



HAL
open science

Développement de méthodes de clonage de génomes entiers chez la levure pour la construction de souches châssis semi-synthétiques de *Bacillus subtilis*.

Gabrielle Guesdon

► **To cite this version:**

Gabrielle Guesdon. Développement de méthodes de clonage de génomes entiers chez la levure pour la construction de souches châssis semi-synthétiques de *Bacillus subtilis*. Microbiologie et Parasitologie. Université de Bordeaux, 2022. Français. NNT : 2022BORD0204 . tel-04485907

HAL Id: tel-04485907

<https://theses.hal.science/tel-04485907v1>

Submitted on 1 Mar 2024

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

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

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

Par Gabrielle GUESDON

**Développement de méthodes de clonage de génomes entiers chez
la levure pour la construction de souches châssis semi-synthétiques
de *Bacillus subtilis*.**

Sous la direction du Pr. Alain BLANCHARD et du Dr. Carole LARTIGUE

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

Soutenue le 29 juin 2022

Membres du jury :

| | |
|--|--------------------|
| M. JULES Matthieu, Professeur à INRAE, AgroParisTech et Université Paris-Saclay | Président |
| M. RODRIGUE Sébastien, Professeur titulaire à l'Université de Sherbrooke (QC Canada) | Rapporteur |
| Mme DABOUSSI Fayza, Directrice de recherche à INRAE-INSA de Toulouse | Rapporteuse |
| M. LABROUSSAA Fabien, Docteur à l'Université de Bern (Suisse) | Examineur |
| M. BLANCHARD Alain, Professeur à l'Université de Bordeaux | Directeur de thèse |
| Mme LARTIGUE Carole, Chargée de Recherche à INRAE de Bordeaux | Co-encadrante |

RÉSUMÉS

Développement de méthodes de clonage de génomes entiers chez la levure pour la construction de souches châssis semi-synthétiques de *Bacillus subtilis*.

Un des défis de la biologie de synthèse (BS), est d'apporter des solutions nouvelles à des enjeux globaux majeurs (thérapeutiques/sanitaires, climatiques), en particulier via la construction de souches de production utiles, performantes et respectueuses de l'environnement.

Bacillus subtilis (*Bsu*) est une bactérie Gram+ non pathogène utilisée en biotechnologie comme plateforme de production de molécules d'intérêt. Or, des études récentes ont établi que des souches de *Bsu* génétiquement modifiées permettaient d'obtenir une plus grande production de protéines recombinantes. Cela suggère que la production de châssis *Bsu* réduits pourrait être une étape importante dans l'amélioration des souches à visée industrielle.

Le projet ANR *Bacillus 2.0* dans lequel s'inscrit cette thèse, vise à transposer à *Bsu* des outils récents du domaine de la BS, et à mettre en place un pipeline efficace pour construire à haut débit des châssis modulables selon l'application biotechnologique. Ce pipeline rassemble les technologies suivantes : (i) design d'un génome synthétique (ii) assemblage de fragments d'ADN chevauchants au sein de la levure *Saccharomyces cerevisiae* (iii) isolement du génome et transplantation dans une cellule receveuse bactérienne et (iv) sélection et caractérisation des souches mutantes.

Les objectifs de cette thèse étaient d'attester la faisabilité de ces méthodes, en tentant de cloner et maintenir le génome réduit de *Bsu* MPG192 (2,86 Mb) dans la levure, puis de le modifier avec les outils d'ingénierie disponibles chez *S. cerevisiae*. Dans un premier temps des stratégies déjà décrites dans la littérature ont été déployées afin de cloner le génome entier de *Bsu*, mais sont restées infructueuses. En nous basant sur une approche de TAR-Cloning, nous avons tenté de cloner plusieurs fragments obtenus par restriction du génome de *Bsu*. Initialement, seuls cinq fragments sur sept ont été clonés. L'incapacité à cloner le plus grand de ces fragments (1,50 Mb) est vraisemblablement due à un manque d'ARS et/ou à sa taille. Concernant le second fragment, un ensemble de facteurs peuvent expliquer notre échec : à nouveau le manque d'ARS mais aussi, la présence de nombreux éléments répétés (7 opérons ribosomiques) ou l'expression délétère de ces gènes. Finalement, d'autres expérimentations ont permis de découper le génome en sous-fragments de tailles variables (6kb à 515kb) et ainsi de cloner en 21 morceaux la totalité du génome de *Bsu* MGP 192. Le TAR-Cloning imposant des contraintes dans le choix des sites de restriction, une nouvelle méthode de clonage, appelée CREasPy-Fusion, a été développée. Elle combine le système d'édition CRISPR-Cas9 et la fusion entre des sphéroplastes de levure et des cellules bactériennes. Comme preuve de concept, nous avons d'abord travaillé avec six espèces de mycoplasmes, et démontré qu'il est possible de cloner et modifier des génomes entiers. Cette approche a ensuite été transposée à *Bsu*, validant pour la première fois la fusion entre des sphéroplastes de levure et des protoplastes de bactéries Gram+. Elle a permis la capture précise d'un fragment d'environ 150 kb.

Bien que le génome entier de *Bsu* n'ait pas été cloné dans la levure à ce jour, plusieurs éléments critiques ont pu être identifiés. Tout d'abord, ces travaux soulignent l'importance de la méthode de clonage à adopter en fonction de l'organisme avec lequel on travaille. Ensuite, ils mettent en exergue l'existence de facteurs biologiques et techniques qui expliquent les difficultés actuelles et qui devront être pris en compte dans la suite des expérimentations. Enfin, ils ont permis le développement d'une nouvelle technique de clonage appelée CREasPy-Fusion, qui vient étoffer le catalogue des méthodes déjà décrites. Par sa versatilité, elle ouvre des perspectives pour la capture de grands fragments de génome, pour l'élimination de *loci* problématiques, ou encore, en appui à l'assemblage de fragments synthétiques.

Mots clés : *Bacillus subtilis*, Clonage de génome, Ingénierie génomique, Levure, Châssis bactérien

UMR 1332 Biologie du Fruit et Pathologie – Equipe Mollicutes

Centre INRAE Bordeaux-Aquitaine Bâtiment IBVM

71 Avenue Edouard Bourlaux, CS 2003233882 Villenave d'Ornon Cedex, FRANCE

Development of in-yeast genome cloning methods for the construction of semi-synthetic *Bacillus subtilis*-derived chassis strains

One of the major challenges in the synthetic biology (SB) field, is to provide new solutions to global issues (therapeutic/sanitary or climatic), in particular through the construction of useful, efficient and environmentally friendly production strains.

The well-characterized, non-pathogenic, Gram+ bacterium *Bacillus subtilis* (*Bsu*), is widely used in industry as a biotechnological workhorse. Recent studies have established that mutant strains with modified genomes are able to produce larger amounts of recombinant proteins. This suggests that the production of rationally designed *Bsu* chassis could be an important step in the improvement of valuable strains for industrial purposes.

This work was performed within the *Bacillus* 2.0's ANR project, which aims at applying SB tools for *Bsu*, and at developing an effective pipeline for the high-throughput construction of versatile *Bsu* chassis strains. Selected SB technologies for the pipeline include (i) the synthetic genome design, (ii) the in-yeast DNA assembly methods using *Saccharomyces cerevisiae*, (iii) the from-yeast whole genome isolation and transplantation (GT) to a recipient bacteria cell and, (iv) the characterization of recombinant strains.

The objectives of this thesis were to ensure the feasibility of these methods using a Gram+ bacterium, by showing, in particular, that it was possible to clone and maintain in *S. cerevisiae* the genome of a minimal *Bsu* strain, MPG192 (2.86 Mb) and to modify it using the large repertoire of yeast genetic tools. Our first attempts to clone the entire *Bsu* genome into yeast using already described methods failed. Using a TAR-Cloning approach, we then attempted to clone large DNA fragments obtained by restriction of the *Bsu* genome. In a first experiment, five out of seven fragments were cloned. Difficulties to clone the largest fragment (1.50 Mb), are presumably related to its size, and/or the lack of ARS elements. Concerning the other fragment, several factors have been proposed to explain the cloning failure: again, an insufficient number of ARS elements, but also, the presence of many repeated sequences (7 ribosomal operons), and/or the deleterious expression of these genes. Finally with other experiments, the whole 2.86 Mb genome was cloned in 21 pieces ranging from 6 kbp to 515 kbp. As TAR-Cloning imposes constraints in the choice of restriction sites, a new cloning method, called CREasPy-Fusion, was developed. This method allows the simultaneous cloning and engineering of mega-sized genome in yeast using the CRISPR-Cas9 system, after direct bacterial cell to yeast spheroplast cell fusion. As a proof of concept, we demonstrated that the method can be used to capture a piece of genome, or to clone and edit the whole genome from six different *Mycoplasma* species. This method was then adapted to *Bsu*, showing for the first-time yeast spheroplast and Gram+ protoplast cell fusion. A fragment of ~150 kb has been successfully cloned in yeast.

Even if, the entire *Bsu* genome has not yet been cloned in yeast, several critical elements have been identified. First of all, this work underlines the importance of the cloning method to be adopted depending on the organism of interest. Then, it emphasizes the existence of both biological and technical factors that explain current difficulties and that will have to be taken into account in subsequent experiments. Finally, it enabled the development of the new in-yeast cloning method called CREasPy-Fusion which expands the catalog of technics already described. Through its versatility, it opens up prospects for the capture of large genome fragments, the suppression of problematic *loci*, and to support the assembly of synthetic fragments.

Key words : *Bacillus subtilis*, Genome cloning, Genome engineering, Yeast, Bacterial chassis

UMR 1332 Biologie du Fruit et Pathologie – Equipe Mollicutes

Centre INRAE Bordeaux-Aquitaine Bâtiment IBVM

71 Avenue Edouard Bourlaux, CS 20032

33882 Villenave d'Ornon Cedex, FRANCE

REMERCIEMENTS

REMERCIEMENTS

Je souhaiterais tout d'abord adresser mes remerciements à Monsieur le Professeur Matthieu Jules, professeur à INRAE, AgroParisTech et à l'Université Paris-Saclay, de me faire l'honneur d'avoir accepté de juger mon travail et de présider mon jury de thèse. Merci également pour tous les précieux conseils afin de m'aider à rentrer dans le monde de Bacillus durant ces années de thèse et ton expertise incontestable.

Je remercie Monsieur le Professeur Sébastien Rodrigue, professeur titulaire à l'Université de Sherbrooke, d'avoir accepté de juger mon travail et de me faire l'honneur de participer à mon jury de thèse en tant que rapporteur.

Je souhaite également remercier Madame la Docteur Fayza Daboussi, Directrice de recherche à INRA-INSA de Toulouse, pour l'honneur que vous me faites en acceptant d'être rapportrice de ma thèse et de juger mon travail.

Je remercie aussi Monsieur le Docteur Fabien Labroussaa, Docteur à l'Université de Bern, d'avoir accepté d'examiner mes travaux et de me faire l'honneur de prendre part à ce jury de thèse.

Je commencerais par remercier infiniment Monsieur le Professeur Alain Blanchard directeur de l'Institut des Sciences de la Vigne et du Vin, pour avoir accepté de m'encadrer tout au long de cette thèse. Je suis extrêmement reconnaissante pour votre soutien sans relâche et vos conseils toujours avisés et très précieux. Merci d'avoir su trouver les mots dans les moments de doute (toujours avec un trait d'humour) et d'avoir su m'accompagner scientifiquement parlant dans l'écriture de ces travaux mais aussi dans mon parcours depuis le Master 2. J'admire énormément votre expertise du monde de la recherche et des mycoplasmes et cela m'a accompagné avec force (surtout avec le petit crochet vers des expériences sur les myco et dans l'écriture du papier). Un grand trugarez vas Alain !

Les mots ne seraient pas assez forts pour exprimer à quel point je souhaite remercier Madame la Docteur Carole Lartigue (autrement connue sous le nom de Carole mon Idole), pour son encadrement inestimable au cours de ces travaux de thèse. Un immense Merci Carole, de m'avoir fait l'honneur d'être ma directrice de thèse et pour tout ce que tu m'as apporté au cours de ces très belles années. Scientifiquement parlant tu es une vraie source d'inspiration pour moi, et grâce à ton expertise et à ton savoir tu m'as souvent donné la force de taper du poing sur la paillasse et d'avancer. J'ai énormément appris à ton contact, tu m'as laissé proposer mes idées tout en calmant parfois mes impulsions quand je fonçais tête baissée. Merci aussi pour ta confiance, car tu m'as laissé me débrouiller seule et prendre des responsabilités avec lesquelles j'ai grandi. Pour moi tu es une vraie leader car non seulement tu t'intéresses à ceux avec qui tu travailles, tu ne les laisses jamais tomber et surtout, tu admets apprendre aussi en retour. Au-delà de la chercheuse brillante j'ai aussi été accompagnée par une femme formidable et profondément bienveillante. Tu as toujours répondu présente quand j'en avais besoin d'un point de vue personnel. Alors merci pour tout, je serai toujours extrêmement fière de dire que j'ai appris auprès de la meilleure directrice de thèse qui puisse être !

Merci également à Monsieur le Docteur Pascal Sirand-Pugnet pour ton soutien, les discussions autour du projet et tous les conseils avisés sur mes travaux depuis mon arrivée dans l'équipe Molli. Je pense sincèrement que l'équipe Molli est chanceuse de pouvoir compter sur un chef d'équipe aussi altruiste et qui prend toujours soin d'avoir un mot ou un conseil pour chacun. J'admire le scientifique que tu es et dont les connaissances nous poussent toujours plus à creuser et à réfléchir sur nos expériences (et stratégie le cas échéant). Merci infiniment Pascal.

Je tiens également à remercier l'ensemble des membres de l'équipe Syber de Micalis avec lesquels j'ai eu la joie et le plaisir d'échanger au cours de notre collaboration sur le projet Bacillus 2.0. Merci de m'avoir accueillie pendant deux semaines pour que je puisse me familiariser avec la bactérie qui a retenu (presque) toute mon attention au cours de ces dernières années. Je tiens à remercier

particulièrement le Pr. Matthieu Jules, le Dr. Teddy Charbonnier et très sincèrement merci au Dr. Etienne Dervyn pour tous les conseils que j'ai pu recevoir. Merci aussi d'avoir pris le temps de répondre à mes milliers de questions et pour votre curiosité sans relâche pour mon travail. Je n'oublie pas non plus le Dr. Vincent Sauveplane et la Dr. Anne Gaelle Planson : merci pour nos échanges et brainstorming au cours de nos réunions !

Je remercie également chaque membre de l'équipe Molli que j'ai pu côtoyer ces dernières années. Si j'ai autant aimé venir travailler au laboratoire c'est en partie grâce à vous. Je commencerais par remercier les membres de l'équipe « Molli-Phyto », Nathalie et Sandrine (promis j'arrête mon bazar à côté de votre bureau avec la sorbonne) merci pour votre bonne humeur et votre bienveillance. Merci aussi à Xavier, Sylvie, Pascal S, Sophie, Laure B, Jean-Paul, Thierry, Marielle, Delphine, Jean-Saïd, Denis, pour les moments de partage scientifique ou tout simplement au cours des évènements Molli ou autour d'un café. Un grand merci à Marie-Pierre (alias Mpi) pour ta gentillesse, ton aide si précieuse, les tips spécial Mpi et pour tout le plaisir que j'ai eu à apprendre à manipuler les levures avec toi. Merci aussi à Sybille d'avoir toujours pris le temps pour discuter (hors hygiène et sécurité) et me redonner du baume au cœur ces derniers temps pendant l'écriture. Un grand merci à Martine sans qui le labo ne tournerait pas aussi bien et parce que j'ai aussi pu compter sur ton oreille attentive ! Je ne t'oublierai pas non plus Suzanne, ta douceur et les discussions que nous avons pu avoir ensemble m'ont apporté du soutien au cours de ces années. Je tiens à remercier aussi nos Wonder-Gestionnaires Isa, Annie et Laurence pour leur rôle crucial et les moments de rire partagés au coin café !

Je remercie particulièrement et très chaleureusement le Dr. Christophe Garcion, tout d'abord merci d'avoir accepté de m'accueillir dans ton bureau pendant mon stage de M2. Tu ne le savais pas encore mais ça n'allait pas être de tout repos (surtout avec Camille ou Baby en plus). Tu es pour moi un brillant chercheur et j'ai énormément appris et cogité au quotidien à ton contact (surtout quand je râlais parce que c'était encore négatif !!) Merci pour ta gentillesse et ton rôle de coach et de soutien, et je soulignerais aussi ton humour spécial qui a amené de la légèreté quand il le fallait et qui m'a fait quand même (beaucoup) rire !

Je ne t'oublie pas non plus Camille pour nos délires en chanson et pour la bonne humeur que tu amenais chaque jour avec toi dans le bureau : merci !

Je poursuis donc avec les membres de l'équipe « Molli-BS », merci à Luis, Paul, Estelle (je suis extrêmement heureuse que l'on se soit redécouvertes à la fin de ta thèse), le duo infernal des Vince'. Merci Vincent T pour toutes les bières partagées autour d'un barbecue et les discussions sans fin (autour des levures ou bien de la vie et des fées ;)) ton énergie contagieuse a participé à donner une grande dynamique au labo que je nommerais les années bonheur ! Merci à toi aussi Vincent Baby (alias Babytos) tu fais désormais partie de mes trois québécois préférés sur terre derrière Cath et Céline (oui oui je parle bien de Céline Dion). Je dois dire que je garderai toujours de très bons souvenirs de nos discussions scientifiques dans le bureau avec lesquelles j'ai beaucoup appris, et puis aussi de tous nos moments partagés à discuter de nos idoles respectives et de la science ! Merci à toi aussi Laure Bataille pour ta bonne humeur, ton humour (Bondouuur) et pour les cocktails !!! Merci d'avoir été là en soutien lorsque cela était nécessaire ! Merci à toi Catherine, tu es la définition même que l'on rencontre parfois des personnes dont il nous semble que l'on se connaît depuis 20 ans. Comme promis ne t'inquiète pas j'arrive bientôt de ton bord de l'Atlantique !! Merci aussi à Flavie ma binôme et ma Flavinette. J'ai adoré former notre duo Flavrielle et travailler avec toi. Tu as énormément contribué et apporté au projet et à mon travail, toujours avec tellement de précision et de soin. Tu m'as aussi beaucoup manqué une fois ton contrat terminé. Je suis heureuse d'avoir rencontré une personne en or comme toi et sache que tu as largement contribué à embellir mes journées au labo !

Merci à toi Maman GG : il n'y aurait pas assez de mot pour te dire à quel point je suis heureuse de t'avoir rencontrée. Tu auras été d'une aide monumentale pendant ma thèse et puis aussi une coach du cœur indispensable. Je sais tu le dis souvent sans toi on y arriverait, mais avec toi c'est tellement mieux et je souhaite à tout le monde d'avoir un jour le bonheur de rencontrer une Géraldine. Pour citer

Agnès Ledig j'ajouterais donc « C'est ça la vie, c'est courir pieds nus, savourer quand le sable est fin ou l'herbe est douce. Accepter le gravier et parfois le verre pilé. Et se dire qu'on avance malgré tout, quelle que soit la surface ». Merci pour tout d'avoir été là quand il fallait marcher sur du verre pilé et aussi d'être toujours là quand dorénavant je cours dans les montagnes vertes !!

Merci également au Docteur Yonathan Arfi, j'ai un immense respect pour le chercheur que tu es (un peu artiste parfois). Je suis admirative de tout ton savoir que tu nous partages sans hésitation lorsque l'on vient te demander de l'aide. Merci de m'avoir supportée pendant les TPs et lorsque je venais te demander des conseils au cours du monitorat. Merci pour les moments de rire (sur fond de blague douteuse), pour ton aide (notamment dans l'écriture de ce manuscrit), pour tes conseils tellement pertinents et aussi réconfortants !

S'il y a bien une personne que j'ai le devoir de remercier c'est toi Dr. Fabien Curtain (Fabinou mon préféré largement en haut du classement) parce que je te dois beaucoup. Déjà je suis extrêmement heureuse d'avoir été ta première stagiaire. Tu m'as tout de suite beaucoup inspirée et grâce à toi et à Chloé (merci infiniment aussi), j'ai encore pris plus goût à la science ! Tu as été ce pilier dans bien des moments et j'ai beaucoup avancé à ton contact. C'est aussi grâce à toi que j'ai pu rencontrer Carole et qui m'a mis un peu sur ce chemin de la thèse. Au-delà du coach, tu es généreux, bienveillant, intelligent dans ton travail et aussi dans ta vie. J'ai adoré apprendre avec toi et travailler avec toi. D'un maître de stage, tu es devenu un ami. Merci pour les énormes soirées qu'on a faites ensemble et les discussions tellement précieuses que l'on a pu avoir. Le seul bémol c'est que tu as mis un boulet sur mon chemin, bon je l'ai ramassé mais merci du cadeau ! Je sais qu'avec la fin de la thèse cela n'empêchera pas que l'on se retrouve autour de moments inoubliables ! Je te souhaite le meilleur et si je me dis souvent que j'ai eu la chance de te rencontrer en M1, ta future Mini-Toi aura la chance d'avoir un excellent Papa !! Je te souhaite le meilleur Fab. « Dans la vie il n'y a pas de hasard, juste de belles rencontres ».

Je remercie aussi les dernières recrues du labo, une équipe de choc, Arthur le maître des quizz, Bastos (sans toi je n'aurais jamais pu faire d'aussi belles images de microscopie), Francesca (ma Topina) je suis très heureuse que tu sois arrivée au labo et j'ai hâte de découvrir ton beau pays ! Merci aussi à toi Djudju, je suis extrêmement fière de la sacrée relève que tu fais et de t'avoir eu en tant que stagiaire. Je vous souhaite le meilleur pour la suite, votre arrivée dans une période très particulière a ramené un nouveau souffle dans notre équipe. « A un moment donné... ».

Je souhaite finir par te remercier Docteur Thomas Ipoutcha. Tout d'abord parce que grâce à toi j'ai passé presque quatre années de thèse dans le rire et la bonne humeur. De nos chamailleries et blagues incessantes, à nos soirées arrosées au Patx, en passant (tout de même) par des heures de travail et papotages scientifiques en tant que voisins de paillasse du labo de biomol ; j'ai aussi pu compter sur toi dans les périodes plus compliquées et vivre de très beaux moments sous les étoiles alignées. D'un stagiaire un peu arrogant mais malin (certes), que j'ai été obligée d'avoir dans les pattes, j'ai découvert un binôme inestimable qui partage mon amour des Sciences. Je sais que le meilleur est à venir pour toi dans ta carrière de brillant chercheur et j'espère aussi partager encore longtemps ta carrière de « metteur dans des cartons ».

PS : un merci aussi à mon deuxième coloc Romain et aux zouaves qui ont illuminé pas mal de mes soirées je nommerais les Dr. Duran & Dr. Gassama. A vous trois vous êtes sur le podium de mes prefs' mais je ne dirais pas dans quel ordre !

Dans cette aventure j'ai aussi pu compter sur le soutien d'amis formidables et complètement fous :

Merci à toi Mathilde, mon Mathou la famille dorénavant ! On en a connu des péripéties, des soirées arrosées et des moments intenses depuis l'internat et les années STL. Merci d'être toujours présente dans ma vie et d'avoir été là dans les bons comme dans les mauvais moments au cours de cette

thèse et bien au-delà. Merci pour tout Je te souhaite le meilleur dans ta thèse, je sais que tu vas briller et je te donne rendez-vous dans un an en Post-Doc sur l'autre continent ;).

Les copains de l'IUT de Brest, bien que tous un peu éloignés géographiquement comment ne pas dire que nos rendez-vous annuels font partie de mes fondamentaux. Merci à Laurie (ma Bibi), La Mama, Marine la plus canon, Jez on my feet, Marcass, et Marot toujours en retard, Queq et Remi les plus grosses tecros que je connaisse, Salom' le beau goss et Flo le ghetto, sans oublier Postec le papa ! On a tous évolué et dans pleins de directions mais je sais que l'on arrivera toujours à se recroiser et à refaire la fête comme dans nos années folles.

Je ne peux sans aucun doute oublier de remercier mes MIMUs préférés : Guénael mon bin-homme, Ana (ma Queen P préférée), Delphine et Dimitri, et Romain (bon toi tu es un MIMU adopté).

Ma Cyrielle, merci à toi ma sœur de cœur pour toutes ces aventures vécues ensemble. Au-delà d'être mon amie de toujours tu es la famille. Tu es celle qui me connaît le mieux et qui me ramène le mieux sur Terre : Merci pour ton honnêteté à toutes épreuves. Bien que loin physiquement j'ai pu énormément compter sur toi durant cette thèse. Il y a aussi un petit bout de bonheur que j'ai vu naître au cours de ces dernières années, Célia dont je suis très fière d'être la marraine, merci à vous deux d'avoir mis tant de magie quand j'ai rencontré des problèmes dans ma thèse (mon bébé à moi). Je sais que le meilleur est encore à venir pour partager tous pleins de beaux moments de vie avec vous !

Je terminerais par remercier ceux qui sont là depuis toujours. Merci infiniment Papou et Mounette pour votre amour et votre aide. Vous m'avez accompagnée et soutenue dans mes choix bien que risqués ou complètement foireux et malgré vos inquiétudes parfois. Alors oui je sais Papa « Les chercheurs ce ne sont pas forcément des trouveurs » mais je sais maintenant que je suis une « trouveuse ». J'ai trouvé ma voie et ce que je voulais vraiment faire et ça, c'est en partie grâce à vous, grâce à votre confiance en moi, et parce que vous avez toujours tout donné et je vous en suis profondément reconnaissante. J'espère vous rendre fière et je tiens à vous certifier que si j'en suis là c'est en très grande partie grâce à vous. Je vous aime. Merci aussi à toi Manu, mon grand frère. Tu m'as toujours protégée et soutenue de près ou de loin, et je trouve toujours autant de réconfort dans tes paroles et conseils. Malgré que bon tu restes un sac à merde soyons honnêtes ;) tu te rattrapes et t'améliores grâce à la meilleure des belles-sœurs j'ai nommé Laetichou. Merci à toi aussi, ton arrivée dans la famille depuis (déjà !) 10 ans est une véritable source de bonheur (et promis pas seulement parce que tu te tâches presque autant que moi haha). Les allers-retours Bordeaux-Paris m'ont permis de bien décompresser quand j'avais besoin de me ressourcer, merci de m'avoir chouchoutée. Grâce à vous je peux aussi savourer le plaisir d'être la marraine et la tata de deux merveilleuses petites Kikinettes Lara et Eloïse. Leur malice et leurs rires ont parfois été les seuls remèdes pour décrocher un peu et m'évader du monde de la recherche quand cela devenait plus que nécessaire.

J'espère n'avoir oublié personne. Je concluais ces remerciements en m'inspirant du poème de Jean d'Ormesson « Le train de la vie » (dédicace pour toi GG), je souhaite à nouveau remercier infiniment toutes les personnes qui ont participé de près ou de loin à mon voyage. Merci d'être montés dans mon train et d'y avoir pris place, d'avoir fait partie de cette aventure. J'espère que de là où vous étiez installés cela aura été un agréable moment passé avec moi, car de ma place c'était un voyage intense que je suis fière d'avoir réalisé.

TABLE DES MATIÈRES

TABLE DES MATIÈRES

| | |
|--|----|
| REMERCIEMENTS | 1 |
| TABLE DES MATIÈRES | 5 |
| LISTE DES ABRÉVIATIONS | 12 |
| AVANT PROPOS | 14 |
| OBJECTIFS DE LA THÈSE | 18 |
| ÉTUDE BIBLIOGRAPHIQUE..... | 20 |
| I. <i>Bacillus subtilis</i> : de l'organisme modèle aux applications biotechnologiques | 20 |
| I.1. Caractéristiques générales | 21 |
| I.2. Sporulation chez <i>B. subtilis</i> | 23 |
| I.3. Compétence naturelle..... | 25 |
| I.4. Biofilms..... | 26 |
| I.5. Caractéristiques génétiques et systémiques..... | 28 |
| I.5.1. Séquençage et annotations du génome de <i>B. subtilis</i> | 28 |
| I.5.2. Autres données « omiques »..... | 30 |
| I.6. Usine bactérienne de production à visée biotechnologique | 30 |
| I.6.1. Ingénierie génétique chez <i>Bacillus subtilis</i> | 31 |
| I.6.1.1. Transformations | 31 |
| I.6.1.2. Outils génétiques..... | 33 |
| I.6.1.3. Vecteurs et contrôle de l'expression de gènes chez <i>Bacillus subtilis</i> | 37 |
| I.6.2. Sécrétion de protéines recombinantes..... | 40 |
| I.6.3. Exemples d'applications industrielles | 42 |
| II. Biologie de synthèse : vers le développement de châssis minimaux | 43 |
| II.1. Généralités sur la biologie de synthèse | 43 |
| II.2. Développement de châssis minimaux en BS..... | 45 |

| | |
|---|----|
| II.2.1. Notion de cellules minimales et d'essentialité de gènes..... | 45 |
| II.2.2. Nouvelle génération de châssis bactériens en BS..... | 47 |
| II.2.3. Réglementation de la biologie de synthèse et traçabilité des châssis..... | 49 |
| III. La levure comme plateforme d'ingénierie de fragments d'ADN ou de génomes microbiens..... | 51 |
| III.1. Méthodes de clonage de génomes dans la levure..... | 53 |
| III.1.1. Méthode de clonage dite « Classique » | 54 |
| III.1.2. Méthode de TAR-Cloning | 55 |
| III.1.3. Méthode de Fusion cellulaire..... | 56 |
| III.1.4. Méthode de clonage par CREasPy-Cloning | 56 |
| III.2. Outils d'ingénierie génomique chez la levure..... | 57 |
| III.2.1. Méthodes TREC et TREC-IN | 57 |
| III.2.2. Méthode dérivée du système CRISPR-Cas9 | 58 |
| III.3. Voies de sortie du cycle d'ingénierie de génome microbien | 60 |
| III.3.1. Transplantation de génome (TG) | 60 |
| III.3.2. Méthodes de transfert de portion de génome | 62 |
| MATÉRIELS ET MÉTHODES | 64 |
| I. Souches microbiennes et conditions de culture | 64 |
| I.1. <i>Escherichia coli</i> | 64 |
| I.2. Les mycoplasmes..... | 65 |
| I.3. <i>Bacillus subtilis</i> | 65 |
| I.4. <i>Saccharomyces cerevisiae</i> | 66 |
| II. Techniques de Biologie moléculaire | 67 |
| II.1. Plasmides..... | 67 |
| II.1.1. Plasmides commerciaux ou précédemment construits..... | 67 |
| II.1.2. Plasmides construits au cours de ces travaux..... | 73 |

| | |
|--|----|
| II.2. Mini-préparation de plasmides | 78 |
| II.3. Midi-préparation de plasmides | 78 |
| II.4. Vérification des plasmides par séquençage | 78 |
| II.5. Extraction d'ADN génomique total | 78 |
| II.5.1. Extraction d'ADN génomique de <i>Bacillus subtilis</i> | 79 |
| II.5.2. Extraction d'ADN génomique de <i>Saccharomyces cerevisiae</i> | 79 |
| II.5.3. Extraction d'ADN génomique de mycoplasmes | 80 |
| II.6. PCRs simplexes et multiplexes | 82 |
| II.6.1. PCR simplexe (Polymerase Chain Reaction ou réaction en chaîne de la polymérase)..... | 82 |
| III.6.2. PCR multiplexe | 83 |
| II.7. Amplification des matrices de recombinaison (pansements)..... | 83 |
| II.7.1. Pansements avec les éléments levures (YCP) | 84 |
| II.7.2. Pansements de jonction | 84 |
| II.8. Purification des ADNs | 85 |
| II.9. Restriction enzymatique des ADNs | 85 |
| II.10. Réaction de Gibson Assembly | 86 |
| II.11. Réaction de modification des ADNs | 86 |
| II.11.1. Phosphorylation des ADNs | 86 |
| II.11.2. Déphosphorylation des ADNs | 86 |
| II.11.3. Ligature des ADNs | 87 |
| II.12. Quantification des ADNs | 87 |
| II.13. Electrophorèse sur gel d'agarose | 88 |
| II.14. Préparation des génomes en blocs d'agarose | 88 |
| II.14.1. Extraction des génomes de mycoplasmes en bloc d'agarose..... | 88 |
| II.14.2. Extraction des génomes de <i>Bacillus subtilis</i> en bloc d'agarose : | 89 |

| | |
|--|-----|
| II.14.3. Extraction des génomes des transformants levures en bloc d'agarose : | 91 |
| II.15. Digestion des blocs d'agarose | 91 |
| II.15.1. Digestion des blocs d'agarose par restriction enzymatique | 92 |
| II.15.2. Digestion des blocs d'agarose par l'enzyme β -agarase | 92 |
| II.16. Migration des grands ADNs par électrophorèse en champs pulsé (PFGE) | 92 |
| II.16.1. Evaluation de la qualité des génomes..... | 93 |
| II.16.2. Analyse des transformants levures propageant des génomes de mycoplasmes | 93 |
| II.16.3 Analyse des transformants levures propageant des fragments de génomes de <i>B. subtilis</i> | 94 |
| III. Transformations microbiennes | 96 |
| III.1. Transformations bactériennes d' <i>Escherichia coli</i> | 96 |
| III.1.1. Transformation par électroporation | 96 |
| III.1.2. Transformation par choc thermique | 96 |
| III.2. Transformations bactériennes des mycoplasmes..... | 97 |
| III.2.1. Transformation de <i>Mycoplasma bovis</i> et <i>Mycoplasma agalactiae</i> | 97 |
| III.2.2. Transformation de <i>Mycoplasma gallisepticum</i> | 98 |
| III.3. Transformation bactérienne de <i>Bacillus subtilis</i> | 98 |
| III.4. Transformations de <i>Saccharomyces cerevisiae</i> | 100 |
| III.4.1. Transformation de plasmides par le Lithium Acétate/ADN simple brin/PEG. 100 | |
| III.4.2. Clonage par transformation de sphéroplastés levures au PEG | 101 |
| III.4.3. Analyses des transformants levures | 104 |
| IV. Ingénierie de génomes bactériens | 104 |
| IV.1. Ingénierie du génome de <i>B. subtilis</i> | 104 |
| IV.2. Ingénierie des génomes de mycoplasmes par édition de base | 105 |
| IV.3. Test d'activité de nucléases de surfaces..... | 105 |

| | | |
|--|--|-----|
| V. | Transplantation de génomes bactériens..... | 106 |
| V.1. | Transplantation de génomes de <i>Mcap</i> et <i>Mmc</i> après CReasPy-Fusion..... | 106 |
| V.2. | Transplantation de génomes de <i>Mccp</i> après CReasPy-Fusion | 107 |
| V.3. | Analyse des transplants | 107 |
| CHAPITRE 1 – Essais de clonage d’un génome entier réduit de <i>B. subtilis</i> dans la levure..... | | 108 |
| Introduction..... | | 108 |
| I. | Essais de clonage par la méthode de CReasPy-Cloning | 110 |
| I.1. | Pré-chargement des levures avec le système CRISPR-Cas9..... | 111 |
| I.2. | Préparation de la matrice de recombinaison..... | 112 |
| I.3. | Préparation des génomes de <i>B. subtilis</i> en bloc d’agarose..... | 112 |
| I.3.1. | Première préparation avec le protocole « ED »..... | 113 |
| I.3.2. | Deuxième préparation et optimisation du protocole « ED »..... | 114 |
| I.3.3. | Troisième préparation et standardisation avec le protocole « Biorad » | 116 |
| I.4. | Essais de clonage du génome entier de <i>B. subtilis</i> | 116 |
| I.5. | Conclusion sur les essais de clonage par la méthode de CReasPy-Cloning | 119 |
| II. | Essais de clonage d’un génome pré-marqué de <i>B. subtilis</i> | 120 |
| II.1. | Insertion des éléments levures dans le génome de <i>B. subtilis</i> par transformation | 120 |
| II.1.1. | Evaluation de la croissance de la souche MGP192 dans les milieux MG1 et MG2 | 121 |
| II.1.2. | Essai de transformation avec de l’ADNg pré-marqué..... | 123 |
| II.1.3. | Nouvel essai de transformation avec différents ADNs pré-marqués | 125 |
| II.2. | Essais de clonage du génome MGP192 par la méthode « classique » | 126 |
| II.3. | Essais de clonage par Fusion | 128 |
| II.3.1. | Production des protoplastes de <i>B. subtilis</i> | 128 |
| II.3.2. | Premier essai de clonage de génome par Fusion <i>B. subtilis</i> -Levure | 129 |
| II.3.3. | Second essai de clonage par Fusion <i>B. subtilis</i> -Levure | 131 |

| | |
|---|-----|
| Discussion - Conclusion | 132 |
| CHAPITRE 2 – Clonage de fragments de génome de <i>B. subtilis</i> par la méthode de TAR-Cloning | 136 |
| Introduction..... | 136 |
| I. Essai de clonage de fragments de génome de <i>B. subtilis</i> | 138 |
| II. Clonage du « Fragment 7 » et analyses bio-informatiques | 141 |
| II.1. Nouvelles tentatives de clonage du « Fragment 7 »..... | 141 |
| II.2. Restriction du « Fragment 7 » en onze sous-fragments par les enzymes FseI et AsiSI | 142 |
| II.3. Bilan des expériences de TAR-Cloning sur le Fragment 7 | 143 |
| II.4. Analyse de gènes toxiques putatifs..... | 145 |
| II.5. Mise en évidence de régions identiques et répétées sur le Fragment 7 | 146 |
| III. Clonage du « Fragment 1 » et analyses bio-informatiques | 148 |
| III.1. Nouvelle tentative de clonage du « Fragment 1 »..... | 148 |
| III.2. Restriction du « Fragment 1 » en sept sous-fragments par l'enzyme FseI..... | 151 |
| III.3. Analyse de la proportion en ARS..... | 152 |
| III.4. Du TAR-Cloning au TAR-Assembly..... | 155 |
| Discussion-Conclusion | 155 |
| CHAPITRE 3 – CReasPy-Fusion : Développement d'une nouvelle méthode de clonage de génomes bactériens dans la levure..... | 159 |
| Introduction..... | 159 |
| I. Clonage et ingénierie de génomes de mycoplasme par CReasPy-Fusion..... | 161 |
| I.1. Clonage par Fusion chez les mycoplasmes | 164 |
| I.2. Preuve de concept de la méthode de CReasPy-Fusion..... | 165 |
| I.3. CReasPy-Fusion appliquée à deux mycoplasmes pathogènes majeurs du cluster mycoïdes : clonage de génomes entiers ou capture de fragments de génomes | 167 |

| | |
|---|-----|
| I.4. Mise en évidence d'un facteur génétique influençant l'efficacité de la méthode à travers le clonage du génome de <i>M. gallisepticum</i> | 169 |
| I.5. Extension de la méthode à d'autres espèces de mycoplasmes appartenant au groupe phylogénétique Hominis | 175 |
| I.6. Bilan et analyses des expériences de CReasPy-Fusion appliquées aux mycoplasmes. | 177 |
| II. Adaptation de la méthode de CReasPy-Fusion à <i>B. subtilis</i> | 182 |
| II.1. Validation de la Fusion par transfert d'un plasmide navette | 183 |
| II.2. Transfert et modification d'un plasmide navette par CReasPy-Fusion | 187 |
| II.3. Capture d'un fragment de génome de <i>B. subtilis</i> | 190 |
| II.3.1. Capture d'un fragment de génome de <i>B. subtilis</i> portant les éléments levures ARS-CEN-HIS3 | 192 |
| II.3.2. Capture d'un fragment de génome de <i>B. subtilis</i> avec contre sélection au 5-FOA | 194 |
| Discussion - Conclusion | 196 |
| DISCUSSION GÉNÉRALE ET CONCLUSION | 201 |
| ANNEXES..... | 211 |
| RÉFÉRENCES BIBLIOGRAPHIQUES | 219 |

LISTE DES ABBRÉVIATIONS

LISTE DES ABRÉVIATIONS

| | |
|---------------|---|
| <i>aadA</i> | gène de résistance à la spectinomycine |
| A, C, T, G, U | adénine, cytosine, thymine, guanine, uracile |
| ADN | acide désoxyribonucléique |
| ADNg | ADN génomique |
| ARN | Acide ribonucléique |
| ARNg | ARN guide |
| ARNr | ARN ribosomique |
| ARS | Autonomously Replicating Sequence |
| ATP | Adénosine Tri-Phosphate |
| BAC | Bacterial Artificial Chromosome |
| BET | Bromure d’Ethidium |
| BS | Biologie de Synthèse (ou Biologie Synthétique) |
| Cas | CRISPR associated protein |
| CEN | centromère |
| crARN | CRISPR ARN |
| CRISPR | Clustered Regularly Interspaced Short Palindromic Repeats |
| DMSO | Diméthylsulfoxyde |
| dNTP | désoxyribonucléotide 5’ triphosphate |
| dpBS | Solution saline dans un tampon phosphate de Dulbecco |
| E | gène essentiel |
| EDTA | éthylène diamine tétra acétate |
| HBSS | Hank’s Balanced Salt Solution |
| HIS | marqueur d’auxotrophie pour l’histidine |
| IS | Insertion Sequence |
| kpb | kilo paire de bases |
| LB | Luria Bertani |
| LYS | marqueur d’auxotrophie pour la lysine |
| Mb | Méga paire de bases |
| MSC | Multi Site de Clonage |
| NE | gène non essentiel |
| NEB | New England Biolab |
| NGS | Séquençage nouvelle génération |
| NHEJ | Non Homologous End Joining |
| <i>oriC</i> | origine de répllication chromosomique |
| Pb | Paire de bases |
| PAM | Protospacer Adjacent Motif |
| PCR | Polymerase Chain Reaction (Réaction en chaîne par polymérase) |
| PEG | Polyéthylène Glycol |
| PFGE | Pulsed Field Gel Electrophoresis |
| PMSF | Fluorure de phénylméthylsulfonyle |
| Qsp | Quantité suffisante pour |
| RPM | Revolution Per Minute |
| RM | Restriction Modification |

| | |
|------|--|
| SDS | Sodium Dodecyl Sulfate |
| SMM | Sucrose Maleate Magnesium |
| SOC | Super Optimal Broth |
| TAE | Tris Acétate EDTA |
| TAR | Transformation Associated Recombination |
| TBE | Tris Borate EDTA |
| TE | Tris EDTA |
| TRP | marqueur d'auxotrophie pour le tryptophane |
| URA | marqueur d'auxotrophie pour l'uracyl |
| UFC | Unité Formant Colonie |
| YAC | Yeast Artificial Chromosome |
| YCP | Yeast Centromeric Plasmid |
| YPDA | Yeast Peptone Dextrose Adénine |

AVANT-PROPOS

AVANT PROPOS

La biologie de synthèse (BS) a pour but de construire de nouveaux systèmes biologiques afin de faciliter le développement d'applications dans le domaine des biotechnologies. Additionnellement à cela, la BS contribue aussi à améliorer les connaissances fondamentales scientifiques.

En se basant sur la réécriture et la fabrication « à façon » de composants ou de systèmes existants (bactéries, algues...), un des objectifs de la BS est de produire des souches industrielles toujours plus performantes et respectueuses de l'environnement. Certaines applications de la BS se basent sur la création d'un « châssis », faisant référence à un organisme de base, modulable à souhait, et utilisé comme hôte d'accueil pour (i) héberger physiquement des composants génétiques, et (ii) permettre leur expression en fournissant les ressources nécessaires à leur fonctionnement. Des micro-organismes modèles comme les bactéries *Escherichia coli*, *Bacillus subtilis* ou encore *Pseudomonas putida* peuvent être utilisés comme organisme de départ pour le développement de telles souches châssis.

Une stratégie prometteuse tend à utiliser des « châssis simplifiés » possédant des génomes réduits ou des « châssis minimaux » ne disposant que de l'équipement indispensable à leur multiplication. Ceci permettrait, en effet, de contourner certaines difficultés rencontrées au cours du processus de fabrication d'un composé d'intérêt, en simplifiant par exemple, les réseaux métaboliques présents dans ces souches « hôtes » et, en favorisant la voie de synthèse du produit souhaité.

B. subtilis est une bactérie non pathogène largement étudiée par la communauté scientifique et considérée comme le modèle bactérien Gram+ par excellence, à l'instar d'*E. coli* pour les bactéries Gram-. *B. subtilis* possède des caractéristiques particulières (statut GRAS, croissance rapide, génétiquement stable, système de sécrétion) qui en font un micro-organisme couramment utilisé en industrie comme plateforme de production de molécules d'intérêt et ce, dans des secteurs aussi divers que la santé, les biotechnologies ou encore de l'environnement. La production de châssis minimaux à partir de *B. subtilis* pourrait constituer une avancée importante dans l'amélioration de souches bactériennes à visée industrielle.

Actuellement, des outils génétiques permettent la réduction du génome de *B. subtilis* par cycles itératifs, mais ces outils restent trop peu efficaces pour envisager une ingénierie génomique à grande échelle. Ainsi, il faut compter environ 1 mois pour supprimer un groupe de gènes et l'efficacité du processus décroît dans les souches à génome réduit. Dès lors, il faudrait des années de travail pour arriver à des souches possédant un génome complètement remanié, ce qui n'est pas compatible avec les exigences de l'innovation industrielle. D'où la nécessité de développer de nouvelles méthodes permettant de réduire, de modifier, voire de redessiner le génome de *B. subtilis*.

Ce projet de thèse s'inscrit au sein d'un projet de recherche ANR « *Bacillus 2.0* », qui vise à transférer des outils de biologie de synthèse, initialement développés pour les mycoplasmes, à *B. subtilis*. L'objectif majeur de ce projet est de promouvoir le développement de méthodes efficaces pour la construction rapide de châssis *B. subtilis* semi-synthétiques et minimaux, qui soient capables de croître dans des milieux dont la composition est définie, et qui soient peu coûteux. L'idée étant de pouvoir programmer ensuite ces derniers grâce aux technologies développées, permettant ainsi de générer des châssis modulables qui s'ajustent aux besoins des industriels.

COLLABORATION ET GENÈSE DE MON PROJET DE THESE

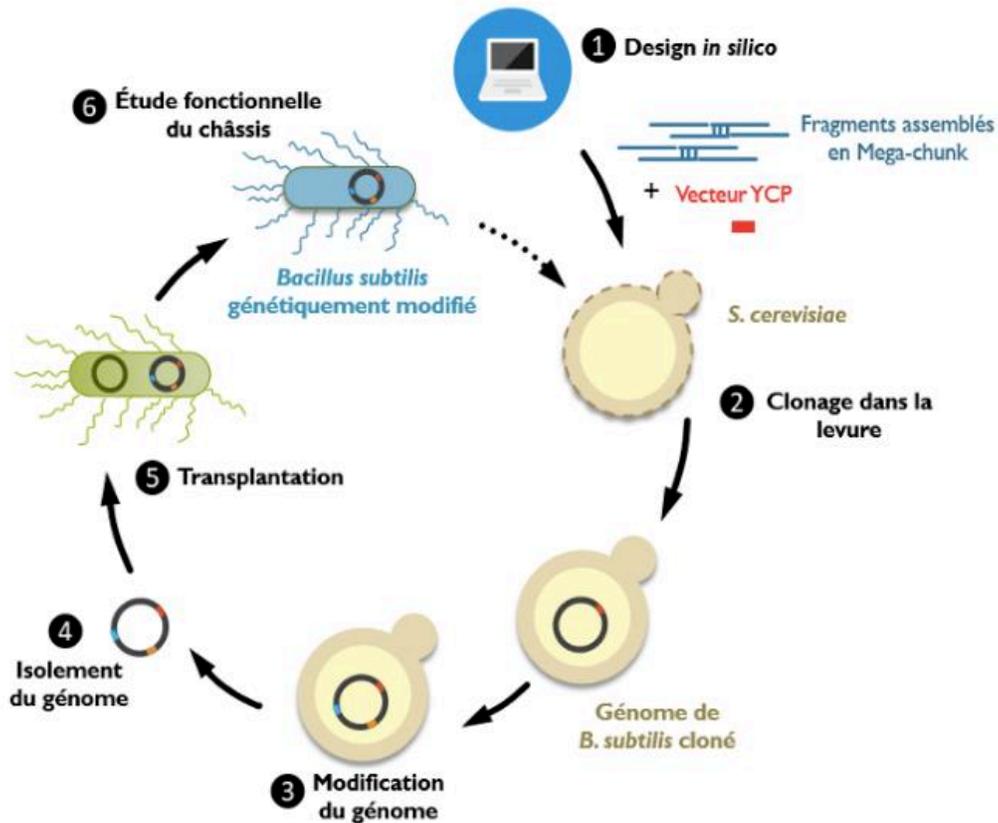
Mon projet de thèse a été développé dans le cadre d'une collaboration entre deux équipes INRAE : l'équipe SyBER (Institut MICALIS, UMR 1319 INRAE / AgroParisTech, Jouy-en-Josas), et l'équipe Mollicutes (UMR 1332 Biologie du Fruit et Pathologie / INRAE, Université de Bordeaux). Le projet de recherche ANR « *Bacillus 2.0* » né de cette collaboration est coordonné par le Professeur Matthieu Jules (Institut MICALIS, UMR 1319 INRAE / AgroParisTech, Jouy-en-Josas) et s'appuie sur les expertises complémentaires des deux équipes : l'équipe SyBER est spécialisée dans les approches de biologie des systèmes et de biologie synthétique pour l'étude et les applications biotechnologiques de *Bacillus subtilis* ; l'équipe Mollicutes développe des recherches sur les bactéries minimales de la classe *Mollicutes* par des approches multiples, notamment des approches innovantes de clonage et de transplantation de génomes bactériens.

Ce projet ANR *Bacillus* 2.0 dans lequel s'est inscrit mon travail de thèse trouve ses racines dans un premier projet collaboratif entre les deux équipes, le projet TBOX4SYNCELL (2017-2019), financé par Toulouse White Biotechnology (TWB) et porté par le Pr. M. Jules et la Dr. C. Lartigue. Ce dernier a permis de jeter les bases du travail, d'évaluer la faisabilité de certaines approches, d'obtenir des résultats préliminaires et finalement de servir de tremplin pour l'obtention d'un financement ANR (AAP 2018).

PROJET ANR « BACILLUS 2.0 » : Construction et transplantation de génomes semi-synthétiques de *Bacillus subtilis* : vers le développement d'une nouvelle génération de châssis bactériens.

Le projet comporte à la fois des dimensions très fondamentales et une visée biotechnologique clairement affichée. Il part du constat que les circuits génétiques synthétiques utilisés en ingénierie microbienne fonctionnent souvent moins bien que prévu en raison d'interactions délétères entre des composants du circuit et de la cellule hôte. Une stratégie pour pallier ce problème repose sur la construction rationnelle de châssis cellulaires simplifiés, mieux connus, et aux propriétés plus prédictibles. *Bacillus* 2.0 a pour but de développer des méthodes efficaces et rapides permettant la construction de génomes artificiels dérivés de *B. subtilis* et leur transplantation dans des protoplastes de *B. subtilis*. Les souches les plus prometteuses seront caractérisées puis soumises à une évolution dirigée afin d'améliorer leur fitness, ce qui représente une condition préalable au développement futur d'usines cellulaires pour les biotechnologies.

Ce projet s'inspire d'une approche innovante développée par une équipe du J. Craig Venter Institute (USA), entre 2007 et 2010, qui regroupe un ensemble de techniques, permettant (i) de synthétiser un génome complet de mycoplasmes à partir d'oligonucléotides, (ii) de répliquer et modifier ce génome dans la levure et, (iii) de le transplanter dans une cellule receveuse. Ces travaux démontraient pour la première fois, la capacité des biologistes à synthétiser des molécules d'ADN à partir d'informations génétiques numérisées puis à les transplanter dans des cellules receveuses (remplaçant toute l'information génétique existante) afin de produire des cellules contrôlées uniquement par l'ADN synthétique nouvellement conçu.



Plateforme d'ingénierie de génomes bactériens chez la levure appliquée au projet ANR *Bacillus* 2.0. Vecteur YCP = Yeast Centromeric Plasmid. Contient les éléments levures permettant le maintien du génome bactérien au sein de la levure (Centromère CEN6, une origine de réplication ARSH4 et un marqueur d'auxotrophie Histidine), ainsi qu'un marqueur de résistance à un antibiotique donné (sélection lors de l'étape de transplantation). **[1]** Design de génome semi-synthétique minimal. **[2]** Assemblage de fragments d'ADN chevauchants au sein de la levure *Saccharomyces cerevisiae*. **[3]** Ingénierie du génome cloné avec les outils disponibles chez la levure (optionnelle). **[4]** Isolement du génome. **[5]** Transplantation de génomes entiers depuis la levure vers une cellule receveuse bactérienne. **[6]** Sélection et caractérisation de nouvelles souches de *B. subtilis* (mutants).

Le projet ANR s'articule en 4 « workpackages » : Clonage d'un génome réduit de *B. subtilis* dans la levure *S. cerevisiae* (WP1) ; transplantation du génome de *B. subtilis* (WP2) ; design et synthèse d'un génome châssis semi-synthétique de *B. subtilis* (WP3) ; évolution dirigée, réorganisation du génome châssis et preuve de concept de production enzymatique (WP4). Ces workpackages associés en un cycle « Design-Build-Test-Learn », forment un pipeline représenté **Figure 1**, qui rassemble dans un ordre défini des technologies de biologie de synthèse pour former une plateforme d'ingénierie de génome bactérien dans la levure.

L'équipe « Syber » de Jouy-en-Josas est en charge du design et de la synthèse du génome semi-synthétique de *B. subtilis* (WP3 et WP4). L'objectif étant de partir du génome de *B. subtilis* 168 à 4,2 Mb et de le réduire au maximum afin d'obtenir un génome semi-synthétique d'environ 2 Mb. En parallèle, l'équipe « Mollicutes », a quant à elle pour objectif de cloner et de modifier le génome de *B. subtilis* dans la levure (WP1). En effet, l'assemblage du génome semi-synthétique se faisant en grande partie dans la levure, il est nécessaire de démontrer que le génome naturel (ou réduit) de *B. subtilis* peut être maintenu chez cet hôte. Cette équipe a aussi pour rôle de développer les méthodes de transplantation de génomes pour *B. subtilis* (WP2).

OBJECTIFS DE LA THÈSE

OBJECTIFS DE LA THÈSE

Pour répondre aux besoins du projet global ANR *Bacillus* 2.0 amorcé au laboratoire depuis fin 2017, ces travaux de thèse ont été effectués au sein de l'équipe Mollicutes (animée par le Dr. Pascal Sirand-Pugnet), sous la direction du Dr. Carole Lartigue et du Pr. Alain Blanchard. Cette thèse avait pour but principal d'adapter des méthodes de biologie de synthèse maîtrisées au laboratoire en travaillant avec le génome de *Bacillus subtilis*. Plus précisément ces travaux se déclinaient selon trois objectifs :

- (i) Développer et repousser les limites des méthodes de clonage de génome entier dans la levure, en clonant un génome réduit de *B. subtilis* à 2,86 Mb. (Étape 2 ; Figure 1)
- (ii) Adapter les outils d'ingénierie disponibles chez la levure pour manipuler aisément le chromosome réduit et avoir à disposition une palette de technologies pour la manufacture de futurs châssis bactériens. (Étape 3 ; Figure 1)
- (iii) Développer des méthodes de transplantation de génome pour *B. subtilis* (Étape 4 et 5 ; Figure 1)

Ce travail de thèse se situe en amont de la production finale de bactéries châssis et de la production de molécules d'intérêt biotechnologique. La mise en place de ces technologies pour un génome tel que celui de *B. subtilis* était très ambitieuse. En effet, son caractère exploratoire ne permettait pas de savoir si nous allions aisément progresser sur ces trois objectifs. Ainsi, les travaux décrits dans ce présent manuscrit correspondent aux résultats obtenus pour les deux premiers objectifs. Cependant, quelques données ont été obtenues pour le troisième objectif et viendront étayer la discussion générale.

Nous avons donc tout d'abord veillé à développer plusieurs méthodes de clonage décrites dans la littérature pour tenter de cloner le génome entier de *B. subtilis* dans la levure. En effet, un prérequis à l'assemblage d'un génome semi-synthétique dans la levure est de démontrer que le génome natif de ce dernier peut être maintenu de façon stable chez cet hôte. Nous avons surtout travaillé à appliquer des protocoles visant à cloner puis à modifier, dans la levure, une version réduite de ce génome (~2,86 Mb, souche *B. subtilis* MGP 192). Ce

travail devait nous permettre d'attester la faisabilité d'une telle approche appliquée à *B. subtilis* (aucun génome de taille >1,8Mb et utilisant un code génétique universel n'ayant été cloné à ce jour dans la levure). Mais aussi, de pouvoir construire et manipuler ce génome cloné avec les outils disponibles chez cet hôte.

Ainsi, le premier chapitre fait état de la mise en place de ces travaux et des divers essais pour cloner le génome réduit dans la levure. Nous nous sommes très vite heurtés à des difficultés et avons affiné notre stratégie en utilisant une méthode pour cloner, cette fois, des fragments de génome. Les données obtenues sont décrites dans le second chapitre et ont permis de mettre en évidence des facteurs potentiellement responsables de notre difficulté à cloner entièrement le chromosome de *B. subtilis*. Finalement, nous avons développé une nouvelle méthode de clonage pour aller au-devant des difficultés rencontrées. Le chapitre 3 présente le développement de la méthode de CReasPy-Fusion, stratégie versatile basée sur le système CRISPR-Cas9 et utilisée pour simultanément cloner et modifier des génomes bactériens. Cette nouvelle méthode permet également de cloner des fragments de génome et pourrait être une aide cruciale pour aller vers le clonage et l'ingénierie du génome de *B. subtilis*.

ÉTUDE BIBLIOGRAPHIQUE

ÉTUDE BIBLIOGRAPHIQUE

Mon projet de thèse avait pour objectif d'appliquer des méthodologies de biologie de synthèse à *Bacillus subtilis* pour permettre le développement de souches châssis dans l'industrie biotechnologique. Cette étude bibliographique se décline donc en trois parties : (i) *B. subtilis* en tant qu'organisme modèle et en tant qu'usine de production de molécules d'intérêt ; (ii) le domaine de la biologie de synthèse et le développement de châssis microbiens et ; (iii) les méthodologies de modification de génomes en utilisant la levure *Saccharomyces cerevisiae* comme plateforme d'ingénierie.

I. *Bacillus subtilis* : de l'organisme modèle aux applications biotechnologiques

Bacillus subtilis est une des bactéries les plus étudiées, et est considérée comme le modèle principal des bactéries Gram+, à l'instar d'*Escherichia coli* pour les Gram-. Elle appartient au Phylum des *Firmicutes*, qui est lui-même divisé en trois classes : les *Clostridia* (regroupant des bactéries du genre *Clostridium*), les *Mollicutes* (genres *Mycoplasma*, *Spiroplasma* et *Phytoplasma*) et les *Bacilli* (genres *Bacillus*, *Listeria*, *Staphylococcus* et *Streptococcus*)¹. Au sein de cette dernière classe, la famille des *Bacillaceae* inclut une très grande diversité de bactéries avec pas moins de 62 genres répertoriés à ce jour². Le genre *Bacillus* est l'un des plus connus et regroupe une très grande diversité de bacilles Gram+, dont la paroi est constituée d'une épaisse couche de peptidoglycane ne laissant pas passer l'éthanol et qui se colore en violet lors de la technique de coloration de Gram³. Les bactéries de ce genre ont aussi la particularité d'être capables de produire des endospores ; structures qui se forment au sein de leur cytoplasme, et permettant à ces dernières de survivre lorsque les conditions environnementales deviennent défavorables (phénomène de sporulation)⁴. Ce genre bactérien regroupe des espèces qui ont fait très tôt l'objet d'une attention particulière, du fait de leur étude dans le domaine de la médecine (e.g. *Bacillus anthracis* et *Bacillus cereus* pathogènes pour l'humain^{5,6}), ou de leur utilisation dans les secteurs biotechnologique et pharmaceutique (e.g. *Bacillus licheniformis* et *B. subtilis* production d'enzymes⁷), et dans l'agriculture (e.g. *Bacillus thuringiensis* production de biopesticide⁸).

Originellement identifié en tant que *Vibrio subtilis* en 1835 par Ehrenberg⁴, c'est en 1872 que *Bacillus subtilis* est renommé en tant que telle par Cohn qui l'associe à la création du genre *Bacillus*. La capacité de *B. subtilis* à produire des endospores en fait rapidement un modèle d'étude de la sporulation. Ce processus soulèvera plus tard des questions de recherche notamment sur la physiologie, la division et la différenciation cellulaire^{9,10}. Au sein de l'espèce *B. subtilis*, une souche est particulièrement étudiée et utilisée au sein des groupes de recherche académique : la souche type 168. Historiquement, elle dérive de la souche « Marburg » isolée en 1930 de la ville éponyme Marburg¹¹ et a été produite au cours d'expériences de mutagenèse aléatoire par Burkholder et Giles de l'université de Yale¹², elle est auxotrophe pour le tryptophane. Une particularité de cette souche 168 « mutante » est qu'elle est capable d'intégrer de l'ADN exogène beaucoup plus facilement que la souche mère, ou bien, que les autres souches parentes obtenues au cours de ces expériences. C'est en 1958 et 1961 qu'Anagnostopoulos et Spizizen exploitent pleinement cette capacité à travers leurs travaux de transformation d'ADN^{13,14}. Ils confirmèrent la place de *B. subtilis* souche 168 comme modèle d'étude car il était désormais possible de la modifier génétiquement. C'est aussi la première bactérie Gram+ dont la séquence du génome a été rendue disponible en 1997¹⁵. Très rapidement l'avènement de nombreuses technologies, ainsi que les efforts de recherche sur ce modèle ont accru les données « omiques » (génomiques, transcriptomiques et protéomiques)¹⁶⁻¹⁹ disponibles et ont notamment facilité son utilisation dans le domaine des biotechnologies en tant que bactérie productrice de molécules d'intérêt²⁰.

I.1. Caractéristiques générales

B. subtilis est une bactérie Gram+ en forme de bacille dont la taille varie entre 2 et 6 µm. C'est une bactérie motile grâce à un système de flagelles recouvrant le pourtour de la cellule (ciliature péritriche). Elle croît en condition aérobie anaérobie facultative (AAF) et de manière rapide (temps de génération de 20 à 30 minutes) en milieu riche type Lysogeny Broth (LB)¹⁰. *B. subtilis* est une bactérie ubiquitaire et saprophyte, principalement retrouvée dans les sols (bactérie tellurique), dans le tractus digestif d'animaux ou des humains, ou encore, dans l'air et sur les plantes²¹. De nombreuses souches (271) sont répertoriées à ce jour permettant d'apprécier une grande diversité génétique chez ce modèle bactérien²². *B. subtilis* est non pathogène et possède le statut GRAS (Generally Recognized As Safe) apposé par la

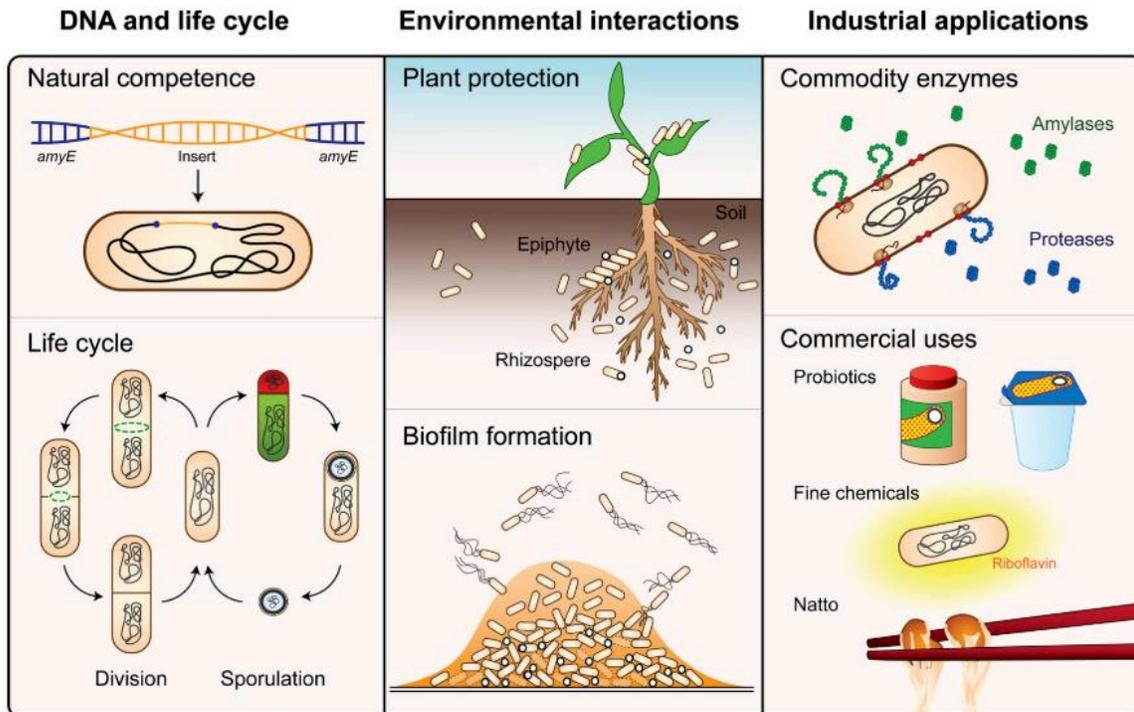


Figure 2 : Profil microbien de *Bacillus subtilis*. Cette figure est issue de Errington & van der Aart, 2020. Elle illustre graphiquement les caractéristiques générales de *B. subtilis*. *Cycle de vie et ADN* : avec la compétence naturelle pour intégrer de l'ADN exogène et avec la sporulation. *Interactions environnementales* : *B. subtilis* est une bactérie symbionte de la rhizosphère de certaines plantes et est capable de former des biofilms. *Applications industrielles* : c'est une bactérie utilisée pour produire des enzymes (amylases et protéases), elle est aussi utilisée commercialement en tant que probiotique, ou pour fermenter des aliments (natto).

Food and Drug Administration (<https://www.fda.gov/food/food-ingredients-packaging/generally-recognized-safe-gras>), ainsi que le statut QPS (Qualified Presumption of Safety) accordé par l'European Food Safety Authority (<https://www.efsa.europa.eu/fr/topics/topic/qualified-presumption-safety-qps>). Elle est également capable de fermentation, c'est-à-dire que c'est une bactérie qui peut se développer en anaérobiose en absence d'accepteur terminal d'électron à partir de glucose, de pyruvate ou d'acides aminés. Cette réaction aboutit ensuite à la production de produits de fermentation (*e.g.* acétate, éthanol, lactate)²³. Cette capacité est utilisée depuis très longtemps pour fermenter des haricots de soja, composant le traditionnel « natto » japonais¹⁰. Ces deux points ont notamment participé à rendre *B. subtilis* prépondérant parmi les bactéries capables de fermentation microbienne et utilisées pour la production industrielle d'enzymes, de vitamines, ou encore, d'antibiotiques²⁴ (**Figure 2**).

B. subtilis est majoritairement retrouvée dans le sol, elle est donc soumise (i) à une forte compétition avec d'autres micro-organismes environnementaux (champignons, autres bactéries...), et (ii) à la fluctuation importante des paramètres physico-chimiques et nutritionnels de son environnement. *B. subtilis* possède un arsenal de processus pour pallier ces contraintes. Aussi, lorsque les conditions environnementales deviennent défavorables, les cellules entrent en état de compétence pour acquérir de l'ADN exogène. Puis plus drastiquement, *B. subtilis* est capable de produire des endospores et donc de survivre dans un milieu ne permettant pas la prolifération des formes végétatives²⁵. C'est aussi une bactérie capable de sécréter des composés antimicrobiens, aussi appelés bactériocines, ce qui lui permet d'assurer sa survie parmi d'autres bactéries. On retrouve par exemple la subtilisine A, un peptide antimicrobien très bien caractérisé²⁶. Finalement, c'est une bactérie capable de former des biofilms, qui lui assure une protection lorsqu'il faut rentrer en compétition avec d'autres populations microbiennes (protection contre des enzymes de dégradation par exemple)²⁷.

Ces caractéristiques particulières ont fait l'objet de nombreuses études plaçant *B. subtilis* comme modèle biologique. Dans la suite de cette partie, les phénomènes de sporulation, compétence naturelle et production de biofilms sont repris plus en détails et des exemples d'applications biotechnologiques sont donnés.

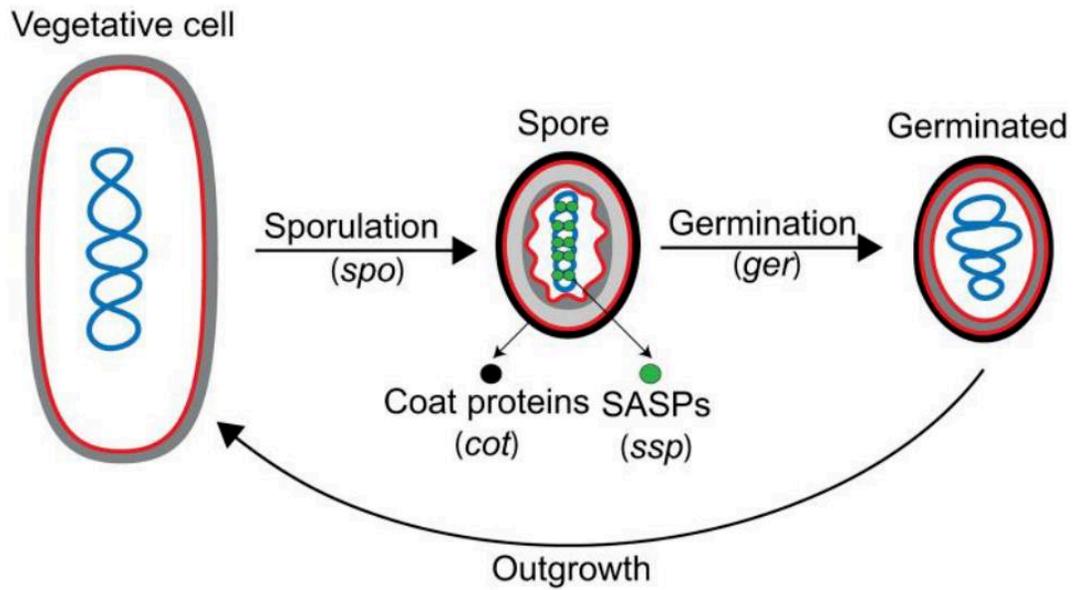


Figure 3 : Diagramme du cycle de sporulation et de germination. Cette figure est issue de Riley *et al.* 2021. Elle illustre la formation d'une spore depuis une cellule végétative via la sporulation ; puis le retour vers une cellule végétative après germination et croissance. Les *loci* impliqués dans ce processus sont également indiqués : les gènes *spo* nécessaires à la formation des spores, les gènes *ger* pour la germination des spores.

1.2. Sporulation chez *B. subtilis*

La première description de la sporulation, par l'observation de la formation d'endospores (aussi couramment appelés spores) date de 1838 avec les travaux d'Ehrenberg²⁸. Les endospores sont des structures de résistance qui se forment au sein du cytoplasme d'une cellule végétative, et qui permettent à la bactérie qui les produit de survivre à des conditions défavorables pendant une très longue durée, grâce à un état physiologique ralenti appelé « dormance »²⁵. La spore contient des acides nucléiques (ADN et ARN) et des protéines. Le cytoplasme est déshydraté et entouré d'une paroi sporale, d'un cortex (composé d'un peptidoglycane spécifique et de dipicolinate de calcium), d'une double tunique constituée de protéines fibreuses et d'une couche externe appelée exosporium²⁹.

Cette capacité de résistance a été démontrée chez *B. subtilis* et *B. anthracis*, respectivement par Cohn et Koch en 1876²⁸. Ces derniers décrivaient des cycles de sporulation puis germination (**Figure 3**) (phénomène inverse de la sporulation permettant le retour vers une cellule bactérienne active), par l'observation successive de corps végétatifs issus de spores ayant survécu à un traitement à haute température. D'autres travaux ont ensuite contribué à étoffer les connaissances sur ce phénomène, comme par exemple la description des premières étapes de la sporulation par Antoinette Ryter en 1965³⁰; mais aussi à développer une méthode de stérilisation nommée tyndallisation par son inventeur John Tyndall³¹ (chauffage discontinu toutes les 24 heures pendant 30 minutes à 60°C éliminant progressivement les formes végétatives après germination des spores).

Puisqu'il est possible de transformer *B. subtilis* avec de l'ADN et donc de mener des études génétiques, ce dernier est devenu un modèle de choix pour l'étude de la sporulation. Ces expériences ont conduit à décrire plus en détail les étapes de ce processus de différenciation cellulaire, mais aussi à identifier les gènes associés, et finalement à les classer en fonction de l'étape de sporulation dans laquelle ils sont impliqués³². D'autres travaux ont contribué à comprendre la régulation de la sporulation et à élucider la dynamique cellulaire liée au processus d'internalisation conduisant à la formation des endospores³³. Le passage des cellules d'une forme végétative vers une endospore est principalement contrôlé par le facteur de transcription Spo0A³⁴. La base de données nommée SporeWeb (<http://sporeweb.molgenrug.nl>) regroupe et détaille les différentes étapes impliquées dans la formation de spores aidant à comprendre la dynamique de ce processus³⁵.

De ces recherches fondamentales d'autres aspects biologiques ont pu être explorés. Des analyses génétiques ont montré que différents facteurs étaient activés au moment de la sporulation. La description de la régulation de ce processus a permis d'identifier des facteurs sigma propres à la cellule végétative (*e.g.* σE ou σF) ou à la spore (*e.g.* σK). L'identification de ces nouveaux facteurs sigma associés à l'ARN polymérase a contribué à étoffer les connaissances en génétique et en transcriptomique chez *B. subtilis*.³⁶

Grâce au statut GRAS (donc non pathogène) de *B. subtilis*, la sporulation présente un intérêt en terme d'applications dans la mesure où les spores peuvent servir en tant que probiotiques³⁷, ou bien, de « vésicule cargo » à visée vaccinale³⁸. L'avantage principal des spores est qu'elles sont capables de résister à l'acidité du tractus digestif, et ainsi de véhiculer les antigènes d'intérêt jusqu'aux intestins, interface permettant des échanges privilégiés entre le système immunitaire et la « lumière » du milieu extracellulaire³⁹. Dès le début des années 2000 des publications font état de l'intérêt d'utiliser ces structures pour favoriser l'expression de protéines recombinantes (et donc d'antigènes) à leur surface⁴⁰, ou encore d'utiliser les spores ainsi modifiées en tant que vaccin ou d'adjuvant à la vaccination⁴¹. A titre d'exemple, des expériences ont été menées chez des souris immunisées avec des spores de *B. subtilis* exprimant la sous-unité B de la toxine labile produite par des souches entérotoxiques d'*E. coli*. Une hausse de la production d'anticorps dirigés contre cet antigène a été observée et a permis de renforcer la protection immunitaire⁴². Plus récemment, une autre étude a souligné les avantages d'une vaccination orale, à moindre coût et de manière moins risquée qu'avec d'autres vaccins (*e.g.* vaccins vivants atténués), en utilisant des spores exposant des antigènes de *B. anthracis*, permettant d'obtenir la réponse immunitaire associée chez la souris⁴³. Pour finir, et en lien avec l'actualité, une étude a démontré qu'il était possible d'induire une réponse immunitaire par l'expression à la surface de spores, de la protéine spike du SARS-Cov2 agent viral de la COVID-19. Que ce soit sur des modèles murins, mais aussi après évaluation chez des volontaires humains, ces essais affichent des résultats prometteurs quant au développement de vaccins oraux efficaces⁴⁴.

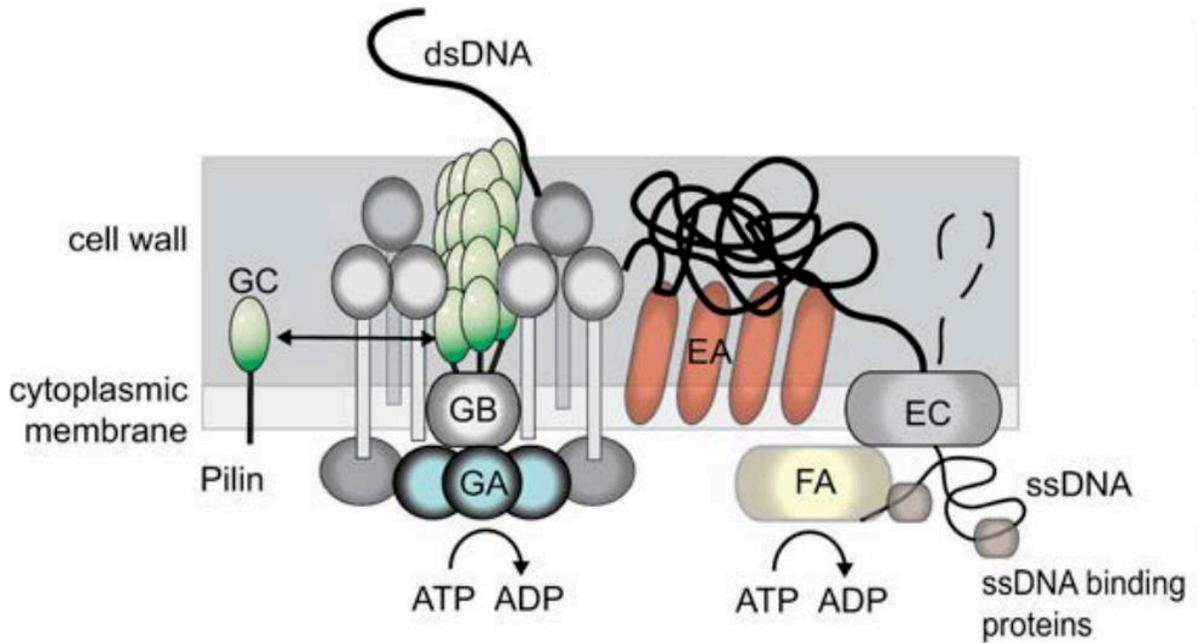


Figure 4 : Modèle de la compétence naturelle. Cette figure est issue de Maier, 2020. Elle illustre la machinerie et les protéines impliquées dans le processus d'import de l'ADN exogène par compétence naturelle. Un pseudo-pilus lie l'ADN du côté extracellulaire, il se compose des protéines ComGC, ComGB et ComGA. ComEA permet de compacter l'ADN (protection contre des DNases) et est impliquée dans l'acheminement de la molécule vers le complexe ComEC. Ce complexe est intégré à la membrane cytoplasmique et, est essentiel pour l'entrée de l'ADN dans le cytoplasme. Il est associé à ComFA qui désassocie les deux brins d'ADN favorisant l'entrée de l'ADN sous forme simple brin (le second étant dégradé). Une fois dans la cellule l'ADN est pris en charge par des protéines chaperonnes (e.g. protéines de liaison à l'ADN simple brin).

1.3. Compétence naturelle

Le terme « compétence » décrit un phénomène qui permet à la bactérie de capter et d'importer de l'ADN exogène depuis l'environnement. Cette compétence peut être naturelle, ou conférée artificiellement par le biais de traitements physico-chimiques des cellules. Si la compétence naturelle est bien connue chez *B. subtilis*, d'autres bactéries sont capables d'entrer dans cet état, comme *Streptococcus pneumoniae*, *Neisseria gonorrhoeae* et *Haemophilus influenzae*⁴⁵. La compétence peut être induite par reconnaissance de signaux spécifiques (motifs d'ADN courts), ou via un effet de quorum sensing (phase de croissance, détection de chitine à la surface de la cellule, composition du milieu extérieur qui varie). Différents mécanismes ont été décrits en fonction du type de paroi (Gram+ ou Gram-) et des machineries d'import de l'ADN associées (pilus ou pseudo-pilus)⁴⁶.

Chez *B. subtilis*, la compétence naturelle est dépendante de la phase de croissance, avec une fenêtre de temps entre la fin de la phase exponentielle et le début de la phase stationnaire propice à l'acquisition d'ADN exogène. C'est notamment grâce à l'identification de cette fenêtre, qu'Anagnostopoulos et Spizizen ont développé un protocole de transformation utilisé aujourd'hui, par l'ensemble de la communauté scientifique travaillant avec *B. subtilis*¹⁴.

Les protéines impliquées dans ce mécanisme, ainsi que son réseau de régulation génétique sont bien connus⁴⁷. Le modèle associé implique plusieurs protéines « Com », retrouvées dans les grandes étapes du processus que sont (**Figure 4**) : (i) l'amarrage et l'acquisition de l'ADN, (ii) le passage de l'ADN à travers la paroi grâce à une structure formant un pseudo-pilus, (iii) le « désappariement » de la structure double brin par une hélicase facilitant ensuite, (iv) le passage de l'ADN sous forme simple brin à travers la membrane plasmique. Ainsi, le pseudo-pilus composé des protéines ComGC est enchâssé dans la paroi de peptidoglycane et ancré à la membrane plasmique par la base grâce à la structure ComGB (protéines trans-membranaires). Cette structure basale est associée à l'ATPase cytoplasmique ComGA, définie comme un « moteur » qui aurait un rôle dans l'assemblage des sous-unités du pilus, mais aussi, dans le mécanisme d'entrée de l'ADN à travers ce dernier. L'ADN est capté de manière non spécifique à la surface de la cellule via cette structure protéique⁴⁸, puis est acheminé à travers l'épaisse couche de peptidoglycane formant la paroi de *B. subtilis*. La protéine ComEA attachée à la membrane cytoplasmique compacte et stabilise cet ADN,

offrant une protection contre d'éventuelles DNases. L'ADN est ensuite clivé en morceaux par l'endonucléase NucA également enchâssée dans la membrane. La molécule est prise en charge par une succession de protéines ComEA pour permettre sa diffusion dans le petit espace situé entre la paroi et la membrane plasmique, aussi appelé périplasma. Cet acheminement se poursuit vers la protéine « hélicase-like » ComEC⁴⁹, elle-même associée à l'ATPase ComFA qui fait office de moteur, assurant l'entrée de la molécule d'ADN sous forme simple brin (ADNsb). Une fois l'ADN introduit dans la cellule, il est soit intégré dans le chromosome de *B. subtilis* via la protéine RecA, soit pris en charge par une multitude de protéines chaperonnes pour permettre par exemple la circularisation des plasmides réplikatifs^{46,47,50,51}. La régulation de ce mécanisme est complexe et fait entrer en jeu plusieurs facteurs et protéines. Il est cependant admis que le facteur ComK est central dans cette régulation. Ce dernier est capable de s'autoréguler en se liant à son propre promoteur. Il est aussi régulé via d'autres signaux transcriptionnels⁵⁰.

La compétence naturelle aurait des rôles dans (i) le transfert horizontal de gènes permettant de capter des gènes de résistance ou tout autre *locus* pouvant conférer un avantage à *B. subtilis* ; (ii) la nutrition via le catabolisme des carbohydrates et ; (iii) le recyclage de fragments d'ADN pour faciliter la réparation de gènes par recombinaison homologue⁵².

1.4. Biofilms

L'étude de la formation de biofilms chez *B. subtilis*, a également contribué à comprendre le processus de morphogénèse, mais aussi plus fondamentalement à cumuler des données de physiologie cellulaire et de biochimie¹⁰. La communauté scientifique a notamment cherché à définir quels étaient les mécanismes impliqués dans la production de biofilm, obtenu après transition subséquente des bactéries d'une forme motile planctonique vers un état inerte²⁷.

