCTAATACGACTCACTATAGG
Gib_lox66-CreT3_R	TACCGTTCGTATAATGTATGCTATACGAAGTTATAATTAACCCTCACTAAAGG
Gib_lox66_Spec_F	ATAACTTCGTATAGCATACATTATACGAACGGTATGCCTCAAGCTAGAGAGTCG
Gib_loxm2/71_Spec_R	ATAACTTCGTATAAGAAACCATATACGAACGGTACCTGAACGAGAAGCTATCAC
Gib pCmR-pTwist	GAAACCATATACGAAGTTATTGACCTCAGAACTCCATCTG
Gib pCmR (fin CmR)-pTwist-N	GATTCTAGTTCCTCCTTAAGTTGTGTGGAATTGTGAGCG
GibpTwist-pCmR (fin CmR)-N	CGCTCACAATCCACACAACCTAAAGGAGGAACTAGAATC
Gib pCmR-promoteur TrmF-N	TATCGGACTTCGGTTTTATGGTTGTGTGGAATTGTGAGCG
Gib promoteur TrmF-pCmR-N	CGCTCACAATCCACACAACCATAAAACCGAAGTCCGATA
Gib promoteur TrmF-pTwist	TGGCCTGCTTCATACCCCATGTTGTTACCTCTCTTTACT
GibpTwist-promoteur TrmF	AGTAAAGAGAGGTGAACAACATGGGGTATGAAGCAGGCCA
Gib pTwist-pCmR	CAGATGGAGTTCGTAGGTCAATAACTTCGTATATGGTTTC
KLD loxm2/66 tetR lox71 F	TTCTTATACGAACGGTACTTTCCGTAAGGGCAAAG
KLD loxm2/66 TetR lox71 R	ACCATATACGAAGTTATGTGTACATTTGGAGCATAG
KLD cre-R	CTAGAGGGAAACCGTTGTG
KLD cre-F	ACTAGTATGCTAGCCCTG
pDR244 R / Ery_F2	GAGCTAAAGAGGTCCATGACGAAAATCGCCATTCGCCAGG
pDR244 F / Ery_R2	CGATCCAGTCCAATAATCGTCGCAATAAACCCCTGCATAGGG
Ery_R2 / pDR244 F	CCCTATGCAAGGGTTTATTGCGACGATTATTGGACTGGATCG
Ery_F2 / pDR244 R	CCTGGCGAATGGCGATTTTCGTCATGGACCTCTTTAGCTC
ARS404 Kana/FrIB	AGAAAAAGTCTAAAAATGGTTTTTTTCATCCAAAATATTAATTTTACTTTTACA GGCGAGAAAGGAGAGATC
kana/YurQ	ATTGAACAGCGAAGCCAGAGGGTTTTTCGCCGTATCTGTG
pDR111/YurQ	CAGCCGCAAACCTGCTTTGTAATCATCCTGTCTCTTGATCC
ORI/FrIB	CTTGCTCCGCTCGTTCTCAAGGGAAATGTGCGCGGAACCC
FrIB/ORI	GGGTTCCGCGCACATTTCCCTTGAGAACGAGCGGAGCAAG
FrIB/ARS404 TetR	TAAAAGTAAATTTAATATTTTGGATGAAAAAACCATTTTTAGACTTTTTCTCTTA AGTCAAGTTCTGGCTT
YurQ/pDR111	GGATCAAGAGACAGGATGAGTACAAAGCAGTTTGGCGGCTG
YurQ/Kana	CACAGATACGGCGAAAAACCCCTCTGGCTTCGCTGTTCAAT
Kana/LYS	GCACTCTCAGTACAATCTGCGGTTTTTCGCCGTATCTGTG
YurQ/LYS	GTGAAATACCGCACAGATGCCTCTGGCTTCGCTGTTCAAT
LYS/YurQ	ATTGAACAGCGAAGCCAGAGGCATCTGTGCGGTATTTTAC
LYS/Kana	CACAGATACGGCGAAAAACCGCAGATTGTAAGTACTGAGAGTGC
GFP/vecteur	GTTTTCCAGTACGACGTTGTAACGACGCGCCAGTGAG
GFP/TEF1	GCATAGCAATCTAATCTAAGCTAGAATGTCTAAAGGTGAA
vecteur/GFP	CTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAAC
TEF1/GFP	TTCACCTTTAGACATTCTAGCTTAGATTAGATTGCTATGC
GFP/SNR52	GCAGTGAAAGATAAATGATCCTAGAATGTCTAAAGGTGAA
SNR52/GFP	TTCACCTTTAGACATTCTAGGATCATTTATCTTTCACTGC
Ery/yurQ	ATTGAACAGCGAAGCCAGAGCGACGATTATTGGACTGGATCG



Nom de l'amorce	Séquence
cre/FrlB	AAGCCAGAACTTGACTTAAGCAATACGCAAACCGCCTCTC
YurQ/Ery	CGATCCAGTCCAATAATCGTCGCTCTGGCTTCGCTGTTCAAT
FrlB/cre	GAGAGGCGGTTTGCATATTGCTTAAGTCAAGTTCTGGCTT
jim/cre	CAATTGACGGCTTGACGGAGCAGACCCCGTAGAAAAGATC
Ery_F2 / pDR244 R	CCTGGCGAATGGCGATTTTCGTCATGGACCTCTTTAGCTC
cre/jim	GATCTTTTCTACGGGGTCTGCTCCGTCAAGCCGTCAATTG
pDR244 R / Ery_F2	GAGCTAAAGAGGTCCATGACGAAAATCGCCATTCGCCAGG
Verif_pShuttle_F2	CTCTTCGCTATTACGCCAGC
Verif_pShuttle_R	CAAGTGTTTCGCTTCGCTCTCA
Verif_sbo_F	CTCGTTCACCGTTTCCGATG
Verif_sbo_R2	TTATCTTGCTCTTCTGCTC
Verif_sbo_SPEC_F2	GAGCCATTATGGATTTCGTCAG
Verif_sbo_R	AGTACAAGGGACCATTACGC
screening_ARS_CEN_HIS_F	GATGTTTCGTTTCCCCGTATG
screening_ARS_CEN_HIS_R	GGACCATTACGCTTCTTGTC
Verif_Fgt5-jonction_F	CGTTTGTCTGATGGTGTCT
Verif_Fgt5-jonction_R	TGGCCCGTCAAACAGAAGAC
ScreenPST_Bsu_sbo_F	GTTATCGGTGTCTTTGCTTC
ScreenPST_Bsu_sbo_R	TGACCGTGTACCGCACAAGTG
Verif_Fgt4_jonction_R	CTTGTTGAGAAACAGCGGAC
T3	GCAATTAACCCTCACTAAAG
T7	TAATACGACTCACTATAGGG
Verif_jonction frgt 3-5 R	GGTGATTGTTCTGAATCTG
Verif_jonction frgt 2-5_R	GCTGTCTGTTGTCTCATCTG
Bsu_Frgt2.2_R	GCGACGTTTGGGTTTTTTGG
Bsu_Frgt2.2_F	CAGTAGGAGCTGTTTTCGGC
screening ARS404 F	CTTCAAAGGGTTGGCCTCCT
Verif_demi-genome-jonction_R2	CGTCAGTATCGGCATCGGTC
PST_pCC1pYAdCreSpec7.2_F	ATGGCTTTTGCCTCATGAACGAGGCTGTTTCAGAAGAGATGATTTTCTACGTTT GGCCGGCCTACCGTTCGTATATGGTTTC
PST_pCC1pYAdCreSpec7.2_R	ATTCGAGGAAGAGAAGATAAAGGAATTCAAAGCTCCTCCGTCTGATAGAG GATGAAGGAGGCTGAACGAGAAGCTATCAC
Verif LYS_F	GAACTTGGTAGCGTCTCAAC
Verif 7.2-ADE_R	GTATCACAAGAGTACAGGCG
Verif 7.2-pCC1_F	GAATTTGTCTGCTTGCAC
Verif-pCC1_R	GGATTTTCTGCTGTGGACA
Verif receveuse F	GCCGACAGACGGATTACATG
Verif receveuse R	GGGAGCTTTTAGGGGAAATC
FE220 receveuse YxeQ F	CGTGCCATCACTTCAACACC
FE220 receveuse YxeP R	CAATACTTTGCGGAGCTGGC
screening landing pad tet F	CCCCATCTGTTTCCCAAAC
seq échange CmR R	GAATCGATCTTTCACAGGCG
Verif_htrc_R	CTTGCCATTGAGATGAAATCGC
ppaC_F	GTTGTGGCTGAGAAGAATC
trme_frgt1_bamHI_R	CAAGGTTTTGCCGTTTAAG



# Références



## Références

- Adetunji, A.I., Olaniran, A.O., 2021. Production strategies and biotechnological relevance of microbial lipases: a review. *Brazilian J. Microbiol.* 52, 1257–1269. <https://doi.org/10.1007/s42770-021-00503-5>
- Ahmadi, M.K., Pfeifer, B.A., 2016. Recent progress in therapeutic natural product biosynthesis using *Escherichia coli*. *Curr. Opin. Biotechnol.* 42, 7–12. <https://doi.org/10.1016/j.copbio.2016.02.010>
- Akinsemolu, A.A., Onyeaka, H., Odion, S., Adebajo, I., 2024. Exploring *Bacillus subtilis*: Ecology, biotechnological applications, and future prospects. *J. Basic Microbiol.* 64. <https://doi.org/10.1002/jobm.202300614>
- Albert, H., Dale, E.C., Lee, E., Ow, D.W., 1995. Site-specific integration of DNA into wild-type and mutant *lox* sites placed in the plant genome. *Plant J.* 7, 649–659. <https://doi.org/10.1046/j.1365-313x.1995.7040649.x>
- Anagnostopoulos, C., Spizizen, J., 1961. Requirements for Transformation in *Bacillus Subtilis*. *J. Bacteriol.* 81, 741–746. <https://doi.org/10.1128/jb.81.5.741-746.1961>
- Arkin, A., Berry, D., Church, G., Arnold, F., Boldt, J., Müller, O., Ellington, A.D., Endy, D., Fussenegger, M., Gold, E.R., Greenwood, J., Lee, S.Y., Lim, W., Minshull, J., Murray, T.H., Poste, G., Prather, K.L.J., El-Samad, H., Smolke, C., Weiss, R., 2009. What's in a name? *Nat. Biotechnol.* 27, 1071–1073. <https://doi.org/10.1038/nbt1209-1071>
- Ashlee M. Earl, 2008. Ecology and genomics of *Bacillus subtilis*. *J. Int. Coll. Surg.* 24, 451–456. <https://doi.org/10.1016/j.tim.2008.03.004>
- Atsumi, S., Hanai, T., Liao, J.C., 2008. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* 451, 86–89. <https://doi.org/10.1038/nature06450>
- AUSUBEL, F.M., 1974. Radiochemical Purification of Bacteriophage  $\lambda$  Integrase. *Nature* 247, 152–154. <https://doi.org/10.1038/247152a0>
- Babasaki, K., Takao, T., Shimonishi, Y., Kurahashi, K., 1985. Subtilosin a, a new antibiotic peptide produced by *Bacillus subtilis* 168: Isolation, structural analysis, and biogenesis. *J. Biochem.* 98, 585–603. <https://doi.org/10.1093/oxfordjournals.jbchem.a135315>
- Badel, C., Da Cunha, V., Oberto, J., 2021. Archaeal tyrosine recombinases. *FEMS Microbiol. Rev.* 45, 1–27. <https://doi.org/10.1093/femsre/fuab004>
- Bais, H.P., Fall, R., Vivanco, J.M., 2004. Biocontrol of *Bacillus subtilis* against Infection of *Arabidopsis* Roots by *Pseudomonas syringae* Is Facilitated by Biofilm Formation and Surfactin Production. *Plant Physiol.* 134, 307–319. <https://doi.org/10.1104/pp.103.028712>
- Barbe, V., Cruveiller, S., Kunst, F., Lenoble, P., Meurice, G., Sekowska, A., Vallenet, D., Wang, T., Moszer, I., Médigue, C., Danchin, A., 2009. From a consortium sequence to a unified sequence: The *Bacillus subtilis* 168 reference genome a decade later. *Microbiology* 155, 1758–1775. <https://doi.org/10.1099/mic.0.027839-0>
- Barrangou, R., Horvath, P., 2017. A decade of discovery: CRISPR functions and applications. *Nat. Microbiol.* 2, 1–9. <https://doi.org/10.1038/nmicrobiol.2017.92>
- Batty, C.J., Bachelder, E.M., Ainslie, K.M., 2021. Historical Perspective of Clinical Nano and Microparticle Formulations for Delivery of Therapeutics. *Trends Mol. Med.* 27, 516–519. <https://doi.org/10.1016/j.molmed.2021.04.002>