L'opéron *tapA* composé des gènes *tapA*, *sipW*, et *tasA*, est responsable chez *B. subtilis* de la formation de biofilms. La protéine TapA est une protéine sécrétée qui participe à l'assemblage et à l'ancrage de TasA. Cette seconde protéine est essentielle dans l'intégrité de la structure du biofilm, et forme des fibres peptidiques composant une matrice⁵³. Enfin, SipW est une signal peptidase impliquée dans la modification post-transcriptionnelle de TapA et de

TasA⁵⁴. D'autres éléments entrent dans la composition des biofilms chez *B. subtilis*, comme la protéine BslA qui permet de consolider la couche de surface, des polysaccharides appelés EPS (exopolysaccharides synthétisés par des enzymes codées par les gènes *epsA* à *epsO*) participant à complexifier la structure, à la rétention d'eau, ou encore à la formation d'une pellicule. De l'ADN extracellulaire ou eDNA est également retrouvé, la présence de cet ADN est cependant peu comprise à ce jour mais reste a priori requise dans les étapes primaires de formation du biofilm. Le processus d'expression de ces composants est graduel dans le temps. La formation de biofilm est initiée par la phosphorylation de Spo0A via le récepteur histidine kinase KinABCD, permettant ensuite d'entraîner une succession de voies de régulation intracellulaires. D'autres facteurs sont impliqués dans la régulation de ce mécanisme, comme les protéines RicA, RicF et RicT, ou encore, le facteur sigma B (σ_B), activé en conditions de stress environnementaux (chaleur, concentration élevée en sels minéraux, faible ressource en nutriments)²⁷.

On peut attribuer à certains composants du biofilm des avantages pratiques lorsque l'on détourne ces éléments à des fins biotechnologiques. Ainsi, la protéine BslA est utilisée pour stabiliser des formulations en émulsion pour des applications biomédicales⁵⁵ ou pour entrer dans la composition de glaces⁵⁶. Plus récemment, une étude fait état de l'utilisation de *B. subtilis* et de sa capacité à former des biofilms, pour entrer dans la composition de matériel biocomposite ou ELM (Engineered Living Materials). Ce genre de composé est un hybride entre des matériaux amorphes (ex : silice), et des cellules bactériennes vivantes. L'idée de ce type d'applications est de proposer des alternatives aux matériaux traditionnels, pour limiter leurs coûts, mais aussi leur impact sur l'environnement⁵⁷.

Pour finir, le rôle des biofilms comme interface synergique entre une collectivité microbienne et un hôte (par exemple une plante) est largement décrit. Ainsi, des études confèrent le statut d'agent de biocontrôle dans l'agriculture à *B. subtilis*, permettant à la fois d'assurer une protection à la plante et de favoriser sa croissance. Associé à la rhizosphère de nombreuses plantes, il assure un gain nutritionnel en produisant par exemple de l'acide indole acétique (ayant des propriétés d'hormones de croissance pour les plantes), et prévient également la colonisation de ces dernières par d'autres microorganismes pathogènes⁵⁸⁻⁶⁰.

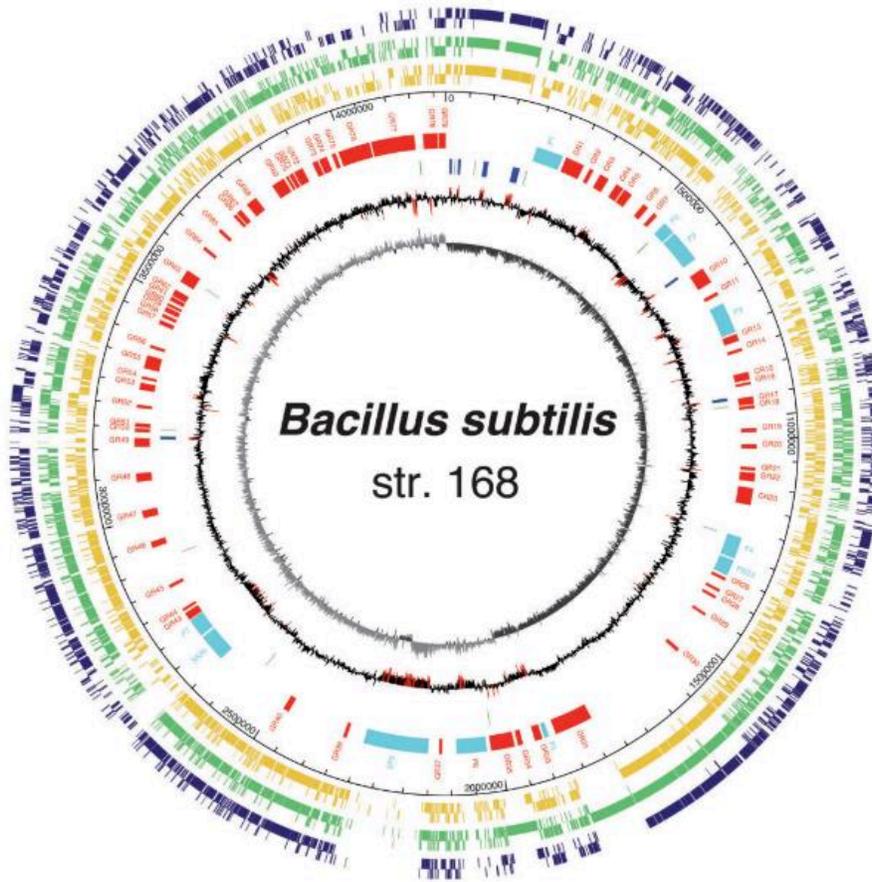


Figure 5 : Représentation de la séquence du génome de *Bacillus subtilis* 168. Cette figure est issue de Barbe *et al.* 2009 et illustre la carte génomique après reséquençage de la souche type. Du centre vers les bords extérieurs sont indiqués : le biais en GC (GC skew), la déviation en GC, les ARNt (en vert) et ADNr (en bleu). Les rectangles bleu clair correspondent aux régions d'origine phagique. Les trois cercles externes en jaune, vert et bleu correspondent aux gènes orthologues présumés retrouvés chez d'autres espèces de *Bacillus*.

I.5. Caractéristiques génétiques et systématiques

I.5.1. Séquençage et annotations du génome de *B. subtilis*

Le génome de *B. subtilis* fut le premier génome d'une bactérie Gram+ à être séquencé en 1997 grâce à la mobilisation de plus de 30 laboratoires européens et japonais regroupés sous la forme d'un consortium¹⁵. Cette première séquence annotée et publiée dans la revue « Nature » est issue d'un effort collectif colossal qui avait pour but de tirer le meilleur parti des informations expérimentales pouvant être associées à la séquence. Ce travail fait état d'un chromosome dont la taille est de 4 214 630 pb avec plus de 4100 séquences codantes (Coding DNA Sequence ou CDS) et un pourcentage en GC de 43,5%. Il a permis d'obtenir une vision plus claire du contenu en gènes (structuraux et régulateurs), de l'organisation de ces gènes autour du génome (synténie), mais également d'assigner une fonction putative à de nombreux gènes. Ces données ont ensuite permis d'initier les premières cartes métaboliques et de mieux appréhender les bases moléculaires du quorum sensing. Elles ont aussi facilité l'identification des voies de signalisation et des systèmes de sécrétion des protéines. Malgré cette percée, ce travail qui émanait de groupes avec des expertises très diverses (et à une époque où le séquençage des génomes était un travail extrêmement difficile) avait un inconvénient majeur : les données étaient très hétérogènes.

A partir de 2005, des techniques de séquençage plus efficaces, plus rapides, plus précises voient le jour. Parallèlement le nombre de séquences dans les banques de référence augmentent fortement et des plateformes d'annotation de haut niveau se développent. Ces avancements permettent d'avoir accès à des séquences plus fiables. Le contexte devenait donc favorable à un travail de reséquençage et réannotation du génome de *B. subtilis* qui a eu lieu en 2009 et a été réalisé cette fois par une seule équipe¹⁸ (**Figure 5**). A l'issue de ces travaux, les auteurs ont ainsi pu confirmer ou invalider certaines portions du génome, révélant notamment des erreurs de séquençage dans la séquence de base. Ils ont également redéfini les groupes de gènes appartenant au « Paleome », c'est à dire les gènes impliqués dans la maintenance, la réplication, la réparation de la cellule. ; mais aussi ceux appartenant au « Cénome » responsables des caractéristiques bactériennes de *B. subtilis*, liant la bactérie à sa niche écologique.

Une dernière annotation a été effectuée en 2018, soit 20 ans après celle de Kunst *et al.*⁶¹. Comme le génome avait été entièrement re-séquencé par des méthodes NGS en 2009, on pouvait s'attendre à ce que (sauf mutations inévitables apparaissant lors de la propagation en laboratoire⁶²), cette séquence corresponde à une séquence exacte, qui n'a pas besoin d'être à nouveau séquencée. En revanche, les annotations de séquences changent inévitablement, car d'une part, l'identification de la fonction des gènes évolue presque quotidiennement et d'autre part, certains éléments génomiques, en particulier des ARN régulateurs non traduits, sont régulièrement découverts. A travers cette étude, Borriss *et al.* ont assigné une fonction putative à des gènes dont les fonctions étaient à l'origine inconnues (plus de 2000 à fonction inconnue sur les 4100 au départ) et ont corrigé certaines erreurs d'annotations artéfactuelles. Plus encore, ils ont affiné l'image globale du génome de *B. subtilis* et de la fonction de ses gènes au sein de sa niche écologique. En effet, les auteurs ont cumulé les données obtenues par comparaison de 36 génomes issus de diverses souches de *B. subtilis*, bien éloignées de la souche 168 domestiquée en laboratoire. Cette nouvelle session de ré-annotations basée sur une nouvelle diversité génomique est cruciale, et souligne l'importance de veiller à mettre à jour les données recueillies pour un modèle bactérien tel que *B. subtilis*. De plus, cette étude met en lumière l'importance de ne pas seulement définir un micro-organisme d'après l'identité d'une seule souche, mais de garder un point de vue global sur les différentes populations/souches ; ceci dans le but de mieux définir la biologie de ce dernier et de mieux tracer son histoire évolutive. Enfin, la comparaison de génomes peut permettre dans les domaines de la BS et des biotechnologies, de mettre en évidence des régions hautement variables (non essentielles) dont il peut être souhaitable de se débarrasser pour construire des châssis bactériens stables.

Pendant plusieurs années la base de données de référence du génome était SubtiList¹⁵, qui permettait d'avoir accès à la séquence du génome de *B. subtilis*, mais aussi aux annotations des gènes⁶³. Depuis 2009, une base de données encore plus fournie nommée SubtiWiki a été créée par l'équipe de Jörg Stülke^{64,65}. Elle regroupe, entre autres, une collection de données de la littérature (dont les références sont reliées à Pubmed), la séquence du génome, mais aussi, des représentations dynamiques des voies de régulation ou des voies métaboliques.

I.5.2. Autres données « omiques »

En partant de la lecture du génome et grâce à l'avènement des technologies à haut débit, d'autres études « omiques » englobant la protéomique, la transcriptomique et la métabolomique, ont pu voir le jour. Cette analyse à grande échelle appelée aussi biologie des systèmes ou intégrative a permis de dessiner une vue globale de la biologie de *B. subtilis*²⁰.

Ainsi, l'étude extensive du protéome a permis de définir les mécanismes guidant par exemple la compétition entre *B. subtilis* et d'autres micro-organismes dans un même environnement, soit le sol, ou plus particulièrement, la rhizosphère comme nous l'avons vu précédemment (§1.2.2.3.)²¹. Ce répertoire permet également de mieux comprendre le rôle des protéines impliquées dans les processus de fermentation, de sécrétion d'autres protéines, crucial pour améliorer les bases de la production industrielle et les applications commerciales qui en découlent⁶⁶.

De telles données ne seraient pas complètes sans une analyse fine de l'architecture transcriptionnelle de *B. subtilis*. Ainsi, un programme européen nommé BaSysBio (*Bacillus System Biology*) a permis de récolter des informations systématiques et quantitatives sur la transcription chez ce modèle bactérien. L'ensemble de ces données est disponible sur le site web associé (<http://genome.jouy.inra.fr/basysbio/bsubtranscriptome>)¹⁷, mais aussi sur la base de données SubtiWiki citée plus haut.

Finalement, des études à grande échelle de modélisation ont également permis de compléter l'image de la dynamique générale du métabolisme de *B. subtilis*^{67,68}. Il est à noter que la base de donnée MetaCys répertorie les voies métaboliques et les enzymes identifiées chez ce modèle⁶⁹. De même, la très complète plateforme Subtiwiki regroupe également ces données et les réseaux métaboliques associés.

I.6. Usine bactérienne de production à visée biotechnologique

Nous avons vu que *B. subtilis* est un modèle bactérien de référence et d'étude en microbiologie depuis environ 150 ans. Grâce à cela, et aussi parce qu'il possède des aptitudes se prêtant à des applications en industrie biotechnologique, il est devenu un modèle incontournable de production de molécules d'intérêt en pharmaceutique, pour l'industrie

(e.g. agroalimentaire) ou pour l'environnement (e.g. biocarburants, traitement des eaux). Ses capacités à croître rapidement (temps de génération court), à faire de la fermentation ou à sécréter des protéines, en ont fait un micro-organisme tout à fait approprié pour la production industrielle de protéines²⁴.

Dans la suite de cette partie nous nous focaliserons sur ce qui permet notamment son utilisation en tant que plateforme microbienne dans l'industrie biotechnologique à savoir, (i) les étapes d'ingénierie génétique (protocoles de transformation, outils génétiques disponibles), (ii) d'expression de gènes, et (iii) de sécrétion de protéines. Nous verrons les avantages qu'offre *B. subtilis* par rapport à d'autres micro-organismes (e.g. *E. coli*) et quelques exemples d'application biotechnologiques.

1.6.1. Ingénierie génétique chez *Bacillus subtilis*

1.6.1.1. Transformations

Plusieurs protocoles existent pour transformer *B. subtilis* avec des molécules d'ADN (plasmides ou ADNg) : la transformation par compétence naturelle, la transformation de protoplastes et la transformation par électroporation⁷⁰.

De par sa simplicité, la technique de transformation par compétence naturelle est la plus utilisée au sein de la communauté scientifique. Son principe repose sur la capacité des cellules de *B. subtilis* à entrer dans un état de compétence naturelle (cf. §1.3.). Les travaux pionniers conduits par Spizizen en 1958¹³ décrivaient la transformation des cellules bactériennes avec de l'ADN. Au cours de ces premiers essais, l'auteur montre (i) que le nombre de colonies augmente lorsque l'on augmente la quantité d'ADN, (ii) que la pureté de cet ADN joue sur l'efficacité de la transformation, et (iii) que l'ajout d'extrait de levure dans le milieu minimum améliore la transformation. En association avec Anagnostopoulos, des travaux supplémentaires sont alors effectués et décrivent plus amplement le protocole et notamment la phase de croissance à atteindre pour que les cellules soient compétentes. Ils déterminent en effet que l'état de compétence est associé à la fin de la phase exponentielle. Ils définissent la composition des deux milieux minimums désormais appelés MG1 et MG2 (cf. Matériel et Méthodes) et démontrent que la culture de *B. subtilis* doit se faire en deux temps : pendant 4h dans le premier milieu, puis pendant 1h30 dans le second milieu moins riche mimant des

conditions environnementales appauvries en nutriments. L'état de compétence est alors atteint et les cellules sont transformées avec de l'ADN¹⁴.

Si ce procédé s'avère être particulièrement aisé à utiliser pour la souche type *B. subtilis* 168, ce n'est pas toujours le cas pour d'autres souches. Certaines études ont donc tenté de pallier ces difficultés en jouant sur l'expression du facteur de transcription central de la compétence ComK. Nijland *et al.*, décrivent par exemple l'induction transitoire de ce facteur à l'aide d'un plasmide thermostable⁷¹. La compétence de la souche transformée est alors suractivée via un promoteur inductible à l'IPTG (PSpac) en amont du gène *comK*. La fréquence de transformation des souches environnementales est alors nettement améliorée. D'autres auteurs ont fait le choix de travailler le protocole plus en détail pour une souche donnée, comme ce fût le cas pour la souche DB104. Les auteurs ont notamment optimisé les temps d'incubation en milieu SM1 et SM2 (correspondant à MG1 et MG2) et ont aussi ajouté une étape supplémentaire de régénération avant étalement et sélection⁷².

La transformation de protoplastes est plus longue à mettre en place, mais reste cependant utile pour transformer des souches non-domestiquées de *B. subtilis*, ou moins compétentes⁷⁰. Initialement, ce protocole consistait à transformer des protoplastes de *B. subtilis* – dont la paroi était digérée au préalable avec du lysozyme – soit avec des lysats de cellules « L-form » (en français « bactérie de forme L »), soit des protoplastes, ou soit avec de l'ADN purifié⁷³. Il est important de bien noter la différence entre ces deux types cellulaires. Les bactéries de forme L sont originellement des mutants spontanés obtenus par traitement avec du NTG (N-méthyl-N-nitro-N-nitrosoguanidine) combiné au lysozyme⁷⁴. Ce sont des variants capables de vivre dans un état dépourvu de paroi aussi nommé la « L-phase ». Cet état pouvant être quelquefois réversible, a fait l'objet d'études approfondies pour comprendre son mécanisme, et pour étudier expérimentalement la synthèse de la paroi et le processus de division cellulaire⁷⁵. Tandis que, les protoplastes sont des cellules dont la paroi a été progressivement digérée par le lysozyme dans un milieu hypertonique prévenant leur lyse. À la différence des bactéries de forme L aucun agent mutagène n'est alors utilisé pour produire ces cellules dépourvues de paroi⁷⁶. Dans les années 1970 et 1980, le protocole a évolué et certains groupes ont inclus l'utilisation de polyéthylène glycol (PEG) pour favoriser la fusion de protoplastes. En effet, des études montrent que cette méthode permet la fusion de protoplastes⁷⁷, la transformation d'ADN plasmidique^{78,79}, mais aussi d'ADN

chromosomique⁸⁰, appliquée à l'espèce *B. subtilis*, ou à d'autres espèces comme *B. licheniformis* ou encore *Bacillus amyloquefaciens*. Très brièvement, les cellules sont cultivées jusqu'à atteindre la moitié de la phase exponentielle, puis traitées dans un milieu SMM (Sucrose, Maleate, Magnesium) additionné de lysozyme. Elles sont ensuite transformées par ajout de PEG avec l'ADN désiré, puis étalées sur un milieu de régénération assurant une protection osmotique. C'est un processus plus long et plus lourd à mettre en place que le protocole basé sur la compétence naturelle. Très récemment, une étude basée sur la fusion entre des protoplastes de différentes espèces de *Bacillus* a permis d'évaluer la dynamique de recombinaison entre différents génomes⁸¹. Ce même protocole a été utilisé pour transformer plusieurs plasmides simultanément dans *B. subtilis*. Plus exactement, ce travail a consisté à fusionner des liposomes (vésicules lipidiques artificielles et vides⁸²) au sein desquels trois plasmides à transformer ont été encapsulés. Ces derniers ont ensuite été introduits dans *B. subtilis* par fusion « liposomes-protoplastes » en présence de PEG 36%⁸³.

Quelques travaux font finalement état qu'il est possible de transformer *B. subtilis* par électroporation mais avec une plus faible efficacité⁸⁴. L'association d'une solution d'électroporation hyper-osmotique (ajout de sorbitol et de mannitol), avec un champ électrique appliqué pour déstabiliser la paroi, a permis d'obtenir des transformants de *B. subtilis*⁸⁵. Ce protocole a également été utilisé pour transformer des souches non domestiquées de *B. subtilis* par électroporation de protoplastes⁸⁶. Cette méthode de transformation n'est cependant pas privilégiée par la communauté scientifique travaillant avec *B. subtilis*.

1.6.1.2. Outils génétiques

- Mutagenèse aléatoire :

Les premiers essais de mutagenèse aléatoire décrits en 1947 par Burkholder et Giles, ont été effectués par irradiation des cellules de *B. subtilis* avec des UV ou des rayons X¹². Cela a permis la production des premiers mutants de *B. subtilis*, dont la souche 168 auxotrophe pour le tryptophane est issue.

Des systèmes d'édition aléatoire par transposition sont également disponibles chez *B. subtilis*. Les transposons sont des éléments génétiques mobiles qui s'intègrent

aléatoirement dans le génome grâce à des séquences IS (séquences d'insertion présentes aux extrémités) et à l'action d'une transposase. Ainsi le transposon, associé à une cassette contenant un marqueur de résistance, peut permettre d'inactiver des gènes par disruption de la séquence codante ou des séquences de régulation. Cette méthode permet de construire, en quelques expériences, de larges banques de mutants à partir desquelles des mutants d'intérêt peuvent être rapidement identifiés⁸⁷. Plusieurs transposons dérivés de bactéries Gram+ se sont révélés fonctionnels chez *B. subtilis*. Parmi ceux-ci, on trouve les transposons Tn916⁸⁸ et Tn1545⁸⁹, qui sont des transposons conjugatifs, ainsi que le Tn917 (et dérivé Tn917ac1), qui sont des transposons de type Tn3 largement utilisés pour la mutagenèse et le clonage chez cet hôte⁹⁰⁻⁹². Des transposons issus de bactéries Gram- ont également été introduits avec succès chez *B. subtilis*, à l'instar du Tn10, un des transposons les plus connus d'*E. coli*⁹³. La transposition du Tn10 et de ses dérivés chez *B. subtilis* se produit de façon aléatoire et avec une fréquence plus élevée que le Tn917 lorsque le gène de la transposase est fusionné à des signaux d'expression appropriés. Les meilleurs résultats ont été obtenus avec le promoteur tac d'*E. coli* et un RBS de *B. subtilis*^{94,95}. Ces deux familles de transposons ont été utilisées pour divers projets, néanmoins, elles montrent des limites. L'insertion régulière du Tn917 à des *locus* préférentiels (« hotspot ») réduit la diversité de la banque⁹⁶ et implique de devoir cribler un grand nombre de clones. Contrairement au Tn917, Tn10 ne semble pas avoir de sites d'insertion préférés dans le chromosome de *B. subtilis* ; cependant, il requiert une séquence particulière de 6 pb comme élément cible⁹⁷. Cela réduit le nombre de sites d'insertion potentiels de Tn10 sur le chromosome de *B. subtilis* et, par conséquent, l'efficacité de Tn10 comme outil de mutagenèse aléatoire.

En 2006, des transposons (TnYLB) dérivant de l'élément transposable marinier Himar1 (isolé à l'origine de la mouche *Haematobia irritans*) ont été adaptés à *B. subtilis*. Ces transposons s'insèrent au hasard dans le génome de *B. subtilis* (pas de « hotspot » d'insertion) et avec une fréquence (10^{-2}) qui est significativement plus élevée que celle rapportée pour les transposons Tn917 et Tn10 (10^{-6} et 10^{-4} , respectivement)⁹⁸. En 2012, des dérivés de TnYLB ont été conçus de façon à élargir le type de criblage pouvant être réalisé avec cet outil. Des transposons TnYLB incluant (i) différents marqueurs antibiotiques (chloramphénicol, spectinomycine, kanamycine), (ii) un promoteur fort (inductible ou pas) orienté vers

l'extérieur (surexpression de gènes adjacents) ou, (iii) un gène *lacZ* sans promoteur (pour générer des fusions transcriptionnelles) sont désormais à disposition de la communauté⁹⁹.

- Mutagenèse dirigée :

La capacité de *B. subtilis* à être transformé avec de l'ADN linéaire associée à l'efficacité de sa machinerie de recombinaison homologue font qu'il est possible d'utiliser des outils d'édition génétique pour inactiver des gènes, effectuer des insertions dans le chromosome, ou encore de produire des mutations de manière ciblée¹⁰⁰.

B. subtilis a la capacité de réaliser la recombinaison homologue par le biais de la protéine RecA, qui est impliquée dans le processus de réparation de l'ADN¹⁰¹. Cette propriété a permis de développer une stratégie d'inactivation de gènes basée sur l'insertion dans le chromosome d'une cassette de recombinaison, portant un marqueur de sélection encadré par des séquences homologues (500pb à 1kpb) au site d'insertion dans le génome. Le gène cible est alors remplacé par la cassette de recombinaison via un double crossing-over (échange allélique) ce qui conduit à la suppression totale du gène.

Chez *B. subtilis*, l'inactivation de gènes peut également se faire par recombinaison homologue via un simple crossing-over (mutagenèse par insertion-duplication). Dans ce cas, l'inactivation du gène cible est réalisée à l'aide d'un plasmide (typiquement le pBR322 ou le pSC101 à faible nombre de copies)¹⁰² qui porte un fragment interne du gène cible. L'intégration complète du plasmide dans le chromosome par recombinaison homologue et simple crossing-over aboutit à la formation de deux copies du gène cible dont l'une est tronquée en 5' et l'autre en 3'. C'est cette stratégie qui a été choisie en 2003, pour procéder à l'inactivation systématique de gènes de *B. subtilis* et estimer l'ensemble minimal de gènes requis pour maintenir la vie bactérienne dans des conditions nutritives¹⁰³. Cette approche d'inactivation gène par gène a été rendu possible grâce à un ensemble des vecteurs pMUTIN¹⁰⁴. Leurs principales propriétés sont : (i) une incapacité à se répliquer dans *B. subtilis*, ce qui permet de faire de la mutagenèse insertionnelle, (ii) la présence d'un gène rapporteur *lacZ* pour mesurer l'expression du gène cible (fusion transcriptionnelle) et (iii) la présence d'un promoteur inductible pour permettre l'expression contrôlée de gènes en aval du gène cible et se trouvant dans le même opéron que celui-ci. Cette dernière exigence a été motivée par

l'observation que la plupart des gènes du génome de *B. subtilis* sont organisés en unités multicistroniques.

D'autres groupes ont cherché à développer des méthodes alternatives basées sur des systèmes de recombinaison phagiques. En 2015, Sun *et al.*, ont décrit pour la première fois l'utilisation de la recombinase GP35 du phage natif SPP1 de *B. subtilis* pour inactiver plusieurs gènes chez cet hôte (*deoD* codant la purine nucleoside phosphorylase, *amyE* codant l'alpha-amylase et *upp* codant l'uracil phosphoribosyltransferase)¹⁰⁵. En utilisant de l'ADN simple brin (ssDNA) comme substrat de recombinaison, dix recombinases provenant de divers phages ont été comparées, GP35 était à chaque fois la plus efficace et montrait la fréquence de recombinaison la plus élevée.

Un point faible de ces méthodes est qu'il est nécessaire de travailler avec un marqueur de sélection différent pour effectuer une modification subséquente, limitant les possibilités d'ingénierie multiple du chromosome¹⁰⁰. Pour pallier ce problème et permettre le recyclage des marqueurs de sélection, d'innombrables stratégies ont été développées ; certaines sont basées sur des systèmes « opérateur-répresseur » (*cl*, *araR*), d'autres sur le métabolisme de la pyrimidine (*upp*, *pyrF*), sur l'auxotrophie (auxotrophie pour la lysine) ou encore sur de la recombinaison site-spécifique (*cre-lox*, *xer/dif*)¹⁰⁰. Plus précisément et à titre d'exemple parmi ces dernières sont retenues :

- La stratégie basée sur le gène répresseur *cl* (issu du bactériophage lambda de chez *E. coli*) et la capacité du répresseur CI à se lier au promoteur P_r du phage lambda et à supprimer son activité. En 1999, Itaya *et al.*, ont développé une méthode de contre-sélection dans laquelle l'expression du gène de résistance à la néomycine fusionné au promoteur P_r, est contrôlée par la présence/absence du répresseur CI¹⁰⁶. La perte du gène répresseur CI rend la souche hôte résistante à la néomycine.
- La stratégie basée sur le gène *upp* (codant l'UPRTase ou Uracil Phosphoribosyl Transferase). Fabret *et al.*, ont décrit le développement d'un système dans lequel le gène *upp* est fusionné avec un gène de résistance aux antibiotiques pour créer une "cassette upp". Un fragment généré par PCR, constitué du gène cible avec la mutation désirée est joint à la cassette upp, puis intégré dans le chromosome d'une souche Δupp par recombinaison homologue (sélection via le marqueur de résistance aux antibiotiques). Ensuite, l'excision de la cassette upp du chromosome se fait par

recombinaison entre de courtes séquences chromosomiques répétées (incluses dans la conception de la molécule d'ADN transformé). Les cellules ayant excisé la cassette se multiplient en présence de 5-fluorouracil, alors que celles n'ayant pas excisé la cassette meurent¹⁰⁷.

- La stratégie basée sur le système *cre-lox*. Yan *et al.*, ont quant à eux proposé un outil d'édition du génome de *B. subtilis* se basant sur un plasmide thermosensible exprimant la recombinase « CRE » issue du bactériophage P1 d'*E. coli*. Après entrée du plasmide dans les cellules, l'enzyme CRE agit sur les sites de reconnaissances *loxP*, permettant par exemple l'excision d'un marqueur de sélection intégré dans le génome¹⁰⁸.

De nombreux outils et stratégies sont disponibles pour modifier génétiquement *B. subtilis*¹⁰⁰, et les descriptions effectuées ci-avant ne sont pas exhaustives mais tendent seulement à donner une vue d'ensemble de la littérature existante. Pour terminer cette rubrique, il est important de préciser que des outils dérivés des systèmes CRISPR-Cas ont très récemment été développés chez *B. subtilis*. Véritable révolution dans le monde scientifique pour l'ingénierie de génomes (bactériens, eucaryotes ou viraux), c'est en 2016 que l'outil CRISPR-Cas9 est adapté chez *B. subtilis*¹⁰⁹. Depuis, une véritable « boîte à outils CRISPR » est désormais disponible chez ce dernier¹¹⁰. Ainsi les systèmes CRISPR-Cas peuvent être utilisés en tant qu'outils d'édition de génome (Cas9, MAD7, Cpf1, Cas9 nickase ou édition de base) mais aussi en tant qu'outils de contrôle d'expression génique (CRISPRa), ou bien, d'interférence par répression de la transcription (CRISPRi). Il est important de noter qu'aucun système CRISPR-Cas endogène n'a été décrit à ce jour chez *B. subtilis*, tous ces outils dérivent donc de systèmes d'autres bactéries, comme par exemple celui de *Streptococcus pyogenes* (SpCas9) premier à avoir été détourné par Jinek *et al.* en 2012, en tant que dispositif d'ingénierie¹¹¹.

1.6.1.3. Vecteurs et contrôle de l'expression de gènes chez Bacillus subtilis

Chez *B. subtilis*, l'expression des gènes codant des protéines recombinantes passe par la transformation de vecteur d'expression ou de cassettes d'expression linéaires qui s'intègrent au chromosome. La souche *B. subtilis* 168 ne contient pas naturellement de

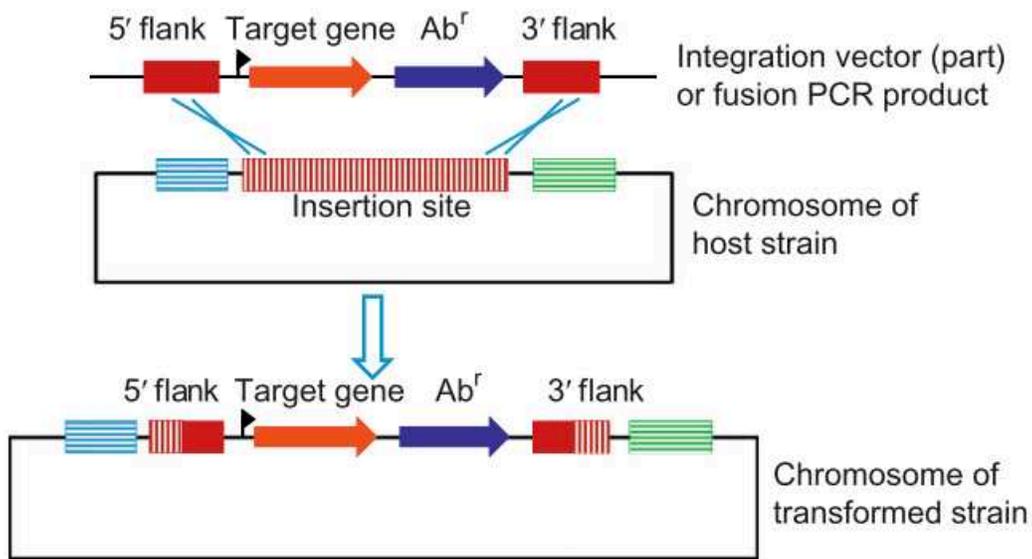


Figure 6 : Stratégie d'intégration par double crossing over dans le génome de *B. subtilis*. Cette figure est issue de Harwood *et al.* 2013. Le vecteur d'intégration possède des régions identiques en 5' et 3' du site d'intégration (en rouge). Ces *locus* encadrent une cassette composée d'un gène cible à exprimer (en orange) avec son promoteur (flèche noire) et d'un gène de résistance à un antibiotique donné (bleu). Après recombinaison homologue par double crossing-over le vecteur (*e.g.* pDR111) est intégré en lieu et place du gène ciblé pour l'intégration (*e.g.* *amyE*).

plasmides endogènes, et la plupart des plasmides présents dans les autres souches de *Bacillus* sont cryptiques. Nous pourrions par exemple mentionner le pLS20, un plasmide conjugatif de 55 kpb, isolé de *B. subtilis natto*¹¹². Ce dernier se maintient de manière stable chez cet organisme et se transfère à une grande variété d'espèces et de souches. En revanche, il est peu utile en tant que vecteur de clonage du fait de sa taille et de l'absence de marqueur de sélection⁷⁰. C'est pourquoi les premiers vecteurs plasmidiques développés pour *B. subtilis* étaient basés sur les plasmides d'autres bactéries Gram+ notamment *Staphylococcus aureus* et *Lactococcus lactis* (pUB110, pC194, pE194, et pWVO1). Ces vecteurs et leurs dérivés ont été largement utilisés⁸⁴ mais présentent des problèmes d'instabilité¹¹³, souvent aggravés par la présence de larges inserts hétérologues¹¹⁴. Cela s'explique par le mode de réplication en cercle roulant (RC) de ces plasmides, similaire à celui de phages filamenteux de bactéries Gram-¹¹⁵. Les plasmides qui utilisent un mode de réplication thêta sont généralement plus stables que les plasmides RC¹¹⁶. Une seconde génération de vecteurs dérivés des plasmides entérococciques (*e.g.* pAM β 1) a donc été développée pour *B. subtilis* comme le plasmide navette *E.coli/B. subtilis* pHV1431¹¹⁷ (fusion entre pBR322 et le pAM β 1) ou pL252¹¹⁸ ou pASM100¹¹⁹. Le clonage dans pHV1431 et pAMS100 est très efficace et les longs inserts sont généralement maintenus de manière stable.

Des problèmes d'instabilité ou d'incompatibilité avec certaines souches, ont également favorisé le développement de vecteurs intégratifs. Ce type de plasmides a notamment permis de produire des plateformes d'expression stable de protéines recombinantes, mais aussi de permettre d'intégrer des gènes rapporteurs très utiles dans des études de fonctions ou d'essentialité de gènes⁷⁰. L'intégration de ces vecteurs peut se faire, comme évoqué au cours du paragraphe précédent, par recombinaison homologue via un simple crossing-over ou un double crossing-over. Les stratégies privilégiant le simple crossing-over, utilisent typiquement des plasmides dérivant du pBR322 ou du pSC101 (vecteurs à faible nombre de copies utiles en cas de toxicité entre des *loci* de *B. subtilis* et de *E. coli*)¹⁰². Celles privilégiant le double crossing-over utilisent certains sites d'intégration préférentiels (**Figure 6**). Nous citerons, par exemple, le pDR111 construit avec des fragments 5'-3' (500pb à 1 kpb) correspondant à des *loci* situés en amont et en aval du gène *amyE* codant pour une amylase. Après transformation, le vecteur est intégré au *locus* cible via un double crossing-over, ce qui aboutit à l'élimination du gène *amyE* et à son remplacement par le marqueur de

sélection et le gène codant la protéine d'intérêt. Ce site d'intégration (*amyE*) présente un avantage car il permet de mettre en place un test phénotypique qui aide au criblage. En effet, les transformants sont cultivés sur milieu gélosé additionné d'amidon, les colonies sont ensuite colorées avec de l'eau iodée (ou lugol). Si le gène *amyE* est intact alors un halo transparent, synonyme de la dégradation de l'amidon, se forme autour des colonies ; dans le cas contraire aucun halo ne se forme, attestant que le mutant souhaité a vraisemblablement été obtenu. Cette méthode est notamment utilisée pour l'expression de gènes hétérologues ou bien pour des études de complémentation⁷⁰.

Pour assurer le bon contrôle de l'expression de gènes, des études ont été menées, d'une part, pour identifier et collecter des banques de promoteurs, et d'autre part, pour caractériser des séquences RBS (RBS en anglais pour ribosome binding site), car le niveau d'expression d'une protéine est influencé par ces deux composants (force du promoteur ou du RBS). Parmi la liste des promoteurs disponibles nous pouvons compter des promoteurs constitutifs (p43, pVeg)^{120,121}, des doubles promoteurs (augmentant la productivité), des promoteurs synthétiques, ou encore, des promoteurs inductibles tel que le PSpac inductible à l'IPTG¹²².

Récemment, une grande étude quantitative visant à déterminer le niveau de transcription associé à une séquence particulière du RBS a été menée et permet désormais de sélectionner au sein d'une large collection (comprenant promoteurs, RBS, ou marqueurs de dégradation des protéines) les éléments requis pour contrôler l'expression d'un gène en fonction de la dynamique que l'on souhaite appliquer¹²³.

En 2013, Radeck *et al.* ont répertorié et réévalué les éléments génétiques (vecteurs, promoteurs et autres composants génétiques) les mieux caractérisés chez *B. subtilis*, dans le but de créer une boîte à outils standardisée appelée « The *Bacillus* BioBrick Box »¹²⁴. En effet, des éléments génétiques standardisés et bien caractérisés, sont une condition préalable à l'assemblage facile et reproductible de nouveaux modules et dispositifs génétiques. Alors que de nombreuses pièces standardisées existaient déjà pour *E. coli*, de tels outils manquaient encore chez *B. subtilis*. L'objectif de cette étude était de pallier ce manque. Une version 2.0 a été publiée en 2017, dans le but d'enrichir la liste des éléments déjà présents et de faciliter l'ingénierie génétique chez *B. subtilis*¹²⁵. Cette dernière est disponible gratuitement et accessible sur le site du « *Bacillus* Genetics Stock Center » (<http://www.bgsc.org/>). Le BGSC a

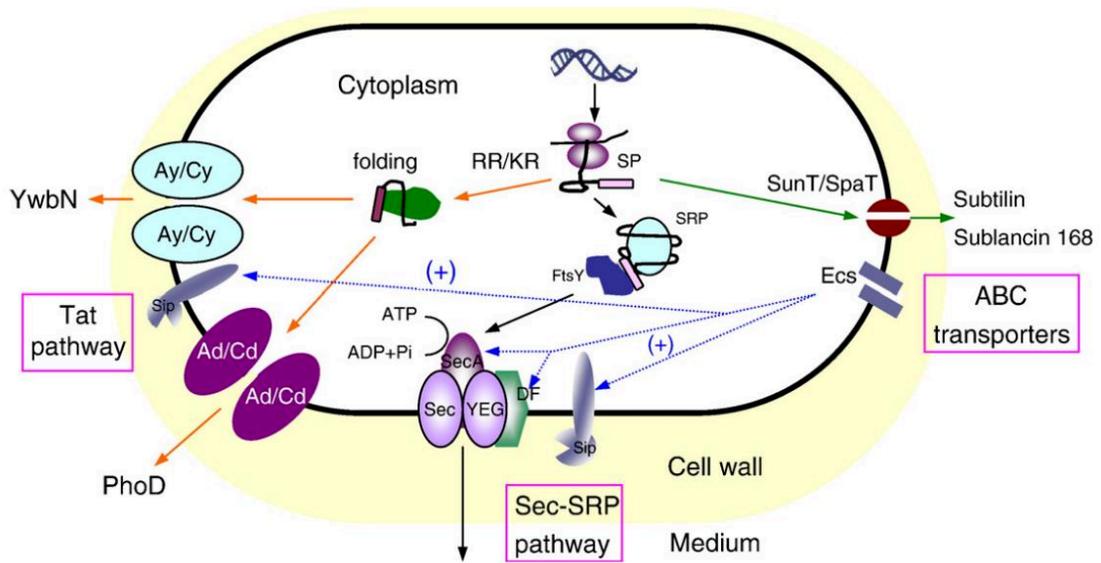


Figure 7 : Représentation schématique des voies de sécrétion de protéines chez *B. subtilis*. Cette figure est issue de Fu *et al.* 2007. Les voies de sécrétion Sec (Sec-SRP pathway), Tat (Tat pathway) et les transporteurs ABC sont représentés.

été créé en 2001 à l'initiative du Dr. Daniel Zeigler. Sa mission principale est de maintenir des souches génétiquement caractérisées, des vecteurs de clonage et des bactériophages pour le genre *Bacillus* et les organismes apparentés et de distribuer ce matériel aux scientifiques du monde entier. Concernant *B. subtilis*, pas moins de 1291 mutants dérivant de la souche 168 ainsi qu'un catalogue extrêmement bien fourni de plasmides décrits dans la littérature y sont retrouvés¹²⁶. Il est à noter que le « The Bacillus BioBrick Box » est également accessible via le registre des éléments de la biologie de synthèse (Registry of Standard Biological Parts ; http://parts.igem.org/Main_Page). Ce dernier est une collection d'éléments génétiques mise à la disposition des chercheurs œuvrant dans le domaine de la biologie de synthèse.

1.6.2. Sécrétion de protéines recombinantes

B. subtilis est un bon candidat pour la sécrétion dans le milieu extracellulaire de protéines hétérologues. Par rapport aux bactéries Gram- (*e.g. E. coli*), l'absence de membrane externe simplifie le processus de sécrétion par des voies dédiées¹²⁷, d'autant que la présence de LPS (lipopolysaccharides) peut-être toxique pour des applications pharmaceutiques²⁴. De plus, cela évite l'accumulation des protéines dans le cytoplasme et la formation de corps d'inclusion¹²⁸. Un autre avantage de la sécrétion protéique directement à l'extérieur de la cellule, et que cela permet d'éviter les coûts associés à la purification (lyse des bactéries, purification des protéines avec des protocoles adéquats)²⁴.

Deux systèmes de sécrétion sont très bien caractérisés et décrits à ce jour dans la littérature : la voie Sec (ou translocase) et la voie Tat (« Twin-arginine translocation »), que l'on retrouve dans tous les domaines du vivant (*i.e. bacteria, archaea, and eukarya*)¹²⁹. D'autres voies de sécrétion protéique moins bien caractérisées et impliquées de manière plus secondaire sont également présentes chez *B. subtilis*. Ainsi, les transporteurs liés à l'ATP ou « ABC transporteurs » (pour ATP-Binding Cassette), mais aussi, une voie de sécrétion liée à la compétence passant par le « pseudopilus Com » peuvent aussi participer à l'export de protéines¹³⁰. Par soucis de clarté, nous ferons seulement un résumé des deux premiers systèmes cités (**Figure 7**).

La voie Sec, la première décrite constitue le mécanisme de transport principal (*i.e. la plus utilisée*) des protéines à travers la membrane puis la paroi bactérienne de *B. subtilis*. Dans

ce mécanisme, la pré-protéine à sécréter est reconnue via son peptide signal en N-terminal par le SRP (particule de reconnaissance du signal) qui interagit avec la protéine FtsY (activité GTPasique) redirigeant la protéine à sécréter vers la translocase enchâssée dans la membrane. Cette macro-structure comprend plusieurs protéines : SecA (ATPase avec activité motrice), SecDF et SecYEG (formant le canal à travers la membrane). Cette voie agit selon deux modes : soit (i) co-traductionnel, soit (ii) post-traductionnel. De nombreuses protéines chaperonnes cytosoliques sont associées à cette voie de sécrétion, comme des « Heat Shock protein », les protéines GroESL, ou encore la protéine CsaA, afin de guider les protéines non-repliées jusqu'à la voie Sec. Suite à la translocation (peu importe le mode), le peptide signal est clivé par des signal peptidases de type I (SipS ou SipT,U,V,W) ou de type II (LspA).

La voie Tat est composée de trois protéines TatA (TatAc, TatAd et TatAy) et de deux protéines TatC (TatCd et TatCy). Le complexe formant un pore à travers la membrane et servant d'amarrage est constitué principalement de l'assemblage de TatAyCy ou bien de TatAdCd. De manière générale, cette voie de sécrétion prend en charge des protéines déjà repliées, à la différence de la voie Sec qui est impliquée dans l'export de pré-protéines non repliées. Ces dernières sont transloquées à travers la membrane suivant un gradient de pH. Le peptide signal est clivé de la même manière que pour la voie Sec à l'aide de signal peptidases.

Après translocation et export des protéines à travers la membrane cytoplasmique, par l'une des voies mentionnées ci-dessus, de nombreux facteurs chaperons (périplasmiques ou bien extracellulaires) protègent les protéines nouvellement sécrétées. On retrouve notamment, la protéines PrsA (lipoprotéine associée à la phase externe de la membrane), mais également, des cations métalliques, des thiol-disulfides oxidoreductases, ou encore des propeptides.^{130,131}

D'un point de vue appliqué, il est intéressant de noter que l'on peut modifier par ingénierie ces systèmes de sécrétion, ou bien les chaperonnes associées. Ainsi, parmi quelques exemples de la littérature Diao *et al.*, ont modifié la voie Sec en co-hybridant la protéine SecA avec la protéine SecB de *E. coli*. Leur étude montre que cette modification, augmente de près de 60% la production de protéines recombinantes¹³². William *et al.*, en 2013, ont quant à eux sur-exprimé la protéine PrsA et considérablement augmenté la sécrétion d'un antigène de

B. anthracis chez *B. subtilis*¹³³. Les connaissances sur ces systèmes sont finalement propices à améliorer les rendements lors d'applications biotechnologiques.

I.6.3. Exemples d'applications industrielles

Au cours de procédés biotechnologiques visant à produire des molécules d'intérêt, *B. subtilis* est une des usines microbiennes les plus utilisées en raison de sa capacité de fermentation, des rendements élevés des produits qu'elle sécrète (20 à 25 grammes par litre) et l'absence de sous-produits toxiques²⁰. Parmi les molécules que *B. subtilis* est capable de produire sont retrouvées, des enzymes, des protéines hétérologues, des antibiotiques, ou encore, des vitamines.^{20,134,135}

B. subtilis est un hôte d'expression idéal pour la production de diverses enzymes industrielles (près de 50% du marché total des enzymes)²⁴, telles que des amylases, des β -galactosidases, ou des cellulases¹³⁶⁻¹³⁸. Ces enzymes jouent un rôle important dans les industries de l'alimentation humaine et animale, des détergents, du textile, du cuir, du papier et pharmaceutique²⁰. En raison de son statut GRAS, les protéases de *B. subtilis* peuvent être utilisées dans diverses applications alimentaires, telles que la préparation d'hydrolysats de soja, l'attendrissement de la viande, la préparation d'hydrolysats de caséine, la coagulation du lait et le traitement des déchets alimentaires¹³⁹.

C'est aussi une bactérie capable de produire des vitamines (B1, B2, B5, B6 et B7)¹³⁵. La vitamine B2 (ou riboflavine) est une vitamine produite en très grande quantité par *B. subtilis* à l'échelle industrielle¹⁴⁰. Elle est largement utilisée pour ses effets antioxydants, anticancéreux, ou de renforcement de l'immunité, ainsi que, pour l'amélioration de l'alimentation humaine et animale¹⁴¹.

B. subtilis peut également sécréter une variété de peptides antimicrobiens et de bactériocines de faible poids moléculaire, tels que la surfactine¹⁴², la bacilysine¹⁴³ et la subtiline¹⁴⁴, qui sont utilisés en ingénierie biomédicale, dans l'alimentation et en agriculture¹⁴⁵. Les peptides antimicrobiens produits par *B. subtilis* sont des outils thérapeutiques prometteurs, en raison de leur large spectre d'activité et de leur capacité à tuer rapidement une variété de pathogènes. Avec les problèmes croissants de résistance

microbienne dus à l'utilisation abusive des antibiotiques conventionnels, les peptides antimicrobiens jouent un rôle important dans le traitement des infections bactériennes.¹⁴⁶

II. Biologie de synthèse : vers le développement de châssis minimaux

II.1. Généralités sur la biologie de synthèse

La biologie de synthèse (BS) est une discipline émergente qui emploie des principes d'ingénierie dans le but de reprogrammer les systèmes vivants¹⁴⁷.

Le terme « biologie synthétique » a été proposé en 1912 par le chimiste français Stéphane Leduc (1853-1939)¹⁴⁸, puis a été repris pour caractériser le domaine de la BS à l'interface entre la biologie moléculaire et les méthodes d'ingénierie¹⁴⁹. Depuis deux décennies, cette discipline a pris toute son ampleur en révolutionnant et conceptualisant les approches d'ingénieries biologiques à l'échelle des systèmes¹⁵⁰. La biologie de synthèse s'appuie sur plusieurs disciplines que sont la physique, la chimie, l'informatique, les mathématiques, la biologie et l'ingénierie¹⁵¹. La définition de la biologie de synthèse n'est pas complètement arrêtée puisqu'elle diverge en fonction des instances, de la réglementation, de son évolution, mais aussi, de partis pris individuels. Par exemple pour George Church (Harvard Medical School, Boston, États-Unis) « Le génie génétique se concentre sur des gènes individuels (généralement le clonage et la surexpression). L'extension logique de cela à l'ensemble du système, c'est l'ingénierie des génomes. Intermédiaire entre les deux est l'ingénierie métabolique, qui implique l'optimisation de plusieurs gènes à la fois. La biologie de synthèse est une "méta" de toutes ces activités qui établit des normes pour les modules génétiques (intentionnellement interopérables dans leur assemblage et leur fonctionnement). Ces propriétés hiérarchiques permettent une conception assistée par ordinateur à différents niveaux d'abstraction : du niveau sous-moléculaire aux niveaux supra-écosystémiques ». En revanche Andrew Ellington (Université du Texas, Austin, États-Unis) considère que « Ces mots [biologie synthétique] n'ont pas beaucoup de sens. La définition d'un nouveau domaine repose soit sur une découverte, soit sur une redéfinition, et - parce que je ne peux pas citer une seule grande découverte dans ce domaine - la biologie synthétique est plutôt une redéfinition de la biotechnologie. [...] Mais je dirais que la principale utilité de la biologie synthétique est d'enthousiasmer les ingénieurs, les étudiants

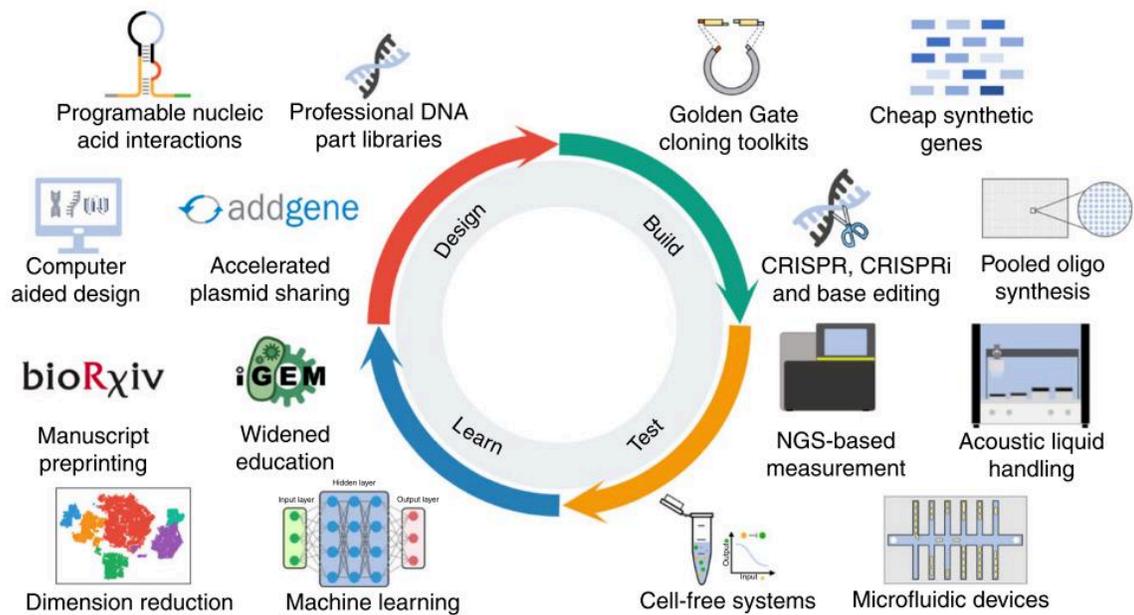


Figure 8 : Schéma représentant les technologies et méthodes qui ont accéléré le cycle « Design – Built – Test -Learn ». Cette figure est issue de Meng & Ellis 2020. Le diagramme représente le cycle d'ingénierie utilisé en biologie synthétique (au centre) avec des dessins illustrant certaines des principales technologies (e.g. NGS, microfluidique, machine learning) et méthodes (e.g. golden gate, CRISPR) qui contribuent à accélérer chaque phase du cycle.

de premier cycle et les organismes de financement. Son principal inconvénient est de créer une hystérie dans la communauté ». ¹⁵² Cependant, d'un point de vue plus global deux objectifs généraux de la biologie de synthèse se dégagent : (i) reproduire et modifier à large échelle des systèmes biologiques déjà existants et (ii) proposer/concevoir de nouveaux systèmes ou organismes qui ne sont pas retrouvés dans la nature.

En 2005, le Dr. Drew Endy donne une image très conceptuelle en comparant la biologie de synthèse à la construction d'un immeuble qui est un processus ancré et normalisé. Dans ce domaine, il existe un ensemble de matériel prédéfini (ciment, brique, outils), des règles d'utilisation ou mode d'emploi pour l'utilisation de ce matériel, et un personnel qualifié et formé pour appliquer ces directives ¹⁵³. Transposée à la biologie et au vivant, cette vision est tout de même plus difficile à appliquer car elle doit prendre en compte la complexité des organismes biologiques, les variations physiques ou biologiques de ces derniers et finalement leur évolution. Pour cela, la communauté scientifique se base notamment sur le cycle « Design Built Test Learn » (DBTL), qui est une approche rétroactive pensée pour associer des données bioinformatiques et expérimentales. Cette méthode d'évolution dirigée permet une amélioration incrémentale du système étudié. La BS s'appuie sur des technologies clés développées ces dernières années telles que les techniques de séquençage et de synthèse, la puissance de calcul de l'informatique, ou la microfluidique et qui ont aidé à l'amélioration de ce cycle DBTL et à l'accélération des processus expérimentaux (**Figure 8**) ¹⁵⁰.

La BS a pour ambition de répondre notamment aux grands enjeux sociétaux tels que la pollution, le réchauffement climatique, les besoins énergétiques ou soigner des maladies en s'inspirant du vivant. Ainsi, plusieurs secteurs sont visés comme l'ingénierie de voies métaboliques existantes ou le clonage de nouvelles voies métaboliques entières. La production d'artémisinine avec une voie métabolique optimisée et synthétique chez *E. coli* est un exemple frappant qui a permis de produire en grande quantité ce médicament contre la malaria ¹⁵⁴⁻¹⁵⁶. D'autres travaux décrivent la modification du métabolisme de *E. coli* en intégrant des voies métaboliques hétérologues, dans le but de produire des hydrocarbures, utilisables comme bio-carburant ^{157,158}.

Un autre secteur de la BS tend à promouvoir la conception et la synthèse de circuits génétiques artificiels pour introduire les principes d'ingénierie d'abstraction (focalisation sur un niveau de complexité de la cellule) et de standardisation des systèmes biologiques ¹⁵³.

L'idée étant de créer un système biologique en utilisant des pièces standardisées pour détourner les signaux interprétés par la cellule et lui faire exprimer quelque chose en particulier (*e.g.* utilisation de cellules biologiques utilisées en tant que biosenseurs). Levskaya *et al.* ont par exemple modifié des cellules de *E. coli* de façon à ce qu'elles puissent répondre à un stimulus lumineux comme un senseur biologique¹⁵⁹. Des efforts significatifs ont été faits aussi pour améliorer les systèmes de régulation d'expression des gènes, notamment en construisant des systèmes synthétiques qui répondent de manière contrôlée à des stimuli extérieurs¹⁶⁰.

D'autres aspects de la BS s'orientent vers la création de systèmes biologiques orthogonaux basés sur une chimie inconnue du vivant et utilisant de nouveaux nucléotides, ou acides aminés¹⁶¹⁻¹⁶³ (xénobiologie).

Enfin, la construction d'une cellule artificielle contrôlée par un génome synthétique publiée par une équipe du JCVI en 2010¹⁶⁴ a ouvert la voie à la création de châssis cellulaires « à la carte ». La conception de châssis rationalisés représentés par des micro-organismes dotés d'un nombre standardisé de gènes, dont la séquence a été modifiée, produite chimiquement et réorganisée pour une application donnée est devenue un des objectifs de la BS¹⁶⁵. Ce dernier point relève encore du défi puisque la complexité des organismes vivants ne permet pas encore d'arriver à ce résultat. Par défaut, une approche consiste à optimiser l'environnement cellulaire des organismes hôtes afin de limiter les interactions avec un système hétérologue synthétique. L'objectif étant de proposer un environnement stable qui serait minimisé¹⁶⁶.

II.2. Développement de châssis minimaux en BS

II.2.1. Notion de cellules minimales et d'essentialité de gènes

La compréhension des principes de base de la vie s'appuie notamment sur le concept de cellule minimale, pour lequel Morowitz propose en 1984 d'étudier les mycoplasmes¹⁶⁷. Les mycoplasmes sont considérés comme les plus petits organismes vivants capable d'auto-réplication et de croissance autonome dans un milieu défini ; et se rapprochant le plus du concept de cellule minimale¹⁶⁸. Ce sont des bactéries dépourvues de paroi avec un génome réduit (0,58 Mb à 1,8 Mb), et donc un métabolisme simplifié. Les mycoplasmes ont un mode

de vie parasitaire et leur évolution réductive a consolidé l'idée que ces organismes pouvaient être un modèle de base pour la définition du jeu minimal de gènes nécessaires à la vie en conditions optimales de laboratoire (milieu défini, absence de stress)¹⁶⁹. À partir de cette compréhension d'une vie cellulaire simplifiée, il deviendrait possible de modéliser et de prédire le comportement des cellules auxquelles on aurait ajouté des gènes et de nouvelles voies métaboliques¹⁷⁰. Pour progresser dans cette vision, une étape préliminaire consiste à identifier et classer les gènes selon trois catégories : (i) les gènes essentiels (E), c'est-à-dire les gènes 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 ; et (iii) les gènes quasi-essentiels (QE) qui ne sont pas strictement essentiels, mais dont l'inactivation altère la croissance des cellules^{171,172}.

En 1995, les efforts des scientifiques pour séquencer le génome d'*Haemophilus influenzae*¹⁷³ puis celui de *Mycoplasma genitalium*¹⁷⁴ permettent d'obtenir les premières séquences complètes de génomes bactériens et ainsi de procéder à l'identification *in silico* d'un ensemble de gènes minimal. En 1996, Mushegian et Koonin, via une première approche de génomique comparée entre ces deux génomes¹⁷⁵, identifient un ensemble de 256 gènes conservés entre ces bactéries appartenant à deux phyla bactériens. Il est alors vraisemblable que ces gènes soient essentiels. Puis des données expérimentales ont permis de préciser ces données. En 1999, Hutchinson *et al.*, ont démontré par des expériences de transposition chez *Mycoplasma genitalium* que 150 gènes non essentiels contre 375 essentiels étaient identifiés pour cette bactérie¹⁷⁶. Cette étude a ensuite été complétée par Glass *et al.* en 2006, sur 482 gènes codant des protéines chez *M. genitalium*, 382 ont été prédits comme étant essentiels contre 100 comme étant non essentiels.¹⁷⁷ D'autres études d'inactivation à grande échelle ont été réalisées, pour tenter de définir quels gènes sont essentiels à la survie cellulaire dans plusieurs modèles bactériens bien caractérisés, tels que *Escherichia coli*, *Bacillus subtilis*, *Staphylococcus aureus* et *Mycoplasma pneumoniae*¹⁷⁸. Ceci a été réalisé en utilisant plusieurs stratégies telles que l'inactivation massive de gènes par transposition, via l'utilisation d'ARN antisens pour inhiber l'expression des gènes, ou via l'inactivation systémique (*e.g.* transposition) de chaque gène individuellement dans le génome^{171,179-181}.

Deux approches se distinguent pour la construction de cellules minimales : l'approche « Top-Down », qui consiste à partir d'un système complexe existant dans la nature,

et lui retirer les composants non essentiels. La réduction du génome à grande échelle a été utilisée pour réduire au minimum les génomes de micro-organismes modèles au laboratoire et dans l'industrie, tels que *E. coli* et *B. subtilis*^{182,183}. Actuellement, la taille minimale du génome de *B. subtilis* par cette approche a été réduite à 2,68 Mb avec 1605 suppressions de gènes, ce qui représente une diminution de 36,5% du génome original¹⁸³. Ce génome réduit contient 18% de gènes aux fonctions inconnues, ce qui est encore loin de l'objectif ultime de compréhension des processus essentiels et, de construire avec ces derniers un châssis cellulaire. D'autres travaux similaires ont démontré la réduction du génome de *B. subtilis* à une taille de 2,73 Mb avec une stratégie itérative nommée « Pop-in/pop-out »¹⁸⁴. De la même manière les auteurs ont conservé 341 gènes aux fonctions inconnues, 82 gènes codant pour des protéines putatives et 18 pseudogènes. Ceci indique bien que des données supplémentaires doivent être produites pour déterminer dans quelle mesure ces régions sont dispensables ou non.

L'approche « Bottom-up » vise, quant à elle, à créer la vie *ab initio* (depuis le début), en assemblant des composants de base (ou biobriques) un à un jusqu'à reconstitution d'un système complet pour copier les fonctions du vivant (cellule artificielle). Un résultat marquant a été obtenu en 2010, avec la création de la première cellule bactérienne contrôlée par un génome synthétique¹⁶⁴. Cette étude démontre qu'il est possible de partir d'informations génétiques numérisées, de synthétiser un nouvel ADN et de le transplanter dans des cellules en remplaçant toutes les informations génétiques existantes, pour aboutir à de nouvelles cellules contrôlées uniquement par cet ADN synthétique. La maîtrise de ces technologies offre des avantages sans précédent, car ce sont des outils puissants permettant de modifier des génomes bactériens à grande échelle, et de les utiliser comme plateforme pour fabriquer des molécules d'intérêt.

II.2.2. Nouvelle génération de châssis bactériens en BS

Un châssis est un cadre rigide ou mobile fait d'une matière résistante, destiné à entourer ou supporter quelque chose (définition du Larousse). Il peut être spécialisé en fonction du domaine dans lequel il est utilisé : par exemple le châssis d'une fenêtre est le cadre

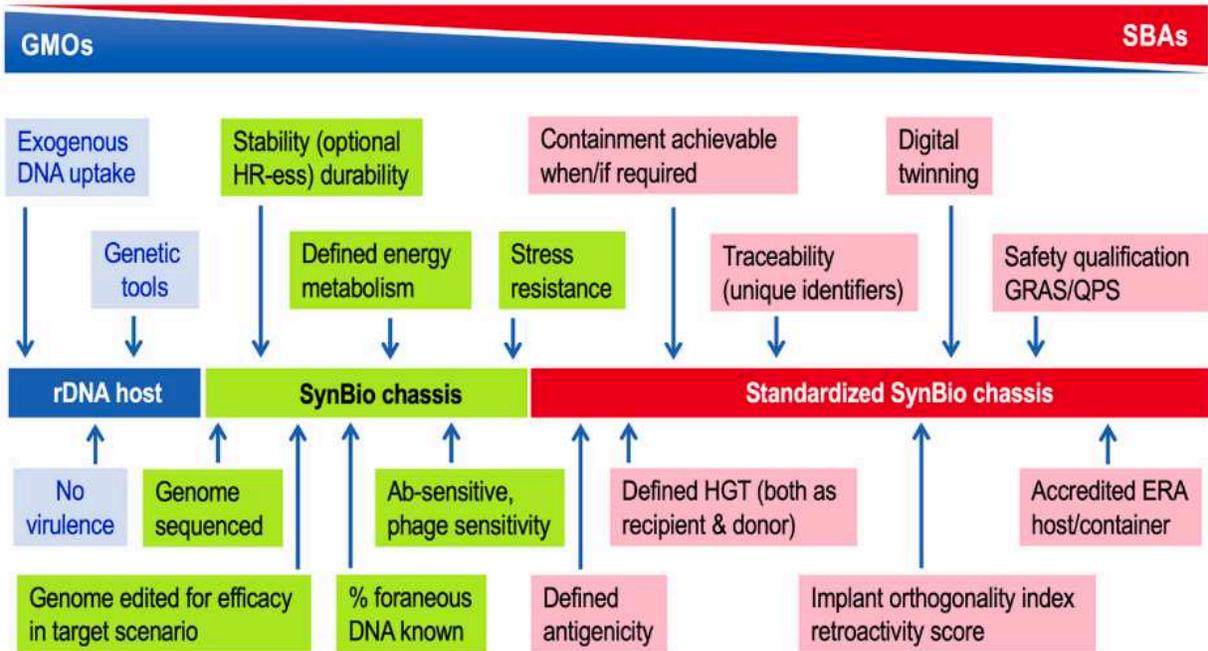


Figure 9 : Graphique représentant les différences entre des organismes génétiquement modifiés (OGMs) et des agents SynBio (SBA). Cette figure est issue de de Lorenzo *et al.* 2021. Le schéma indique la nature des informations qui devraient être disponibles pour chaque catégorie de souches. En bleu sont indiquées les caractéristiques des souches « hôtes d’ADN », en vert des « châssis SynBio » et en rouge des « châssis SynBio standardisés ».

qui supporte le vitrage, tandis que pour une automobile c'est une structure destinée à supporter les roues, la suspension et la carrosserie d'un véhicule.

Le terme de « châssis » est largement utilisé en biologie de synthèse par la communauté scientifique, si bien que son utilisation est étendue à des entités différentes (cellules modifiées, cellules avec un génome synthétique, systèmes sans cellules « cell-free system », ou même vésicules) et pour des fonctions différentes selon le contexte. La diversification de l'emploi de ce mot, rend de ce fait son véritable sens assez flou. De Lorenzo *et al.* ont dernièrement proposé une définition de ce terme châssis en biologie de synthèse comme étant « une plateforme biologique modulable à façon et réutilisable, avec un génome codant un nombre de fonctions basiques pour sa propre maintenance, sa croissance et l'utilisation de ses fonctions de manière optimale ; tout en étant capable de traiter la tâche pour laquelle elle aura été déterminée et ce, dans un environnement prédéfini ». La notion de réduction de taille (du génome) y est sous-jacente à travers l'idée d'un nombre de fonctions basiques allouées au bon fonctionnement de l'hôte d'accueil appelé châssis SynBio. Les auteurs repoussent plus loin encore la réflexion sur ce terme, en définissant des limites entre les organismes « hôte d'accueil d'ADN recombinant », « châssis SynBio » et « châssis SynBio standardisés ». Ils indiquent par ailleurs un certain nombre de paramètres pour différencier les organismes génétiquement modifiés (OGMs), des châssis SynBio standardisés et proposent même une nouvelle appellation sous le terme d'agent SynBio (SBAs) (**Figure 9**)¹⁵¹.

A l'heure actuelle, les avancées dans le domaine de la biologie de synthèse ne permettent pas encore de parler d'agent SynBio ultra-standardisés, mais plutôt de châssis SynBio pour lesquels des organismes modèles de laboratoire (*e.g.* *E. coli* pour les procaryotes et *S. cerevisiae* pour les eucaryotes) sont déjà largement utilisés pour des applications en biotechnologies^{185,186}. La production de nouveaux châssis Synbio à partir d'espèces microbiennes moins conventionnelles, requiert un certain nombre de conditions pour que ces derniers soient considérés comme idéaux¹⁸⁷. Par exemple, il est nécessaire de disposer de connaissances suffisantes sur le micro-organisme ciblé, et d'avoir à disposition des outils de manipulation génétiques et des protocoles de transformation pour pouvoir faire de l'ingénierie. De plus, cet organisme doit avoir des besoins nutritionnels simples, être robuste (*e.g* grâce à une paroi par exemple) et être capable de croître rapidement. Il doit aussi résister

à des stress physico-chimiques et à certaines conditions (température élevée ou pH acide). Parmi les bactéries proposées pour faire partie de la nouvelle génération de châssis microbiens sont retrouvées *Pseudomonas putida*¹⁸⁸, *Streptomyces sp*¹⁸⁹, ou bien, *Bacillus subtilis*¹⁹⁰.

B. subtilis est particulièrement approprié comme modèle bactérien pour devenir un châssis SynBio en raison de son statut non pathogène (GRAS), mais aussi, car c'est une bactérie étudiée de longue date et utilisée comme hôte pour la sécrétion d'enzymes et d'antibiotiques (cf. Etude bibliographique §I). Les développements actuels en matière de construction de châssis de *B. subtilis* se concentrent sur les aspects suivants : identification de l'ensemble de gènes minimal, réduction du génome à grande échelle, élimination de voies métaboliques pour éviter la co-utilisation de sources de carbone multiples et ingénierie des composants cellulaires pour la synthèse de produits spécifiques¹⁹¹. En tant que tel, c'est un candidat idéal pour le développement d'un châssis minimal bien caractérisé dans lequel on peut mettre en œuvre avec précision des circuits métaboliques complexes. Des répertoires de gènes et d'intervalles non essentiels sont disponibles^{103,184,192} et des approches de suppression de portion de son génome à grande échelle ont déjà été mises en œuvre avec succès pour évaluer le potentiel des souches à génome réduit pour des applications biotechnologiques^{183,193}. Deux souches dont le génome a été réduit d'environ 20% ont montré une productivité accrue pour la production de protéines recombinantes¹⁹³ et une production de guanosine et de thymidine multipliée par cinq¹⁹⁴. Plus récemment, une souche appelée « *MiniBacillus* » (avec un génome minimisé à 2,7 Mb) a permis l'expression de l'antigène A de *Staphylococcus aureus* (IsaA). Cependant, malgré que la protéine ait été sécrétée avec succès, les auteurs ont constaté qu'une certaine quantité d'IsaA était toujours présente dans le cytosol et dans la fraction membranaire, entraînant une forte réponse au stress et un ajustement de l'ensemble des transporteurs de la cellule^{195,196}.

II.2.3. Réglementation de la biologie de synthèse et traçabilité des châssis

L'adoption croissante des technologies, des souches et des produits SynBio par l'industrie biotechnologique a placé leur identification dans la ligne de mire des organismes de réglementation. Le jargon SynBio (y compris le terme "châssis") peut être utilisé dans le

domaine universitaire avec une grande flexibilité, mais les frontières entre les OGM traditionnels et les agents SynBio sont assez floues et il est difficile d'établir des critères objectifs pour les distinguer¹⁵¹.

Un certain nombre de mesures peuvent être adoptées pour assurer la lutte biologique : certaines sont inhérentes à l'organisme (auxotrophie artificielle ; utilisation d'un code génétique non standard *e.g.* les mycoplasmes), tandis que d'autres peuvent être ajoutées (kill-switches¹⁹⁷, incorporation d'acides aminés non naturels dans les protéines essentielles).

Chez *E. coli*, une approche a notamment permis de réduire l'utilisation du code génétique à 63 codons démontrant la possibilité d'expansion du code génétique¹⁹⁸. D'autres travaux ont démontré la reconstruction d'un génome synthétique et redessiné pour supprimer l'utilisation de trois codons sur les 64 codons du code génétique¹⁹⁹. Un tel recodage permet non seulement, de modifier les cellules afin d'insérer facilement des acides aminés non standard dans les protéines, mais aussi de proposer un code alternatif pour différencier les organismes SynBio de ceux naturels.

D'autres idées plus complexes faisant appel à des pare-feu génétiques pour contenir les organismes génétiquement modifiés et les agents SynBio proposent de mettre en place des systèmes de code barre génétique en faisant appel à de la stéganographie sur l'ADN. Cette méthodologie a été mise au point pour cacher des messages dans des régions variables (polymorphismes de nucléotides simples)²⁰⁰. Ce genre d'approches pourrait assurer un suivi de traçabilité et compléter les mesures de confinement déjà utilisées pour la manipulation de micro-organismes modifiés génétiquement¹⁶⁵.

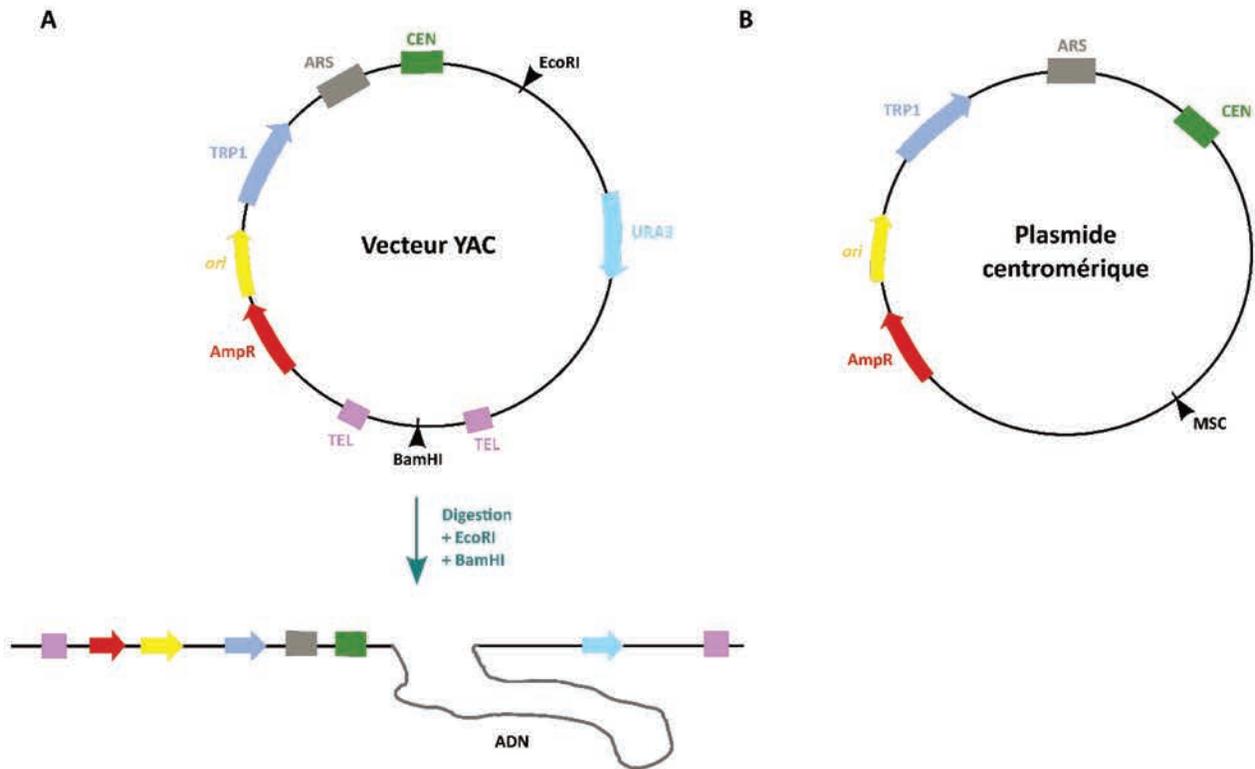


Figure 10 : Schéma représentant la différence entre un vecteur YAC et un plasmide centromérique. TEL signifie télomère, AmpR correspond au marqueur de résistance à l'ampicilline, *ori* correspond à une origine de réplication chez *E. coli* (e.g. *colE1*), TRP1 correspond au marqueur d'auxotrophie pour le tryptophane, ARS correspond à une origine de réplication levure, CEN signifie centromère, URA3 correspond au marqueur d'auxotrophie pour l'uracyle.

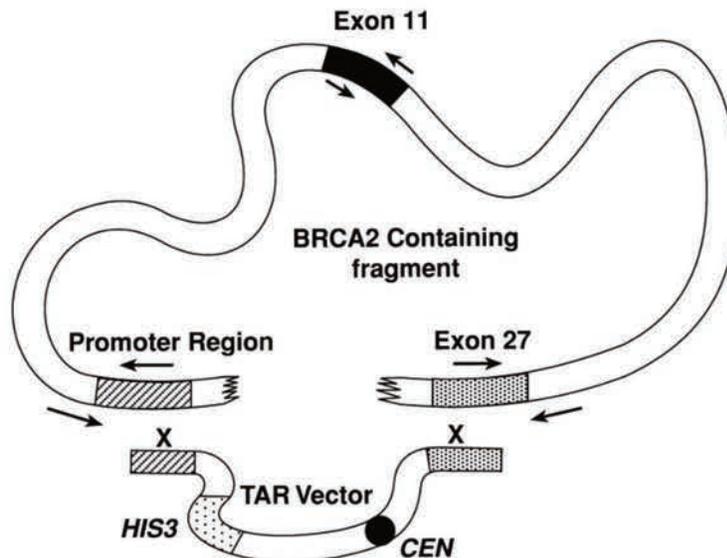


Figure 11 : Capture du gène humain *brca2* en tant que YAC circulaire. Cette figure est issue de Larionov *et al.* 1997. La recombinaison entre les séquences du vecteur TAR et l'ADN génomique de BRCA2 conduit à l'établissement d'un fragment circulaire (appelé YAC par les auteurs). CEN signifie centromère et HIS3 correspond au marqueur d'auxotrophie pour l'histidine.

III. La levure comme plateforme d'ingénierie de fragments d'ADN ou de génomes microbiens

La levure *Saccharomyces cerevisiae* est un organisme utilisé en tant que modèle d'étude des fonctions biologiques eucaryotes incluant la génomique, les voies métaboliques, ou encore le vieillissement cellulaire²⁰¹. C'est aussi un micro-organisme permettant l'exploration des interactions entre protéines et qui est utilisé comme usine de production de molécules pharmaceutiques ou d'intérêt biotechnologique²⁰².

C'est également un hôte de choix pour le clonage et la propagation de molécules d'ADN en tant que chromosome artificiel de levure (Yeast Artificial Chromosome ou YAC) ou de plasmides centromériques (**Figure 10**). L'utilisation de YAC est référencée pour la première fois à la fin des années 1980 par Burke *et al.*, démontrant la possibilité de cloner des molécules d'ADN dont la taille peut atteindre 400 kpb²⁰³. Les YAC sont des constructions linéaires composées de la séquence d'ADN capturée, additionnée d'un centromère (CEN), d'une origine de réplication levure (ARS), de deux marqueurs d'auxotrophie (à l'origine URA3 et TRP1) positionnés de part et d'autre de la séquence, et flanquée de deux télomères levures (TEL) à chaque extrémité. Cette approche a ouvert la voie au clonage de fragments de génomes de toute une gamme d'organismes, incluant des eucaryotes, des bactéries et des virus²⁰⁴⁻²⁰⁶. Les applications en découlant sont nombreuses et comprennent notamment des analyses de fonctions de gènes ou encore la cartographie de génomes complexes²⁰⁷⁻²⁰⁹.

A travers la méthode de Transformation-Associated Recombinaison (TAR) traduite de l'anglais par « transformation associée à la recombinaison », Larionov *et al.* ont démontré en 1997, la capture du gène *brca2* humain impliqué dans le cancer du sein (90 kpb) comme un « YAC circulaire » (**Figure 11**). Le vecteur de clonage TAR (linéaire) utilisé pour cette étude avait une structure simplifiée. Il ne possédait ni ARS, ni séquences télomériques. Il était constitué d'un centromère et du marqueur HIS3, flanqué en 5' par une partie du promoteur du gène *brca2* (669pb) et en 3' par la séquence terminale (exon 27) du gène (308pb)²¹⁰. Cette organisation du vecteur TAR permettait pour la première fois de capturer un fragment d'intérêt sous forme circulaire, qui s'apparentait désormais davantage à un grand plasmide centromérique (Ycp, **Figure 10B**) qu'à un YAC. L'année suivante, une version modifiée du vecteur TAR a été proposée se basant cette fois sur un vecteur TAR « YAC-BAC » (Bacterial Artificial Chromosome ou chromosome artificiel bactérien) permettant non seulement la

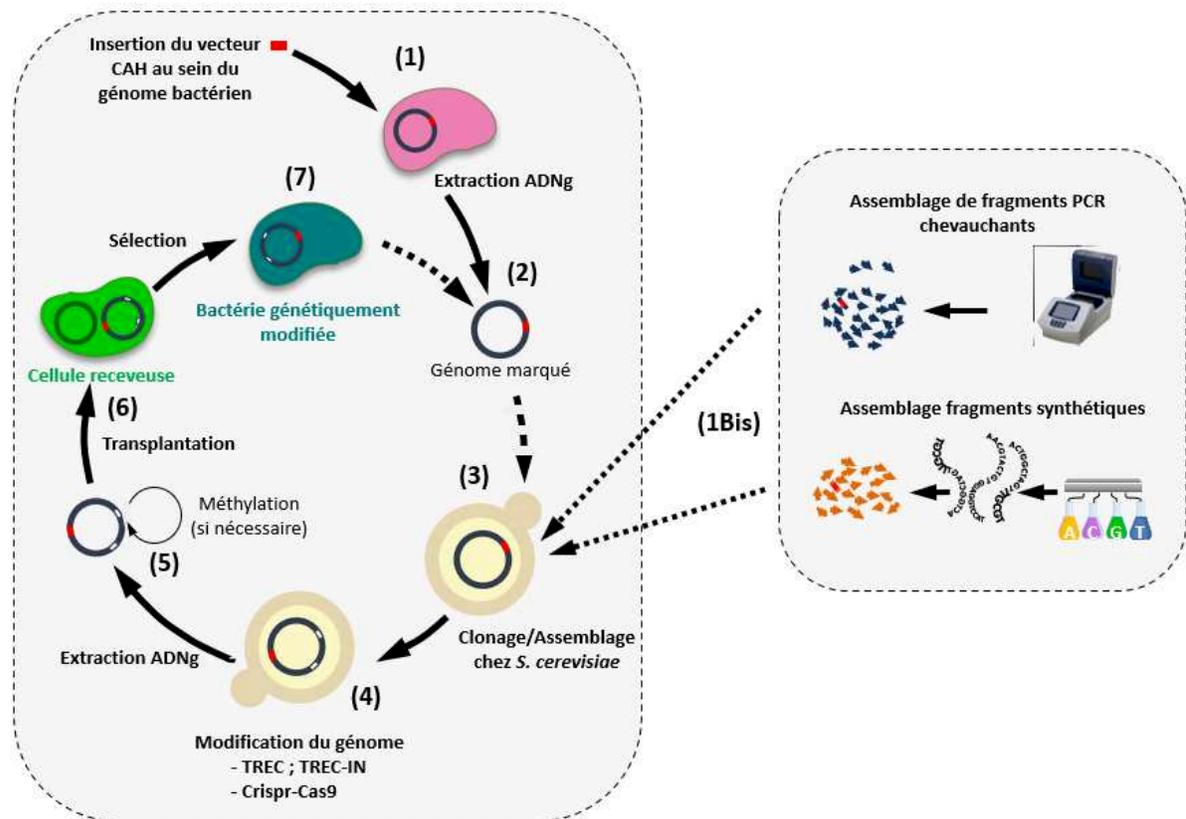


Figure 12 : Schéma illustrant le cycle d'ingénierie de génome bactérien développé chez les mycoplasmes. Adapté d'après Lartigue *et al.* 2009. **(1)** Un vecteur est inséré au sein du génome bactérien afin d'assurer la réplication et le maintien du génome dans la levure. Ce vecteur CAH contient des éléments levures (un Centromère CEN6, une origine de réplication ARSH4 et un marqueur d'auxotrophie Histidine), ainsi qu'un marqueur de résistance. **(2 à 4)** Le génome marqué est ensuite isolé, cloné/assemblé **(1Bis)** chez *S. cerevisiae* puis modifié grâce aux outils génétiques disponibles chez la levure (TREC, TREC-IN, CRISPR-Cas9). **(5-6)** Le génome nouvellement modifié est extrait, méthylé *in vitro* puis transplanté dans une cellule receveuse. La méthylation est nécessaire afin de protéger le génome donneur des systèmes de restriction présents dans la cellule receveuse. **(7)** Une bactérie génétiquement modifiée est alors obtenue et peut-être caractérisée au niveau génotypique et phénotypique.

capture d'un morceau d'ADN humain circulaire pouvant être propagé chez la levure, mais aussi chez *E. coli* grâce à la réplication unidirectionnelle du facteur F (*oriS* et *repE*) ajoutée dans la « partie BAC » du vecteur²¹¹.

C'est au cours de travaux réalisés au John Craig Venter Institute (JCVI, Rockville - Etats Unis) que son utilisation en tant que réelle plateforme d'ingénierie de génome bactérien a pris toute son ampleur. Cette équipe avait pour ambition de construire le premier génome synthétique bactérien (582 kpb). Les premières étapes ont été réalisées par Gibson Assembly²¹² et clonage des fragments assemblés *in vitro* dans *E. coli* et ont permis d'assembler des petits fragments d'ADN d'environ 5 à 7 kpb en fragments de plus en plus grands de taille environ égale à 144 kpb (~quarts de génome). Puis, face à l'impossibilité d'assembler et de cloner de plus grands fragments chez *E. coli* (demi-génomes et génome entier), les auteurs ont eu recours à la levure *S. cerevisiae* pour l'assemblage finale. Les auteurs ont ainsi démontré qu'il était possible d'assembler et de cloner dans la levure le génome synthétique de *Mycoplasma genitalium* JCVI-1.0 par recombinaison de six fragments chevauchants pour produire un génome circulaire de 582 kpb²¹³. Cette équipe a repoussé les limites de la méthode en transformant simultanément la levure avec 25 fragments chevauchants, dont les tailles variaient de 17 à 35 kpb, permettant l'assemblage de ce même génome synthétique en une seule étape²¹⁴.

Suite à ces travaux, une approche innovante d'assemblage et de clonage de génomes bactériens naturels ou synthétiques dans la levure a été proposée. Cette alternative, dans laquelle la levure *S. cerevisiae* est utilisée comme plateforme d'édition d'ADN, est composée de trois étapes (**Figure 12**) : (i) le clonage/assemblage du génome bactérien à éditer, entier, dans la levure^{213,215} ; (ii) l'édition de ce génome cloné/assemblé en utilisant les outils d'ingénierie génétique disponibles chez la levure²¹⁶ ; et (iii) la transplantation du génome modifié de la levure vers une cellule receveuse, afin d'obtenir une cellule bactérienne mutée^{216,217}. Ces travaux démontraient, pour la première fois, la capacité des biologistes à synthétiser des molécules d'ADN à partir d'informations génétiques numérisées puis à les greffer dans des cellules receveuses (remplaçant toute l'information génétique existante) afin de produire des cellules contrôlées uniquement par l'ADN synthétique nouvellement conçu²¹⁸. Le premier organisme auto répliquatif contrôlé par un génome

Tableau 1 : Exemples de génomes microbiens clonés ou édités dans la levure. Tableau issu de Vashee *et al.* 2020.

| Table 1 | | | | | | |
|--|------------|---------|--------------|-------------------------------|---------------------|-------------------------|
| Key examples of microbial genomes cloned or edited in yeast ^a . | | | | | | |
| Organism | Size (Mbp) | % G + C | Genetic code | Cloning strategy ^b | Rescue strategy | References ^c |
| Mollicutes | | | | | | |
| <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 ferriuminatoris</i> | 1.2 | 24 | Nonstandard | Cloning | Transplantation | Pers. com. 2019 |
| Proteobacteria | | | | | | |
| <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] |
| Cyanobacteria | | | | | | |
| <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] |
| Diatoms | | | | | | |
| <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] |
| Viruses | | | | | | |
| 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 | 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] |

^a Sorting is done by year of publication.

^b Cloning refers to any method described in the text.

^c The assembly was performed in mammalian cells.

synthétique d'environ 1,08 Mpb, inspiré du génome de *Mycoplasma mycoides* subsp. *capri* (*Mmc*), a été appelé *Mmc* JCVI-syn1.0.

Depuis 2009, ce cycle d'ingénierie de génome est exploité pour cloner et/ou assembler, puis modifier de nombreux chromosomes incluant ceux de bactéries, de virus, ou d'organismes eucaryotes²¹⁹. Plus précisément des génomes natifs bactériens de mollicutes^{215,220–222} (e.g. *Mycoplasma genitalium*, *Mycoplasma pneumoniae*, *Acholeplasma laidlawii*, *Mesoplasma florum*), de protéobactérie et de cyanobactérie (e.g. *Haemophilus influenzae*, *Prochlorococcus marinus*)^{223,224} ont été clonés (**Tableau 1**). Des versions synthétiques et assemblées sont également référencées à l'instar des génomes JCVI-Syn1.0²¹⁸ et JCVI-Syn3.0¹⁷², ou encore ceux d'*Escherichia coli* (version réduite du génome à 1,03 Mb)²²⁵ et de *Caulobacter ethensis 2.0* (version réduite de *Caulobacter crescentus* à 0,78 Mb)²²⁶ (**Tableaux 1 et 2**). Pas moins de douze génomes viraux clonés ou assemblés après synthèse sont également répertoriés, comme celui du virus de type 2 de la dengue²²⁷, ou très récemment celui du virus émergent SARS-CoV-2 agent de la COVID19²²⁸ (**Tableau 1**).

La levure ainsi exploitée comme plateforme d'ingénierie de génomes microbiens offre de nombreuses perspectives pour l'étude d'organismes difficilement modifiables, ou bien pour la construction de nouveaux organismes vivants. Associés à ce cycle, les méthodes de clonage dans la levure, les outils d'ingénierie génomique disponibles chez cet hôte, ainsi que, les voies « de sortie » du génome modifié et transféré vers une cellule receveuse (e.g. transplantation), sont plus amplement décrits dans la suite de cette partie en se concentrant sur leur application à des génomes bactériens.

III.1. Méthodes de clonage de génomes dans la levure

Plusieurs méthodes de clonage dans la levure sont disponibles dans la littérature, pour cloner au choix des fragments de génomes, des génomes entiers, mais également, pour permettre l'assemblage de multiples morceaux d'ADN (amplifiés par PCR ou bien produits par synthèse chimique) en une plus grande molécule²²⁹ (**Figure 13**).

Tableau 2 : Chromosomes synthétiques construits à ce jour. Tableau issu de Baby *et al.* 2019.

| | Génome synthétisé | Taille (kpb) | Référence |
|--------------------------|---|---------------------|------------------|
| Virus | Poliovirus PV1(M) | 7,6 | [14] |
| | Bactériophage PhiX174 | 5,4 | [15] |
| | Virus de la grippe espagnole de 1918 | ~13,0 | [48] |
| | HERV-K(HML-2) | 9,5 | [55] |
| | Bat SARS-like-SCoV | 29,7 | [56] |
| | VIH-1-C | 11,0 | [57] |
| | Bactériophage G4 | 5,6 | [58] |
| | Virus de l'hépatite C sous-type 1a | 9,6 | [59] |
| | Bactériophage S13-like | 5,4 | [60] |
| | CHIKV | 11,9 | [61] |
| | Virus de la mosaïque du tabac | 6,4 | [62] |
| | GALV-NF | 4,7 | [63] |
| | PRRSV | 15,4 | [64] |
| | AcMNPV | 145,3 | [65] |
| | Virus de la variole équine | 212,8 | [49] |
| Procaryotes | <i>Mycoplasma genitalium</i> | 583,0 | [8,10] |
| | <i>Mycoplasma mycoides subsp. capri</i> | 1077,9 | [20] |
| | JCVI-syn3.0 | 531,5 | [3] |
| Eucaryotes et organelles | ADNm de souris | 16,3 | [28] |
| | <i>Saccharomyces cerevisiae chrIII</i> | 272,8 | [33] |
| | <i>Saccharomyces cerevisiae chrII</i> | 770,0 | [34] |
| | <i>Saccharomyces cerevisiae chrX</i> | 707,4 | [37] |
| | <i>Saccharomyces cerevisiae chrXII</i> | 976,0 | [2] |
| | <i>Saccharomyces cerevisiae chrV</i> | 536,0 | [35] |
| | <i>Saccharomyces cerevisiae chrVI</i> | 242,7 | [36] |

Toutes ces méthodes requièrent l'ajout (en amont ou au cours de l'expérience) d'une cassette composée :

- (i) D'un centromère (CEN) obligatoire pour permettre la partition lors de la division cellulaire du chromosome bactérien (ou fragment d'ADN) cloné en tant que molécule surnuméraire dans la levure.
- (ii) D'une origine de réplication (Autonomously Replicating Sequence ou ARS), élément optionnel (ARS nécessaire pour des génomes avec un GC% > 40 ; ARS non nécessaire pour des génomes avec un GC% < 40).
- (iii) D'un marqueur de sélection (complémentation d'auxotrophie).

Cette cassette portant « les éléments levure » permet la réplication et le maintien de la molécule d'intérêt dans la levure. Le choix de la méthode s'effectue en fonction des caractéristiques du génome à cloner (*e.g.* taille, pourcentage en GC de la séquence, présence de sites de restriction), mais aussi des caractéristiques du micro-organisme d'origine (si ce dernier est cultivable ou transformable).

III.1.1. Méthode de clonage dite « Classique »

Cette méthode, décrite par Benders *et al.*, en 2010²¹⁵, a été utilisée pour la première fois pour cloner les génomes bactériens entiers de trois mycoplasmes : *Mycoplasma genitalium*, *Mycoplasma pneumoniae* et *Mycoplasma mycoides* subsp. *capri*. Elle consiste à pré-intégrer la cassette dite « éléments levure » dans le chromosome du micro-organisme que l'on souhaite cloner (**Figure 13A**). Chez les mycoplasmes, l'intégration de cette cassette se fait traditionnellement par transposition^{216,230}. Cette intégration aléatoire ne nécessite pas de connaître la séquence du génome à cloner. En revanche, il est nécessaire d'avoir la capacité de transformer la bactérie avant clonage de son génome dans la levure. Une fois marqué le génome est ensuite extrait intact en bloc d'agarose afin d'éviter au maximum les cassures mécaniques de la molécule d'ADN à cloner, puis directement transformé dans des sphéroplastes de levures (levures dont la paroi a été au préalable digérée de manière enzymatique) grâce à un protocole utilisant le PEG²¹⁵.

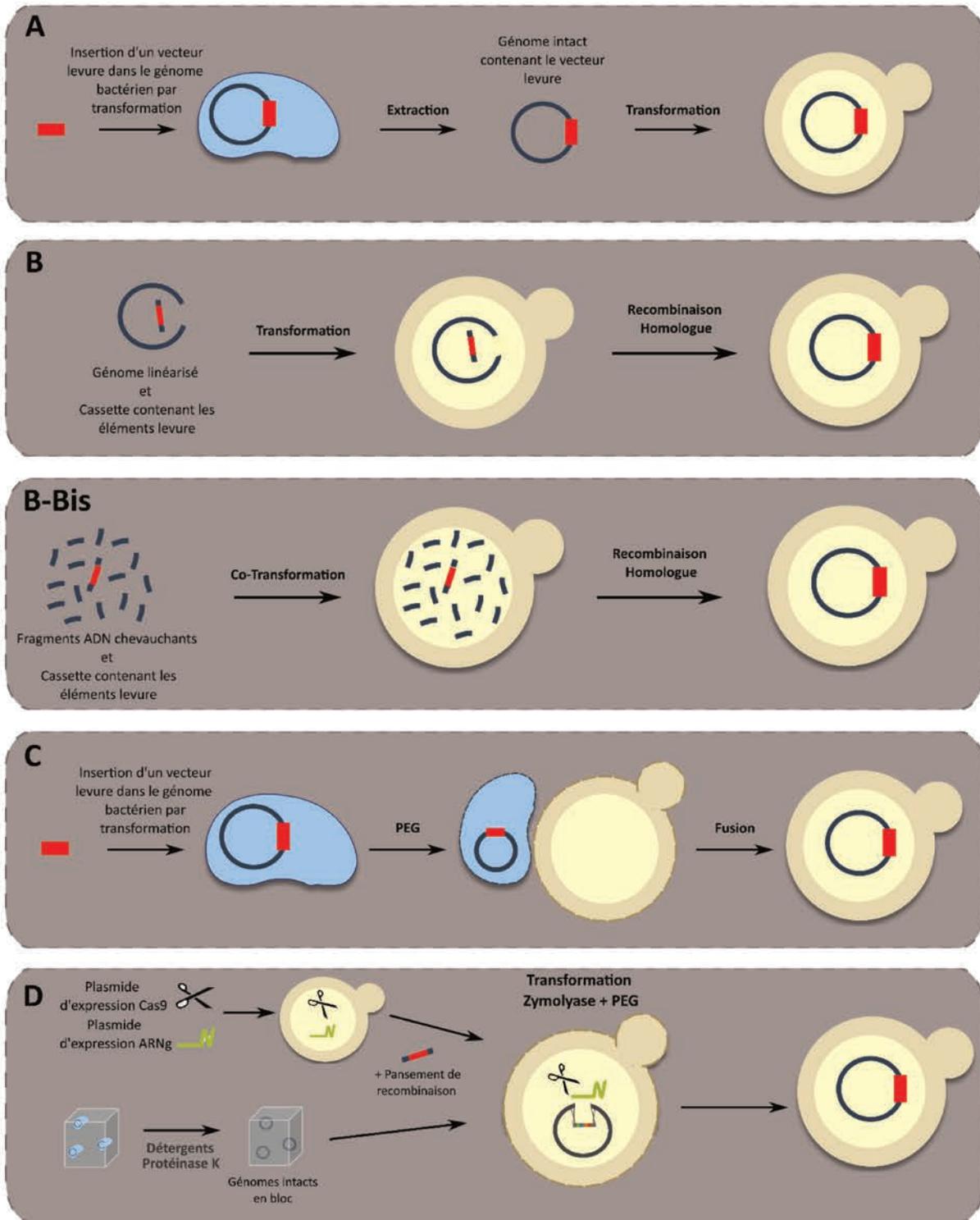


Figure 13 : Schéma illustrant les différentes méthodes de clonage de génome dans la levure. [A] Transformation dans la levure d'un génome bactérien pré-marqué par la méthode Classique. **[B]** Méthode de TAR-Cloning. **[B-Bis]** TAR-Cloning en utilisant de multiples fragments d'ADN chevauchants (naturels ou synthétiques). **[C]** Clonage par la méthode de Fusion. **[D]** Transformation et modification simultanée d'un génome bactérien par CREasPy-Cloning.

III.1.2. Méthode de TAR-Cloning

La méthode de TAR-Cloning est basée sur le mécanisme de recombinaison homologue de la levure en association avec la transformation. C'est une méthode qui permet de cloner des fragments de chromosomes (jusqu'à 300 kpb) sous forme de YAC ou plasmides centromériques comme expliqué plus haut. Initialement, le groupe ayant développé cette méthode s'attachait à cloner de l'ADN génomique (ADNg) humain dans le but de caractériser des *loci* de grande taille, ou bien toxiques chez *E. coli*²³¹. Ils ont également développé des protocoles alternatifs pour maximiser l'efficacité de clonage. Par exemple, ils ont modifié le vecteur TAR correspondant à la cassette portant les éléments génétiques levures (décrite §III.1.), en ajoutant un système de contre sélection avec le marqueur URA3. Brièvement, les extrémités identiques à celles de la région à capturer sont placées entre un promoteur *pADHI* (issu de *Schizosaccharomyces pombe*) et le gène *URA3*. Une fois la séquence cible circularisée avec le vecteur TAR, l'expression du gène *URA3* est abolie et confère à la levure la résistance à l'acide 5-fluoroorootique (5-FOA). Dans le cas contraire, si le vecteur se circularise seul, le gène *URA3* est exprimé et génère des enzymes convertissant le 5-FOA en 5-fluorouracil toxique pour la levure²³². En 2015, ils ont également associé le système CRISPR-Cas9 afin de pré-digérer *in vitro* l'ADNg plutôt que d'avoir recours aux enzymes de restriction. Avec cette amélioration du protocole, ils ont ainsi démontré une augmentation de l'efficacité de capture de fragments d'ADN humain (de 2% à 32%)²³³.

Parallèlement à ces mises au point, des groupes se sont intéressés à la méthode du TAR-Cloning dans le but de cloner des génomes bactériens entiers dans la levure. A ce jour, quatre génomes bactériens ont été clonés en utilisant cette stratégie : *Mycoplasma genitalium*, *Mycoplasma hominis*, *Acholeplasma laidlawii*, et *Prochlorococcus marinus*^{215,221,224,234}. Le clonage s'effectue en co-transformant la levure avec le génome linéarisé par une enzyme de restriction et un vecteur TAR spécifique, conçu de manière à ce que ses extrémités (60 pb) soient identiques à celles bordant le site de clivage sur le génome bactérien. Une fois dans la levure, la molécule à capturer est assemblée au vecteur TAR par recombinaison homologue (**Figure 13B**). La même stratégie permet également l'assemblage de multiples fragments chevauchants synthétiques (**Figure 13B-Bis**), en co-transformant plusieurs fragments avec le vecteur TAR.

III.1.3. Méthode de Fusion cellulaire

Cette méthode de clonage originale et décrite dans la littérature par Karas *et al*, repose sur un phénomène de fusion cellulaire, permettant le transfert du chromosome bactérien depuis la bactérie vers la levure (**Figure 13C**). À la différence des autres méthodes (Classique et TAR-Cloning), cette stratégie ne nécessite pas d'isoler le génome bactérien, via une étape de purification en utilisant une matrice d'agarose (bloc d'agarose), évitant ainsi des étapes supplémentaires de manipulation tout en prévenant d'éventuelles fragmentations mécaniques^{235,236}. Des travaux supplémentaires de cette même équipe, précisent également que ce procédé de transfert s'effectuerait, soit par endocytose, soit par fusion des membranes et déversement du contenu bactérien dans la cellule levure²³⁷. Grâce à cette méthode, les génomes de *Mycoplasma mycoides* subsp. *capri*, de *Mycoplasma capricolum* subsp. *capricolum*, et de *Haemophilus influenzae* ont été clonés. Tout comme pour la méthode Classique, cette dernière requiert la capacité de transformer la bactérie avant fusion pour intégrer les éléments génétiques levures nécessaires à sa propagation (*cf.* §III.1.1.). Une fois le génome pré-marqué, les cellules bactériennes sont directement fusionnées avec des sphéroplastes de levures (dont la paroi a été au préalable digérée de manière enzymatique), grâce à un protocole au PEG²¹⁵. Dans le cas particulier de *H. influenzae*, la paroi bactérienne Gram- (fine couche de peptidoglycane et avec une membrane externe), est également digérée par du lysozyme pour favoriser la fusion avec la levure. Ce traitement n'est pas nécessaire pour les mycoplasmes qui ne possèdent pas de paroi.

III.1.4. Méthode de clonage par CREasPy-Cloning

La méthode de CREasPy-Cloning développée par Ruiz *et al.*, en 2019 allie le système d'édition de génome CRISPR-Cas9 au protocole de clonage de génome dans la levure (**Figure 13D**). Elle a notamment été développée en réponse à un besoin de s'affranchir de l'étape de pré-transformation des éléments levures dans le chromosome bactérien (comme pour les méthodes Classique et de Fusion), mais également de l'utilisation d'enzymes de restriction pour linéariser le génome (TAR-Cloning). Les génomes de *Mycoplasma pneumoniae*, *Mycoplasma leachii*, *Mycoplasma mycoides* subsp. *mycoides* et de *Mycoplasma feriruminatoris* ont été clonés et également modifiés grâce à cette méthode^{220,238}. En effet,

cette stratégie consiste à transformer et simultanément modifier le génome bactérien donneur selon le schéma suivant : (i) la levure est pré-transformée avec un plasmide d'expression de l'endonuclease Cas9 (pCas9), et avec un plasmide d'expression d'un ARN guide (pARNg). Cet ARNg chimérique est composé de la fusion du crARN (séquence spécifique du gène cible) et du tracrARN (composant structural)^{111,239} ; ensuite, (ii) les levures pré-chargées sont co-transformées avec le génome bactérien purifié en bloc d'agarose, et un pansement de recombinaison portant la cassette « éléments levures » et un marqueur de sélection, encadrés par deux régions identiques au *locus* ciblé par le complexe Cas9-ARNg ; par la suite, (iii) ce complexe détecte le site de clivage et génère une coupure double brin qui linéarise le chromosome co-transformé et stimule le mécanisme de réparation par recombinaison homologue chez la levure ; enfin (iv) le génome ainsi réparé est circularisé avec la cassette « éléments levures » intégrée à la place du *locus* ciblé. Cette méthode n'implique donc pas une insertion aléatoire des éléments levures mais bien de manière définie et permet la délétion d'un *locus* donné dès l'étape de clonage. De plus, il est également possible de multiplier les délétions (jusqu'à trois à ce jour) en transformant des levures pré-chargées avec le pCas9 et un pARNg construit de manière à exprimer deux ou trois ARNg en même temps. La réparation de chaque coupure induite par les complexes Cas9-ARNg (1, 2 ou 3) s'effectue à ce moment-là avec les pansements adéquats fournis (un pansement portant les éléments levures et un ou deux pansements correspondants à des séquences d'ADN de 90 pb)²²⁰.

III.2. Outils d'ingénierie génomique chez la levure

Il existe de nombreux outils d'édition chez la levure, en revanche ils ne sont pas tous adaptés à l'ingénierie de génomes bactériens clonés²⁴⁰. De nouveaux outils ont donc été spécifiquement créés pour éditer les génomes bactériens dans la levure. Dans cette section sont décrits les outils les plus utilisés pour cette étape : les méthodes TREC et TREC-IN, mais aussi une méthode basée sur l'outil d'édition de génome CRISPR-Cas9.

III.2.1. Méthodes TREC et TREC-IN

La méthode TREC (« Tandem Repeat coupled with Endonuclease Cleavage ») a été spécialement développée pour l'ingénierie de génomes clonés dans la levure²⁴¹. Cette

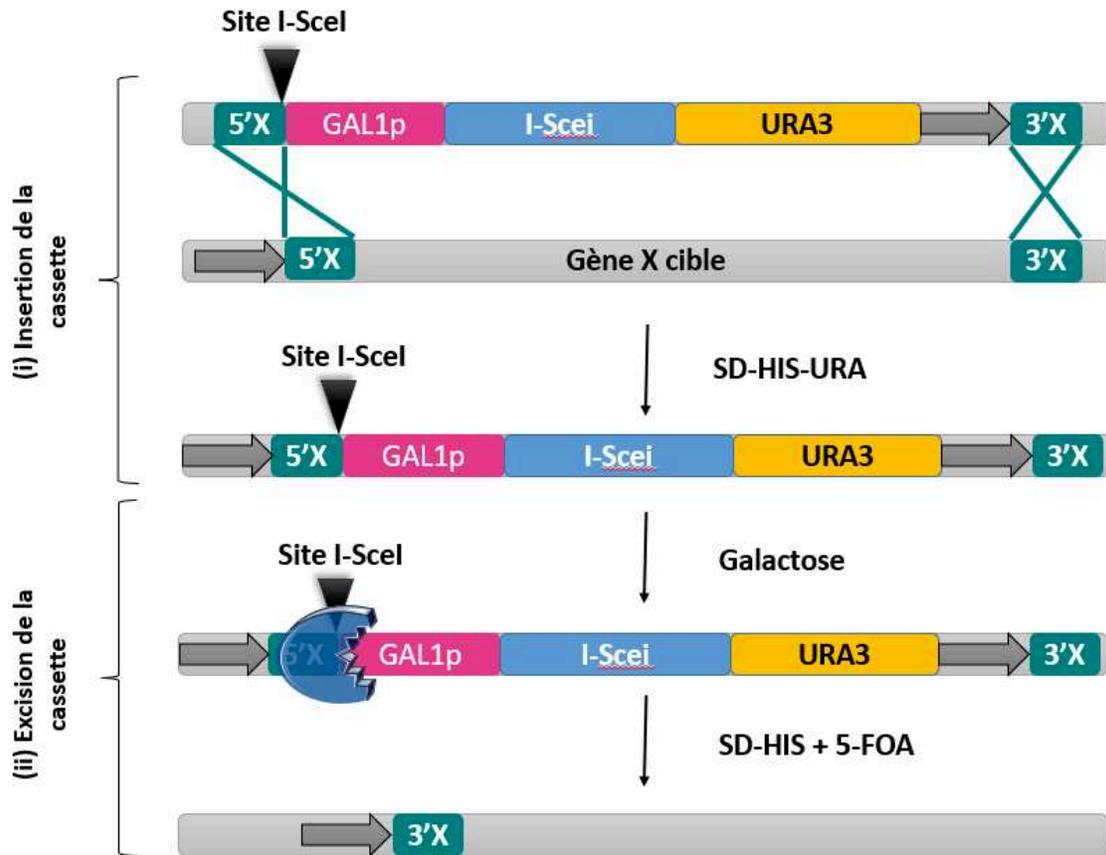


Figure 14 : Schéma illustrant la méthode TREC. La cassette de mutagenèse est constituée d'une région CORE (contenant le site de reconnaissance de l'enzyme I-SceI de 18 pb, le gène *I-sceI* (en bleu) sous le contrôle du promoteur GAL1 (en rose), et le marqueur d'auxotrophie URA3 (en jaune)) et d'un fragment identique à la région située en amont du gène cible (indiqué par une flèche grise). La recombinaison homologue génère l'apparition de deux séquences répétées en tandem encadrant la région CORE (flèches grises). L'ajout de galactose induit l'expression du gène *sceI* et la production de l'endonuclease I-SceI, générant une coupure double brin au niveau du site I-SceI. Cette cassure double brin provoque des événements de recombinaison homologue intramoléculaires entre les séquences répétées, conduisant à l'excision de la cassette sans laisser de trace.

technique permet de supprimer un ou des gènes cibles sans laisser de trace. Elle repose sur deux évènements clés (**Figure 14**) :

- (i) La région cible est remplacée par une cassette constituée d'un site de reconnaissance de l'enzyme I-SceI, du gène *sceI* sous le contrôle du promoteur GAL1, du marqueur de sélection URA3 et d'un *locus* homologue de ~300 pb en amont (5') du gène cible.
- (ii) L'ajout de galactose induit l'expression d'I-SceI qui génère une coupure double brin sur son site de reconnaissance. Cette cassure va stimuler la recombinaison homologue entre les séquences répétées obtenues au cours de la primo-intégration : la cassette de mutagénèse est alors proprement excisée.

La méthode dérivée TREC-IN (« TREC-assisted gene knock-in ») permet d'insérer ou de remplacer un ou des gènes cibles sans laisser de trace au site d'insertion. Elle suit approximativement le même principe que la méthode TREC mais inclut une étape supplémentaire qui permet la restauration du marqueur de sélection kanMX. La cassette TREC-IN porte de plus un gène hétérologue et une région identique à la séquence en amont du gène cible d'environ 50 pb.²⁴².

Chez les mycoplasmes, la méthode TREC a permis de générer plusieurs mutants déficients pour une ou plusieurs fonctions (*e.g. Mcap Δ trmfo, Mmc Δ glf, Mmc Δ 68*)²⁴³⁻²⁴⁵. La méthode TREC-IN a quant à elle, permis de générer de manière dirigée le premier mycoplasme thermosensible (*MmcTS+*)²⁴⁶ et d'aider à la construction de souche *Mmc* avec un génome réduit¹⁷².

III.2.2. Méthode dérivée du système CRISPR-Cas9

Certaines bactéries et archées possèdent un système de défense adaptatif appelé CRISPR (Cluster Regularly Interspaced Short Palindromic Repeats), en association avec une endonucléase Cas (CRISPR-associated)²⁴⁷. Ce système protège les cellules d'une invasion par des éléments mobiles exogènes (phages et plasmides). Il est composé d'une région alternant des séquences répétées (R) et des séquences uniques appelées « spacer » (S1, S2, S3...), et d'une région codant des gènes *cas* impliqués dans des fonctions hélicases et endonucléases

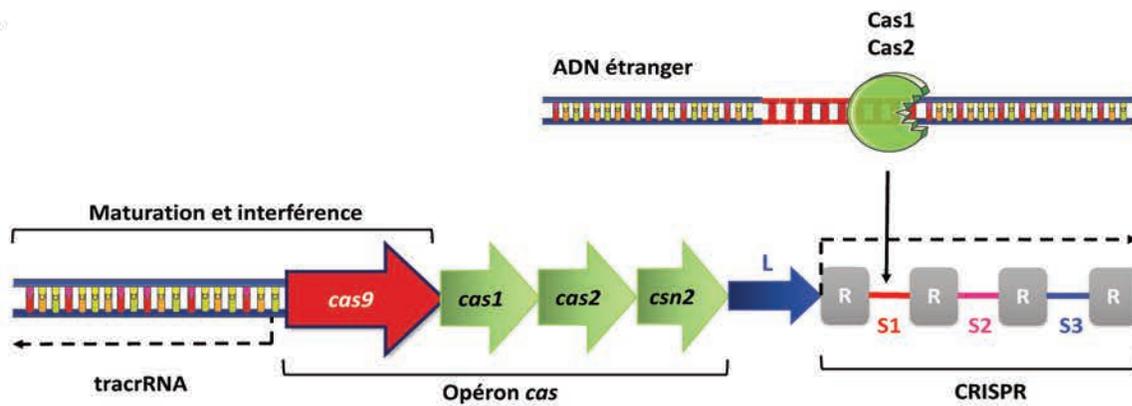
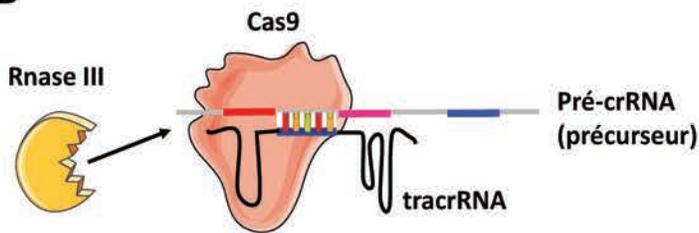
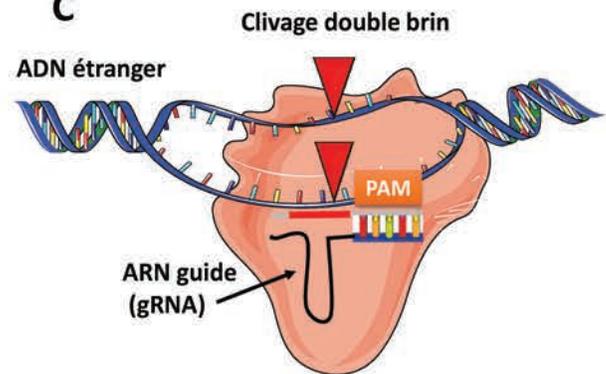
A**B****C**

Figure 15 : Schéma du système CRISPR-Cas9 de type II. Adapté d'après Sorek *et al.* 2013 et Doudna & Charpentier. 2014. **[A]** Etape d'immunisation contre l'ADN étranger. **[B]** Biogénèse du crRNA (CRISPR RNA). **[C]** Clivage de l'ADN cible : le duplex crRNA-tracrRNA s'hybride avec sa cible et après recrutement de l'endonucléase Cas9, les deux brins d'ADN sont clivés de manière site spécifique.

(*cas9, cas1, cas2, csn2*) (**Figure 15**). Ce système est un mécanisme d'interférence qui requiert trois étapes²⁴⁸ : (i) L'étape d'immunisation contre l'ADN étranger, c'est une étape d'adaptation permettant l'acquisition de nouvelles séquences « spacers » ; (ii) La biogénèse d'un crRNA (CRISPR RNA) : le pré-crRNA, un long transcrit primaire est généré à partir du *locus* CRISPR, s'apparie avec un tracrRNA (trans-activating crRNA) au niveau des répétitions, déterminant ainsi les zones de clivage pour la RNase III (ribonucléase double brin) ; (iii) Le clivage de l'ADN cible : le duplex crRNA-tracrRNA s'hybride avec sa cible et après recrutement de l'endonucléase Cas9, les deux brins d'ADN sont clivés de manière site spécifique. Ce complexe crRNA-tracrRNA-Cas9 reconnaît la séquence cible via le crRNA dont la séquence est complémentaire, puis la Cas9 et le dinucléotide GG de la séquence PAM (Protospacer Adjacent Motif). La coupure double-brin produite laisse des extrémités libres qui vont stimuler les mécanismes de réparation de l'ADN (HR : recombinaison homologue ou NHEJ : Non Homologous End-Joining (Jonctions d'extrémités non-homologues)).

Le potentiel du système CRISPR-Cas9 en termes d'édition de génomes est tel que le système a été détourné pour en faire un véritable outil biotechnologique applicable aussi bien chez la bactérie que chez la levure, les plantes ou l'homme. Il a notamment été montré que le système à deux composantes (crRNA:tracrRNA), utilisé pour recruter Cas9, pouvait être remplacé par un ARN guide chimère (sgRNA) simplifié mais gardant les mêmes spécificités et la même structure que le duplex initial (la partie double brin permettant la fixation de Cas9 et la partie simple brin permettant l'appariement à la séquence cible)¹¹¹.

Suite aux travaux de Dicarlo *et al.* en 2013 qui ont adapté la technologie CRISPR-Cas9 chez la levure, une méthode d'ingénierie de génome cloné chez la levure basée sur le système CRISPR-Cas9 a été proposée. Dans cet exemple, les auteurs ont transformé les différents plasmides codant pour l'endonucléase Cas9 et l'ARN guide, ainsi qu'un pansement de recombinaison, dans des levures propageant le génome de *Mycoplasma mycoides* subsp. *capri* GM12. Ils ont ainsi créé une délétion sans trace du gène *glpO* codant pour l'oxydase glycerol-3-phosphate²⁴⁹. Cette méthode a été reprise pour modifier le génome cloné de *Mycoplasma hominis* dans la levure, en ciblant le gène *vaa* codant pour une adhésine putative chez ce mycoplasme pathogène opportuniste chez l'humain²³⁴. Elle a dernièrement été utilisée pour l'ingénierie successive du génome de *Mycoplasma feriruminatoris* (*Mferi*), dans le but de supprimer des gènes codant des facteurs de virulence potentiels (*glpOKF, gtsABCD* et les *loci*

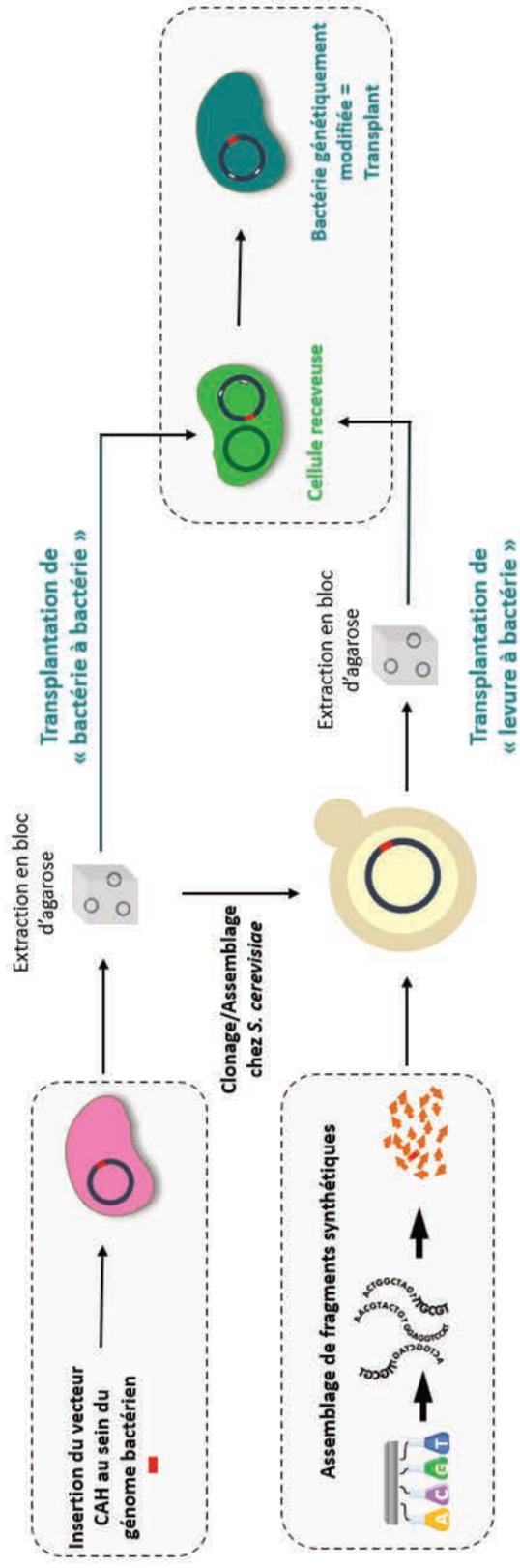


Figure 16 : Schéma de la transplantation de génome (TG). Adapté d'après Lartigue *et al.* 2007 et 2009, et Labrousseau *et al.* 2017. Les cellules du génome du donneur sont d'abord transformées pour intégrer les éléments levures et un marqueur de sélection à un antibiotique. Les génomes nouvellement marqués et isolés en bloc d'agarose sont soit, directement transplantés dans une cellule receveuse correspondant à la « Transplantation de bactérie à bactérie » ; soit transformés dans la levure pour être modifiés puis transplantés dans une cellule receveuse correspondant à la « Transplantation de levure à bactérie ». Alternativement un génome synthétique peut-être assemblé dans la levure puis transplanté dans une cellule receveuse.

MIB-MIP) et de produire un châssis *Mferi* non virulent (perte de 6% du génome). Des gènes d'une autre espèce de mycoplasme pathogène de caprins (*Mycoplasma capricolum* subsp. *capripneumoniae*) ont ensuite été ajoutés afin de construire un châssis bactérien pouvant être utilisé pour présenter des antigènes et stimuler le système immunitaire de l'hôte²⁵⁰.

Avec une approche similaire, Kannan *et al.* ont modifié la petite sous-unité de l'ARN 16S de *M. mycoides* une fois son génome cloné dans la levure. Au cours de cette étude l'endonucléase Cas9 exprimée constitutivement chez la levure (car intégrée dans le chromosome) est guidée par un ARNg au préalable transcrit de manière *in vitro*. A la différence des études de Dicarlo *et al.* et de Tsarmopoulos *et al.* les auteurs ont éliminé l'utilisation de deux plasmides pour exprimer le systèmes CRISPR-Cas9²⁵¹.

III.3. Voies de sortie du cycle d'ingénierie de génome microbien

Que ce soit pour des génomes bactériens ou viraux, une fois assemblés et/ou clonés, puis modifiés dans la levure, il est important de pouvoir relocaliser la molécule d'ADN éditée (fragment ou chromosome entier) vers un compartiment propice à l'expression de la séquence encodée. Dans la plupart des cas, les virus peuvent être reconstitués simplement en transfectant directement (ou bien après transcription *in vitro* pour les virus à ARN) leurs génomes modifiés ou par fusion dans des cellules hôtes^{229,252}. Pour les génomes bactériens le processus est en revanche plus complexe. Néanmoins plusieurs méthodes sont aujourd'hui référencées pour transférer un génome entier ou un fragment de génome depuis la levure vers une cellule bactérienne receveuse.

III.3.1. Transplantation de génome (TG)

La troisième étape du cycle d'ingénierie de génome présenté **Figure 16**, consiste à extraire le génome bactérien cloné et modifié dans la levure afin de le réintroduire dans une cellule receveuse. Le génome donneur (naturel ou synthétique) transféré confère un nouveau génotype et donc phénotype à la cellule, entraînant l'apparition de bactéries génétiquement modifiées, appelées « transplants ».

Cette étape est nommée transplantation de génomes ou TG, et a été développée spécifiquement pour l'étude de génomes de mycoplasmes après clonage dans la levure par un groupe du JCVI. Dans un premier temps, ce procédé a été réalisé pour transplanter le génome de *Mycoplasma mycoides* subsp. *capri* (cellule donneuse *Mmc*) dans *Mycoplasma capricolum* subsp. *capricolum* (cellule receveuse *Mcap*) illustrant la transplantation dite de « bactérie à bactérie »²¹⁷. Puis ces mêmes auteurs ont ensuite montré en 2009, qu'il était possible de transplanter dans la cellule receveuse *Mcap*, un génome *Mmc* préalablement cloné et modifié dans la levure²¹⁶. L'année suivante sa version synthétique *Mmc* JCVI-Syn1.0 assemblée dans la levure a été transplantée avec succès, générant ainsi les premières bactéries contrôlées par un génome artificiel¹⁶⁴. Cette approche a permis, 6 ans plus tard, de construire puis transplanter une version synthétique et réduite *Mmc* JCVI-Syn3.0 démontrant pour la première fois la production de cellules vivantes quasi-minimales contrôlées par un génome synthétique arborant un ensemble de gènes défini comme quasi-essentiel (473 gènes pour 531 kpb)¹⁷² (**Figure 16**). Au cours de ces expériences de transplantations de génomes dites « levure-bactérie », les auteurs ont mis en évidence qu'il est nécessaire d'inactiver les systèmes de restriction-modification présents dans la cellule receveuse, pour éviter que le génome donneur ne soit dégradé. Ils ont également démontré qu'il était possible de méthyliser au préalable le génome donneur (via des extraits cellulaires de la cellule receveuse ou bien des méthyltransférases purifiées) dans le but de le protéger²¹⁶.

D'autres travaux ont souligné que la distance phylogénétique entre le génome donneur et la cellule receveuse impacte fortement la réussite d'une telle méthode. Labroussaa *et al.* ont démontré que plus le génome donneur est proche de celui de la receveuse, plus l'efficacité de transplantation est élevée. Ceci s'explique notamment par le fait que la bactérie receveuse doit être capable de transcrire et de traduire les gènes du génome entrant avec une machinerie moléculaire compatible. Leur étude indique notamment que si le pourcentage de similarité du « core-protéome » est supérieur à 90%, il est fort probable que l'initialisation de la lecture du génome donneur se fasse alors correctement²³⁰. L'exemple particulier de *Mycoplasma mycoides* subsp. *mycoides* proche phylogénétiquement de *Mcap* complexifie en revanche cette règle. Bien que très proche de *Mcap* d'un point de vue phylogénétique, son génome n'a, à ce jour, jamais pu être transplanté depuis la levure vers la cellule receveuse *Mcap*. Ce mécanisme de transplantation est encore peu compris et les

données de la littérature suggèrent que des facteurs (spécifiques d'espèces) pourraient influencer son efficacité ou sa faisabilité, par exemple, des nucléases, des facteurs du soi et du non-soi (e.g. systèmes de restriction-modification) ou encore, des séquences génétiques mobiles (e.g. séquences d'insertion IS)²⁵³ ; cette liste n'est pas exhaustive.

La TG est pour le moment exclusivement applicable à quelques espèces du genre *Mycoplasma*²⁵³, avec l'exception de *Mesoplasma florum*, une espèce de Mollicutes pour laquelle la TG a aussi été effectuée avec succès²²².

III.3.2. Méthodes de transfert de portion de génome

Des alternatives à la transplantation de génome ont été développées pour d'autres bactéries. Ces approches suivent le même schéma, à savoir cloner de grands fragments génomiques dans la levure pour les modifier puis à les réintégrer dans les génomes bactériens « natifs ».

Chez *Escherichia coli*, des travaux ont montré l'assemblage de fragments synthétiques de 100 kpb chez la levure, intégrés ensuite dans le génome original par remplacement successif de *locus* via les méthodes REXER (Replicon EXcision Enhanced Recombination) et GENESIS (GENomE Stepwise Interchange Synthesis). Grâce à ce procédé, Fredens *et al.* ont généré une souche d'*E. coli* pour laquelle le code génétique a été redéfini ; elle utilise désormais 61 codons pour la synthèse protéique au lieu des 64 codons standards¹⁹⁹.

Une autre approche a été utilisée chez *Salmonella typhimurium* dans laquelle des fragments de génomes synthétiques et dessinés *in silico* (de 2 à 4 kpb), sont d'abord assemblés dans la levure formant des fragments plus grands (de 10 à 25 kpb). Ces derniers ont ensuite été linéarisés et amplifiés par la méthode d'amplification par cercle roulant puis intégrés dans une souche recombinante de *S. typhimurium* exprimant le système de recombinaison site spécifique « lambda Red ». Ce système issu du phage lambda a été détourné comme outil de recombinaison et utilise de l'ADN linéaire simple brin ou double brin (servant de substrat). L'expression des gènes *gam*, *exo* et *beta* sont régulés via un promoteur inductible permettant d'initier la recombinaison à un temps donné. Plusieurs remplacements effectués de manière itérative par une méthode nommée SIRCAS (Stepwise Integration of Rolling Circle Amplified

Segment), ont permis d'obtenir *in fine* une souche *S. typhimurium* dont le génome avait été recodé²⁵⁴.

Pour finir une méthode récente proposée comme alternative à la TG a été développée chez *Mycoplasma pneumoniae* par Garcia-Morales *et al.* en 2020 et s'appuie sur le système de recombinaison Cre-lox²⁵⁵. Nommée RAGE pour Recombinase-Associated Genomic Exchange, cette stratégie est inspirée d'autres travaux (RAGE et CRAGE décrit par Santos *et al.* et Wang *et al.*²⁵⁶⁻²⁵⁸) et se décline en trois étapes: (i) la levure est utilisée comme plateforme d'assemblage de fragments de génome et associe le *locus* à capturer à des « landing pad » constitués de sites loxP incompatibles entre eux (lox66 et loxm2/71 situés de part et d'autre du locus), un vecteur YAC-BAC et le gène codant la recombinase Cre ; (ii) la molécule d'ADN construite est transférée chez *E. coli*, amplifiée grâce à un système de copie contrôle puis purifiée ; (iii) elle est ensuite transformée dans une cellule receveuse au préalable modifiée par intégration de deux sites loxP (lox71 et loxm2/66) incompatibles entre eux, mais compatibles avec les sites présents sur la molécule donneuse à intégrer. En utilisant cette stratégie, Garcia-Morales *et al.* ont montré qu'il était possible d'insérer un fragment de 15 kpb à un *locus* spécifique du génome de *M. pneumoniae* mais aussi de remplacer un *locus* de 38 kpb soit par une version modifiée (dans la levure ou dans *E. coli*) soit par une version synthétique et réduite construite *in vitro*. C'est ainsi qu'une souche de *M. pneumoniae* débarrassée de 13 gènes non essentiels a pu être générée et propagée *in vitro*.

MATÉRIELS ET MÉTHODES

MATÉRIELS ET MÉTHODES

I. Souches microbiennes et conditions de culture

I.1. *Escherichia coli*

Plusieurs souches commerciales d'*Escherichia coli* ont été utilisées afin de mener à bien les expérimentations de biologie moléculaire (propagation et amplification de plasmides).

- *E. coli* DH10B [F- *mcrA* Δ (*mrr-hsdRMS-mcrBC*) Φ 80*lacZ* Δ M15 Δ *lacX74* *recA1 endA1 araD139* Δ (*ara-leu*)7697 *galU galk* λ - *rpsL* (Str^R) *nupG*] (initialement fournie par Thermo Fisher, puis propagée au laboratoire).
- *E. coli* NEB[®]5- α [*fhuA2* (*argF-lacZ*)U169 *phoA glnV44* Φ 80 (*lacZ*)M15 *gyrA96 recA1 relA1 endA1 thi-1 hsdR17*] (fournie par NEB).
- *E. coli* NEB[®]10- β [Δ (*ara-leu*) 7697 *araD139 fhuA* Δ *lacX74 galk16 galE15 e14-* (Φ 80 Δ *lac* Δ M15) *recA1 relA1 endA1 nupG rpsL* (Str^R) *rph spoT1* Δ (*mrr-hsdRMS-mcrBC*)] (fournie par NEB).
- *E. coli* Stellar[™] [F-, *endA1, supE44, thi-1, recA1, relA1, gyrA96, phoA, Φ 80d lacZ* Δ M15, Δ (*lacZYA-argF*) U169, Δ (*mrr-hsdRMS-mcrBC*), Δ *mcrA*, λ -] (fournie par Clontech-Takara).
- *E. coli* ElectroMAX[™] Stbl4[™] [*mcrA* Δ (*mcrBC-hsdRMS-mrr*) *recA1 endA1 gyrA96 gal thi-1 supE44* λ -*relA1* Δ (*lac-proAB*)/F' *proAB⁺lacI^qZ* Δ M15 *Tn10* (Tet^R)] (fournie par Invitrogen). Cette souche est utilisée pour cloner des fragments d'ADN instables. Au cours de ce travail, elle a été utilisée pour cloner des gènes ou *locus* du génome de *B. subtilis* connus pour être toxiques ou instables dans des *E. coli* plus classiques comme celles citées ci-dessus.

Ces souches ont été cultivées en milieu LB à 30°C (*E. coli* ElectroMAX[™] Stbl4[™]) ou à 37°C sous agitation (190 rpm). En fonction de la sélection à appliquer, les concentrations en antibiotique qui ont été utilisées sont les suivantes : ampicilline 100 μ g/mL, spectinomycine 50 μ g/mL, chloramphénicol 6 μ g/mL et érythromycine 200 μ g/mL.

Tableau 3 : Composition du milieu SP5

| Base autoclavée (20min à 120°C) | |
|---|------------|
| Mycoplasma Broth Base (Fisher Scientific) | 3,5 g |
| Bacto Tryptone (Difco) | 10 g |
| Bacto Peptone (Difco) | 5,3 g |
| Agar noble (milieu solide) | 10 g |
| pH | 7,5 |
| Qsp H ₂ O | 600 mL |
| Complément filtré (0,45 µm) | |
| Sérum de veau foetal (Eurobio, décomplémenté 1h à 56°C) | 170mL |
| Yeastolate 2% (m/v) | 100 mL |
| CMRL 1066 10X (Gibco) | 50 mL |
| Extrait de levure (Gibco) | 35 mL |
| Glucose 20% (m/v) | 25 mL |
| Bicarbonate de sodium 7,5% (m/v) | 14,6 mL |
| Glutamine 200 mM | 5 mL |
| Rouge de phénol 1% (m/v) | 2 mL |
| Ampicilline 10 mg/mL | 1 mL |
| Volume final | 1 L |

Tableau 4 : Composition du milieu Hayflick modifié (Mccp ou Mgal).

†Dans le cas de Mccp l'extrait de levure est préparé frais.

*Le pyruvate de sodium est ajouté au complément seulement pour Mccp.

| Base autoclavée (20min à 120°C) | |
|--|------------|
| PPLO Broth (Difco) | 21 g |
| Glucose (Sigma) | 5 g |
| Agar noble (milieu solide) | 10 g |
| pH | 7,6-7,8 |
| Qsp H ₂ O | 695 mL |
| Complément filtré (0,45 µm) | |
| Extrait de levure (Gibco)† | 100 mL |
| Pyruvate de sodium* | 10 g |
| Sérum de cheval (Gibco, décomplémenté 1h à 56°C) | 200 mL |
| Rouge de phénol 1% (m/v) | 2 mL |
| Ampicilline 100mg/mL | 1 mL |
| Volume final | 1 L |

Tableau 5 : Composition du milieu SP4-Bovis

| Base autoclavée (20min à 120°C) | |
|---|------------|
| PPLO Broth (Difco) | 3,5 g |
| Bacto Tryptone (Difco) | 10 g |
| Bacto Peptone (Difco) | 5,325 g |
| Agar noble (milieu solide) | 9 g |
| pH | 7,8 |
| Qsp H ₂ O | 626 mL |
| Complément filtré (0,45 µm) | |
| Sérum de veau foetal (Eurobio, décomplémenté 1h à 56°C) | 170 mL |
| CMRL 1066 10X (Gibco) | 50 mL |
| Yeastolate 2% (m/v) | 100 mL |
| Glucose 20% (m/v) | 25 mL |
| Rouge de phénol 1% (m/v) | 2 mL |
| Ampicilline 10 mg/mL | 1 mL |
| Volume final | 1 L |

1.2. Les mycoplasmes

Différentes espèces de mycoplasmes ont été utilisées afin de développer la méthode de clonage de génomes CReasPy-Fusion.

Mycoplasma capricolum subsp. *capricolum* (*Mcap*), *Mycoplasma mycoides* subsp. *capri* (*Mmc*) et *Mycoplasma mycoides* subsp. *mycoides* (*Mmm*) ont été cultivés dans du milieu SP5²³⁰ (**Tableau 3**) à 37°C sans agitation. Le suivi de la croissance est effectué par mesure du pH et par l'observation d'un virage coloré du rouge au jaune grâce au rouge de phénol présent dans le milieu (acidification du milieu).

Mycoplasma capricolum subsp. *capripneumoniae* (*Mccp*) a été cultivé en milieu Hayflick modifié (m-Hayflick)²⁵⁹ (**Tableau 4**) à 37°C sans agitation. Le suivi de la croissance est effectué par visualisation de la turbidité de la culture.

Mycoplasma gallisepticum (*Mgal*) a été cultivé en milieu Hayflick modifié²⁵⁹ (**Tableau 4**) à 37°C sans agitation. Le suivi de la croissance est effectué par mesure du pH et par l'observation d'un virage coloré du rouge au jaune grâce au rouge de phénol présent dans le milieu (acidification du milieu).

Mycoplasma bovis (*Mbov*) et *Mycoplasma agalactiae* (*Maga*) ont été cultivés en milieu SP4-Bovis²⁵⁹ (**Tableau 5**) à 37°C sans agitation. Le suivi de la croissance est effectué par mesure du pH et par l'observation d'un virage coloré du rouge au jaune grâce au rouge de phénol présent dans le milieu (acidification du milieu).

1.3. *Bacillus subtilis*

Quatre souches de *Bacillus subtilis* ont été utilisées au cours de ma thèse. Ces souches nous ont été données par nos collaborateurs de l'équipe Syber, unité Micalis, INRAE dans le cadre du projet *Bacillus 2.0*. Elles ont été propagées au laboratoire dans les milieux adéquats et stockées à -80°C.

- *B. subtilis* souche 168 qui est la souche de référence (ATCC 6051). La séquence du génome est annotée et disponible dans la base de donnée Subtiwiki^{64,65}

(<http://subtiwiki.uni-goettingen.de/v4/>). Le génome de cette souche a une taille de 4,214 Mb.

- *B. subtilis* souche Master Strain (MS) [*trpC2 ΔICEBs1, ΔPBSX, ΔSPB, Δskin, Δ(gtaB-yvzH,ggaB,ggaA,yvzI,yvzE), Δupp::[lambdaPr-neo]*]. Cette souche dérive de la souche de *B. subtilis* 168 par réduction itérative du génome¹⁸⁴. Le génome de cette souche a une taille de 3,981 Mb.
- *B. subtilis* souche FE220 [*trpC2 ΔICEBs1, ΔPBSX, ΔSPB, Δskin, Δ(gtaB-yvzH,ggaB,ggaA,yvzI,yvzE), Δupp::[lambdaPr-neo], Δ(sboA-albG), Δ(ydzT-ydjC)::[cl-*upp*-Phleo]*]. Cette souche dérive de la souche MS et se différencie par l'absence du prophage 3 (région entre les gènes *ydzT* et *ydjC*) qui contient les gènes *ydiO* et *ydiP* codant le système de restriction modification BsuM. Le génome de cette souche a une taille de 3,964 Mb.
- *B. subtilis* souche MGP 192 (Minimal Genome Project), cette souche dite « minimale » dérive de *B. subtilis* souche 168 par réduction itérative du génome¹⁸⁴. Le génome de cette souche a une taille de 2,868 Mb.

Ces différentes souches ont été cultivées en milieu LB à 37°C sous agitation (190-200 rpm). En fonction du marqueur de sélection abrité par ces souches, les concentrations en antibiotique qui ont été utilisées pour la culture sont les suivantes : spectinomycine 100 µg/mL, chloramphénicol 5 µg/mL, érythromycine 1 µg/mL puis 5 µg/mL (pour les repiquages) et tétracycline 7,5 µg/mL.

1.4. *Saccharomyces cerevisiae*

Les souches de levures utilisées au cours des expériences de clonage de plasmides ou de génomes entiers dans la levure sont des souches importées du laboratoire de N. Kouprina (NIH, Bethesda, USA) au J. Craig Venter Institute (Rockville, USA). Elles ont ensuite été rapportées par C. Lartigue au laboratoire.

- *S. cerevisiae* souche VL6-48 (ATCC MYA-3666) caractérisée par le génotype [*MATα his3-Δ200 trp1-Δ1 ura3-52 lys2 ade2-1 met14 cir⁰*].

- *S. cerevisiae* souche VL6-48N [*MAT α* , *his3- Δ 200*, *trp1- Δ 1*, *ura3- Δ 1*, *lys2*, *ade2-101*, *met14*, *cir^o*]. Cette souche a été produite à partir de la souche VL6-48 par substitution du gène *URA3-52* par une cassette KanMX.
- *S. cerevisiae* souche W303a (ATCC MYA-151) caractérisée par le génotype [*MAT α* {*leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*} [*phi⁺*]].

Ces souches de levures ont été cultivées en milieu riche YPDA (Takara Bio - Clontech) à 30°C sous agitation (190 rpm). Les clones levures hébergeant le plasmide p414-TEF-pcas9-cyc1 seul ou en combinaison avec le plasmide p426-SNR52p-gRNA-AarI-SUP4t (version URA3 ou ADE2), ont été respectivement cultivés en milieu synthétique minimum (Minimal SD Base Takara Bio - Clontech) sans tryptophane (SD-TRP), sans tryptophane et sans uracile (SD-TRP-URA), ou encore sans tryptophane et sans adénine (SD-TRP-ADE) à 30°C sous agitation. Les clones hébergeant un génome de mycoplasme ou bien des fragments de génome de *B. subtilis* ont, quant à eux, été cultivés en milieu sans histidine (SD-HIS) pour la souche W303a ou sans histidine et sans leucine (SD-HIS-LEU) pour les souches VL6-48 et VL6-48N. Pour certaines expériences, elles ont été cultivées en milieu sans histidine et sans lysine (SD-HIS-LYS ; souche W303a) ou sans histidine, sans leucine et sans lysine (SD-HIS-LEU-LYS ; souche VL6-48 et VL6-48N).

II. Techniques de Biologie moléculaire

II.1. Plasmides

II.1.1. Plasmides commerciaux ou précédemment construits

Cette section décrit les différents plasmides déjà conçus au laboratoire, disponibles commercialement ou obtenus par échanges académiques, qui ont été utilisés au cours de ces travaux.

pGEM[®]-T Easy :

Ce plasmide commercialisé par Promega[®] est un vecteur permettant de cloner des produits PCR générés à l'aide d'une polymérase thermostable qui ajoute une déoxyadénosine à l'extrémité 3' du fragment (méthode de clonage appelée TA-cloning). Il contient une origine

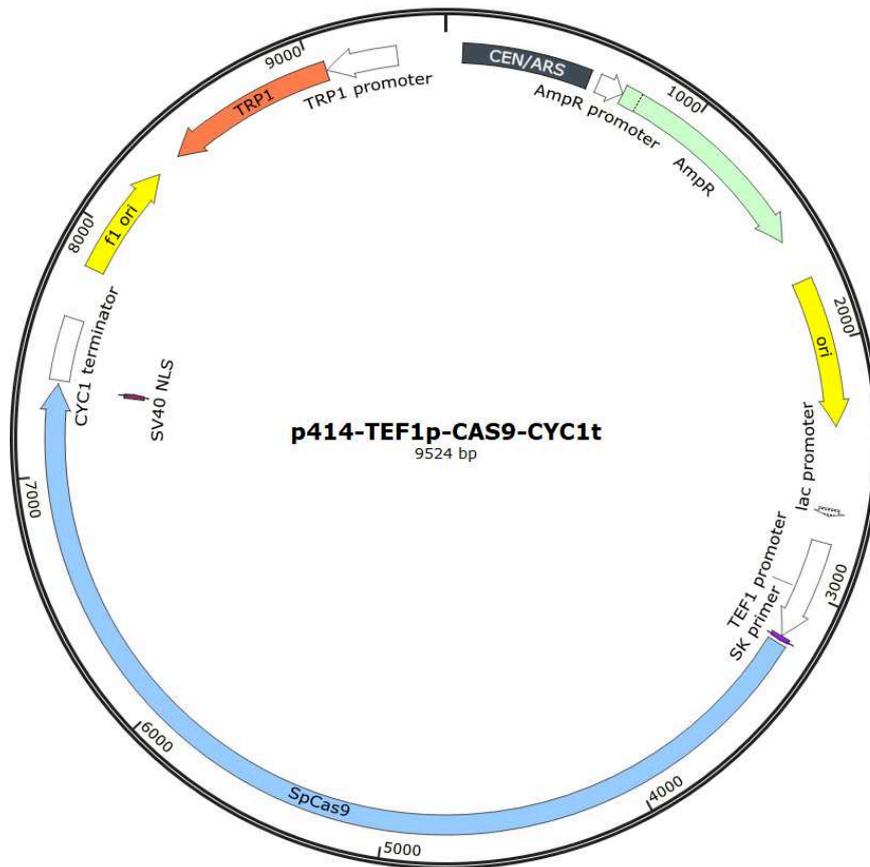


Figure 17 : Carte du plasmide p414-TEF1p-Cas9-CYC1t portant le marqueur d'auxotrophie TRP1 et permettant l'expression constitutive du gène codant la protéine SpCas9 dans la levure.

de réplication *E. coli* (*ColE1*), ainsi qu'un marqueur de sélection *ampR*, conférant la résistance à l'ampicilline.

pICE-Ade2 :

Ce plasmide, développé par Flagg *et al.*, 2019²⁶⁰, est distribué par Addgene (#122876). Il contient un centromère (CEN), une origine de réplication levure (ARS) et un marqueur d'auxotrophie pour l'adénine (ADE2). Il possède également le marqueur de sélection *ampR*, conférant la résistance à l'ampicilline. Ce plasmide a été utilisé comme matrice pour amplifier la cassette ADE2 nécessaire à la construction de nouveaux plasmides recombinants.

pICE-Lys2 :

Ce plasmide, développé par Flagg *et al.*, 2019²⁶⁰, est distribué par Addgene (#122879). Il contient un centromère (CEN), une origine de réplication levure (ARS) et un marqueur d'auxotrophie pour la lysine (LYS2). Il contient également le marqueur de sélection *ampR*, conférant la résistance à l'ampicilline. Ce plasmide a servi de matrice pour amplifier la cassette LYS2 nécessaire à la construction de nouveaux plasmides recombinants ou pour amplifier des pansements de recombinaison utilisés au cours de ces travaux.

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* (**Figure 17**). Il contient une origine de réplication *E. coli* (*ColE1*), ainsi qu'un marqueur de résistance à l'ampicilline, *ampR*. Il dispose également d'un centromère levure (CEN), d'une origine de réplication levure (ARS), et d'un marqueur d'auxotrophie pour le tryptophane (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. Le plasmide p414-TEF1p-Cas9-CYC1t a été développé par DiCarlo *et al.*, 2013²⁶¹ et, est distribué par Addgene (#43802).

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 d'enzymes de restriction. Ce MCS est flanqué de part et d'autre par les promoteurs T3 et T7. Il contient

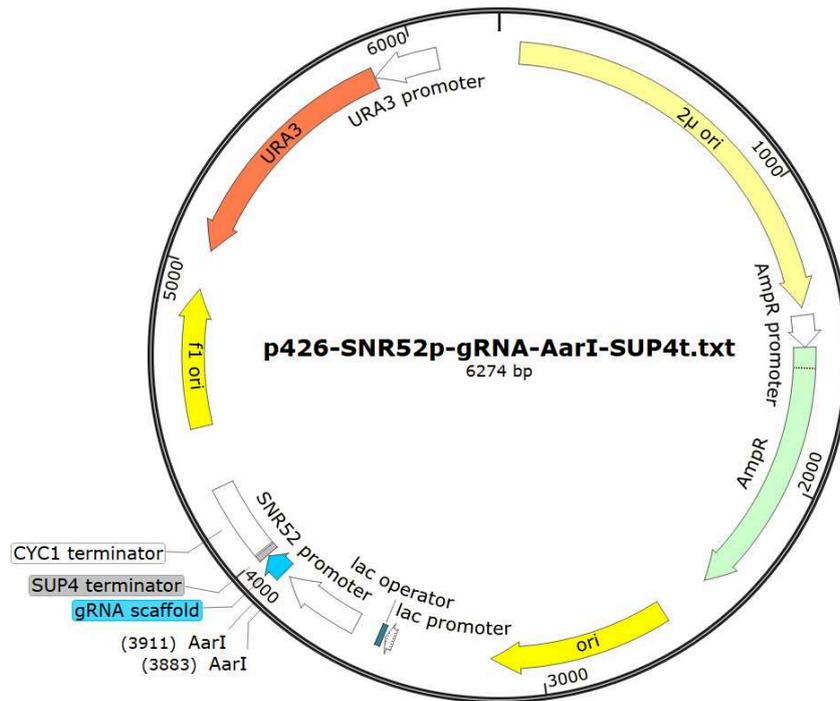


Figure 18 : Carte du plasmide p426-SNR52p-gRNA-AarI-SUP4t portant le marqueur d'autotrophie URA3 et permettant l'expression constitutive d'un gène codant un ARNg dans la levure.

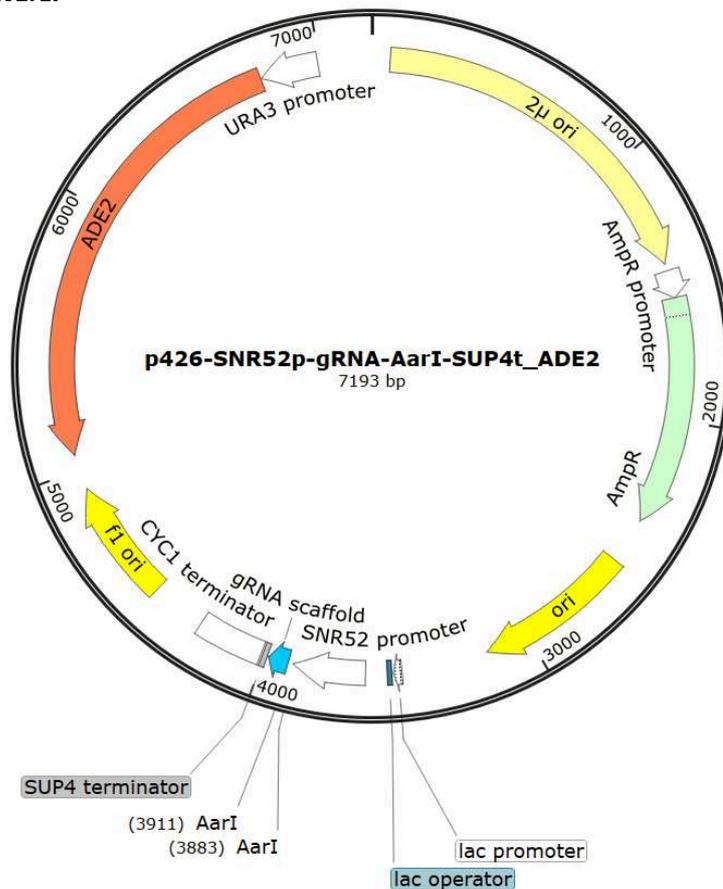


Figure 19 : Carte du plasmide p426-SNR52p-gRNA-AarI-SUP4t_Ade2 portant le marqueur d'autotrophie ADE2 et permettant l'expression constitutive d'un gène codant un ARNg dans la levure.

une origine de répllication *E. coli* (*ColE1*), ainsi qu'un marqueur de sélection *ampR*, conférant la résistance à l'ampicilline.

p426-SNR52p-gRNA-Aar1-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 constitutive, sous le contrôle du promoteur constitutif SNR52 (**Figure 18**). Il contient une origine de répllication *E. coli* (*ColE1*), ainsi qu'un marqueur de résistance à l'ampicilline, *ampR*. C'est un plasmide à grand nombre de copies chez la levure grâce à l'origine de répllication 2 μ . Il dispose du marqueur d'auxotrophie pour l'uracile (URA3) qui permet sa sélection. Il a été optimisé par Tsarmopoulos *et al.*²⁴⁹ à partir du vecteur p426-SNR52p-gRNA.CAN1.Y-SUP4t développé par DiCarlo *et al.* 2013²⁶¹ et distribué par 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 de l'enzyme AarI. Ce site de clonage peut être remplacé par le spacer de notre choix, par restriction et ligature d'un duplex d'oligonucléotides. Au cours des expérimentations et en fonction des cibles choisies, plusieurs spacers ont été intégrés (*cf.* Annexes **Tableau S1**).

p426-SNR52p-gRNA-Aar1-SUP4t Ade2 :

Ce plasmide est un dérivé du p426-SNR52p-gRNA-Aar1.Y-SUP4t dans lequel le marqueur d'auxotrophie URA3 a été remplacé par celui pour l'adénine (ADE2) (**Figure 19**).

pBSB9 :

Le pBSB9 nous a été fourni par nos collaborateurs de l'équipe Syber, unité Micalis, INRAE (**Figure 20**). Il contient une origine de répllication *E. coli* (*ColE1*), ainsi qu'un marqueur de résistance à l'ampicilline *ampR*. Le marqueur de sélection *cat* conférant la résistance au chloramphénicol et un gène d'expression de la protéine fluorescente « Green Fluorescent Protein » (GFP) sont également présents. Au cours de ces travaux, ce plasmide a servi de base pour amplifier la cassette *cat*.

pDR111 :

Ce plasmide est intégratif chez *B. subtilis* et nous a été fourni par nos collaborateurs de l'équipe Syber, unité Micalis, INRAE (**Figure 21**). Il contient une origine de répllication *E. coli*

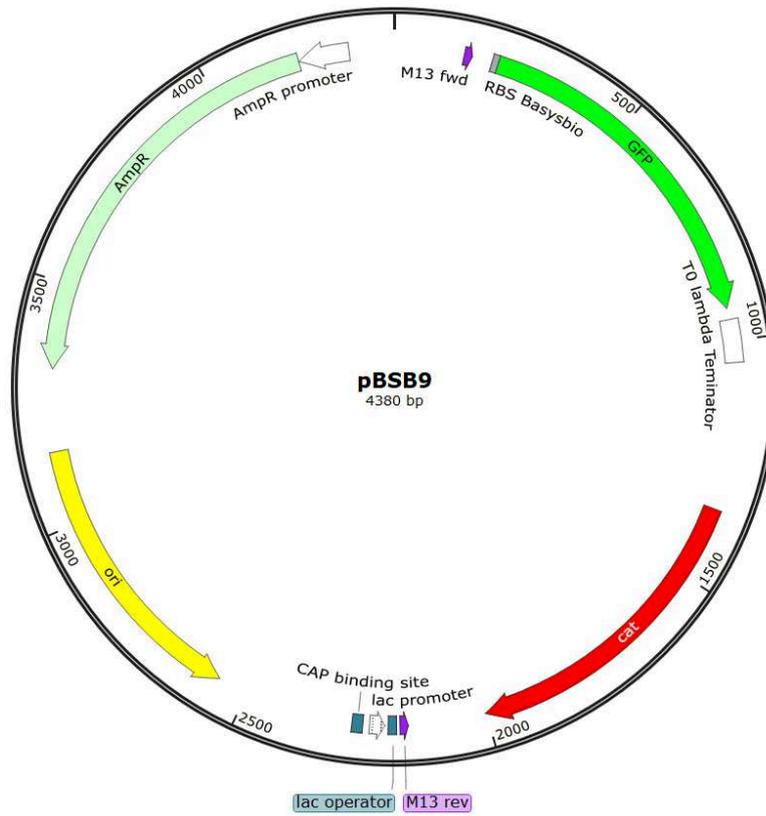


Figure 20 : Carte du plasmide pBSB9

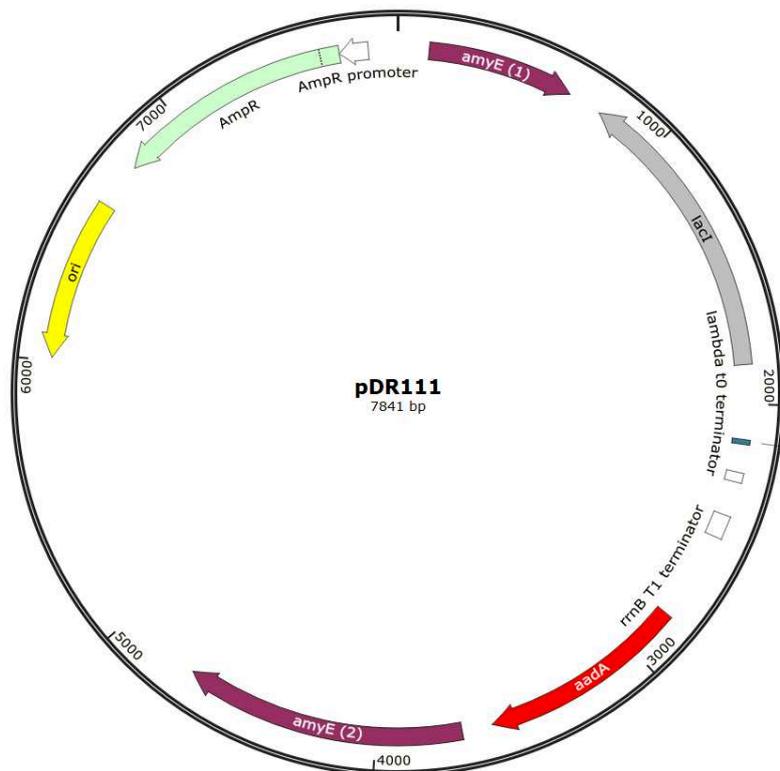


Figure 21 : Carte du plasmide intégratif pDR111

(*ColE1*), ainsi qu'un marqueur de résistance à l'ampicilline *ampR*. Deux cassettes correspondant au début et à la fin du gène *amyE* encadrent le marqueur de sélection *aadA* conférant la résistance à la spectinomycine et le represser *lacl*. Ces cassettes permettent l'intégration du marqueur *aadA* par double recombinaison au sein du *locus amyE*. Il a servi de matrice pour amplifier la cassette *aadA*, mais aussi comme plasmide contrôle au cours des expériences de transformation chez *B. subtilis*.

pJIM :

Ce plasmide est répliatif chez *B. subtilis* (**Figure 22**). Il a été obtenu via l'équipe Syber Unité Micalis d'INRAE de Jouy-en-Josas. Il contient une origine de répliation *E. coli* (*ColE1*), ainsi qu'un marqueur de résistance *ermB* conférant la résistance à l'érythromycine. Les séquences *copF*, *repD* et *repE* permettent la répliation du plasmide chez *B. subtilis*.

pCM185 :

Ce plasmide est un vecteur d'expression chez la levure de type centromérique (YCp). C'est un don du Pr. François Doignon du Laboratoire de Biogénèse Membranaire (LBM) de l'UMR 5200 du CNRS (campus INRAE, Villenave d'Ornon). Il contient une origine de répliation *E. coli* (*ColE1*), ainsi qu'un marqueur de résistance *ampR* conférant la résistance à l'ampicilline. Il dispose également d'un centromère levure (CEN), d'une origine de répliation levure (ARS416 ou ARS1), et d'un marqueur d'auxotrophie pour le tryptophane (TRP1). L'expression de protéines recombinantes chez la levure est induite par le promoteur tetO-CYC1. Il est commercialisé par l'ATCC® via le numéro d'accession ATCC#87659. Il a servi à isoler la séquence de l'ARS416 afin de construire le plasmide pGG03 utilisé comme matrice des pansements pour les expériences de TAR-Assembly.

pFL39 :

Ce plasmide est un vecteur navette chez *E. coli* et chez la levure *S. cerevisiae*. Il nous a été donné par le Pr. François Doignon du Laboratoire de Biogénèse Membranaire (LBM) de l'UMR 5200 du CNRS (campus INRAE, Villenave d'Ornon). Il contient une origine de répliation *E. coli* (*ColE1*), ainsi qu'un marqueur de résistance *ampR* conférant la résistance à l'ampicilline. Il dispose également d'une origine de répliation levure (ARS413) et d'un marqueur d'auxotrophie pour le tryptophane (TRP1). Il est commercialisé par l'ATCC® via le numéro

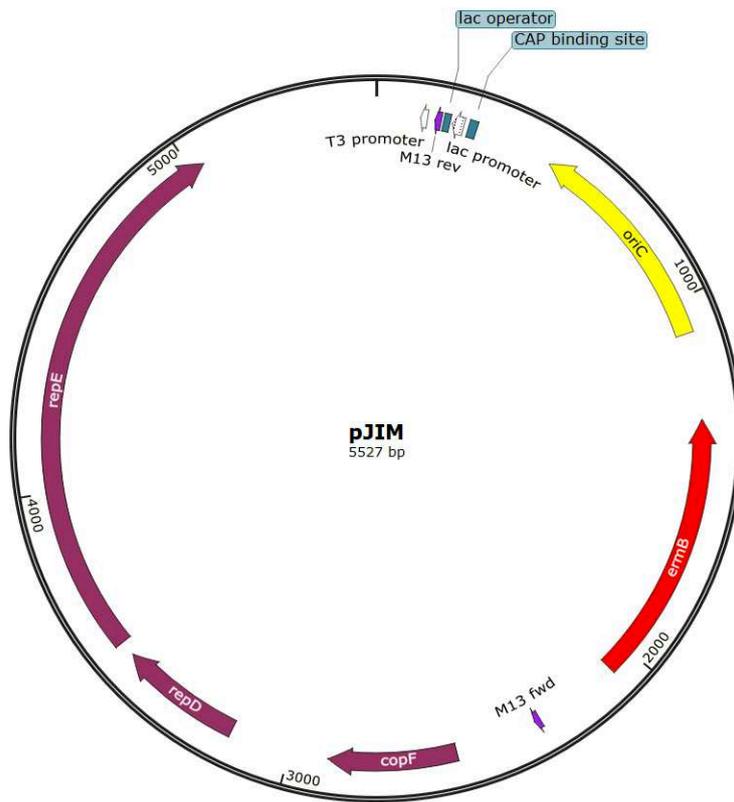


Figure 22 : Carte du plasmide réplcatif pJIM

d'accèsion ATCC#77204. Il a servi à isoler la séquence de l'ARS413 afin de construire le pGG01 utilisé comme matrice des pansements pour les expériences de TAR-Assembly.

pMT85PStetM-PSlacZ-pRS313 :

Ce plasmide intégratif chez les mycoplasmes a été développé par Labroussaa *et al.*, 2016²³⁰ (**Figure 23**). Il contient une cassette *PSlacZ*, composée du gène *lacZ* codant la β -galactosidase sous le contrôle du promoteur constitutif PS (trouvé en amont du gène codant la spiraline chez *Spiroplasma citri*), une cassette contenant les éléments levure (CEN6, ARSH4, HIS3), ainsi qu'un système de transposition composé de séquences d'insertion répétées inversées (IR) et des gènes *TnA* et *TnB* codant une transposase. Il contient également le marqueur de sélection *tet(M)*, conférant la résistance à la tétracycline. Enfin, il contient une origine de répllication *E. coli* (*ColE1*), ainsi qu'un marqueur de résistance *ampR* conférant la résistance à l'ampicilline. Il a servi à amplifier différentes matrices de recombinaison ou pansement avec les éléments levure pour les expériences de clonage de génomes chez la levure.

pMT85PStetM-PSlacZ-pVC604 :

Ce plasmide intégratif chez les mycoplasmes est identique au pMT85PStetM-PSlacZ-pRS313 à la seule différence qu'il ne possède pas la séquence ARSH4. Il a servi à amplifier différentes matrices de recombinaison ou pansement avec les éléments levure mais sans séquence ARS pour les expériences de clonage de génomes chez la levure.

pMiniTetM :

Ce plasmide intégratif chez les mycoplasmes a été développé par Pour-El *et al.*, en 2002²⁶² et publié sous le nom de pISM31.1. Il contient une origine de répllication *E. coli* (*oriC*), ainsi qu'un marqueur de résistance *ampR* à l'ampicilline et un système de transposition composé de séquences d'insertion répétées inversées (IR) et des gènes *TnA* et *TnB* codant une transposase. Il contient également le marqueur de sélection *tet(M)*, conférant la résistance à la tétracycline.

pTi4.0 SpCas9 pmcDA1 :

Ce plasmide intégratif développé par Ipoutcha *et al.* 2022²⁶³ a été utilisé afin de mener à bien les expériences d'inactivation de gènes par désamination (**Figure 24**). Ce plasmide contient une origine de répllication *E. coli* (*ColE1*), ainsi qu'un marqueur de résistance à l'ampicilline, *ampR*. Il permet notamment l'expression chez *Mycoplasma gallisepticum* (*Mgal*) de la

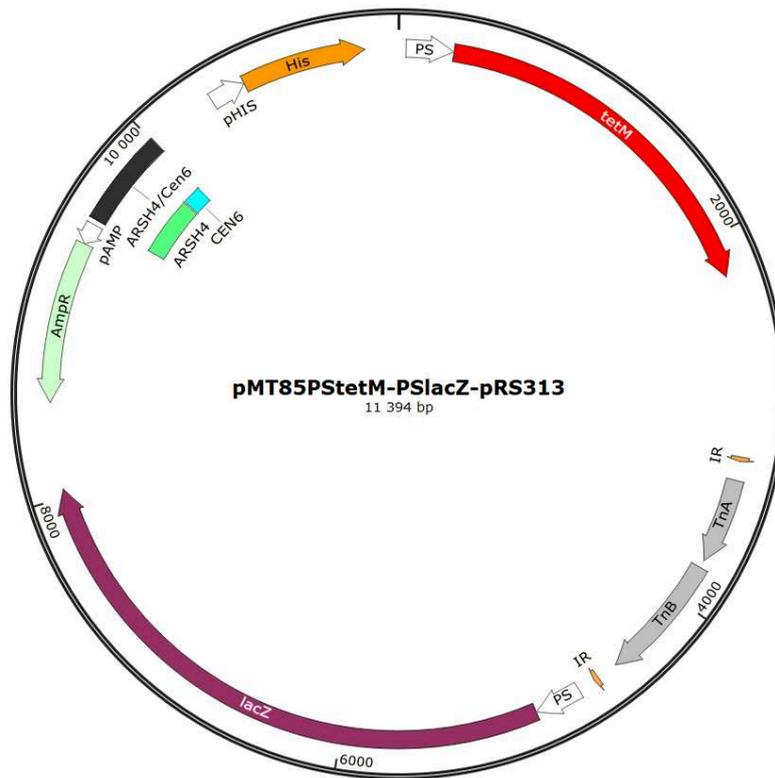


Figure 23 : Carte du plasmide pMT85PstetM-PSlacZ-pRS313

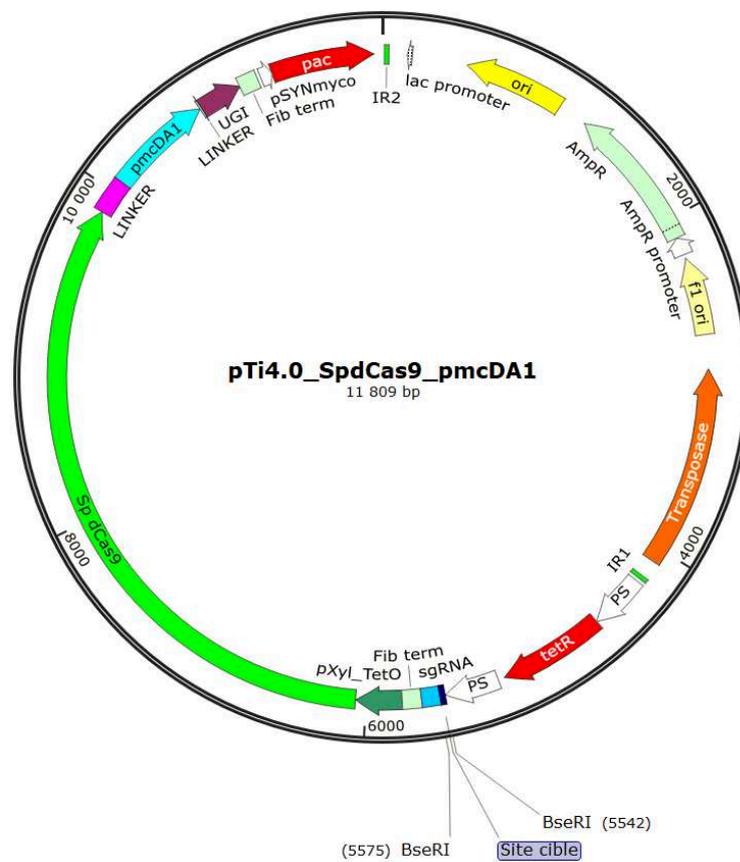


Figure 24 : Carte du plasmide pTi4.0_SpdCas9_pmcDA1

protéine SpdCas9 (dead SpCas9) fusionnée avec la cytosine desaminase pmcDA1 et l'inhibiteur UGI (uracil glycosylase inhibitor). L'expression de cette protéine Cas9 hybride est contrôlée par le promoteur inducible *Pxy/tetO2*. Il permet aussi l'expression d'un ARNg sous le contrôle du promoteur PS constitutif chez certaines espèces de mollicutes. Cet ARNg se compose d'une séquence de 20 bases spécifiques du gène cible (« spacer ») qui est fusionnée avec un composant structural (tracrARN). Le site de clonage « Site cible » peut être remplacé par le spacer de notre choix, par restriction avec l'enzyme BseRI et ligature d'un duplex d'oligonucléotides. Au cours de ces travaux, le spacer intégré a permis de cibler le gène *mnuA* codant une nucléase ; il est répertorié **Tableau S1** (cf. Annexes). Des régions inversées et répétées (IR-1 et IR-2) représentées en vert, encadrent l'ensemble de la cassette [pPSTetR-sgRNA-SpdCas9-*pac*] et permettent son intégration dans le génome de *Mgal* grâce à l'action de la transposase. Le marqueur de résistance à la puromycine, *pac*, permet la sélection après transposition.

pMT85_SpdCas9_pmcDA1 :

Ce plasmide intégratif développé par Ipoutcha *et al.* 2022²⁶³ a été utilisé afin de mener à bien les expériences d'inactivation de gènes par désamination (**Figure 25**). Ce plasmide contient une origine de répllication *E. coli* (*ColE1*), ainsi qu'un marqueur de résistance à l'ampicilline, *ampR*. Il permet notamment l'expression chez *Mycoplasma bovis* (*Mbov*) ou chez *Mycoplasma agalactiae* (*Maga*) de la protéine SpdCas9 (dead SpCas9) fusionnée avec la cytosine desaminase pmcDA1 et l'inhibiteur UGI (uracil glycosylase inhibitor). L'expression de cette protéine Cas9 hybride est contrôlée par le promoteur inducible *Pxy/tetO2*. Il permet aussi l'expression d'un ARNg sous le contrôle du promoteur PS constitutif chez certaines espèces de mollicutes. Cet ARNg se compose d'une séquence de 20 bases spécifiques du gène cible (« spacer ») qui est fusionnée avec un composant structural (tracrARN). Le site de clonage « Site cible » peut être remplacé par le spacer de notre choix, par restriction avec l'enzyme BseRI et ligature d'un duplex d'oligonucléotides. Au cours de ces travaux, les spacers intégrés ont permis de cibler respectivement le gène *mnuA* codant la nucléase MnuA de *M. bovis* ou de *M. agalactiae* ; ils sont répertoriés **Tableau S1** (cf. Annexes). Des régions inversées et répétées (IR-1 et IR-2) représentées en vert, encadrent l'ensemble de la cassette [Res-PSTetR-sgRNA-SpdCas9-*aacA-aphD-RES*] et permettent son intégration dans les génomes de *Mbov* et de

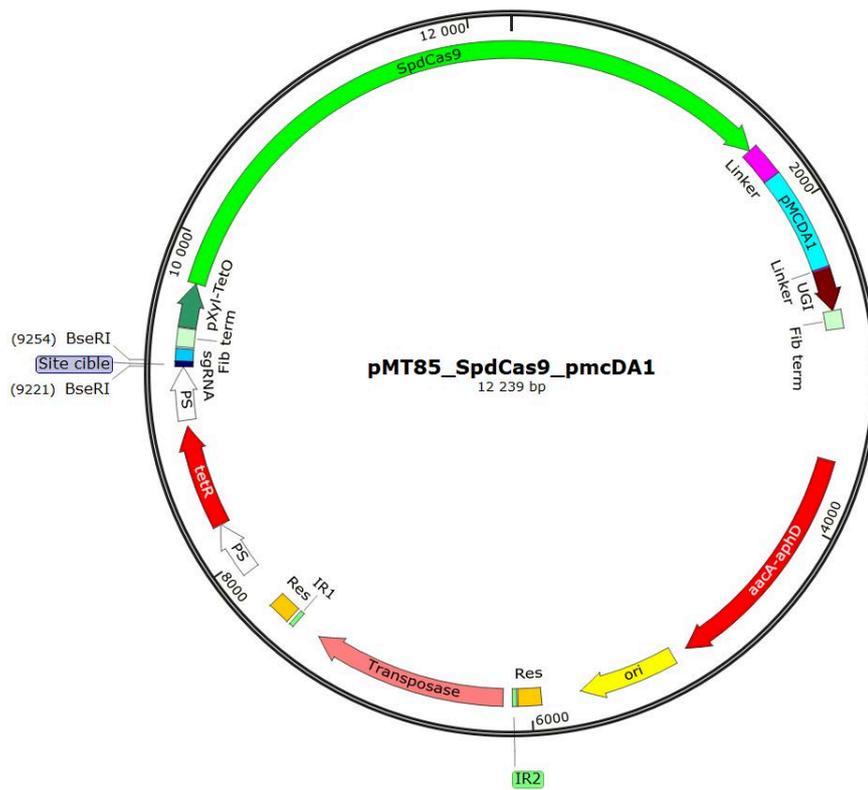


Figure 25 : Carte du plasmide pMT85_SpdCas9_pmcDA1

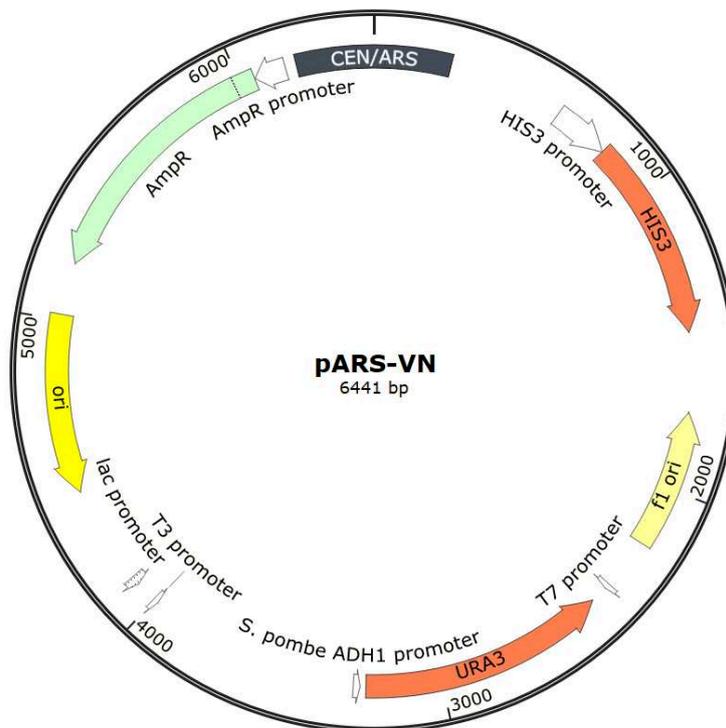


Figure 26 : Carte du pARS-VN

Mgal grâce à l'action de la transposase. Le marqueur de résistance à la gentamycine, *aacA-aphD*, permet la sélection après transposition.

pARS-VN :

Ce plasmide développé par Noskov *et al.* 2003²³², nous a été transmis par N. Kouprina (NIH, Bethesda, USA) (**Figure 26**). Il contient une origine de répllication *E. coli* (*ColE1*), ainsi qu'un marqueur de résistance *ampR* conférant la résistance à l'ampicilline. Il contient un centromère (CEN), une origine de répllication levure (ARS) et un marqueur d'auxotrophie pour l'histidine (HIS3). Un second marqueur d'auxotrophie pour l'uracil (URA3) est également présent sous le contrôle du promoteur *ADH1* de chez *Schizosaccharomyces pombe*. Il a servi à amplifier des matrices de recombinaison (ou pansement) avec les éléments levure pour les expériences de capture de fragments de génome de *B. subtilis* par CREasPy-Fusion avec une contre sélection par le 5-FOA.