- Benders, G.A., Noskov, V.N., Denisova, E.A., Lartigue, C., Gibson, D.G., Assad-Garcia, N., Chuang, R.Y., Carrera, W., Moodie, M., Algire, M.A., Phan, Q., Alperovich, N., Vashee, S., Merryman, C., Venter, J.C., Smith, H.O., Glass, J.I., Hutchison, C.A., 2010. Cloning whole bacterial genomes in yeast. *Nucleic Acids Res.* 38, 2558–2569. <https://doi.org/10.1093/nar/gkq119>
- BERCHE, P., 2016. Histoire de la Biologie. *Popul.* (French Ed. 43, 481. <https://doi.org/10.2307/1532842>
- Bernheim, A., 2017. Incontournables et pourtant rares : les déterminants de la distribution des systèmes CRISPR-Cas dans les génomes bactériens. *Biol. Aujourd'hui.* 211, 255–264. <https://doi.org/10.1051/jbio/2018005>
- Billingsley, J.M., DeNicola, A.B., Tang, Y., 2016. Technology development for natural product biosynthesis in *Saccharomyces cerevisiae*. *Curr. Opin. Biotechnol.* 42, 74–83. <https://doi.org/10.1016/j.copbio.2016.02.033>
- Blake, C., Christensen, M.N., Kovacs, A.T., 2021. Molecular aspects of plant growth promotion and protection by *Bacillus subtilis*. *Mol. Plant-Microbe Interact.* 34, 15–25. <https://doi.org/10.1094/MPMI-08-20-0225-CR>
- Blattner, F.R., Plunkett, G., Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B., Shao, Y., 1997. The complete genome sequence of *Escherichia coli* K-12. *Science* (80-. ). 277, 1453–1462. <https://doi.org/10.1126/science.277.5331.1453>
- Boni, L.T., Hui, S.W., 1987. The Mechanism of Polyethylene Glycol-Induced Fusion in Model Membranes BT - Cell Fusion, in: Sowers, A.E. (Ed.), . Springer US, Boston, MA, pp. 301–330. [https://doi.org/10.1007/978-1-4757-9598-1\\_14](https://doi.org/10.1007/978-1-4757-9598-1_14)
- Bordes, P., Genevaux, P., 2017. Addiction des systèmes toxine-antitoxine envers un chaperon chez *Mycobacterium tuberculosis*. *Medecine/Sciences* 33, 486–488. <https://doi.org/10.1051/medsci/20173305009>
- Breuer, M., Earnest, T.M., Merryman, C., Wise, K.S., Sun, L., Lynott, M.R., Hutchison, C.A., Smith, H.O., Lapek, J.D., Gonzalez, D.J., Crécy-Lagard, V. de, Haas, D., Hanson, A.D., Labhsetwar, P., Glass, J.I., Luthey-Schulten, Z., 2019. Essential metabolism for a minimal cell. *Elife* 8, 1–75. <https://doi.org/10.7554/ELIFE.36842>
- Brno, I.G.E.M.T., Mendel, G., 2020. How to handle *Bacillus subtilis*.
- Brown, W.R.A., Lee, N.C.O., Xu, Z., Smith, M.C.M., 2011. Serine recombinases as tools for genome engineering. *Methods* 53, 372–379. <https://doi.org/10.1016/j.ymeth.2010.12.031>
- Burke, D.T., Carle, G.F., Olson, M. V., 1987. Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science* (80-. ). 236, 806–812. <https://doi.org/10.1126/science.3033825>
- Burkholder, P.R., Giles, N.H., 1947. Induced Biochemical Mutations in *Bacillus subtilis*. *Am. J. Bot.* 34, 345–348. <https://doi.org/10.2307/2437147>
- Burley, S.K., Berman, H.M., Bhikadiya, C., Bi, C., Chen, L., Di Costanzo, L., Christie, C., Duarte, J.M., Dutta, S., Feng, Z., Ghosh, S., Goodsell, D.S., Green, R.K., Guranovic, V., Guzenko, D., Hudson, B.P., Liang, Y., Lowe, R., Peisach, E., Periskova, I., Randle, C., Rose, A., Sekharan, M., Shao, C., Tao, Y.P., Valasatava, Y., Voigt, M., Westbrook, J., Young, J., Zardecki, C., Zhuravleva, M., Kurisu, G., Nakamura, H., Kengaku, Y., Cho, H., Sato, J., Kim, J.Y., Ikegawa, Y., Nakagawa, A., Yamashita, R., Kudou, T., Bekker, G.J., Suzuki, H., Iwata, T., Yokochi, M., Kobayashi, N., Fujiwara, T., Velankar, S., Kleywegt, G.J., Anyango, S., Armstrong, D.R., Berrisford, J.M., Conroy, M.J., Dana,



- J.M., Deshpande, M., Gane, P., Gáborová, R., Gupta, D., Gutmanas, A., Koča, J., Mak, L., Mir, S., Mukhopadhyay, A., Nadzirin, N., Nair, S., Patwardhan, A., Paysan-Lafosse, T., Pravda, L., Salih, O., Sehnal, D., Varadi, M., Vřeková, R., Markley, J.L., Hoch, J.C., Romero, P.R., Baskaran, K., Maziuk, D., Ulrich, E.L., Wedell, J.R., Yao, H., Livny, M., Ioannidis, Y.E., 2019. Protein Data Bank: The single global archive for 3D macromolecular structure data. *Nucleic Acids Res.* 47, D520–D528. <https://doi.org/10.1093/nar/gky949>
- Cai, M.Z., Chen, P.T., 2021. Novel combined Cre-Cas system for improved chromosome editing in *Bacillus subtilis*. *J. Biosci. Bioeng.* 132, 113–119. <https://doi.org/10.1016/j.jbiosc.2021.04.005>
- Cameron, D.E., Bashor, C.J., Collins, J.J., 2014. A brief history of synthetic biology. *Nat. Rev. Microbiol.* 12, 381–390. <https://doi.org/10.1038/nrmicro3239>
- Campbell, R.E., Tour, O., Palmer, A.E., Steinbach, P.A., Baird, G.S., Zacharias, D.A., Tsien, R.Y., 2002. A monomeric red fluorescent protein. *Proc. Natl. Acad. Sci. U. S. A.* 99, 7877–7882. <https://doi.org/10.1073/pnas.082243699>
- Carrasco, B., Ayora, S., Lurz, R., Alonso, J.C., 2005. *Bacillus subtilis* RecU Holliday-junction resolvase modulates RecA activities. *Nucleic Acids Res.* 33, 3942–3952. <https://doi.org/10.1093/nar/gki713>
- Caulier, S., Nannan, C., Gillis, A., Licciardi, F., Bragard, C., Mahillon, J., 2019. Overview of the antimicrobial compounds produced by members of the *Bacillus subtilis* group. *Front. Microbiol.* 10, 1–19. <https://doi.org/10.3389/fmicb.2019.00302>
- Cautereels, C., Smets, J., De Saeger, J., Cool, L., Zhu, Y., Zimmermann, A., Steensels, J., Gorkovskiy, A., Jacobs, T.B., Verstrepen, K.J., 2024. Orthogonal LoxP sites allow multiplexed site-specific recombination in prokaryotic and eukaryotic hosts. *Nat. Commun.* 15. <https://doi.org/10.1038/s41467-024-44996-8>
- Chandran, S., Noskov, V.N., Segall-Shapiro, T.H., Ma, L., Whiteis, C., Lartigue, C., Jores, J., Vashee, S., Chuang, R.Y., 2014. TREC-IN: Gene knock-in genetic tool for genomes cloned in yeast. *BMC Genomics* 15, 1–11. <https://doi.org/10.1186/1471-2164-15-1180>
- Chang, S., Cohen, S.N., 1979. High frequency transformation of *Bacillus subtilis* protoplasts by plasmid DNA. *MGG Mol. Gen. Genet.* 168, 111–115. <https://doi.org/10.1007/BF00267940>
- Citti, C., Blanchard, A., 2013. Mycoplasmas and their host: Emerging and re-emerging minimal pathogens. *Trends Microbiol.* 21, 196–203. <https://doi.org/10.1016/j.tim.2013.01.003>
- Cohn, 1872. Identity Of *Bacillus Subtilis*. *Beitrag Zur Biol. Der Pflanz* 1, 175.
- Conn, H.J., 1930. The Identity of *Bacillus Subtilis*. *J. Infect. Dis.* 46, 341–350.
- Consortium, T.U., 2007. The Universal Protein Resource (UniProt). *Nucleic Acids Res.* 35, 193–197. <https://doi.org/10.1093/nar/gkl929>
- Cormack, B.P., Bertram, G., Egerton, M., Gow, N.A.R., Falkow, S., Brown, A.J.P., 1997. Yeast-enhanced green fluorescent protein (yEGFP): A reporter of gene expression in *Candida albicans*. *Microbiology* 143, 303–311. <https://doi.org/10.1099/00221287-143-2-303>
- Cormack, B.P., Valdivia, R.H., Falkow, S., 1996. FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* 173, 33–38. [https://doi.org/10.1016/0378-1119\(95\)00685-0](https://doi.org/10.1016/0378-1119(95)00685-0)
- Courvalin, P., Carlier, C., 1987. Tn 1545: A conjugative shuttle transposon. *MGG Mol. Gen. Genet.* 206, 259–264. <https://doi.org/10.1007/BF00333582>
- Cui, W., Han, L., Suo, F., Liu, Z., Zhou, L., Zhou, Z., 2018. Exploitation of *Bacillus subtilis* as a robust



- workhorse for production of heterologous proteins and beyond. *World J. Microbiol. Biotechnol.* 34, 1–19. <https://doi.org/10.1007/s11274-018-2531-7>
- Dershowitz, A., Newlon, C.S., 1993. The Effect on Chromosome Stability of Deleting Replication Origins. *Mol. Cell. Biol.* 13, 391–398. <https://doi.org/10.1128/mcb.13.1.391-398.1993>
- Dervyn, E., Planson, A.-G., Tanaka, K., Chubukov, V., Guérin, C., Derozier, S., Lecointe, F., Sauer, U., Yoshida, K.-I., Nicolas, P., Noirot, P., Jules, M., 2023. Greedy reduction of *Bacillus subtilis* genome yields emergent phenotypes of high resistance to a DNA damaging agent and low evolvability. *Nucleic Acids Res.* 1–19. <https://doi.org/10.1093/nar/gkad145>
- Deshpande, A.M., Newlon, C.S., 1992. The ARS consensus sequence is required for chromosomal origin function in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 12, 4305–4313. <https://doi.org/10.1128/mcb.12.10.4305>
- Deveau, H., Garneau, J.E., Moineau, S., 2010. CRISPR/Cas system and its role in phage-bacteria interactions. *Annu. Rev. Microbiol.* 64, 475–493. <https://doi.org/10.1146/annurev.micro.112408.134123>
- DiCarlo, J.E., Norville, J.E., Mali, P., Rios, X., Aach, J., Church, G.M., 2013. Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res.* 41, 4336–4343.
- Donati, S., Barbier, I., García-Soriano, D.A., Grasso, S., Handal-Marquez, P., Malci, K., Marlow, L., Westmann, C., Amara, A., 2022. Synthetic biology in Europe: current community landscape and future perspectives. *Biotechnol. Notes* 3, 54–61. <https://doi.org/10.1016/j.biotno.2022.07.003>
- Dong, H., Zhang, D., 2014. Current development in genetic engineering strategies of *Bacillus* species. *Microb. Cell Fact.* 13, 1–11. <https://doi.org/10.1186/1475-2859-13-63>
- Dubnau, D., 1999. DNA UPTAKE IN BACTERIA. *Public Health.*
- Elenga-Wilson, P.S., Kayath, C.A., Mokemiabeka, N.S., Nzaou, S.A.E., Nguimbi, E., Ahombo, G., 2021. Profiling of Indigenous Biosurfactant-Producing *Bacillus* Isolates in the Bioremediation of Soil Contaminated by Petroleum Products and Olive Oil. *Int. J. Microbiol.* 2021. <https://doi.org/10.1155/2021/9565930>
- Engel, S.R., Dietrich, F.S., Fisk, D.G., Binkley, G., Balakrishnan, R., Costanzo, M.C., Dwight, S.S., Hitz, B.C., Karra, K., Nash, R.S., Weng, S., Wong, E.D., Lloyd, P., Skrzypek, M.S., Miyasato, S.R., Simison, M., Cherry, J.M., 2014. The Reference Genome Sequence of *Saccharomyces cerevisiae*: Then and Now. *G3 Genes, Genomes, Genet.* 4, 389–398. <https://doi.org/10.1534/g3.113.008995>
- Fabret, C., Ehrlich, S.D., Noirot, P., 2002. A new mutation delivery system for genome-scale approaches in *Bacillus subtilis*. *Mol. Microbiol.* 46, 25–36. <https://doi.org/10.1046/j.1365-2958.2002.03140.x>
- Fernández, S., Ayora, S., Alonso, J.C., 2000. *Bacillus subtilis* homologous recombination: Genes and products. *Res. Microbiol.* 151, 481–486. [https://doi.org/10.1016/S0923-2508\(00\)00165-0](https://doi.org/10.1016/S0923-2508(00)00165-0)
- Flagg, M.P., Kao, A., Hampton, R.Y., 2019. Integrating after CEN Excision (ICE) Plasmids: Combining the ease of yeast recombination cloning with the stability of genomic integration. *Yeast* 36, 593–605. <https://doi.org/10.1002/yea.3400>
- Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J.F., Dougherty, B.A., Merrick, J.M., 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* 269, 496–512. <https://doi.org/10.1126/science.7542800>



- Flórez, L.A., Roppel, S.F., Schmeisky, A.G., Lammers, C.R., Stülke, J., 2009. A community-curated consensual annotation that is continuously updated: The *Bacillus subtilis* centred wiki SubtiWiki. Database 2009, 1–9. <https://doi.org/10.1093/database/bap012>
- Franke, A.E., Clewell, D.B., 1981. Evidence for a chromosome-borne resistance transposon (Tn916) in *Streptococcus faecalis* that is capable of “conjugal” transfer in the absence of a conjugative plasmid. *J. Bacteriol.* 145, 494–502. <https://doi.org/10.1128/jb.145.1.494-502.1981>
- Fraser, C.M., Gocayne, J.D., White, O., Adams, M.D., Clayton, R.A., Fleischmann, R.D., Bult, C.J., Kerlavage, A.R., Sutton, G., Kelley, J.M., Fritchman, R.D., Weidman, J.F., Small, K. V., Sandusky, M., Fuhrmann, J., Nguyen, D., Utterback, T.R., Saudek, D.M., Phillips, C.A., Merrick, J.M., Tomb, J.F., Dougherty, B.A., Bott, K.F., Hu, P.C., Lucier, T.S., Peterson, S.N., Smith, H.O., Hutchison, C.A. 3rd, Venter, J.C., 1995. The minimal gene complement of *Mycoplasma genitalium*. *Science* 270, 397–403. <https://doi.org/10.1126/science.270.5235.397>
- Fredens, J., Wang, K., de la Torre, D., Funke, L.F.H., Robertson, W.E., Christova, Y., Chia, T., Schmied, W.H., Dunkelmann, D.L., Beránek, V., Uttamapinant, C., Llamazares, A.G., Elliott, T.S., Chin, J.W., 2019. Total synthesis of *Escherichia coli* with a recoded genome. *Nature* 569, 514–518. <https://doi.org/10.1038/s41586-019-1192-5>
- Galanie, S., Thodey, K., Trenchard, I.J., Interrante, M.F., Smolke, C.D., 2015. Complete biosynthesis of opioids in yeast. *Science* (80- ). 349, 1095–1100. <https://doi.org/10.1126/science.aac9373>
- Gamba, P., Jonker, M.J., Hamoen, L.W., 2015. A Novel Feedback Loop That Controls Bimodal Expression of Genetic Competence. *PLoS Genet.* 11, 1–32. <https://doi.org/10.1371/journal.pgen.1005047>
- Garcia-Morales, L., Ruiz, E., Gourgues, G., Rideau, F., Piñero-Lambea, C., Lluch-Senar, M., Blanchard, A., Lartigue, C., 2020. A RAGE Based Strategy for the Genome Engineering of the Human Respiratory Pathogen *Mycoplasma pneumoniae*. *ACS Synth. Biol.* 9, 2737–2748. <https://doi.org/10.1021/acssynbio.0c00263>
- Geissler, A.S., Anthon, C., Alkan, F., González-Tortuero, E., Poulsen, L.D., Kallehauge, T.B., Breüner, A., Seemann, S.E., Vinther, J., Gorodkin, J., 2021. Bsgatlas: A unified *Bacillus subtilis* genome and transcriptome annotation atlas with enhanced information access. *Microb. Genomics* 7. <https://doi.org/10.1099/mgen.0.000524>
- Gibbons, N.E., Murray, R.G.E., 1978. Proposals concerning the higher taxa of bacteria. *Int. J. Syst. Bacteriol.* 28, 1–6. <https://doi.org/10.1099/00207713-28-1-1>
- Gibson, D.G., 2011. Enzymatic assembly of overlapping DNA fragments. *Methods Enzymol.* 498, 349–361. <https://doi.org/10.1016/B978-0-12-385120-8.00015-2>
- Gibson, D.G., Benders, G.A., Andrews-Pfannkoch, C., Denisova, E.A., Baden-Tillson, H., Zaveri, J., Stockwell, T.B., Brownley, A., Thomas, D.W., Algire, M.A., Merryman, C., Young, L., Noskov, V.N., Glass, J.I., Venter, J.C., Hutchison, C.A., Smith, H.O., 2008a. Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science* (80- ). 319, 1215–1220. <https://doi.org/10.1126/science.1151721>
- Gibson, D.G., Benders, G.A., Axelrod, K.C., Zaveri, J., Algire, M.A., Moodie, M., Montague, M.G., Venter, J.C., Smith, H.O., Hutchison, C.A., 2008b. One-step assembly in yeast of 25 overlapping DNA fragments to form a complete synthetic *Mycoplasma genitalium* genome. *Proc. Natl. Acad. Sci. U. S. A.* 105, 20404–20409. <https://doi.org/10.1073/pnas.0811011106>
- Gibson, D.G., Glass, J.I., Lartigue, C., Noskov, V.N., Chuang, R.Y., Algire, M.A., Benders, G.A., Montague, M.G., Ma, L., Moodie, M.M., Merryman, C., Vashee, S., Krishnakumar, R., Assad-



- Garcia, N., Andrews-Pfannkoch, C., Denisova, E.A., Young, L., Qi, Z.N., Segall-Shapiro, T.H., Calvey, C.H., Parmar, P.P., Hutchison, C.A., Smith, H.O., Venter, J.C., 2010. Creation of a bacterial cell controlled by a chemically synthesized genome. *Science* (80-. ). 329, 52–56. <https://doi.org/10.1126/science.1190719>
- Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A., Smith, H.O., 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* 6, 343–345. <https://doi.org/10.1038/nmeth.1318>
- Gietz, R.D., Schiestl, R.H., Willems, A.R., Woods, R.A., 1995. Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. *Yeast* 11, 355–360. <https://doi.org/10.1002/yea.320110408>
- Gilbert Chu, Douglas Vollrath, R.W.D., 1985. *Science* Volume 234 issue 4783 1986 [doi 10.1126\_science.3538420] Chu, G; Vollrath, D; Davis, R. -- Separation of large DNA molecules by contour-clamped homogeneous electric fields 801.
- Gilbertson, L., 2003. Cre-lox recombination: Cre-ative tools for plant biotechnology. *Trends Biotechnol.* 21, 550–555. <https://doi.org/10.1016/j.tibtech.2003.09.011>
- Glass, J.I., Merryman, C., Wise, K.S., Iii, C.A.H., Smith, H.O., 2017. Minimal Cells — Real and Imagined 1–12.
- Goffeau, A., Barrell, B.G., Bussey, H., Davis, R.W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J.D., Jacq, C., Johnston, M., Louis, E.J., Mewes, H.W., Murakami, Y., Philippsen, P., Tettelin, H., Oliver, S.G., 1996. Life with 6000 genes. *Science* 274, 546,563-567. <https://doi.org/10.1126/science.274.5287.546>
- Goosens, V.J., Monteferrante, C.G., Van Dijl, J.M., 2014. The Tat system of Gram-positive bacteria. *Biochim. Biophys. Acta - Mol. Cell Res.* 1843, 1698–1706. <https://doi.org/10.1016/j.bbamcr.2013.10.008>
- Gourgues, G., Manso-Silvan, L., Chamberland, C., Sirand-Pugnet, P., Thiaucourt, F., Blanchard, A., Baby, V., Lartigue, C., 2024. A toolbox for manipulating the genome of the major goat pathogen, *Mycoplasma capricolum* subsp. *capripneumoniae*. *Microbiol. (United Kingdom)* 170, 1–17. <https://doi.org/10.1099/mic.0.001423>
- Green, E.R., Mecsas, J., 2016. Bacterial secretion systems: An overview. *Virulence Mech. Bact. Pathog.* 213–239. <https://doi.org/10.1128/9781555819286.ch8>
- Grindley, N.D.F., Whiteson, K.L., Rice, P.A., 2006. Mechanisms of site-specific recombination. *Annu. Rev. Biochem.* 75, 567–605. <https://doi.org/10.1146/annurev.biochem.73.011303.073908>
- Guesdon, G., Gourgues, G., Rideau, F., Ipoutcha, T., Manso-Silvan, L., Jules, M., Sirand-Pugnet, P., Blanchard, A., Lartigue, C., 2023. Combining Fusion of Cells with CRISPR-Cas9 Editing for the Cloning of Large DNA Fragments or Complete Bacterial Genomes in Yeast. *ACS Synth. Biol.* 12, 3252–3266. <https://doi.org/10.1021/acssynbio.3c00248>
- Hahn, J., Kong, L., Dubnau, D., 1994. The regulation of competence transcription factor synthesis constitutes a critical control point in the regulation of competence in *Bacillus subtilis*. *J. Bacteriol.* 176, 5753–5761. <https://doi.org/10.1128/jb.176.18.5753-5761.1994>
- Hamoen, L.W., Van Werkhoven, A.F., Venema, G., Dubnau, D., 2000. The pleiotropic response regulator DegU functions as a priming protein in competence development in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. U. S. A.* 97, 9246–9251. <https://doi.org/10.1073/pnas.160010597>
- Hamoen, L.W., Venema, G., Kuipers, O.P., 2003. Controlling competence in *Bacillus subtilis*: Shared



- use of regulators. *Microbiology* 149, 9–17. <https://doi.org/10.1099/mic.0.26003-0>
- Hoa, T.T., Tortosa, P., Albano, M., Dubnau, D., 2002. Rok (YkuW) regulates genetic competence in *Bacillus subtilis* by directly repressing comK. *Mol. Microbiol.* 43, 15–26. <https://doi.org/10.1046/j.1365-2958.2002.02727.x>
- Hochman, L., Segev, N., Sternberg, N., Cohen, G., 1983. Site-specific recombinational circularization of bacteriophage P1 DNA. *Virology* 131, 11–17. [https://doi.org/10.1016/0042-6822\(83\)90528-7](https://doi.org/10.1016/0042-6822(83)90528-7)
- Hong, H.A., Khaneja, R., Tam, N.M.K., Cazzato, A., Tan, S., Urdaci, M., Brisson, A., Gasbarrini, A., Barnes, I., Cutting, S.M., 2009. *Bacillus subtilis* isolated from the human gastrointestinal tract. *Res. Microbiol.* 160, 134–143. <https://doi.org/10.1016/j.resmic.2008.11.002>
- Huang, Y., Swarge, B.N., Roseboom, W., Bleeker, J.D., Brul, S., Setlow, P., Kramer, G., 2024. Integrative Metabolomics and Proteomics Allow the Global Intracellular Characterization of *Bacillus subtilis* Cells and Spores. *J. Proteome Res.* 23, 596–608. <https://doi.org/10.1021/acs.jproteome.3c00386>
- Hutchison, C.A., Chuang, R.Y., Noskov, V.N., Assad-Garcia, N., Deerinck, T.J., Ellisman, M.H., Gill, J., Kannan, K., Karas, B.J., Ma, L., Pelletier, J.F., Qi, Z.Q., Richter, R.A., Strychalski, E.A., Sun, L., Suzuki, Y., Tsvetanova, B., Wise, K.S., Smith, H.O., Glass, J.I., Merryman, C., Gibson, D.G., Venter, J.C., 2016. Design and synthesis of a minimal bacterial genome. *Science* (80-. ). 351. <https://doi.org/10.1126/science.aad6253>
- Hutchison, C.A., Peterson, S.N., Gill, S.R., Cline, R.T., White, O., Fraser, C.M., Smith, H.O., Venter, J.C., 1999. Global transposon mutagenesis and a minimal mycoplasma genome. *Science* (80-. ). 286, 2165–2169. <https://doi.org/10.1126/science.286.5447.2165>
- Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., Nakamura, A., 1987. Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isoenzyme conversion in *Escherichia coli*, and identification of the gene product. *J. Bacteriol.* 169, 5429–5433. <https://doi.org/10.1128/jb.169.12.5429-5433.1987>
- Itaya, M., 1999. Effective Cloning of Unmarked DNA Fragments in the *Bacillus subtilis* 168 Genome 1–17.
- Jacob, F., Monod, J., 1961. Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* 3, 318–356. [https://doi.org/10.1016/S0022-2836\(61\)80072-7](https://doi.org/10.1016/S0022-2836(61)80072-7)
- Jansen, R., Van Embden, J.D.A., Gaastra, W., Schouls, L.M., 2002. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.* 43, 1565–1575. <https://doi.org/10.1046/j.1365-2958.2002.02839.x>
- Jayaram, M., Ma, C.H., Kachroo, A.H., Rowley, P.A., Guga, P., Fan, H.F., Voziyanov, Y., 2015. An overview of tyrosine site-specific recombination: From an *flp* perspective. *Mob. DNA III* 43–71. <https://doi.org/10.1128/9781555819217.ch2>
- Ji, Y., Li, J., Liang, Y., Li, L., Wang, Y., Pi, L., Xing, P., Nomura, C.T., Chen, S., Zhu, C., Wang, Q., 2024. Engineering the Tat-secretion pathway of *Bacillus licheniformis* for the secretion of cytoplasmic enzyme arginase. *Appl. Microbiol. Biotechnol.* 108, 1–12. <https://doi.org/10.1007/s00253-023-12917-2>
- Ji, Y., Zhang, B., Van Horn, S.F., Warren, P., Woodnutt, G., Burnham, M.K.R., Rosenberg, M., 2001. Identification of critical staphylococcal genes using conditional phenotypes generated by antisense RNA. *Science* (80-. ). 293, 2266–2269. <https://doi.org/10.1126/science.1063566>
- Jiang, W., Bikard, D., Cox, D., Zhang, F., Marraffini, L.A., 2013. RNA-guided editing of bacterial