II.1.2. Plasmides construits au cours de ces travaux

Différents plasmides ont été développés et produits au cours de la thèse. Ces derniers dérivent de ceux décrits dans la section précédente, et ont été assemblés en utilisant des méthodes standards de biologie moléculaire décrites ci-après dans les sections II.8 à II.13.

p426-gRNA-Ura3-double-target :

Ces plasmides ont été construits par Gibson Assembly (*cf.* §II.10.). Le fragment « insert » composé des éléments promoteur *SNR52*, gRNA scaffold, *SUP4* terminator et *CYC1* terminator est amplifié par PCR grâce aux amorces *p426_insert_F* et *p426_insert_R* en utilisant le plasmide *p426-SNR52p-gRNA1-Aarl-SUP4t* comme matrice (**Figure 18**). Le fragment « vecteur » correspondant à l'ensemble du plasmide *p426-SNR52p-gRNA-Aarl-SUP4t* est amplifié grâce aux amorces *Div_p426_vecteur_F* et *Div_p426_vecteur_R* placées en sens inverse en utilisant un second plasmide matrice avec un gRNA différent au premier, le *p426-SNR52p-gRNA2-Aarl-SUP4t* (**Figure 18**). Les plasmides *p426-gRNA-double-target* permettent l'expression constitutive de deux gRNAs sur un même plasmide ciblant simultanément deux cibles différentes. Plusieurs plasmides ont été construits au cours de ces travaux sur le modèle

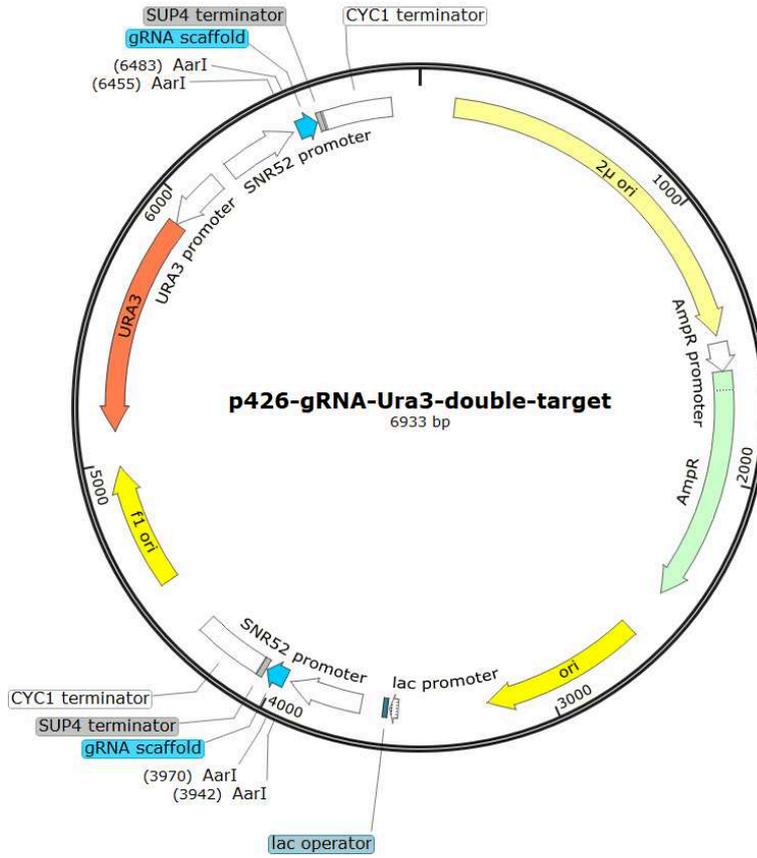


Figure 27 : Carte du p426-gRNA-Ura3-double-target

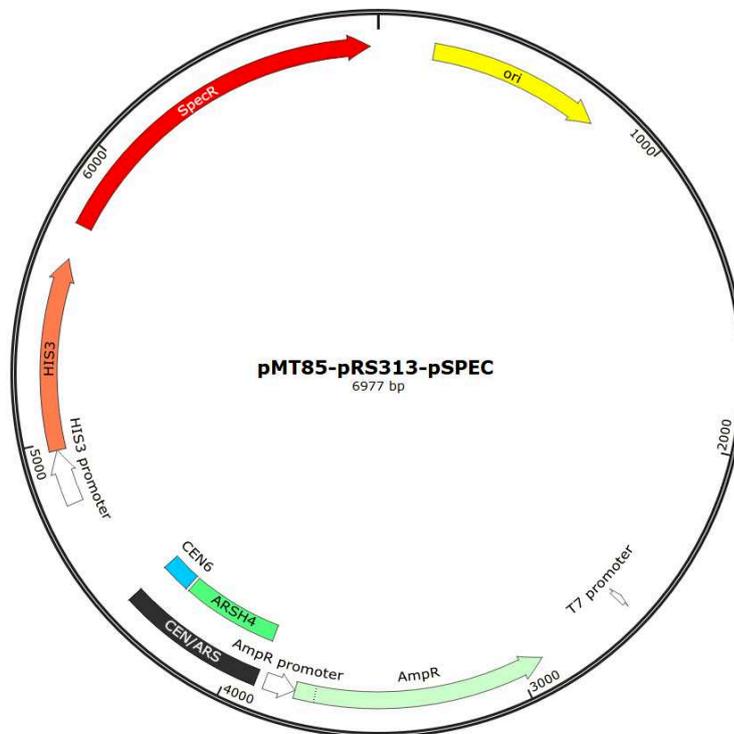


Figure 28 : Carte du plasmide matrice pMT85-pRS313-pSPECK

de la **Figure 27** et sont répertoriés dans le **Tableau S1** (cf. Annexes). Les amorces utilisées pour les amplifications PCR sont quant à elles indiquées dans le **Tableau S2** (cf. Annexes).

p426-gRNA-Ade2-double-target :

Ces plasmides ont été construits de la même manière que pour les p426-gRNA-Ura3-double-target mais avec des plasmides matrice p426-SNR52p-gRNA+AarI-SUP4t_Ade2. La différence entre ces vecteurs réside dans le marqueur d'auxotrophie : ADE2 au lieu de URA3. Plusieurs plasmides ont été construits au cours de ces travaux et sont répertoriés dans le **Tableau S2** (cf. Annexes).

pMT85PStetM-pRS313 sans lacZ :

Ce plasmide intermédiaire à la construction des pMT85-pRS313-pSPEC ou pMT85-pRS313-pCAT a été produit afin d'éliminer le gène *lacZ* et de réduire la taille du plasmide. Le pMT85PStetM-PS*lacZ*-pRS313 (11 394 pb) a été digéré par l'enzyme EcoRI (NEB) afin d'éliminer le gène *PSlacZ* (3346 pb). Les produits de digestion sont ensuite purifiés et quantifiés. La ligation permettant la recirculation du plasmide pMT85PStetM-pRS313 sans le gène *lacZ* a été réalisée en utilisant la T4 DNA Ligase (Promega©), selon les instructions du fournisseur (sur la nuit à 16°C). Le mélange de ligation (1 µL) est ensuite transformé par électroporation dans des bactéries *E. coli* DH10B. Après étalement des cellules transformées sur milieu de sélection, une colonie est isolée puis mise en culture liquide. Le plasmide est extrait par mini-préparation, quantifié, vérifié puis stocké à -20°C. La vérification du plasmide est réalisée par analyse du profil de restriction obtenu en électrophorèse après digestion par différentes enzymes de restriction (EcoRI et KpnI), et par séquençage Sanger.

pMT85-pRS313-pSPEC :

Ce plasmide (**Figure 28**) dérive du plasmide pMT85tetM-pRS313 sans *lacZ* précédemment décrit, après remplacement du marqueur de sélection *tet(M)* par le marqueur *aadA* conférant la résistance à la spectinomycine. Brièvement, la cassette *aadA* est amplifiée à partir du plasmide pDR111 en utilisant des amorces (cf. Annexes **Tableau S3**) qui contiennent des sites de restriction pour les enzymes AgeI et NsiI. La PCR est réalisée avec la polymérase Advantage 2 (Clontech). Les produits PCR sont analysés par électrophorèse sur gel d'agarose, puis purifiés et quantifiés. Le plasmide pMT85tetM-pRS313 sans *lacZ* (1 µg) et la cassette *aadA* purifiée (1 µg) sont digérés avec les enzymes de restriction AgeI-HF et NsiI-HF (NEB). Les

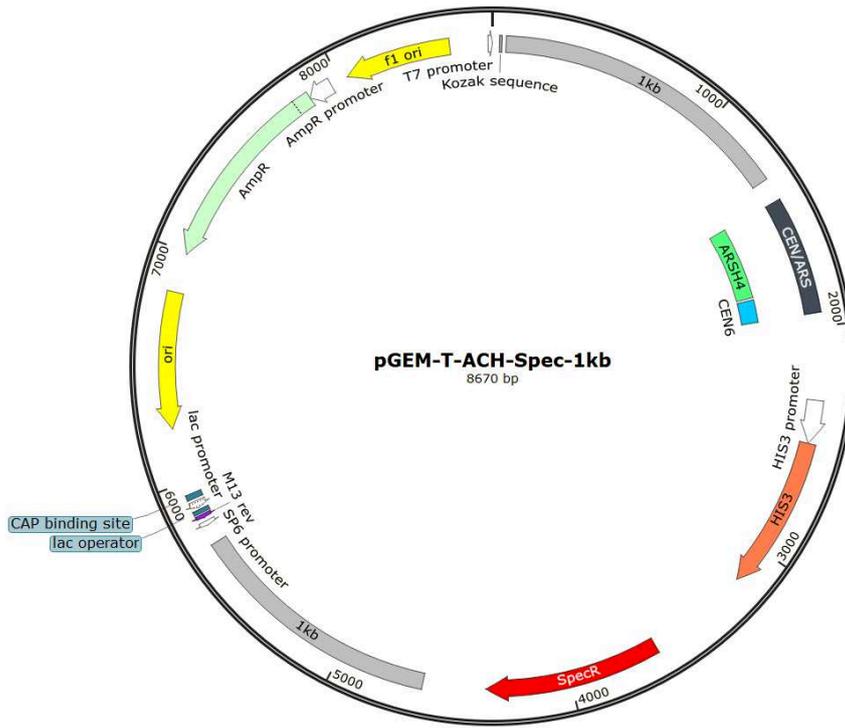


Figure 29 : Carte du plasmide matrice pGEM-T-ACH-Spec-1kb

produits de digestion sont ensuite purifiés et quantifiés. La ligature entre le plasmide pMT85tetM-pRS313 sans lacZ et la cassette *aadA*, tous deux digérés, est réalisée en utilisant la T4 DNA Ligase (Promega®), selon les instructions du fournisseur (sur la nuit à 16°C). Le mélange de ligature (1 µL) est ensuite transformé par choc thermique dans des bactéries chimiocompétentes *E. coli* Stellar™ Clontech commerciales. Après étalement des cellules transformées sur milieu de sélection, une colonie est isolée puis mise en culture liquide. Le plasmide est extrait par mini-préparation, quantifié, vérifié puis stocké à -20°C. La vérification du plasmide est réalisée par analyse du profil de restriction obtenu en électrophorèse après digestion par différentes enzymes de restriction (HindIII, AgeI/NdeI, EcoRI et KpnI), et par séquençage Sanger.

pGEM-T-ACH-Spec-1kb :

Ce plasmide a été construit par hybridation de trois fragments PCR chevauchants clonés par TA cloning dans le pGEM®-T Easy (**Figure 29**). Chacun des fragments correspond respectivement aux éléments levure (ARS-CEN-HIS), et à deux régions de 1 kpb du génome de *B. subtilis* qui bordent le « Fragment 1 » (*cf.* expériences TAR-Cloning avec l'enzyme de restriction SgrDI). Brièvement, les fragments au préalable purifiés et dosés sont hybridés par cycle PCR en utilisant la polymérase Q5 de chez NEB. Le produit d'hybridation (correspondant à un insert) est vérifié sur gel d'agarose et dosé au spectrophotomètre à 260nm (*cf.* §II.12.). Une seconde réaction PCR est effectuée en utilisant la Taq ADN polymérase de chez NEB afin de rajouter des bases « déoxyadénosine ou A » aux extrémités de l'insert. Ce dernier est ensuite ajouté par ligation au pGEM®-T Easy (Promega) linéarisé et possédant des résidus « déoxythymidine ou T » exposés en 3', en suivant les recommandations du fournisseur. Le mélange de ligature (1 µL) est ensuite transformé par électroporation dans des bactéries électrocompétentes *E. coli* ElectroMAX™ Stbl4™ (Invitrogen). Après étalement des cellules transformées sur milieu de sélection, une colonie est isolée puis mise en culture liquide. Le plasmide est extrait par mini-préparation, quantifié, vérifié puis stocké à -20°C. La vérification du plasmide est réalisée par analyse du profil de restriction obtenu en électrophorèse après digestion par différentes enzymes de restriction (HindIII, EcoRI et KpnI), et par séquençage Sanger. Ce plasmide a servi de base pour produire par réaction PCR la matrice de recombinaison nécessaire aux essais de clonage du « Fragment 1 – SgrDI » du génome de

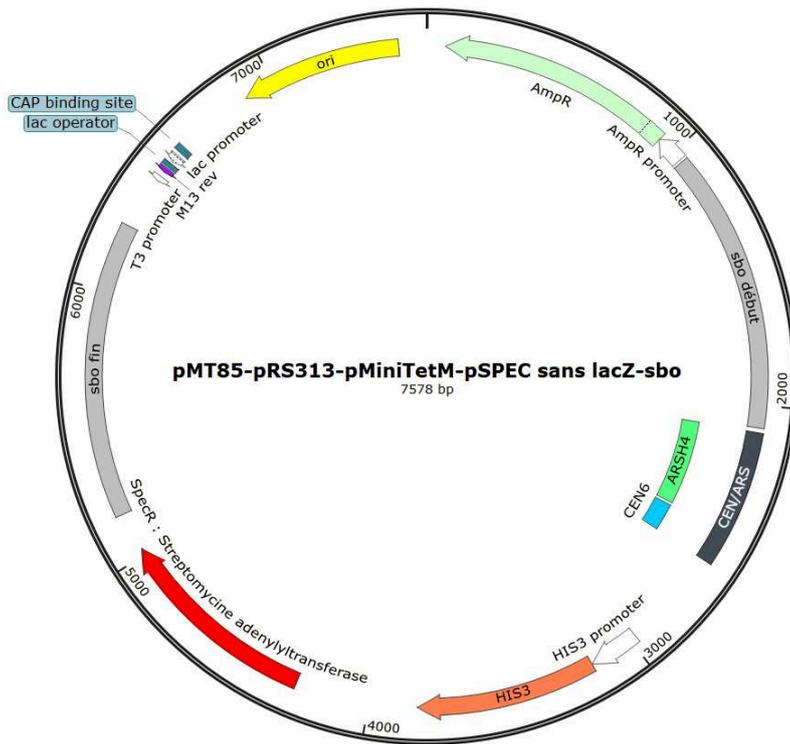


Figure 30 : Carte du plasmide intégratif pMT85-pRS313-pMiniTetM-pSPEC sans lacZ-sbo

B. subtilis souche MGP192 par TAR-Cloning. Les amorces utilisées pour les amplifications PCR sont indiquées dans le **Tableau S3** (cf. Annexes).

pMT85-pRS313-pMiniTetM-pSPEC sans lacZ-sbo :

Ce plasmide a été construit par réaction de Gibson Assembly (cf. §II.10.) entre quatre fragments issus de différents plasmides ou d'ADNg (**Figure 30**). Le 1^{er} fragment contient les éléments levure ARS-CEN-HIS et le marqueur de sélection pour la spectinomycine (amplifiés à partir du pMT85-pRS313-pSPEC-sbo sans lacZ). Le 2nd fragment porte une origine de répllication *E. coli* et un marqueur de sélection pour l'ampicilline (à partir du pMiniTetM). Les 3^{ème} et 4^{ème} fragments correspondent respectivement à des *locus* du génome de *B. subtilis* encadrant l'opéron *sboAX-albAG* et amplifiés à partir de l'ADNg de *B. subtilis*. Le mélange réactionnel (1 µL) est ensuite transformé par électroporation dans des bactéries électrocompétentes *E. coli* ElectroMAX™ Stbl4™ (Invitrogen). Après étalement des cellules transformées sur milieu de sélection, une colonie est isolée puis mise en culture liquide. Le plasmide est extrait par mini-préparation, quantifié, vérifié puis stocké à -20°C. La vérification du plasmide est réalisée par analyse du profil de restriction obtenu en électrophorèse après digestion par différentes enzymes de restriction (NdeI Nspl et KpnI), et par séquençage Sanger. Ce plasmide a servi de base pour produire par réaction PCR ou par digestion la matrice de recombinaison nécessaire au marquage du génome de *B. subtilis* par recombinaison homologue afin d'intégrer au préalable les éléments levures avant les essais de clonage de génome entier par les méthodes de clonage classique²¹⁵ ou par fusion²²³. Les amorces utilisées pour les amplifications PCR sont indiquées dans le **Tableau S3** (cf. Annexes).

Les plasmides pGG :

Ces plasmides ont été produits par restriction puis ligation (cf. §II.9. et §II.11.3.) de trois cassettes (une ARS, un marqueur d'auxotrophie et un marqueur antibiotique) dans le multi site de clonage (MSC) du plasmide pBluescript. Deux plasmides pGG01 et pGG03 ont été construits en utilisant des amorces correspondantes pour amplifier chaque cassette (cf. Annexes **Tableau S4**) et qui contiennent des sites de restriction pour les enzymes BamHI, PstI, EcoRI et HindIII. La PCR est réalisée avec la polymérase Advantage 2 (Clontech). Les produits PCR sont analysés par électrophorèse sur gel d'agarose, puis purifiés et quantifiés. Le plasmide pBluescript (1 µg) et les « cassettes insert » purifiées (1 µg) sont digérés avec les

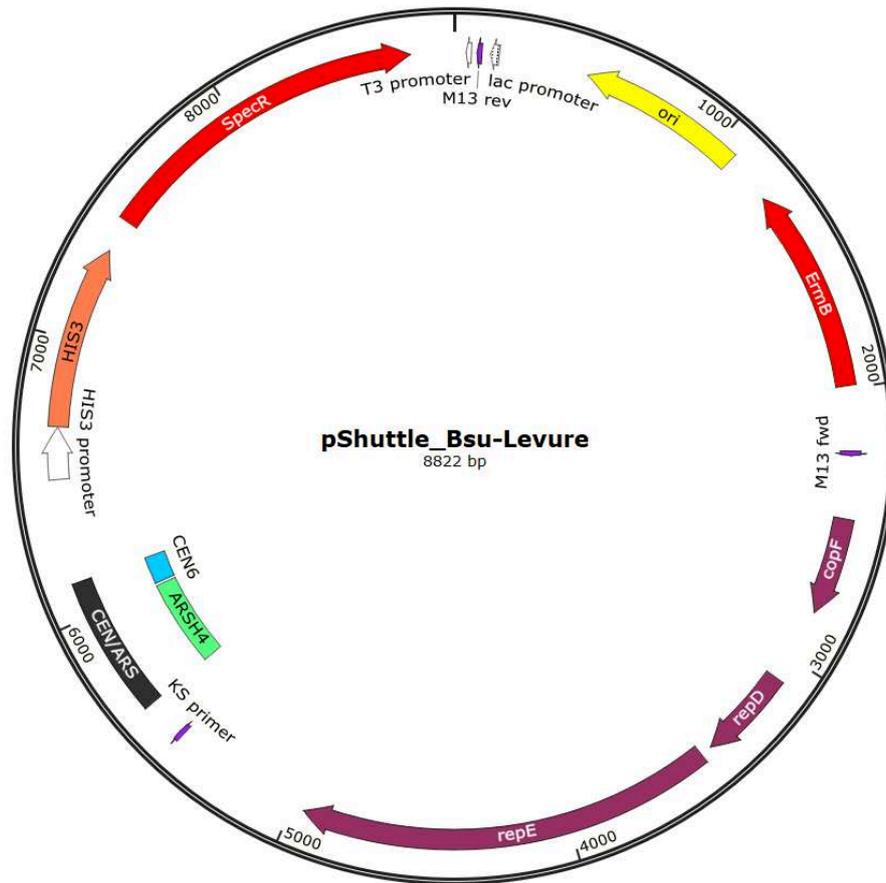


Figure 31 : Carte du plasmide navette pShuttle_Bsu-Levure

enzymes de restriction BamHI-HF, PstI-HF, EcoRI-HF et HindIII-HF. Les produits de digestion sont ensuite purifiés et quantifiés. La ligature entre le plasmide vecteur digéré et les cassettes, est réalisée en utilisant la T4 DNA Ligase (Promega®), selon les instructions du fournisseur (sur la nuit à 16°C). Le mélange de ligature (1 µL) est ensuite transformé par choc thermique dans des bactéries chimiocompétentes *E. coli* NEB® 5-alpha commerciales. Après étalement des cellules transformées sur milieu de sélection, une colonie est isolée puis mise en culture liquide. Le plasmide est extrait par mini-préparation, quantifié, vérifié puis stocké à -20°C. La vérification du plasmide est réalisée par analyse du profil de restriction obtenu en électrophorèse après digestion par différentes enzymes de restriction (HindIII, KpnI et BamHI), et par séquençage Sanger. Ces plasmides ont été construits pour servir de matrice à la production des pansements nécessaires aux expériences de TAR-Assembly.

pShuttle *Bsu*-Levure :

Ce plasmide a été produit par restriction puis ligation (*cf.* §II.9. et §II.11.3.) d'un insert dans le multi site de clonage (MSC) du plasmide répliatif pJIM (**Figure 31**). Brièvement, la cassette insert composée des éléments levures ARS-CEN-HIS et du marqueur de sélection *aadA* (SpecR), est amplifiée à partir du plasmide pMT85-pRS313-pSPEC (**Figure 28**) en utilisant des amorces (*cf.* Annexes **Tableau S4**) qui contiennent des sites de restriction pour les enzymes BamHI et SpeI. La PCR est réalisée avec la polymérase Advantage 2 (Clontech). Les produits PCR sont analysés par électrophorèse sur gel d'agarose, puis purifiés et quantifiés. Le plasmide pJIM (1 µg) et la cassette insert purifiée (1 µg) sont digérés avec les enzymes de restriction BamHI-HF et SpeI-HF (NEB). Les produits de digestion sont ensuite purifiés et quantifiés. La ligature entre le pJIM digéré et la cassette insert, tous deux digérés, est réalisée en utilisant la T4 DNA Ligase (Promega®), selon les instructions du fournisseur (sur la nuit à 16°C). Le mélange de ligature (1 µL) est ensuite transformé par choc thermique dans des bactéries chimiocompétentes *E. coli* NEB® 5-alpha commerciales. Après étalement des cellules transformées sur milieu de sélection, une colonie est isolée puis mise en culture liquide. Le plasmide est extrait par mini-préparation, quantifié, vérifié puis stocké à -20°C. La vérification du plasmide est réalisée par analyse du profil de restriction obtenu en électrophorèse après digestion par différentes enzymes de restriction (KpnI, PvuII et PvuII+EcoRV), et par séquençage Sanger. Ce plasmide navette répliatif chez *B. subtilis* mais pouvant aussi se maintenir chez la levure, a servi aux essais de mise au point de la méthode de CReasPy-Fusion.

II.2. Mini-préparation de plasmides

La mini-préparation de plasmide permet d'extraire une petite quantité (100 ng à 500 ng) de plasmides au préalable assemblés et/ou clonés par transformation chez *E. coli*. Basée sur le principe d'extraction d'ADN plasmidique par lyse alcaline (renaturation brutale et différentielle de l'ADN chromosomique et plasmidique par neutralisation rapide du pH puis séparation par centrifugation), cette extraction est réalisée en utilisant le kit « NucleoSpin® Plasmid EasyPure » (Macherey-Nagel), en suivant les instructions du fournisseur.

II.3. Midi-préparation de plasmides

La midi-préparation de plasmide permet d'extraire une plus grande quantité de plasmide (au-delà de 500 ng). Le principe de la purification est le même que celui du kit de mini-préparation. L'ADN plasmidique est extrait en utilisant le kit « NucleoBond® Xtra Midi » (Macherey-Nagel) en suivant les instructions du fournisseur. Après élution, les plasmides purifiés sont quantifiés et conservés à -20°C.

II.4. Vérification des plasmides par séquençage

Les plasmides sont séquencés par la société prestataire Genewiz® Azenta Life Sciences (séquençage Sanger), afin de vérifier qu'ils ne présentent pas de mutations indésirées et que les séquences sont identiques aux cartes plasmidiques théoriques. Les régions d'intérêt sont ainsi analysées en suivant les recommandations du prestataire. Lorsque les chromatogrammes résultants des analyses de séquençage sont reçus, la qualité du base-calling est étudiée et validée (score de qualité et longueur continue de lecture). Les séquences générées sont ensuite analysées par alignement avec les séquences théoriques en utilisant le logiciel ApE (A plasmide Editor).

II.5. Extraction d'ADN génomique total

Plusieurs protocoles d'extraction ont été utilisés au cours de ces travaux de thèse, en fonction des organismes mais aussi de la qualité des ADNg requise pour mener à bien les

analyses ultérieures (e.g. criblage PCR, séquençage « short-reads », séquençage « long-reads »).

II.5.1. Extraction d'ADN génomique de *Bacillus subtilis*

L'ADNg total de cellules bactériennes de *Bacillus subtilis* a été extrait en suivant le protocole décrit dans le kit Wizard® Genomic DNA Purification Kit de chez Promega©. À partir d'une culture en milieu liquide cultivée sur la nuit, un aliquot est prélevé en tube eppendorf et les cellules sont culotées par centrifugation. *B. subtilis* étant une bactérie Gram Positive, les cellules sont remises en suspension dans de l'EDTA 50 mM puis traitées 1h à 37°C avec du lysozyme à 2,5 mg/mL final. Cette lyse enzymatique permet de digérer l'épaisse couche de peptidoglycane formant la paroi de ces bactéries. Le protocole se décline ensuite en quatre étapes : (i) les cellules sont culotées par centrifugation puis remises en suspension dans la « Nuclei Lysis Solution ». Cette lyse chimique suivie d'une lyse thermique à 80°C permet de rompre les membranes cellulaires, (ii) le lysat est traité à la RNase pendant 1h à 37°C, (iii) les protéines cellulaires sont ensuite précipitées avec la « Protein Precipitation Solution » pendant 5 min sur glace, (iv) enfin, l'ADN génomique toujours présent en solution dans le surnageant après centrifugation, est concentré et purifié en éliminant les sels par une dernière étape de précipitation volume à volume par l'isopropanol. Après séchage pour enlever toutes traces de solvant, l'ADN est remis en suspension dans de l'eau ou du TE 1X puis dosé par spectrophotométrie au Nanodrop (cf. §II.12.). L'ADN est conservé à 4°C ou à -20°C en attendant les analyses supplémentaires.

II.5.2. Extraction d'ADN génomique de *Saccharomyces cerevisiae*

- Protocole « classique » :

L'ADN total des levures transformées est classiquement extrait selon le protocole de Kouprina et Larionov, 2008²⁶⁴. Des colonies isolées présentes sur les boîtes de transformation sont piquées et striées sur milieu sélectif puis incubées pendant 48h à 30°C. Un « patch » est ensuite réalisé à partir d'un sous-clone sur milieu sélectif afin d'augmenter le matériel microbien à extraire. Les patches sont incubés pendant 48h à 30°C. L'ADN des levures est extrait

à partir des patchs fraîchement réalisés (soit 48h). L'extraction résulte de la combinaison successive de plusieurs types de lyse : enzymatique, chimique et thermique. Pour chaque clone à extraire, une petite quantité de levures (~1mm³) est remise en suspension dans 90 µL d'une solution de zymolyase (Zymolyase® 100-T, 10 mg/mL ; Glycérol 50% (v/v)) additionnée de β-mercaptoethanol (activateur de la zymolyase). Après une incubation d'1h à 37°C, 10 µL de SDS 2% (concentration finale à 0,2%) est ajouté et le mélange est placé 15 min à 70°C (dénaturation des protéines et lyse des membranes). Puis 11 µL d'acétate de potassium à 5 mol.L⁻¹ (concentration finale à 0,55 mol.L⁻¹) sont ensuite ajoutés au lysat qui est réservé pendant 15 min sur glace. Les débris cellulaires et les protéines sont culotés par centrifugation à 14000 g pendant 10 min à 4°C. L'ADN soluble dans le surnageant est précipité par l'ajout de 90 µL d'isopropanol (ratio 1 : 1), puis culoté lors d'une seconde centrifugation à 14000g pendant 10 min à 4°C. Le culot d'ADN est remis en suspension après séchage dans du TE1X puis conservé à -20°C.

- Protocole « Quick & Dirty REXER » :

L'ADN total des levures transformées peut aussi être extrait selon un protocole plus rapide et simple de Robertson *et al.*, 2021²⁶⁵, à partir d'une colonie ou à partir de matériel levure concentré (aussi appelé patch) remis en suspension dans 30 µL d'eau milli-Q. Pour chaque clone (ou colonie) 6 µL de suspension cellulaire sont mélangés avec 14 µL de NaOH 0,02M et chauffés individuellement en tubes PCR à 95°C pendant 45 min à l'aide d'un thermocycleur. Les tubes PCR sont ensuite transférés dans une centrifugeuse de paillasse pour culoter et éliminer les débris cellulaires (30s, température ambiante). L'analyse PCR (polymérase One Taq, NEB) est effectuée en utilisant 1 µL de surnageant.

II.5.3. Extraction d'ADN génomique de mycoplasmes

- Protocole Wizard® Genomic DNA Purification Kit :

L'ADNg total de cellules bactériennes de mycoplasmes a été extrait en suivant le protocole décrit dans le kit Wizard® Genomic DNA Purification Kit de chez Promega®. A partir d'une culture liquide en fin de phase exponentielle, les cellules sont culotées par centrifugation puis resuspendues dans la « Nuclei Lysis Solution ». Cette lyse chimique suivie d'une lyse thermique à 80°C permet de rompre les membranes cellulaires. Le lysat est traité

à la RNase pendant 1h à 37°C. Les protéines cellulaires sont ensuite précipitées par la « Protein Precipitation Solution » pendant 5 min dans la glace. Enfin, l'ADN génomique toujours présent en solution dans le surnageant après centrifugation, est concentré et purifié en éliminant les sels par une dernière étape de précipitation volume à volume par l'isopropanol. Après séchage pour enlever toutes traces de solvant, l'ADN est remis en suspension dans de l'eau ou du TE 1X puis dosé par spectrophotométrie au Nanodrop. L'ADN ainsi préparé est conservé à 4°C ou à -20°C en attendant les analyses supplémentaires.

- Protocole « Quick & Dirty » :

Ce protocole permet l'extraction et l'analyse rapide de colonies de mycoplasmes après des expériences de transformation ou de transplantation. 200 µL d'une culture liquide en fin de phase exponentielle sont prélevés et centrifugés 10 min à 8000g afin de culoter les cellules de mycoplasmes. Le culot est remis en suspension dans 100 µL de TE 1X et chauffé 10 min à 95°C. Le lysat cellulaire est dilué au 1/10^{ème} dans de l'eau et 1 µL de préparation est utilisé pour préparer un mélange réactionnel PCR de 25 µL contenant 0,5 µL de polymérase Advantage 2 de chez Takara (*cf.* §II.6.1.).

- Qiagen® Genomic-tips 100/G :

Ce kit, basé sur l'utilisation d'une colonne échangeuse d'anions, permet l'extraction et la purification d'ADN de haut poids moléculaire (100-200 kpb). Cette technologie permet l'obtention d'échantillons d'ADN d'une grande pureté et d'une grande qualité (moins fragmentés qu'avec des protocoles plus classiques) car elle génère peu de cassures mécaniques lors de l'étape de capture de l'ADN. Ce protocole est particulièrement indiqué lorsque les ADN génomiques doivent être soumis à une analyse par séquençage à haut débit (séquençage NGS). L'extraction est réalisée à partir de 10 mL de culture pour un clone à analyser et en suivant les instructions du fabricant. L'ADN génomique purifié est élué dans de l'eau ultra-pure et conservé à 4°C.

II.6. PCRs simplexes et multiplexes

II.6.1. PCR simplexe (Polymerase Chain Reaction ou réaction en chaîne de la polymérase)

En fonction des analyses effectuées plusieurs polymérases et kit associés ont été utilisés pour réaliser les réactions de PCR : que ce soit comme moyen de détection (validation lors d'un criblage de la présence d'un *locus* particulier), ou bien, que ce soit comme moyen d'amplification d'un fragment d'ADN (production des pansements de recombinaison).

- Kit Advantage[®] 2 Polymerase Mix (Takara Bio) :

Ce kit est principalement utilisé lors de l'analyse des transformants levures après les expériences de clonage de génomes et extraction d'ADN génomique total de *S. cerevisiae* par le protocole classique (cf. §II.5.2.). Le mélange réactionnel ainsi que les paramètres des cycles d'amplification sont ajustés en suivant les recommandations du fournisseur.

- Kit Q5[®] High Fidelity DNA Polymerase (NEB) :

Ce kit est utilisé lors de l'analyse des transformants de *B. subtilis* après transformation par compétence naturelle, mais aussi pour l'amplification d'amplicons longs (>1 kb). La polymérase de ce kit étant extrêmement processive, elle permet la production de fragments d'ADN avec un taux d'erreurs très faible. Les matrices de recombinaison ou certaines cassettes nécessaires à la construction de plasmides sont amplifiées en utilisant prioritairement ce kit. Le mélange réactionnel ainsi que les paramètres des cycles d'amplification sont ajustés en suivant les recommandations du fournisseur.

- Kit OneTaq[®] DNA Polymerase (NEB) :

Ce kit est utilisé pour une analyse rapide des transformants levures après les expériences de clonage de génomes et extraction d'ADN génomique total de *S. cerevisiae* par le « Quick & Dirty REXER » (cf. §II.5.2.). La polymérase de ce kit est moins performante que celles décrites précédemment mais elle est cependant moins coûteuse et se révèle utile pour

des PCRs de détection sur des matrices de plus faible qualité. Le mélange réactionnel ainsi que les paramètres des cycles d'amplification sont ajustés en suivant les recommandations du fournisseur.

III.6.2. PCR multiplexe

Cette PCR permet d'amplifier simultanément plusieurs fragments d'ADN afin d'analyser l'intégrité d'un génome bactérien cloné dans la levure, ou bien, de vérifier l'identité d'un transplant après acquisition d'un génome donneur. En fonction du génome analysé, un set d'amorces spécifique est utilisé. Ce dernier est composé de plusieurs couples d'amorces répartis de manière homogène sur le chromosome circulaire. Les PCRs multiplexes sont réalisées à l'aide du kit Qiagen® Multiplex PCR en suivant les recommandations du fournisseur et après mise au point pour nos analyses. Brièvement, la réaction PCR est composée du tampon Qiagen Multiplex PCR Master Mix (1X final), de 0,025 µM de chaque amorce et de 40 à 60 ng de matrice d'ADN, dans un volume final de 15 µL. Le cycle PCR comporte une étape de dénaturation initiale de 15 min à 95°C, suivie de 35 cycles composés d'une phase de dénaturation d'1 min à 95°C, d'une phase d'hybridation de 2 min et 30 s à 50°C, et d'une phase d'élongation d'1 min et 30 s à 72°C, et se termine par une étape d'extension finale de 10 min à 68°C.

II.7. Amplification des matrices de recombinaison (pansements)

Les expériences de clonage de génome entier de *B. subtilis* ou de mycoplasmes par CREasPy-Cloning ou CREasPy-Fusion, ou bien, de capture de fragments de génome par TAR-Cloning sont basées sur l'utilisation du système de recombinaison homologue de la levure. Afin de favoriser le fonctionnement de ce système, il est important de fournir à la cellule un fragment d'ADN double brin linéaire qui peut être utilisé comme matrice de recombinaison. Cette matrice sert à réparer la cassure double brin causée suite à l'action du système CRISPR-Cas9 (*in vivo*) ou d'une enzyme de restriction (*in vitro*) ; Elle est dénommée « pansement ».

II.7.1. Pansements avec les éléments levures (YCP)

En fonction des expériences, de nombreux pansements de recombinaison composés de cassettes contenant les éléments levure CEN6, HIS3 ou ARSH4 ont été produits à partir de diverses matrices. Les cassettes sont flanquées de séquences identiques au *locus* à modifier, ou bien, encadrant le fragment à capturer. La taille de ces séquences est variable. Usuellement 60 pb sont rajoutées aux extrémités des cassettes par amplification PCR et en utilisant des amorces ayant une queue flottante en 5'. Les produits PCR sont analysés par électrophorèse sur gel d'agarose, puis purifiés, quantifiés et conservés à -20°C. Dans le cas particulier où des bras de recombinaison de 1 kpb sont ajoutés, le pansement est produit également par réaction PCR à partir du plasmide pGEM-T-ACH-Spec-1kb.

II.7.2. Pansements de jonction

- Pansement de 90 pb :

Il s'agit de pansements de recombinaison composés uniquement de deux régions juxtaposées de 45 pb, encadrant des *loci* à supprimer ou permettant la recircularisation aux jonctions de fragments de génome capturés. Ils sont obtenus par hybridation de deux oligonucléotides complémentaires de 90 nucléotides. La réaction est composée de 10 µL de chaque oligonucléotide concentré à 100 µM, et de 2,5 µL de tampon 10X Advantage 2 PCR Buffer. Le mélange est incubé pendant 5 min à 95°C dans un thermocycleur, puis la température est diminuée jusqu'à 16°C à un rythme de 0,1°C par seconde, afin de permettre l'hybridation des oligonucléotides. Les oligonucléotides hybridés sont conservés à -20°C.

- Pansement de 2 ou 10 kpb :

Ces pansements sont des amplicons PCR correspondants à des régions juxtaposées de 1 ou 5 kpb permettant la recircularisation aux jonctions de fragments de génome capturés. Ils sont obtenus en utilisant le kit Q5® High Fidelity DNA Polymerase (NEB) décrit précédemment. Les produits PCR sont analysés par électrophorèse sur gel d'agarose, puis purifiés, quantifiés et conservés à -20°C.

- Pansement de jonction avec ARS et marqueurs d'auxotrophie :

Ces pansements sont des amplicons PCR correspondants aux patches utilisés au cours des expériences de TAR-Assembly, et produits avec des amorces possédant des queues flottantes et à partir des plasmides matrice pGG01 et pGG03. Ils sont obtenus en utilisant le kit Q5[®] High Fidelity DNA Polymerase (NEB) décrit précédemment avec les amorces. Les produits PCR sont analysés par électrophorèse sur gel d'agarose, puis purifiés, quantifiés et conservés à -20°C.

II.8. Purification des ADNs

Les ADNs (matrices de recombinaison ou fragments nécessaires à la construction de plasmides) produits par réaction enzymatique (PCR, restriction, phosphorylation, déphosphorylation), sont purifiés grâce au kit de chez GE Healthcare « Illustra GFX PCR DNA and Gel Band Purification Kit ». Ce protocole permet de purifier et de concentrer des ADN allant de 50 pb à 10 kpb à l'aide d'une colonne avec une membrane de silice. Brièvement, l'ADN est adsorbé sur la membrane grâce à un tampon de capture (ou Capture Buffer) dont le pH est inférieur à 7,5. Des cycles de centrifugation permettent successivement le chargement de l'ADN à purifier sur la colonne, le lavage par un tampon de lavage (ou Wash Buffer) afin d'éliminer les sels et contaminants, et enfin le décrochage de l'ADN par élution avec au choix un tampon à base de Tris-HCl (10 mM pH8) ou avec de l'eau stérile dépourvue de nucléases.

II.9. Restriction enzymatique des ADNs

Les expériences de digestions par restriction sont réalisées en suivant les recommandations des fournisseurs pour chaque enzyme donnée. Le mélange réactionnel contient le tampon approprié à une concentration finale de 1X, additionné de 1 à 5 unités d'enzyme et de 0,2 à 2 µg d'ADN à cliver. Les réactions se déroulent généralement dans un volume final compris entre 10 et 50 µL dont la quantité suffisante pour (ou qsp) est ajustée avec de l'eau. Le mélange est incubé 1h à 3h à la température optimale de l'enzyme de restriction choisie.

II.10. Réaction de Gibson Assembly

Pour la construction de certains plasmides par réaction de Gibson Assembly²¹², le kit NEBuilder® Hifi DNA Assembly Master Mix de chez NEB a été utilisé en suivant les instructions du fournisseur. Cette méthode permet d'assembler *in vitro* des fragments d'ADN chevauchants en utilisant l'activité synchrone de trois enzymes (exonucléase 5'-3', ADN polymérase et ADN ligase).

II.11. Réaction de modification des ADNs

II.11.1. Phosphorylation des ADNs

Certains ADNs utilisés au cours de ces travaux ont été phosphorylés en utilisant la T4 Polynucléotide Kinase de chez NEB. Cette enzyme permet de rajouter des groupements phosphates aux extrémités 5' de fragments d'ADN ou d'ARN avant ligation. Les réactions de phosphorylation sont réalisées en suivant les recommandations du fournisseur. Le mélange réactionnel contient le tampon T4 PNK Reaction Buffer à une concentration finale de 1X, additionné de 10 unités d'enzyme pour 300 pmol d'ADN (maximum) et 10 mM d'ATP. Les réactions se déroulent généralement dans un volume final allant de 20 µL à 50 µL dont la qsp est ajustée avec de l'eau. Le mélange est incubé 30 min à 37°C. La réaction est ensuite arrêtée par inactivation par la chaleur pendant 20 min à 65°C.

II.11.2. Déphosphorylation des ADNs

L'enzyme Antartic phosphatase (AnP) de chez NEB est utilisée pour catalyser la déphosphorylation des extrémités 5' ou 3' de fragments d'ADN ou d'ARN ayant des groupements phosphates résiduels après coupure enzymatique par exemple. Au cours de ces travaux, la déphosphorylation de certains plasmides après digestion par des enzymes de restriction a été effectuée de façon à prévenir la recircularisation des plasmides au cours de réaction de ligation et favoriser l'incorporation d'un insert d'intérêt dans son vecteur. Les réactions de déphosphorylation sont réalisées en suivant les recommandations du fournisseur. Le mélange réactionnel contient le tampon Antartic Phosphatase Reaction Buffer à une concentration finale de 1X, additionné de 5 unités d'enzyme pour 1 pmol d'ADN. Les

réactions se déroulent généralement dans un volume final de 20 μL dont la qsp est ajustée avec de l'eau. Le mélange est incubé 30 min à 37°C. La réaction est ensuite arrêtée par inactivation via la chaleur pendant 2 min à 80°C.

II.11.3. Ligature des ADNs

L'enzyme T4 DNA Ligase de chez NEB est utilisée pour catalyser la formation de ponts phosphodiester entre deux ADNs dont les extrémités 5' phosphate et 3' hydroxyl se juxtaposent. Elle a notamment été utilisée pour la construction de certains plasmides au cours de ces travaux, après digestion de deux ou plusieurs fragments d'ADN et/ou de plasmide par des enzymes de restriction. Les réactions de ligature sont réalisées en suivant les recommandations du fournisseur. Le mix réactionnel contient le tampon T4 DNA Ligase Buffer à une concentration finale de 1X, additionné de 1 μL d'enzyme. Les réactions se déroulent généralement dans un volume final de 20 μL dont la qsp est faite avec de l'eau. La quantité d'ADN est à ajuster selon un ratio molaire le plus souvent d'1 mole de vecteur pour 3 moles d'insert en fonction de la taille des ADN. Le vecteur est généralement considéré comme un plasmide ou bien comme le plus grand des fragments à assembler. L'outil NEBioCalculator module « ligation » disponible sur le site NEB (<https://nebiocalculator.neb.com/#!/ligation>) est utilisé pour calculer les ratios molaires à ajouter pour chaque fragment. Le mélange peut être incubé 10 min à température ambiante ou toute la nuit à 16°C (pour les fragments dont les extrémités digérées sont cohésives) ; ou bien, à température ambiante pendant 2h ou également toute la nuit à 16°C (pour les fragments dont les extrémités digérées sont franches). La réaction peut être arrêtée par inactivation via la chaleur pendant 20 min à 65°C. Le mélange est ensuite maintenu dans la glace avant d'être utilisé pour transformer des cellules bactériennes d'*E. coli* compétentes.

II.12. Quantification des ADNs

La concentration d'ADN produit ou extrait est mesurée par spectrophotométrie à une densité optique de 260nm. Pour chaque échantillon, la mesure s'effectue par dépôt de 2 μL et lecture à l'aide de la plaque Take3 (Biotek) associée à l'appareil Epoch (Biotek). La

concentration en ADN est calculée à partir de la densité optique en utilisant les paramètres par défaut du logiciel Gen5 (Biotek) équipant le spectrophotomètre.

II.13. Electrophorèse sur gel d'agarose

Les ADN produits (plasmides digérés, matrice de recombinaison, produits PCR) sont analysés par électrophorèse sur gel d'agarose dont le maillage est ajusté entre 0,8 et 2,5% (m/v) en fonction de la taille des fragments à analyser. La migration est effectuée dans un tampon TAE 1X (Tris-Acétate 40 mM ; EDTA 1 mM ; pH 8.0), avec un voltage allant de 50 à 100 volts en fonction de la vitesse de migration désirée. Les gels sont chargés avec les échantillons d'ADN mélangés à un tampon de charge concentré à 6X (Gel Loading Dye Purple NEB) (compter 5 volumes d'ADN pour un volume de tampon de charge). Le Ficoll présent dans ce tampon permet le dépôt des échantillons au fond des puits du gel. Afin d'apprécier la taille des ADNs analysés, des marqueurs de taille sont ajoutés : le 1 kpb Plus (Invitrogen) ou le 100 pb (Promega et NEB). La présence d'acides nucléiques dans le gel est révélée par incubation des gels dans un bain de bromure d'éthidium (BET) (à 2 µg/mL final) ; après coloration, les fragments d'ADN sont visualisés par fluorescence, avec le système d'imagerie Vilbert Lourmat E-BOX VX2.

II.14. Préparation des génomes en blocs d'agarose

II.14.1. Extraction des génomes de mycoplasmes en bloc d'agarose

L'extraction des génomes de mycoplasme est réalisée selon un protocole décrit par Lartigue *et al.* 2007²¹⁷. En fonction de l'espèce de mycoplasme, 100mL de culture sont incubés à 37°C en milieu adéquat (*cf.* §I.2.). A pH~7 (soit environ 10⁹ cellules/mL), du chloramphénicol est ajouté à une concentration finale de 100 µg/mL et la culture est incubée pendant 1h à 2h30 à 37°C. Ce temps additionnel d'incubation est ajusté en fonction du temps de génération de l'espèce de mycoplasme cultivée. Les cellules sont ensuite lavées avec un tampon de lavage (le plus souvent avec le tampon S/T = Saccharose 0,5 M ; Tris-HCl 0,01 M ; pH 6,5) puis centrifugées à 5000g pendant 15 min à 10°C. Les cellules sont remises en suspension dans 500 µL de tampon de lavage et une gamme de dilutions au 1/2 est ensuite réalisée (non dilué, 1/2, 1/4, 1/8 et 1/16). Les cellules sont incubées 15 min à 50°C puis mélangées volume à volume à une solution d'agarose (agarose low-melting 2% (m/v), TAE 1X en surfusion à 50°C.

Tableau 6 : Composition des solutions du protocole « ED »

| Tampon de rinçage | |
|---------------------------|---------|
| Tris-HCl 1M pH = 8 | 1 mL |
| NaCl 5M | 3 mL |
| EDTA 0,5M pH = 9 | 20 mL |
| Filtration | 0,22µm |
| Qsp H ₂ O | 100mL |
| Tampon de lyse | |
| Triton X100 | 0,1 g |
| NaCl 5M | 20 mL |
| EDTA 0,5M pH = 9 | 20 mL |
| N-lauryl-sarkosyl | 1 g |
| Filtration | 0,22µm |
| Qsp H ₂ O | 100mL |
| Tampon de stockage | |
| EDTA 0,5M pH = 9 | 200 mL |
| N-lauryl-sarkosyl | 10 g |
| Qsp H ₂ O | 1000 mL |

Tableau 7 : Composition des solutions du protocole « Biorad »

| Cell Suspension Buffer Kit Biorad | |
|--|--------|
| Tris-HCl pH = 7.2 | 10 mM |
| NaCl | 20 mM |
| EDTA | 50 mM |
| Lysozyme Buffer Kit Biorad | |
| Tris-HCl pH = 7.2 | 10 mM |
| NaCl | 50 mM |
| Deoxycholate de sodium | 0,2% |
| Sodium Lauryl sarcosine | 0,5% |
| Proteinase K solution Kit Biorad | |
| EDTA | 100 mM |
| Deoxycholate de sodium | 0,2% |
| Sodium Lauryl sarcosine pH = 8 | 1% |
| Washing Buffer Kit Biorad | |
| Tris-HCl pH = 8 | 20 mM |
| EDTA | 50 mM |

Le mélange cellule/agarose est coulé dans des moules Bio-Rad (100 µl par moule). Les blocs sont solidifiés 20 min à 4°C, puis démoulés et soumis à l'action de détergents et de protéinase K pendant 24h à 50°C (kit CHEF Mammalian Genomic DNA Plug, Bio-Rad). Quatre lavages d'1h sous agitation, dans du tampon TE (Tris-HCl 20 mM pH 8 ; EDTA (Acide Ethylène Diamine Tétracétique) 50 mM) sont effectués puis les blocs sont stockés à 4°C dans ce même tampon.

II.14.2. Extraction des génomes de *Bacillus subtilis* en bloc d'agarose

Deux protocoles d'extraction de génomes de *B. subtilis* en bloc d'agarose ont été utilisés au cours de ces travaux. Ils diffèrent notamment par l'utilisation de tampons et solutions différentes qui sont répertoriées **Tableau 6** pour le protocole « ED » et **Tableau 7** pour le protocole « Biorad ».

- Protocole « ED » :

Ce protocole d'extraction des chromosomes bactériens de *B. subtilis* est un protocole dit « maison » transmis par nos collaborateurs de l'équipe Syber Micalis à Jouy-en-Josas. Il a été repris et adapté pour les besoins de ma thèse. Les cellules de *B. subtilis* sont cultivées dans 100 mL de milieu LB additionné de streptomycine à 0,75 µg/mL finale. Lorsque la DO à 600nm est comprise entre 1,5 et 1,6 (soit environ 10⁷ cellules/mL), du chloramphénicol est ajouté à une concentration finale de 180 µg/mL et la culture est incubée pendant 55 min à 37°C. Afin d'obtenir une gamme de blocs avec différentes concentrations en génome, la culture est répartie selon différents volumes en tubes Falcon (2 mL, 5 mL, 10 mL et 20 mL). En fonction du nombre de blocs à préparer pour une culture donnée, les volumes sont à ajuster (ici préparation décrite pour 5 blocs). Les cellules sont culotées par centrifugation à 5000g pendant 15 min à 20°C, lavées dans 5 mL de tampon de rinçage, centrifugées à nouveau et remises en suspension dans 300 µL de tampon de rinçage. Les cellules sont incubées 5 min à 37°C puis mélangées volume à volume à une solution d'agarose (agarose low-melting 2% (m/v), tampon de rinçage en surfusion à 50°C). Le mélange cellule/agarose est coulé dans des moules Bio-Rad (100 µl par moule). Les blocs sont solidifiés 20 min à température ambiante, puis démoulés et traités pendant 24h à 37°C avec une solution de lysozyme (2,5 mL de tampon de lyse + 250 µL de Lysozyme stock à 50mg/mL soit 5mg/mL de concentration finale + 250 µL

de RNase A à 200 µg/mL soit 20 µg/mL de concentration finale), afin de digérer la paroi formée de peptidoglycane des cellules bactériennes. Les blocs sont lavés avec le tampon de stockage, puis soumis à l'action de détergents et de protéinase K durant 48h à 50°C avec un changement de solution après 24h (2 mL de tampon de stockage + 2 mg de Proteinase K en poudre soit 1 mg/mL de concentration finale). Pour finir, les blocs d'agarose sont lavés 2 fois 1h dans du tampon de stockage et stockés à 4°C.

- Protocole « Biorad » :

L'extraction des génomes de *B. subtilis* est également réalisée selon un protocole développé au laboratoire (inspiré du protocole « ED ») et utilisant le kit CHEF Bacterial Genomic DNA Plug de chez Bio-Rad. Les cellules de *B. subtilis* sont cultivées dans 100 mL de milieu LB additionné de streptomycine à 0,75 µg/mL finale. Lorsque la DO à 600nm est comprise entre 1,5 et 1,6 (soit environ 10^7 cellules/mL), du chloramphénicol est ajouté à une concentration finale de 180 µg/mL et la culture est incubée pendant 55 min à 37°C. Afin d'obtenir une gamme de blocs avec différentes concentrations en génome la culture est répartie selon différents volumes en tubes Falcon (2 mL, 5 mL, 10 mL et 20 mL). En fonction du nombre de blocs à préparer pour une culture donnée, les volumes sont à ajuster (ici préparation décrite pour 5 blocs). Les cellules sont culotées par centrifugation à 5000g pendant 15 min à 20°C, puis remises en suspension dans 300 µL de Cell Suspension Buffer. Les cellules sont incubées 10 min à 50°C puis mélangées volume à volume à une solution d'agarose (agarose low-melting 2% (m/v), TAE 1X en surfusion à 50°C). Le mélange cellule/agarose est coulé dans des moules Bio-Rad (100 µl par moule). Les blocs sont solidifiés 20 min à température ambiante, puis démoulés et traités pendant 2h à 37°C avec une solution de lysozyme (1,25 mL de Lysozyme Buffer + 50µL de Lysozyme stock à 25 mg/mL soit 1 mg/mL de concentration finale), afin de digérer la paroi formée de peptidoglycane des cellules bactériennes. Les blocs sont lavés avec de l'eau stérile, puis soumis à l'action de détergents et de protéinase K durant 48h à 50°C avec un changement de solution après 24h (1,25 mL de Proteinase K Reaction Buffer + 50 µL de Proteinase K à 25 mg/mL soit 1 mg/mL de concentration finale). Pour finir, les blocs d'agarose sont lavés 4 fois 1h dans du Tris-HCl 20 mM, pH 8.0 ; EDTA 50 mM et stockés à 4°C.

II.14.3. Extraction des génomes des transformants levures en bloc d'agarose

L'ADNg total des levures hébergeant des génomes bactériens ou des fragments de génomes bactériens est isolé en bloc d'agarose en suivant le protocole « Preparation of agarose embedded yeast DNA » extrait du manuel Bio-Rad CHEF-DR III^{215,216}. Les cellules sont cultivées dans 100 mL de milieu minimal SD additionné du « DO complement » adéquat (cf. §II.4.) à 30°C. Lorsque la DO à 600nm est comprise entre 3 et 4, les cellules sont centrifugées, et les culots repris dans 10 mL d'EDTA froid (50 mM, pH 8, 4°C). La concentration cellulaire est déterminée sur cellule de Malassez et le volume correspondant à 6×10^9 cellules est prélevé et centrifugé à 5000g pendant 10 min et à 4°C. En fonction du nombre de blocs à préparer pour une culture de levure donnée, les volumes et les pesées sont à ajuster. Pour la préparation de 10 blocs d'agarose, le culot cellulaire est remis en suspension dans 500 µL de tampon (Tris 10 mM pH 7.2 ; NaCl 20 mM ; EDTA 50 mM) et incubé 10 min à 50°C. 5 mg de zymolyase 100T sont ensuite ajoutés à la suspension cellulaire ainsi que 500 µL de solution d'agarose (agarose low-melting 2% (m/v), TAE 1X) maintenue en surfusion à 50°C. Le mélange cellule/agarose (1000 µL) est coulé dans des moules Bio-Rad (100 µL par moule) puis laissé 20 min à 4°C. Après polymérisation, les blocs sont incubés 2h à 37°C dans une solution de Tris-HCl 1 M pH 7,5 ; EDTA 50 mM pH 8 contenant de la zymolyase 100T (7,5 mg). Les blocs sont lavés avec de l'eau stérile, puis soumis à l'action de détergents et de protéinase K durant 48h avec un changement de solution après 24h (kit CHEF Mammalian Genomic DNA Plug, Bio-Rad). Pour finir, les blocs d'agarose sont lavés 4 fois 1h dans du Tris-HCl 20 mM, pH 8.0 ; EDTA 50 mM et stockés à 4°C.

II.15. Digestion des blocs d'agarose

Au cours des expériences de clonage de fragments de génomes de *B. subtilis* par TAR-Cloning ou bien, lors des analyses par électrophorèse en champs pulsé, il a été nécessaire de digérer les génomes emprisonnés en blocs d'agarose par des enzymes de restriction. Par ailleurs, avant les expériences de clonage de génomes entiers ou de fragments de génome chez la levure, et avant les essais de transplantation, les blocs ont été digérés avec l'enzyme β -agarase afin de libérer l'ADN bactérien contenu dans les blocs.

II.15.1. Digestion des blocs d'agarose par restriction enzymatique

La digestion de l'ADN contenu dans des blocs d'agarose par des enzymes de restriction se fait généralement sur des moitiés de blocs (volume 50 μ L). Les blocs sont préparés en suivant une série de lavages à température ambiante et sous agitation douce avec différents tampons. Brièvement, chaque $\frac{1}{2}$ bloc est lavé une fois pendant 1h avec du tampon de lavage 1X (Tris-HCl 20 mM, EDTA 50 mM, pH8) additionné de PMSF (fluorure de phénylméthylsulfonyl) à 1 mM final, puis un second lavage avec du tampon de lavage 1X est effectué pendant 1h. Deux lavages d'1h dans du tampon de lavage 0,1X (Tris-HCl 2 mM pH8 ; EDTA 5 mM) sont ensuite effectués. Les blocs sont ensuite équilibrés pendant 1h avec le tampon de restriction à 1X final approprié à l'enzyme de restriction choisie. Pour finir, l'ADN présent dans chaque demi-bloc est digéré toute la nuit à la température optimale de l'enzyme sélectionnée et avec un mélange réactionnel composé de 30 à 50 unités d'enzyme de restriction, du tampon de restriction 1X adéquat et d'eau (qsp un volume final de 500 μ L).

II.15.2. Digestion des blocs d'agarose par l'enzyme β -agarase

Chaque bloc est lavé deux fois 30 min dans du tampon de lavage 0,1X (Tris-HCl 2 mM pH8 ; EDTA 5 mM). Après ajout de 10 μ L par bloc de tampon β -agarase I Buffer 10X (NEB, Tris-HCl 10 mM ; EDTA 1 mM, pH6,5), les blocs sont successivement placés 5 min à 42°C, 8 min à 65°C et 10 min à 42°C. La β -agarase I est ensuite ajoutée (3U/bloc) et le mélange réactionnel est incubé 5h à 37°C. Les blocs digérés sont conservés à température ambiante avant d'être utilisés.

II.16. Migration des grands ADNs par électrophorèse en champs pulsé (PFGE)

L'électrophorèse en champs pulsés (PFGE) est une méthode de migration qui permet de séparer et d'estimer la taille de très grandes molécules d'ADN (>50 kpb). Lors de cette électrophorèse, l'orientation du champ électrique varie au cours du temps ; chaque changement du champ électrique permet de réorienter la molécule d'ADN linéaire et ainsi sa migration à travers le maillage du gel d'agarose. Au cours de ces travaux, cette méthode a été utilisée afin de vérifier la qualité des génomes bactériens emprisonnés en blocs d'agarose, ou

bien, afin d'estimer la taille des fragments génomiques ou génomes entiers clonés dans la levure *S. cerevisiae*.

II.16.1. Evaluation de la qualité des génomes

Les blocs d'agarose contenant l'ADN de mycoplasmes ou l'ADN de *B. subtilis* sont tout d'abord lavés puis digérés suivant les descriptions du paragraphe II.9. En fonction de la séquence du génome et du profil de restriction souhaité une ou plusieurs enzymes de restriction sont utilisées. Les blocs d'agarose sont ensuite chargés sur un gel d'agarose Pulsed Field Certified (Bio-Rad) 1% réalisé avec du TBE 0,5X (Tris 44,5 mM, Acide borique 44,5 mM, EDTA 1 mM, pH 8,3, Bio-Rad). La migration est effectuée pendant 22 à 24h dans deux litres de tampon TBE 0,5X à 14°C avec le système d'électrophorèse en champ pulsé CHEF-DR III (Bio-Rad). Le tampon est mis en circulation via un circuit fermé qui le fait passer dans un module de refroidissement permettant de le maintenir à température constante (14°C). L'appareil est réglé avec les paramètres suivants : angle de 120°, voltage de 6V/cm, temps de commutation initial = 50s, temps de commutation final = 90s. Le gel est révélé par coloration au SYBR™Gold 1X (Invitrogen) et la présence des molécules d'ADN est détectée grâce au système d'imagerie Vilbert Lourmat E-BOX VX2. Le marqueur de taille utilisé est le CHEF DNA Size Marker *S. cerevisiae* Ladder (Bio-Rad).

II.16.2. Analyse des transformants levures propageant des génomes de mycoplasmes

Cette analyse par électrophorèse en champ pulsé est réalisée sur les demi-blocs d'agarose contenant l'ADN de levures transformées avec un génome de mycoplasme. Elle consiste à exécuter une série d'étapes dans le but d'éliminer l'ADN de levure, puis à faire migrer le génome de mycoplasme selon sa taille (linéarisation ou bien digestion du génome en plusieurs morceaux).

Pour la plupart des génomes de mycoplasmes clonés dans la levure et analysés au cours de ces travaux, les blocs d'agarose sont traités avec un cocktail d'enzyme de restriction (AsiSI, FseI, RsrII) hydrolysant exclusivement l'ADN génomique de levure. Ils sont ensuite chargés sur gel d'électrophorèse classique (agarose 1% (p/v), TAE1 X, 120V, 120min) afin

Tableau 8 : Enzymes de restriction utilisées lors de l'analyse en PFGE des génomes de mycoplasme clonés dans la levure.

| Génome de mycoplasme | Enzyme(s) utilisée(s) | Fragments obtenus |
|--|-----------------------|--------------------------------|
| <i>M. capricolum</i> subsp. <i>capricolum</i> | BssHII | 1010 kb (génomme linéarisé) |
| <i>M. mycoides</i> subsp. <i>capri</i> | SfoI | 1084 kb (deux fragments) |
| <i>M. capricolum</i> subsp. <i>capripneumoniae</i> | BssHII | 1016 kb (génomme linéarisé) |
| <i>M. mycoides</i> subsp. <i>mycoides</i> | BssHII | 1200 kb (génomme linéarisé) |
| | EagI | 561 et 627 kb (deux fragments) |
| <i>M. gallisepticum</i> | SacII | 449 et 536 kb (deux fragments) |
| <i>M. agalactiae</i> | AscI | 877 kb (génomme linéarisé) |

d'éliminer une grande partie de l'ADN levure et d'éviter que ce dernier ne masque le profil de digestion du génome bactérien que l'on souhaite analyser. L'analyse des demi-blocs contenant l'ADN de levures transformées avec un génome de *M. gallisepticum* fait exception au suivi de cette procédure. Cette étape d'élimination est effectuée directement avec une première migration par PFGE (cf. §II.16.1.), le cocktail d'enzymes étant non utilisable, les sites de restriction des enzymes AsiSI et RsrII sont présents sur le génome de *M. gallisepticum*. Au cours de cette étape, les seize chromosomes linéaires de la levure migrent tandis que le génome circulaire reste dans les blocs d'agarose. Ces derniers sont ensuite retirés du gel, puis lavés et digérés suivant les instructions décrites §II.15.1. Le choix de l'enzyme de restriction est fonction de la présence d'un ou plusieurs sites compatibles sur le génome de mycoplasme à analyser et du profil de restriction souhaité (**Tableau 8**). Les génomes présents dans les blocs sont finalement analysés par PFGE de la même manière que celle décrite pour l'évaluation de la qualité des génomes (cf. §II.16.1.).

II.16.3 Analyse des transformants levures propageant des fragments de génomes de *B. subtilis*

L'analyse par électrophorèse en champ pulsé réalisée sur les demi-blocs d'agarose contenant l'ADN de levures transformées avec des fragments de génome de *B. subtilis* par TAR-Cloning suit un protocole dont la mise au point a été effectuée au cours de cette thèse. La procédure a dû être modifiée à cause de contraintes liées à la présence de beaucoup de sites de restriction pour un grand nombre d'enzymes sur le génome de *B. subtilis*. Le protocole classiquement utilisé au laboratoire sur des clones levures hébergeant des génomes de mycoplasmes requiert 2 étapes. La 1^{ère} étape consiste à identifier des enzymes de restriction capables de cliver spécifiquement les chromosomes de levure afin de débarrasser les blocs d'agarose de cet ADN « contaminant » au cours d'une 1^{ère} électrophorèse. Cette étape ne peut pas être réalisée sur des clones levures hébergeant un génome de *B. subtilis* (2,86 Mb) car ce dernier contient des sites de restriction variés et en très grand nombre. Seulement 4 enzymes clivant spécifiquement l'ADN levure ont pu être identifiées (I-PopI, I-Scel, PI-PSpl, PI-Scel) ; elles n'ont pas été utilisées car elles sont peu classiques et possèdent des sites de reconnaissance d'une dizaine de paires de bases (sites rares).

Tableau 9 : Enzymes de restriction et paramètres de migration utilisés pour l'analyse des différents fragments de génome de *B. subtilis* clonés dans la levure.

| Fragments du génome de <i>Bsu</i> | Enzyme(s) utilisée(s) | Paramètres du PFGE | Marqueur de taille | Fragments attendus |
|-----------------------------------|-----------------------|-------------------------|--|---------------------|
| Fragment 1 | SgrDI | Paramètres ^a | <i>S. cerevisiae</i> Ladder (Bio-Rad) | 1499 kb |
| Fragment 2 | SgrDI | Paramètres ^a | <i>S. cerevisiae</i> Ladder (Bio-Rad) | 390 kb |
| Fragment 3 | SgrDI | Paramètres ^b | MidRange PFG (NEB) | 214 kb |
| | Ascl | | | 169,5 kb et 48,4kb |
| Fragment 4 | SgrDI | Paramètres ^b | MidRange PFG (NEB) et LowRange PFG (NEB) | 141 kb |
| | AsiSI | Paramètres ^b | | |
| Fragment 5 | SgrDI | Paramètres ^b | MidRange PFG (NEB) | 133 kb |
| | Ascl | | | 121,8 kb et 14,6 kb |
| Fragment 7 | SgrDI | Paramètres ^a | <i>S. cerevisiae</i> Ladder (Bio-Rad) | 482 kb |
| Sous-fragment 1.1 | Fsel | Paramètres ^a | <i>S. cerevisiae</i> Ladder (Bio-Rad) | 367 kb |
| Sous-fragment 1.2 | Fsel | Paramètres ^b | MidRange PFG (NEB) | 83 kb |
| Sous-fragment 1.3 | Fsel | Paramètres ^a | <i>S. cerevisiae</i> Ladder (Bio-Rad) | 514 kb |
| Sous-fragment 1.4 | Fsel | Paramètres ^b | MidRange PFG (NEB) | 159 kb |
| Sous-fragment 1.5 | Fsel | Paramètres ^b | MidRange PFG (NEB) | 111 kb |
| Sous-fragment 1.6 | Fsel | Paramètres ^b | MidRange PFG (NEB) | 130 kb |
| Sous-fragment 1.7 | Fsel | Paramètres ^a | <i>S. cerevisiae</i> Ladder (Bio-Rad) | 269 kb |
| Sous-fragment 7.2 | Fsel | Paramètres ^b | MidRange PFG (NEB) | 44 kb |
| Sous-fragment 7.3 | Fsel | Paramètres ^b | MidRange PFG (NEB) | 63 kb |
| Sous-fragment 7.5 | AsiSI | Paramètres ^b | MidRange PFG (NEB) | 252 kb |
| Sous-fragment 7.6 | AsiSI | Paramètres ^b | MidRange PFG (NEB) | 128 kb |
| Sous-fragment 7.9 | AsiSI | Paramètres ^b | MidRange PFG (NEB) | 74 kb |
| Sous-fragment 7.11 | AsiSI | Paramètres ^b | MidRange PFG (NEB) | 141 kb |

Paramètres ^a : Temps de commutation : 50-90s ; Temps de migration : 24h ; Angle : 120° ; Voltage 6V/cm.
 Paramètres ^b : Temps de commutation : 1-25s ; Temps de migration : 21h ; Angle : 120° ; Voltage 6V/cm

Le protocole mis en place requiert 6 jours et trois migrations sont nécessaires pour visualiser correctement le génome ou des fragments de génomes de *B. subtilis* cloné dans la levure sur gel d'agarose :

- Chaque demi-bloc est lavé 1h dans 1 mL de tampon de lavage 1X (Tris-HCl 20 mM, pH 7.5 ; EDTA 50 mM, pH 7.5) puis chargé directement sur un gel d'électrophorèse en champ pulsé afin d'effectuer une première migration telle que décrite §II.16.1. Au cours de cette première électrophorèse, l'ADN levure est éliminé des blocs, car seul l'ADN linéaire peut migrer hors des blocs d'agarose et être propagé dans le gel.
- Afin d'optimiser l'élimination de l'ADN levure, les blocs subissent une seconde migration par PFGE (§II.16.1.) après avoir subi au préalable, une série de lavages et un traitement à la désoxyribonucléase Plasmid-Safe (Epicentre). Cette enzyme digère exclusivement l'ADN double brin linéaire, ce qui permet d'éliminer l'ADN levure restant dans les blocs. Brièvement, les blocs sont lavés 1h dans 1 mL de tampon de lavage 1X (Tris-HCl 20 mM, pH 7.5 ; EDTA 50 mM, pH 7.5) contenant 1 mM de fluorure de phénylméthylsulfonyl (PMSF), puis 1h dans 1 mL de tampon de lavage 1X, et enfin deux fois 1h dans 1 mL de tampon de lavage 0,1X. Avant d'être soumis à l'action de la DNase, l'ADN contenu dans les blocs est lavé une fois 1h dans 0,5 mL de tampon de réaction (Tris-Acétate 33 mM, pH 7,5 ; Acétate de potassium 66 mM ; Acétate de magnésium 10 mM ; Dithiothréitol 0,6 mM ; Adénosine-Triphosphate 1 mM), et digéré avec 50 unités d'enzyme Plasmid-Safe sur la nuit à 37°C.
- Enfin, une troisième migration est finalement effectuée afin de faire migrer les fragments de génomes de *B. subtilis*. Les blocs sont lavés et digérés suivant les instructions décrites §II.15.1. Le choix de l'enzyme de restriction est fonction de la présence d'un ou plusieurs sites compatibles sur le fragment à analyser et du profil de restriction souhaité. Les blocs sont analysés par PFGE en ajustant les paramètres de migration et avec plusieurs marqueurs de taille afin de mieux évaluer la taille des différents fragments de génome analysés (**Tableau 9**).

Il est important de préciser qu'à la fin de chacune des migrations, les blocs sont à chaque fois retirés du gel pour être lavés puis traités, avant d'être rechargés sur un nouveau gel propice à la migration suivante.

III. Transformations microbiennes

III.1. Transformations bactériennes d'*Escherichia coli*

Les souches d'*E. coli* décrites paragraphe I.1. sont transformées suivant deux types de protocoles et en suivant les recommandations des fournisseurs.

III.1.1. Transformation par électroporation

Les cellules d'*E. coli* sont tout d'abord décongelées dans la glace, puis 20 µL de cellules sont transférés dans un tube eppendorf contenant l'ADN à transformer (maximum 5 ng de plasmide ou mélange de ligation). Le mélange de transformation est transféré dans une cuvette d'électroporation (de 1 mm ou de 2 mm) maintenue dans la glace. En fonction de la souche utilisée, les cellules sont soumises à un choc électrique à l'aide de l'électroporateur Gene Pulser Xcell (Bio-Rad) en suivant les paramètres préconisés par chaque fournisseur (épaisseur de la cuvette, voltage, résistance et capacitance). Après transformation, les bactéries sont incubées 1h à 37°C dans 500 µL de milieu SOC (NEB). 250 µL de cette culture sont étalés sur boîtes LB agar contenant l'antibiotique approprié. Les boîtes sont incubées au moins une nuit à 37°C, jusqu'à apparition des colonies.

III.1.2. Transformation par choc thermique

Les cellules d'*E. coli* sont tout d'abord décongelées dans la glace, puis 20 µL de cellules sont transférés dans un tube eppendorf contenant l'ADN à transformer (maximum 5 ng de plasmide ou mélange de ligation). Le mélange de transformation est placé 30 min dans la glace puis les cellules subissent un choc thermique pendant 45s à 42°C au bain-marie. Le mélange est momentanément réservé sur la glace (5 min) puis 500 µL de SOC (NEB) sont ajoutés afin de favoriser le processus de récupération des bactéries après transformation. Les cellules ainsi reprises sont transférées dans des tubes de 10 mL et incubées à 37°C sous agitation à 180rpm pendant 1h. 250 µL des bactéries transformées sont étalés sur milieu LB gélosé sélectif ; l'antibiotique choisi et sa concentration seront fonction du marqueur de résistance présent sur la construction ou le plasmide cloné. Les boîtes sont incubées au moins une nuit à 37°C, jusqu'à apparition des colonies.