- genomes using CRISPR-Cas systems. *Nat. Biotechnol.* 31, 233–239. <https://doi.org/10.1038/nbt.2508>
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., Charpentier, E., 2012. A Programmable Dual-RNA – Guided 337, 816–821.
- Johnston, C., Martin, B., Fichant, G., Polard, P., Claverys, J.P., 2014. Bacterial transformation: Distribution, shared mechanisms and divergent control. *Nat. Rev. Microbiol.* 12, 181–196. <https://doi.org/10.1038/nrmicro3199>
- Kaishima, M., Ishii, J., Matsuno, T., Fukuda, N., Kondo, A., 2016. Expression of varied GFPs in *Saccharomyces cerevisiae*: Codon optimization yields stronger than expected expression and fluorescence intensity. *Sci. Rep.* 6, 1–15. <https://doi.org/10.1038/srep35932>
- Kakeshtia, H., Kageyama, Y., Ara, K., Ozaki, K., Nakamura, K., 2010. Enhanced extracellular production of heterologous proteins in *Bacillus subtilis* by deleting the c-terminal region of the SecA secretory machinery. *Mol. Biotechnol.* 46, 250–257. <https://doi.org/10.1007/s12033-010-9295-0>
- Kant, R., Blom, J., Palva, A., Siezen, R.J., de Vos, W.M., 2011. Comparative genomics of *Lactobacillus*. *Microb. Biotechnol.* 4, 323–332. <https://doi.org/10.1111/j.1751-7915.2010.00215.x>
- Karas, B.J., Jablanovic, J., Irvine, E., Sun, L., Ma, L., Weyman, P.D., Gibson, D.G., Glass, J.I., Venter, J.C., Hutchison, C.A., Smith, H.O., Suzuki, Y., 2014. Transferring whole genomes from bacteria to yeast spheroplasts using entire bacterial cells to reduce DNA shearing. *Nat. Protoc.* 9, 743–750.
- Karas, B.J., Jablanovic, J., Sun, L., Ma, L., Goldgof, G.M., Stam, J., Ramon, A., Manary, M.J., Winzeler, E.A., Venter, J.C., Weyman, P.D., Gibson, D.G., Glass, J.I., Hutchison, C.A., Smith, H.O., Suzuki, Y., 2013. Direct transfer of whole genomes from bacteria to yeast. *Nat. Methods* 10, 410–412.
- Karas, B.J., Moreau, N.G., Deerinck, T.J., Gibson, D.G., Venter, J.C., Smith, H.O., Glass, J.I., 2019. Direct Transfer of a *Mycoplasma mycoides* Genome to Yeast Is Enhanced by Removal of the *Mycoides* Glycerol Uptake Factor Gene *glpF*. *ACS Synth. Biol.* 8, 239–244. <https://doi.org/10.1021/acssynbio.8b00449>
- Karas, B.J., Tagwerker, C., Yonemoto, I.T., Hutchison, C.A., Smith, H.O., 2012. Cloning the *Acholeplasma laidlawii* pg-8a genome in *Saccharomyces cerevisiae* as a yeast centromeric plasmid. *ACS Synth. Biol.* 1, 22–28. <https://doi.org/10.1021/sb200013>
- Karr, J.R., Sanghvi, J.C., MacKlin, D.N., Gutschow, M. V., Jacobs, J.M., Bolival, B., Assad-Garcia, N., Glass, J.I., Covert, M.W., 2012. A whole-cell computational model predicts phenotype from genotype. *Cell* 150, 389–401. <https://doi.org/10.1016/j.cell.2012.05.044>
- Ketner, G., Spencer, F., Tugendreich, S., Connelly, C., Hieter, P., 1994. Efficient manipulation of the human adenovirus genome as an infectious yeast artificial chromosome clone. *Proc. Natl. Acad. Sci. U. S. A.* 91, 6186–6190. <https://doi.org/10.1073/pnas.91.13.6186>
- Kidane, D., Carrasco, B., Manfredi, C., Rothmaier, K., Ayora, S., Tadesse, S., Alonso, J.C., Graumann, P.L., 2009. Evidence for different pathways during horizontal gene transfer in competent *Bacillus subtilis* cells. *PLoS Genet.* 5. <https://doi.org/10.1371/journal.pgen.1000630>
- Kim, J., Salvador, M., Saunders, E., González, J., Avignone-Rossa, C., Jiménez, J.I., 2016. Properties of alternative microbial hosts used in synthetic biology: Towards the design of a modular chassis. *Essays Biochem.* 60, 303–313. <https://doi.org/10.1042/EBC20160015>
- Kitano, S., Lin, C., Foo, J.L., Chang, M.W., 2023. Synthetic biology: Learning the way toward high-precision biological design. *PLoS Biol.* 21, 2–5. <https://doi.org/10.1371/journal.pbio.3002116>



- Kobayashi, K., Ehrlich, S.D., Albertini, A., Amati, G., Andersen, K.K., Arnaud, M., Asai, K., Ashikaga, S., Aymerich, S., Bessieres, P., Boland, F., Brignell, S.C., Bron, S., Bunai, K., Chapuis, J., Christiansen, L.C., Danchin, A., Débarbouillé, M., Dervyn, E., Deuerling, E., Devine, K., Devine, S.K., Dreesen, O., Errington, J., Fillinger, S., Foster, S.J., Fujita, Y., Galizzi, A., Gardan, R., Eschevins, C., Fukushima, T., Haga, K., Harwood, C.R., Hecker, M., Hosoya, D., Hullo, M.F., Kakeshita, H., Karamata, D., Kasahara, Y., Kawamura, F., Koga, K., Koski, P., Kuwana, R., Imamura, D., Ishimaru, M., Ishikawa, S., Ishio, I., le Coq, D., Masson, A., Mauël, C., Meima, R., Mellado, R.P., Moir, A., Moriya, S., Nagakawa, E., Nanamiya, H., Nakai, S., Nygaard, P., Ogura, M., Ohanan, T., O'Reilly, M., O'Rourke, M., Pragai, Z., Pooley, H.M., Rapoport, G., Rawlins, J.P., Rivas, L.A., Rivolta, C., Sadaie, A., Sadaie, Y., Sarvas, M., Sato, T., Saxild, H.H., Scanlan, E., Schumann, W., Seegers, J.F.M.L., Sekiguchi, J., Sekowska, A., Séror, S.J., Simon, M., Stragier, P., Studer, R., Takamatsu, H., Tanaka, T., Takeuchi, M., Thomaidis, H.B., Vagner, V., van Dijl, J.M., Watabe, K., Wipat, A., Yamamoto, H., Yamamoto, M., Yamamoto, Y., Yamane, K., Yata, K., Yoshida, K., Yoshikawa, H., Zuber, U., Ogasawara, N., 2003. Essential *Bacillus subtilis* genes. *Proc. Natl. Acad. Sci. U. S. A.* 100, 4678–4683. <https://doi.org/10.1073/pnas.0730515100>
- Komatsu, M., Uchiyama, T., Omura, S., Cane, D.E., Ikeda, H., 2010. Genome-minimized *Streptomyces* host for the heterologous expression of secondary metabolism. *Proc. Natl. Acad. Sci. U. S. A.* 107, 2646–2651. <https://doi.org/10.1073/pnas.0914833107>
- Konkol, M.A., Blair, K.M., Kearns, D.B., 2013. Plasmid-encoded comI inhibits competence in the ancestral 3610 strain of *Bacillus subtilis*. *J. Bacteriol.* 195, 4085–4093. <https://doi.org/10.1128/JB.00696-13>
- Koo, B.M., Kritikos, G., Farelli, J.D., Todor, H., Tong, K., Kimsey, H., Wapinski, I., Galardini, M., Cabal, A., Peters, J.M., Hachmann, A.B., Rudner, D.Z., Allen, K.N., Typas, A., Gross, C.A., 2017. Construction and Analysis of Two Genome-Scale Deletion Libraries for *Bacillus subtilis*. *Cell Syst.* 4, 291-305.e7. <https://doi.org/10.1016/j.cels.2016.12.013>
- Kouprina, N., Graves, J., Cancilla, M.R., Resnick, M.A., Larionov, V., 1997. Specific isolation of human rDNA genes by TAR cloning. *Gene* 197, 269–276. [https://doi.org/10.1016/S0378-1119\(97\)00271-0](https://doi.org/10.1016/S0378-1119(97)00271-0)
- Kouprina, N., Kim, J.H., Larionov, V., 2021. Highly Selective, CRISPR/Cas9-Mediated Isolation of Genes and Genomic Loci from Complex Genomes by TAR Cloning in Yeast. *Curr. Protoc.* 1, 1–27. <https://doi.org/10.1002/cpz1.207>
- Kouprina, N., Larionov, V., 2023. Transformation-associated recombination (TAR) cloning and its applications for gene function; genome architecture and evolution; biotechnology and biomedicine. *Oncotarget* 14, 1009–1033. <https://doi.org/10.18632/oncotarget.28546>
- Kouprina, N., Larionov, V., 2008. Selective isolation of genomic loci from complex genomes by transformation-associated recombination cloning in the yeast *Saccharomyces cerevisiae*. *Nat. Protoc.* 3, 371–377. <https://doi.org/10.1038/nprot.2008.5>
- Kouprina, N., Noskov, V.N., Larionov, V., 2020. Selective isolation of large segments from individual microbial genomes and environmental DNA samples using transformation-associated recombination cloning in yeast. *Nat. Protoc.* 15, 734–749. <https://doi.org/10.1038/s41596-019-0280-1>
- Kresge, N., Simoni, R.D., Hill, R.L., 2009. Total Synthesis of a Tyrosine Suppressor tRNA: the Work of H. Gobind Khorana. *J. Biol. Chem.* 284, e5–e7. [https://doi.org/10.1016/s0021-9258\(19\)82144-0](https://doi.org/10.1016/s0021-9258(19)82144-0)
- Kubo, Y., Rooney, A.P., Tsukakoshi, Y., Nakagawa, R., Hasegawa, H., Kimura, K., 2011. Phylogenetic analysis of *Bacillus subtilis* strains applicable to natto (fermented soybean) production. *Appl. Environ. Microbiol.* 77, 6463–6469. <https://doi.org/10.1128/AEM.00448-11>



- Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G., Azevedo, V., Bertero, M.G., Bessières, P., Bolotin, A., Borchert, S., Borriss, R., Boursier, L., Brans, A., Braun, M., Brignell, S.C., Bron, S., Brouillet, S., Bruschi, C. V., Caldwell, B., Capuano, V., Carter, N.M., Choi, S.-K., Codani, J.-J., Connerton, I.F., Cummings, N.J., Daniel, R.A., Denizot, F., Devine, K.M., Dusterhöft, A., Ehrlich, S.D., Emmerson, P.T., Entian, K.D., Errington, J., Fabret, C., Ferrari, E., Foulger, D., Fritz, C., Fujita, M., Fujita, Y., Fuma, S., Galizzi, A., Galleron, N., Ghim, S.-Y., Glaser, P., Goffeau, A., Golightly, E.J., Grandi, G., Guiseppi, G., Guy, B.J., Haga, K., Haiech, J., Harwood, C.R., Hénaut, A., Hilbert, H., Holsappel, S., Hosono, S., Hullo, M.-F., Itaya, M., Jones, L., Joris, B., Karamata, D., Kasahara, Y., Klaerr-Blanchard, M., Klein, C., Kobayashi, Y., Koetter, P., Koningstein, G., Krogh, S., Kumano, M., Kurita, K., Lapidus, A., Lardinois, S., Lauber, J., Lazarevic, V., Lee, S.-M., Levine, A., Liu, H., Masuda, S., Mauël, C., Médigue, C., Medina, N., Mellado, R.P., Mizuno, M., Moestl, D., Nakai, S., Noback, M., Noone, D., O'Reilly, M., Ogawa, K., Ogiwara, A., Oudega, B., Park, S.-H., Parro, V., Pohl, T.M., Portetelle, D., Porwollik, S., Prescott, A.M., Presecan, E., Pujic, P., Purnelle, B., Rapoport, G., Rey, M., Reynolds, S., Rieger, M., Rivolta, C., Rocha, E., Roche, B., Rose, M., Sadaie, Y., Sato, T., Scanlan, E., Schleich, S., Schroeter, R., Scoffone, F., Sekiguchi, J., Sekowska, A., Seror, S.J., Serror, P., Shin, B.-S., Soldo, B., Sorokin, A., Tacconi, E., Takagi, T., Takahashi, H., Takemaru, K., Takeuchi, M., Tamakoshi, A., Tanaka, T., Terpstra, P., Tognoni, A., Tosato, V., Uchiyama, S., Vandenbol, M., Vannier, F., Vassarotti, A., Viari, A., Wambutt, R., Wedler, E., Wedler, H., Weitzenegger, T., Winters, P., Wipat, A., Yamamoto, H., Yamane, K., Yasumoto, K., Yata, K., Yoshida, K., Yoshikawa, H.-F., Zumstein, E., Yoshikawa, H., Danchin, A., 1997. The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*. *Nature* 390, 249–256.
- Kusaoka, H., Hayashi, Y., Kadowaki, Y., Kimoto, H., 1989. Optimum Conditions for Electric Pulse-Mediated Gene Transfer to *Bacillus Subtilis* Cells. *Agric. Biol. Chem.* 53, 2441–2446. <https://doi.org/10.1080/00021369.1989.10869689>
- Kuspa, A., Vollrath, D., Cheng, Y., Kaiser, D., 1989. Physical mapping of the *Myxococcus xanthus* genome by random cloning in yeast artificial chromosomes. *Proc. Natl. Acad. Sci. U. S. A.* 86, 8917–8921. <https://doi.org/10.1073/pnas.86.22.8917>
- Labroussaa, F., Baby, V., Rodrigue, S., Lartigue, C., 2019. La transplantation de génomes. *Médecine/Sciences* 35, 761–770. <https://doi.org/10.1051/medsci/2019154>
- Labroussaa, F., Lebaudy, A., Baby, V., Gourgues, G., Matteau, D., Vashee, S., Sirand-Pugnet, P., Rodrigue, S., Lartigue, C., 2016. Impact of donor-recipient phylogenetic distance on bacterial genome transplantation. *Nucleic Acids Res.* 44, 8501–8511. <https://doi.org/10.1093/nar/gkw688>
- Langer, S.J., Ghafoori, A.P., Byrd, M., Leinwand, L., 2002. A genetic screen identifies novel non-compatible loxP sites. *Nucleic Acids Res.* 30.
- Larionov, V., Kouprina, N., Graves, J., Chen, X.N., Korenberg, J.R., Resnick, M.A., 1996. Specific cloning of human DNA as yeast artificial chromosomes by transformation-associated recombination. *Proc. Natl. Acad. Sci. U. S. A.* 93, 491–496. <https://doi.org/10.1073/pnas.93.1.491>
- Larionov, V., Kouprina, N., Nikolaishvili, N., Resnick, M.A., 1994. Recombination during transformation as a source of chimeric mammalian artificial chromosomes in yeast (YACs). *Nucleic Acids Res.* 22, 4154–4162. <https://doi.org/10.1093/nar/22.20.4154>
- Larionov, V., Kouprina, N., Solomon, G., Barrett, J.C., Resnick, M.A., 1997. Direct isolation of human *brca2* gene by transformation-associated recombination in yeast. *Proc. Natl. Acad. Sci. U. S. A.* 94, 7384–7387. <https://doi.org/10.1073/pnas.94.14.7384>
- Lartigue, C., Glass, J.I., Alperovich, N., Pieper, R., Parmar, P.P., Hutchison, C.A., Smith, H.O., Venter,



- J.C., 2007. Genome transplantation in bacteria: Changing one species to another. *Science* (80- ). 317, 632–638. <https://doi.org/10.1126/science.1144622>
- Lartigue, C., Vashee, S., Algire, M.A., Chuang, R.Y., Benders, G.A., Ma, L., Noskov, V.N., Denisova, E.A., Gibson, D.G., Assad-Garcia, N., Alperovich, N., Thomas, D.W., Merryman, C., Hutchison, C.A., Smith, H.O., Venter, J.C., Glass, J.I., 2009. Creating bacterial strains from genomes that have been cloned and engineered in yeast. *Science* (80- ). 325, 1693–1696. <https://doi.org/10.1126/science.1173759>
- Le Breton, Y., Mohapatra, N.P., Haldenwang, W.G., 2006. In vivo random mutagenesis of *Bacillus subtilis* by use of TnYLB-1, a mariner-based transposon. *Appl. Environ. Microbiol.* 72, 327–333. <https://doi.org/10.1128/AEM.72.1.327-333.2006>
- Lechat, P., Hummel, L., Rousseau, S., Moszer, I., 2008. GenoList: An integrated environment for comparative analysis of microbial genomes. *Nucleic Acids Res.* 36, 469–474. <https://doi.org/10.1093/nar/gkm1042>
- Leduc, S., 1912. *La biologie synthétique, Études de biophysique; II TA - TT -*. A. Poinat, Paris SE - 1 vol. (217 p.) : ill. ; 23 cm. <https://doi.org/LK> - <https://worldcat.org/title/492450889>
- Lee, N.K., Kim, W.S., Paik, H.D., 2019. *Bacillus* strains as human probiotics: characterization, safety, microbiome, and probiotic carrier. *Food Sci. Biotechnol.* 28, 1297–1305. <https://doi.org/10.1007/s10068-019-00691-9>
- Lenhart, J.S., Schroeder, J.W., Walsh, B.W., Simmons, L.A., 2012. DNA Repair and Genome Maintenance in *Bacillus subtilis*. *Microbiol. Mol. Biol. Rev.* 76, 530–564. <https://doi.org/10.1128/mubr.05020-11>
- Lévi-Meyrueis, C., Fodor, K., Schaeffer, P., 1980. Polyethyleneglycol-induced transformation of *Bacillus subtilis* protoplasts by bacterial chromosomal DNA. *MGG Mol. Gen. Genet.* 179, 589–594. <https://doi.org/10.1007/BF00271749>
- Li, N., Gou, S., Wang, J., Zhang, Q., Huang, X., Xie, J., Li, L., Jin, Q., Ouyang, Z., Chen, F., Ge, W., Shi, H., Liang, Y., Zhuang, Z., Zhao, X., Lian, M., Ye, Y., Quan, L., Wu, H., Lai, L., Wang, K., 2021. CRISPR/Cas9-Mediated Gene Correction in Newborn Rabbits with Hereditary Tyrosinemia Type I. *Mol. Ther.* 29, 1001–1015. <https://doi.org/10.1016/j.ymthe.2020.11.023>
- Li, X., Heyer, W.D., 2008. Homologous recombination in DNA repair and DNA damage tolerance. *Cell Res.* 18, 99–113. <https://doi.org/10.1038/cr.2008.1>
- Liachko, I., Youngblood, R.A., Keich, U., Dunham, M.J., 2013. High-resolution mapping, characterization, and Optimization of Autonomously Replicating Sequences in Yeast 698–704. <https://doi.org/10.1101/gr.144659.112.Freely>
- Ling Lin Fu, Zi Rong Xu, Wei Fen Li, Jiang Bing Shuai, Ping Lu, Chun Xia Hu, 2007. Protein secretion pathways in *Bacillus subtilis*: Implication for optimization of heterologous protein secretion. *Biotechnol. Adv.* 25, 1–12. <https://doi.org/10.1016/j.biotechadv.2006.08.002>
- Liu, D., Huang, C., Guo, J., Zhang, P., Chen, T., Wang, Z., Zhao, X., 2019. Development and characterization of a CRISPR/Cas9n-based multiplex genome editing system for *Bacillus subtilis*. *Biotechnol. Biofuels* 12, 1–17. <https://doi.org/10.1186/s13068-019-1537-1>
- Liu, Y., Liu, L., Li, J., Du, G., Chen, J., 2019. Synthetic Biology Toolbox and Chassis Development in *Bacillus subtilis*. *Trends Biotechnol.* 37, 548–562. <https://doi.org/10.1016/j.tibtech.2018.10.005>
- Lluch-Senar, M., Delgado, J., Chen, W., Lloréns-Rico, V., O'Reilly, F.J., Wodke, J.A., Unal, E.B., Yus, E., Martínez, S., Nichols, R.J., Ferrar, T., Vivancos, A., Schmeisky, A., Stülke, J., van Noort, V., Gavin,



- A., Bork, P., Serrano, L., 2015. Defining a minimal cell: essentiality of small ORFs and ncRNAs in a genome-reduced bacterium. *Mol. Syst. Biol.* 11, 1–7. <https://doi.org/10.15252/msb.20145558>
- Lu, Y.P., Zhang, C., Lv, F.X., Bie, X.M., Lu, Z.X., 2012. Study on the electro-transformation conditions of improving transformation efficiency for *Bacillus subtilis*. *Lett. Appl. Microbiol.* 55, 9–14. <https://doi.org/10.1111/j.1472-765X.2012.03249.x>
- Maier, B., 2020. Competence and transformation in *Bacillus subtilis*. *Curr. Issues Mol. Biol.* 37, 57–76. <https://doi.org/10.21775/CIMB.037.057>
- Makarova, K.S., Wolf, Y.I., Iranzo, J., Shmakov, S.A., Alkhnbashi, O.S., Brouns, S.J.J., Charpentier, E., Cheng, D., Haft, D.H., Horvath, P., Moineau, S., Mojica, F.J.M., Scott, D., Shah, S.A., Siksnyts, V., Terns, M.P., Venclovas, Č., White, M.F., Yakunin, A.F., Yan, W., Zhang, F., Garrett, R.A., Backofen, R., van der Oost, J., Barrangou, R., Koonin, E. V., 2020. Evolutionary classification of CRISPR–Cas systems: a burst of class 2 and derived variants. *Nat. Rev. Microbiol.* 18, 67–83. <https://doi.org/10.1038/s41579-019-0299-x>
- Mandic-Mulec, I., Stefanic, P., van Elsas, J.D., 2016. Ecology of bacillaceae. *Bact. Spore From Mol. to Syst.* 59–85. <https://doi.org/10.1128/9781555819323.ch3>
- Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.J., Chen, Z., Dewell, S.B., Du, L., Fierro, J.M., Gomes, X. V., Godwin, B.C., He, W., Helgesen, S., Ho, C.H., Irzyk, G.P., Jando, S.C., Alenquer, M.L.I., Jarvie, T.P., Jirage, K.B., Kim, J.B., Knight, J.R., Lanza, J.R., Leamon, J.H., Lefkowitz, S.M., Lei, M., Li, J., Lohman, K.L., Lu, H., Makhijani, V.B., McDade, K.E., McKenna, M.P., Myers, E.W., Nickerson, E., Nobile, J.R., Plant, R., Puc, B.P., Ronan, M.T., Roth, G.T., Sarkis, G.J., Simons, J.F., Simpson, J.W., Srinivasan, M., Tartaro, K.R., Tomasz, A., Vogt, K.A., Volkmer, G.A., Wang, S.H., Wang, Y., Weiner, M.P., Yu, P., Begley, R.F., Rothberg, J.M., 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437, 376–380. <https://doi.org/10.1038/nature03959>
- Marschall, P., Malik, N., Larin, Z., 1999. Transfer of YACs up to 2.3 Mb intact into human cells with polyethylenimine. *Gene Ther.* 6, 1634–1637. <https://doi.org/10.1038/sj.gt.3300975>
- Martin V, Pitera D, Withers S, Newman J, Keasling J, 2003. Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. *Nat. Biotechnol.* 21, 796–802.
- Mathiasen, D.P., Lisby, M., 2014. Cell cycle regulation of homologous recombination in *Saccharomyces cerevisiae*. *FEMS Microbiol. Rev.* 38, 172–184. <https://doi.org/10.1111/1574-6976.12066>
- Matz, M. V, Fradkov, A.F., Labas, Y.A., Savitsky, A.P., Zaraisky, A.G., Markelov, M.L., Lukyanov, S.A., 1999. Fluorescent proteins from nonbioluminescent Anthozoa species. *Nat. Biotechnol.* 17, 969–973. <https://doi.org/10.1038/13657>
- McAdams, H.H., Arkin, A., 2000. Gene regulation: Towards a circuit engineering discipline. *Curr. Biol.* 10, 318–320. [https://doi.org/10.1016/S0960-9822\(00\)00440-1](https://doi.org/10.1016/S0960-9822(00)00440-1)
- Meinke, G., Bohm, A., Hauber, J., Pisabarro, M.T., Buchholz, F., 2016. Cre Recombinase and Other Tyrosine Recombinases. *Chem. Rev.* 116, 12785–12820. <https://doi.org/10.1021/acs.chemrev.6b00077>
- Moger-Reischer, R.Z., Glass, J.I., Wise, K.S., Sun, L., Bittencourt, D.M.C., Lehmkuhl, B.K., Schoolmaster, D.R., Lynch, M., Lennon, J.T., 2023. Evolution of a minimal cell. *Nature* 620, 122–127. <https://doi.org/10.1038/s41586-023-06288-x>
- Morange, M., 2000. A history of molecular biology. Harvard University Press.