III.2. Transformations bactériennes des mycoplasmes

Les trois espèces de mycoplasme *M. bovis* (*Mbov*), *M. agalactiae* (*Magal*) et *M. gallisepticum* (*Mgal*) ont été transformées au cours de ces travaux afin de produire des mutants désaminés et donc mutés sur certains *loci*. Les mutants obtenus ont été utilisés pour la mise en place de la méthode de clonage par CReasPy-Fusion.

III.2.1. Transformation de *Mycoplasma bovis* et *Mycoplasma agalactiae*

Les espèces de mycoplasmes *Mbov* et *Maga* sont transformées suivant un protocole similaire et de la même manière que la méthode décrite par Zhu *et al.* 2020²⁶⁶. A partir d'un stock de *Mbov* ou de *Maga* conservé au congélateur -80°C, les cellules sont cultivées en milieu SP4-Bovis (*cf.* §1.2.) pendant 48h avant transformation. Une pré-culture diluée en milieu SP4-Bovis est ensemencée dans un premier temps, puis diluée le lendemain dans un volume plus important. Lorsque la culture ensemencée la veille atteint une phase de croissance tardive (pH environ égal à 7, les cellules sont collectées par centrifugation pendant 20 min à 10000g et à 10°C. Après élimination du surnageant, le culot est lavé deux fois avec 12 mL de tampon froid dPBS 1X (Dulbecco's phosphate-buffered saline, Gibco refroidi sur glace au préalable). Au cours de ces étapes de lavage et de centrifugation successives à 10000g et à 10°C, les cellules ne sont pas remises en suspension. Les bactéries collectées sont resuspendues dans 1 mL de CaCl₂ à 0,1 mol.L⁻¹ et incubées sur glace pendant 30 min. Pour une condition de transformation, 100 µL de cellules de mycoplasme sont mélangés avec 3 à 10 µg de plasmide à transformer additionnés de 10 µg d'ARNt de levure (Thermo Fisher, faisant office de « carrier »). Elles sont ensuite transformées par ajout de 1 mL d'une solution de PEG (PEG 8000 Thermo Fisher, 50% (m/v) dans du tampon dPBS 1X), pendant 1 min à température ambiante. Le contact avec le PEG est ensuite stoppé en ajoutant 5 mL de milieu SP4-Bovis, puis le mélange est incubé pendant 3h à 37°C. Les bactéries sont ensuite centrifugées pendant 8 min à 8000g et remises en suspension dans 1 mL de SP4-Bovis, puis étalées sur des boîtes SP4-Bovis contenant l'antibiotique approprié à la sélection du plasmide transformé. Les transformants sont visibles et peuvent être dénombrés après 3 à 7 jours d'incubation à 37°C. Les colonies sont ensuite dénombrées puis repiquées dans 1 mL de milieu SP4-Bovis pendant trois passages successifs avant analyse (un passage correspondant à une dilution de la culture au 1/100^{ième} suivi d'une incubation de 48h à 37°C).

III.2.2. Transformation de *Mycoplasma gallisepticum*

A partir d'un stock de *Mgal* conservé au congélateur -80°C , les cellules sont cultivées pendant 36h avant transformation. Une pré-culture diluée au 1/10ème en milieu Hayflick modifié (m-Hayflick) estensemencée dans un premier temps, puis diluée le lendemain au 1/10ème dans un volume plus important. Lorsque le pH de la culture ensemencée la veille est compris entre 6,2 et 6,5, des aliquots de culture de 10 mL sont centrifugés pendant 15 min à $6000g$ et à 10°C . Un aliquot de 10 mL correspond à une condition de transformation. Après élimination du surnageant le culot est lavé avec 5 mL de tampon HBSS 1X (Solution commerciale HBSS 10X Sigma), puis une seconde centrifugation est effectuée. Les cellules collectées sont resuspendues dans $250\ \mu\text{L}$ de CaCl_2 à $0,1\ \text{mol.L}^{-1}$ et incubées sur glace pendant 30 min. Les bactéries traitées sont mélangées avec $10\ \mu\text{g}$ de plasmide à transformer additionnés de $10\ \mu\text{g}$ d'ARNt de levure (Thermo Fisher, faisant office de « carrier »). Elles sont ensuite transformées par ajout de 2 mL d'une solution de PEG (PEG 6000 Sigma, 40% (m/v) dans du tampon HBSS 1X), pendant 2 min à température ambiante. Le contact avec le PEG est ensuite stoppé en ajoutant 20 mL de tampon HBSS 1X. Le mélange est centrifugé pendant ($12000g$, 15 min, 10°C), puis les cellules collectées sont remises en suspension dans 1 mL de milieu m-Hayflick au préalable réchauffé à 37°C et incubées 2h à 37°C . Les bactéries sont ensuite étalées sur des boîtes m-Hayflick contenant l'antibiotique approprié à la sélection du plasmide transformé. Les transformants sont visibles et peuvent être dénombrés après 10 à 15 jours d'incubation à 37°C dans une étuve à 5% de CO_2 . Les colonies sont ensuite dénombrées puis repiquées dans 1 mL de milieu m-Hayflick pendant trois passages successifs avant analyse (un passage correspondant à une dilution de la culture au 1/100^{ième} suivi d'une incubation de 48h à 37°C).

III.3. Transformation bactérienne de *Bacillus subtilis*

Classiquement *B. subtilis* est transformé grâce à sa capacité à être naturellement compétent en début de phase stationnaire. Au cours de ces travaux, deux protocoles ont été utilisés.

Tableau 10 : Composition des solutions utilisées pour la transformation de *B. subtilis* avec le protocole MG1-MG2

| MM 10X | |
|---|--------|
| Sulfate d'ammonium (NH ₄) ₂ SO ₄ (VWR Chemicals) | 2 g |
| Tri sodium Citrate dihydraté Na ₃ C ₆ H ₅ O ₇ (VWR Chemicals) | 1 g |
| Potassium Phosphate Dibasique K ₂ HPO ₄ anhydre (Sigma) | 10,7 g |
| Potassium Phosphate Monobasique KH ₂ PO ₄ (Sigma) | 6 g |
| Qsp H ₂ O | 100 mL |
| Autoclavage 20 min 120°C | |
| MG0 | |
| MM10X | 10 mL |
| Glucose 20% (m/v) | 2,5 mL |
| Sulfate de Magnésium MgSO ₄ 1M | 160 µL |
| Qsp H ₂ O | 100 mL |
| MG1 | |
| MG0 | 50 mL |
| Hydrolysate de caséine 5% (m/v) | 250 µL |
| Extrait de levure 5% (m/v) | 1 mL |
| Tryptophane 0,5% (m/v) si souche auxotrophe | 2 mL |
| Filtration 0,22µm – Stock à -20°C | |
| MG2 | |
| MG0 | 50 mL |
| Hydrolysate de caséine 5% (m/v) | 125 µL |
| Extrait de levure 5% (m/v) | 250 µL |
| Sulfate de Magnésium MgSO ₄ 1M | 125µL |
| Filtration 0,22µm – Stock à -20°C | |
| Nitrate de calcium Ca(NO ₃) ₂ 0,5M filtré 0,22µm | 250 µL |

Tableau 11 : Composition des solutions utilisées pour la transformation de *B. subtilis* avec le protocole MC

| MC 10X | |
|---|---------|
| Glucose 20% (m/v) | 10 mL |
| Glutamate de Potassium 10% (m/v) | 2 mL |
| Citrate de Sodium Trisodique 0,1M | 3 mL |
| Citrate d'Ammonium Ferrique 10g/L | 220 µL |
| Hydrolysate de caséine 5% (m/v) | 2 mL |
| Potassium Phosphate Dibasique K ₂ HPO ₄ anhydre (Sigma) | 1,06 g |
| Potassium Phosphate Monobasique KH ₂ PO ₄ (Sigma) | 0,52 g |
| Tryptophane 0,5% (m/v) si souche auxotrophe | 2 mL |
| Qsp H ₂ O | 100 mL |
| Filtration 0,22µm – Stock à -20°C | |
| MC 1X | |
| MC 10X | 1 mL |
| Sulfate de Magnésium MgSO ₄ 1M (extemporanément) | 33,5 µL |
| Qsp H ₂ O | 10 mL |

- Protocole MG1-MG2 :

B. subtilis est transformé suivant un protocole développé par Anagnostopoulos et Spizizen en 1961¹⁴. Au cours de ce protocole, les cellules bactériennes sont rendues compétentes par croissances successives dans un milieu minimum enrichi « MG1 » puis après dilution dans un second milieu carencé « MG2 ». 2 mL de milieu MG1 sontensemencés à partir d'une colonie fraîche et cultivés pendant minimum 4h30 à 37°C sous agitation (220 rpm). Puis 200 µL de cette culture sont transférés dans 1,8 mL de milieu MG2 (dilution au 1/10^{ème}) et cultivés pendant 1h30 à 37°C sous agitation : c'est à cette étape que les cellules deviennent compétentes. Pour une condition de transformation, 100 µL de culture en MG2 sont ajoutés à l'ADN à transformer (10 ng à 2 µg d'ADNg total ou 0,5 à 4 µg de plasmide circulaire ou linéaire) puis incubés 20 min à 37°C au bain-marie. Après cette dernière étape, le mélange est étalé à l'aide de billes de verre sur boîtes gélosées LB additionnées de l'antibiotique approprié. Les boîtes sont incubées au minimum une nuit à 37°C jusqu'à apparition des colonies. La préparation des milieux « MG1 » et « MG2 » est effectuée à partir d'une solution mère MM 10X. Cette solution est la base du milieu secondaire MG0 qui est scindé en deux : 50 mL pour la solution MG1 et 50 mL pour la solution MG2. Les solutions MG1 et MG2 sont filtrées grâce à un filtre 0,22 µm. Il est important de noter que l'ajout du nitrate de calcium doit se faire en dernier après filtration. Un précipité doit apparaître et ce dernier est crucial pour assurer l'efficacité du protocole de transformation. Les milieux sont conservés à -20°C et sont décongelés avant chaque transformation sur la nuit à 4°C (**Tableau 10**).

- Protocole MC « simple » :

Ce protocole de transformation simplifié de *B. subtilis* est caractérisé par l'utilisation d'un seul milieu minimum MC, et a notamment été décrit par Konkol *et al.* en 2013²⁶⁷. 2 mL de milieu MC 1X sontensemencés avec une colonie fraîche et cultivés pendant minimum 6h à 37°C sous agitation (220 rpm). Pour une condition de transformation, 200 µL de cette culture sont ajoutés à l'ADN à transformer (10 ng à 2 µg ADNg total ou 0,5 à 4 µg de plasmide circulaire ou linéaire) puis incubés 2h à 37°C au bain-marie. Le mélange est étalé à l'aide de billes de verres sur boîtes gélosées LB additionnées de l'antibiotique approprié. Les boîtes sont incubées au minimum une nuit à 37°C jusqu'à apparition des colonies.

La préparation du milieu MC 1X se fait extemporanément à partir d'une solution mère MC 10X (**Tableau 11**) et complétement avec du MgSO_4 1mol.L^{-1} . Le milieu de compétence MC 10X est conservé à -20°C et doit être décongelé avant chaque transformation une nuit à 4°C .

III.4. Transformations de *Saccharomyces cerevisiae*

III.4.1. Transformation de plasmides par le Lithium Acétate/ADN simple brin/PEG

Ce protocole de transformation de levures par des plasmides ou de l'ADN double brin linéaire est adapté d'après la méthode de Gietz *et al.* 1995²⁶⁸. Il s'agit d'une méthode de transformation au polyéthylène glycol (PEG) au cours de laquelle les levures sont traitées avec de l'acétate de lithium afin de fragiliser leurs parois, puis subissent un choc thermique en présence de PEG 4000 et d'ADN simple brin (appelé ADN « carrier ») favorisant l'entrée de l'ADN dans la levure. En fonction de la souche de levure à transformer, les cellules sont cultivées à 30°C en milieu riche YPDA ou minimal SD-TRP/TRP-URA jusqu'à obtenir une densité optique (DO) à 600nm égale à 1. La culture est diluée au quart dans un volume final de 20 mL de milieu YPDA, puis incubée jusqu'à 4h à 30°C sous agitation (190 rpm). Lorsque la DO atteint à nouveau 1, les levures sont culotées par centrifugation à 5000g pendant 2 min. 2,5 mL de culture sont utilisés pour une condition de transformation. Le culot est lavé avec 1 mL de solution d'acétate de lithium-TE (1X Tris-HCl 1M pH7,5 ; EDTA 0,5M pH7,5 ; acétate de lithium 0,1M), puis une seconde centrifugation de 5000g pendant 2 min est effectuée afin de collecter à nouveau les cellules. Ces dernières sont reprises dans 1 mL de solution d'acétate de lithium-TE et incubées 1h dans une étuve à 30°C . Après incubation, les levures sont centrifugées et remises en suspension dans 50 μL de solution d'acétate de lithium-TE de manière à obtenir une suspension cellulaire à $\sim 2 \times 10^9$ cellules/mL. Les cellules et l'ADN plasmidique (0,5 à 2 μg) sont mélangés en présence d'ADN simple brin dit « ADN carrier » de sperme de saumon (10 mg/mL) au préalable dénaturé pendant 10 min à 100°C puis conservé dans la glace. La réaction de transformation est enclenchée par ajout de 0,5 mL d'une solution de PEG 40%-Acétate de lithium-TE (PEG 4000 40% (m/v) ; 1X Tris-HCl 1M pH7,5 ; EDTA 0,5M pH7,5 ; acétate de lithium 0,1M) et de 10% (v/v) de DMSO. Les tubes sont incubés successivement 30 min à 30°C sans agitation, puis pendant 25 min à 42°C sans agitation. Les cellules sont finalement

Tableau 12 : Composition des solutions de transformation de sphéroplastes levures au PEG : méthodes de clonage classique, de TAR-Cloning et de CReasPy-Cloning

| SPEM | |
|--|---------|
| Sorbitol (Sigma) | 91 g |
| Na ₂ HPO ₄ , 2H ₂ O | 0,69 g |
| NaH ₂ PO ₄ , 2H ₂ O | 0,18 g |
| EDTA 0,5 M pH = 7,5 | 10 mL |
| Qsp H ₂ O | 500 mL |
| Filtration 0,22µm | |
| STC | |
| Sorbitol (Sigma) | 18,2 g |
| Tris-HCl 1M pH = 7,5 | 1 mL |
| CaCl ₂ 1M | 1 mL |
| Qsp H ₂ O | 100 mL |
| Filtration 0,22µm | |
| SOS | |
| Sorbitol (Sigma) | 9,1 g |
| Extrait de levure (Gibco) | 0,125 g |
| Bacto peptone (Difco) | 0,25 g |
| CaCl ₂ 1M | 300µL |
| Qsp H ₂ O | 50 mL |
| Filtration 0,22µm | |
| PEG | |
| PEG 8000 (Thermo Fisher) | 3 g |
| Tris-HCl 1M pH = 7,5 | 150 µL |
| CaCl ₂ 1M | 150 µL |
| Qsp H ₂ O | 15 mL |
| Filtration 0,22µm | |

centrifugées pendant 3 min à 5000g et reprises dans 200 μ L d'eau et étalées sur milieu sélectif adéquat à 30°C pendant 48h au minimum.

III.4.2. Clonage par transformation de sphéroplastes levures au PEG

- Protocole appliqué pour les méthodes de clonage classique, de TAR-Cloning et de CReasPy-Cloning :

Ce protocole de transformation dérive de celui de Kouprina et Larionov, 2008²⁶⁴ et consiste à transformer des sphéroplastes de levure obtenus par traitement à la zymolyase, en utilisant un agent fusogène le PEG. Cette méthode permet de transformer la levure *S. cerevisiae* avec des molécules d'ADN de grande taille (de l'ordre du méga paire de bases), tels que de grands plasmides, des fragments de génome de *B. subtilis*, ou encore, des génomes entiers de mycoplasmes.

Les souches de *S. cerevisiae* sont mises en culture à partir d'une pré-culture de 5 mL de milieu riche YPDA ou de milieu minimum SD en fonction de l'auxotrophie à appliquer. Puis les cellules sont diluées dans 100 mL de culture et cultivées sur la nuit à 30°C. Lorsque la DO à 600nm est comprise entre 3,5 et 4, les levures sont centrifugées à 2500g pendant 5 min et remises en suspension dans 20 mL de sorbitol 1M puis conservées à 4°C pendant environ 18h. Le lendemain, jour de la transformation, le sorbitol est éliminé par centrifugation à 2500g pendant 5 min et le culot cellulaire est remis en suspension dans 10 mL de solution SPEM (Sorbitol 1M ; sodium phosphate 10 mM ; Na₂EDTA 10 mM, pH 7,5). A cette étape, les levures subissent un traitement par ajout de 20 μ L de solution de zymolyase (Zymolyase 100-T 10 mg/mL ; Glycérol 50% (v/v)) et 20 μ L de β -mercapto-éthanol et par incubation au bain-marie à 30°C. Un suivi de DO_{600nm} est réalisé au cours du temps afin d'apprécier la production des sphéroplastes. Les cellules sont prélevées régulièrement et diluées au 1/10^{ème} dans du sorbitol 1M ou du SDS 2%. Lorsque le rapport DO_{Sorbitol}/DO_{SDS} est compris entre 3 et 4, la réaction est stoppée par ajout immédiat de 40 mL de Sorbitol 1M et le mélange est centrifugé pendant 5 min à 1500g à 4°C. Les cellules sont lavées précautionneusement avec 50 mL de sorbitol 1M puis sont finalement collectées par centrifugation et reprises dans 1 à 2 mL de solution STC (Sorbitol 1 M ; Tris-HCl 10 mM, pH 7,5 ; CaCl₂ 10 mM). Les levures ainsi traitées sont réservées pendant 10 min à température ambiante et remises en suspension par

Tableau 13 : Composition des solutions de transformation de sphéroplastes levures au PEG : méthodes de clonage par Fusion et CReasPy-Fusion

| SPEM | |
|--|----------|
| Sorbitol (Sigma) | 91 g |
| Na ₂ HPO ₄ , 2H ₂ O | 0,69 g |
| NaH ₂ PO ₄ , 2H ₂ O | 0,18 g |
| EDTA 0,5 M pH = 7,5 | 10 mL |
| Qsp H ₂ O | 500 mL |
| Filtration 0,22µm | |
| STC | |
| Sorbitol (Sigma) | 18,2 g |
| Tris-HCl 1M pH = 7,5 | 1 mL |
| CaCl ₂ 1M | 1 mL |
| MgCl ₂ 1M | 225 µL |
| Qsp H ₂ O | 100 mL |
| Filtration 0,22µm | |
| SOS | |
| Sorbitol (Sigma) | 9,1 g |
| Extrait de levure (Gibco) | 0,125 g |
| Bacto peptone (Difco) | 0,25 g |
| CaCl ₂ 1M | 300µL |
| Qsp H ₂ O | 50 mL |
| Filtration 0,22µm | |
| PEG | |
| PEG 8000 (Thermo Fisher) | 3 g |
| Tris-HCl 1M pH = 7,5 | 150 µL |
| CaCl ₂ 1M | 150 µL |
| MgCl ₂ 1M | 75 µL |
| Qsp H ₂ O | 15 mL |
| Filtration 0,22µm | |
| Resuspension Buffer | |
| Sucrose (Sigma) | 17,115 g |
| Tris-HCl 1M pH = 7,5 | 1 mL |
| CaCl ₂ 1M | 1 mL |
| MgCl ₂ 1M | 225 µL |
| Qsp H ₂ O | 100 mL |
| Filtration 0,22µm | |
| SMM 1X (Sucrose Magnesium Maleate) | |
| Sucrose (Sigma) | 8,55 g |
| Maleate Buffer 0,2M | 5 mL |
| MgCl ₂ pH = 7,5 1M | 1 mL |
| Qsp H ₂ O | 50 mL |
| Filtration 0,22µm | |

agitation douce du tube sans pipetages. Pour une réaction de transformation, 100 μ L de sphéroplastes sont mélangés avec 20 μ L d'ADN génomique préparé en bloc d'agarose (cf. §II.14.1. et §II.14.2.), additionné ou non de 300 ng de pansement de réparation (cf. §II.7). Le mélange est incubé 10 min à température ambiante puis traité pendant 20 min avec 0,5 mL de solution de PEG (PEG 8000, 20% (m/v) ; Tris-HCl 10 mM, pH 7,5 ; CaCl₂ 10 mM). Les sphéroplastes sont collectés par centrifugation pendant 5 min à 5000g et remis en suspension dans 0,7 mL de solution SOS (Sorbitol 1 M, CaCl₂ 6,5 mM, extrait de levure 0,25% (m/v), peptone 0,5% (m/v)) pour être incubés 2h à 30°C, sans agitation. Les cellules sont ensuite mélangées à 12 mL de Top-Agar (milieu sélectif SD contenant 1 M de sorbitol et 3% (m/v) d'agar, conservé à 55°C en surfusion au bain-marie jusqu'à utilisation) et immédiatement étalées sur milieu sélectif SD contenant 1M de sorbitol. Les transformants sont visibles et peuvent être dénombrés après 2 à 5 jours d'incubations à 30°C.

- Protocole appliqué pour les méthodes de clonage par Fusion et de CReasPy-Fusion :

Ce protocole de transformation diffère de celui décrit précédemment et dérive de celui de Karas *et al.*, 2013. Il est basé sur le même principe mais l'apport du matériel génétique (génome ou fragment de génome) est directement effectué par fusion cellulaire. Ce procédé est permis par mise en contact des sphéroplastes de levures et des cellules bactériennes (mycoplasmes sans paroi ou protoplastes de *B. subtilis*) en présence de PEG.

Préparation des cellules bactériennes :

Les mycoplasmes sont cultivés dans 50 mL de milieu adéquat (cf. §I.2.) jusqu'à atteindre un pH environ égal à 7 (soit environ 10⁹ cellules/mL). Du chloramphénicol est ajouté à une concentration finale de 100 μ g/mL et la culture est incubée pendant 1h à 2h30 à 37°C. Les cellules sont ensuite collectées par centrifugation pendant 10 min à 8000g et à 10°C, puis lavées dans 25 mL de tampon de resuspension (Sucrose 0,5 M ; Tris-HCl 10 mM pH = 7,5 ; MgCl₂ 2,25 mM ; CaCl₂ 10 mM). Une seconde centrifugation est effectuée et le culot est finalement repris dans 1 mL de tampon de resuspension. La préparation est incubée 15min à 49°C juste avant mélange avec les sphéroplastes.

Les souches de *B. subtilis* sont cultivées dans 50 mL de milieu adéquat sous agitation à 37°C (cf. §I.3.) jusqu'à atteindre une DO à 600nm comprise entre 0,4 et 1 (milieu de phase

exponentielle). Du chloramphénicol est ajouté à une concentration finale de 180 µg/mL et la culture est incubée pendant 55 min à 37°C sous agitation (190 rpm). Les cellules sont ensuite collectées par centrifugation pendant 5 min à 7000g et à température ambiante, puis lavées dans 5 mL de tampon SMM 1X (Sucrose 0,5 M ; maleate 0,02 M ; MgCl₂ 0,02 M ; pH = 6,5). Du lysozyme est ajouté à une concentration finale de 2 mg/mL et la culture est incubée pendant 1h à 2h à 37°C sous agitation très douce (50 rpm). A cette étape la paroi de peptidoglycane est dégradée afin d'obtenir des protoplastes. A la fin du traitement, les protoplastes sont collectés par centrifugation pendant 15 min à 2500g et à température ambiante. Le culot est lavé par ajout de 5 mL de tampon SMM 1X et une seconde centrifugation est effectuée. Les protoplastes sont finalement repris dans 1 mL de tampon SMM 1X avant mélange avec les sphéroplastes.

Protocole de transformation :

Les souches de *S. cerevisiae* sont cultivées à partir d'une pré-culture de 5 mL de milieu riche YPDA ou de milieu minimum SD en fonction de l'auxotrophie à appliquer. Puis les cellules sont diluées dans 50 mL de culture et cultivées sur la nuit à 30°C. Lorsque la DO à 600nm est comprise entre 2,5 et 3, les levures sont centrifugées à 2500g pendant 5 min à 10°C. Elles sont successivement lavées et centrifugées (2500g, 5min, 10°C) avec 50 mL d'eau stérile puis avec 50 mL de sorbitol 1M. Le culot cellulaire est conservé à 4°C pendant environ 18h. Le lendemain, jour de la transformation, les levures sont remises en suspension dans 20 mL de solution SPEM (Sorbitol 1M ; sodium phosphate 10 mM ; Na₂EDTA 10 mM, pH 7,5). A cette étape les levures subissent un traitement par ajout de 40 µL de solution de zymolyase (Zymolyase 20-T 10 mg/mL ; Glycérol 50% (v/v)) et 30 µL de β-mercaptoéthanol et par incubation au bain-marie à 30°C. Un suivi de DO_{600nm} est réalisé au cours du temps afin d'apprécier la production des sphéroplastes. Les cellules sont prélevées régulièrement et diluées au quart (200 µL de culture dans 800 µL) dans du sorbitol 1M ou dans de l'eau. Lorsque le rapport DO_{Sorbitol}/DO_{H₂O} est compris entre 1,8 et 2, la réaction est stoppée par ajout immédiat de 30 mL de Sorbitol 1M et le mélange est centrifugé pendant 5 min à 1500g à 10°C. Les cellules sont lavées précautionneusement avec 40 mL de sorbitol 1M puis sont finalement collectées par centrifugation et reprises dans 2 mL de solution STC (Sorbitol 1 M ; Tris-HCl 10 mM, pH 7,5 ; CaCl₂ 10 mM). Les levures ainsi traitées sont réservées pendant 10 min à température ambiante et remises en suspension par agitation douce du tube sans pipetages.

Pour une réaction de transformation, 200 μ L de sphéroplastes sont mélangés avec 50 μ L de cellules de mycoplasmes ou de protoplastes de *B. subtilis*, additionné ou non de 1 μ g de pansement de réparation (cf. §II.7). Le mélange est incubé 5 min à température ambiante puis traité pendant 20 min avec 1 mL de solution de PEG (PEG 8000, 20% (m/v) ; Tris-HCl 10 mM, pH 7,5 ; CaCl₂ 10 mM). Les sphéroplastes sont collectés par centrifugation pendant 7 min à 1500 g et remis en suspension dans 0,8 mL de solution SOS (Sorbitol 1 M, CaCl₂ 6,5 mM, extrait de levure 0,25% (m/v), peptone 0,5% (m/v)) pour être incubés 2h à 30°C, sans agitation. Les cellules sont ensuite mélangées à 12 mL de Top-Agar (milieu sélectif SD contenant 1 M de sorbitol et 3% (m/v) d'agar, conservé à 55°C en surfusion au bain-marie jusqu'à utilisation) et immédiatement étalées sur milieu sélectif SD contenant 1M de sorbitol. Dans le cas où les transformants sont sélectionnés sur milieu SD avec une contre sélection 5-FOA, de l'acide 5-fluoroorotique est ajouté à une concentration finale de 1 g/L. Les transformants sont visibles et peuvent être dénombrés après 4-5 jours d'incubation à 30°C.

III.4.3. Analyses des transformants levures

Des colonies individuelles sont piquées puis striées sur milieu sélectif, et sont incubées 2 jours à 30°C. Une colonie isolée par strie est ensuite repiquée sur le même milieu sélectif et incubée 2 jours de plus à 30°C. Les cellules repiquées sont analysées par la suite afin d'identifier les transformants positifs. L'ADNg total est extrait à l'aide des protocoles d'extraction décrits paragraphe II.5.2, puis analysés par PCR simplexe et multiplexe (cf. §II.6.). Le cas échéant, les clones validés au cours de l'étape de vérification par PCR (c'est-à-dire avec le génome bactérien d'intérêt cloné), sont ensuite cultivés pour extraire le génome en bloc d'agarose (cf. §II.14.3.). La taille de la molécule d'intérêt clonée est finalement estimée par analyse par électrophorèse en champs pulsé (cf. §II.16.2. et §II.16.3).

IV. Ingénierie de génomes bactériens

IV.1. Ingénierie du génome de *B. subtilis*

Les mutants de *B. subtilis* utilisés au cours de ces travaux ont été obtenus par transformation de plasmides intégratifs dans le génome. Plus précisément, en utilisant les

protocoles de compétence naturelle décrit paragraphes III.3.1. ou III.3.2., les cellules de *B. subtilis* sont modifiées par intégration d'une cassette donnée dans le génome par recombinaison homologue¹⁰¹. Le plasmide intégratif pMT85-pRS313-pMiniTetM-pSPEC sans lacZ-sbo (cf. §II.1.2.) a la particularité d'avoir des zones de recombinaison nécessaires à l'intégration de la cassette [ARS-CEN-HIS-*aadA*(SpecR)] à la place de l'opéron *sboAX-albAG*. Les mutants *B. subtilis* MSΔ*sbo-alb*::ACH-Spec et MGP192Δ*sbo-alb*::ACH-Spec ont été obtenus et utilisés pour les essais de clonage de génome entier par la méthode classique ou par fusion.

IV.2. Ingénierie des génomes de mycoplasmes par édition de base

Les mutants *M. bovis*, *M. agalactiae* et *M. gallisepticum* utilisés au cours de ces travaux ont été obtenus par édition de base (ou CRISPR Base Editor) en utilisant des outils développés très récemment au laboratoire²⁶³. Brièvement, les cellules de mycoplasmes sont transformées avec 20 µg de plasmide intégratif à l'aide des protocoles de transformation médiés par le PEG décrits paragraphes III.2.1. et III.2.2.. Ces plasmides intégratifs pMT85_SpdCas9_pmcDA1 et pTi4.0_SpCas9_pmcDA1 (cf. §II.1.1.) développés spécifiquement pour cette méthode d'ingénierie sont respectivement utilisés pour modifier *M. bovis* et *M. agalactiae* (pMT85), et *M. gallisepticum* (pTi4.0). Après transformation, les cellules sont remises en suspension dans 1 mL de milieu de culture approprié (cf. §I.2.) et incubées pendant 2h à 37°C. Elles sont finalement étalées sur milieu sélectif solide (gentamycine 100 µg/mL pour *M. bovis* et *M. agalactiae* ; puromycine 10 µg/mL pour *M. gallisepticum*) et incubées à 37°C pendant 3 à 10 jours jusqu'à obtention de colonies isolées. L'expression du système d'édition de base est induite par ajout de 0,5 µg/mL⁻¹ d'anhydrotetracycline (aTC) sur des colonies précédemment isolées et cultivées sur la nuit. La présence de la mutation sur la cible provoquant la formation d'un codon stop (CAA →TAA), est finalement démontrée par amplification PCR et séquençage Sanger.

IV.3. Test d'activité de nucléases de surface

Ce test a été utilisé au cours de ces travaux dans le but d'attester le phénotype des mutants de mycoplasmes produits en amont des expériences de CReasPy-Fusion. L'activité nucléasique des cellules de *M. bovis* WT et *M. bovis*Δ*mnuA* a été évaluée et rapportée par

Tableau 14 : Composition des milieux et tampons utilisés pour le protocole de transplantation

| SOB + | |
|---|---------|
| SOB (Difco) | 28 g |
| Glucose (Sigma) | 10 g |
| Sérum de veau foetal (Eurobio, décomplémenté 1h à 56°C) | 170 mL |
| Rouge de phénol 1% (m/v) | 2 mL |
| Ampicilline 100 mg/mL | 1 mL |
| pH | 7.5 |
| Qsp H ₂ O | 1000 mL |
| Filtration 0,45µm | |
| Tampon de Fusion 2X | |
| Trisma base (Sigma) | 1,21 g |
| MgCl ₂ -6H ₂ O (Sigma) | 2,03 g |
| NaCl (Sigma) | 14,61 g |
| PEG 6000 (Fluka) | 50 g |
| pH | 6.5 |
| Qsp H ₂ O | 500 mL |
| Filtration 0,45µm | |

Ipoutcha *et al.* 2022²⁶³ en suivant un protocole publié par Sharma *et al.* 2015²⁶⁹. La même analyse a été effectuée pour les souche sauvages et les mutants des espèces *M. gallisepticum* et *M. agalactiae*. Les cellules sont cultivées jusqu'à la fin de la phase logarithmique et après centrifugation à 7000g pendant 10 min à 10°C, elles sont remises en suspension dans 500 µL de tampon de réaction à la nucléase (25 mM Tris-HCl, pH 8.8, 10 mM CaCl₂, 10 mM MgCl₂). Les préparations cellulaires sont co-incubées avec de l'ADN plasmidique (utilisé comme contrôle dans les expériences de CReasPy-Fusion) ou avec de l'ADN double brin linéaire (matrice de recombinaison portant des éléments de levure), pendant 5 ou 60 min à 37°C. À chaque fois, 10 µL sont prélevés et la réaction est arrêtée par addition d'EDTA à une concentration finale de 20 mM. Les analyses de chaque échantillon sont effectuées par migration des mélanges sur gel d'électrophorèse à 1% TAE (Tris-Acétate 40 mM ; EDTA 1 mM ; pH 8.0). La dégradation de l'ADN est appréciée après coloration au bromure d'éthidium (concentration finale 2 µg/mL) et révélation aux UV.

V. Transplantation de génomes bactériens

Afin de valider la méthode de clonage CReasPy-Fusion développée au cours de ces travaux de thèse, les génomes de *Mcap*, *Mmc* et *Mccp* au préalable clonés et modifiés dans la levure, ont été réintroduits dans une cellule receveuse de mycoplasme par transplantation. Ce protocole a été développé et décrit par Lartigue *et al.*, 2007 et 2009^{216,217}.

V.1. Transplantation de génomes de *Mcap* et *Mmc* après CReasPy-Fusion

Les cellules receveuses *Mcap*ΔRE sont cultivées en milieu SOB+ (**Tableau 14**) à 30°C sur la nuit. Lorsque le pH est environ égal à 6,5, 12 mL de cellules sont centrifugés à 5800g pendant 15 min à 10°C, puis lavés avec 6 mL de tampon Wash Buffer (10 mM Tris-HCl, 250 mM NaCl, pH = 6,5). Après une seconde centrifugation (5800g, 15 min, 10°C), les bactéries sont remises en suspension dans 400 µL de CaCl₂ 0,1 M puis laissées dans la glace pendant 30 min. 100 µL de préparation de génomes donneurs issus de blocs d'agarose traités à la β-agarase (*cf.* §II.15.2) sont ensuite délicatement déposés au contact des 1200 µL de suspension cellulaire (400 µL de cellules receveuses + 800 µL de milieu SP5 sans sérum) en présence 1300 µL d'un tampon de Fusion 2X. Après 90 min à 30°C, la réaction est stoppée et les cellules

sont diluées dans 5 mL de milieu SP5. Elles sont ensuite collectées par centrifugation (5800g, 15 min, 10°C) et remises en suspension dans 500 µL de milieu SP5. Finalement, les cellules sont étalées sur milieu gélosé SP5 supplémenté de tétracycline 5 µg/mL et incubées à 37°C en présence de CO₂ 5%, pendant 3 à 5 jours.

V.2. Transplantation de génomes de *Mccp* après CReasPy-Fusion

D'après Gourgues *et al.*, 2022 (en cours d'écriture) le protocole de transplantation effectué avec des génomes de *Mccp* au préalable clonés et modifiés dans la levure dans une cellule receveuse de *Mcap*ΔRE suit le même déroulé que celui décrit §V.1. Cependant il diffère après l'étape de mise en contact des cellules receveuses et des génomes donneurs en présence de tampon de fusion. Après 90 min à 30°C, la réaction est stoppée et les cellules sont diluées dans 5 mL de milieu Hayflick modifié. Elles sont ensuite collectées par centrifugation (5800g, 15 min, 10°C) et remises en suspension dans 1 mL de milieu Hayflick modifié. Finalement, les cellules sont étalées sur milieu gélosé Hayflick modifié supplémenté de gentamycine 300 µg/mL et incubées à 37°C en présence de CO₂ 5%, pendant 3 à 5 jours.

V.3. Analyse des transplants

Après transplantation et incubation pendant plusieurs jours à 37°C, les colonies sont dénombrées puis repiquées dans 1mL de milieu adéquat. Trois passages successifs avant analyse sont effectués (un passage correspondant à une dilution de la culture au 1/100^{ième} suivi d'une incubation de 24 à 48h à 37°C). L'ADNg total des transplants est extrait grâce au protocole « Quick & Dirty » décrit paragraphe II.5.3., et les clones sont analysés par PCR simplexe et multiplexe (*cf.* §II.6.) afin de vérifier que ces derniers possèdent bien le génome nouvellement transplanté. Après validation par PCR, l'ADNg est à nouveau extrait avec le protocole Qiagen® Genomic-tips 100/G (*cf.* II.5.3.) approprié à l'analyse finale des transplants par séquençage NGS.

CHAPITRE 1 – Essais de clonage d'un génomme entier réduit de *B. subtilis* dans la levure

CHAPITRE 1 – Essais de clonage d'un génome entier réduit de *B. subtilis* dans la levure

Introduction

Le projet ANR *Bsu2.0* propose de mettre en place un ensemble de méthodes de biologie de synthèse (BS) efficaces qui permettraient d'assembler et modifier à façon des génomes semi-synthétiques de *B. subtilis* dans la levure *S. cerevisiae*. Le pipeline mis en place rassemblerait, dans un ordre défini, les méthodologies suivantes : (i) conception *in silico* de nouveaux génomes, (ii) synthèse et assemblage des génomes conçus dans la levure, (iii) modification (optionnelle) des génomes dans la levure, (iv) et transplantation de génomes nouvellement conçus, depuis la levure vers une cellule receveuse.

Adapter l'ensemble de ces technologies (déjà disponibles chez les mycoplasmes) à *B. subtilis*, requiert, au-delà de nombreuses mises au point, d'attester de la faisabilité de certaines étapes. Par exemple, un prérequis à la synthèse et à l'assemblage de génomes semi-synthétiques de *B. subtilis* dans la levure dont les tailles pourraient atteindre ~2 Mb (coût estimé à environ 10 000 euros) est de démontrer que le génome de *B. subtilis* natif (ou du moins, une version réduite de ce génome) agrémenté d'une origine de réplication levure (ARS), d'un centromère (CEN) et d'un marqueur d'auxotrophie peut être maintenu de façon stable chez cet hôte et ceci pendant plusieurs passages.

La première étape de notre travail a donc consisté à mettre en place une série de protocoles visant à cloner le génome de *B. subtilis* dans la levure *S. cerevisiae*. La question initiale qui s'est alors posée concernait le choix de la souche à utiliser. En effet, comme mentionné dans l'introduction (*cf.* §III), au cours des 15 dernières années, de nombreux génomes bactériens (mollicutes, protéobactéries, cyanobactéries) ont été clonés dans la levure. Certains de ces génomes ont été clonés sans encombre ; c'est le cas de la plupart des génomes de mollicutes (18 génomes clonés au total) qui possèdent des petites tailles (0,58 à 1,8 Mb), des taux en G+C faibles (24 et 40%) et utilisent un code génétique alternatif²¹⁹. C'est également le cas, des génomes naturels de la protéobactérie *Haemophilus influenzae* (1,8 Mb et 38% G+C)²²³ et de la cyanobactérie *Prochlorococcus marinus* (1,6 Mb et 31% G+C)²²⁴ qui ont également été clonés avec succès chez *S. cerevisiae*.

Pour d'autres génomes (tailles plus grandes, taux en G+C >40% et/ou code génétique standard), quelques ajustements ont, en revanche, été nécessaires pour stabiliser leur maintien sous forme de YCP au sein de la levure. C'est le cas du génome d'*Acholeplasma laidlawii* (1,5Mb, 32% G+C), mollicutes utilisant un code génétique standard, et dont le génome n'a pu être cloné qu'après suppression d'un gène codant une endonucléase toxique pour la levure²²¹. C'est aussi le cas, des versions synthétiques et réduites des génomes des protéobactéries *Caulobacter crescentus* (0,79 Mb et 57% G+C)²⁷⁰ et *E. coli* (1,03 Mb et 51% G+C)²²⁵, pour lesquelles il a été nécessaire d'introduire une dizaine de séquences ARS, afin de les maintenir de façon stable dans la levure. La tentative de clonage du génome naturel de *Synechococcus elongatus* dont la taille et le GC% sont tous deux élevés (2,7 Mb et 55% G+C) a, quant à elle, échoué²⁷¹.

Au démarrage du projet, aucun génome de taille supérieure à 1,8 Mb et utilisant un code génétique universel n'avait donc été cloné dans la levure. Aussi, il nous est apparu essentiel d'amorcer le travail avec une souche possédant une version réduite du génome plutôt qu'avec la souche de référence *B. subtilis* 168 (4,2 Mb) ou la souche MS (3,98 Mb) plus couramment utilisées par nos collaborateurs. C'est ainsi que nous avons sélectionné la souche *B. subtilis* MPG192 (2,86 Mb) qui possède un des génomes les plus réduits de la collection « Minimal Genome Project » de l'équipe MICALIS¹⁸⁴. Malgré des pertes successives de matériel génétique (suppression d'1,34 Mb), le pourcentage en G+C du génome MGP192 reste relativement inchangé (~44%) et, est proche du seuil critique des 40%. De même, la proportion des ARS putatives dans le génome de *B. subtilis* MPG192 (recherchées via le motif ACS pour ARS Consensus Sequence, 5'-WTTTAYRTTWTW-3')^{272,273} reste remarquablement proche de celle de *B. subtilis* 168 (144 contre 221) ; et leur distribution est relativement homogène le long du génome MGP192.

De nombreuses méthodes ont été décrites dans la littérature pour cloner des génomes bactériens dans la levure. Toutes requièrent, l'ajout de séquences nécessaires au maintien du génome bactérien chez *S. cerevisiae*, à savoir, un centromère, un marqueur de sélection et pour les génomes à fort taux en G+C (> 40%), une ou des origine(s) de réplication de levure. Le choix de la méthode (présentées §III de l'introduction) dépend essentiellement de l'organisme avec lequel on souhaite travailler et de ses caractéristiques (transformable ou

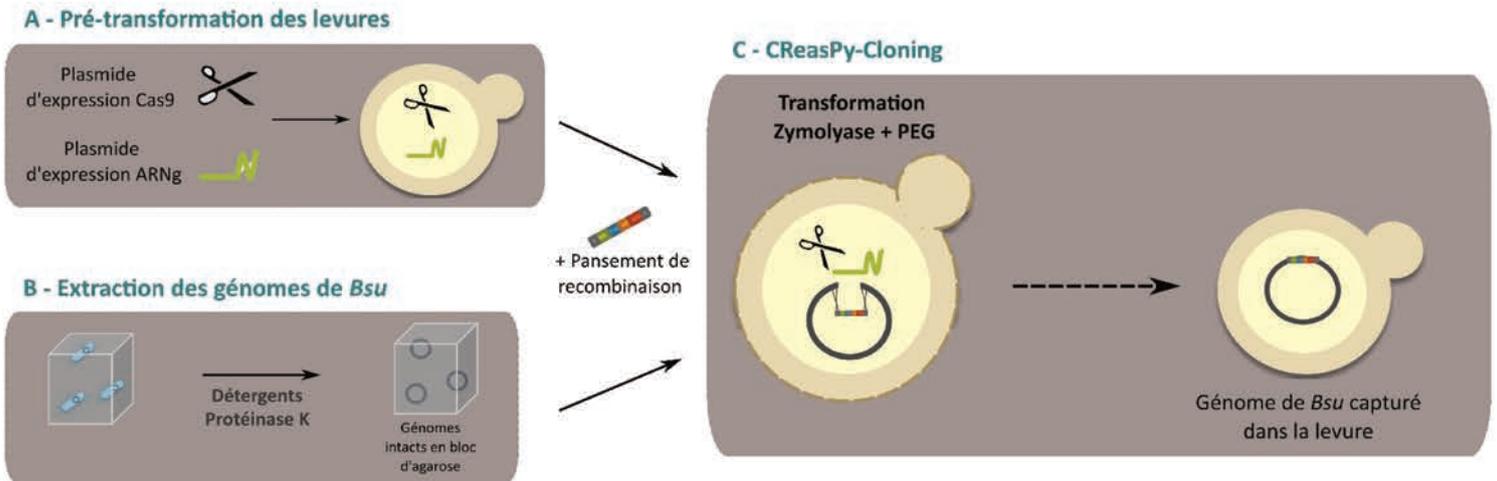


Figure 32 : Clonage du génome de *B. subtilis* dans *S. cerevisiae* par CRasPy-Cloning. [A] Les levures sont transformées avec les plasmides pCas9 (p414-TEF1p-CAS9-CYC1t) et pARNg (p426-SNR52p-gRNA+AarI-SUP4t) permettant respectivement l'expression de la protéine Cas9 et d'un ARNg spécifique du *locus* cible. [B] Le génome de *B. subtilis* est extrait et purifié en bloc d'agarose. [C] La levure est finalement co-transformée avec le génome nu et intact, et une matrice d'ADN linéaire aussi appelée « pansement de recombinaison ». Cette dernière est composée d'une cassette « éléments levure » et d'un marqueur de sélection, et possède à ses extrémités deux régions identiques au *locus* ciblé par l'ARNg et à réparer. Le complexe Cas9-ARNg présent dans la levure détecte le site de clivage dans le génome bactérien, et génère une coupure double brin qui linéarise le chromosome transformé. Cette coupure stimule un des systèmes de réparation de l'ADN présents chez la levure. Plus précisément, la présence du « pansement » enclenche le mécanisme de recombinaison homologue « HR ». Après réparation, le chromosome bactérien est de nouveau circulaire, et porte à présent la cassette « éléments levure » insérée au *locus* précis défini par l'expérimentateur.

non, outils génétiques disponibles, présence de sites de restriction (nature et nombre). Pour la souche *B. subtilis* MGP192, nous avons choisi de tester trois méthodes :

- (i) La méthode de CReasPy-Cloning développée au laboratoire et qui permet le clonage et la modification simultanée du génome d'intérêt. L'avantage de cette méthode est qu'il n'est pas nécessaire de pré-transformer la bactérie pour y ajouter les éléments levures. Cet avantage semble important, car il s'avère que la souche MGP192 est moins compétente et facile à transformer que la souche naturelle.
- (ii) La méthode dite « Classique » pour laquelle le génome pré-marqué est extrait en bloc d'agarose puis directement transformé dans la levure avec un protocole au PEG. L'avantage de cette méthode réside dans sa simplicité et sa robustesse puisqu'elle a permis le clonage de nombreux génomes.
- (iii) La méthode de clonage par Fusion qui permet le transfert d'un génome pré-marqué, directement par contact de cellules à cellules entre des bactéries et des sphéroplastes de levure. L'avantage de cette méthode est qu'il n'est pas nécessaire d'extraire au préalable les chromosomes bactériens en bloc d'agarose.

Au cours de ce chapitre, les résultats obtenus pour les essais de clonage du génome entier de *B. subtilis* MGP192 seront présentés, en suivant séquentiellement la mise en place de ces méthodes.

I. Essais de clonage par la méthode de CReasPy-Cloning

Les premiers essais de clonage du génome de *B. subtilis* dans la levure ont été effectués en utilisant la méthode de CReasPy-Cloning (**Figure 32**). Cette méthode a été développée par Ruiz *et al.* en 2019, et a permis de cloner et d'éditer plusieurs génomes de mycoplasmes avec une grande efficacité²²⁰. La mise en place de cette méthode passe par la préparation (i) de levures pré-chargées avec un plasmide pCas9 et avec un plasmide pARNg, (ii) d'une matrice de recombinaison (« pansement de réparation ») composée des éléments levures et spécifique du *locus* cible à réparer, et (iii) du génome bactérien purifié en bloc d'agarose. Ce n'est que lorsque tout ce matériel biologique est prêt que l'expérience de CReasPy-Cloning, à proprement parler, peut être réalisée.

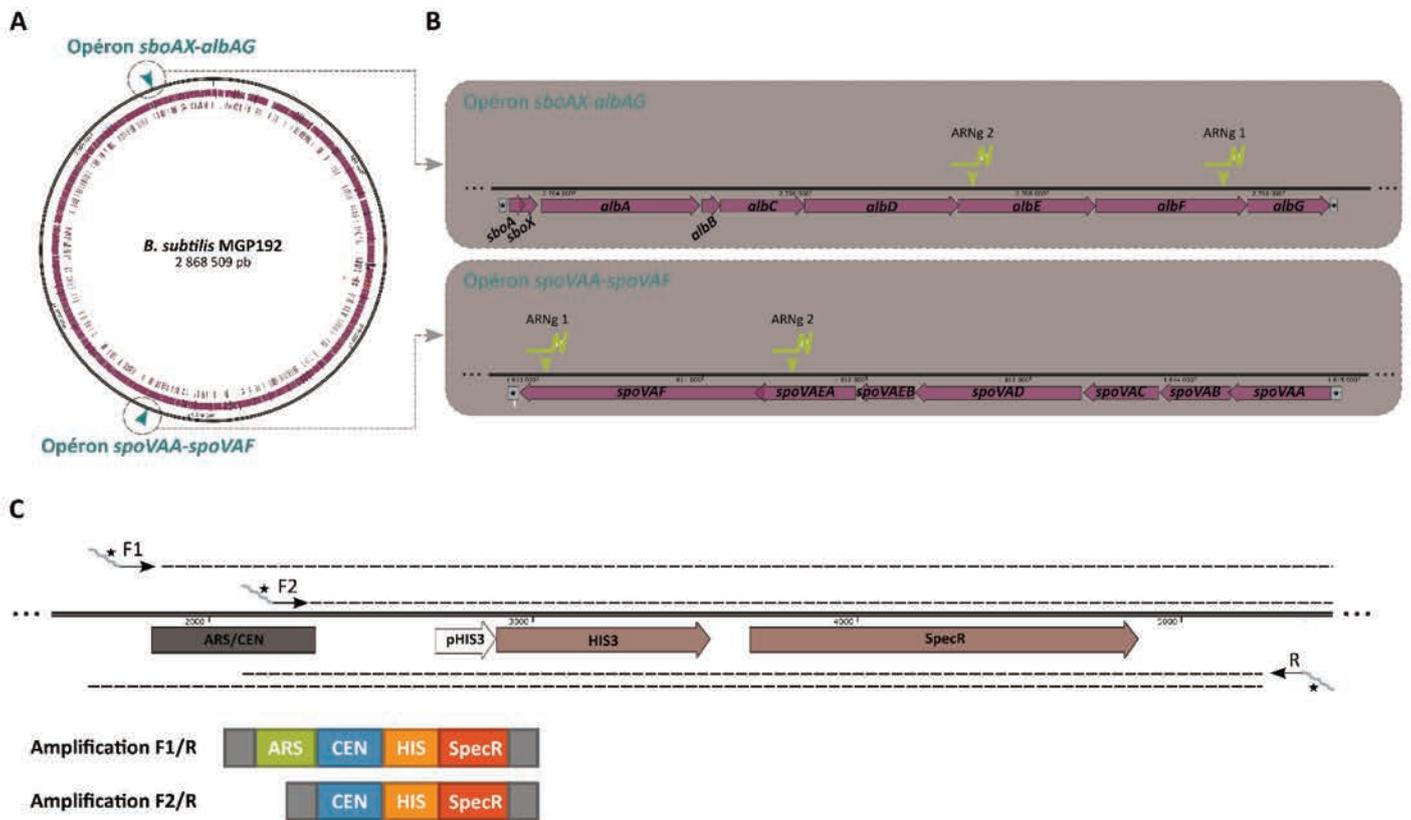


Figure 33 : Représentation schématique des opérons ciblés par le système CRISPR et les matrices de réparation associées. [A] Carte du génome à 2,86 Mb de *B. subtilis* souche MGP192. La localisation génomique des opérons ciblés choisis est indiquée à l'aide d'une flèche bleue. **[B]** Un zoom sur chacun des opérons est représenté dans les encadrés en gris. L'ordre séquentiel, la taille et le nom de chaque gène est indiqué (en rose). Les régions bordant les opérons sont indiquées par un rectangle gris et une étoile. Elles correspondent aux séquences choisies pour les extrémités des pansements de recombinaison. La localisation des ARNg dessinés est indiquée par une flèche verte. **[C]** L'amplification des pansements ARS-CEN-HIS-SpecR ou CEN-HIS-SpecR est représentée par les lignes en pointillées avec les amorces sens (« F1 » et « F2 ») ou antisens (« R ») associées. Les queues flottantes des amorces sont représentées en gris avec une étoile noire et sont identiques aux séquences des extrémités des opérons cibles. La représentation schématique des pansements après amplification est également indiquée.

Tableau 15 : Transformation des levures VL6-48N-SpCas9 avec les plasmides pARNg

| Levures | pARNg | UFC | Criblage des transformants | | |
|--------------------|---------------|-------|----------------------------|--------------|------------|
| | | | PCR 1 (ARNg) | PCR 2 (Cas9) | Séquençage |
| | - | 0 | - | - | - |
| VL6-48N + pCas9 | ARNg1_sbo-alb | >1000 | 5/5 | 2/2 | Ok |
| | ARNg2_sbo-alb | >1000 | 5/5 | 2/2 | Ok |
| | ARNg1_spoVAA | >1000 | 5/5 | 2/2 | Ok |
| | ARNg2_spoVAA | >1000 | 5/5 | 2/2 | Ok |

I.1. Pré-chargement des levures avec le système CRISPR-Cas9

Ruiz *et al.*, ont démontré que pour être efficace, la méthode de CReasPy-Cloning devait être réalisée en deux temps : la levure doit d'abord être transformée avec les plasmides pCas9 et pARNg ; puis être co-transformée avec le génome bactérien et le pansement de recombinaison²²⁰. La première étape de mon travail consistait donc à produire des souches levures hébergeant ces deux types de plasmides.

Au cours des travaux réalisés au laboratoire, les auteurs avaient généré plusieurs stocks de souches de levure (VL6-48N et W303a) déjà transformées avec le pCas9 (p414-TEF1p-Cas9-CYC1t) portant le marqueur d'auxotrophie TRP1. Nous avons utilisé ce matériel disponible, comme point de départ et avons introduit les pARNg *ad hoc*, en utilisant le protocole de transformation à l'acétate de lithium (*cf.* §III.4.1. Matériels et Méthodes). Ces expérimentations ont été menées seulement avec la souche levure VL6-48N car de très bons résultats venaient d'être obtenus au laboratoire en utilisant cette souche. Quatre plasmides pARNg (basés sur le p426-SNR52p-gRNA-Aarl-SUP4t portant le marqueur d'auxotrophie URA3) ont été conçus puis validés par séquençage. Chaque plasmide porte une séquence codant un ARNg spécifique capable de cibler un *locus* précis dans le génome de *B. subtilis*. Le choix des gènes cibles s'est porté sur deux opérons non essentiels (**Figure 33A-B**) :

- l'opéron *sboAX-albAG* (localisation 2 703 458 à 2 710 395 pb ; taille 6938 pb) dont les gènes sont impliqués dans la production de la subtilisine A, qui est un peptide antimicrobien²⁷⁴.
- l'opéron *spoVAA-spoVAF* (localisation 1 609 893 à 1 614 825 pb ; taille 4933 pb) dont les gènes codent des protéines impliquées dans la maturation des spores de *B. subtilis*²⁷⁵.

Dans le cas du premier opéron, nous avons ciblé plus précisément les gènes *albF* (ARNg 1) et *albE* (ARNg 2). Pour le second opéron, nous avons ciblé les gènes *spoVAF* (ARNg 1) et *spoVAEA* (ARNg 2). Au cours de l'expérience représentée **Figure 32A** et résumée dans le **Tableau 15**, la souche de levure VL6-48N-Cas9, a été transformée avec les différents plasmides pARNg puis étalées sur milieu sélectif SD-TRP-URA.

Après 48 h d'incubation à 30°C, aucune colonie n'est apparue sur les boîtes « témoin négatif » (transformation sans ADN), alors que des milliers de colonies sont apparues sur les boîtes « essais ». L'ADNg total des transformants a été extrait pour cinq clones par condition, puis des PCR ont été effectuées afin de vérifier : (i) la présence du plasmide permettant la

production des ARNg 1 ou 2 pour chacun des opérons et (ii) de confirmer la présence du pCas9. Les résultats du tableau montrent que cette étape de pré-chargement fonctionne très bien puisque la totalité des clones vérifiés à l'aide des amorces gRNA_verif_F/R ont été validés par la PCR 1 « ARNg ». Le plasmide a donc été transformé avec succès. De plus, la présence du pCas9 a été confirmée grâce aux amorces T3-T7_F/R par la PCR 2 « Cas9 ». Enfin, pour chaque essai, les amplicons « ARNg » de deux clones ont été purifiés puis vérifiés par séquençage Sanger afin de nous assurer de l'absence de mutation délétère au bon fonctionnement de la méthode. Les clones ainsi validés ont été stockés au congélateur à -80°C, en attente de l'étape de co-transformation avec le génome de *B. subtilis* MGP192 et du pansement de réparation.

1.2. Préparation de la matrice de recombinaison

Les pansements de recombinaison, éléments nécessaires à la réparation du génome après coupure par le duplex Cas9-gRNA, ont été produits par PCR, purifiés puis dosés. Ils sont généralement constitués d'une origine de répllication levure (ARS), d'un centromère (CEN), d'un marqueur d'auxotrophie pour l'histidine (HIS3), et d'un gène conférant la résistance à un antibiotique que nous appellerons SpecR (gène *aadA* conférant une résistance à la spectinomycine chez *B. subtilis*). Ce dernier élément est rajouté afin d'anticiper l'étape de transplantation, et de permettre la sélection des transplants avec un génome modifié. Pour nos expériences, nous avons finalement produit deux types de pansements : ARS-CEN-HIS-SpecR (amorces F1/R), et CEN-HIS-SpecR (amorces F2/R) (**Figure 33C**). Les queues flottantes des amorces correspondaient aux séquences encadrant les opérons cibles *sboAX-albAG* ou bien *spoVAA-VAf*.

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

Un prérequis au succès du clonage de grands fragments de génome, ou de génomes entiers, est l'apport de molécules d'ADN purifiées et non endommagées. Cela passe par l'extraction des chromosomes bactériens à cloner en bloc d'agarose. Cette méthode favorise la purification des génomes tout en diminuant le nombre de cassures mécaniques possibles (généralement dues aux pipetages répétés). La qualité de ces préparations est un point crucial pour la réussite des méthodes de clonage de génomes dans la levure. Elle inclut la quantité

Tableau 16 : Paramètres testés au cours des préparations du génome de *B. subtilis* en bloc d'agarose

| Paramètres | Série 1 : Première préparation | Série 2 : Deuxième préparation | Série 3 : Troisième préparation |
|---------------------------|---|---|--|
| Protocole | Protocole "ED" | Protocole "ED" | Protocole Biorad |
| Phase de croissance | DO = 0,8 ou 1,6 | DO = 1,6 | DO = 1,6 |
| Volumes de culture (mL) | 1,5 ; 7,5 ; 15 ; 30 | 2 ; 5 ; 10 ; 20 | 2 ; 5 ; 10 ; 20 |
| Concentration en lysozyme | 5mg/mL | 5mg/mL | 1mg/mL |
| Agent compactant | - | +/- Hexamine cobalt(III) chloride | +/- Hexamine cobalt(III) chloride |

d'ADN présente dans les blocs, la pureté des molécules, l'état de l'ADN (molécules intactes non fragmentées et entièrement répliquées) et l'état de condensation de l'ADN.

Au début de ces travaux, nous n'avions pas au laboratoire de protocole disponible pour la préparation de blocs à partir de bactéries dont la paroi est Gram+ et qui nécessitent un traitement enzymatique par le lysozyme. Il a donc été important de veiller à la mise au point d'une procédure efficace et adaptée à l'extraction des génomes de *B. subtilis*.

Deux protocoles ont alors été comparés :

- Le protocole « ED » qui est un protocole transmis par nos collaborateurs de l'équipe de Jouy-en-Josas. Il nous a paru primordial de tester un protocole élaboré spécifiquement pour *B. subtilis*, notamment pour la culture des cellules ou pour les conditions de lyse de la paroi Gram+.

- Le protocole « Biorad » issu du protocole associé au kit commercial CHEF Bacterial Genomic DNA Plug du fournisseur Bio-Rad qui offre une approche plus standardisée des préparations, qu'il était utile de tester pour garder une répétabilité dans le temps. C'est aussi un protocole très proche de celui couramment utilisé au laboratoire pour la purification des génomes de mycoplasmes²¹⁶ avant ou après clonage dans la levure (kit CHEF Mammalian Genomic DNA Plug). Il diverge seulement par la présence d'une étape de lyse spécifique des parois bactériennes.

Au cours du travail de mise au point, nous avons testé la phase de culture et la quantité de cellules prélevée, la quantité de lysozyme, et enfin, l'utilisation d'un agent chimique pour condenser l'ADN (**Tableau 16**).

I.3.1. Première préparation avec le protocole « ED »

Lartigue *et al.* 2009, ont montré que l'état physiologique des bactéries (lié à la phase de croissance) ainsi que la quantité de cellules à inclure dans un bloc (qui sera directement liée à la quantité de génomes présents) sont des paramètres déterminants pour l'efficacité du clonage de génomes de mycoplasmes dans la levure. En nous basant sur cette expérience, la série n°1 de blocs a été préparée à partir de cultures de *B. subtilis* dont la DO_{600nm} correspondait soit au début ($DO_{600nm} = 0,8$) ou au milieu ($DO_{600nm} = 1,6$) de la phase

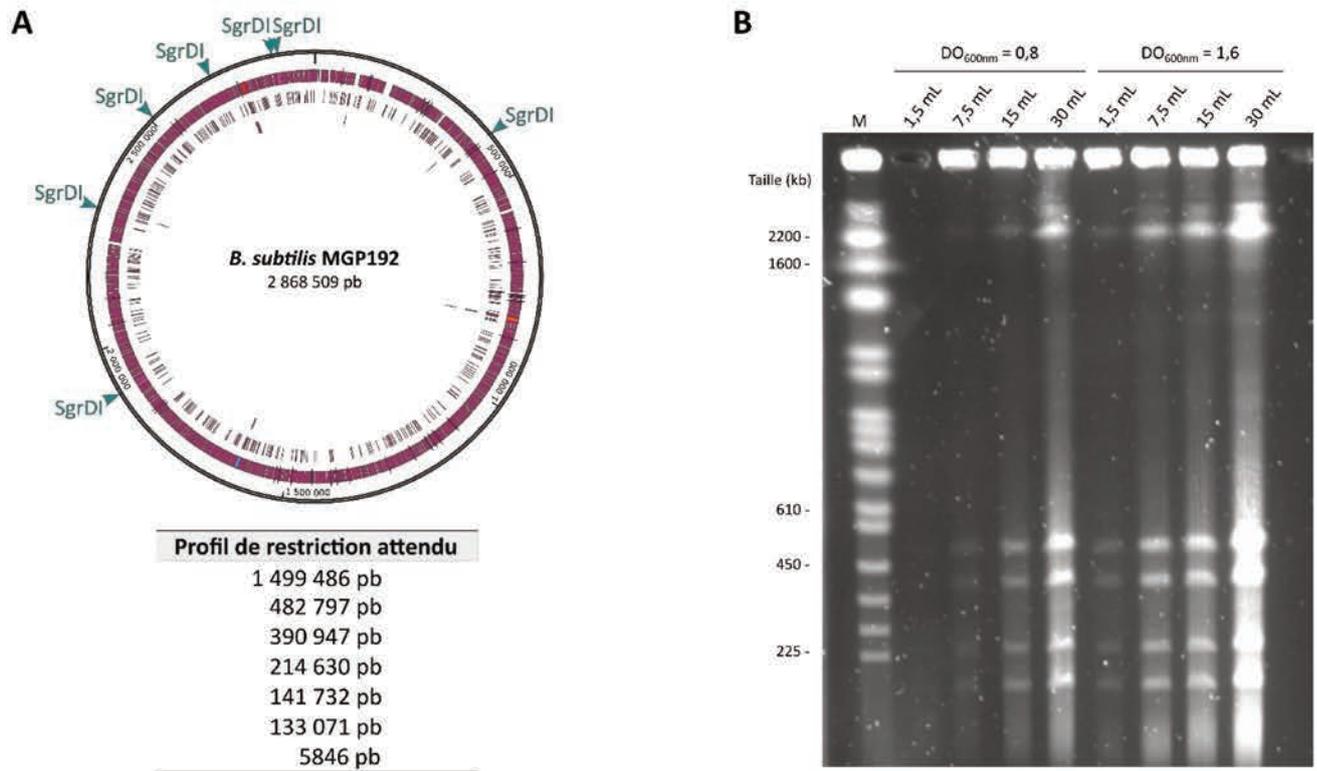


Figure 34 : Restriction du génome de *B. subtilis* MGP192 et analyse par PFGE. [A] Carte du génome de *B. subtilis* souche MGP192 à 2,86 Mb. La localisation génomique des sites de coupure de l'enzyme SgrDI est indiquée en bleu. Le profil de restriction attendu avec la taille des fragments produits correspondant est également répertorié. **[B]** Les chromosomes préparés en blocs d'agarose issus de la série 1 ont été digérés avec l'enzyme de restriction SgrDI puis analysés en PFGE. Le profil attendu est composé de sept fragments dont les tailles varient de 1 499 486 pb à 5846 pb. "M": *S. cerevisiae* chromosomal DNA ladder Bio-Rad (225-2200 kb).

exponentielle de croissance. Pour chacune d'entre elles, plusieurs volumes ont été prélevés (1,5 mL, 7,5 mL, 15 mL et 30 mL) afin de plus ou moins concentrer les blocs en cellules (et donc en ADNg).

En appliquant le protocole « ED », les blocs produits ont été analysés par restriction avec l'enzyme SgrDI puis migration par PFGE (**Figure 34A-B**). La présence de 7 sites de restriction sur l'ensemble du génome de la souche MGP192 devait permettre l'hydrolyse de la molécule en 7 morceaux linéaires (allant de 5846 pb à 1 499 486 pb). Le résultat obtenu (**Figure 34B**) montre un profil avec les bandes aux tailles attendues, mais avec seulement 5 bandes visibles sur les 7 théoriques. Les conditions de migration choisies expliquaient ce résultat : la bande à 5846 pb de petite taille, a vraisemblablement été éliminée du gel, tandis que les bandes à 133 071 et 141 732 pb (proche en taille) ne pouvaient être distinguées individuellement.

L'intensité des bandes révélait une augmentation des quantités d'ADN en fonction de la phase de croissance mais aussi des volumes de culture initiaux. La condition [DO = 0,8 et volume 1,5 mL] semblait toutefois très peu concentrée et pouvait ne pas suffire pour apporter la quantité de génomes nécessaire au clonage (**Figure 34B**). De plus, la qualité des blocs était comparable pour toutes les conditions mais une augmentation du bruit de fond en corrélation avec l'augmentation des volumes témoignait potentiellement de la présence de génomes fragmentés de façon aléatoire. La condition [DO = 1,6 et volume 30 mL] pouvait donc également être de mauvaise qualité pour des essais de clonage (**Figure 34B**). Ce résultat étant plutôt encourageant, nous avons alors décidé de conserver ces préparations pour un futur essai de clonage par CReasPy-Cloning. Cependant, dans le but de mettre toutes les chances de notre côté pour tenter de cloner le génome de *B. subtilis*, nous avons décidé de tester d'autres paramètres pour optimiser le protocole.

1.3.2. Deuxième préparation et optimisation du protocole « ED »

Le protocole « ED » très spécifique à *B. subtilis* est un protocole originellement utilisé pour l'analyse des génomes par PFGE (typage moléculaire). Il n'était cependant pas forcément le plus approprié pour fournir des génomes purifiés adéquats au clonage dans la levure. Notre expertise pour la préparation de tels blocs est basée sur un protocole mis au point pour le

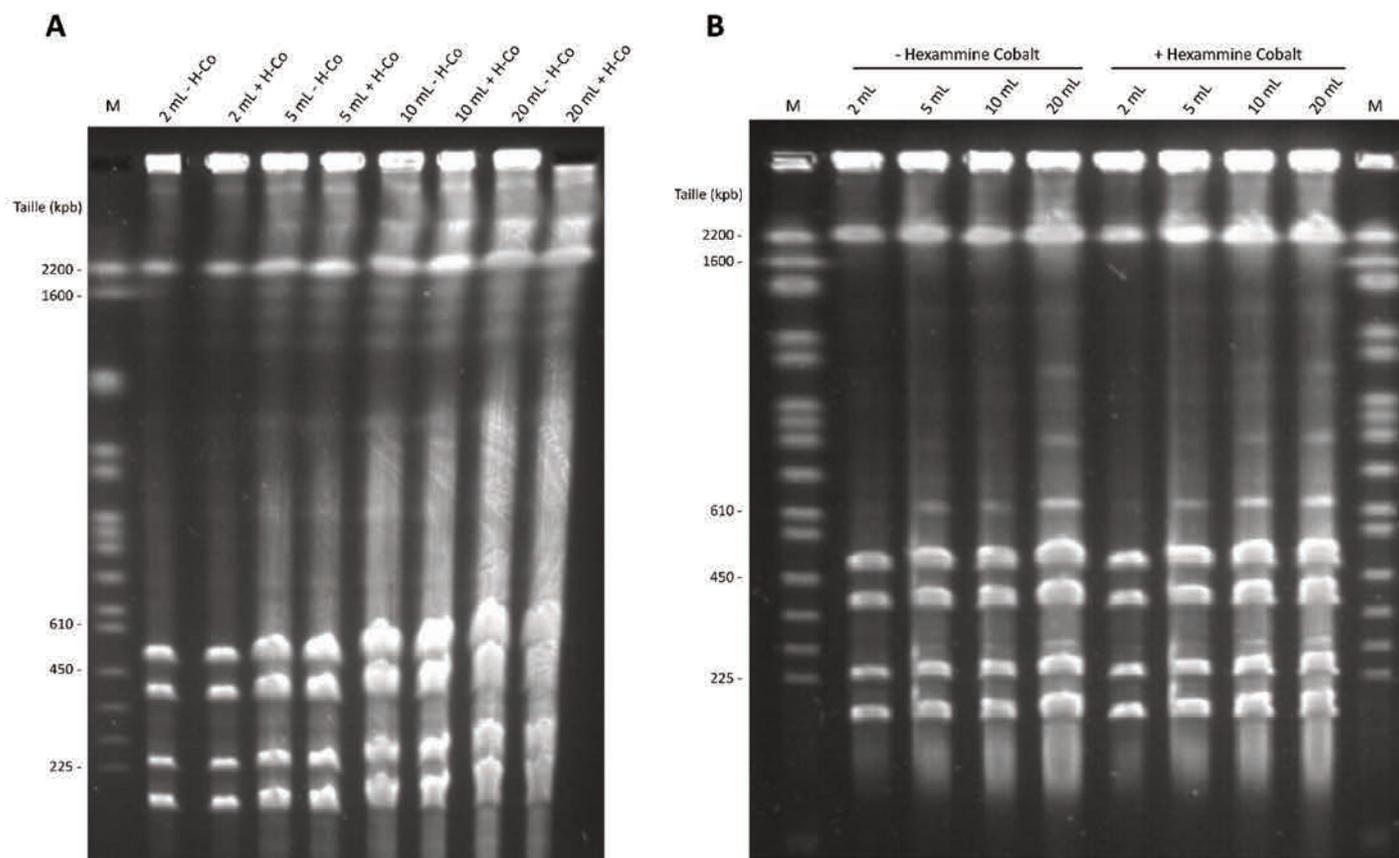


Figure 35 : Analyse par PFGE des préparations des génomes de *B. subtilis* MGP192 en blocs d'agarose. Les génomes de *Bsu* préparés en blocs d'agarose issus **[A]** de la série 2 ou **[B]** de la série 3, ont été digérés avec l'enzyme de restriction SgrDI puis analysés en PFGE. Le profil attendu est composé de sept fragments dont les tailles varient de 1 499 486 pb à 5846 pb. "M": *S. cerevisiae* chromosomal DNA ladder Bio-Rad (225-2200 kb). "H-Co" : Hexamine cobalt(III) chloride.

clonage de génomes de mollicutes dont les tailles sont plus petites (0,56 Mb à 1,8 Mb). Il était donc nécessaire de tester de nouveaux paramètres qui pourraient être plus adaptés au génome de *B. subtilis*. Par exemple, la condensation de l'ADN à l'intérieur des blocs pouvait être un facteur important voire limitant pour le clonage de génomes beaucoup plus grands comme celui de la souche MGP192. Sous forme décondensée (induite par la déprotéinisation par le traitement à la protéinase K), il était possible que les chromosomes ne puissent pas physiquement entrer dans les cellules de levure. Chez les mycoplasmes, il a été montré par exemple que traiter l'ADN avec de l'hexamine cobalt (III) chlorure (agent chimique favorisant la compaction de l'ADN²⁷⁶) aidait l'entrée et l'installation des chromosomes dans des cellules receveuses lors des expériences de transplantation de génomes²¹⁶. Aussi, une série n°2 de blocs a été préparée avec quelques modifications apportées au protocole initial :

- De nouveaux volumes ont été testés ajustant les écarts précédemment choisis : soit 2 mL, 5 mL, 10 mL et 20 mL. La DO à 600nm a quant à elle été fixée à 1,6.
- Lorsque les cultures de *B. subtilis* avaient atteint la $DO_{600nm} = 1,6$, une incubation additionnelle de 55 minutes avant collecte a été effectuée en présence de chloramphénicol à $180 \mu\text{g}\cdot\text{mL}^{-1}$. Le chloramphénicol est un antibiotique qui inhibe la traduction en se fixant sur la sous-unité 50S du ribosome. Dans notre contexte, l'ajout de chloramphénicol avait pour but d'obtenir des génomes à la fois circulaires doubles brins (avec peu ou pas d'intermédiaires de réplication) et condensés. En effet, des travaux ont montré d'une part qu'en présence de chloramphénicol, la synthèse protéique est stoppée et aucun nouveau cycle de réplication ne peut être initié²⁷⁷, et d'autre part que certains antibiotiques (tels que le chloramphénicol) peuvent avoir un effet « compactant » sur l'ADN²⁷⁸.
- Après le traitement à la protéinase K et pendant toutes les étapes de lavages, les blocs ont été incubés en présence ou absence d'hexamine cobalt (III) chlorure à 0,1 mM final²⁷⁹.

À nouveau, la qualité des blocs a été évaluée par restriction avec l'enzyme SgrDI puis migration par PFGE (**Figure 35A**). Les profils de digestions étaient très similaires à ceux obtenus lors de la série n°1. Si aucune différence significative n'était observée entre les blocs traités ou non avec l'hexamine cobalt (III) chlorure, les bandes obtenues avaient une intensité légèrement accrue lorsque le volume de culture prélevé était plus grand. Les différents traitements ne semblaient pas avoir d'effets indésirables sur les préparations de blocs puisque l'hydrolyse du

Tableau 17 : Essais de clonage du génome complet de *B. subtilis* par CReasPy-Cloning

| Expériences | Cibles | ARNg | Matrices de recombinaison | Série de blocs | Concentrations Blocs | Criblage des transformants | | | |
|--------------|--------------------------|--------|---------------------------|----------------|----------------------|----------------------------|-----------------|--|--|
| | | | | | | UFC | Analyse par PCR | | |
| Essai 1 | Opéron <i>sbo-alb</i> | ARNg 1 | - | Série 1 | - | 0 | - | | |
| | | | ACH-Spec | | - | 13 | - | | |
| | | | | | 7,5 mL (20µL) | 5 | 0/5 | | |
| | | | | | 15 mL (20µL) | 12 | 0/12 | | |
| | | | | | 30 mL (20µL) | 8 | 0/8 | | |
| | | ARNg 2 | - | | 0 | - | | | |
| | | | CH-Spec | | - | 14 | - | | |
| | | | | | 7,5 mL (20µL) | 7 | 0/7 | | |
| | | | | | 15 mL (20µL) | 6 | 0/6 | | |
| | | | | | 2 | 0/2 | | | |
| Total | | | | | | 40 | 0/40 | | |
| Essai 2 | Opéron <i>spoVAA-VAF</i> | ARNg 1 | - | Série 1 | - | 0 | - | | |
| | | | ACH-Spec | | - | 13 | - | | |
| | | | | | 15mL (20µL) | 9 | 0/9 | | |
| | | | | | 30mL (20µL) | 17 | 0/17 | | |
| | | | | | | | | | |
| | | ARNg 2 | - | | 0 | - | | | |
| | | | CH-Spec | | - | 52 | - | | |
| | | | | | 15mL (20µL) | 24 | 0/24 | | |
| | | | | | 30mL (20µL) | 65 | 0/65 | | |
| Total | | | | | | 115 | 0/115 | | |
| Essai 3 | Opéron <i>spoVAA-VAF</i> | ARNg 1 | - | Série 2 | - | 0 | - | | |
| | | | ACH-Spec | | - | 225 | - | | |
| | | | | | 2 mL (20µL) | 240 | 0/10 | | |
| | | | | | 5 mL (20µL) | 135 | 0/10 | | |
| | | | | | 10 mL (20µL) | 60 | 0/10 | | |
| | | ARNg 2 | - | | 0 | - | | | |
| | | | CH-Spec | | - | 272 | - | | |
| | | | | | 2 mL (20µL) | 241 | 0/10 | | |
| | | | | | 5 mL (20µL) | 146 | 0/10 | | |
| | | | | | 160 | 0/10 | | | |
| Total | | | | | | 982 | 0/60 | | |

génomique était nette. La « condensation artificielle » potentiellement appliquée sur cette préparation n'était pas visible directement sur ce gel (c'est également le cas avec les génomes de mycoplasmes), mais nous avons en revanche émis l'hypothèse qu'elle pourrait être perçue au cours des essais de CREasPy-Cloning. Nous avons donc conservé cette seconde série de blocs afin de les tester.

1.3.3. Troisième préparation et standardisation avec le protocole « Biorad »

Nous souhaitons obtenir des préparations de génomes de *B. subtilis* répétables dans le temps avec des solutions commerciales et calibrées. A contrario, le protocole ED fait appel à des tampons préparés au laboratoire et soumis à plus de variabilité (précision de la préparation, pureté des composants, variations lot à lot). (cf. §II.14.2. Matériels et Méthodes). Nous avons donc décidé de préparer une série n°3 avec le protocole commercial associé au kit Biorad. Ce protocole est proche de celui utilisé pour la préparation des blocs de la série n°2 ; la seule différence se situe dans la concentration de lysozyme à utiliser pour la lyse enzymatique de la paroi de *B. subtilis* : 5 mg.mL⁻¹ (protocole précédent) vs 1 mg.mL⁻¹ (protocole Biorad). Les résultats obtenus ont montré que la digestion n'était potentiellement pas complète (**Figure 35B**). En effet, quelques bandes supplémentaires sont visibles, notamment pour les volumes 5, 10 et 20 mL avec des tailles intermédiaires à ~285, 620, 750 et 900 kpb. Cette troisième et dernière préparation de blocs, moins propre, a été conservée mais n'a pas été utilisée lors des essais de clonage par CREasPy-Cloning.

1.4. Essais de clonage du génome entier de *B. subtilis*

Au total trois essais de clonage de génomes entiers de *B. subtilis* par CREasPy-Cloning ont été tentés (**Tableau 17**). Les levures pré-chargées avec le système CRISPR-Cas9 [VL6-48N +pCas9 +pARNg1 ou 2] ciblant soit l'opéron *sboAX-albAG*, soit l'opéron *spoVAA-VAF* ont été co-transformées avec deux matrices de recombinaison (**Figure 33C**) et avec les séries de blocs n°1 ou n°2. Le protocole original de transformation au PEG décrit en 2010 par Benders *et al.* a été appliqué²¹⁵.

Pour chaque essai, deux contrôles devaient être validés. Le premier était un contrôle dit « levure » qui permet de vérifier que les sphéroplastes ne sont pas capables de pousser sur la sélection SD-HIS si le pansement avec les éléments ARS-CEN-HIS3 ou CEN-HIS3 et le génome à cloner n'ont pas été acquis. Ainsi, aucune colonie n'a été observée sur les boîtes correspondantes. Le second contrôle consistait à transformer les levures uniquement avec les pansements de recombinaison ACH-SpecR ou CH-SpecR. En théorie, aucune colonie ne doit apparaître sur ce contrôle, mais en pratique des colonies sont quasiment toujours présentes sur ces boîtes. Cette expérience ne dérogeait pas à la règle, avec un total de 8 à 272 colonies en fonction du pansement ACH-Spec seul ou bien CH-Spec seul (**Tableau 17**). Plusieurs hypothèses ont été émises pour tenter d'expliquer ces observations : (i) la présence résiduelle de plasmide matrice ayant servi à l'amplification par PCR du pansement de recombinaison ; (ii) la circularisation du pansement de recombinaison par les systèmes de réparation de la levure (recombinaison homologue (HR) ou par jonction d'extrémités non homologues (non-homologous end-joining, NHEJ)) ; et /ou (iii) l'intégration du pansement de recombinaison dans le génome de la levure. Des expériences complémentaires menées au laboratoire nous ont permis de démontrer que lorsque les deux bras de recombinaison d'un même pansement ont des séquences courtes similaires, il peut dans certains cas se circulariser et se maintenir au sein de la levure. Ce phénomène a également été décrit dans la littérature sans informations précises sur les causes menant à la circularisation du pansement²³⁴.

Si la présence de colonies sur ce contrôle indique que des faux positifs sont certainement présents sur les boîtes essais, il s'avère que de « vrais » transformants peuvent également être présents sur ces boîtes, et parfois en très grand nombre. Aussi, nous avons donc pris pour habitude d'outrepasser ce contrôle, et de cribler en grand nombre les colonies présentes sur les boîtes essais.

L'essai n°1 réalisé avec des ARNg ciblant l'opéron *sbo-alb* et la série de blocs *B. subtilis* n°1 a permis d'obtenir au total 40 colonies toutes conditions confondues (**Tableau 17**). La quantité d'ADN des blocs ne semblait pas significativement influencer le nombre de transformants obtenus sur boîte. Après vérification par PCR, nous n'avons pas détecté la présence du génome de *B. subtilis* cloné dans les levures.

Suite à cet échec, deux points ont été soulevés :

- La séquence cible et le choix des ARNg : il n'est pas rare en effet d'observer des différences d'efficacité entre des ARNg(s) sélectionnés pour une même cible ou pour des cibles différentes. Nous nous sommes alors demandé dans quelle mesure la zone où nous souhaitions intégrer les éléments levure, mais aussi, les 2 ARNg sélectionnés, pouvaient avoir un impact sur le résultat du clonage.

- La quantité de génomes transformés et leur qualité : le génome de *B. subtilis* étant bien plus grand que ceux avec lesquels nous avons l'habitude de travailler, il se pouvait alors qu'il soit nécessaire d'en ajouter plus (ou moins) pour favoriser l'évènement de capture. Plus une molécule d'ADN est grande, plus elle est fragile (nombre de cassures mécaniques plus élevé).

L'essai n°2 a donc été réalisé en ciblant un second *locus*, l'opéron *spoVAA-VAF*. Lors de cette expérience, 115 colonies au total ont pu être dénombrées sur boîte, soit deux fois plus que dans l'expérience précédente (**Tableau 17**). D'autre part, une corrélation positive directe a pu être observée entre le nombre de transformants obtenus et la quantité d'ADN utilisé. Cependant, après vérification par PCR aucun transformant propageant le génome de *B. subtilis* n'a été détecté. Ce résultat suggérait que ni le *locus* cible, ni les ARNg sélectionnés ne semblaient être la cause de notre premier échec.