- Morimoto, T., Kadoya, R., Endo, K., Tohata, M., Sawada, K., Liu, S., Ozawa, T., Kodama, T., Kakeshita, H., Kageyama, Y., Manabe, K., Kanaya, S., Ara, K., Ozaki, K., Ogasawara, N., 2008. Enhanced recombinant protein productivity by genome reduction in *Bacillus subtilis*. *DNA Res.* 15, 83–91. <https://doi.org/10.1093/dnares/dsm032>
- Morowitz, H.J., 1984. The completeness of molecular biology. *Isr. J. Med. Sci.* 20, 750–753.
- Mullis, K., 1987. Process for amplifying nucleic acid sequences. United States Pat.
- Nakano, M.M., Zuber, P., 1998. ANAEROBIC GROWTH OF A “STRICT AEROBE” (*BACILLUS SUBTILIS*). *Annu. Rev. Microbiol.* 52, 165–1990.
- Natalay Kouprina and Vladimir Larionov, 2016. TAR cloning for genomic studies and synthetic biology. *Chromosoma* 125, 621–632. <https://doi.org/10.1007/s00412-016-0588-3>. Transformation-associated
- NESTER, E.W., STOCKER, B.A., 1963. Biosynthetic Latency in Early Stages of Deoxyribonucleic Acid transformation in *Bacillus subtilis*. *J. Bacteriol.* 86, 785–796.
- Nicolas, P., Mäder, U., Dervyn, E., Rochat, T., Leduc, A., Pigeonneau, N., Bidnenko, E., Marchadier, E., Hoebeke, M., Aymerich, S., Becher, D., Bisicchia, P., Botella, E., Delumeau, O., Doherty, G., Denham, E.L., Fogg, M.J., Fromion, V., Goelzer, A., Hansen, A., Härtig, E., Harwood, C.R., Homuth, G., Jarmer, H., 2012. Condition-Dependent Transcriptome Reveals High-Level Regulatory Architecture in *Bacillus subtilis*. *Science* (80-. ). 335, 1103–1106.
- Noskov, V.N., Karas, B.J., Young, L., Chuang, R.Y., Gibson, D.G., Lin, Y.C., Stam, J., Yonemoto, I.T., Suzuki, Y., Andrews-Pfannkoch, C., Glass, J.I., Smith, H.O., Hutchison, C.A., Venter, J.C., Weyman, P.D., 2012. Assembly of large, high G+C bacterial DNA fragments in yeast. *ACS Synth. Biol.* 1, 267–273. <https://doi.org/10.1021/sb3000194>
- Noskov, V.N., Koriabine, M., Solomon, G., Randolph, M., Barrett, J.C., Leem, S.H., Stubbs, L., Kouprina, N., Larionov, V., 2001. Defining the minimal length of sequence homology required for selective gene isolation by TAR cloning. *Nucleic Acids Res.* 29, 32. <https://doi.org/10.1093/nar/29.6.e32>
- Noskov, V.N., Kouprina, N., Leem, S.H., Ouspenski, I., Barrett, C., Larionov, V., 2003. A general cloning system to selectively isolate any eukaryotic or prokaryotic genomic region in yeast. *BMC Genomics* 4, 1–11. <https://doi.org/10.1186/1471-2164-4-16>
- Noskov, V.N., Ma, L., Chen, S., Chuang, R.Y., 2015. Recombinase-mediated cassette exchange (RMCE) system for functional genomics studies in *Mycoplasma mycoides*. *Biol. Proced. Online* 17, 1–8. <https://doi.org/10.1186/s12575-015-0016-8>
- Noskov, V.N., Segall-Shapiro, T.H., Chuang, R.Y., 2010. Tandem repeat coupled with endonuclease cleavage (TREC): A seamless modification tool for genome engineering in yeast. *Nucleic Acids Res.* 38, 2570–2576. <https://doi.org/10.1093/nar/gkq099>
- Nottelet, P., Bataille, L., Gourgues, G., Anger, R., Lartigue, C., Sirand-Pugnet, P., Marza, E., Fronzes, R., Arfi, Y., 2021. The mycoplasma surface proteins MIB and MIP promote the dissociation of the antibody-antigen interaction. *Sci. Adv.* 7, 1–15. <https://doi.org/10.1126/sciadv.abf2403>
- Ohama, T., Inagaki, Y., Bessho, Y., Osawa, S., 2008. Evolving genetic code. *Proc. Japan Acad. Ser. B Phys. Biol. Sci.* 84, 58–74. <https://doi.org/10.2183/pjab.84.58>
- Ohta, K., Ichihashi, N., 2019. Liposome fragment-mediated introduction of multiple plasmids into *Bacillus subtilis*. *Biochem. Biophys. Reports* 18, 100646. <https://doi.org/10.1016/j.bbrep.2019.100646>



- Olmos, J., Acosta, M., Mendoza, G., Pitones, V., 2020. *Bacillus subtilis*, an ideal probiotic bacterium to shrimp and fish aquaculture that increase feed digestibility, prevent microbial diseases, and avoid water pollution. *Arch. Microbiol.* 202, 427–435. <https://doi.org/10.1007/s00203-019-01757-2>
- Oren, A., Garrity, G.M., 2021. Valid publication of the names of forty-two phyla of prokaryotes. *Int. J. Syst. Evol. Microbiol.* 71. <https://doi.org/10.1099/ijsem.0.005056>
- Paddon, C.J., Westfall, P.J., Pitera, D.J., Benjamin, K., Fisher, K., McPhee, D., Leavell, M.D., Tai, A., Main, A., Eng, D., Polichuk, D.R., Teoh, K.H., Reed, D.W., Treynor, T., Lenihan, J., Jiang, H., Fleck, M., Bajad, S., Dang, G., Dengrove, D., Diola, D., Dorin, G., Ellens, K.W., Fickes, S., Galazzo, J., Gaucher, S.P., Geistlinger, T., Henry, R., Hepp, M., Horning, T., Iqbal, T., Kizer, L., Lieu, B., Melis, D., Moss, N., Regentin, R., Secrest, S., Tsuruta, H., Vazquez, R., Westblade, L.F., Xu, L., Yu, M., Zhang, Y., Zhao, L., Lievense, J., Covello, P.S., Keasling, J.D., Reiling, K.K., Renninger, N.S., Newman, J.D., 2013. High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* 496, 528–532. <https://doi.org/10.1038/nature12051>
- Pargellis, C.A., Nunes-Düby, S.E., de Vargas, L.M., Landy, A., 1988. Suicide recombination substrates yield covalent lambda integrase-DNA complexes and lead to identification of the active site tyrosine. *J. Biol. Chem.* 263, 7678–7685. [https://doi.org/10.1016/s0021-9258\(18\)68552-7](https://doi.org/10.1016/s0021-9258(18)68552-7)
- Petit, M.A., Bruand, C., Janniere, L., Ehrlich, S.D., 1990. Tn10-derived transposons active in *Bacillus subtilis*. *J. Bacteriol.* 172, 6736–6740. <https://doi.org/10.1128/jb.172.12.6736-6740.1990>
- Potapov, V., Ong, J.L., 2017. Examining sources of error in PCR by single-molecule sequencing. *PLoS One* 12, 1–19. <https://doi.org/10.1371/journal.pone.0169774>
- Pozsgai, E.R., Blair, K.M., Kearns, D.B., 2012. Modified mariner transposons for random inducible-expression insertions and transcriptional reporter fusion insertions in *Bacillus subtilis*. *Appl. Environ. Microbiol.* 78, 778–785. <https://doi.org/10.1128/AEM.07098-11>
- Provedj, R., Dubnau, D., 1999. ComEA is a DNA receptor for transformation of competent *Bacillus subtilis*. *Mol. Microbiol.* 31, 271–280. <https://doi.org/10.1046/j.1365-2958.1999.01170.x>
- Rahmer, R., Heravi, K.M., Altenbuchner, J., 2015. Construction of a super-competent *Bacillus subtilis* 168 using the PmtIA-comKS inducible cassette. *Front. Microbiol.* 6, 1–11. <https://doi.org/10.3389/fmicb.2015.01431>
- Ravikumar, V., Nalpas, N.C., Anselm, V., Krug, K., Lenuzzi, M., Šestak, M.S., Domazet-Lošo, T., Mijakovic, I., Macek, B., 2018. In-depth analysis of *Bacillus subtilis* proteome identifies new ORFs and traces the evolutionary history of modified proteins. *Sci. Rep.* 8, 1–11. <https://doi.org/10.1038/s41598-018-35589-9>
- Reuß, D.R., Altenbuchner, J., Mäder, U., Rath, H., Ischebeck, T., Sappa, P.K., Thürmer, A., Guérin, C., Nicolas, P., Steil, L., Zhu, B., Feussner, I., Klumpp, S., Daniel, R., Commichau, F.M., Völker, U., Stülke, J., 2017. Large-scale reduction of the *Bacillus subtilis* genome: consequences for the transcriptional network, resource allocation, and metabolism. *Genome Res.* 27, 289–299. <https://doi.org/10.1101/gr.215293.116>
- Rideau, F., Le Roy, C., Descamps, E.C.T., Renaudin, H., Lartigue, C., Bébéar, C., 2017. Cloning, Stability, and Modification of *Mycoplasma hominis* Genome in Yeast. *ACS Synth. Biol.* 6, 891–901. <https://doi.org/10.1021/acssynbio.6b00379>
- Roebroek, A.J.M., Gordts, P.L.S.M., Reekmans, S., 2011. Generation of a series of knock-in alleles using RMCE in ES cells. *Methods Mol. Biol.* 693, 277–281. [https://doi.org/10.1007/978-1-60761-974-1\\_16](https://doi.org/10.1007/978-1-60761-974-1_16)