Au cours du dernier essai, nous avons conservé l'idée de cibler l'opéron *spoVAA-VAF* mais cette fois en co-transformant les levures avec la préparation de blocs issue de la série n°2. Notre objectif ici était de tester la qualité de cette seconde série de blocs et d'essayer d'évaluer l'effet des traitements avec le chloramphénicol et l'hexamine cobalt (III) chlorure. De façon surprenante, un total de 982 colonies sont apparues sur boîte, soit 24 et 8 fois plus que les expériences précédentes (**Tableau 17**). Cependant, le nombre de clones sur les contrôles « pansement » (225 pour ACH-Spec et 272 pour CH-Spec) fût considérable et supérieur à toutes les autres conditions testées. Ce résultat nous a amené à seulement analyser un nombre restreint de colonies (60 transformants). Bien que nous ayons augmenté le nombre de transformants obtenus sur boîte (plus de 1000) pour ce dernier essai, le génome cloné dans la levure n'a pas été détecté par PCR.

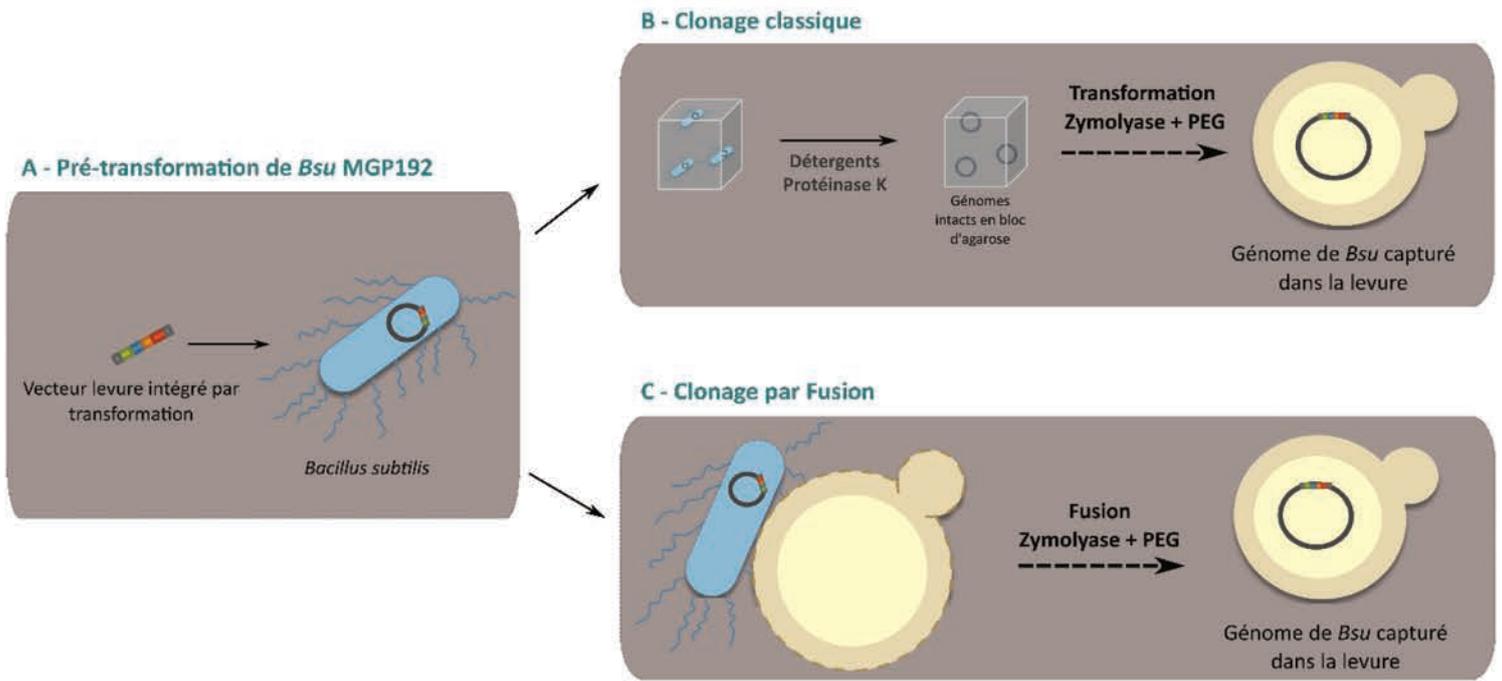


Figure 36 : Schéma des méthodes de clonage classique et par fusion adaptées à *B. subtilis*. [A] *B. subtilis* est transformé par compétence naturelle avec le plasmide intégratif pMT85-pRS313-pMiniTetM-pSPEC sans *lacZ-sbo* qui correspond au vecteur levure. Cette étape commune aux deux méthodes permet l'intégration des éléments levures « vecteur levure » dans le génome. [B] Les chromosomes marqués de *B. subtilis* sont purifiés et extraits intacts en blocs d'agarose, puis transformés dans la levure avec un protocole au PEG. [C] Des protoplastes de *B. subtilis* marqués sont directement fusionnés à des sphéroplastes de levures suivant un protocole modifié de transformation au PEG par fusion des cellules.

1.5. Conclusion sur les essais de clonage par la méthode de CReasPy-Cloning

Nos tentatives de clonage du génome de *B. subtilis* dans la levure par la méthode CReasPy-Cloning développée au laboratoire pour des génomes de mycoplasmes sont toutes restées infructueuses. A ce stade, plusieurs hypothèses pouvaient expliquer nos résultats : (i) la préparation des génomes en blocs restait sub-optimale malgré nos efforts ; (ii) la méthode de clonage n'était pas adaptée ; et/ou (iii) la taille du génome était trop élevée pour être clonée dans la levure. Chacun de ces points sont repris ci-dessous.

La qualité de la préparation de blocs étant fondamentale pour la réussite du clonage, nous ne pouvons pas exclure que malgré une mise au point poussée du protocole, cela ne soit pas suffisant et que la présence de contaminants délétères persistants puisse expliquer nos échecs répétés. Pour s'affranchir de ce point, il devenait donc essentiel de démontrer que l'ADN inclus dans les blocs était intrinsèquement « clonable » dans la levure, indépendamment des autres paramètres qui pouvaient expliquer les résultats négatifs de ces premières expériences.

La technique de clonage CReasPy-Cloning est une méthode de clonage efficace mais complexe. En effet, elle repose sur une cascade d'évènements successifs incluant l'entrée d'un génome entier dans la levure, le clivage de ce génome par le système d'édition CRISPR-Cas9, puis la réparation de la cassure double brin occasionnée par recombinaison homologue ; conduisant à l'intégration des éléments levures essentiels au maintien du génome et à l'expression de trois marqueurs de sélection. Cette méthode s'est montrée d'une réelle efficacité lors de clonages de plusieurs génomes de mycoplasmes (*Mycoplasma pneumoniae*, *Mycoplasma feriruminatoris*, *Mycoplasma mycoides* subsp. *mycoides*, *Mycoplasma leachii*). Cependant, elle comportait peut-être trop d'étapes critiques pour permettre le clonage d'un génome tel que celui de *B. subtilis*, qui est plus grand, qui possède un % en bases G+C plus élevé (43,5%) et, qui provient de bactéries possédant une paroi Gram+. De ce fait, revenir vers un protocole plus simple tel que le « clonage Classique », (**Figure 36B**), nous est apparu plus judicieux dans le cadre de nos objectifs.

Enfin, la taille du génome de la souche MGP192 à 2,86 Mb restait un de nos questionnements majeurs : est-il possible de cloner dans la levure (sous forme de YCp) des molécules d'ADN de 2,86 Mb ? Est-il possible que la levure puisse physiquement acquérir puis

maintenir des molécules d'ADN aussi grandes et provenant des blocs d'agarose (c'est-à-dire peu condensées) ? Si l'équipe avait déjà réussi à cloner un génome à 1,8 Mb, le projet *Bacillus 2.0* proposait de repousser cette limite. Afin de contourner cet aspect lié à la taille, ainsi qu'à l'état de relâchement des génomes *B. subtilis* en bloc d'agarose, nous nous sommes orientés vers une stratégie de clonage encore peu utilisée, appelée « clonage par Fusion » (**Figure 36C**). Cette stratégie est tout-à-fait originale puisqu'elle permettrait de « déverser » le contenu cellulaire (incluant le génome sous sa forme la plus compacte) de la bactérie vers la levure après contact membrane à membrane.

Le développement pour *B. subtilis* de ces deux nouvelles méthodes (clonage Classique et par Fusion) a été initié en parallèle. Les expérimentations mises en œuvre sont exposées dans la suite de ce chapitre.

II. Essais de clonage d'un génome pré-marqué de *B. subtilis*

Les deux méthodes de clonage « Classique » et par « Fusion » requièrent l'incorporation au préalable des éléments levures dans le génome bactérien, avant isolement en bloc d'agarose (stratégie Classique) ; ou avant fusion entre cellules de levures et cellules bactériennes (stratégie Fusion) (**Figure 36**). La souche MGP192 a donc dû être transformée en amont des essais de clonage.

II.1. Insertion des éléments levures dans le génome de *B. subtilis* par transformation

La transformation de *B. subtilis* est basée sur un protocole qui utilise sa compétence naturelle en début de phase stationnaire. Les cellules bactériennes sont rendues compétentes lors d'une croissance en milieu minimum enrichi appelé « MG1 », suivie d'une dilution dans un milieu carencé appelé « MG2 »¹⁴. Les éléments levures ont été intégré dans le génome de la souche MGP192 via le plasmide intégratif pMT85-pRS313-pMiniTetM-pSPEC sans *lacZ-sbo*. Cependant, par comparaison avec la souche type *B. subtilis* 168, il est plus difficile d'obtenir des transformants pour cette dernière. Nous avons donc effectué une mise au point du protocole spécialement pour la souche MGP192 en tentant (i) d'évaluer sa croissance en

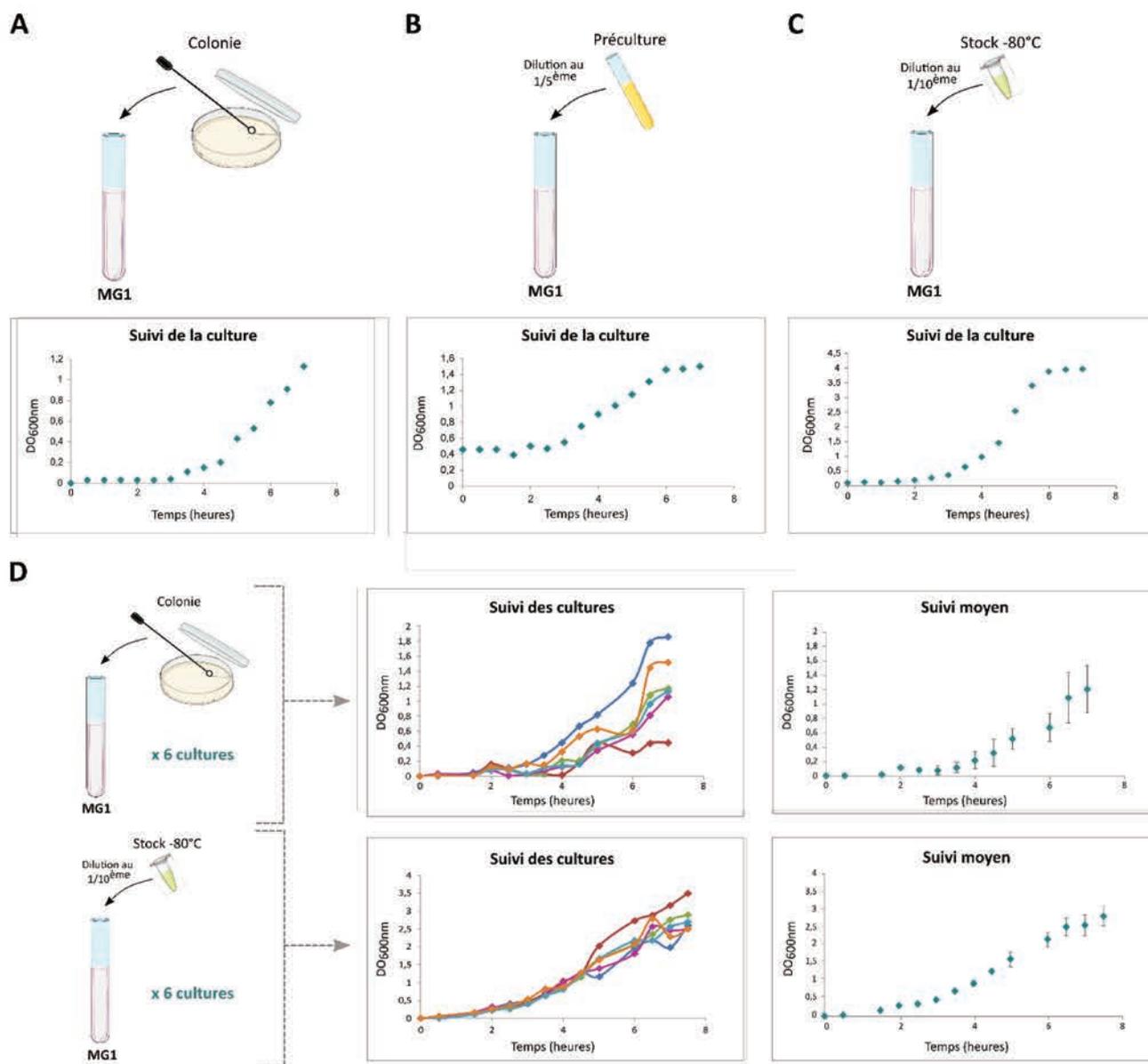


Figure 37 : Evaluation de la croissance de *B. subtilis* souche MGP192 dans le milieu MG1. Trois représentations graphiques de suivi de la DO à 600nm en fonction du temps sont représentées. Elles diffèrent par la manière dont le milieu MG1 a été ensemencé : **[A]** avec une colonie, **[B]** par dilution au 1/5^{ème} d'une préculture, ou **[C]** par dilution au 1/10^{ème} d'un stock glycérol à -80°C. **[D]** Des répliques ont été effectuées pour les conditions « Colonie » ou « Stock -80°C » : deux expérimentations avec trois répliques sont représentées. Les graphiques des six suivis de DO_{600nm} en fonction du temps sont indiqués par des couleurs différentes pour une culture donnée. La représentation graphique moyenne de ces répliques est aussi visualisée et le mode d'ensemencement (colonie ou stock -80°C) est comparé. Figure réalisée à l'aide de schémas issus du site : <https://smart.servier.com>.

milieu MG1 ou MG2 afin d'ajuster les temps d'incubation et (ii) d'utiliser différents ADN (ADNg total ou plasmide) et quantités de ces mêmes ADN pour transformer la souche.

II.1.1. Evaluation de la croissance de la souche MGP192 dans les milieux MG1 et MG2

Les courbes de suivi de DO_{600nm} en fonction du temps qui ont été effectuées ne sont pas de réelles courbes de croissance. En effet, ici nous souhaitons seulement mesurer la turbidité croissante du milieu (et non estimer le nombre d'UFC obtenues sur boîte), afin d'identifier le temps au bout duquel on atteignait la phase stationnaire (plus d'augmentation de la mesure de DO).

Une première expérience de mesure à DO_{600nm} en fonction du temps (dernière mesure à $T = 7h$ d'incubation) a été effectuée en comparant tout d'abord l'ensemencement de la culture : 3 cultures de 10 mL de milieu MG1 ont été ensemencées à partir, (i) d'une colonie fraîche repiquée d'un milieu LB gélosé ; (ii) d'une pré-culture de 10 mL démarrée la veille et incubée toute la nuit à $37^{\circ}C$ (ensemencement au $1/5^{ème}$) ; et (iii) d'un stock congelé de la souche MGP192 conservée à $-80^{\circ}C$ (ensemencement au $1/10^{ème}$). Les résultats sont présentés **Figure 37**.

Après ensemencement avec une colonie (méthode usuelle, **Figure 37A**), la phase de latence était plutôt longue (3h-3h30) et au bout de 7h d'incubation, la DO mesurée était seulement de 1,13 (soit environ au début de la phase exponentielle). Pour atteindre la fin de la phase exponentielle, nécessaire à l'état de compétence, il était donc nécessaire dans cette configuration, d'incuber la culture bien au-delà des 4h30 préconisées. Cela rendait complexe la mise en place d'un protocole sur la journée et pouvait expliquer les difficultés à appliquer le protocole original à cette souche.

En nous basant sur une mise au point similaire pour une autre souche de *B. subtilis* (DB104), pour laquelle les auteurs avaient fixé une DO_{600nm} de départ à $0,5^{72}$, nous avons effectué une seconde courbe de croissance à partir d'une pré-culture cultivée sur la nuit (**Figure 37B**). Le lendemain matin, la pré-culture avait une $DO_{600nm} = 2,45$ et a été diluée au $1/5^{ème}$ ($DO_{600nm} = 0,49$), puis des mesures au cours du temps ont été effectuées pendant 7h. Cette fois, la phase de latence semblait un peu plus courte (2h30), laissant place à un « semblant » de phase exponentielle, puis à une phase stationnaire qui stagnait à partir de 6h

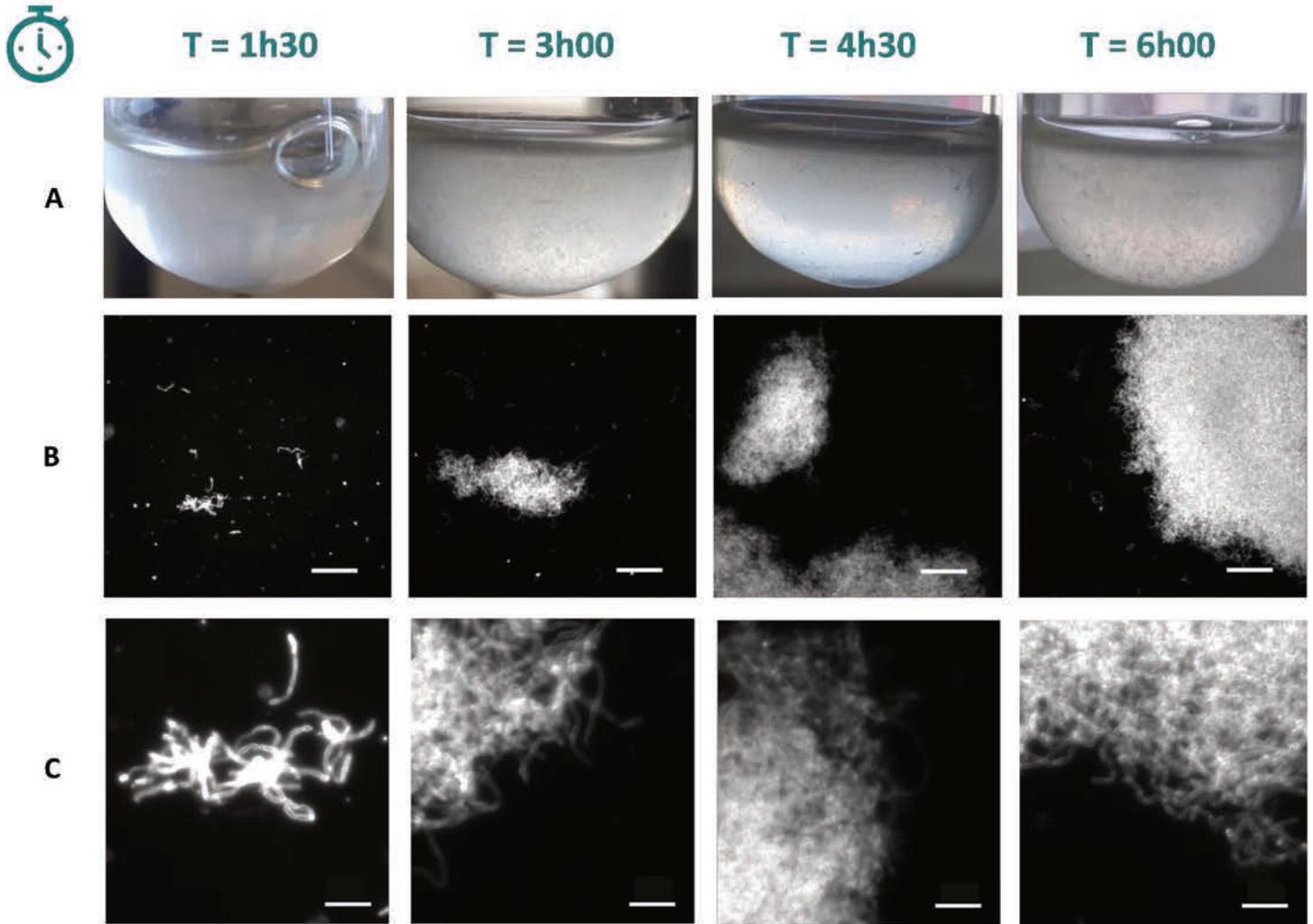


Figure 38 : Observation de la formation d'agrégats de cellules de *B. subtilis* MGP192 dans le milieu MG1 au cours du temps. [A] La formation des agrégats dans les tubes de culture en milieu MG1 est visualisée au cours du temps (de T = 1h30 à T = 6h00). **[B]** Observation en microscopie à fond noir grossissement x400, de la formation d'agrégats de plus en plus gros au cours du temps. L'échelle (bande blanche) correspond à 50 µm. **[C]** Image en zoom des observations en microscopie à fond noir au grossissement x400. L'échelle correspond à 10 µm.

de culture avec une DO_{600nm} de seulement 1,5 environ. Etant donné que précédemment, nous avons estimé qu'à cette DO nous pouvions encore nous trouver dans la phase exponentielle (*cf.* Suivi de la culture à partir d'une colonie), et que la souche peut normalement atteindre des mesures bien supérieures, nous avons estimé que cette méthode était inadaptée. En effet, la pré-culture sur la nuit devait déjà avoir largement atteint la phase stationnaire impliquant la présence de nombreuses cellules bactériennes en sénescence. La culture effectuée semblait donc ne pas être de bonne qualité et expliquer cet arrêt précoce de la croissance.

Finalement, le dernier suivi de croissance à partir d'une dilution au 1/10^{ème} d'un stock au -80°C était celui qui a donné les résultats les plus probants (**Figure 37C**). En effet, la phase de latence de 2h30 était rapidement remplacée par le début de la phase exponentielle qui atteint son pic au bout de 6h ($DO_{600nm} = 3,89$) avant de laisser place à la phase stationnaire. Cette DO finale mesurée avant le plateau reflétait une bonne quantité de cellules. Nous avons donc avec cette première démonstration un bon point de départ quant à la manière d'ensemencer la culture afin d'obtenir une culture dense et de qualité en finalité.

Si ces résultats montraient qu'il était nécessaire d'augmenter le temps de culture, un point majeur était à noter quant au phénotype de la souche MGP192 en milieu MG1 : elle formait de gros agrégats (de plus en plus nombreux au cours du temps) dans le milieu et il était fort possible que cela gêne les mesures de DO (**Figure 38**). C'est pourquoi l'expérience a été répétée deux autres fois et les cultures ont été réalisées en triplica (**Figure 37D**). Au vu des résultats obtenus lors de la 1^{ère} expérience, la condition « pré-culture » (**Figure 37B**) n'a pas été incluse dans ces deux nouvelles expériences.

Pour toutes les cultures ensemencées à partir de colonies, nous avons noté que la phase de latence s'était allongée jusqu'à 4h avant le début de la phase exponentielle (**Figure 37D**). A nouveau, au bout de 7h de culture, la fin de la phase exponentielle n'était toujours pas atteinte. De plus, la très grande variabilité observée dans les mesures du fait certainement des agrégats et ce, malgré une homogénéisation appliquée avant prélèvement, pouvait aussi venir de la taille de la colonie repiquée (*cf.* graphiques « Suivi des culture » et « Suivi moyen »). De ce fait, cette méthode a été abandonnée.

La méthode après dilution du stock au -80°C fût à nouveau la plus convaincante. Si cette fois, la phase stationnaire ne semblait pas être tout à fait atteinte, la similarité entre les

courbes représentées **Figure 37D**, ainsi que le profil de la courbe moyenne et des écart-types entre les points, montrait une répétabilité dans la croissance des bactéries en milieu MG1. C'est donc cette manière d'ensemencer qui a été validée et fixée pour tous les essais de transformation qui ont suivi.

Il faut noter que le même type d'expérience de mesure de DO a été effectué après dilution dans le milieu MG2 mais n'a pas donné de résultat concluant. La particularité de ce milieu est qu'il se compose d'un précipité de nitrate de calcium $\text{Ca}(\text{NO}_3)_2$ essentiel au succès du protocole de transformation. Cela rend difficile la lecture de l'absorbance à 600nm et la croissance de la souche en MG2 n'avait alors pas pu être appréciée aussi bien que dans le premier milieu.

II.1.2. Essai de transformation avec de l'ADNg pré-marqué

D'après les résultats précédemment décrits et grâce à l'expertise du Dr. Etienne Dervyn, (Chargé de Recherche à INRAE, Jouy-en-Josas), nous avons donc tenté un premier essai de transformation.

Comme évoqué un peu plus haut, nous souhaitons, à cette étape, intégrer en lieu et place de l'opéron *sboAX-albAG* du génome MGP192, une cassette portant les éléments levure ARS-CEN-HIS3 ainsi que le marqueur SpecR. Pour ce faire, nous avons construit un plasmide pMT85-pRS313-pMiniTetM-pSPEC sans lacZ-sbo qui porte la cassette d'intérêt encadrée par deux séquences de 1 kpb identiques aux régions encadrant l'opéron cible (cf. §II.1.2. et Figure 30 Matériels et Méthodes). Initialement, nous avons pour objectif d'introduire directement ce plasmide sous forme circulaire ou digérée par l'enzyme de restriction SmaI (libération de la cassette) dans *B. subtilis* MGP192 (**Figure 39**). Mais à cette étape, nos collaborateurs nous ont conseillé d'appliquer une approche différente. La souche MGP192 minimale (2,86 Mb) est difficile à transformer contrairement à d'autres souches de *B. subtilis* que nous avons à disposition au laboratoire comme la souche de référence *B. subtilis* 168 (wild-type : 4,2 Mb) ou la souche MS (3,98 Mb) (cf. §I.3. Matériels et Méthodes). Il est généralement plus efficace de transformer d'abord *B. subtilis* 168 ou MS avec le plasmide d'intérêt puis, d'utiliser « l'ADNg marqué » de ces souches pour transformer la souche MGP192 et obtenir des transformants avec la modification désirée.

Tableau 18 : Expérience n°1 de transformation de *B. subtilis* MGP192 avec le protocole MG1-MG2

| Condition | Incubation MG1 | Incubation MG2 | ADN transformé | Quantité | Criblage des transformants | |
|-----------|----------------|----------------|---|----------|----------------------------|-----------------|
| | | | | | UFC | Analyse par PCR |
| 1 | 4h30 | 0h30 | ADNg MSΔ <i>sbo- alb::ACH-Spec</i> | - | 0 | - |
| 2 | | | | 10 ng | 0 | - |
| 3 | | | | 100 ng | 1 | 0/1 |
| 4 | | 1000 ng | | 5 | 0/5 | |
| 5 | | 1h00 | | 10 ng | 0 | - |
| 6 | | | | 100 ng | 0 | - |
| 7 | | | | 1000 ng | 8 | 0/8 |
| 8 | | 1h30 | | 10 ng | 0 | - |
| 9 | | | | 100 ng | 0 | - |
| 10 | | | | 1000 ng | 1 | 0/1 |
| 11 | | 2h00 | | 10 ng | 0 | - |
| 12 | | | | 100 ng | 1 | 0/1 |
| 13 | | | | 1000 ng | 5 | 0/5 |
| 14 | | 2h30 | | 10 ng | 0 | - |
| 15 | | | | 100 ng | 1 | 0/1 |
| 16 | | | | 1000 ng | 4 | 0/4 |
| 17 | | 3h00 | | 10 ng | 0 | - |
| 18 | | | | 100 ng | 2 | 0/2 |
| 19 | 1000 ng | | 5 | 0/5 | | |
| 20 | 6h00 | 0h30 | - | 0 | - | |
| 21 | | | 10 ng | 0 | - | |
| 22 | | | 100 ng | 0 | - | |
| 23 | | 1000 ng | 5 | 0/5 | | |
| 24 | | 1h00 | 10 ng | 0 | - | |
| 25 | | | 100 ng | 1 | 0/1 | |
| 26 | | | 1000 ng | 2 | 0/2 | |
| 27 | | 1h30 | 10 ng | 1 | 0/1 | |
| 28 | | | 100 ng | 1 | 0/1 | |
| 29 | | | 1000 ng | 5 | 0/5 | |
| 30 | | 2h00 | 10 ng | 0 | - | |
| 31 | | | 100 ng | 1 | 0/1 | |
| 32 | | | 1000 ng | 6 | 0/6 | |
| 33 | | 2h30 | 10 ng | 0 | - | |
| 34 | | | 100 ng | 0 | - | |
| 35 | | | 1000 ng | 2 | 0/2 | |
| 36 | | 3h00 | 10 ng | 0 | - | |
| 37 | | | 100 ng | 1 | 0/1 | |
| 38 | 1000 ng | | 2 | 0/2 | | |

Nos collaborateurs ayant insisté sur le fait que l'efficacité de transformation est nettement supérieure avec de l'ADNg total qu'avec un plasmide (circulaire ou linéaire), nous avons donc appliqué cette stratégie. Au cours du premier essai (**Tableau 18**), nous avons tenté de transformer l'ADNg de la souche MS, issue de l'approche « Top-down » de réduction du génome de *B. subtilis*¹⁸⁴. Cette souche a été transformée sans difficulté avec le plasmide intégratif pMT85-pRS313-pMiniTetM-pSPEC sans lacZ-sbo permettant l'obtention d'un extrait d'ADNg marqué « ADNgMSΔsbo-*alb*::ACH-Spec ».

Nous avons poursuivi notre expérience en tentant de transformer cet ADNg marqué dans la souche *B. subtilis* MGP192. Nous avons fixé les paramètres de transformation suivants :

- Culture en milieu MG1 : pendant 4h30 (sur les conseils de notre collègue E. Dervyn) et 6h (d'après notre étude)
- Dilution en milieu MG2 et prélèvement des cellules toutes les 30 minutes pour transformation de l'ADNg

Les cellules de *B. subtilis* transformées ont été étalées sur milieu LB + Spec100μg.mL⁻¹. Après une incubation de 24h à 37°C, le nombre de colonies sur boîtes a été compté. Il est apparu qu'au total 33 UFC ont été obtenues après culture en MG1 pendant 4h30, contre 27 UFC après culture pendant 6h. Nous avons aussi observé, que plus la quantité d'ADNg transformée est importante (1000 ng) plus on obtient de colonies. L'ADNg de tous les transformants a été extrait et analysé par PCR. Si le résultat était plutôt encourageant car c'était la première fois que nous obtenions des colonies sur boîte après transformation de la souche MGP192, aucune colonie ne fût cependant validée. A ce stade, nous n'avions donc pas de mutant MGP192Δsbo-*alb*::ACH-Spec.

Lors de cette expérience, nous avons pu observer que le nombre de colonies obtenues en fonction des conditions de culture ne suivait pas de logique. En effet, les données obtenues indiquaient qu'il n'y avait pas forcément de différence entre la croissance de la souche pendant 4h30 ou 6h en milieu MG1 ; de même, aucune condition de transformation en milieu MG2 ne se distinguait des autres. Conscients du phénotype particulier de cette souche en culture (nombreux agrégats formant des « pelotes de laine » impossibles à désagréger par simple pipetage ; cf. §II.1.1 + **Figure 38**), ce résultat ne fût pas une réelle surprise, d'autant qu'il pouvait être conforté par des données bibliographiques. En effet, une publication récente

Tableau 19 : Expérience n°2 de transformation de *B. subtilis* MGP192 avec le protocole MG1-MG2.

| Condition | Incubation MG1 | Incubation MG2 | ADN transformé | Quantité | UFC | Criblage des transformants | | | |
|-----------|----------------|----------------|---------------------|---------------------|--------|----------------------------|------|------------|---|
| | | | | | | Analysés | PCR | Séquençage | |
| 1 | 4h30 | 0h30 | - | - | - | - | - | - | |
| 2 | | | ADNg | 1 µg | 0 | - | - | - | |
| 3 | | | | 2 µg | 42 | 30* | 2/30 | - | |
| 4 | | | Plasmide circulaire | 0,2 µg | 0 | - | - | - | |
| 5 | | | | 2 µg | 0 | - | - | - | |
| 6 | | | Plasmide linéarisé | 0,2 µg | 0 | - | - | - | |
| 7 | | | | 2 µg | 0 | - | - | - | |
| 8 | | | 1h00 | ADNg | 1 µg | 0 | - | - | - |
| 9 | | | | | 2 µg | 1 | 0/1 | 0/1 | - |
| 10 | | | | Plasmide circulaire | 0,2 µg | 0 | - | - | - |
| 11 | | | | | 2 µg | 0 | - | - | - |
| 12 | | | | Plasmide linéarisé | 0,2 µg | 0 | - | - | - |
| 13 | | 2 µg | | | 0 | - | - | - | |
| 14 | | 1h30 | ADNg | 1 µg | 0 | - | - | - | |
| 15 | | | | 2 µg | 0 | - | - | - | |
| 16 | | | Plasmide circulaire | 0,2 µg | 0 | - | - | - | |
| 17 | | | | 2 µg | 0 | - | - | - | |
| 18 | | | Plasmide linéarisé | 0,2 µg | 0 | - | - | - | |
| 19 | | | | 2 µg | 0 | - | - | - | |
| 20 | | 2h00 | ADNg | 1 µg | 0 | - | - | - | |
| 21 | | | | 2 µg | 0 | - | - | - | |
| 22 | | | Plasmide circulaire | 0,2 µg | 0 | - | - | - | |
| 23 | | | | 2 µg | 0 | - | - | - | |
| 24 | | | Plasmide linéarisé | 0,2 µg | 0 | - | - | - | |
| 25 | | | | 2 µg | 1 | 1/1 | 1/1 | 1/1 | |
| 26 | | 2h30 | ADNg | 1 µg | 220 | 30* | 0/30 | - | |
| 27 | | | | 2 µg | 0 | - | - | - | |
| 28 | | | Plasmide circulaire | 0,2 µg | 0 | - | - | - | |
| 29 | | | | 2 µg | 43 | 30* | 0/30 | - | |
| 30 | | | Plasmide linéarisé | 0,2 µg | 108 | 30* | 0/30 | - | |
| 31 | | | | 2 µg | 0 | - | - | - | |
| 32 | | 3h00 | ADNg | 1 µg | 0 | - | - | - | |
| 33 | | | | 2 µg | 0 | - | - | - | |
| 34 | | | Plasmide circulaire | 0,2 µg | 60 | 30* | 0/30 | - | |
| 35 | | | | 2 µg | 0 | - | - | - | |
| 36 | | | Plasmide linéarisé | 0,2 µg | 1 | 1 | 0/1 | - | |
| 37 | | | | 2 µg | 0 | - | - | - | |

*Clones extraits en pool

a décrit que les protéines impliquées dans la liaison de l'ADN à la surface de la cellule de *B. subtilis* se situeraient majoritairement aux pôles du bacille²⁸⁰. Dans notre cas, les extrémités étant probablement difficilement accessibles pour l'ADN, le nombre de cellules « transformées » doit être extrêmement faible, voire nul.

II.1.3. Nouvel essai de transformation avec différents ADN pré-marqués

Nous avons à nouveau tenté de transformer la souche MGP192. Afin de diminuer la quantité d'agrégats dans la culture MG1 (qui augmente avec le temps) et comme nous n'avions pas observé de différence entre les deux temps d'incubation testés, les cellules ont été cultivées pendant 4h30 en milieu MG1. Après dilution en milieu MG2, toutes les 30 minutes, les bactéries ont été pipetées abondamment puis transformées avec différents ADN : (i) l'ADNg MSΔ*sbo-alb*::ACH-Spec (1 µg ou 2 µg), (ii) le plasmide pMT85-pRS313-pMiniTetM-pSPEC sans *lacZ-sbo* (0,2 µg ou 2 µg) circulaire ou (iii) hydrolysé par restriction avec l'enzyme *Sma*I (excision de la cassette).

Après incubation, pendant 48h à 37°C afin d'obtenir un maximum de transformants, le nombre de colonies sur boîtes a été répertorié (**Tableau 19**). Nous avons considérablement augmenté l'efficacité de la transformation, puisqu'il est apparu 476 UFC au total toutes conditions confondues. Au cours de cet essai, seules 7 conditions ont permis d'obtenir entre 1 et 220 colonies. Les conditions où le plus grand nombre d'UFC a été observé sont les suivantes [2h30 de culture en MG2 puis ajout d'1 µg d'ADNg] et [2h30 de culture en MG2 puis ajout de 0,2 µg de plasmide linéarisé]. L'ADNg de plusieurs transformants a été extrait et analysé par PCR. Les conditions dont les nombres de colonies obtenues étaient les plus grands [3, 26, 29, 30 et 34] ont été analysées par pool (3 pools de 10 clones) pour un ensemble de 30 UFC au total. Les clones uniques obtenus pour les conditions [9], [25] et [36] ont aussi été analysés. Sur tous les pools mis en culture, seulement deux sont devenus troubles, révélant qu'un grand nombre de colonies faussement positives étaient finalement apparues sur les boîtes. De la même manière, seules deux cultures effectuées à partir des clones individuels (11.7 et 33.1) avaient poussé. Les ADNg totaux ont donc été extraits et analysés par PCR (**Figures 39 et 40**). Deux analyses ont été effectuées afin de vérifier les jonctions (PCR 1 et PCR 2) synonymes d'intégration de la cassette ARS-CEN-HIS3-SpecR au bon endroit dans le

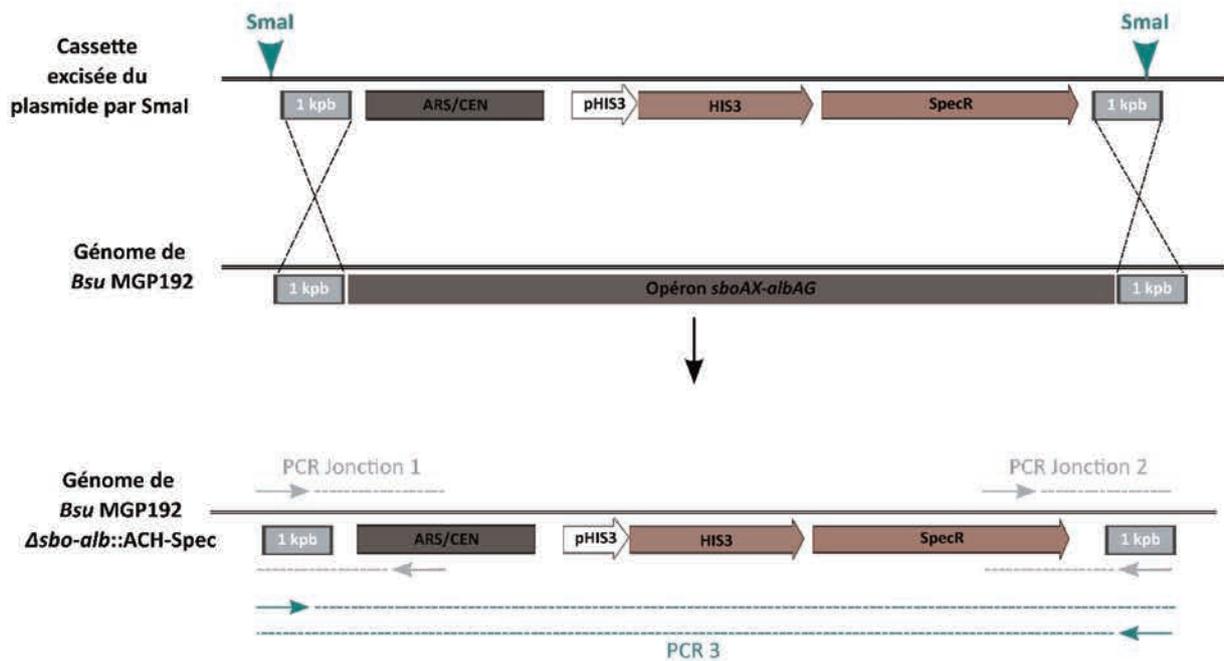


Figure 39 : Schéma de l'intégration des éléments levures dans le génome de *B. subtilis* MGP192. L'intégration de la cassette [ARS-CEN-HIS3-specR] (excisée du plasmide par restriction avec l'enzyme SmaI) au locus cible par recombinaison homologue est représentée. Les PCR ayant permis les analyses après transformation sont indiquées par des flèches grises et pointillés associés (PCR jonction 1 et 2), et par des flèches bleues et pointillés correspondants (PCR 3).

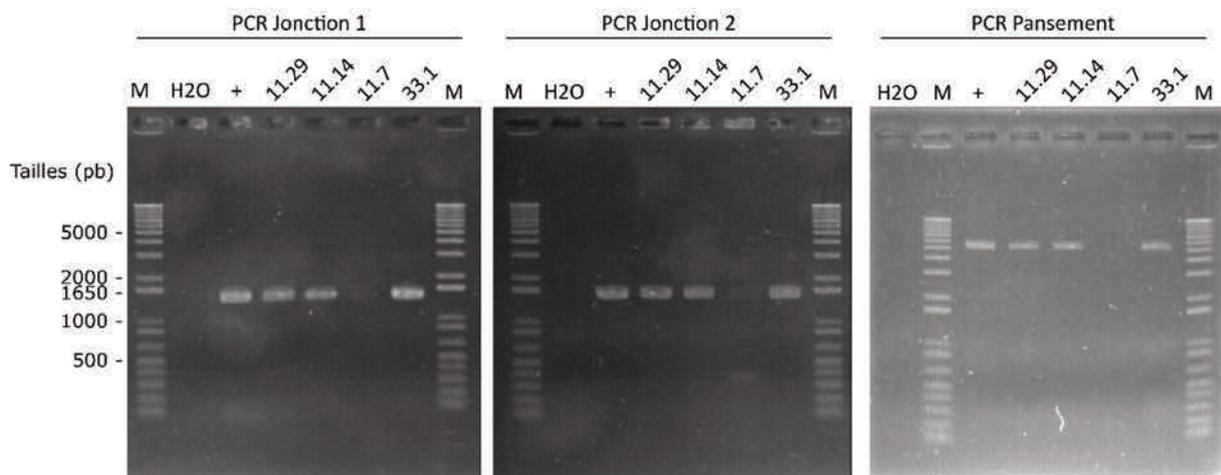


Figure 40 : Criblage des transformants de *B. subtilis* MGP192 obtenus après transformation par compétence naturelle. Cette figure regroupe les résultats de PCR simplexe pour l'analyse des transformants de *B. subtilis*. "M": DNA Ladder 1kb+ Invitrogen (100-12000 pb). "H2O": témoin négatif sans ADN. "+": témoin positif avec l'ADNg de *B. subtilis* MGP192. L'intégrité du pansement intégré est évaluée par trois PCRs « Jonction 1 », « Jonction 2 » et « PCR3 ». Elle est validée lorsque les clones arborent respectivement des amplifiats à 1453 pb, 1558 pb et 5306 pb, à l'instar du témoin positif.

génomique de *B. subtilis*, c'est-à-dire à la place de l'opéron *sboAX-albAG*. Une troisième analyse (PCR 3) a permis, quant à elle, de vérifier l'intégration complète de la cassette au locus cible. Les deux pools « 11.29 et 11.14 » (issus de la condition 3) et le clone 33.1 (issu de la condition 25) ont été validés avec un profil identique au résultat attendu pour chaque PCR. Etant donné que pour poursuivre nos expériences, seul un clone suffisait, nous avons approfondi l'analyse du cl33.1 en séquençant l'amplicon après purification de la « PCR 3 ». Le séquençage a permis de valider que le clone 33.1 présentait le génotype MGP192 Δ *sbo-alb*::ACH-Spec, ouvrant ainsi la voie à nos expériences de clonage classique ou par fusion.

II.2. Essais de clonage du génome MGP192 par la méthode « classique »

Deux préparations en blocs d'agarose ont été effectuées afin de purifier le génome de *B. subtilis* MGP192 Δ *sbo-alb*::ACH-Spec cl33.1. A l'instar des résultats présentés précédemment pour *B. subtilis* MGP192 non marqué (séries 2 et 3), les protocoles « ED » et « Biorad » ont été utilisés afin d'obtenir des chromosomes bactériens *B. subtilis* MGP192 marqués, purs et intacts.

A partir de ces préparations, nous avons tenté de cloner le génome de *B. subtilis* MGP192 marqué avec les éléments levures en utilisant le protocole « standard » de transformation de sphéroplastes de levure au PEG²⁸¹ (**Figure 36B**). De la même manière que pour les essais de clonage par la méthode de CReasPy-Cloning, nous avons utilisé la souche VL6-48N.

Si le contrôle négatif (condition 1) montrait bien que les levures VL6-48N n'étaient pas capables de pousser sur la sélection SD-HIS (pas de colonies sur boîte), nous pouvions également observer qu'aucune colonie n'était visible sur les boîtes de transformation et ceci quelle que soit la condition (**Tableau 20**). L'un des avantages de cette méthode réside en l'intégration au préalable des éléments levures dans le chromosome bactérien que l'on souhaite cloner. Ici il était donc peu probable d'avoir des faux positifs sans qu'un moins un morceau du génome ne soit rentré dans les cellules de levure. Ce premier résultat était donc très surprenant puisque l'absence totale de colonies pouvait signifier que, (i) l'expérience n'avait pas fonctionné (l'absence de contrôle positif ne nous permettait pas de valider

Tableau 20 : Essai n°1 de clonage du génome complet de *B. subtilis* par la méthode classique

| Condition | Souche de levure | Série de blocs | Quantité Blocs | Traitement hexamine cobalt(III) chloride | UFC | |
|-----------|------------------|---|---|--|-------------|---|
| 1 | VL6-48N | - | - | - | 0 | |
| 2 | | MGP192Δ <i>sbo-alb</i> ::ACH-Spec protocole "ED" | 2 mL (20μL) | - | 0 | |
| 3 | | | 5 mL (20μL) | | 0 | |
| 4 | | | 10 mL (20μL) | | 0 | |
| 5 | | | 20 mL (20μL) | | 0 | |
| 6 | | | 2 mL (20μL) | | + | 0 |
| 7 | | | 5 mL (20μL) | | | 0 |
| 8 | | | 10 mL (20μL) | | | 0 |
| 9 | | | 20 mL (20μL) | | | 0 |
| 10 | | | MGP192Δ <i>sbo-alb</i> ::ACH-Spec protocole "Biorad" | | 2 mL (20μL) | - |
| 11 | | 5 mL (20μL) | | 0 | | |
| 12 | | 10 mL (20μL) | | 0 | | |
| 13 | | 20 mL (20μL) | | 0 | | |
| 14 | | 2 mL (20μL) | | + | 0 | |
| 15 | | 5 mL (20μL) | | | 0 | |
| 16 | | 10 mL (20μL) | | | 0 | |
| 17 | | 20 mL (20μL) | | | 0 | |

Tableau 21 : Essai n°2 de clonage du génome complet de *B. subtilis* par la méthode classique

| Condition | Souche de levure | Plasmide | Série de blocs | Quantité Blocs | Traitement Hexamine cobalt(II) chloride | UFC |
|-----------|------------------|----------|---|----------------|---|-------|
| 1 | VL6-48N | - | - | - | - | 0 |
| 2 | | 300 ng | - | - | - | >1000 |
| 3 | | - | MGP192Δ <i>sbo-alb</i> ::ACH-Spec protocole "ED" | 5 mL (20μL) | - | 0 |
| 4 | | | | 5 mL (20μL) | + | 0 |
| 5 | | | MGP192Δ <i>sbo-alb</i> ::ACH-Spec protocole "Biorad" | 5 mL (20μL) | - | 0 |
| 6 | | | | 5 mL (20μL) | + | 0 |
| 7 | W303a | - | - | - | - | 0 |
| 8 | | 300 ng | - | - | - | >300 |
| 9 | | - | MGP192Δ <i>sbo-alb</i> ::ACH-Spec protocole "ED" | 5 mL (20μL) | - | 0 |
| 10 | | | | 5 mL (20μL) | + | 0 |
| 11 | | | MGP192Δ <i>sbo-alb</i> ::ACH-Spec protocole "Biorad" | 5 mL (20μL) | - | 0 |
| 12 | | | | 5 mL (20μL) | + | 0 |

l'expérience), ou bien, que (ii) l'apport du génome de 2,86 Mb via des préparations en blocs d'agarose ne pouvait permettre son entrée dans les sphéroplastes.

Nous avons donc préféré réitérer l'expérience, en ajoutant un contrôle positif (plasmide pMT85-pRS313-pMiniTetM-pSPEC sans *lacZ-sbo*) et en transformant, en plus de la souche VL6-48N, la souche W303a, afin d'écartier la possibilité que ce premier échec soit lié à la souche VL6-48N. Pour rappel, ces deux souches ont des génotypes différents, mais des travaux de clonage et d'ingénierie de génome bactériens, ont montré qu'elles pouvaient toutes deux être utilisées avec succès^{220,230}. En ajoutant, la souche W303a, il devenait difficile de conserver toutes les conditions sans que le plan de l'expérimentation ne soit trop conséquent. C'est pourquoi nous avons réduit nos conditions de transformation en sélectionnant uniquement des blocs préparés à partir de 5 mL de culture de *B. subtilis*.

Les résultats obtenus et présentés **Tableau 21** montraient clairement que cette fois encore le génome n'avait pas été cloné dans la levure, et dans une plus grande mesure n'était peut-être pas capable de rentrer dans les sphéroplastes. En effet, aucune colonie sur boîte n'est apparue lorsque les levures ont été traitées seules (conditions 1 et 7, contrôles négatifs) ou traitées avec des génomes (conditions 3 à 6 et 9 à 12). Au contraire 300 à 1000 colonies ont été obtenues après transformation de 300 ng de plasmide (conditions 2 et 8, contrôles positifs), que ce soit respectivement pour la souche W303a ou pour la VL6-48N. Ceci confirmait que les éléments levures étaient bien fonctionnels sur ce plasmide, et qu'il devait nous être possible de sélectionner le génome marqué (si tant est qu'aucune mutation ne soit apparue une fois les éléments intégrés chez *B. subtilis*). Ce contrôle validait également que l'expérience s'était bien déroulée et que notre échec à cloner le génome de *B. subtilis* pouvait potentiellement résider ailleurs. Nous nous sommes à nouveau demandé si notre incapacité à cloner le génome était due à la taille du génome, à la préparation des blocs d'agarose dont la qualité n'était pas suffisante, ou bien si finalement la méthode de clonage classique était, elle aussi, non adaptée au clonage du génome de *B. subtilis*.

Il nous est alors apparu, que l'hypothèse émise lors des essais de clonage par CReasPy-Cloning, selon laquelle le génome préparé en bloc d'agarose et déprotéinisé pouvait être trop relâché, et empêcher sa transformation était plausible. En effet, il était vraiment étonnant de ne voir apparaître aucune colonie sur boîte au cours de ces deux expériences ; nous aurions pu à minima, sélectionner quelques clones levures hébergeant un morceau de génome

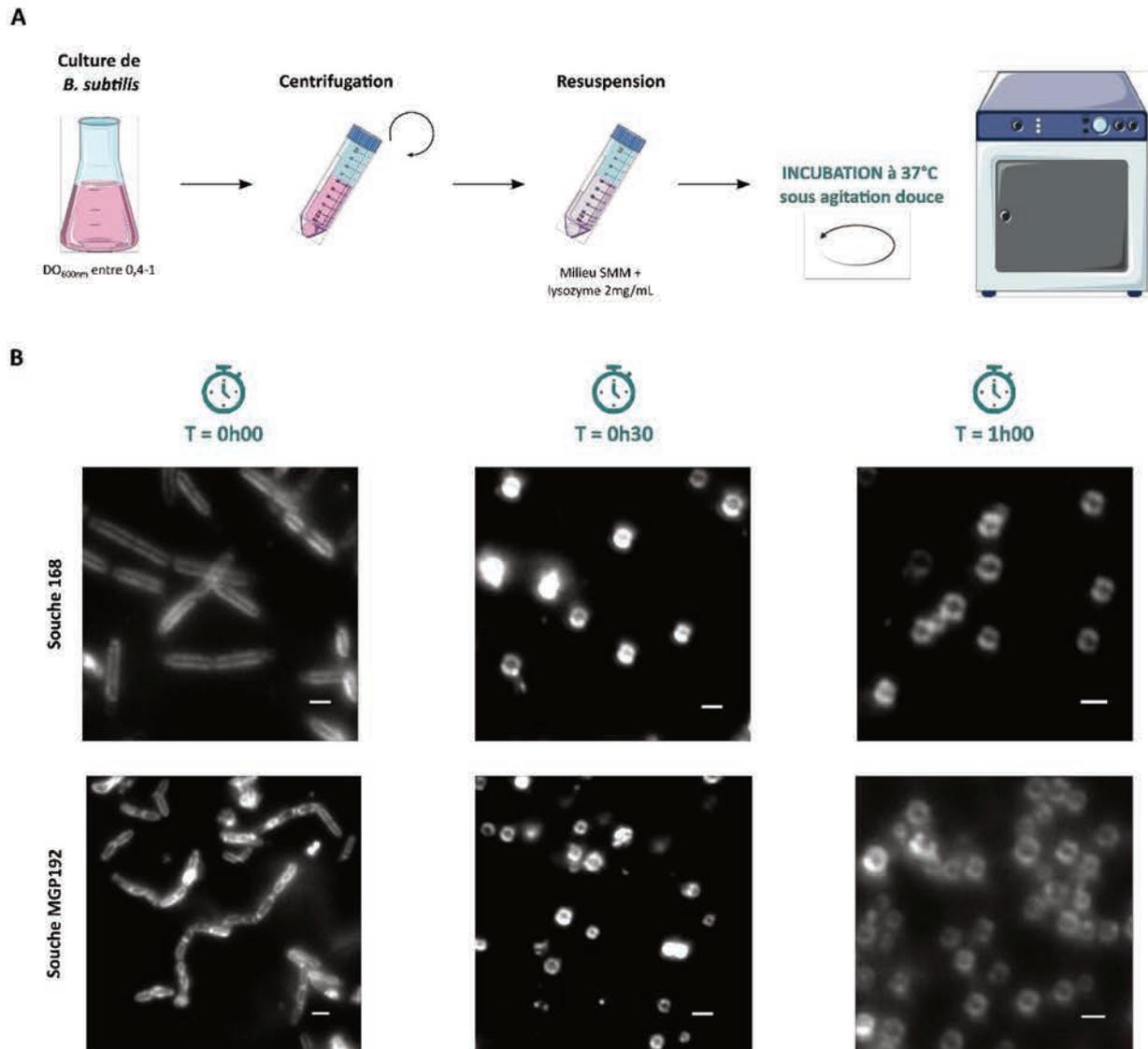


Figure 41 : Observation au microscope à fond noir de la formation de protoplastes de *B. subtilis* MGP192 au cours du temps. Figure réalisée à l'aide de schémas issus du site : <https://smart.servier.com>. **[A]** Schéma du protocole de production de protoplastes. Une culture de *B.subtilis* est cultivée à 37°C dans du LB jusqu'à atteindre une DO_{600nm} comprise entre 0,4 et 1 (soit milieu de phase exponentielle). Les cellules sont centrifugées et resuspendues dans le milieu hypertonique SMM additionné de 2mg/mL de lysozyme. Puis elles sont incubées à 37° dans un incubateur sous agitation douce. **[B]** Image en zoom des observations en microscopie à fond noir à l'objectif x400. L'échelle (bande blanche) correspond à 1 µm. Les cellules de *B. subtilis* souche 168 ou MGP192 sont visualisées à T = 0h00 sans traitement, à T = 0h30 soit au bout de 30 minutes d'incubation et à T = 1h00 au bout d'une heure d'incubation.

B. subtilis MGP192 de plus ou moins grande taille contenant les éléments ARS-CEN-HIS3-SpecR, or, ce ne fût pas le cas.

II.3. Essais de clonage par Fusion

Les résultats obtenus et jusqu'alors décrits tendent à montrer que pour cloner le génome de *B. subtilis* dans la levure, nous sommes techniquement limités par sa taille et par la méthode que nous devons utiliser. Aussi, les méthodes de clonage testées et basées sur l'extraction du génome en bloc d'agarose ne semblaient pas les plus adaptées. C'est pourquoi nous nous sommes efforcés de mettre en place au laboratoire la méthode de clonage « par Fusion » qui consiste à transférer le génome bactérien que l'on veut cloner, par fusion entre des cellules bactériennes et des sphéroplastes de levure (**Figure 36C**).

II.3.1. Production des protoplastes de *B. subtilis*

Le protocole de Fusion tel que décrit par Karas *et al.*, implique pour la bactérie Gram-*Haemophilus influenzae* un traitement de la paroi avec du lysozyme afin de produire des protoplastes²²³. La paroi de *B. subtilis* étant Gram+, il était alors nécessaire d'appliquer un traitement adéquat. En nous appuyant sur des travaux relatant la fusion de protoplastes de *B. subtilis* au PEG^{77,78}, nous avons mis au point au laboratoire un protocole spécifique de production de protoplastes. Le but était de trouver les conditions pour lesquelles la paroi bactérienne serait suffisamment dégradée sans pour autant que les cellules ne soient lysées. Ces dernières ont été cultivées jusqu'à atteindre la phase exponentielle puis reprises dans un milieu hypertonique appelé tampon SMM. Elles ont ensuite été traitées par ajout de lysozyme à 2 mg.mL⁻¹ et incubation à 37°C sous agitation douce (50 rpm) (**Figure 41A**). Au cours de ce traitement, la formation de protoplastes a été observée au cours du temps par microscopie à fond noir. Les protoplastes de *B. subtilis* étaient visibles dès 30 min de traitement ; à T30, il était en effet possible d'observer la présence de cellules bactériennes dont la forme est ronde plutôt que bacillaire. Que ce soit pour la souche de référence *B. subtilis* 168 ou bien pour notre souche d'intérêt MGP192, nous avons pu valider qu'au bout de 1h la totalité des bactéries avait une forme ronde caractéristique des protoplastes, voire de donuts avec un centre plus foncé. Bien que la souche MGP192 se divise de manière désordonnée et forme des

Tableau 22 : Essai n°1 de clonage du génome complet de *B. subtilis* par la méthode de fusion

| Condition | Levure | Volume Sphéroplastés | <i>B. subtilis</i> | Volume Protoplastés | UFC | Criblage des transformants | | | |
|-----------|---------|----------------------|-----------------------------------|---------------------|-----|----------------------------|--------------|----------------|----------------------------|
| | | | | | | Analysés | PCR simplexe | PCR multiplexe | PCR multiplexe Sous-clones |
| 1 | VL6-48N | 100 µL | - | - | 0 | - | - | - | - |
| 2 | | - | - | 100 µL | 0 | - | - | - | - |
| 3 | | 100 µL | MGP192Δ <i>sbo-alb</i> ::ACH-Spec | 25 µL | 8 | 8 | 4/8 | 0/4 | - |
| 4 | | | | 50 µL | 7 | 7 | 3/7 | 0/3 | - |
| 5 | | 200 µL | SMM + 20mM MgCl ₂ | 100 µL | 3 | 3 | 3/3 | 1/3 | 0/20 |
| 6 | | | | 25 µL | 13 | 9 | 6/9 | 0/6 | - |
| 7 | | | | 50 µL | 12 | 7 | 5/7 | 0/5 | - |
| 8 | | | | 100 µL | 15 | 12 | 12/12 | 1/12 | 0/20 |
| 9 | | 100 µL | MGP192Δ <i>sbo-alb</i> ::ACH-Spec | 25 µL | 3 | 1 | 0/1 | - | - |
| 10 | | | | 50 µL | 6 | 6 | 4/6 | 0/4 | - |
| 11 | | | | 100 µL | 3 | 3 | 2/3 | 0/2 | - |
| 12 | | 200 µL | SMM + 2,5mM MgCl ₂ | 25 µL | 8 | 8 | 5/8 | 0/5 | - |
| 13 | | | | 50 µL | 3 | 3 | 1/3 | 0/1 | - |
| 14 | | | | 100 µL | 12 | 10 | 7/10 | 1/7 | 0/20 |

« pelotes de laine », le traitement semblait tout aussi efficace pour obtenir des protoplastes isolés. **Figure 41B**, l'image sélectionnée pour cette souche correspond à un reste d'agrégats désolidarisé et composé de cellules isolées. Nous pouvons également noter que les cellules traitées au bout d'1h étaient moins denses, signifiant certainement que la paroi était moins épaisse et donc mieux digérée qu'après 30 minutes de traitement.

Un point important à noter était que le milieu hypertonique SMM, ainsi que, l'agitation douce nous avait permis d'obtenir une préparation sans lyse cellulaire (c'est-à-dire sans viscosité due au relargage de l'ADN) ; le respect scrupuleux de la phase de croissance aidant également à obtenir une préparation de qualité. En effet, au cours d'autres préparations, nous nous sommes aperçus qu'il s'avérait néfaste de travailler avec des cultures trop poussées (DO entre 2 et 2,5). Nous avons donc ainsi validé notre protocole : culture des cellules jusqu'à atteindre une DO_{600nm} comprise entre 0,6 et 1 ; et temps de traitement minimal avec le lysozyme fixé à 1h. Ces paramètres permettaient d'avoir un maximum de bactéries traitées et dont la paroi est convenablement digérée.

II.3.2. Premier essai de clonage de génome par Fusion *B. subtilis*-Levure

En se basant sur le protocole développé par Karas *et al.* en 2013, des sphéroplastes de levure VL6-48N ont été fusionnés avec des protoplastes de *B. subtilis* MGP192 Δ sbo-alb::ACH-Spec. Quelques modifications ont néanmoins été apportées au protocole. Nous avons notamment décidé de tester différents volumes de sphéroplastes et de protoplastes, et de faire varier la concentration en $MgCl_2$ du milieu SMM. En effet, par comparaison avec le milieu de resuspension utilisé dans le protocole original²³⁶, la concentration en $MgCl_2$ requise dans le SMM de *B. subtilis* était largement supérieure (2,5 mM protocole de Fusion original vs 20 mM protocole protoplastes de *B. subtilis*). Nous nous étions alors demandé si cela pouvait avoir un réel impact sur la qualité des protoplastes.

Après transformation et sélection des clones sur milieu sélectif SD-HIS, aucune colonie n'est apparue sur les boîtes « témoin négatif » que ce soit pour les sphéroplastes étalés seuls (condition 1), ou pour les protoplastes de *B. subtilis* (condition 2) qui ne sont pas capables de pousser avec cette sélection. Une centaine de colonies, toutes conditions confondues, a été obtenue sur le reste des boîtes (conditions 3 à 14). Un nombre de colonies légèrement plus

grand a été obtenu lorsque 200 µL de sphéroplastes ont été fusionnés plutôt que 100 µL. De même, que lorsque les protoplastes de *B. subtilis* ont été repris dans du SMM 20 mM de MgCl₂ (58 clones (conditions 3 à 8) contre 35 avec le SMM 2,5 mM de MgCl₂ (conditions 9 à 14) (**Tableau 22**).

Ce premier essai s'est révélé tout à fait informatif puisque des colonies sont apparues sur boîtes de sélection contrairement à la précédente méthode (« Classique »). Après extraction et analyse par PCR simplexe avec le couple d'amorces Bsu14F/R (**Figure 42B**), nous avons pu pour la première fois détecter la présence du génome de *B. subtilis* dans des transformants levure. Sur 77 clones repiqués et analysés, 52 se sont avérés positifs avec un amplicon PCR à la taille attendue (442 pb) conformément au témoin positif. Au cours de l'analyse suivante (PCR, multiplexe, **Figure 42C et Figure 44**), 2 clones sur 52 positifs (cl8.10 et cl8.11), ont arboré un profil valide en PCR multiplexe (profil avec 9 bandes identique au témoin positif ADNg *B. subtilis* MGP192). Ce résultat prometteur signifiait que non seulement le génome de *B. subtilis* avait pu rentrer dans les sphéroplastes de levure, mais qu'il était potentiellement « entier » dans ces 2 clones. D'autres clones affichaient un profil complet (cl5.1, 6.9, cl7.1, cl10.2, cl12.3, cl14.8) mais avec des bandes d'intensité assez hétérogènes, la bande à 1020 bp étant la plus intense (Bsu20F/R). Ceci laissait supposer que l'on pouvait avoir en réalité un mélange de clones, ou bien, que le génome se dégradait. L'étape de sous-clonage (stries, **Figure 42A**) est ordinairement effectuée avant confection des patchs et extraction. Cette étape ajoutant 48h d'attente avant analyse des transformants, nous avons dans un premier temps voulu nous affranchir de ce temps additionnel et obtenir le résultat de la méthode le plus rapidement possible. Cependant, nous sommes revenus sur cette étape afin d'éviter d'analyser ce qui pourrait s'apparenter à un « mélange de population ». Nous avons procédé à une troisième analyse (**Figure 42D, Figure 44**) et cette fois, nous avons pu constater que deux bandes persistaient majoritairement dans quasi tous les clones analysés : la bande à 911 bp pour le set d'amorces n°1 (Bsu9F2/R2) et la bande à 1020 bp pour le set d'amorces n°2 (Bsu20F/R) et sans surprise, ces couples se situent à proximité du site d'insertion de la cassette ARS-CEN-HIS3-SpecR (position 2,7 Mb du génome MGP192, **Figure 44**), suggérant qu'un morceau de génome *B. subtilis* de taille environ égale à 300 kpb subsiste dans ces clones. Aussi, nous avons pu remarquer que pour le clone 3.31, plus aucune bande n'est

Tableau 23 : Essai n°2 de clonage du génome complet de *B. subtilis* par la méthode de fusion

| Condition | Levure | Volume Sphéroplastés | <i>B. subtilis</i> | Volume Protoplastés | UFC | Criblage des transformants | | |
|-----------|--------|----------------------|-----------------------------------|---------------------|-----|----------------------------|--------------|----------------|
| | | | | | | Analysés | PCR simplexe | PCR multiplexe |
| 1 | W303a | 200 µL | - | - | 0 | - | - | - |
| 2 | | - | MGP192Δ <i>sbo-alb</i> ::ACH-Spec | 50 µL | 0 | - | - | - |
| 3 | | 200 µL | | 50 µL | 33 | 20 | 20/20 | 0/20 |
| 4 | | 200 µL | 100 µL | 43 | 20 | 20/20 | 0/20 | |
| 5 | | - | MSΔ <i>sbo-alb</i> ::ACH-Spec | 50µL | 0 | - | - | - |
| 6 | | 200 µL | | 50 µL | 0 | - | - | - |
| 7 | | 200 µL | | 100 µL | 0 | - | - | - |

présente (si ce n'est un léger amplicon aspécifique également présent sur le témoin ADN-levure), signifiant que l'érosion du génome a été encore plus forte dans ces clones).

Afin d'être surs de ce que nous observions, nous avons réitéré l'expérience mais cette fois en tentant de fusionner les protoplastes de *B. subtilis* avec des sphéroplastes de la souche W303a. La première souche étant décrite comme plus recombinante que cette dernière, nous nous sommes demandés dans quelle mesure le génome pourrait être plus stable dans cette nouvelle souche.

II.3.3. Second essai de clonage par Fusion *B. subtilis*-Levure

Avec ce second essai, nous avons fusionné la souche de levure W303a, avec des protoplastes de *B. subtilis* provenant des souches MGP192 Δ *sbo-alb*::ACH-Spec et MS Δ *sbo-alb*::ACH-Spec. Nous avons ajouté cette seconde souche afin d'observer dans quelle mesure son génome à 3,8 Mb pouvait être transféré par fusion. Dans le but de simplifier l'expérience, nous avons fixé le volume de sphéroplastes à 200 μ L, puisque c'est avec ce dernier que les plus grands nombres de transformants avaient été obtenus. Nous avons également fixé les volumes de protoplastes à 50 μ L et 100 μ L, et la concentration en MgCl₂ du SMM à 20 mM (**Tableau 23**).

Les contrôles « levures seules » et « protoplastes de *B. subtilis* seuls », ont permis de valider l'expérience car aucune colonie n'est apparue sur ces boîtes (conditions 1 et 2). Sur les « boîtes essais », 76 colonies ont été dénombrées au total pour la souche MGP192, et aucune colonie pour la souche MS. Plusieurs raisons peuvent expliquer ce résultat : (i) le génome de cette dernière est plus grand et ne peut pas être transféré et/ou maintenu dans la levure ; (ii) le génome est transféré mais est maintenu de façon très courte du fait de l'expression de gènes toxiques dans les 1,34 Mb de séquence génomique supplémentaires ; ou (iii) des métabolites toxiques absents chez la souche MGP192 sont présents dans le cytoplasme de la MS et déversés dans les cellules de levure.

Avec ces quelques colonies, nous validons à nouveau le fait qu'il était déjà possible d'obtenir des « transformants » sur boîte avec un génome de *B. subtilis* dans lequel les éléments génétiques ARS-CEN-HIS-SpecR avait déjà été insérés. Les analyses présentées **Figure 43** révélaient qu'après sous-clonage par strie, puis extraction à partir de patch, nous

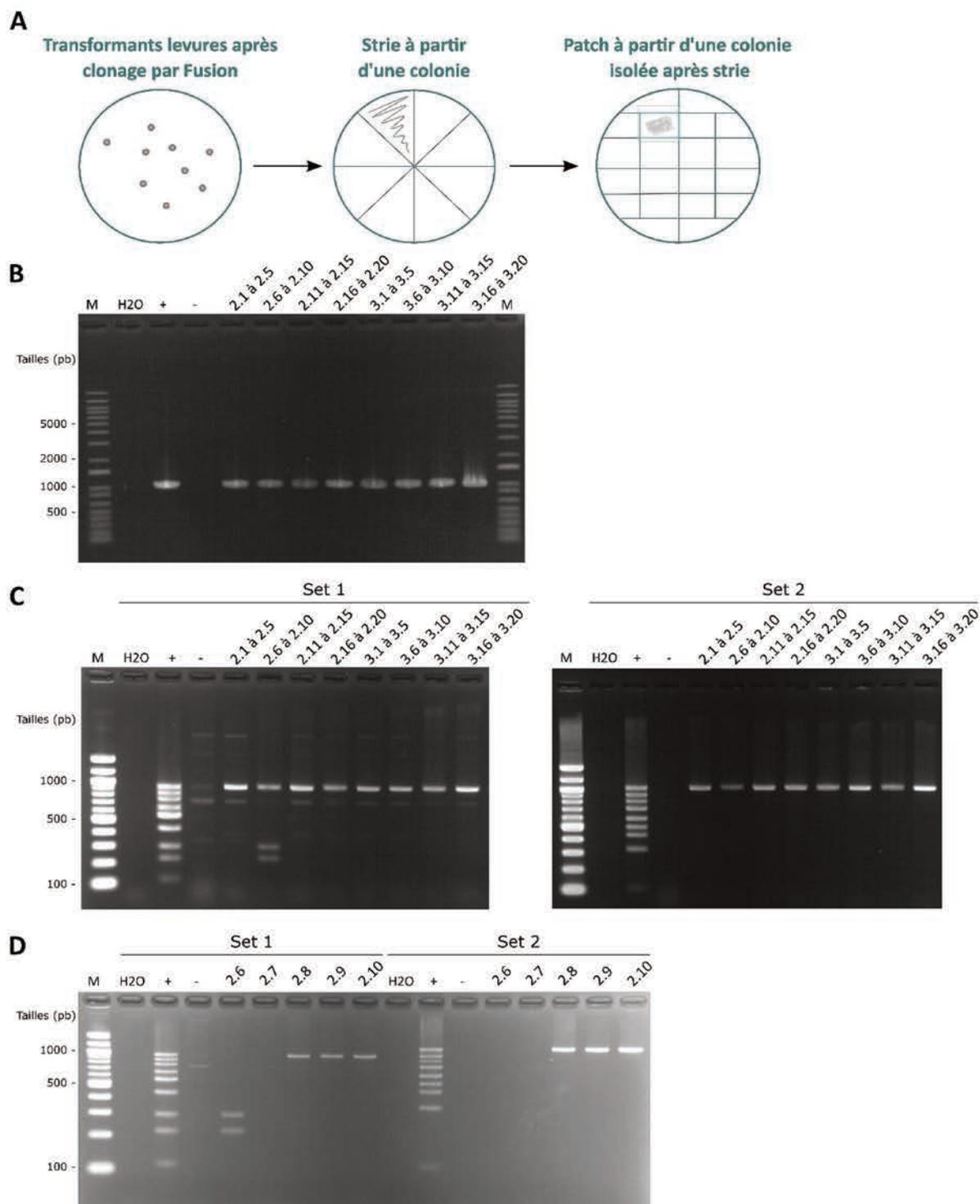


Figure 43 : Criblage des transformants levures après essai n°2 de clonage par Fusion du génome de *B. subtilis* MGP192. [A] Schéma de la procédure effectuée au cours de l'analyse des transformants: les colonies sont striées puis des patches sont effectués à partir des colonies isolées. **[B]** La présence du génome cloné dans la levure est détectée par PCR à l'aide d'amorces générant un amplifiat à 1020 pb (couple *Bsu20F/R*). **[C]** L'intégrité du génome cloné dans les clones levures est évaluée par PCR multiplexe sur des pools de clones grâce aux sets d'amorces 1 et 2 distribués sur l'ensemble du génome (106 à 1020 pb). **[D]** L'intégrité du génome cloné dans les clones levures est évaluée par PCR multiplexe sur des clones individuels grâce aux sets d'amorces 1 et 2 distribués sur l'ensemble du génome (106 à 1020 pb). "M": PCR simplexe = DNA Ladder 1kb+ Invitrogen (100-12000 pb) ; PCR multiplexe = DNA Ladder 100bp NEB (100-1517pb). "H2O" : témoin négatif sans ADN. "+" : témoin positif avec l'ADNg de *B. subtilis* MGP192. "-" : témoin négatif avec l'ADNg de *S. cerevisiae* W303a.

détections bien par PCR (en utilisant le couple d'amorces Bsu20F/R), le génome de la souche MGP192 dans les transformants analysés (analyse en pool, **Figure 43A et 43B**). En revanche, l'analyse par PCR multiplexe sur ces mêmes pools de transformants révélait que le génome cloné n'était pas maintenu de manière stable (**Figure 43C**). Seuls les couples d'amorces 9F/R et 20F/R les plus proches des éléments levures (**Figure 44**) ont permis d'obtenir un amplicon PCR. Ceci corrobore les résultats obtenus au cours de l'essai 1. En revanche, le pool de transformants cl2.6 à 2.10 analysés, montrait un profil un peu différent des autres. Nous avons donc extrait individuellement chaque clone de ce pool et répété les analyses avec les deux sets d'amorces de PCR multiplexe (**Figure 43D**). Il est apparu de manière assez intéressante pour le cl2.6, que les couples 2F/R et 3F/R donnaient un signal positif. Ceci signifiait qu'une autre portion de génome était maintenue et pourtant bien éloignée des éléments ARS-CEN-HIS-SpecR situés à ~ 2,7 Mb. Il s'avère que ces deux couples encadrent une ACS étendue de 17 nucléotides (5'-WWWTTTAYRTTTWGTT-3')²⁸² (**Figure 44**), laissant supposer que cette partie du génome pourrait être stable compte tenu de la présence de cette ACS décrite comme « parfaite ».

Discussion - Conclusion

Avant ces travaux, nous avions peu de données pour nous aiguiller sur la possibilité de cloner un génome bactérien tel que celui de *B. subtilis*. Nous avons donc choisi d'appliquer trois méthodes différentes pour tenter de cloner ce génome. Aucune de nos tentatives n'a permis de démontrer que nous étions capables de cloner et de maintenir stablement le chromosome de *B. subtilis* dans la levure. Cependant, de nouvelles informations ont été récoltées et des hypothèses ont pu être émises pour nous aider à poursuivre nos investigations.

En effet, avec la méthode de CReasPy-Cloning (Figure 1) testée en premier, nous avons obtenu des transformants sur boîte mais aucun ne propageait le génome de *B. subtilis*. Il s'est avéré que ce que nous observions était seulement des faux positifs induits par l'apport connexe du vecteur de recombinaison portant les éléments levures. Ce phénomène est courant dans nos expériences de clonage au laboratoire avec des génomes de mycoplasmes. Si le contrôle associé est un indicateur du nombre de faux positifs présents, il demeure

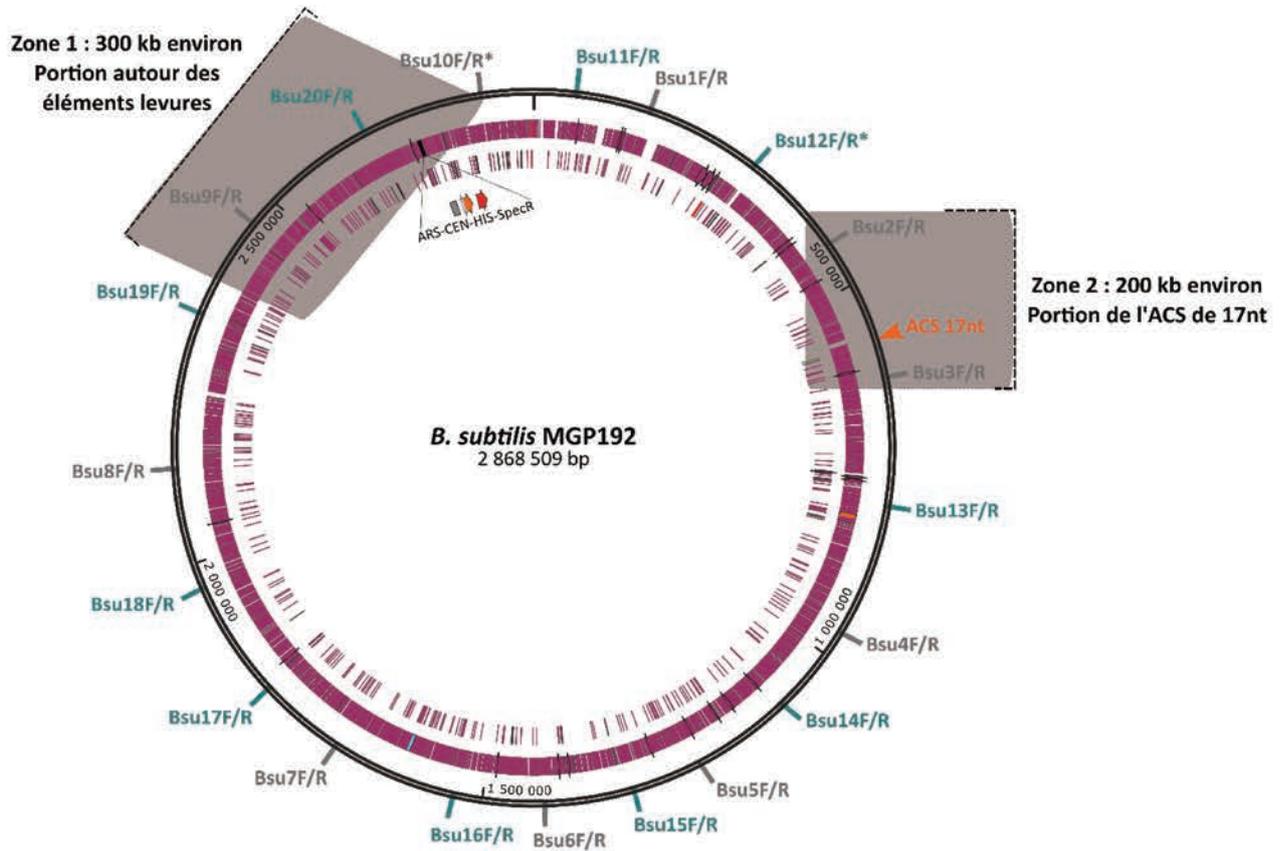


Figure 44 : Répartition des sets d'amorces de PCR multiplexe sur la carte génomique de *B. subtilis*. Le set d'amorce n°1 (en gris) est composé des couples d'amorces Bsu1F/R à Bsu9F/R. Le set d'amorce n°2 en bleu est composé des couples d'amorces Bsu11F/R à 20F/R. Les couples arborant une (*) ne fonctionnent pas en multiplexe et sont utilisés individuellement. Les zones 1 et 2 (gris et pointillés noirs) sont des appréciations des loci conservés par la levure après transformation (méthode de clonage par Fusion).

pourtant possible de détecter la molécule d'ADN d'intérêt (fragment ou génome) clonée dans une proportion raisonnable de transformants. Ce phénomène peut être dû, à la présence du plasmide matrice utilisé pour produire par PCR le pansement de recombinaison ; à l'intégration des éléments génétiques levure dans le chromosome levure ou bien ; à la circularisation de ce pansement au niveau des bras de recombinaison. Ce dernier point a été démontré au laboratoire et également décrit dans la littérature²³², et peut-être limité grâce à une contre sélection avec le marqueur *URA3* associé à l'ajout de 5-FOA (acide 5-fluoroorotique). Le design du vecteur levure est tel, que si ce dernier se circularise, il permet alors l'expression du gène *URA3* réassocié au promoteur *pADH1*. Cette conformation rend alors les transformants sensibles au 5-FOA, et permet la sélection des clones avec la bonne construction. Cette méthode de contre sélection aurait pu être envisagée pour limiter dans nos expériences la présence de faux positifs. Cependant, il nous semblait que d'autres facteurs pouvaient davantage bloquer le succès du clonage. A savoir, que la préparation des chromosomes en blocs n'était pas de qualité suffisante, que la taille du génome allait au-delà des limites biologiques de cet hôte, ou que la méthode de clonage n'était tout simplement pas adaptée.

Nous avons donc entrepris de poursuivre avec deux autres méthodes qui requièrent communément l'intégration des éléments levures ARS-CEN-HIS3 dans le génome de *B. subtilis*. Il est apparu que nous étions incapables d'obtenir des colonies sur boîtes avec la méthode classique, et ceci par deux fois. Tandis qu'avec la méthode de Fusion nous étions capables d'obtenir des transformants après clonage. Un aspect plutôt technique lié à l'extraction des génomes en bloc d'agarose et à l'incapacité pour ces génomes peu condensés de « traverser » la membrane plasmique (barrière physique) semblait alors se confirmer. Ces deux méthodes avaient pourtant permis le clonage, sous forme de YCP, de génomes dont la taille atteignait 1,8 Mb, à savoir celui de *Spiroplasma citri*²³⁰ (méthode classique), et celui d'*Haemophilus influenzae*²⁸³ (méthode par Fusion). Avec nos données, nous étions peut-être face à une limite adjointe à la taille du génome de la souche MGP192 de 2,86 Mb.

Finalement, les essais de clonage par la méthode de Fusion, ont non seulement permis d'obtenir des colonies sur boîte après transformation, mais aussi de détecter des portions du génome MGP192 clonées dans la levure. Ces résultats montraient pour la première fois qu'il était possible d'appliquer un tel protocole pour une bactérie dont la paroi est Gram+. Ils

Tableau 24 : Exemples de génomes bactériens clonés ou non clonés dans la levure

| Organismes | Tailles génomes (Mpb) | % G+C | Code génétique | ACS | Clonage dans la levure | Références |
|---|-----------------------|-------|----------------|---------|------------------------|---------------------------------|
| <i>Mycoplasma pneumoniae</i> M129 | 0,81 | 41 | Non standard | 19 | ✓ | Benders <i>et al</i> , 2010 |
| <i>Mycoplasma mycoides</i> subsp. <i>capri</i> GM12 | 1,07 | 24 | Non standard | 301 + 1 | ✓ | Lartigue <i>et al</i> , 2009 |
| <i>Acholeplasma laidlawi</i> PG8 | 1,49 | 32 | Universel | 175 + 2 | ✓ | Karas <i>et al</i> , 2012 |
| <i>Spiroplasma citri</i> GII-3 | 1,84 | 26 | Non standard | 330 + 5 | ✓ | Labrousseau <i>et al</i> , 2016 |
| <i>Bacillus subtilis</i> MGP192 | 2,86 | 44 | Universel | 144 + 1 | ✗ | - |
| <i>Haemophilus influenzae</i> | 1,83 | 38 | Universel | 148 + 1 | ✓ | Karas <i>et al</i> , 2013 |
| <i>Escherichia coli</i> (reduced genome) | 1,03 | 51 | Universel | NK | ✓ | Zhou <i>et al</i> , 2016 |
| <i>Caulobacter ethensis</i> 2.0 (<i>C. crescentus</i>) | 0,78 | 57 | Universel | 0 | ✓ | Venetz <i>et al</i> , 2019 |
| (<i>E. coli</i> génome réduit : 10 ARS ajoutées ; <i>C. ethensis</i> 2.0 : 5 marqueurs d'auxotrophie et 10 ARS ajoutées) | | | | | | |
| <i>Prochlorococcus marinus</i> | 1,65 | 31 | Universel | 328 + 4 | ✓ | Tagwerker <i>et al</i> , 2012 |
| <i>Synechococcus elongatus</i> | 2,69 | 55 | Universel | 7 | ✗ | Noskov <i>et al</i> , 2012 |
| (<i>S. elongatus</i> : seuls des fragments ~150kb ont été clonés) | | | | | | |

Mollicutes

Proteob.

Cyanob.

semblaient également confirmer que nous étions capables de transférer le génome de *B. subtilis*, depuis des protoplastes vers des sphéroplastes levures. Les analyses par PCR simplexe et multiplexe ont cependant montré que seuls des morceaux de génomes étaient maintenus et que le génome était donc rapidement dégradé (**Figures 42, 43 et 44**). Des études de stabilité effectuées avec les génomes de *Mycoplasma hominis*²³⁴ et de *Mycoplasma pneumoniae* (Thèse Dr. Estelle Ruiz) montrent qu'à partir de 10 passages, la levure élimine progressivement le chromosome surnuméraire. Les résultats obtenus avec le génome de *B. subtilis* montraient qu'au bout d'à peine 2 ou 3 passages, la quasi-totalité du génome était éliminée. Cela posait un problème d'autant plus important que dans le cadre du projet, les châssis bactériens synthétiques assemblés dans la levure se doivent d'être suffisamment stables pour être modifiés à souhait.

Pour conclure, différentes hypothèses ont été émises afin d'expliquer pourquoi nous n'étions pas parvenus à cloner le génome entier réduit de *B. subtilis* MGP192.

- (i) Des hypothèses dites « biologiques », intrinsèques aux caractéristiques du génome et à la biologie de *B. subtilis* : la taille du génome, le GC% combiné à la présence d'ACS, la toxicité d'un *locus* en lien avec le code génétique universel, des événements de recombinaison dus à des séquences répétées.
- (ii) Des hypothèses dites « techniques », liées directement aux conditions expérimentales et aux méthodes de clonage : la taille du génome à nouveau, la qualité des préparations de blocs, et l'état de condensation de l'ADN dans ces préparations.

Par comparaison avec les données de la littérature (**Tableau 24**), et d'après les résultats présentés au cours de ce chapitre, la taille du génome ainsi que la proportion de séquences ACS (en lien avec le GC%) semblaient être des facteurs qui combinés, influençaient la stabilité du génome dans la levure. Pour rappel, la taille du génome réduit de *B. subtilis* à 2,86 Mb reste largement supérieure à celle des plus grands génomes bactériens clonés dans la levure (*H. influenzae* 1,8 Mb et *S. citri* 1,8 Mb). De plus, la proportion en ACS sur le génome MGP192 nous semblait suffisamment importante (144 ACS motif 11nt + 1 ACS motif 17nt), mais cela ne signifiait pas que ces dernières soient fonctionnelles et permettent de maintenir de façon stable le génome dans la levure. D'autant que lorsque l'on compare avec d'autres génomes clonés comme celui de *Mycoplasma mycoides* subsp. *capri*²¹⁵ (301 ACS motif 11nt + 1 ACS

motif 17nt) ou bien de, *Prochlorococcus marinus*²²⁴ (328 + 4 ACS), le nombre d'ACS est deux fois plus élevées pour des tailles de génomes inférieures (1,07 Mb pour *M. mycoides* subsp. *capri* et 1,65 Mb pour *P. marinus*). Néanmoins, ces chiffres sont à ramener au prorata par rapport au pourcentage en GC, plus ce dernier augmente, plus le nombre d'ACS diminue car ce sont des séquences riches en A et en T. Il n'est donc pas exclu qu'il faille rajouter des ARS fonctionnelles dans le génome de *B. subtilis*, à l'instar de ce qui a été fait pour les génomes réduits et synthétiques d'*Escherichia coli* (1,08 Mb)²²⁵ et de *Caulobacter ethensis* 2.0 (0,78 Mb)²⁷⁰.

Nous ne pouvons pas non plus exclure qu'un ou plusieurs *locus* de *B. subtilis* pouvant s'exprimer chez *S. cerevisiae*, nuisent au clonage. Si les résultats décrits au cours de ce chapitre ne permettaient pas de conclure quant à cette possibilité, nous avons souhaité investiguer plus particulièrement la question avec des expériences complémentaires. Ainsi, la méthode de Transformation-Associated Recombination ou TAR-Cloning, a été mise en place dans le but de capturer individuellement des fragments de génome de *B. subtilis* et d'évaluer ainsi leur stabilité ou toxicité. Cette méthode nous permettait non seulement de répondre à nos hypothèses « biologiques » mais également, à celles « techniques » puisqu'elle servirait à évaluer la qualité des préparations de chromosomes en bloc d'agarose. Les expérimentations effectuées, ainsi que les données recueillies sont détaillées Chapitre 2.

CHAPITRE 2 – Clonage de fragments de génomme de *B. subtilis* par la méthode de TAR-Cloning

CHAPITRE 2 – Clonage de fragments de génome de *B. subtilis* par la méthode de TAR-Cloning

Introduction

Dans le Chapitre 1, nous avons décrit nos tentatives de clonage du génome entier de *B. subtilis* MGP192 (2,86 Mb) dans la levure, en utilisant des méthodes déjà décrites dans la littérature. Mais malgré nos efforts, aucune des trois méthodes testées n'a donné les résultats escomptés : (i) en utilisant la méthode Classique, aucune colonie n'est apparue sur milieu de sélection après transformation, (ii) avec la méthode CReasPy-Cloning, des colonies sont apparues sur milieu de sélection mais le génome de *B. subtilis* MGP192 n'a pas été détecté *in fine* dans les transformants levures, (iii) avec la méthode de clonage par Fusion, des colonies sont apparues, le génome de *B. subtilis* MGP192 a été détecté dans les transformants levures mais les données expérimentales semblent indiquer que ce dernier n'y est pas maintenu dans le temps.

De nombreuses raisons peuvent expliquer ces échecs répétés. Nous les avons classées en deux catégories : des raisons d'ordre « technique » (protocoles mal adaptés, paramétrages mal définis) et des raisons d'ordre « biologique » intrinsèques à la biologie de l'organisme que nous manipulons et aux caractéristiques de son génome. Parmi elles, nous pouvons citer :

- La taille du génome de *B. subtilis* MGP192 à 2,86 Mb ; qui reste largement supérieure à celle des plus grands génomes bactériens clonés dans la levure (*e.g. Haemophilus influenzae* 1,8 Mb et *Spiroplasma citri* 1,8 Mb)^{223,230}.
- Le pourcentage en GC du génome à 44% et le nombre d'ACS (ARS Consensus Sequence) présents dans le génome. Nous en avons comptabilisé 144 au total (*cf.* Chapitre 1 **Tableau 24**), mais cela ne signifie pas que ces dernières soient fonctionnelles et permettent de maintenir de façon stable le génome dans la levure²⁸⁴.
- La présence de *loci* dont l'expression serait toxique pour la levure ; *B. subtilis* MGP192 utilisant le code génétique universel tout comme *S. cerevisiae*.
- Et enfin, la présence d'éléments répétés qui pourraient favoriser les événements de recombinaison et la perte de certaines régions du génome bactérien.

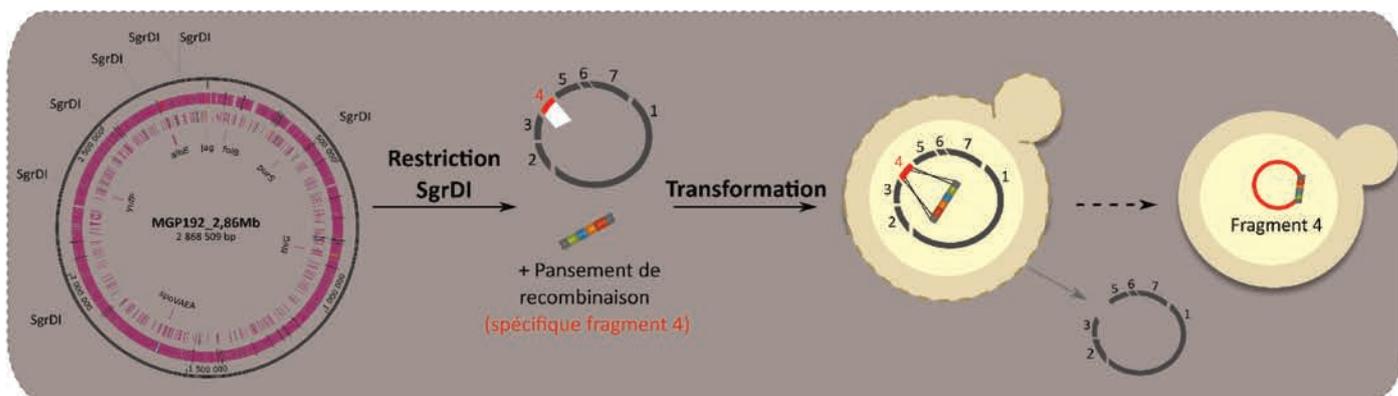


Figure 45 : Clonage du génome de *B. subtilis* par TAR-Cloning. Le génome extrait en bloc d'agarose est digéré par l'enzyme de restriction SgrDI. Les levures sont co-transformées avec le génome extrait en bloc d'agarose digéré en morceaux, et le pansement de recombinaison portant les éléments levures (ARS-CEN-HIS3). Dans cet exemple, le Fragment 4 est spécifiquement capturé grâce à la matrice de recombinaison qui possède à ses extrémités des séquences identiques à celles encadrant ce fragment.

Tableau 25 : Enzymes de restriction et nombre de sites respectifs présents sur le génome de *B. subtilis* MGP192.

| Enzyme de restriction et site de digestion | Nombre de sites | Taille des fragments |
|--|-----------------|--|
| SgrDI 5'-CG/TCGACG-3' | 7 | De 6 kb à 1500 kb |
| I-CeuI 5'-TAACTATAACGGTCCTAA/GGTAGCGA-3' | 10 | De 5 kb à 1600 kb 3 fragments à 5kb 2 fragments à 60 kb |
| AbsI 5'-CC/TCGAGG-3' | 11 | De 14 kb à 659 kb 2 fragments à 14 et 19 kb 2 fragments 657 et 659 kb |
| MauBI 5'-CG/CGCGCG-3' | 14 | De 2kb à 1063 kb 3 fragments à 2 kb, 3 kb et 4kb |
| FseI 5'-GGCCGG/CC-3' | 17 | De 0,168 kb à 514 kb 4 fragments à 42 kb à 63 kb 3 fragments de 100 à 117 kb |
| SfiI 5'-GGCCNNNN/NGGCC-3' | 21 | De 3 kb à 519 kb |
| AscI 5'-GG/CGCGCC-3' | 25 | De 3 kb à 343 kb |
| AsiSI 5'-GCGAT/CGC-3' | 32 | De 2 kb à 327 kb |
| Sbfl 5'-CCTGCA/GG-3' | 36 | De 5 kb à 542 kb |

Afin de mieux appréhender les raisons pour lesquelles nous n'arrivions pas à cloner le génome de *B. subtilis* MGP192 entier dans la levure, nous avons décidé d'en cloner des fragments en utilisant la méthode de TAR-Cloning²⁸⁵ (Transformation-Associated Recombination). Cette approche a notamment permis, dans le cas de l'espèce bactérienne *Acholeplasma laidlawii*, de localiser un gène codant une endonucléase toxique pour la levure et responsable à elle seule des échecs de clonage du génome entier chez cet hôte²²¹.

La méthode de TAR-Cloning permet de cloner des fragments d'ADN de tailles très diverses (quelques kpb à 1,6 Mb)²⁸⁵. Elle se base principalement sur la capacité des levures à incorporer des fragments d'ADN puis à les recombiner entre eux dès lors qu'ils possèdent des extrémités identiques. En pratique, le génome d'intérêt isolé en bloc d'agarose, est d'abord hydrolysé par des enzymes de restriction. Il est ensuite co-transformé dans la levure avec un segment d'ADN de réparation spécifique (appelé « vecteur TAR » ou « pansement ») possédant à ses extrémités des séquences de 60 pb (« bras d'homologie ») identiques aux séquences bordant le fragment cible. Après transformation, ce fragment cible linéaire est « réparé » par le système de recombinaison homologue de la levure en utilisant comme matrice le vecteur TAR fourni. Ce dernier porte les éléments ARS, CEN et HIS3 nécessaires à la réplication et au maintien de la molécule nouvellement circularisée (**Figure 45**).

Cette méthode nous paraissait tout-à-fait appropriée pour capturer individuellement des fragments de génomes de *B. subtilis* MGP192 de différentes tailles, mais aussi, pour évaluer leur stabilité au cours de leur propagation dans les levures. Elle nous paraissait également être une méthode de choix pour explorer la taille maximale qu'il est possible de cloner de manière stable dans la levure et d'étudier la répartition des ACS putatives sur le chromosome. Nous étions d'autant plus confiants qu'en 1993 Azevedo *et al.*²⁰⁸ avaient réussi à produire une banque de levures propageant des segments linéaires du génome de *B. subtilis* (de 60 kpb à 190 kpb) sous forme de chromosomes artificiels de levure (YAC) et couvrant environ 98% du génome original (4,2 Mb).

Au cours de ce chapitre, nous présentons les diverses expériences de TAR-Cloning effectuées, ainsi que les analyses associées.

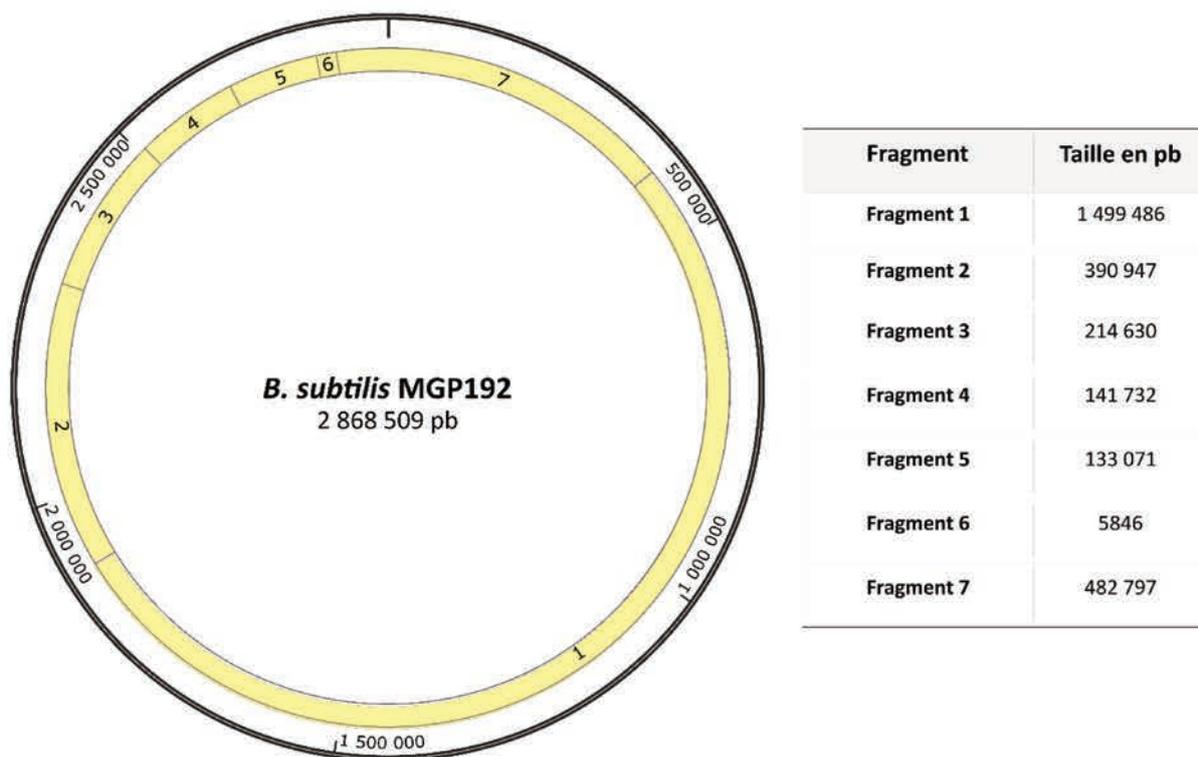


Figure 46 : Carte du génome de *B. subtilis* souche MGP192 avec les différents fragments issus de la digestion par l'enzyme de restriction SgrDI. En jaune sont représentés les fragments 1 à 7 dont les tailles sont répertoriées dans le tableau associé à la carte.

Tableau 26 : Résultats de l'expérience de TAR-Cloning après restriction du génome par l'enzyme SgrDI.

| Fragment de génome de <i>B. subtilis</i> MGP 192 | Taille en kpb | Colonies (UFC) | Clones positifs / Clones analysés ^a | | |
|--|---------------|----------------|--|----------------------|-------------------|
| | | | PCR simplexe | Autres PCR simplexes | PFGE ^b |
| Fragment 1 | 1499 | 45 | 0/45 | - | - |
| Fragment 2 | 390 | 37 | 7/37 | 7/7 | 1/7 |
| Fragment 3 | 214 | 54 | 26/54 | 26/26 | 3/3 |
| Fragment 4 | 141 | 46 | 25/46 | 25/25 | 4/4 |
| Fragment 5 | 133 | 23 | 13/23 | 13/13 | 3/3 |
| Fragment 6 | 5,8 | 37 | 15/37 ^c | - | - |
| Fragment 7 | 482 | 20 | 0/20 | - | - |

^aRésultats issus d'une seule expérience de clonage et obtenus pour chaque fragment. Le nombre de transformants levures analysés par PCR, PCR multiplexe et PFGE est indiqué, ainsi que le nombre de clones positifs obtenus.

^b Les analyses par PFGE n'ont pas été effectuées sur tous les clones positifs en PCR multiplexe, seulement sur un échantillon.

^c Les clones propageant le « Fragment 6 » ont été analysés avec une PCR amplifiant la totalité de sa séquence.

I. Essai de clonage de fragments de génome de *B. subtilis*

Pour mettre en place les premières expériences de TAR-Cloning, il nous fallait au préalable lister les enzymes de restriction capables de digérer le génome de *B. subtilis* MGP192, et identifier celles qui dans cette liste pourraient étayer nos hypothèses.

Nous avons donc commencé par réaliser une étude *in silico* des enzymes susceptibles d'hydrolyser le génome en plusieurs fragments en tenant compte des critères suivants :

- (a) Nombre de sites de restriction présents sur le génome et nombre de fragments produits après digestion
- (b) Distribution des sites de restriction autour du génome et taille des fragments après digestion
- (c) Position du site de restriction (région intergénique, région codante, région promotrice *cf. Tableau S5 – Annexes*)

Neuf enzymes ont été présélectionnées (**Tableau 25**) et notre choix s'est porté, dans un premier temps, sur l'enzyme de restriction SgrDI (site de restriction : CG/TCGACG) qui permet d'hydrolyser le génome en sept fragments allant de 6 kpb à 1500 kpb (**Figure 46**). Il s'agit de l'enzyme de restriction possédant le nombre de sites de coupure le plus faible. Étonnement, aucune autre enzyme ne permet de scinder le génome en un nombre de fragments inférieur à sept. Nous avons donc examiné les tailles de fragments générés par SgrDI, qui semblaient assez bien convenir à notre première étude pour les raisons suivantes :

- La gamme des tailles de fragments était large et il était raisonnable de penser, qu'à minima, certains petits fragments puissent être clonés dans la levure.
- Les fragments 4 (141 kpb) et 5 (133 kpb) possédaient des tailles proches des fragments clonés sous forme de YAC par Azevedo *et al.* en 1993. En parvenant à cloner ces fragments, nous approcherions alors des résultats générés par ce groupe.
- Les fragments 3 (214 kpb), 2 (390 kpb) et 7 (482 kpb) avaient des tailles supérieures à celles des fragments clonés dans l'étude antérieure. Nous pouvions alors franchir une étape, en clonant des fragments plus grands que ce que les auteurs avaient pu décrire. De plus, le fragment 7 (le plus grand) portait des séquences particulières comme l'origine bactérienne de réplication (*ori*) et la majorité des opérons ribosomiques (7 copies sur les 10 présentes sur l'ensemble du génome).

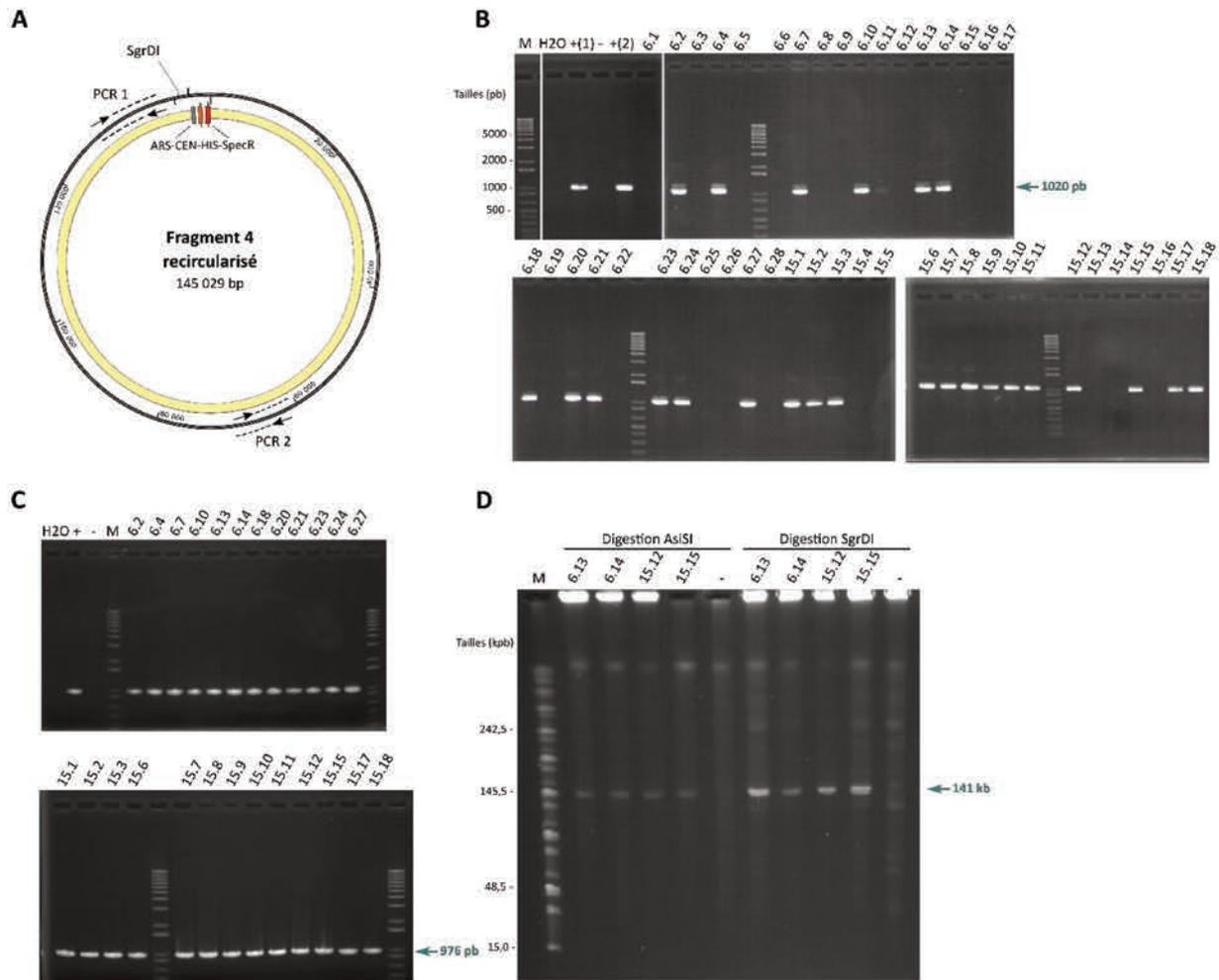


Figure 47 : Résultats du criblage des transformants levures obtenus après l'expérience de TAR-Cloning pour le Fragment 4. Cette figure regroupe les résultats de PCR simplexe et PFGE des analyses des clones propageant le « Fragment 4 ». **[A]** Le fragment est détecté à l'aide d'une PCR simplexe permettant d'amplifier un amplicon de 1020 pb à l'instar du témoin positif. **[B]** Une seconde PCR simplexe permet la validation des transformants en affichant un amplicon de 976 pb. **[C]** Les quatre clones sélectionnés pour l'analyse par PFGE arborent un profil avec une bande à 141 kb après linéarisation avec les enzymes AsiSI ou SgrDI. "M" : PCR simplexe = DNA Ladder 1kb+ Invitrogen (100-12000 pb) ; PCR multiplexe : DNA Ladder 100bp NEB (100-1517pb) ; PFGE = *S. cerevisiae* chromosomal DNA ladder Bio-Rad (225-2200 kb). "H2O" : témoin négatif sans ADN. "+(1)" : témoin positif avec l'ADNg de *B. subtilis* MGP192. "+(2)" : témoin positif avec l'ADNg de *B. subtilis* MGP192 mélangé à celui de la levure *S. cerevisiae* VL6-48. "-" : témoin négatif avec l'ADNg de *S. cerevisiae* VL6-48.

- Le fragment 1 (1500 kpb) avait une taille avoisinant celle de plusieurs génomes bactériens déjà clonés dans la levure (*A. laidlawii* 1,5 Mb, *S. citri* 1,8 Mb, et *H. influenzae* 1,8 Mb). Nous espérions donc raisonnablement parvenir à cloner ce fragment, et potentiellement valider que la moitié du génome de *B. subtilis* MGP192 était clonable dans la levure.

Au cours de la première expérience, des sphéroplastes de *S. cerevisiae* souche VL6-48 ont été co-transformés avec (i) des génomes de *B. subtilis* MGP192 isolés en blocs d'agarose (issus de la série 2 cf. Chapitre 1) digérés par l'enzyme SgrDI et, (ii) avec un « vecteur TAR » spécifique du fragment à capturer, qui portait entre les deux « bras d'homologie » des éléments levures ARS-CEN-HIS3 ainsi que le marqueur de sélection *aadA* conférant la résistance à la spectinomycine (SpecR).

Au total, sept essais ont été réalisés en parallèle, un pour chacun des sept fragments. Après transformation, des colonies sont apparues sur milieu de sélection pour toutes les conditions. Les nombres de colonies obtenues, entre 20 et 54 suivant les conditions (**Tableau 26**), ne semblaient être corrélés ni à la quantité d'ADN introduit, ni à la taille des fragments à capturer. Les transformants ont été repiqués afin d'être analysés par PCR simplexe, puis par PFGE. La **Figure 47** illustre les résultats du criblage effectué pour l'analyse des clones levures après capture du « Fragment 4 ». Sur les 46 colonies obtenues et analysées, 25 ont été validées en PCR simplexe en utilisant deux couples d'amorces spécifiques du « Fragment 4 » (**Figure 47A**). La PCR n°1 (**Figure 47B**) a permis de détecter un premier amplicon de 1020 pb identique au témoin positif (ADNg de *B. subtilis*) pour 25 clones sur 46 analysés. La totalité de ces clones a ensuite été validée avec une seconde PCR amplifiant une séquence située à l'opposé de la première sur le fragment circularisé. Une bande à 976 pb est obtenue pour tous les clones (**Figure 47C**), suggérant que ces transformants hébergeaient potentiellement le « Fragment 4 » entier. Suite à ces PCRs, quatre clones ont été sélectionnés (cl. 6.13, 6.14, 5.12 et 5.15) puis analysés par PFGE afin de s'assurer que la molécule clonée avait une taille cohérente avec celle attendue, à savoir 141 732 pb. Le résultat est présenté **Figure 47D**. Après digestion par des enzymes de restriction à site unique (AsiSI ou SgrDI), une seule bande de taille avoisinant 141 kpb fût visible sur gel pour chacun des quatre clones, confirmant la présence du « Fragment 4 » entier dans ces clones.

Les six autres fragments (1, 2, 3, 5, 6 et 7) ont été analysés en suivant la même procédure (données non montrées). Au total 5 fragments entiers sur 7 ont pu être clonés dans la levure (**Tableau 26**). Si l'on additionne l'ensemble de ces fragments [fragments 2+3+4+5+6], ce sont presque 900 kpb du génome *B. subtilis* MGP192 qui ont été clonés sans difficulté particulière dans la levure (soit ~1/3 du génome). Parmi ces derniers, le fragment le plus grand (Fragment 2) avait une taille d'environ 390 kpb ; nous avons donc franchi une étape (en terme de taille) par comparaison au travail réalisé par Azevedo *et al.*.

En revanche, deux fragments ont posé problème : le « Fragment 1 », qui est aussi le plus grand (1500 kpb) ; et le « Fragment 7 » (482 kpb). Pour ces deux fragments, aucune des PCRs effectuées n'a donné d'amplicon et donc aucune des colonies analysées n'a été validée. La présence de colonies sur ces boîtes peut notamment être expliquée, par une circularisation du vecteur TAR, phénomène déjà décrit dans des études connexes²³⁴ (*cf.* §1.4 – Chapitre 1) ou bien, hypothétiquement par intégration du vecteur TAR dans un chromosome levure.

Malgré des résultats mitigés, cette première expérience a permis de fournir des éléments qui ont pu nourrir et faire progresser notre réflexion. En effet, avant la mise en œuvre des expériences de TAR-Cloning, nous avons formulé plusieurs hypothèses pouvant expliquer notre incapacité à cloner le génome de *B. subtilis* MGP192 dans la levure (problèmes d'ordre biologique ou d'ordre technique). Cette expérience nous a permis de démontrer clairement qu'il était possible de cloner des morceaux de génomes *B. subtilis* MGP192 dans la levure et ceci en utilisant les blocs d'agarose spécifiquement mis au point pour nos expériences (*cf.* Chapitre 1). L'hypothèse selon laquelle la qualité des préparations de chromosomes de *B. subtilis* en bloc d'agarose n'était pas suffisante semblait désormais peu plausible. En revanche, les causes d'ordre biologique restaient toujours valides. En effet, tout comme pour le génome entier, l'impossibilité de cloner ces deux molécules pouvait être liée (i) à un manque d'ARS, et/ou, (ii) à l'expression délétère d'un gène ou groupement de gènes pour la levure ou encore, (iii) à la présence d'éléments répétés favorisant des événements de recombinaison.

La méthode de TAR-Cloning est régulièrement décrite comme peu efficace, que ce soit pour la capture de régions chromosomiques (0,5-2%)^{233,281} ou le clonage de génomes entiers (3%)²³⁴. Il est donc généralement nécessaire de cribler un grand nombre de clones. A l'issue, de cette expérience liminaire, nous avons donc décidé de persévérer et de tenter de cloner à

Tableau 27 : Expérience de clonage du « Fragment 7 » par TAR-Cloning avec des blocs issus du protocole « ED ».

| Condition | Pansement | Quantité Blocs | Traitement hexamine cobalt | Colonies (UFC) | Clones positifs / Clones analysés ^a | | |
|--------------|------------|----------------|----------------------------|----------------|--|----------------------|------|
| | | | | | PCR simplexe n°1 | Autres PCR simplexes | PFGE |
| 1 | | - | - | 0 | - | - | - |
| 2 | | - | - | >100 | - | - | - |
| 3 | | 2 mL (20µL) | | 375 | 2/20 | 0/2 | - |
| 4 | | 5 mL (20µL) | | 130 | 2/20 | 1/2 | 0/1 |
| 5 | | 10 mL (20µL) | - | 149 | 3/20 | 0/3 | - |
| 6 | Fragment 7 | 20 mL (20µL) | | 41 | 3/20 | 1/3 | 0/1 |
| 7 | | 2 mL (20µL) | | 386 | 2/20 | 0/2 | - |
| 8 | | 5 mL (20µL) | | 246 | 0/20 | 0/0 | - |
| 9 | | 10 mL (20µL) | + | 187 | 3/20 | 0/3 | - |
| 10 | | 20 mL (20µL) | | 83 | 2/20 | 0/2 | - |
| Total | | | | 1597 | 17/160 | 2/17 | 0/2 |

^aRésultats issus d'une expérience de clonage. Le nombre de transformants levures analysés par un ensemble de PCR simplexes et par PFGE est indiqué, ainsi que le nombre de clones positifs obtenu.

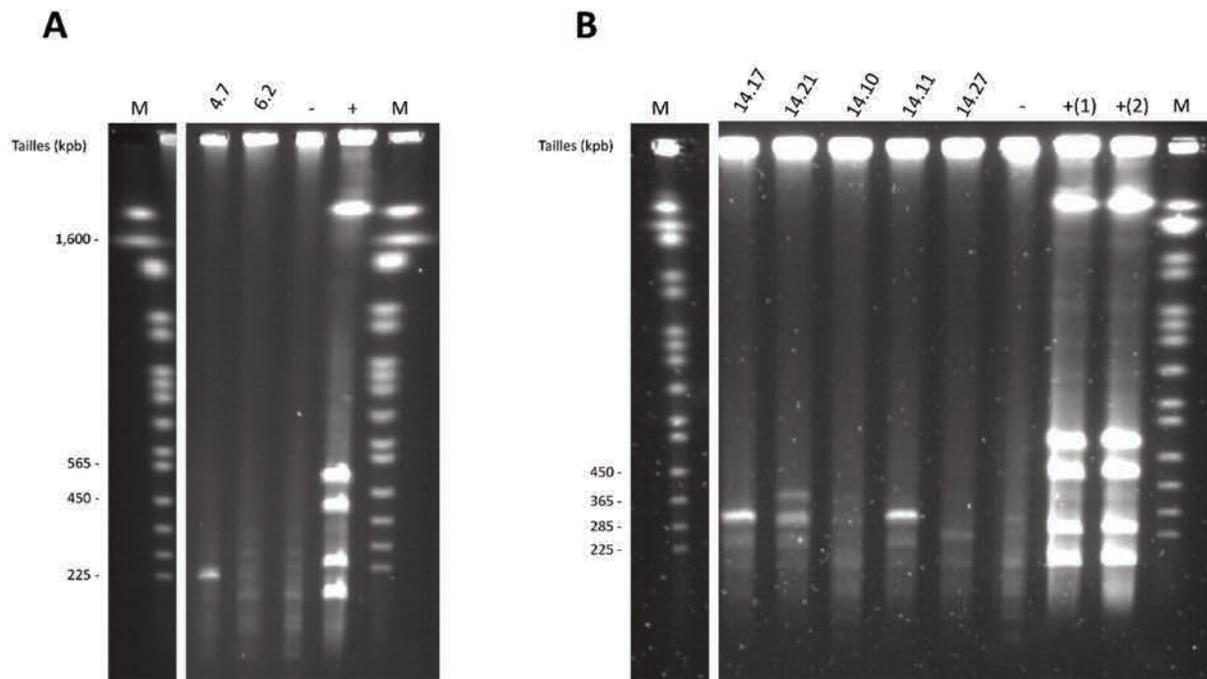


Figure 48 : Criblage par PFGE des transformants levures issus des expériences de TAR-Cloning pour la capture du « Fragment 7 ». [A] Résultats obtenus après l'analyse en PFGE de transformants issus de l'expérience avec les blocs du protocole « ED ». La taille attendue correspond à la bande à 482 kb observée pour le témoin positif. [B] Résultats obtenus après l'analyse en PFGE de transformants issus de l'expérience avec les blocs du protocole « Biorad ». La taille attendue correspond à la bande à 482 kb observée pour les témoins positifs. "M" : *S. cerevisiae* chromosomal DNA ladder Bio-Rad (225-2200 kb). "-" : témoin négatif avec les chromosomes de *S. cerevisiae* VL6-48 extraits en bloc d'agarose et digérés par l'enzyme SgrDI. "+" ou "+(1)" : témoin positif avec les chromosomes de *B. subtilis* MGP192 extraits en bloc d'agarose et digérés par l'enzyme SgrDI (quantité 5 mL). "+(2)" : témoin positif avec les chromosomes de *B. subtilis* MGP192 extraits en bloc d'agarose et digérés par l'enzyme SgrDI (quantité 10 mL).

nouveau les « Fragments 1 et 7 » entiers dans la levure. De plus, comme cela a été fait pour *A. laidlawii*, nous avons également entrepris de diviser ces deux régions du chromosome de *B. subtilis* en sous-fragments afin d'affiner notre analyse et d'identifier potentiellement un *locus* qui pourrait empêcher le clonage. Dans la suite de ce chapitre, les analyses *in silico* et les résultats associés seront présentés en deux temps : d'abord pour le « Fragment 7 », puis pour le « Fragment 1 ».

II. Clonage du « Fragment 7 » et analyses bio-informatiques

II.1. Nouvelles tentatives de clonage du « Fragment 7 »

Cette nouvelle expérience de TAR-Cloning, résumée dans le **Tableau 27**, a été réalisée en utilisant des préparations de chromosomes en blocs d'agarose effectuées avec le protocole « ED ». Ces blocs contenaient des quantités de cellules allant de 2 mL à 20 mL et, avaient ou non été traitées à l'hexamine cobalt (III) chlorure. Par comparaison avec les résultats obtenus au cours de l'expérience initiale pour ce « Fragment 7 » (*cf.* §II.1.1.), nous avons considérablement augmenté le nombre de transformants, puisque pas moins de 1597 colonies ont été dénombrées au total. Nous avons sélectionné 20 clones par condition et procédé aux analyses avec plusieurs PCR simplexes, puis par PFGE. Sur 160 colonies analysées, 17 clones ont été validés dans un premier temps, en utilisant un couple d'amorces spécifiques du « Fragment 7 » (Bsu_10F/R *cf.* Figure 44 Chapitre 1). Pour compléter ce premier criblage, deux autres couples d'amorces spécifiques de la séquence ont été utilisés (Bsu_11F/R et Bsu_1F/R), et seulement deux clones ont été identifiés comme propageant potentiellement le « Fragment 7 » entier. L'analyse par PFGE n'a pas confirmé ces résultats. En effet, une bande d'environ 200 kpb a été visualisée sur gel pour le cl4.7, au lieu des 482 kpb attendus et aucune bande n'était visible pour le cl6.2 (**Figure 48A**).

Nous avons réitéré cette expérience (**Tableau 28**), en utilisant cette fois la préparation de génome réalisée avec le protocole « Biorad ». Ces blocs contenaient des quantités d'ADN extraites à partir de volumes de cellules allant de 2 mL à 20 mL, mais avaient tous été traitée avec l'hexamine cobalt (III) chlorure. Étonnamment, la quantité de transformants obtenus sur boîte (89 UFC au total) était beaucoup plus faible que lors des expériences précédentes. Nous avons décidé d'analyser la totalité des colonies et sur 7 transformants testés positifs au

Tableau 28 : Expérience de clonage du « Fragment 7 » par TAR-Cloning avec des blocs issus du protocole « Biorad ».

| Condition | Pansement | Quantité Blocs | Traitement hexamine cobalt | Colonies (UFC) | Clones positifs / Clones analysés ^a | | |
|--------------|------------|----------------|----------------------------|----------------|--|--------------------|------|
| | | | | | PCR | Autres PCR simples | PFGE |
| 1 | | - | - | 0 | - | - | - |
| 2 | | - | - | >100 | - | - | - |
| 3 | Fragment 7 | 2 mL (20µL) | + | 11 | 0/10 | - | - |
| 4 | | 5 mL (20µL) | | 7 | 0/7 | - | - |
| 5 | | 10 mL (20µL) | | 41 | 2/41 | 0/2 | - |
| 6 | | 20 mL (20µL) | | 30 | 5/29 | 2/5 | 0/2 |
| Total | | | | 89 | 7/89 | 2/7 | 0/2 |

^aRésultats issus d'une expérience de clonage. Le nombre de transformants levures analysés par PCR, PCR multiplexe et PFGE est indiqué, ainsi que le nombre de clones positifs obtenu.

cours de la première PCR, seuls deux ont été validés avec les PCRs suivantes. L'analyse par PFGE a révélé que seules des portions du fragment étaient maintenues dans ces clones. Les cl14.17 et 14.11 semblaient héberger un fragment d'environ 250 kpb et le cl14.21 semblait héberger deux molécules, une aux alentours de 300 kpb et l'autre vers 250 kpb (**Figure 48B**). Cela pouvait s'expliquer par un mélange de population au sein de ce clone avec des cellules propageant le fragment à 300 kpb, et d'autres cellules le fragment à 250 kpb ; signe que le « Fragment 7 » était en cours de dégradation.

Finalement, tout comme pour l'expérience de clonage avec les blocs « ED », la totalité du « Fragment 7 » n'était pas conservée. Nous nous sommes alors demandés si une/des zone(s) de ce fragment étaient préférentiellement éliminée(s). Pour répondre à nos questionnements, nous avons mis en place de nouvelles expériences impliquant la division de ce « Fragment 7 » en sous-fragments.

II.2. Restriction du « Fragment 7 » en onze sous-fragments par les enzymes FseI et AsiSI

Pour cette étude, nous avons choisi d'effectuer l'expérience de TAR-Cloning avec deux conceptions différentes, permettant de capturer des sous-fragments se chevauchant partiellement ou totalement. Ainsi, nous avons travaillé avec les enzymes de restriction (i) FseI qui permettait de produire les sous-fragments 7.1 à 7.4 (en bleu), divisant le « Fragment 7 » en 4 morceaux de 44 à 322 kpb et (ii) AsiSI qui permettait de produire les sous-fragments 7.5 à 7.11 (en saumon), divisant le « Fragment 7 » en 7 morceaux, dont les tailles variaient de 6 à 252 kpb (**Figure 49A**). L'expérience a été effectuée en duplicat pour les 11 sous-fragments et a mis en jeu les deux types de préparations de génomes de *B. subtilis* « ED » et « Biorad ».

Les résultats des deux expériences sont regroupés et résumés dans le **Tableau 29**. Une centaine de colonies a été dénombrée pour l'ensemble des conditions, avec des nombres allant de 1 (Fragments 7.4 et 7.7) à 32 (Fragment 7.3). La quantité de colonies obtenues sur boîte n'était pas corrélée avec la taille du fragment capturé. En effet, très peu de colonies ont été obtenues pour de très grands sous-fragments comme le « 7.1 » à 264 kpb, ou bien le « 7.4 » à 322 kpb ; mais aussi pour le plus petit, le « 7.7 » à 6 kpb. Il était donc probable que des *loci* problématiques puissent se trouver préférentiellement sur ces derniers.

Tableau 29 : Essai de clonage par TAR-Cloning de « Sous-fragments » après restriction par les enzymes FseI et AsiSI.

| Fragment de génome de <i>B. subtilis</i> MGP 192 | Taille en kpb | Colonies (UFC) | Clones positifs / Clones analysés ^a | | |
|--|---------------|----------------|--|----------------------|-------------------|
| | | | PCR simplexe n°1 | Autres PCR simplexes | PFGE ^b |
| Fragment 7.1 | 264 | 2 | 0/2 | - | - |
| Fragment 7.2 | 44 | 9 | 4/9 | 4/4 | 3/4 |
| Fragment 7.3 | 63 | 32 | 13/26 | 13/13 | 4/4 |
| Fragment 7.4 | 322 | 1 | 0/1 | - | - |
| Fragment 7.5 | 252 | 4 | 2/4 | 2/2 | 2/2 |
| Fragment 7.6 | 128 | 14 | 9/14 | 9/9 | 4/4 |
| Fragment 7.7 | 6 | 1 | 1/1 | 1/1* | - |
| Fragment 7.8 | 10 | 11 | 5/11 | 5/5* | - |
| Fragment 7.9 | 74 | 10 | 5/11 | 5/5 | 3/3 |
| Fragment 7.10 | 20 | 14 | 6/14 | 6/6* | - |
| Fragment 7.11 | 141 | 18 | 6/18 | 6/6 | 4/4 |

^aL'expérience de clonage a été effectuée en dupliqué et les résultats obtenus sont additionnés. Le nombre de transformants levures analysés par PCR et PFGE est indiqué, ainsi que le nombre de clones positifs obtenu.

^bLes analyses par PFGE n'ont pas été effectuées sur tous les clone positifs en PCR mais seulement sur un échantillon.

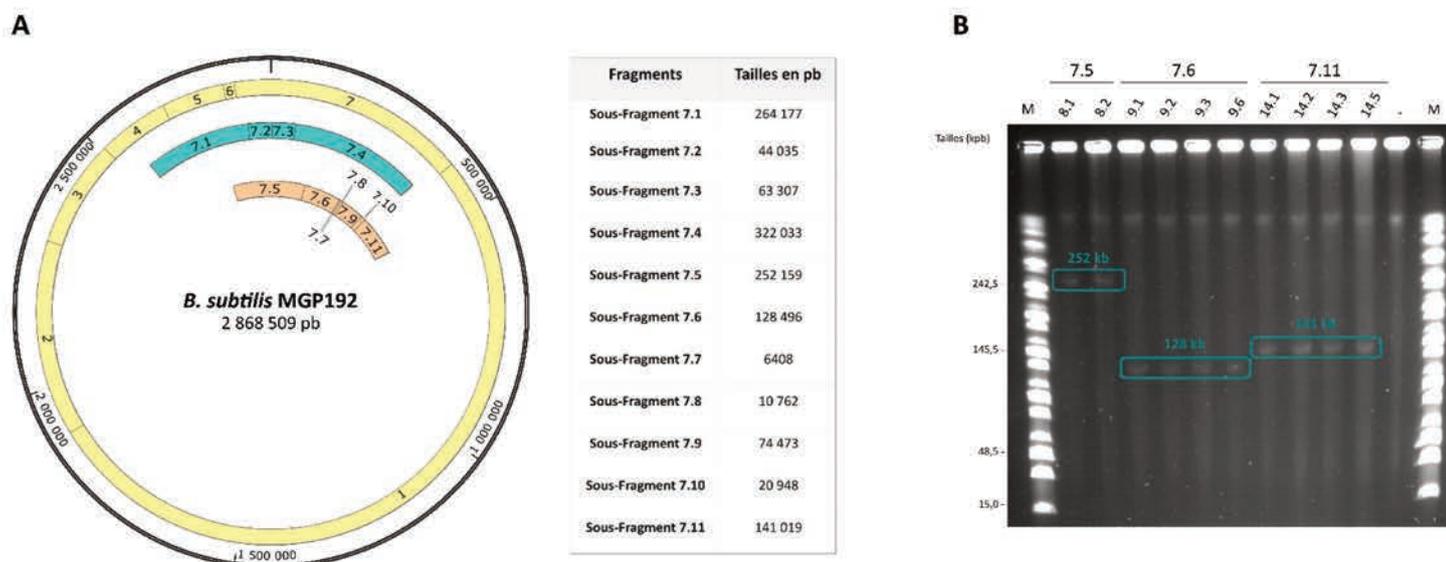


Figure 49 : Essai de clonage de « Sous-fragments » après restriction par les enzymes FseI et AsiSI. [A] Carte du génome de *B. subtilis* souche MGP192 avec les différents fragments issus de la restriction par l'enzyme SgrDI (en jaune), par l'enzyme FseI (en bleu), ou par l'enzyme AsiSI (en saumon). Les tailles des « Sous-Fragments 7.1 à 7.11 » sont répertoriées dans le tableau associé à la carte. **[B]** Exemple d'un résultat de gel obtenu après l'analyse en PFGE de transformants propageant les « Sous-fragment 7.5, 7.6 et 7.11 ». Les tailles attendues correspondent respectivement aux encadrés en bleu, indiquant 252 kb, 128 kb et 141 kb. "M" : *S. cerevisiae* chromosomal DNA ladder Bio-Rad (225-2200 kb). "-" : témoin négatif avec les chromosomes de *S. cerevisiae* VL6-48 extraits en bloc d'agarose et digérés par l'enzyme AsiSI.

La totalité des colonies obtenues a été analysée par de multiples PCRs et par PFGE (**Tableau 29**). À titre d'exemple, le gel de PFGE obtenu pour valider la taille des sous-fragments « 7.5 », « 7.6 » et « 7.11 » hébergés par les clones levures sélectionnés après PCRs est présenté **Figure 49B**. Pour chacun des clones analysés, une bande à la taille attendue a été observée, soit 252 kpb pour le « Sous-fragment 7.5 », 128 kpb pour le « Sous-fragment 7.6 » et 141 kpb pour le « Sous-fragment 7.11 ». Les transformants propageant les sous-fragments les plus petits, à savoir les sous-fragments « 7.7 » (~6 kpb), « 7.8 » (~10 kpb) et « 7.10 » (~20 kpb), n'ont pas été validés en PFGE ; des PCRs amplifiant des fragments contigus ayant été réalisées de façon à mettre en évidence la totalité de ces sous-fragments.

Finalement, sur les 11 sous-fragments, 9 ont pu être capturés et propagés de manière stable chez la levure. Les clones testés positifs lors de la première étape de PCR l'étaient également avec les PCRs suivantes et lors du PFGE le cas échéant. L'ensemble de ces données nous a permis de démontrer qu'il était possible de cloner la région chromosomique correspondant au « Fragment 7 », lorsque cette dernière est divisée en multiple sous-fragments. Cependant, les données recueillies nous interrogeaient fortement, puisque parmi les 11 sous-fragments, les deux sous-fragments « 7.1 » (264 kpb) et « 7.4 » (322 kpb) n'ont jamais pu être clonés. Cela est d'autant plus étonnant que ces sous-fragments recouvraient des régions présentes dans d'autres fragments facilement capturés (**Figure 49A**) ; le « Fragment 7.1 » par exemple (en bleu sur la carte) recouvre les « Fragments 5 » (133 kpb) et « 6 » (6 kpb) ainsi qu'une partie des fragments « 4 » (141 kpb) et « 7.5 » (252 kpb), tous clonés assez facilement dans la levure.

II.3. Bilan des expériences de TAR-Cloning sur le Fragment 7

Nous avons initialement posé différentes hypothèses pour expliquer notre incapacité à cloner le « Fragment 7 » dans la levure : manque d'ARS, expression délétère d'un ou plusieurs gènes, présence d'éléments répétés. Ce fragment a donc été découpé en sous-fragments dans le but de faire progresser notre réflexion.

Cette nouvelle série d'expériences n'a pas permis d'identifier de façon claire les raisons pour lesquelles le Fragment 7 n'est pas « clonable » en l'état dans la levure. Elle tend néanmoins à montrer que l'instabilité de cette région est vraisemblablement à relier (outre le

Tableau 30 : Comparaison des « gaps » observés dans l'article d'Azavedo *et al.*⁶ avec les résultats des expériences de TAR-Cloning.

| Gaps | Gap 1 (<9kb) | Gap 2 (<14kb) | Gap 3 (<36kb) | Gap 4 (<20kb) |
|--|---|---|---|---|
| Identification par Azavedo <i>et al.</i> 1993 | Système restriction-modification (RE) | Gène <i>dinR</i> impliqué dans la réponse SOS (dommage de l'ADN) | Recombinase site spécifique | Entre <i>gerB</i> et <i>spoIIID</i> |
| Locus correspondants | <i>ydiR</i> et <i>ydiS</i> ^{11,12} | <i>dinrR</i> ou <i>lexA</i> ¹³ | <i>spoIVCA</i> ¹⁴ | - |
| Localisations | vers 650kb (<i>Bsu</i> 168) vers 330kb (<i>Bsu</i> MGP192) | vers 1,9Mb (<i>Bsu</i> 168) vers 1,3Mb (<i>Bsu</i> MGP192) | vers 2,65Mb (<i>Bsu</i> 168) Absent (<i>Bsu</i> MGP192) | vers 3,7Mb (<i>Bsu</i> 168) Absent (<i>Bsu</i> MGP192) |
| Fragments | Fragment 7 et Sous-Fragment 7.9 | Fragment 1 et Sous-Fragment 1.3 | - | - |

Bsu 168 : *Bacillus subtilis* souche 168

Bsu MGP192 : *Bacillus subtilis* souche MGP192

manque possible d'ARS dans cette région) : (i) à l'expression ou co-expression létale de certains gènes (facteurs agissant en trans), ou (ii) à la présence même de certaines séquences (éléments répétés ou facteurs agissant en cis). En effet, certains sous-fragments de cette série (e.g. « 7.1 » et « 7.4 ») ont non seulement généré très peu de colonies sur milieu de sélection (une à deux au maximum), mais ces colonies se sont avérées être de faux positifs signifiant qu'à aucun moment de l'expérimentation, l'un ou l'autre des sous-fragments n'a pu être détecté dans les levures et ceci même à nombre de passage très faible. Il semblerait donc que dès l'instant où ces sous-fragments pénètrent dans les levures, les cellules meurent ou bien s'en débarrassent ; dans un cas comme dans l'autre, les levures ne peuvent plus croître en milieu de sélection, à moins qu'elles n'aient intégré au chromosome ou circularisé le vecteur TAR.

Nous avons alors comparé nos résultats avec ceux publiés par Azevedo *et al.* en 1993. La banque produite par ces auteurs est constituée de 59 inserts qui ont pu être ordonnés en 4 contigs couvrant ~98% du génome de *B. subtilis* souche 168, signifiant que seulement 4 portions de génomes (« gaps ») n'avaient pas été clonées dans la levure. Une hypothèse proposée par les auteurs pour expliquer leur résultat fait référence à la présence potentielle de gènes toxiques pour la levure. Nous avons donc cherché à savoir si ces *loci* pouvaient être assimilés aux régions qui nous faisaient défaut. Ces analyses sont répertoriées **Tableau 30** et montrent qu'au cours de nos expériences, nous avons été capables de cloner deux de ces régions (Gap 1 et Gap2). Concernant les deux autres régions (Gap3 et Gap4), elles ne sont plus présentes dans le génome de la souche MGP192. Ainsi, les gènes identifiés ne pouvaient pas être incriminés dans nos recherches. En revanche, un point important dans cette étude révélait que le « YAC 15-426 » était tronqué d'une partie de sa séquence. De manière intéressante, ce YAC s'inter-hybridait avec d'autres YAC sur lesquels étaient présents des opérons codant des ARN ribosomiques. Il se trouve justement que ce « YAC 15-426 » correspond à une partie du « Fragment 7 ». Après analyse, nous avons constaté que ces opérons sont nombreux sur le « Fragment 7 ». D'une part, ils constituent des zones fortement répétées sur une région restreinte du génome, et peuvent donc être le siège d'évènements de recombinaison homologue favorisant le remaniement de cette région dans la levure. D'autre part, on ne peut écarter l'hypothèse que l'expression des opérons ribosomiques

B. subtilis MGP192 chez la levure puisse être problématique pour son hôte, un phénomène déjà décrit dans la littérature pour d'autres espèces²⁸⁶.

En partant de ces constats, nous avons poursuivi nos analyses (i) à la recherche de gènes candidats potentiellement toxiques, et (ii) en portant une attention particulière sur les régions hautement répétées du « Fragment 7 ».

II.4. Analyse de gènes toxiques putatifs

Karas *et al.* en 2011 ont démontré que le gène codant l'endonucléase de surface ACL0117 devait être inactivé en amont du clonage du génome d'*A. laidlawii*, afin de prévenir son effet toxique pour la levure. Dans le but d'identifier ce problème, les auteurs ont scindé le génome en fragments puis en sous-fragments. Ils ont utilisé pour les régions les plus petites, de l'ADN de sources différentes : (i) de l'ADNg d'*A. laidlawii* digéré par des enzymes de restriction, (ii) de l'ADNg non digéré, et enfin (iii) des amplicons PCR. Étonnement, pour deux sous-fragments, ils notèrent que les régions étaient « clonables » dans la levure lorsque que ces dernières étaient issues d'une amplification par PCR, mais ne l'étaient plus lorsque l'ADNg était utilisé. Ils ont alors supposé qu'au cours de la réaction PCR, des mutations avaient été introduites dans des gènes potentiellement toxiques pour la levure. Après plusieurs vérifications (séquençage, clonage de la région chez *E. coli etc...*), l'hypothèse s'est confirmée.

Suivant la même idée, nous avons cherché à savoir si des gènes mutés étaient présents sur des fragments ou sous-fragments que nous avons réussi à cloner. Certains fragments étant grands, nous avons retenu (avec nos collaborateurs de l'équipe Syber) trois *loci* sur l'ensemble du « Fragment 7 » :

- (i) Les gènes *endB* et *ndoA* codant un système toxine-antitoxine MazE/MazF²⁸⁷. Ce *locus* est présent sur les « Sous-fragments 7.4 et 7.9 ».
- (ii) Le gène *trmE* (ou *thdF*) codant une GTPase impliquée dans la modification des ARN de transfert (ARNt)²⁸⁸. Ce gène est présent sur le « Sous-fragment 7.5 ».
- (iii) Le gène *ywaC* (ou *sasA*) codant une ppGpp synthétase²⁸⁹. Ce gène est présent sur les « Sous-fragments 7.1 et 7.5 ».

Le choix de ces candidats a essentiellement été motivé par le fait que ce type de gènes avait déjà montré des incompatibilités chez *E. coli*^{290,291}. Après amplification par PCR et analyse par séquençage Sanger, aucune mutation n'a été observée sur ces gènes, rendant notre hypothèse difficile à défendre. Cependant, des analyses plus poussées réalisées avec l'aide de nos collaborateurs sont actuellement en cours. Pour chaque fragment et sous-fragment, plusieurs clones levures ont été sélectionnés. L'ADN total de ces clones a été extrait et sera séquencé par une méthode de séquençage nouvelle génération. De plus, à l'instar de ce qui a été fait pour le génome de *Mesoplasma florum* cloné dans la levure²²², des analyses de transcriptomique ont été entamées. L'idée est de déterminer si des transcrits bactériens présents dans la levure peuvent conduire à des incompatibilités de clonage. Le cas échéant, la nature de ces transcrits sera déterminée ainsi que leur niveau d'expression.

II.5. Mise en évidence de régions identiques et répétées sur le Fragment 7

Une série de 45 PCRs différentes réparties sur l'ensemble du « Fragment 7 » a été effectuée sur les ADN_g des transformants levures cl4.7, 14.17, 14.11 et 14.21 (cf. §II.1). Les résultats bruts des 45 PCRs ne sont pas montrés ; en revanche, ils ont été extrapolés de façon à visualiser sur une carte les portions du « Fragment 7 » présentes ou absentes dans chaque clone (**Figure 50A**). Tous les profils ne sont pas identiques mais en superposant les cartes (**Figure 50B**), il apparaît de façon claire que certaines portions du fragment sont maintenues plus facilement que d'autres.

En effet, trois profils se distinguent (profil 4.7 ; profil 14.17 et profil 14.11/14.21) : tous incluent un *locus* de 200 kpb englobant le vecteur TAR portant les éléments levures ainsi qu'un *locus* de 33 kpb environ. Les clones 4.7 et 14.17 semblent héberger en plus de ces séquences des portions supplémentaires. Cela n'est pas totalement en accord avec les profils observés sur les gels d'électrophorèse en champ-pulsé (**Figure 48**). Pour rappel, le clone 4.7 semblait héberger un fragment de ≈225 kpb, les clones 14.11 et 14.17 un fragment de ≈285 kpb/300 kpb et le cl14.21 un fragment de ≈285 kpb/300 kpb et un second de ≈365 kpb. Cela peut provenir d'un mélange de populations dans les clones avec certaines cellules portant des fragments plus dégradés que d'autres. La PCR étant beaucoup plus sensible, elle

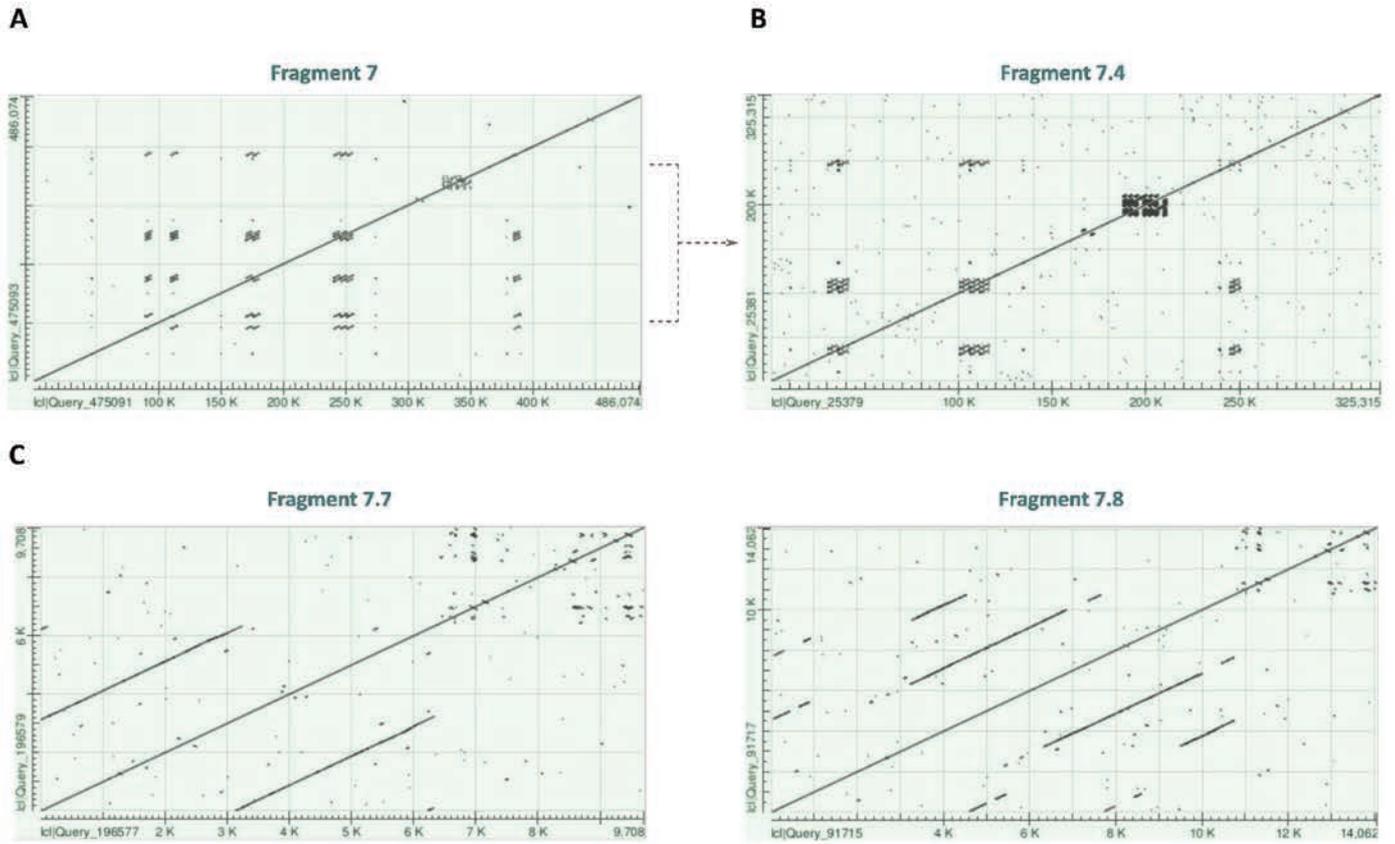


Figure 51 : Représentation en Dot-Plot de l'alignement des séquences du « Fragment 7 » et du « Sous-fragment 7.4 ». L'alignement a été effectué à l'aide de l'outil Blastn sur le portail NCBI (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). **[A]** Dot-Plot obtenu après auto-alignement du « Fragment 7 ». **[B]** Dot-Plot obtenu après auto-alignement du « Fragment 7.4 ». **[C]** Dot-Plot obtenu après auto-alignement des « Fragment 7.7 » et « Fragments 7.8 ».

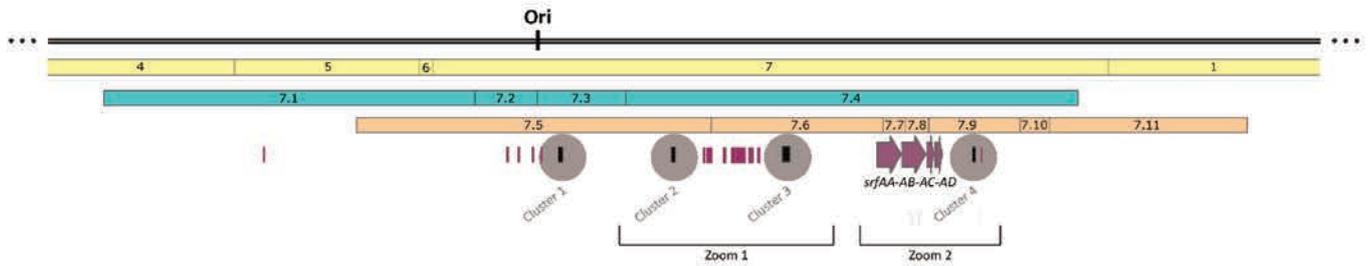
permet peut-être de détecter des restes de séquences en mélange avec les fragments majoritaires visibles en PFGE.

Malgré ces quelques incohérences, cette analyse met en lumière un résultat marquant : en regardant plus précisément la section manquante entre 180 kpb et 400 kpb nous nous sommes aperçus qu'elle correspondait au « Fragment 7.4 » que nous n'avons jamais réussi à cloner.

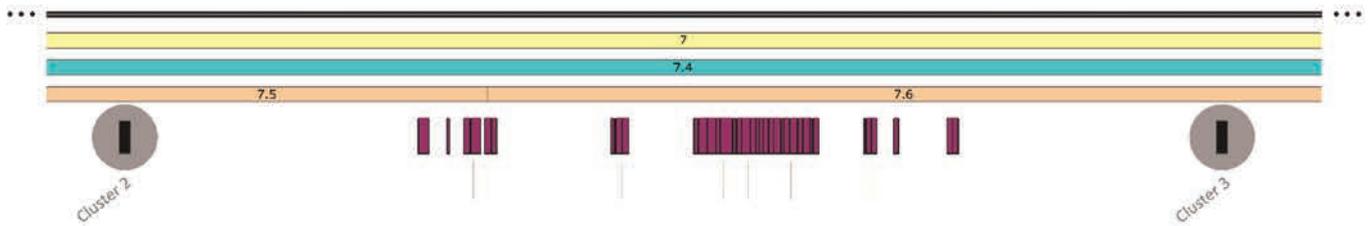
Afin de comprendre pourquoi cette région du fragment était préférentiellement éliminée, nous avons aligné les « Fragment 7 » et « Fragment 7.4 » sur eux-mêmes. Les résultats, représentés sous forme de Dot-Plot (**Figure 51**), indiquent que des répétitions en tandem sont présentes vers 200 kpb, 250 kpb et 350 kpb sur le « Fragment 7 » (**Figure 51A**). Lorsque l'on fait un zoom sur ce *locus* (**Figure 51B**), nous observons que ces répétitions sont aussi retrouvées sur le « Fragment 7.4 ». Or, la présence de ce genre de structure peut favoriser le remaniement des molécules d'ADN clonées. En effet, une étude menée sur le clonage et la stabilité du génome de *M. hominis* fait état que ce sont toujours les mêmes régions qui sont éliminées préférentiellement dans le temps²³⁴. Les auteurs précisent que des séquences répétées ou répétées en tandem sont la cause de ces observations.

Par comparaison et grâce à l'outil en ligne Reputer (<https://bibiserv2.cebitec.uni-bielefeld.de/reputer>), nous avons répertorié au total pas moins de 50 séquences répétées (taille supérieure à 400 pb) sur le « Fragment 7.4 ». Comme décrit par Azevedo *et al.* et souligné paragraphe II.3., ces dernières correspondent à des gènes codant des protéines ribosomiques et des ARN ribosomiques chez *B. subtilis*. Sept copies d'ADNr (*rrns*) sur les dix présentes sur l'ensemble du génome de *B. subtilis* MGP192 sont concentrées dans cette région chromosomique. Ces copies sont groupées en quatre clusters (**Figure 52A**)¹⁵. Le cluster 1 est inclus dans les « Sous-fragment 7.3 et 7.5 ». Les clusters 2, 3 et 4 sont tous retrouvés sur le « Fragment 7.4 » et par correspondance sur les « Sous-fragments 7.5, 7.6 et 7.9 ». Or, des travaux ont déjà fait état que les opérons d'ADNr pouvaient être à l'origine de réarrangements causés par des évènements de recombinaison intra-chromosomiques chez *B. subtilis*^{292,293}. Il est donc tout-à-fait possible que de tels évènements surviennent également chez *S. cerevisiae*, un organisme dans lequel la machinerie de recombinaison homologue est hautement performante. De plus, parmi les 40 gènes codant les sous-unités des protéines

A - Région chromosomique de *B. subtilis* MGP192 autour de l'oriC



B - Zoom 1



C - Zoom 2

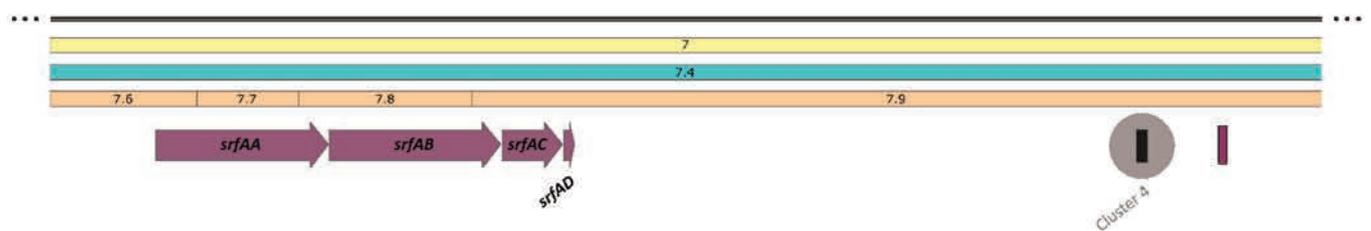


Figure 52 : Carte du locus proche de l'origine de réplication du chromosome de *B. subtilis*. Les fragments « 4 à 1 » (SgrDI) sont indiqués en jaune, les sous-fragments « 7.1 à 7.4 » (FseI) sont représentés en bleu tandis que les sous-fragments « 7.5 à 7.11 » sont en saumon. Les gènes codant pour les protéines ribosomiques sont représentés par des rectangles en violet foncé et les clusters regroupant les opérons ADNr sont représentés avec des rectangles noirs entourés d'un cercle gris. Les gènes inclus dans l'opéron *srfAA-srfAD* sont indiqués par des flèches en rose foncé. **[A]** Vue d'ensemble du locus. **[B]** Zoom n°1 représentant spécifiquement la région entre les clusters 2 et 3 sur les sous-fragments « 7.4, 7.5 et 7.6 ». **[C]** Zoom n°2 représentant spécifiquement l'opéron *srfAA-AD* et le cluster 4 sur les sous-fragments « 7.4, 7.6, 7.7, 7.8 et 7.9 ».

ribosomiques²⁹⁴, une grande majorité se localise à l'intersection des « Sous-fragments 7.5 et 7.6 » (**Figure 52B**). Chacun de ces sous-fragments avaient été clonés individuellement, en revanche une fois les gènes regroupés sur un même *locus* (à savoir, le « Fragment 7.4 »), il devenait impossible à cloner. La présence de ces nombreuses répétitions sur le « Fragment 7 » et le « Sous-fragment 7.4 » semble être un point clé pour expliquer l'instabilité de ces régions (remaniements). Néanmoins, on ne peut écarter l'hypothèse que l'expression de ces gènes dans la levure soit également problématique (*cf.* §II.4 Analyses de gènes toxiques putatifs). Les gènes codant les ARN ribosomiques et protéines ribosomiques, de par leur nombre et leur niveau d'expression, pourraient ainsi, être une double source d'instabilité.

Hormis ces gènes codant les ARN et les protéines ribosomiques, nous avons noté la présence sur le « Fragment 7.4 » d'un grand opéron (s'étalant sur ~25 kpb) impliqué dans la production de la surfactine synthétase²⁹⁵ (**Figure 52C**). Ce dernier est réparti sur plusieurs « Sous-fragments » (7.6, 7.7, 7.8 et 7.9), que nous avons réussi à cloner. Pour les « Sous-fragments 7.7 » et « Sous-fragments 7.8 », nous avons procédé à l'alignement des séquences de ces derniers sur elles-mêmes et obtenu des représentations en dot-plot (**Figure 51C**). Les données montrent que de larges répétitions sont présentes sur les gènes *srfAA* et *srfAB*. Ces répétitions pourraient expliquer que ces *loci* soient éliminés préférentiellement. Cette hypothèse est d'autant plus plausible que le « Sous-Fragment 7.7 » qui est assez petit (~7,7 kpb) a été difficilement obtenu avec seulement un transformant levure validé. Enfin, tout comme pour les potentiels gènes toxiques, et pour les gènes codant les ARNr et protéines ribosomiques, des analyses complémentaires sont à effectuer par séquençage NGS afin de vérifier si des mutations peuvent être à l'origine de ce succès et diminuer un éventuel effet toxique.

III. Clonage du « Fragment 1 » et analyses bio-informatiques

III.1. Nouvelle tentative de clonage du « Fragment 1 »

Tout comme pour le « Fragment 7 », des expériences ont été menées pour tenter de cloner à nouveau le « Fragment 1 » (1500 kpb) dans la levure et comprendre les causes de notre insuccès (*cf.* §I).

Tableau 31 : Essai de clonage du « Fragment 1 » par TAR-Cloning.

| Condition | Pansement | Série de blocs | Quantité Blocs | Traitement hexamine cobalt | Clones positifs / Clones analysés ^a | | | | |
|--------------|--|---------------------------|----------------|----------------------------|--|-------|----------------|------|-----|
| | | | | | Colonies (UFC) | PCR | PCR Multiplexe | PFGE | |
| 1 | Spécifique du Fragment 1 (bras homologie de "1kb") | - | - | - | 0 | - | - | - | |
| 2 | | - | - | - | 93 | - | - | - | |
| 3 | | MGP192 protocole "ED" | 2 mL (20µL) | - | - | 22 | 0/22 | - | - |
| 4 | | | 5 mL (20µL) | | | 63 | 0/63 | - | - |
| 5 | | | 10 mL (20µL) | | | 22 | 1/22 | 1/1 | 0/1 |
| 6 | | | 20 mL (20µL) | | | 36 | 0/36 | - | - |
| 7 | | | 2 mL (20µL) | | | 46 | 0/46 | - | - |
| 8 | | | 5 mL (20µL) | | | 10 | 0/10 | - | - |
| 9 | | MGP192 protocole "Biorad" | 10 mL (20µL) | + | + | 9 | 0/9 | - | - |
| 10 | | | 20 mL (20µL) | | | 15 | 0/15 | - | - |
| 11 | | | 2 mL (20µL) | | | 76 | 0/76 | - | - |
| 12 | | | 5 mL (20µL) | | | 38 | 0/38 | - | - |
| 13 | | | 10 mL (20µL) | | | 39 | 0/39 | - | - |
| 14 | | | 20 mL (20µL) | | | 36 | 0/36 | - | - |
| 15 | | - | 2 mL (20µL) | - | - | 20 | 0/20 | - | - |
| 16 | | | 5 mL (20µL) | | | 18 | 0/18 | - | - |
| 17 | | | 10 mL (20µL) | | | 16 | 0/16 | - | - |
| 18 | | | 20 mL (20µL) | | | 25 | 0/25 | - | - |
| Total | | | | | 491 | 1/491 | 1/1 | 0/1 | |

^aRésultats issus d'une expérience de clonage. Le nombre de transformants levures analysés par PCR, PCR multiplexe et PFGE est indiqué, ainsi que le nombre de clones positifs obtenu.

Au vu des difficultés rencontrées lors de la première expérience, des modifications ont été apportées pour mettre en place ce nouvel essai. Comme précédemment, les sphéroplastes de levure VL6-48 ont été co-transformés avec le génome de la souche MGP192 préparé en blocs d'agarose et digérés par SgrDI, et un vecteur TAR spécifique du « Fragment 1 ». Cependant pour cette expérience, la conception de ce pansement a été modifiée de telle façon que ses « bras de recombinaison » situés aux extrémités aient une taille égale à 1 kpb au lieu des 60 pb utilisés précédemment. En effet, au cours de travaux connexes réalisés au laboratoire, nous avons pu observer qu'en augmentant la taille des bras de recombinaison nous augmentions la probabilité de capturer la molécule d'ADN ciblée²²⁰. D'autre part, pour maximiser nos chances de succès, nous avons également décidé d'utiliser des blocs d'agarose (i) préparés à partir de différents protocoles (protocole « ED » vs « Biorad », cf Chapitre 1), (ii) renfermant différentes quantités de cellules et donc d'ADN (2, 5, 10 et 20 mL), et (iii) traitées ou non avec l'hexamine cobalt (III) chloride (agent compactant de l'ADN) (**Tableau 31**).

Après transformation, les résultats obtenus sur les boîtes contrôles « levures seules » (condition 1 : pas de colonies,) et « levures transformées avec le pansement » (condition 2 : 93 colonies) étaient conformes à ce que nous obtenons habituellement, et ont été validés. Concernant les « conditions 3 à 18 », un total de 491 UFC était comptabilisé sur l'ensemble des boîtes, avec un dénombrement minimal (9 colonies) pour la condition 9 et maximal (76 colonies) pour la condition 11, qui encore une fois, n'était pas corrélé avec la quantité d'ADN transformé.

La totalité des clones a été analysée et seulement un transformant, le clone 5.15, a été validé successivement par PCR simplexe et PCR multiplexe (**Figure 53B et 53C**). L'analyse avec le set de multiplexe n°1 a permis d'obtenir le profil attendu avec 6 bandes visibles sur gel avec des tailles comprises entre 216 pb à 736 pb. L'analyse effectuée avec le set n°2 a permis de compléter les résultats en montrant un profil à 5 bandes allant de 327 pb à 720 pb. Ceci révélait que non seulement nous étions capables de détecter la présence du « Fragment 1 » dans un clone levure, mais que ce dernier semblait intègre dans ce clone. Les deux sets de PCR multiplexe, répartis de manière homogène sur l'ensemble du génome de *B. subtilis* et donc, *a fortiori*, sur ce fragment, ayant permis d'obtenir l'ensemble des amplicons attendus (**Figure 53A**). Les conditions expérimentales testées au cours de cette expériences (protocoles ED ou Biorad, quantité de cellules dans les blocs, traitement avec un agent compactant), n'ont

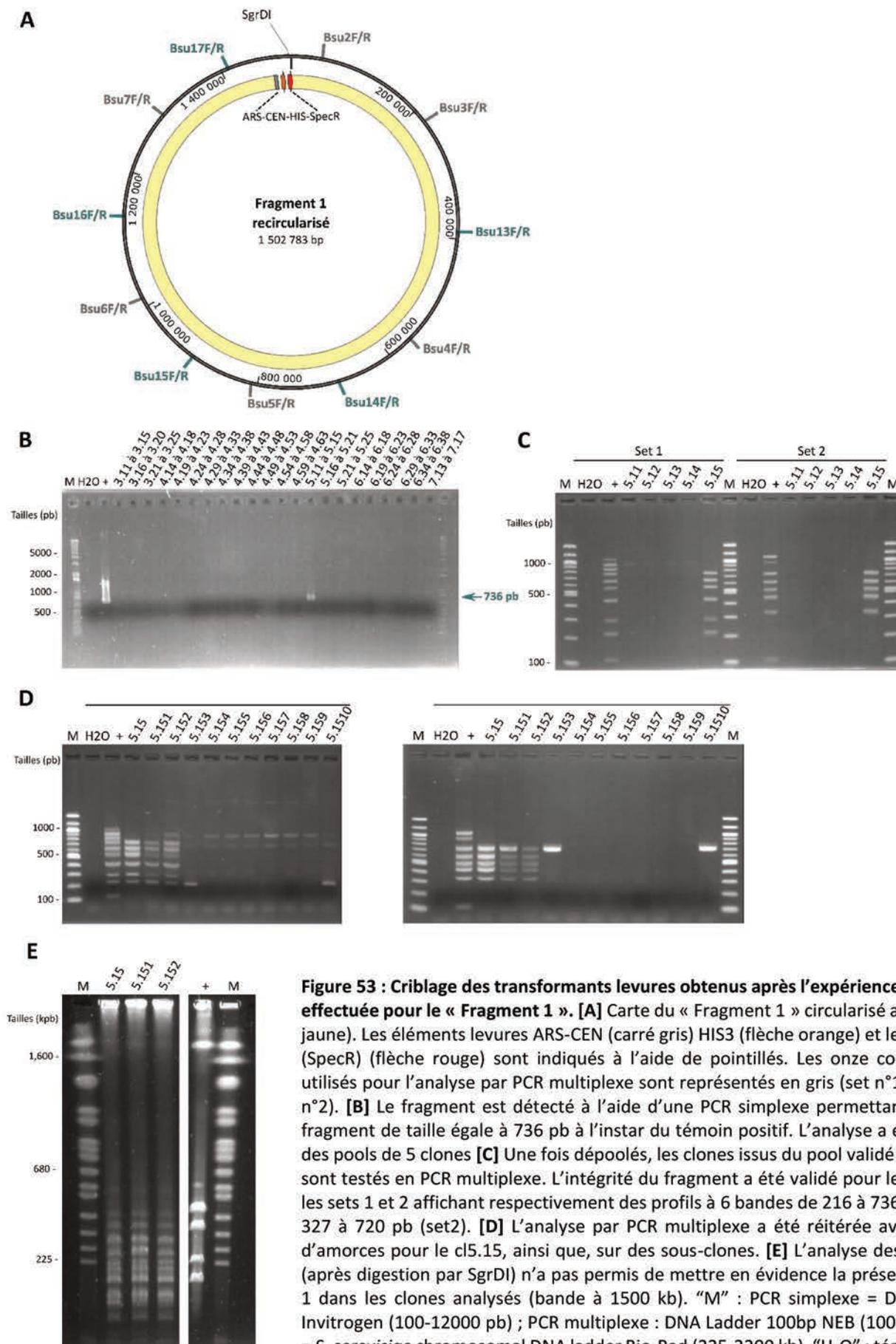


Figure 53 : Criblage des transformants levures obtenus après l'expérience de TAR-Cloning effectuée pour le « Fragment 1 ». [A] Carte du « Fragment 1 » circularisé après capture (en jaune). Les éléments levures ARS-CEN (carré gris) HIS3 (flèche orange) et le marqueur *aadA* (SpecR) (flèche rouge) sont indiqués à l'aide de pointillés. Les onze couples d'amorces utilisés pour l'analyse par PCR multiplexe sont représentés en gris (set n°1) et en bleu (set n°2). [B] Le fragment est détecté à l'aide d'une PCR simplex permettant d'amplifier un fragment de taille égale à 736 pb à l'instar du témoin positif. L'analyse a été effectuée sur des pools de 5 clones [C] Une fois dépoolés, les clones issus du pool validé en PCR simplex sont testés en PCR multiplexe. L'intégrité du fragment a été validé pour le clone 5.15 avec les sets 1 et 2 affichant respectivement des profils à 6 bandes de 216 à 736 pb (set 1), et de 327 à 720 pb (set2). [D] L'analyse par PCR multiplexe a été réitérée avec les deux sets d'amorces pour le cl5.15, ainsi que, sur des sous-clones. [E] L'analyse des clones en PFGE (après digestion par SgrDI) n'a pas permis de mettre en évidence la présence du fragment 1 dans les clones analysés (bande à 1500 kb). "M" : PCR simplex = DNA Ladder 1kb+ Invitrogen (100-12000 pb) ; PCR multiplexe : DNA Ladder 100bp NEB (100-1517 pb) ; PFGE = *S. cerevisiae* chromosomal DNA ladder Bio-Rad (225-2200 kb). "H₂O" : témoin négatif sans ADN. "+(1)" : témoin positif avec l'ADNg de *B. subtilis* MGP192. "+(2)" : témoin positif avec l'ADNg de *B. subtilis* MGP192 mélangé à celui de la levure *S. cerevisiae* VL6-48. "+" : témoin positif avec les chromosomes de *B. subtilis* extraits en bloc d'agarose et digérés par l'enzyme SgrDI. "-" : témoin négatif avec l'ADNg de *S. cerevisiae* VL6-48.

pas pu être comparées puisqu'un seul clone a été obtenu. Il est difficile d'attester que la condition idéale pour cloner ce fragment soit celle associée au clone 5.15 [ED-10mL-sans agent compactant]. Cependant, la conception du vecteur TAR avec des bras de recombinaison de 1 kpb a potentiellement aidé à obtenir ce résultat. Il aurait pu être intéressant de répéter l'expérience pour éventuellement conclure sur ce sujet.