- Rossignol, J.L., 1990. La Recombinaison Homologue: Mecanismes Et Consequences. *Medecine/Sciences* 6. <https://doi.org/10.4267/10608/4227>
- Rothstein, R., 1983. Is Digested With a Restriction Enzyme That Cleaves Within the. *Methods* 101, 202–211.
- Ruiz, E., Talenton, V., Dubrana, M.P., Guesdon, G., Lluch-Senar, M., Salin, F., Sirand-Pugnet, P., Arfi, Y., Lartigue, C., 2019. CReasPy-Cloning: A Method for Simultaneous Cloning and Engineering of Megabase-Sized Genomes in Yeast Using the CRISPR-Cas9 System. *ACS Synth. Biol.* 8, 2547–2557. <https://doi.org/10.1021/acssynbio.9b00224>
- Santos, C.N.S., Regitsky, D.D., Yoshikuni, Y., 2013. Implementation of stable and complex biological systems through recombinase-assisted genome engineering. *Nat. Commun.* 4, 1–10. <https://doi.org/10.1038/ncomms3503>
- Santos, C.N.S., Yoshikuni, Y., 2014. Engineering complex biological systems in bacteria through recombinase-assisted genome engineering. *Nat. Protoc.* 9, 1320–1336. <https://doi.org/10.1038/nprot.2014.084>
- Sapranauskas, R., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., Siksnys, V., 2011. The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res.* 39, 9275–9282. <https://doi.org/10.1093/nar/gkr606>
- Sarvas, M., Harwood, C.R., Bron, S., Van Dijl, J.M., 2004. Post-translocational folding of secretory proteins in Gram-positive bacteria. *Biochim. Biophys. Acta - Mol. Cell Res.* 1694, 311–327. <https://doi.org/10.1016/j.bbamcr.2004.04.009>
- Sauer, B., 1987. Functional Expression of the cre-lox Site-Specific Recombination System in the Yeast *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 7, 2087–2096. <https://doi.org/10.1128/mcb.7.6.2087-2096.1987>
- Sauer, B., Henderson, N., 1988. Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. *Proc. Natl. Acad. Sci. U. S. A.* 85, 5166–5170. <https://doi.org/10.1073/pnas.85.14.5166>
- Schaeffer, P., Hotchkiss, R.D., 1978. Fusion of Bacterial Protoplasts. *Methods Cell Biol.* 20, 149–158. [https://doi.org/10.1016/S0091-679X\(08\)62017-8](https://doi.org/10.1016/S0091-679X(08)62017-8)
- Schallmeyer, M., Singh, A., Ward, O.P., 2004. Developments in the use of *Bacillus* species for industrial production. *Can. J. Microbiol.* 50, 1–17. <https://doi.org/10.1139/w03-076>
- Schetelig, M.F., Yan, Y., Zhao, Y., Handler, A.M., 2019. Genomic targeting by recombinase-mediated cassette exchange in the spotted wing drosophila, *Drosophila suzukii*. *Insect Mol. Biol.* 28, 187–195. <https://doi.org/10.1111/imb.12537>
- Serrano, E., Carrasco, B., Gilmore, J.L., Takeyasu, K., Alonso, J.C., 2018. RecA regulation by RecU and DprA During *Bacillus subtilis* natural plasmid transformation. *Front. Microbiol.* 9, 1–20. <https://doi.org/10.3389/fmicb.2018.01514>
- Serrano, E., Ramos, C., Alonso, J.C., Ayora, S., 2021. Recombination proteins differently control the acquisition of homeologous DNA during *Bacillus subtilis* natural chromosomal transformation. *Environ. Microbiol.* 23, 512–524. <https://doi.org/10.1111/1462-2920.15342>
- Serrano, E., Ramos, C., Ayora, S., Alonso, J.C., 2020. Viral SPP1 DNA is infectious in naturally competent *Bacillus subtilis* cells: inter- and intramolecular recombination pathways. *Environ. Microbiol.* 22, 714–725. <https://doi.org/10.1111/1462-2920.14908>
- Serror, P., Sonenshein, A.L., 1996. CodY is required for nutritional repression of *Bacillus subtilis*



- genetic competence. *J. Bacteriol.* 178, 5910–5915. <https://doi.org/10.1128/jb.178.20.5910-5915.1996>
- Shaw, D., Serrano, L., Lluch-Senar, M., 2021. *Lox'd* in translation: Contradictions in the nomenclature surrounding common *lox*-site mutants and their implications in experiments. *Microbiol. (United Kingdom)* 167, 1–11. <https://doi.org/10.1099/mic.0.000997>
- Shen, B., Zhang, J., Wu, H., Wang, J., Ma, K., Li, Z., Zhang, X., Zhang, P., Huang, X., 2013. Generation of gene-modified mice via Cas9/RNA-mediated gene targeting. *Cell Res.* 23, 720–723. <https://doi.org/10.1038/cr.2013.46>
- Sherman, F., Azubalis, D.A., Fjellstedt, T.J., Mehvert, D., Ogur, M., Sherman, F., Lacroute, F., Lacroute, F., Fink, G.R., Rose, M., Grisafi, P., Botstein, D., Fink, G.R., Lawrence, C., Harbor, C.S., 1987. 5 - Fluoroorotic Acid as a Selective Agent in Yeast Molecular Genetics 154, 164–175.
- SHIMOMURA, O., JOHNSON, F.H., SAIGA, Y., 1962. Extraction, purification and properties of aequorin, a bioluminescent. *J. Cell. Comp. Physiol.* 59, 223–239. <https://doi.org/10.1002/jcp.1030590302>
- Smith, M.C.M., Thorpe, H.M., 2002. Diversity in the serine recombinases. *Mol. Microbiol.* 44, 299–307. <https://doi.org/10.1046/j.1365-2958.2002.02891.x>
- Song, Y., He, S., Jopkiewicz, A., Setroikromo, R., van Merkerk, R., Quax, W.J., 2022. Development and application of CRISPR-based genetic tools in *Bacillus* species and *Bacillus* phages. *J. Appl. Microbiol.* 133, 2280–2298. <https://doi.org/10.1111/jam.15704>
- Soupene, E., Van Heeswijk, W.C., Plumbridge, J., Stewart, V., Bertenthal, D., Lee, H., Prasad, G., Paliy, O., Charernnoppakul, P., Kustu, S., 2003. Physiological studies of *Escherichia coli* strain MG1655: Growth defects and apparent cross-regulation of gene expression. *J. Bacteriol.* 185, 5611–5626. <https://doi.org/10.1128/JB.185.18.5611-5626.2003>
- Spizizen, J., 1958. Transformation of Biochemically Deficient Strains of *Bacillus Subtilis* By Deoxyribonucleate. *Proc. Natl. Acad. Sci.* 44, 1072–1078. <https://doi.org/10.1073/pnas.44.10.1072>
- Stark, W.M., 2015. The serine recombinases. *Mob. DNA* III 73–89. <https://doi.org/10.1128/9781555819217.ch3>
- Sternberg, N., Hamilton, D., 1981. Bacteriophage P1 site-specific recombination. I. Recombination between *loxP* sites. *J. Mol. Biol.* 150, 467–486. [https://doi.org/10.1016/0022-2836\(81\)90375-2](https://doi.org/10.1016/0022-2836(81)90375-2)
- Su, Y., Liu, C., Fang, H., Zhang, D., 2020. *Bacillus subtilis*: A universal cell factory for industry, agriculture, biomaterials and medicine. *Microb. Cell Fact.* 19, 1–12. <https://doi.org/10.1186/s12934-020-01436-8>
- Sumi, C.D., Yang, B.W., Yeo, I.C., Hahm, Y.T., 2015. Antimicrobial peptides of the genus *Bacillus*: A new era for antibiotics. *Can. J. Microbiol.* 61, 93–103. <https://doi.org/10.1139/cjm-2014-0613>
- Sutyak, K.E., Wirawan, R.E., Aroutcheva, A.A., Chikindas, M.L., 2008. Isolation of the *Bacillus subtilis* antimicrobial peptide subtilisin from the dairy product-derived *Bacillus amyloliquefaciens*. *J. Appl. Microbiol.* 104, 1067–1074. <https://doi.org/10.1111/j.1365-2672.2007.03626.x>
- Tagwerker, C., Dupont, C.L., Karas, B.J., Ma, L., Chuang, R.Y., Benders, G.A., Ramon, A., Novotny, M., Montague, M.G., Venepally, P., Bрами, D., Schwartz, A., Andrews-Pfannkoch, C., Gibson, D.G., Glass, J.I., Smith, H.O., Venter, J.C., Hutchison, C.A., 2012. Sequence analysis of a complete 1.66 Mb *Prochlorococcus marinus* MED4 genome cloned in yeast. *Nucleic Acids Res.* 40, 10375–10383. <https://doi.org/10.1093/nar/gks823>



- Talenton, V., Baby, V., Gourgues, G., Mouden, C., Claverol, S., Vashee, S., Blanchard, A., Labroussaa, F., Jores, J., Arfi, Y., Sirand-Pugnet, P., Lartigue, C., 2022. Genome Engineering of the Fast-Growing *Mycoplasma feriruminatoris* toward a Live Vaccine Chassis. *ACS Synth. Biol.* 11, 1919–1930. <https://doi.org/10.1021/acssynbio.2c00062>
- Tanaka, K., Henry, C.S., Zinner, J.F., Jolivet, E., Cohoon, M.P., Xia, F., Bidnenko, V., Ehrlich, S.D., Stevens, R.L., Noirot, P., 2013. Building the repertoire of dispensable chromosome regions in *Bacillus subtilis* entails major refinement of cognate large-scale metabolic model. *Nucleic Acids Res.* 41, 687–699. <https://doi.org/10.1093/nar/gks963>
- Thanh Bien, T.L., Tsuji, S., Tanaka, K., Takenaka, S., Yoshida, K.I., 2014. Secretion of heterologous thermostable cellulases in *Bacillus subtilis*. *J. Gen. Appl. Microbiol.* 60, 175–182. <https://doi.org/10.2323/jgam.60.175>
- Tindall, B.J., 2019. When treated as heterotypic synonyms the names *Caryophanaceae* Peshkoff 1939 (Approved lists 1980) and *Caryophanales* Peshkoff 1939 (approved lists 1980) have priority over the names *Planococcaceae* Krasil'nikov 1949 (approved lists 1980) and *bacillales* prev. *Int. J. Syst. Evol. Microbiol.* 69, 2187–2195. <https://doi.org/10.1099/ijsem.0.003354>
- Tjalsma, H., Haike Antelmann, Jan D H Jongbloed, Peter G Braun, Elise Darmon, Ronald Dorenbos, Jean-Yves F Dubois, Helga Westers, Geeske Zanen, Wim J Quax, Oscar P Kuipers, Sierd Bron, Michael Hecker, J.M.V.D., Tjalsma, H., Antelmann, H., Jongbloed, J.D.H., Braun, P.G., Darmon, E., Dorenbos, R., Dubois, J.-Y.F., Westers, H., Zanen, G., Quax, W.J., Kuipers, O.P., Bron, S., Hecker, M., Van Dijk, J.M., 2004. Proteomics of protein secretion by *Bacillus subtilis*: separating the “secrets” of the secretome. *Microbiol. Mol. Biol. Rev.* 68, 207–233. <https://doi.org/10.1128/MMBR.68.2.207>
- Torres, R., Serrano, E., Alonso, J.C., 2019. *Bacillus subtilis* RecA interacts with and loads RadA/Sms to unwind recombination intermediates during natural chromosomal transformation. *Nucleic Acids Res.* 47, 9198–9215. <https://doi.org/10.1093/nar/gkz647>
- Tsarnpopoulos, I., Gourgues, G., Blanchard, A., Vashee, S., Jores, J., Lartigue, C., Sirand-Pugnet, P., 2016. In-yeast engineering of a bacterial genome using CRISPR/Cas9. *ACS Synth. Biol.* 5, 104–9.
- Tully, J.G., Rose, D.L., Yunker, C.E., Carle, P., Bove, J.M., Williamson, D.L., Whitcomb, R.F., 1995. *Spiroplasma ixodetis* sp. nov., a new species from *Ixodes pacificus* ticks collected in Oregon. *Int. J. Syst. Bacteriol.* 45, 23–28. <https://doi.org/10.1099/00207713-45-1-23>
- Turan, S., Galla, M., Ernst, E., Qiao, J., Voelkel, C., Schiedlmeier, B., Zehe, C., Bode, J., 2011. Recombinase-mediated cassette exchange (RMCE): Traditional concepts and current challenges. *J. Mol. Biol.* 407, 193–221. <https://doi.org/10.1016/j.jmb.2011.01.004>
- Turan, S., Zehe, C., Kuehle, J., Qiao, J., Bode, J., 2013. Recombinase-mediated cassette exchange (RMCE) - A rapidly-expanding toolbox for targeted genomic modifications. *Gene* 515, 1–27. <https://doi.org/10.1016/j.gene.2012.11.016>
- Turgay, K., Hahn, J., Burghoorn, J., Dubnau, D., 1998. Competence in *Bacillus subtilis* is controlled by regulated proteolysis of a transcription factor. *EMBO J.* 17, 6730–6738. <https://doi.org/10.1093/emboj/17.22.6730>
- Turgay, K., Hamoen, L.W., Venema, G., Dubnau, D., 1997. Biochemical characterization of a molecular switch involving the heat shock protein ClpC, which controls the activity of ComK, the competence transcription factor of *Bacillus subtilis*. *Genes Dev.* 11, 119–128. <https://doi.org/10.1101/gad.11.1.119>
- Vallenet, D., Labarre, L., Rouy, Z., Barbe, V., Bocs, S., Cruveiller, S., Lajus, A., Pascal, G., Scarpelli, C.,



- Médigue, C., 2006. MaGe: A microbial genome annotation system supported by synteny results. *Nucleic Acids Res.* 34, 53–65. <https://doi.org/10.1093/nar/gkj406>
- Van Dijl, J.M., Braun, P.G., Robinson, C., Quax, W.J., Antelmann, H., Hecker, M., Müller, J., Tjalsma, H., Bron, S., Jongbloed, J.D.H., 2002. Functional genomic analysis of the *Bacillus subtilis* Tat pathway for protein secretion. *J. Biotechnol.* 98, 243–254. [https://doi.org/10.1016/S0168-1656\(02\)00135-9](https://doi.org/10.1016/S0168-1656(02)00135-9)
- van Duyne, G.D., 2015. Cre recombinase. *Mob. DNA* III 119–138. <https://doi.org/10.1128/9781555819217.ch5>
- Van Houdt, R., Leplae, R., Lima-Mendez, G., Mergeay, M., Toussaint, A., 2012. Towards a more accurate annotation of tyrosine-based site-specific recombinases in bacterial genomes. *Mob. DNA* 3, 1–11. <https://doi.org/10.1186/1759-8753-3-6>
- van Houten, J.V., Newlon, C.S., 1990. Mutational Analysis of the Consensus Sequence of a Replication Origin from Yeast Chromosome III. *Mol. Cell. Biol.* 10, 3917–3925. <https://doi.org/10.1128/mcb.10.8.3917-3925.1990>
- Van Sinderen, D., Luttinger, A., Kong, U., Dubnau, D., Venema, G., Hamoen, L., 1995. comK encodes the competence transcription factor, the key regulatory protein for competence development in *Bacillus subtilis*. *Molecular Microbiol.* 15, 455–462.
- Van Sinderen, D., Venema, G., 1994. comK Acts as an autoregulatory control switch in the signal transduction route to competence in *Bacillus subtilis*. *J. Bacteriol.* 176, 5762–5770. <https://doi.org/10.1128/jb.176.18.5762-5770.1994>
- Vashee, S., Arfi, Y., Lartigue, C., 2020. Budding yeast as a factory to engineer partial and complete microbial genomes. *Curr. Opin. Syst. Biol.* 24, 1–8. <https://doi.org/10.1016/j.coisb.2020.09.003>
- Venetz, J.E., Medico, L. Del, Wölfle, A., Schächle, P., Bucher, Y., Appert, D., Tschan, F., Flores-Tinoco, C.E., Van Kooten, M., Guennoun, R., Deutsch, S., Christen, M., Christen, B., 2019. Chemical synthesis rewriting of a bacterial genome to achieve design flexibility and biological functionality. *Proc. Natl. Acad. Sci. U. S. A.* 116, 8070–8079. <https://doi.org/10.1073/pnas.1818259116>
- Vlamakis, H., Chai, Y., Beaugard, P., Losick, R., Kolter, R., 2013. Sticking together: Building a biofilm the *Bacillus subtilis* way. *Nat. Rev. Microbiol.* 11, 157–168. <https://doi.org/10.1038/nrmicro2960>
- Vos, P., Garrity, G., Jones, D., Krieg, N.R., Ludwig, W., Rainey, F.A., Schleifer, K.H., Whitman, W.B., 2009. *Bergey's Manual of Systematic Bacteriology: Volume 3: The Firmicutes*, *Bergey's Manual of Systematic Bacteriology*. Springer New York.
- Wang, G., Zhao, Z., Ke, J., Engel, Y., Shi, Y.-M., Robinson, D., Bingol, K., Zhang, Z., Bowen, B., Louie, K., Wang, B., Evans, R., Miyamoto, Y., Cheng, K., Kosina, S., De Raad, M., Silva, L., Luhrs, A., Lubbe, A., Hoyt, D.W., Francavilla, C., Otani, H., Deutsch, S., Washton, N.M., Rubin, E.M., Mouncey, N.J., Visel, A., Northen, T., Cheng, J.-F., Bode, H.B., Yoshikuni, Y., 2019. CRAGE enables rapid activation of biosynthetic gene clusters in undomesticated bacteria. *Nat. Microbiol.* 4, 2498–2510. <https://doi.org/10.1038/s41564-019-0573-8>
- Wang, J.P., Yeh, C.M., Tsai, Y.C., 2006. Improved subtilisin YaB production in *Bacillus subtilis* using engineered synthetic expression control sequences. *J. Agric. Food Chem.* 54, 9405–9410. <https://doi.org/10.1021/jf061982f>
- Wang, K., Fredens, J., Brunner, S.F., Kim, S.H., Chia, T., Chin, J.W., 2016. Defining synonymous codon compression schemes by genome recoding. *Nature* 539, 59–64.



<https://doi.org/10.1038/nature20124>

- Watzlawick, H., Altenbuchner, J., 2019. Multiple integration of the gene *ganA* into the *Bacillus subtilis* chromosome for enhanced  $\beta$ -galactosidase production using the CRISPR/Cas9 system. *AMB Express* 9. <https://doi.org/10.1186/s13568-019-0884-4>
- Westbrook, A.W., Moo-Young, M., Chou, C.P., 2016. Development of a CRISPR-Cas9 tool kit for comprehensive engineering of *Bacillus subtilis*. *Appl. Environ. Microbiol.* 82, 4876–4895. <https://doi.org/10.1128/AEM.01159-16>
- Wozniak, K.J., Simmons, L.A., 2022. Genome Editing Methods for *Bacillus subtilis*. *Methods Mol. Biol.* 2479, 159–174. [https://doi.org/10.1007/978-1-0716-2233-9\\_11](https://doi.org/10.1007/978-1-0716-2233-9_11)
- Wright, B.Y.G.G., Mandell, G.L., 1986. ANTHRAX TOXIN BLOCKS PRIMING OF NEUTROPHILS BY LIPOPOLYSACCHARIDE AND BY MURAMYL DIPEPTIDE 164.
- Wu, Y., Liu, Y., Lv, X., Li, J., Du, G., Liu, L., 2020. CAMERS-B: CRISPR/Cpf1 assisted multiple-genes editing and regulation system for *Bacillus subtilis*. *Biotechnol. Bioeng.* 117, 1817–1825. <https://doi.org/10.1002/bit.27322>
- Xue, G.P., Johnson, J.S., Dalrymple, B.P., 1999. High osmolarity improves the electro-transformation efficiency of the gram-positive bacteria *Bacillus subtilis* and *Bacillus licheniformis*. *J. Microbiol. Methods* 34, 183–191. [https://doi.org/10.1016/S0167-7012\(98\)00087-6](https://doi.org/10.1016/S0167-7012(98)00087-6)
- Yan, X., Yu, H.J., Hong, Q., Li, S.P., 2008. Cre/lox system and PCR-based genome engineering in *Bacillus subtilis*. *Appl. Environ. Microbiol.* 74, 5556–5562. <https://doi.org/10.1128/AEM.01156-08>
- Youngman, P.J., Perkins, J.B., Losick, R., 1983. Genetic transposition and insertional mutagenesis in *Bacillus subtilis* with *Streptococcus faecalis* transposon Tn917. *Proc. Natl. Acad. Sci. U. S. A.* 80, 2305–2309. <https://doi.org/10.1073/pnas.80.8.2305>
- Zhang, J., Hansen, L.G., Gudich, O., Viehrig, K., Lassen, L.M.M., Schrübbers, L., Adhikari, K.B., Rubaszka, P., Carrasquer-Alvarez, E., Chen, L., D'Ambrosio, V., Lehka, B., Haidar, A.K., Nallapareddy, S., Giannakou, K., Laloux, M., Arsovska, D., Jørgensen, M.A.K., Chan, L.J.G., Kristensen, M., Christensen, H.B., Sudarsan, S., Stander, E.A., Baidoo, E., Petzold, C.J., Wulff, T., O'Connor, S.E., Courdavault, V., Jensen, M.K., Keasling, J.D., 2022. A microbial supply chain for production of the anti-cancer drug vinblastine. *Nature* 609, 341–347. <https://doi.org/10.1038/s41586-022-05157-3>
- Zhang, N., Wang, D., Liu, Y., Li, S., Shen, Q., Zhang, R., 2014. Effects of different plant root exudates and their organic acid components on chemotaxis, biofilm formation and colonization by beneficial rhizosphere-associated bacterial strains. *Plant Soil* 374, 689–700. <https://doi.org/10.1007/s11104-013-1915-6>
- Zhang, X.Z., Yan, X., Cui, Z.L., Hong, Q., Li, S.P., 2006. *mazF*, a novel counter-selectable marker for unmarked chromosomal manipulation in *Bacillus subtilis*. *Nucleic Acids Res.* 34. <https://doi.org/10.1093/nar/gkl358>
- Zhou, J., Wu, R., Xue, X., Qin, Z., 2016. CasHRA (Cas9-facilitated Homologous Recombination Assembly) method of constructing megabase-sized DNA. *Nucleic Acids Res.* 44. <https://doi.org/10.1093/nar/gkw475>
- Zhu, B., Stülke, J., 2018. SubtiWiki in 2018: From genes and proteins to functional network annotation of the model organism *Bacillus subtilis*. *Nucleic Acids Res.* 46, D743–D748. <https://doi.org/10.1093/nar/gkx908>



## **Clonage, ingénierie et transfert de grands fragments de génome chez *Bacillus subtilis*.**

L'ingénierie des génomes des micro-organismes est devenue un standard dans les biotechnologies microbiennes. En 2010, des technologies prometteuses de biologie de synthèse utilisant la levure comme plateforme pour l'assemblage et l'ingénierie de génomes synthétiques bactériens suivi de leur transplantation dans une cellule receveuse ont vu le jour. Ces technologies ont conduit à la création des premières cellules synthétiques et ouvert de nouvelles voies vers la construction de cellules aux propriétés biologiques entièrement contrôlées. Le transfert de ces outils à des micro-organismes d'intérêt industriel comme la bactérie Gram+ *Bacillus subtilis* (*Bsu*), modèle dans le secteur des biotechnologies, constituerait une avancée majeure. C'est précisément l'objectif du projet ANR « Bacillus 2.0 », qui réunit deux équipes INRAE et qui se propose d'adapter l'ensemble de ces outils de biologie de synthèse à *Bsu* afin d'être en mesure de partir d'une conception, assistée par ordinateur, de génomes semi-synthétiques de *Bsu* jusqu'à l'obtention de nouvelles souches industrielles. Cependant, les premiers travaux réalisés sur ce projet ont montré que le génome entier de *Bsu* ne pouvait pas être cloné et maintenu en l'état dans la levure. Ces résultats risquaient de remettre en question la faisabilité du projet dans son ensemble et en particulier la pertinence d'utiliser la levure comme plateforme d'assemblage du génome semi-synthétique de *Bsu*.

L'objectif de ma thèse a consisté à démontrer que la levure restait un hôte pertinent pour le projet « Bacillus 2.0 ». Elle s'est déclinée en 3 parties. Dans la première partie, une méthode de clonage de génome récemment développée au laboratoire et dénommée CREasPy-Fusion, a progressivement été adaptée à *Bsu*. Les résultats obtenus ont montré (i) le transfert possible d'ADN plasmidique entre protoplastes bactériens et sphéroplastes de levure, (ii) l'efficacité d'un système CRISPR-Cas9 porté par les cellules de levure pour capturer/modifier cet ADN plasmidique pendant la fusion *Bsu*/levure, puis (iii) l'efficacité de ce même système pour capturer des fragments de génome d'une centaine de kb à partir de trois souches différentes. Des observations en microscopie à fluorescence ont également été réalisées et ont mis en évidence deux types d'interactions qui permettraient de passer d'un contact protoplastes/sphéroplastes à un ADN bactérien cloné dans la levure. Dans la seconde partie de ma thèse, la méthode CREasPy-Fusion a été mise à profit pour tenter de cloner de grands fragments du génome de *Bsu* dans la levure. Des fragments génomiques jusqu'à ~1 Mb ont pu être clonés dans la levure, mais leur capture a nécessité l'ajout préalable d'un grand nombre d'ARS sur le génome de *Bsu* pour stabiliser les constructions génétiques. La dernière partie a été l'adaptation de la méthode RAGE à *Bsu*. Cette méthode permet le transfert, non pas d'un génome entier mais de portions de génomes bactériens depuis la levure vers la bactérie à éditer. Une preuve de concept a été réalisée avec l'échange d'un premier fragment de génome de 155 kb par une version réduite de 44 kb.

En conclusion, les travaux réalisés au cours de cette thèse ont montré la pertinence d'utiliser la levure comme plateforme d'ingénierie dans les modifications à grande échelle du génome de *Bsu*. D'une part, nous avons montré que des fragments d'une centaine de kb peuvent être clonés dans la levure, modifiés et transférés dans une cellule receveuse de façon à générer des *Bsu* mutants. Cette stratégie offre une véritable alternative à la transplantation de génome. D'autre part, nous avons montré que de grands fragments du génome de *Bsu* (jusqu'à 1Mb) peuvent également être clonés dans la levure à condition de contenir de nombreux ARS dans leurs séquences. Grâce à ces résultats, le clonage d'un génome réduit de *Bsu* chez la levure est redevenu un objectif réalisable.

**Mots clés :** *Bacillus subtilis*, Clonage de génome dans la levure, Ingénierie de génome, méthode RAGE, Biologie de Synthèse, ARS

**UMR 1332 Biologie du Fruit et Pathologie – Equipe Mollicutes**

Centre INRAE Nouvelle-Aquitaine Bordeaux, Bâtiment IBVM 71 Avenue Edouard Bourloux, CS20032  
33882 Villenave d'Ornon Cedex, France