Dans le but de vérifier la stabilité du « Fragment 1 », nous avons repiqué le cl5.15 et effectué de nouvelles analyses sur 10 sous-clones. Les résultats, présentés **Figure 53D**, indiquent de façon claire que le fragment est rapidement éliminé par la levure. En effet, sur 10 sous-clones, seuls les clones 5.151 et 5.152 montrent un profil complet avec toutefois des bandes moins intenses que pour le clone parental cl5.15. Des amplicons PCR correspondants aux couples Bsu_2F/R (216 pb, pour le set n°1) et Bsu_17F/R (720 pb, pour le set n°2) sont également visibles sur gel pour deux autres sous-clones, les cl5.153 et 5.1510. Il est intéressant de noter que ces deux couples d'amorces, situés aux bornes du « Fragment 1 », sont normalement éloignés l'un de l'autre. Ce n'est plus le cas après capture et circularisation du fragment dans la levure ; une configuration dans laquelle ces deux couples se retrouvent à proximité des « éléments levure » portés par le vecteur TAR (**Figure 53A**). Il semblait donc que le fragment n'était pas maintenu entier dans ces clones et que seules les séquences proches du vecteur TAR étaient préservées.

Le clone 5.15 et les sous-clones 5.151 et 5.152 ont été analysés par PFGE dans le but d'évaluer plus précisément la taille des fragments hébergés par ces clones. Le profil obtenu sur gel après coloration au Sybr Gold (**Figure 53E**), était constitué seulement de bandes aspécifiques (reliques de digestion de l'ADNg total de la levure) suggérant fortement que le « Fragment 1 » était dégradé ou perdu au cours de la propagation des levures.

L'ensemble de ces données montrent que le « Fragment 1 » à 1500 kpb, peut être capturé dans la levure tout comme les fragments 2, 3, 4, 5 et 6 (expérience de TAR-Cloning initiale du paragraphe I), mais n'est pas stable au cours du temps. En effet, après plusieurs passages *in vivo*, le fragment n'est plus détecté. L'instabilité de ce fragment peut être due à un manque d'ARS fonctionnelles ou bien à la présence d'un gène toxique pour la levure. Afin d'affiner notre analyse, nous avons entrepris de découper le « Fragment 1 » en sous fragments.

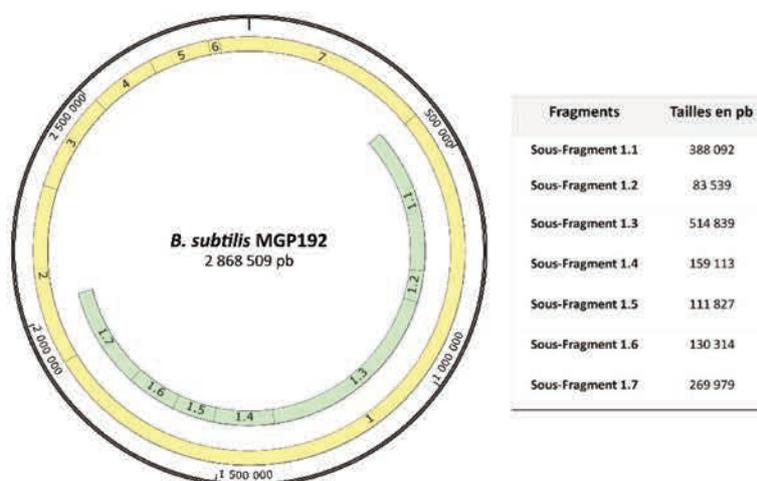
Tableau 32 : Essai de clonage par TAR-Cloning de « Sous-fragments » après digestion par l'enzyme de restriction FseI

| Fragments de génome de <i>B. subtilis</i> MGP 192 | Taille en kpb | Colonies (UFC) | Clones positifs / Clones analysés ^a | | |
|---|---------------|----------------|--|----------------------|-------------------|
| | | | PCR simplexe n°1 | Autres PCR simplexes | PFGE ^b |
| Fragment 1.1 | 388 | 27 | 13/26 | 13/13 | 4/4 |
| Fragment 1.2 | 83 | 105 | 29/40 | 29/29 | 3/3 |
| Fragment 1.3 | 514 | 15 | 1/14 | 1/1 | 1/1 |
| Fragment 1.4 | 159 | 63 | 34/40 | 34/34 | 4/4 |
| Fragment 1.5 | 111 | 67 | 32/40 | 31/32 | 3/4 |
| Fragment 1.6 | 130 | 136 | 28/40 | 28/28 | 4/4 |
| Fragment 1.7 | 269 | 87 | 19/40 | 19/19 | 4/4 |

^aRésultats issus d'une expérience de clonage. Le nombre de transformants levures analysés par un ensemble de PCR simplexes et par PFGE est indiqué, ainsi que le nombre de clones positifs obtenu.

^b Les analyses par PFGE n'ont pas été effectuées sur tous les clones positifs en PCR mais seulement sur un échantillon.

A



B

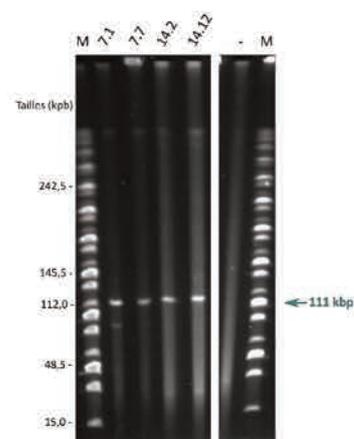


Figure 54 : Essai de clonage de « Sous-fragments » après restriction par l'enzyme FseI. [A] Carte du génome de *B. subtilis* souche MGP192 avec les différents fragments issus de la digestion par l'enzyme de restriction SgrDI (en jaune), ou par l'enzyme de restriction FseI (en vert). Les tailles des « Sous-Fragments 1.1 à 1.7 » sont répertoriées dans le tableau associé à la carte. [B] Exemple d'un résultat de gel obtenu après l'analyse en PFGE de transformants propageant le « Sous-fragment 1.5 ». La taille attendue correspond à la bande observée à 111 kb. "M" : *S. cerevisiae* chromosomal DNA ladder Bio-Rad (225-2200 kb). "-" : témoin négatif avec les chromosomes de *S. cerevisiae* VL6-48 extraits en bloc d'agarose et digérés par l'enzyme FseI.

III.2. Restriction du « Fragment 1 » en sept sous-fragments par l'enzyme FseI

Pour cette dernière expérience de TAR-Cloning, nous avons choisi d'hydrolyser le « Fragment 1 » en sept sous-fragments, appelés « Sous-fragments 1.1 à 1.7 », grâce à l'enzyme de restriction FseI, et de les capturer individuellement dans la levure à l'aide de nouveaux vecteurs TAR (**Figure 54A**). Les résultats obtenus sont rassemblés **Tableau 32**.

Après transformation, 500 UFC ont été comptabilisés sur l'ensemble des conditions. Pour les plus grands sous fragments, le « Fragment 1.1 » à 388 kpb et le « Fragment 1.3 » à 514 kpb, les nombres de colonies étaient les plus faibles : 27 et 15 UFC respectivement. Mais, ces deux conditions mises à part, le nombre de colonies obtenues sur milieu de sélection n'était pas corrélé avec la taille du sous fragment à capturer.

Quarante colonies ont été repiquées par condition (excepté pour les fragments « 1.1 » - 27 colonies, et « 1.3 » - 15 colonies, pour lesquels nous avons repiqué la totalité des clones) afin d'être analysées par un ensemble de PCR simplexe, puis par PFGE. Au cours de ce criblage, plusieurs points ont retenu notre attention. Tout d'abord, nous avons pu noter que tous les clones testés positifs après la première PCR simplexe, étaient également positifs avec les PCRs complémentaires (deuxième étape du criblage) ainsi qu'en PFGE (troisième étape du criblage) ; un seul sous-fragment dérogea à cette règle (le « Sous-fragment 1.5 »), pour lequel 31 transformants sur 32 examinés furent positifs au second criblage par PCR et, 3 sur 4 furent positifs en PFGE (**Figure 54B**). Ensuite, nous avons pu noter que pour le plus grand des sous-fragments, le « Sous-fragment 1.3 » à 514 kpb, un seul transformant était positif sur les 14 analysés ; soit un taux de succès (nombre de clones positifs/nombre de clones analysés) environ égal à 7%, largement inférieur aux taux de succès des 6 autres sous-fragments (estimé entre 47% et 85%). Cela se justifie certainement par la taille de ce fragment et possiblement par un rapport [nombre d'ARS fonctionnelle/taille de la séquence] proche du seuil toléré par la levure. Cependant, de notre point de vue, un taux de succès égal à 7% reste suffisamment élevé et plus qu'acceptable pour une expérience de TAR-Cloning (dont les taux de succès décrits dans la littérature avoisinent plutôt 0,5-2%)^{233,281}.

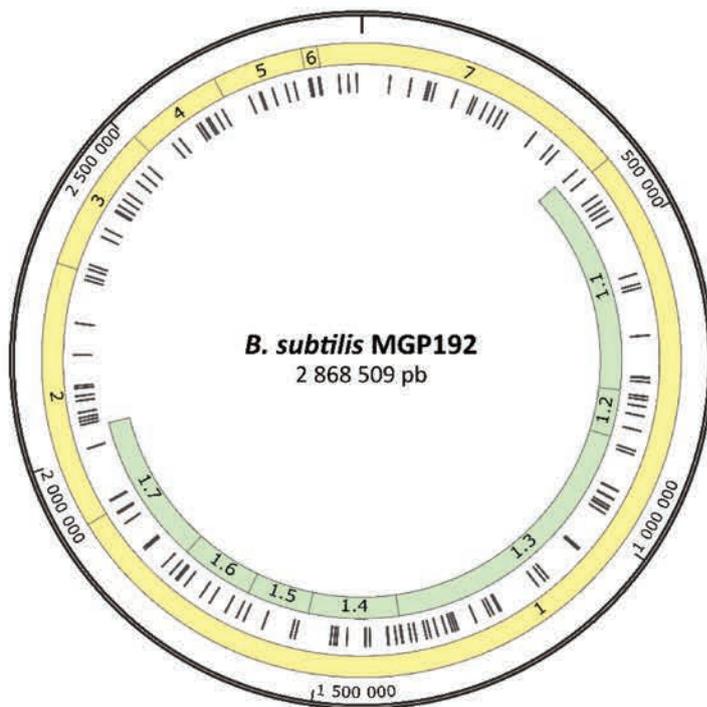
Pour conclure, cette expérience a été un véritable succès. Les sept sous-fragments ont été clonés dans la levure sans difficulté, y compris le « Sous-fragment 1.3 » ayant une taille égale à 514 kpb. A ce stade de notre étude, il s'agissait du plus grand fragment de génome de

B. subtilis MGP192 jamais cloné dans la levure et maintenu stable chez cet hôte. Pour ce fragment, il pourrait être intéressant de faire une étude de la stabilité à plus long terme (sur une centaine de générations), comme cela a été fait pour certains génomes²³⁴(cf. Thèse Dr Estelle Ruiz).

Ces données suggèrent fortement que les gènes présents dans cette région du chromosome de *B. subtilis* ne sont pas toxiques pour la levure, même si l'hypothèse d'une co-létalité, issue de l'expression simultanée de gènes situés sur des sous-fragments différents, ne peut être totalement écartée. De même, il est possible que les cassures double-brin générées par FseI aient inactivé un gène dont l'expression est toxique pour la levure (cf. **Tableau S6** – Annexes). Ces hypothèses restent plausibles, néanmoins la conduite de cette expérience nous laisse penser que l'hypothèse la plus probable pour expliquer notre difficulté à cloner le « Fragment 1 » (1500 kpb) et à le maintenir de façon stable dans la levure est le manque d'ARS sur ce fragment. Nous avons donc effectué des analyses *in silico* pour évaluer plus précisément la proportion et répartition de ces séquences sur ce fragment, mais aussi plus largement à l'échelle du génome de *B. subtilis* souche MGP192. Ce faisant, cela nous permettait d'identifier les zones pauvres en ARS dans lesquelles ces séquences pourraient être rajoutées artificiellement. L'idée étant de stabiliser la séquence et d'ensuite cloner des morceaux de plus en plus larges de notre génome d'intérêt, voire de parvenir à cloner le génome entier.

III.3. Analyse de la proportion en ARS

Il est admis dans la littérature que les ARS prédites comme fortes sont potentiellement capables de stabiliser des séquences d'ADN de 100 à 300 kpb dans la levure²⁹⁶. Concernant les expériences de clonage de certains chromosomes bactériens chez la levure, il a été montré qu'au-delà d'un certain pourcentage en G+C, il était nécessaire de rajouter ces dites séquences tous les 100-150 kpb de manière à favoriser leur maintien²⁷¹. La recherche des motifs consensus 5'-WTTTATRTTTW-3' (11nt) ou 5'-WWWTTTAYRTTTWGTT-3' (17nt), aussi appelés ARS consensus sequence (ACS), sur la séquence d'un génome bactérien à cloner peut aider à estimer la proportion en ARS (Autonomous Replicating Sequence, séquence de taille environ égale à 100 pb, riche en AT et contenant le motif ACS). Cependant, cela ne permet pas



| Fragments | ACS |
|------------|--------|
| Fragment 1 | 77 + 1 |
| Fragment 2 | 15 |
| Fragment 3 | 12 |
| Fragment 4 | 10 |
| Fragment 5 | 6 |
| Fragment 6 | 2 |
| Fragment 7 | 22 |

Figure 55 : Proportion d'ACS sur le génome après recherche manuelle. Carte du génome de *B. subtilis* souche MGP192 avec les différents fragments issus de la digestion par l'enzyme de restriction SgrDI (en jaune), ou par l'enzyme FseI (en vert). Les ACS sont indiquées par des traits noirs. Le nombre d'ACS sur les « Fragments 1 à 7 » est répertorié dans le tableau associé à la carte.

Tableau 33 : Nombre d'ACS et d'ARS putatives identifiées *in silico*

| Fragments | Tailles en pb | GC% | ACS ^a | Ori-Finder ^b | Clonage |
|-------------------|---------------|-----|------------------|-------------------------|---------|
| Fragment 1 | 1 499 486 | 43 | 77 + 1 | 18 | Non |
| Sous-Fragment 1.1 | 388 092 | 45 | 14 + 1 | 2 | Oui |
| Sous-Fragment 1.2 | 83 539 | 43 | 6 | 1 | Oui |
| Sous-Fragment 1.3 | 514 839 | 43 | 30 | 10 | Oui |
| Sous-Fragment 1.4 | 159 113 | 42 | 10 | 2 | Oui |
| Sous-Fragment 1.5 | 111 827 | 43 | 4 | 1 | Oui |
| Sous-Fragment 1.6 | 130 314 | 43 | 8 | 1 | Oui |
| Sous-Fragment 1.7 | 269 979 | 44 | 9 | 1 | Oui |

^aRecherche effectuée manuellement avec les motifs ACS sur les séquences d'ADN.

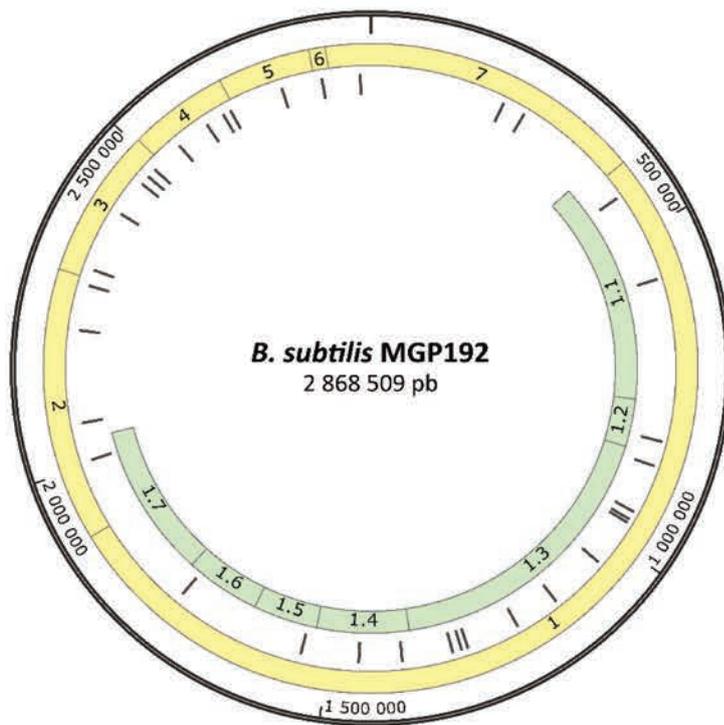
^bRecherche d'ARS putatives effectuée avec le logiciel Ori-Finder3.

de valider que les ARS soient fonctionnelles. Pour le génome de *B. subtilis*, nous avons recherché puis positionné manuellement ces séquences (**Figure 55**); nous avons comptabilisé au total 144 ACS (motif 5'-WTTTATRTTTW-3' (11nt)) et 1 ACS (motif 5'-WWWTTTAYRTTTWGTT-3' (17nt)).

Nous avons souhaité affiner cette analyse à l'aide de l'outil de prédiction Ori-finder3²⁹⁷ (<http://tubic.tju.edu.cn/Ori-Finder3/public/index.php>). Grâce à ce logiciel en ligne, l'identification d'ARS putatives se fait par comparaison avec une banque de données répertoriant toutes les séquences ARS fonctionnelles identifiées dans une collection de génomes de *S. cerevisiae*. Un score associé à chaque prédiction est indiqué, plus ce dernier est proche de 1, plus il y a de chance que l'ARS identifiée soit fonctionnelle. Par exemple, le score associé à l'ARSH4 est en moyenne compris entre 0,8 et 0,95.

Cette analyse a d'abord été réalisée sur le « Fragment 1 » et sur les sous-fragments correspondants (**Tableau 33**). Tous les scores obtenus pour cette section du génome étaient compris entre 0,15 et 0,78. Étrangement, une majorité de ces scores n'était pas associée aux motifs ACS identifiés sur le génome. Aussi, dans le but de tenter d'unifier cette analyse, nous avons conservé seulement les 20 premières entrées obtenues pour chaque fragment analysé. Parmi ces dernières, nous avons ensuite retenus seulement ceux associés à une ACS. En suivant cette démarche, nous avons observé que le score moyen était assez faible et atteignait à peine 0,4. En nous basant sur ce score, nous avons complété le **Tableau 33** (colonne 5), en ne comptabilisant que les séquences associées à des scores supérieurs à 0,4 et incluant une ACS. Parmi ces séquences, nous avons retrouvé celle contenant l'ACS de 17 nucléotides présente dans le « Sous-Fragment 1.1 » du génome de *B. subtilis* MGP192 ; elle présentait un score de 0,69, un des plus hauts obtenus lors de cette analyse.

De façon plus générale, nous pouvons observer que le nombre d'ARS prédit est bien plus faible lorsque la recherche est faite via Ori-finder3. En effet, seulement 18 séquences sont prédites dans le « Fragment 1 » contre 77+1 lorsque la recherche est réalisée manuellement (**Tableau 33** et **Figure 55**) ; une baisse qui se confirme également pour tous les sous-fragments. Au total, 18 séquences ont été identifiées dans les sous-fragments « 1.1 » à « 1.7 », un chiffre identique au nombre de séquences prédites pour le « Fragment 1 » (18 au total). De plus, elles se positionnent toutes au niveau des séquences prédites sur ce dernier. Ces proportions indiquent que l'on retrouverait, pour certains sous-fragments (« 1.2 » ou « 1.3 » par exemple)



| Fragments | Ori Finder |
|------------|------------|
| Fragment 1 | 18 |
| Fragment 2 | 4 |
| Fragment 3 | 5 |
| Fragment 4 | 4 |
| Fragment 5 | 1 |
| Fragment 6 | 1 |
| Fragment 7 | 3 |

Figure 56 : Proportion d'ARS sur le génome après recherche via Ori-finder3. Carte du génome de *B. subtilis* souche MGP192 avec les différents fragments issus de la digestion par l'enzyme de restriction SgrDI (en jaune), ou par l'enzyme FseI (en vert). Les ARS sont indiquées par des traits noirs. Le nombre d'ARS sur les « Fragments 1 à 7 » est répertorié dans le tableau associé à la carte.

une ARS environ tous les 40-50 kpb ; pour d'autres, (1.5, 1.6 et 1.7), ce serait plutôt tous les 170 kpb (**Tableau 33** et **Figure 56**). Pour ces derniers, il pourrait être utile de rajouter de telles séquences en modifiant le génome de *B. subtilis* MGP192. En effet, dans nos expériences de capture de fragments de génomes bactériens, nous ajoutons (via le pansement ARS-CEN-HIS3), l'ARSH4, une origine de réplication de levure bien caractérisée et efficace, permettant l'initiation de la réplication de l'ADN²⁹⁸. La présence de cette ARS associée à quelques ARS moins efficaces provenant du génome bactérien suffit peut-être à promouvoir la réplication de sous-fragments d'une centaine de kilobases. En revanche, il est possible que cet ensemble ne soit plus suffisant lorsqu'il s'agit de promouvoir la réplication et le maintien de fragments plus grands comme le « Fragment 1 » (1500 kpb).

Cette étude a été élargie à l'ensemble du génome (**Figure 56**). Alors que la recherche des motifs ACS 5'-WTTTATRTTTW-3' (11nt) ou 5'-WWWTTTAYRTTTWGTT-3' (17nt) avaient indiqué la présence au total de 145 ACS sur le génome de *B. subtilis* MGP192 (2,86 Mb), Ori-Finder3 en prédit seulement 36 soit plus de quatre fois moins. En comparaison, 38 ARS (score moyen d' $\sim 0,56$) sont prédites pour le génome d'*Haemophilus influenzae* (1,8 Mb) et 44 ARS (avec un score moyen d' $\sim 0,7$) pour le génome de *M. mycoides* subsp. *mycoides* (1,2 Mb), deux génomes clonés sans aucune difficulté dans la levure. Ainsi, même si le nombre d'ARS prédites via Ori-Finder3 reste raisonnable (36 pour un génome de 2,86 Mb soit une toutes les 80 kpb) les scores associés restent bas ($\sim 0,4$) signifiant que ces dernières sont potentiellement peu efficaces. Aussi, leur répartition est hétérogène (**Figure 56**). En plus de ce que nous avons décrit pour le « Fragment 1 » et les sous-fragments associés, des zones clairsemées sont aussi observées au niveau du « Fragment 7 » et du « Fragment 2 » ; signifiant qu'il faudrait peut-être aussi rajouter d'autres ARS sur l'ensemble du chromosome pour le stabiliser.

Pour terminer, au cours de cette étude, nous avons parfois noté des variations de scores, en fonction de la taille de la séquence, de sa composition et de son orientation (sens ou anti-sens). Ces données ont donc une simple valeur prédictive et sont là pour nous orienter. Seules des données expérimentales pourraient véritablement attester de la fonctionnalité d'une ARS (*i.e.* clonage d'une séquence dans un plasmide et étude de la capacité de cette séquence à promouvoir la réplication du plasmide).

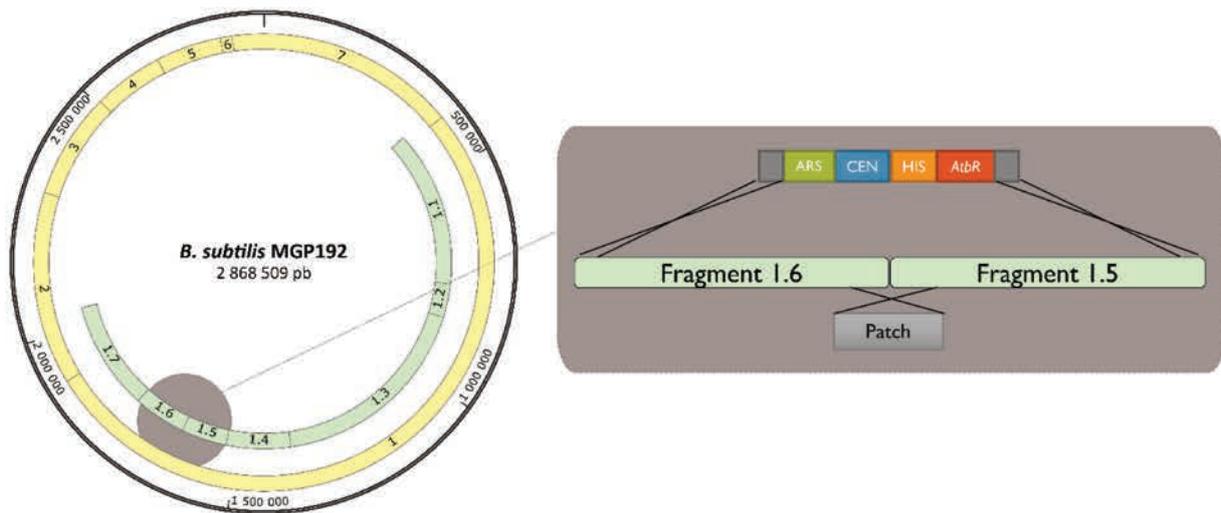


Figure 57 : Schéma de la méthode de TAR-Assembly. Carte du génome de *B. subtilis* souche MGP192 avec les différents fragments issus de la digestion par l'enzyme de restriction SgrDI (en jaune), ou par l'enzyme FseI (en vert). Les fragments 1.5 et 1.6 (en vert) à assembler sont ciblés avec un cercle gris dont un agrandissement est représenté dans l'encadré gris. L'ensemble est assemblé grâce à un patch à la jonction et simultanément recircularisé avec le pansement de recombinaison portant les éléments levures ARS-CEN-HIS3 et un marqueur de sélection produisant une résistance à un antibiotique.

Tableau 34 : Essai d'assemblage rapide de Fragments par TAR-Assembly

| | Patches | Colonies (UFC) | Clones positifs / Clones analysés ^a |
|---|---------------------|----------------|--|
| | | | PCR simplexe |
| Jonction simple | Jonction – 2kb | >300 | 0/100 |
| | Jonction – 10 kb | >300 | 0/100 |
| Jonction + ARS | ARS413 – 60 pb | >300 | 0/100 |
| | ARS413 – 2 kb | >300 | 0/100 |
| | ARS416 – 60 pb | >300 | 0/100 |
| | ARS416 – 2kb | >300 | 0/100 |
| Jonction + ARS + marqueur d'auxotrophie | ARS413/Ade2 – 60 pb | >300 | 0/100 |
| | ARS416/Lys2 – 60 pb | >300 | 0/100 |

^aRésultat issu d'une expérience de clonage. Le nombre de transformants levures analysés par PCR simplexe est indiqué, ainsi que le nombre de clones positifs obtenu.

III.4. Du TAR-Cloning au TAR-Assembly

En partant du constat que la quantité d'ARS fonctionnelles pouvait être insuffisante dans le maintien du génome de *B. subtilis* dans la levure, nous avons tenté de mettre en place une méthode rapide pour rajouter ces dites séquences. Ainsi, une expérience nommée TAR-Assembly a été effectuée. Sur le même principe que le TAR-Cloning, la méthode consiste à hydrolyser le génome grâce à une enzyme de restriction, puis à le co-transformer dans les sphéroplastes de levure avec deux fragments PCR (aussi appelés pansements) (**Figure 57**). Le premier pansement correspond au vecteur levure ARS-CEN-HIS3 ou « vecteur TAR » dont les bras de recombinaison correspondent aux extrémités d'un côté du « Fragment 1.5 », et l'autre, au « Fragment 1.6 » à assembler. Le second pansement (ou patch sur le schéma **Figure 57**) correspond à la jonction entre les deux sous fragments. Nous avons choisi de travailler soit avec (i) une jonction simple (2 kpb ou 10 kpb), (ii) une jonction additionnée d'une ARS (ARS413²⁹⁹ ou ARS416³⁰⁰), ou (iii) une jonction portant une ARS et un marqueur d'auxotrophie (ADE2 et LYS2). Les bras de recombinaison des pansements avec ajout d'ARS ou « ARS+marqueur » avaient des tailles de 60 pb (le minimum utilisé dans nos expériences), ou bien de 2 kpb. L'idée étant de favoriser potentiellement la recombinaison à cette jonction. L'expérience, effectuée avec les « Sous-fragments 1.5 » (111 828 pb) et « 1.6 » (130 314 pb), est résumée **Tableau 34**. Après transformation, nous avons obtenu un nombre considérable de colonies sur boîtes. Mais après criblage de plusieurs centaines de colonies, nous n'avons pas pu obtenir de levure hébergeant l'assemblage des deux sous-fragments. Seul le « Sous-Fragment 1.5 » a été détecté dans quelques clones.

Au vu des résultats, il semblerait que cette méthode ne soit pas la meilleure approche pour atteindre nos objectifs. Le nombre de clones « faux positifs » est bien trop élevé. Nous nous sommes donc tournés vers d'autres types de développements.

Discussion-Conclusion

Au cours de cette étude, la totalité du génome de *B. subtilis* a pu être cloné de manière stable en fragments dont les tailles varient de 6 kpb à 512 kpb (**Figure 58**). Avec ces résultats, nous avons pu produire une banque de levure propageant différents fragments et sous-fragments du génome. Cette banque vient compléter celle générée par Azevedo *et al.* en 1993,

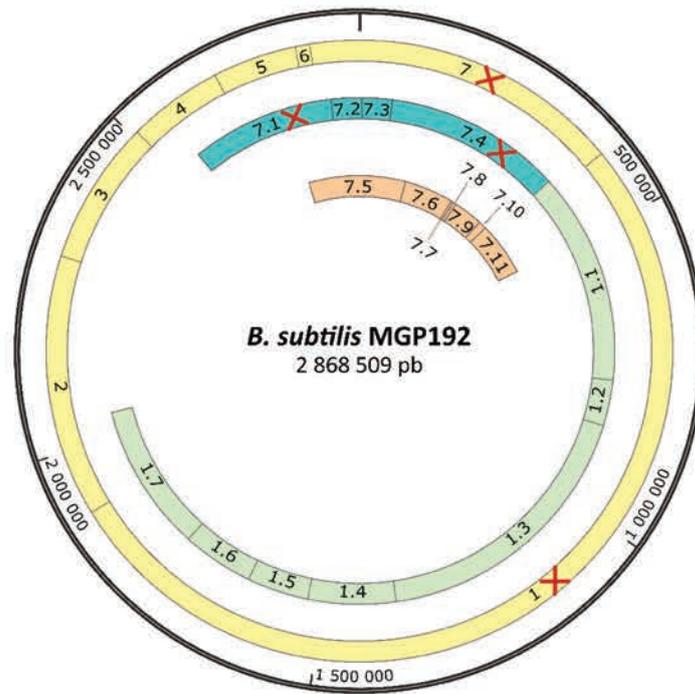


Figure 58 : Bilan des expériences de TAR-Cloning. Carte du génome de *B. subtilis* souche MGP192 avec les différents fragments issus de la digestion par l'enzyme de restriction SgrDI (en jaune), ou par l'enzyme FseI (en vert ou bleu), ou bien, par l'enzyme AsiSI (en saumon). Les fragments ou sous-fragments n'ayant pas été clonés sont marqués d'une croix rouge.

en y apportant des fragments du génome de *B. subtilis* plus grands que ceux portés par les YAC produits par ces auteurs (>190 kpb)²⁰⁸. Même si elle couvre la quasi-totalité du génome (2,86 Mb), notre banque n'est pas aussi complète que nous l'aurions souhaité puisque quatre fragments ou sous-fragments n'ont pas été correctement maintenus dans la levure : le « Fragment 1 » (1500 kpb), le « Fragment 7 » (482 kpb), le « Sous-fragment 7.1 » (264 kpb) et le « Sous-fragment 7.4 » (322 kpb).

Les expériences de TAR-Cloning décrites dans ce chapitre ont permis de répondre en partie aux problématiques d'ordre techniques et biologiques auxquelles nous faisons face pour cloner le génome de *B. subtilis* entier dans la levure. Tout d'abord, nous avons pu valider que la qualité des préparations de chromosomes de *B. subtilis* en bloc d'agarose (cf. Chapitre 1) semblait suffisante pour cloner des fragments de génomes. À ce jour, nous n'avons pas vu de différence majeure (nombre de transformants positifs) entre les préparations réalisées à partir du protocole « ED » ou « Biorad », et traitées ou pas avec un agent compactant (hexamine cobalt (III) chloride). Cela étant, nous avons pu valider et utiliser les protocoles mis au point au cours du Chapitre 1. Ensuite, il semble désormais assez clair qu'un ensemble de facteurs sont impliqués et nuisent à nos objectifs. Parmi eux le manque d'ARS, la présence de *loci* répétés, et/ou la présence de gènes toxiques. Des analyses sont actuellement en cours dans le but de compléter les données collectées.

Concernant le « Fragment 7 » (482 kpb), nous avons mis en évidence que des séquences répétées ou des *loci* induisant une potentielle toxicité, peuvent favoriser son instabilité. Les perspectives de ce travail pourraient consister à éliminer ces éléments. En effet, des études ont répertorié que pas moins de 22 protéines ribosomiques parmi les 50 codées par le génome de *B. subtilis* ne sont pas essentielles, du moins pour la prolifération cellulaire^{301,302}. En revanche, concernant les ADNr (10 opérons trouvés dans le génome de *B. subtilis* souche 168 et conservés dans *B. subtilis* souche MGP192), des travaux indiquent que produire des mutants avec seulement une seule copie d'ADNr entraîne des difficultés de croissance chez *B. subtilis*³⁰³. Cependant, dans le cas de l'opéron *rno*, ces mêmes auteurs montrent qu'il est possible de contre balancer cet effet en mutant le suppresseur associé³⁰⁴. Ces données suggèrent que différentes stratégies peuvent être tentées pour contourner nos difficultés : suppression de certains opérons rDNAs, modulation des niveaux d'expression de ces opérons ou encore, relocalisation de ces opérons vers d'autres régions du chromosome

afin d'éviter qu'ils ne soient tous concentrés en un même point. De même, les gènes non-essentiels et dont l'expression serait potentiellement gênante pour la levure ou dont la séquence possède également des répétitions (comme ceux de l'opéron non-essentiel *srfAA-AD* codant la surfactine) pourraient être supprimés afin de voir si cela stabiliserait la séquence.

Concernant le « Fragment 1 » (et plus largement l'ensemble du génome), notre plan de travail consiste désormais à ajouter par génie génétique plusieurs ARS dans le génome de *B. subtilis*. Ce travail est réalisé en collaboration avec l'équipe Syber (INRAE UMR MICALIS) ; leur connaissance du génome nous permet d'intégrer ces éléments à des positions précises sur le chromosome n'interférant pas, entre autres, avec les unités de transcription. Ce travail nous semble nécessaire et est conforté par les travaux de clonage des génomes de *Caulobacter crescentus* (0,79 Mb et 57% G+C)²⁷⁰ et *E. coli* (1,03 Mb et 51% G+C)²²⁵ pour lesquels l'introduction d'un dizaine de séquences ARS a été nécessaire pour qu'ils soient maintenus de façon stable dans la levure.

La méthode du TAR-Cloning utilisée au cours de ce chapitre et par extension celle du TAR-assembly, sont toutes deux soumises à la présence des sites de restriction et ne permettaient pas de choisir strictement la portion chromosomique que nous souhaitions capturer. Cela ne permettait pas un design précis des expériences et a parfois rendu l'interprétation des résultats délicate. De plus, l'efficacité globale du TAR-Cloning est faible, il est souvent nécessaire de cribler beaucoup de colonies pour avoir peu de transformants positifs. Nous avons déjà souligné au cours de ce chapitre (mais aussi au cours du Chapitre 1), que nous observions beaucoup de colonies faussement positives résultant probablement de la circularisation du vecteur TAR. Des données expérimentales personnelles (Lartigue *et al.*), mais aussi, issues de la littérature ont également confirmé ce phénomène^{221,234,264}. Les travaux pionniers sur cette méthode (Kouprina, Larionov et Noskov), apportent cependant des solutions à ce problème. Leurs données décrivent largement l'utilisation d'un vecteur nommé « pARS-VN » dont la conception permet d'effectuer une contre-sélection avec le marqueur URA3 associé à l'ajout de 5-FOA dans le milieu, et ainsi de limiter la sélection des clones propageant un vecteur circularisé²³². Plus récemment, ce groupe a publié une nouvelle version de la méthode de TAR-Cloning en proposant non pas de pré-digérer les chromosomes *in vitro* avec des enzymes de restriction, mais d'utiliser à la place le système CRISPR-Cas9 afin de générer une coupure double-brin. Si cette digestion *in vitro* améliore l'efficacité du TAR-

Cloning (plus de 20 fois), en stimulant ensuite la recombinaison *in vivo* chez la levure, elle permet également de choisir de manière raisonnée le *locus* que l'on souhaite cloner³⁰⁵.

Nous aurions pu mettre en place ces solutions, mais nous avons finalement préféré développer une méthode alliant l'outil d'édition de génome CRISPR-Cas9 à la méthode de Fusion (*cf.* Chapitre 1). L'objectif était de pouvoir cloner et modifier des fragments de génomes de plus en plus grands et d'aller jusqu'à la capture du génome entier de *B. subtilis* MGP192. Pour rappel, nous avons vu au cours du Chapitre 1, que la méthode de Fusion semblait plus propice au transfert d'un tel génome dans la levure. Techniquement, les méthodes basées sur l'extraction et la purification des chromosomes en bloc d'agarose se trouvaient limitées pour une taille de génome de 2,86 Mb. De plus, l'association avec le système CRISPR-Cas9 permettait de simultanément capturer des fragments tout en rajoutant des ARS, ou bien, en éliminant des *loci* problématiques. Cette méthode appelée CREasPy-Fusion a été développée à la fin de ces travaux de thèse et sa mise au point fait l'objet du Chapitre 3.

CHAPITRE 3 – CReasPy-Fusion :
Développement d'une nouvelle méthode
de clonage de génomes bactériens dans la
levure

CHAPITRE 3 – CReasPy-Fusion : Développement d'une nouvelle méthode de clonage de génomes bactériens dans la levure.

Introduction

Nous avons vu au cours du Chapitre 1 que le génome réduit de *B. subtilis* MGP192 n'avait pas pu être cloné entier dans la levure par des méthodes déjà décrites dans la littérature. Au cours de ce travail, un résultat notable fût néanmoins obtenu : nous avons en effet démontré que le génome de cette bactérie Gram+ d'une taille égale à 2,86 Mb peut bel et bien être transféré intact dans la levure et s'y maintenir entier sur quelques générations. Une seule méthode parmi les 3 mises en œuvre a permis de révéler ce fait majeur pour notre étude : la méthode de clonage par Fusion. Les deux autres méthodes laissaient penser que le génome ne pénétrait peut-être même pas dans la cellule (absence de transformants sur boîte ou faux positifs). Ce résultat marquant a fortement impacté notre recherche :

- D'une part, nous avons compris qu'une fois entré dans la levure le génome était rapidement dégradé. Au cours du Chapitre 2, nous avons donc cherché à identifier les facteurs biologiques et/ou techniques responsables de cette instabilité. Pour cela, nous avons utilisé la méthode du TAR-Cloning afin de capturer des morceaux de génomes de différentes tailles. Ces expérimentations ont permis de cloner la totalité du génome de *B. subtilis* MGP192 en 25 fragments de tailles variables (6 kpb à 514 kpb). Elles ont également permis d'établir qu'un ensemble de facteurs étaient très certainement responsables de l'instabilité globale du génome dans la levure. Parmi eux, le manque d'ARS fonctionnelles sur certaines régions du chromosome, la présence de nombreuses copies d'opérons ribosomiques concentrées en un endroit du génome pouvant favoriser la survenue d'évènements de recombinaison, et/ou un effet toxique en cas d'expression connexe de ces gènes.
- D'autre part, nous avons également compris que pour des molécules d'ADN de cette taille (2,86 Mb), la méthode par Fusion constituait vraisemblablement la voie d'entrée la plus appropriée. En effet, les techniques de clonage Classique ou par CReasPy-Cloning mettent en jeu des génomes intacts isolés en bloc d'agarose ayant subi, entre autres, des traitements à la protéinase K et à divers détergents. Il s'agit donc de

génomés très peu condensés, potentiellement difficiles à transformer. Au contraire, dans la méthode de Fusion, le contenu cellulaire de la bactérie serait entièrement « déversé » dans la levure (soit après fusion protoplastes/sphéroplastes, soit après endocytose de la bactérie puis décomposition), délivrant un génome hautement compacté directement dans le cytoplasme de la cellule hôte. Le génome doit ensuite migrer jusqu'au noyau mais cette étape n'a pas été documentée.

Forts de ces connaissances et souhaitant poursuivre notre objectif initial (*i.e.* cloner le génome de *B. subtilis* MGP192 entier dans la levure), nous avons imaginé une nouvelle stratégie de clonage, nommée CREasPy-Fusion qui allie la fusion de cellules bactériennes avec des sphéroplastes de levure et le système d'édition de génome CRISPR-Cas9. À travers une conception précise et adaptable à souhait, pour au choix cloner des fragments de génome, ajouter des ARS, ou enlever des *loci* non désirés, notre idée était de développer une méthode de clonage puissante qui soit plus qu'une simple alternative aux méthodes déjà décrites dans la littérature (et testées au cours du Chapitre 1) ; notamment par sa flexibilité mais aussi son ouverture vers une gamme de microorganismes plus larges (possibilités de cloner tout aussi bien des petits génomes que des grands génomes).

La méthode de CREasPy-Fusion est la combinaison de deux stratégies déjà existantes : la méthode de clonage par CREasPy-Cloning²²⁰ et la méthode de clonage par Fusion²⁸³ (*cf.* §III - Introduction). En effet, bien que spectaculaire d'un point de vue fondamental, cette dernière possède un inconvénient majeur : sa mise en œuvre requiert de savoir transformer la bactérie dont on souhaite cloner le génome. Or, de tels protocoles ne sont pas toujours disponibles ce qui limite fortement son utilisation. Nous avons alors eu l'idée de surpasser cette barrière en y associant, lors des premières étapes, le système d'édition de génome CRISPR-Cas9, à l'instar de ce qui est fait en CREasPy-Cloning. Dès lors, la méthode se décline en trois grandes étapes successives :

(1) Les levures sont transformées avec un plasmide pCas9 et un plasmide pARNg spécifique (étape empruntée au CREasPy-Cloning) ;

(2) Les levures pré-chargées en plasmides pCas9 et pARNg sont fusionnées avec les bactéries d'intérêt (étape empruntée au clonage par Fusion), et en présence d'un vecteur spécifique de recombinaison (amplicon PCR linéaire appelé également « pansement ») ;

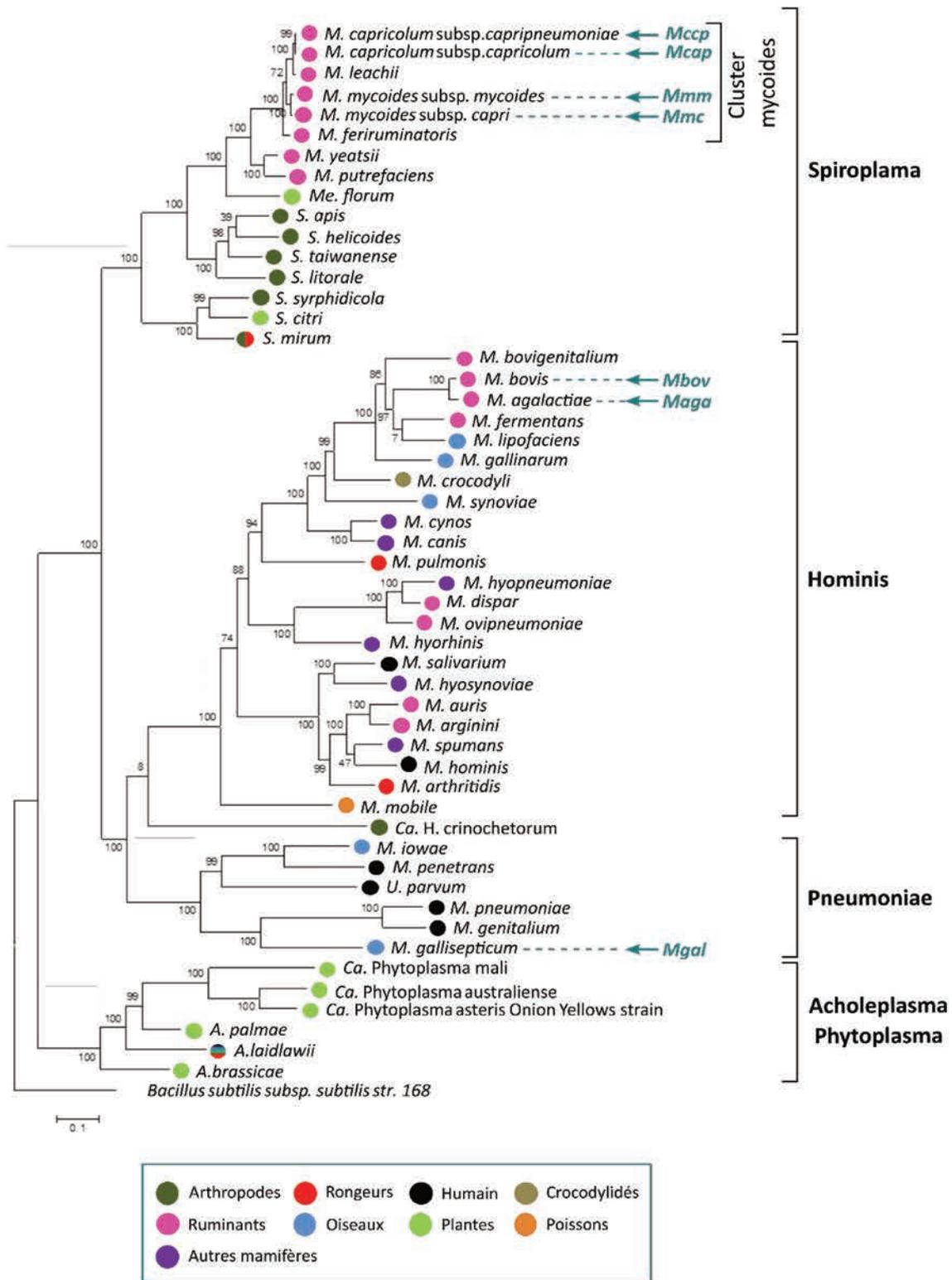


Figure 59 : Arbre phylogénétique des principaux Mollicutes. Cet arbre a été adapté d'après celui publié dans Ipoutcha *et al.* 2019. Il est basé sur la comparaison de 62 séquences protéiques orthologues impliquées dans la traduction grâce à la méthode Maximum Likelihood (maximum de vraisemblance). Les principaux groupes phylogénétiques Spiroplasma, Hominis, Pneumoniae, Acholeplasma/Phytoplasma sont indiqués, ainsi que le cluster mycoides. Les espèces utilisées pour la mise en place de la méthode de clonage par CRISPR-Cas9 sont indiquées par une flèche bleu turquoise. Les hôtes de chaque espèce sont indiqués par des ronds colorés suivant le code couleur précisé dans l'encadré. Il faut noter que le modèle bactérien *Bacillus subtilis* est utilisé dans cet arbre comme groupe externe (outgroup).

(3) Après entrée du génome dans la levure, le duplex Cas9-ARNg induit une cassure double brin sur la cible choisie, qui est réparée grâce au pansement fourni, intégrant par la même occasion les éléments génétiques (ARS)-CEN-HIS3.

Le Chapitre 3 décrit le développement de cette nouvelle méthode, en utilisant dans un premier temps des génomes de mycoplasmes, puis dans un second temps celui de *B. subtilis* MGP192. Les mycoplasmes sont des bactéries sans paroi qui possèdent des génomes de petites tailles (<1,5Mb), à faible G+C% (<40%) et utilisant un code génétique non standard. Elles ont été utilisées comme modèles pour mettre au point toutes les méthodes de clonage de génomes dans la levure, décrites jusqu'ici dans la littérature (*cf.* §III - Introduction). Au vu des nombreuses difficultés rencontrées avec le génome de *B. subtilis*, il nous est apparu judicieux de valider la faisabilité de cette nouvelle approche avec des génomes de mycoplasmes (facilement « clonables » dans la levure) et de ne l'appliquer à notre organisme d'intérêt qu'une fois son efficacité prouvée. Pour la mise en place de la méthode chez *B. subtilis* MGP192, nous avons défini une stratégie en plusieurs étapes qui a consisté à (i) revalider la Fusion entre des protoplastes de bactérie Gram+ et la levure ; puis à (ii) transférer (et modifier) de *B. subtilis* vers la levure un plasmide navette ; et enfin à (iii) capturer après fusion un fragment de génome de *B. subtilis* dans la levure.

I. Clonage et ingénierie de génomes de mycoplasme par CReasPy-Fusion

Les expériences de clonage par CReasPy-Fusion présentées dans la première partie de ce chapitre ont été mises au point avec sept espèces de mycoplasmes (**Figure 59**) que l'on peut répartir en deux groupes :

- Les espèces *Mycoplasma capricolum* (*Mcap*), *Mycoplasma mycoides* subsp. *capri* (*Mmc*), *Mycoplasma mycoides* subsp. *mycoides* (*Mmm*), et *Mycoplasma capricolum* subsp. *capripneumoniae* (*Mccp*) appartenant au « Cluster mycoides » (**Figure 59**, groupe phylogénétique Spiroplasma). Leurs génomes avaient déjà été clonés avec succès dans la levure^{215,230} (Gourgues *et al.* non publié) en utilisant des méthodes décrites dans la littérature.
- Les espèces *Mycoplasma gallisepticum* (*Mgal*), *Mycoplasma agalactiae* (*Maga*) et *Mycoplasma bovis* (*Mbov*) appartenant aux groupes phylogénétiques Hominis et Pneumoniae (**Figure 59**) et dont les génomes n'avaient jamais été clonés dans la levure auparavant.

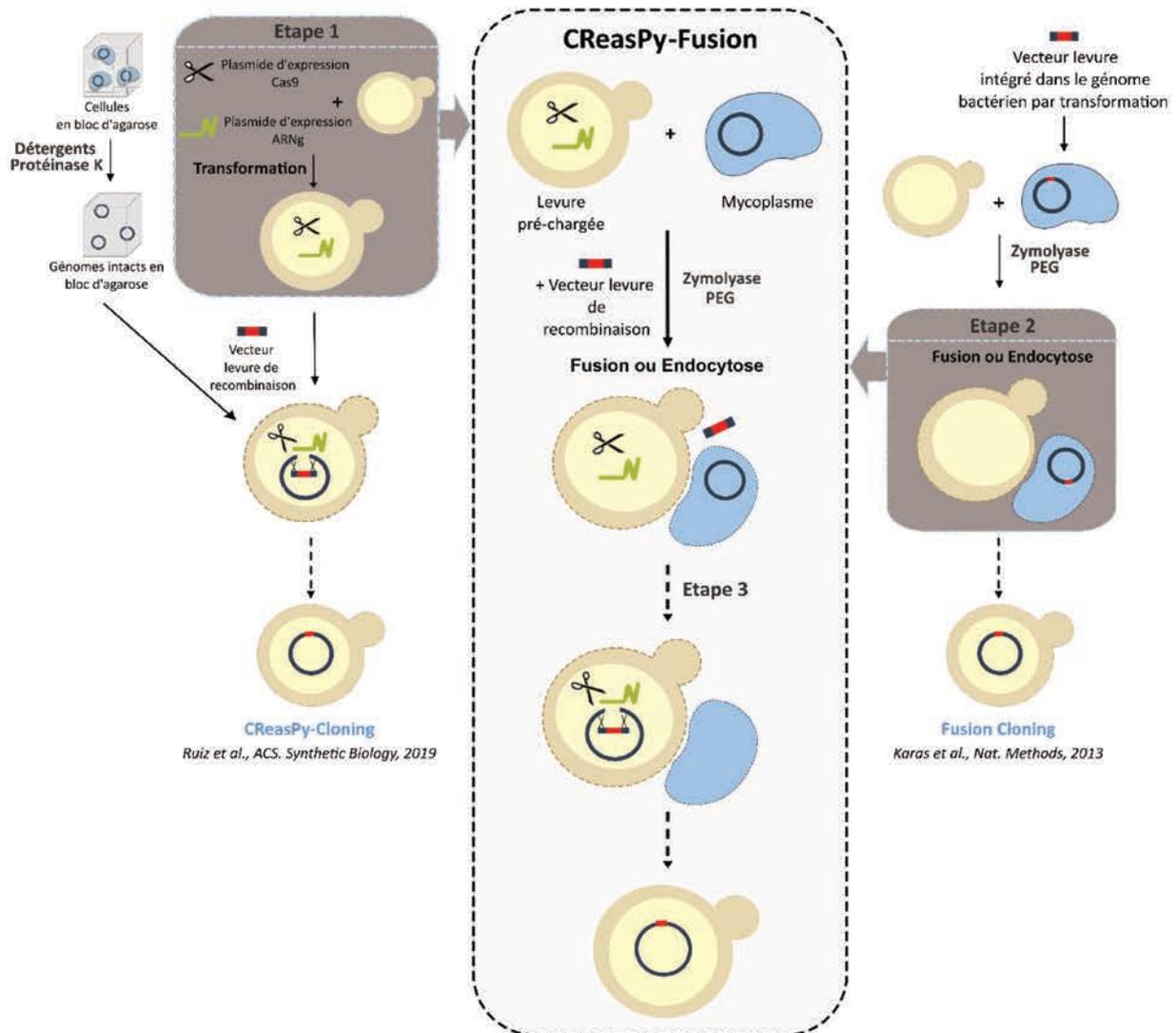


Figure 60 : Schéma général de la méthode de CReasPy-Fusion appliquée aux mycoplasmes. Les levures sont transformées avec les plasmides pCas9 (p414-TEF1p-CAS9-CYC1t) et pARNg (p426-SNR52p-gRNA+AarI-SUP4t) permettant respectivement l'expression de la protéine Cas9 et d'un ARNg spécifique du *locus* cible (Etape 1). Les levures pré-chargées sont co-transformées avec des cellules de mycoplasmes et le vecteur levure de recombinaison, soit les éléments levures CEN-HIS3, avec ou sans ARS, encadrés par deux bras de recombinaison identiques aux séquences encadrant le *locus* cible, suivant un protocole modifié de transformation au PEG par fusion des cellules (Etape 2). Le génome est pris en charge par la levure, clivé par Cas9 et simultanément réparé par le système de recombinaison homologue chez la levure grâce au pansement de réparation fourni. Le génome bactérien ainsi modifié porte les éléments génétiques levure intégrés à la place du *locus* cible et est propagé par cet hôte en tant que chromosome artificiel surnuméraire et circulaire (Etape 3).

Ces sept espèces sont présentées en détail **Tableau S7** (souches, isolats, taille du génome, %GC). Les gènes ciblés au cours des expériences de CReasPy-Fusion sont également indiqués dans le tableau pour chacune d'entre elles (noms des gènes, mnémoniques et localisation génomique). Tout comme pour la méthode de clonage par CReasPy-Cloning, le protocole de CReasPy-Fusion (**Étape 1, Figure 60**), nécessite une étape de pré-chargement des levures avec le système CRISPR-Cas9. Dans le but d'optimiser le processus expérimental, nous avons décidé de recycler du matériel biologique déjà à disposition au laboratoire, notamment des clones levures transformés avec le pCas9 et un pARNg spécifique d'un gène cible. Nous avons pu en particulier, recycler des clones levures permettant de cloner les génomes des espèces *Mcap*, *Mmc*, *Mccp* et *Mmm* en ciblant les gènes codant l'enzyme de restriction CCATC (*Mcap* CK, *Mmc* GM12 et *Mmc* 95010), la peptidase S41 (*Mccp* 14020, *Mccp* 95043 et *Mcap* 14232) et l'opéron *glpOKF* (*Mmm* T1/44). Ils étaient non essentiels et pour certains potentiellement impliqués dans la virulence (*e.g.* *glpOKF*) (avant dernière colonne du **Tableau S7**). En effet, ces espèces sont très proches d'un point de vue phylogénétique et ont de nombreux gènes en commun avec des séquences parfois (quasiment) identiques. Au cours de nos expérimentations, un même clone levure a donc pu être utilisé pour cloner le génome de plusieurs mycoplasmes.

Concernant les espèces *Mgal*, *Mbov* et *Maga*, nous avons choisi de cibler des gènes potentiellement impliqués dans la virulence (*mnuA* ou *cysP*). Les ARNg contenant les protospacers ciblant les différents *loci* ont été conçus, clonés dans le plasmide pARNg et validés par séquençage (*cf.* **Tableau S1** et §II.1. Matériels et Méthodes). Puis, de la même manière que celle décrite Chapitre 1 §II.1.1., nous avons ensuite transformé la souche VL6-48N-Cas9 avec les pARNg nouvellement construits pour chacune des trois espèces. En parallèle, les pansements de recombinaison correspondants ont été produits et purifiés (*cf.* §II.7). Pour rappel, ces pansements de recombinaison sont des éléments nécessaires à la réparation du génome après coupure par le duplex Cas9-gRNA. Ils ont été produits par PCR, purifiés puis dosés. Ils sont constitués d'un centromère (CEN), d'un marqueur d'auxotrophie pour l'histidine (HIS3), éventuellement d'une origine de réplique levure (ARS) (version avec et sans ARS), et d'un gène conférant la résistance à un antibiotique (tétracycline ou gentamycine). Ce dernier élément est ajouté afin d'anticiper l'étape de transplantation de

génomés depuis la levure vers une bactérie receveuse, et permettre la sélection de transplants avec un génome modifié.

Une fois, le matériel biologique prêt et dûment sélectionné, l'étape 2 de CReasPy-Fusion a pu être amorcée (**Figure 60**). Elle consiste à fusionner les levures préchargées en pCas9/pARNg avec les bactéries cibles. Cette étape a été réalisée en appliquant le protocole décrit par Karas *et al.*, en 2013 pour fusionner des levures (non transformées) avec des bactéries d'intérêt. Cinq contrôles ont été réalisés à chaque fois :

- Le contrôle « levure » attestant que les sphéroplastes transformés par pCas9/pARNg ne sont pas capables de se multiplier dans un milieu de sélection SD-HIS.
- Le contrôle « mycoplasme » prouvant qu'après CReasPy-Fusion les cellules de mycoplasme ne sont pas capables de se multiplier dans un milieu levure.
- Le contrôle « pansement », qui consiste à transformer les levures pCas9/pARNg avec les matrices de recombinaison. En théorie, aucune colonie ne devrait apparaître sur ce contrôle. Mais comme mentionné à plusieurs reprises au cours de ces travaux, des colonies sont observées de façon quasiment systématique avec ce dernier.
- Un contrôle « plasmide » qui consiste à transformer les levures avec un plasmide portant les éléments levures ARS-CEN-HIS3 (en général, le plasmide ayant servi de matrice à la PCR mise en œuvre pour produire le pansement de recombinaison). Il permet de valider l'expérience (contrôle +).
- Un contrôle « plasmide + mycoplasmes » mettant en présence (i) les sphéroplastes de levures, (ii) les cellules de mycoplasmes et (iii) un plasmide purifié. Ce témoin a été imaginé afin de vérifier que nous obtenions des transformants levures sur milieu de sélection même en présence de mycoplasmes. Plus clairement, nous nous demandions dans quelle mesure une proportion de plasmide pénétrait dans les cellules bactériennes ; et donc étaient en quantité limitante pour être captée par les cellules levures. Ce témoin était à chaque fois comparé au contrôle « plasmide ».

Au total, treize expérimentations de CReasPy-Fusion (décrites ci-après) ont été effectuées au cours de ce travail ; pour chacune de ces expérimentations, aucune colonie n'est apparue sur le contrôle « levure » ainsi que sur le contrôle « mycoplasme ». Des colonies ont,

Tableau 35 : Clonage du génome de *Mcap*ΔRE par Fusion.

| Volume de sphéroplastes de levures | Volume de <i>M. capricolum</i> subsp <i>capricolum</i> ΔRE | Colonies (CFU) | Clones positifs / Clones analysés ^a | | |
|------------------------------------|--|----------------|--|-----------------------------|-------------------|
| | | | PCR simplexe | PCR multiplexe ^b | PFGE ^b |
| 100 μL | 50 μL | 88 | 10/10 | 5/5 | 2/2 |
| | 100 μL | 97 | 9/10 | 4/4 | 2/2 |
| 200 μL | 50 μL | 202 | 10/10 | 5/5 | 2/2 |
| | 100 μL | 185 | 9/10 | 5/5 | 2/2 |

^aRésultats issus d'une expérience de clonage. Les résultats obtenus sont résumés dans ce tableau. Le nombre de transformants levures analysés par PCR, PCR multiplexe et PFGE sont indiqués, ainsi que le nombre de clones positifs obtenus.

^bLes analyses par PCR multiplexe et PFGE n'ont pas été effectuées sur tous les clones positifs en PCR simplexe mais seulement sur un échantillon.

en revanche été observées sur les contrôles « pansement », mais cela était attendu (voir plus haut). Par soucis de clarté, nous ne reviendrons pas sur ces points pour chaque expérimentation. Néanmoins, dans certains cas, les différences observées entre les contrôles « plasmide » et « plasmide + mycoplasmes » sont discutées.

1.1. Clonage par Fusion chez les mycoplasmes

Avant de mettre en place les expérimentations de CReasPy-Fusion, nous avons voulu nous assurer que nous étions capables de reproduire les résultats obtenus en 2013 par Karas *et al.*²³⁶ (*i.e.* clonage des génomes *Mcap* et *Mmc* marqués dans la levure par la méthode de Fusion). Nous avons donc effectué une expérience préliminaire en utilisant la souche *Mcap* California Kid (CK) dépourvue du système de restriction modification CCATC par délétion du gène MCAP_RS00270 (ancien mnémonique : MCAP_0050) via l'intégration de la cassette ARS-CEN-HIS3-PSpuro dans le *locus*. Cette souche a été construite par Dr. C. Lartigue au cours de ses travaux de post-doctorat au JCVI et est à disposition au laboratoire. Elle est couramment appelée *McapΔRE* et sera dénommée comme tel dans la suite du chapitre^{216,217}. Ce protocole a été effectué avec la levure VL6-48N.

Au cours de cet essai, différents volumes de sphéroplastes de levures (100 μ L ou 200 μ L) et de cellules de *McapΔRE* (50 μ L ou 100 μ L) ont été testés (**Tableau 35**). Il est apparu que le nombre de transformants était quasiment doublé en ajoutant 200 μ L de sphéroplastes plutôt que 100 μ L. A l'inverse, la quantité de *McapΔRE* (50 μ L ou 100 μ L) ne semblait pas influencer sur le nombre de colonies obtenues. Cette première information nous a permis de fixer le ratio volume de sphéroplastes/volume de mycoplasmes à 200 μ L pour 50 μ L. Ceci était en accord avec les quantités suggérées par les travaux de Karas *et al.*²²³

Pour chaque condition de transformation, 10 colonies ont été repiquées et analysées. Les analyses par PCR simplexe avec le couple d'amorces MCAPCKF1/MCAPCKR1 permettant de détecter la présence du génome de *McapΔRE* dans la levure ont permis d'observer une bande à 272 pb identique à celle obtenue pour le témoin positif (ADNg *McapΔRE*) dans la quasi-totalité des clones (9 à 10 suivant les conditions), indiquant que le génome de *McapΔRE* avait certainement été cloné avec une grande efficacité (**Figure 61B**).

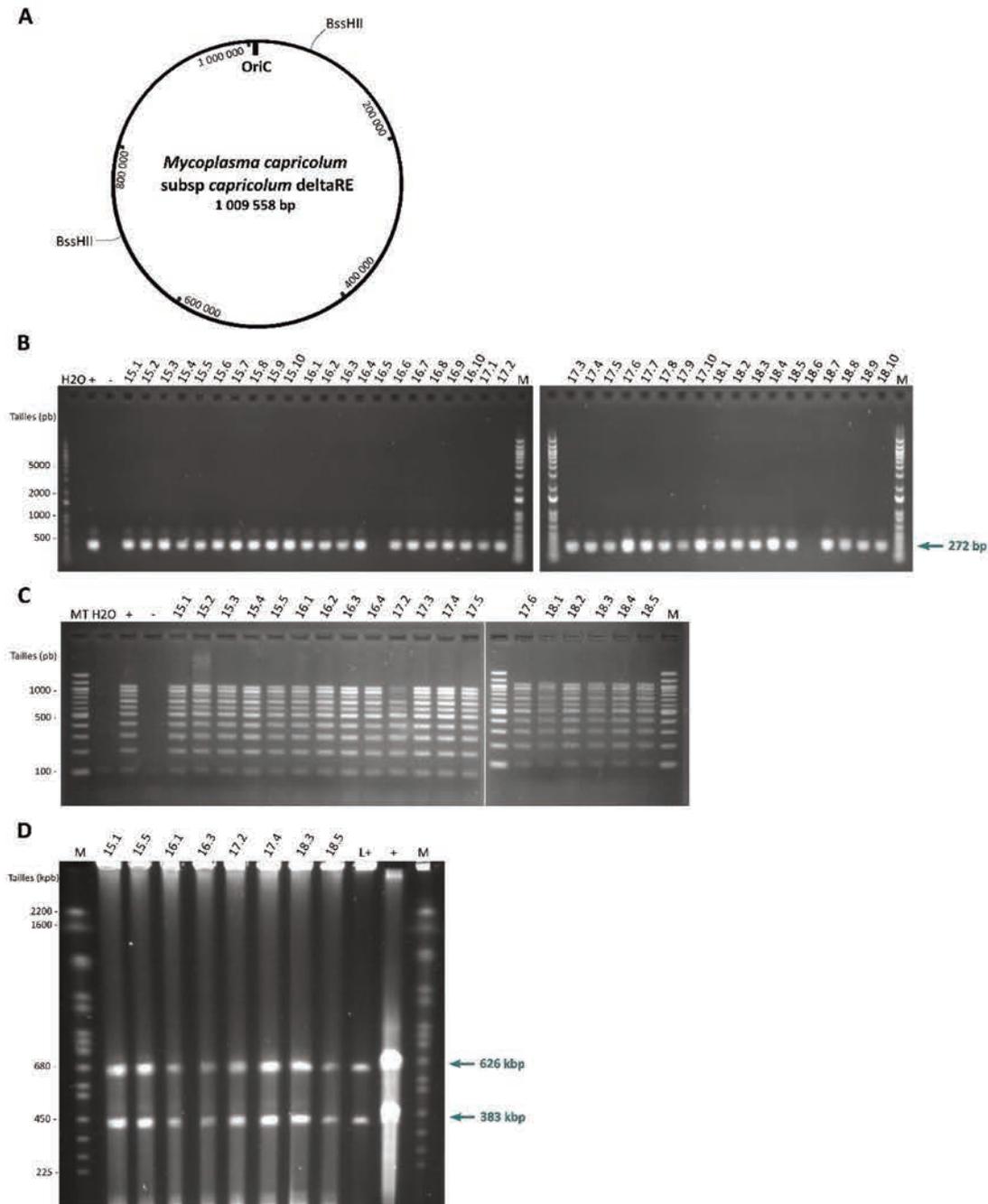


Figure 61 : Criblage des transformants levures après clonage du génome de *McapΔRE* par la méthode de Fusion. Cette figure regroupe les résultats obtenus après clonage du génome de *McapΔRE* par la méthode de Fusion et analyse des transformants par PCR simplexe, PCR multiplexe et par PFGE. **[A]** La présence du génome de *McapΔRE* est détectée par PCR à l'aide d'amorces générant un amplicon de 272 pb identique au témoin positif. **[B]** L'intégrité du génome de *McapΔRE* cloné est évaluée par PCR multiplexe grâce à un ensemble d'amorces distribué de façon homogène sur le génome (de 100 pb à 1000 pb). Les clones propageant des génomes non tronqués montrent un profil à 10 bandes identique à celui obtenu pour le témoin positif. **[C]** La taille du génome de *McapΔRE* est finalement validée par restriction enzymatique et analyse par PFGE. La coupure par l'enzyme BssHII permet de générer deux fragments d'ADN de 626 kpb et 383 kpb. "M" : PCR simplexe = DNA Ladder 1kb+ Invitrogen (100-12000 pb) ; PCR multiplexe = DNA Ladder 100 pb NEB (100-1517 pb) ; PFGE = *S. cerevisiae* chromosomal DNA ladder Bio-Rad (225-2200 kpb). "H2O" : témoin négatif sans ADN. "+" : témoin positif avec l'ADNg de *McapΔRE*. "L+" : témoin positif avec l'ADNg d'un transformant *S. cerevisiae* propageant le génome de *McapΔRE* ; "-" : témoin négatif levure avec l'ADNg de *S. cerevisiae* VL6-48N.

Pour chaque condition, cinq clones ont ensuite été analysés par PCR multiplexe, dans le but de vérifier que le chromosome était potentiellement cloné entier dans la levure. Au cours de cette analyse, tous les transformants testés montraient un profil à 10 bandes (allant de 100 pb à 1000 pb), identique à celui du témoin positif et ont donc tous été validés (**Figure 61C**). Finalement, deux clones par condition ont été sélectionnés pour effectuer l'analyse par électrophorèse en champs pulsé (PFGE) et évaluer la taille de la molécule d'ADN clonée. Le résultat présenté **Figure 61D**, montre que tous les clones levures arboraient un profil identique à celui obtenu pour le témoin positif « + » *McapΔRE* et à celui obtenu pour le second témoin positif « L+ » (transformant levure propageant le génome de *McapΔRE* issu d'une expérience antérieure) ; à savoir un profil à deux bandes dont les tailles (626 kpb et de 383 kpb) correspondaient aux tailles théoriques attendues lorsque le génome *McapΔRE* est clivé par l'enzyme BssHII.

Lors de cette analyse, 70 à 100% des clones levures sélectionnés se sont révélés positifs, validant ainsi (i) la mise en place du protocole au laboratoire, et (ii) la méthode puisque nos résultats corroboraient ceux décrits dans la littérature.

1.2. Preuve de concept de la méthode de CReasPy-Fusion

A cette étape, nous avons complexifié la méthode de Fusion en ajoutant le système d'édition de génome CRISPR-Cas9 dans le partenaire « levure ». Nous avons travaillé avec deux espèces de mycoplasmes, *Mcap* souche CK et *Mmc* souche GM12 dont les génomes avaient été clonés par la méthode de Fusion par Karas *et al.*²⁸³ (et validées au laboratoire pour *Mcap* CK).

Des sphéroplastes de levures VL6-48N-pCas9-pARNg ont été fusionnés avec soit des cellules de *Mcap* CK, soit des cellules de *Mmc* GM12 au cours d'expériences indépendantes (**Tableau 36**). Pour ces deux espèces, nous avons ciblé les gènes codant l'endonucléase de restriction de type II CCATC, respectivement MCAP_RS00270 et MMCAP2_RS00520 (**Figure 62**). Le ratio de cellules a été fixé à 200 µL de sphéroplastes pour 50 µL de cellules bactériennes et la quantité de matrice de recombinaison à 1 µg (au lieu des 300 ng habituellement utilisé dans le protocole de CReasPy-cloning²²⁰). Les raisons expliquant ce choix étaient que (i) nous travaillions avec plus de cellules au cours de la CReasPy-Fusion (au

Tableau 36 : Bilan du clonage des génomes de *Mcap* CK et *Mmc* GM12 par CReasPY-Fusion.

| Mycoplasme | Clones positifs / Clones analysés ^a | | |
|--|--|----------------|------|
| | PCR simplexe | PCR multiplexe | PFGE |
| <i>M. capricolum</i> subsp. <i>capricolum</i> CK (<i>Mcap</i> 27343 CK) | 4/22 | 3/4 | 2/3 |
| <i>M. mycoides</i> subsp. <i>capri</i> GM12 (<i>Mmc</i> GM12) | 6/43 | 4/6 | 2/4 |

^aRésultats issus de deux expériences indépendantes pour chaque espèce de mycoplasme faisant office de répliques. Les résultats obtenus sont résumés dans ce tableau. Le nombre de transformants levures analysés par PCR, PCR multiplexe et PFGE sont indiqués, ainsi que le nombre de clones positifs obtenus.

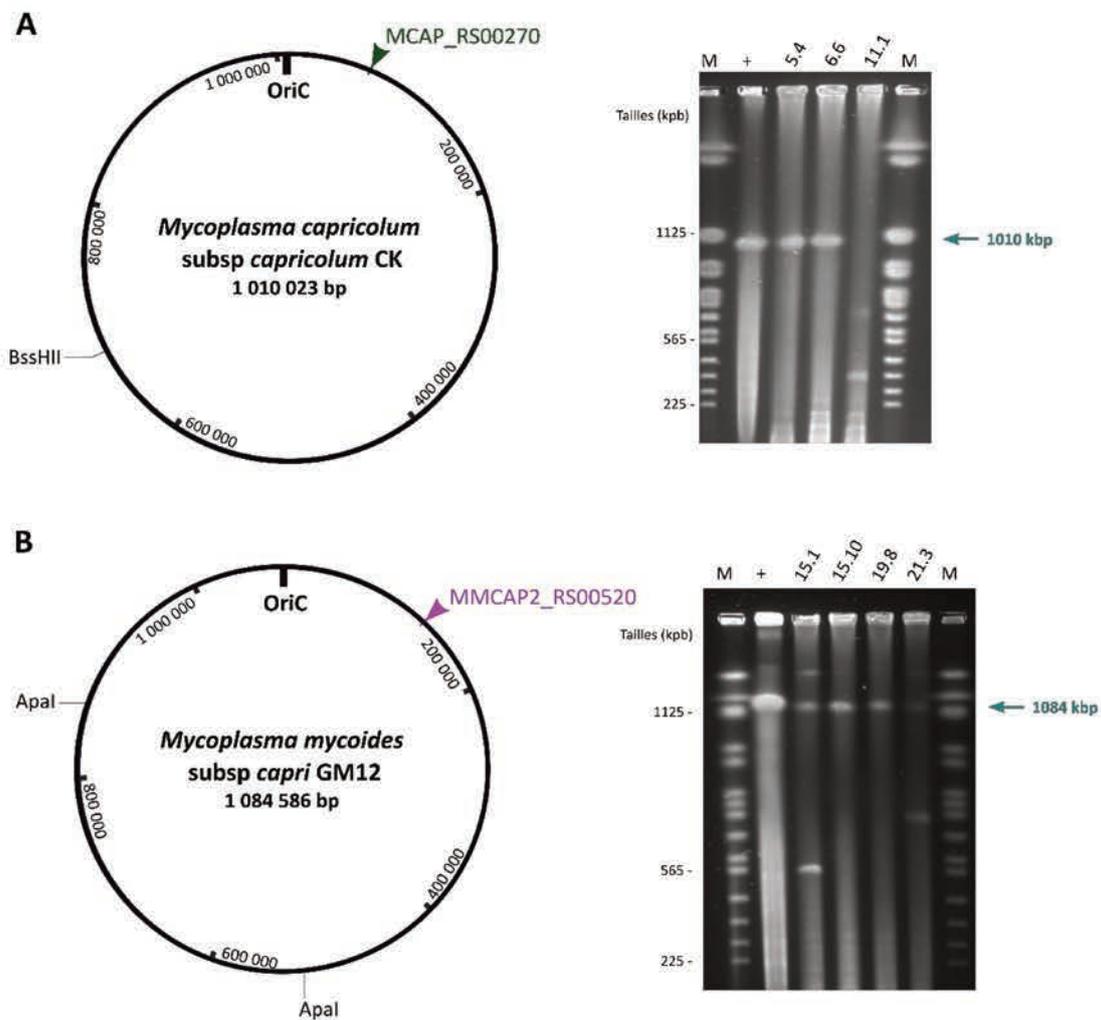


Figure 62 : Criblage par PFGE des transformants levures après clonage des génomes de *Mcap* CK et *Mmc* GM12 par CReasPy-Fusion. Cette figure regroupe les résultats obtenus pour évaluer le succès de la méthode de clonage avec les génomes de *Mycoplasma capricolum* subsp. *capricolum* CK (*Mcap* CK) et de *Mycoplasma mycoides* subsp. *capri* GM12 (*Mmc* GM12). **[A]** La carte du génome de *Mcap* CK est présentée avec son locus cible MCAP_RS00270 (flèche verte). Sa taille après clonage dans la levure est validée par restriction enzymatique et analyse par PFGE. La coupure par l'enzyme BssHIII permet de linéariser le génome en un fragment de 1010 kbp. **[B]** La carte du génome de *Mmc* GM12 est présentée avec son locus cible MMCAP2_RS00520 (flèche rose). Sa taille après clonage dans la levure est finalement validée par restriction enzymatique et analyse par PFGE. La coupure par l'enzyme SfoI permet de linéariser le génome en un fragment de 1084 kbp. "M" : *S. cerevisiae* chromosomal DNA ladder Bio-Rad (225-2200 kbp). "+" : témoin positif avec l'ADNg de *Mcap* ou de *Mmc*.

total 250 μ L contre 100 μ L en CReasPy-cloning), et que (ii) nous supposons qu'une proportion du pansement pouvait ne pas être « captée » par les levures mais par les mycoplasmes.

Deux expériences de CReasPy-Fusion ont été effectuées pour chacune des espèces, les résultats obtenus sont résumés **Tableau 36**. Le nombre de colonies dénombrées sur boîte pour *Mcap* CK est \sim 10 fois inférieur avec cette méthode (22 transformants) comparé à celui obtenu avec le protocole de Fusion « simple » (\sim 200 transformants, cf. §II.1.2). Cela peut s'expliquer par le fait que le génome doit, non seulement rentrer dans la cellule hôte mais également être modifié pour se maintenir de façon stable. Ainsi, pour obtenir des transformants propageant un génome de *Mcap* une cascade d'évènements doit avoir lieu : (i) transfert du génome dans la levure par fusion et acquisition concomitante du pansement, (ii) coupure double brin du génome par le duplex Cas9-ARNg, (iii) réparation de la cassure par le système de recombinaison homologue de la levure en utilisant le pansement comme matrice (double crossing-over) et, (iv) maintien et propagation du génome bactérien (portant les éléments de la levure insérés à un *locus* précis) en tant que plasmide centromérique de levure.

Sur les 22 transformants obtenus sur boîte, 4 clones arboraient des profils correspondant à la suppression du *locus* cible *MCAP_RS00270* et à son remplacement par le pansement. Sur ces 4 clones, 3 ont ensuite été validés en PCR multiplexe puis analysés en PFGE : 2 transformants sur les 3 analysés ont été identifiés comme propageant le génome de *Mcap* CK à la taille attendue de \sim 1 Mb (**Figure 62A**). Il s'agit des clones 5.4 et 6.6 qui présentent une bande à 1010 kpb identique au témoin positif. Le cl11.1 affiche quant à lui, un profil dégradé avec des bandes surnuméraires de plus petites tailles (\sim 700 kpb et \sim 300 kpb) et n'a pas été validé.

Concernant *Mmc* GM12, le nombre de colonies sur boîtes était plus élevé que pour *Mcap* CK (43 transformants, soit presque 2 fois plus). Cependant, après analyses par PCR et PFGE, il est apparu que ce nombre masquait une proportion de faux positifs et que les clones levures propageant le génome de *Mmc* GM12 étaient un peu moins nombreux. En effet, seulement 6 transformants sur 43 ont été validés par PCR simplexe, puis, 4 clones sur 6 en PCR multiplex (profil à 10 bandes) et enfin, 2 clones sur 4 en PFGE (**Figure 62B**). Il s'agit des transformants levure cl 15.10 et 19.8 qui montrent un profil linéaire (1084 kpb) après restriction par *Sfol*, identique au témoin positif (*ADNg Mmc* GM12). Les autres clones (cl 15.1

Tableau 37 : Bilan du clonage des génomes de *Mccp* 95043, *Mccp* 14020 et *Mmm* T1/44.

| Mycoplasme | Taille de génome ou de fragment capturé | Clones positifs / Clones analysés ^a | | |
|---|---|--|----------------|-------------------|
| | | PCR simplexe | PCR multiplexe | PFGE ^b |
| <i>M. capricolum</i> subsp. <i>capripneumoniae</i> 95043 Niger (<i>Mccp</i>) | 1,016 Mb | 3/14 | 3/3 | 3/3 |
| <i>M. capricolum</i> subsp. <i>capripneumoniae</i> 14020 Tanzanie (<i>Mccp</i>) | 1,016 Mb | 4/15 | 4/4 | 3/3 |
| <i>M. mycoides</i> subsp. <i>mycoides</i> T1/44 (<i>Mmm</i>) | 1,188 Mb | 51/60 | 39/51 | 6/10 |
| | 0,596 Mb | 64/70 | 39/64 | 11/12 |

^aRésultat issu d'une expérience de clonage pour *Mccp* et de trois expériences indépendantes pour *Mmm*. Les résultats obtenus sont résumés dans ce tableau. Le nombre de transformants levures analysés par PCR, PCR multiplexe et PFGE sont indiqués, ainsi que le nombre de clones positifs obtenus.

^bLes analyses par PFGE n'ont pas été effectuées sur tous les clones positifs en PCR multiplex mais seulement sur un échantillon.

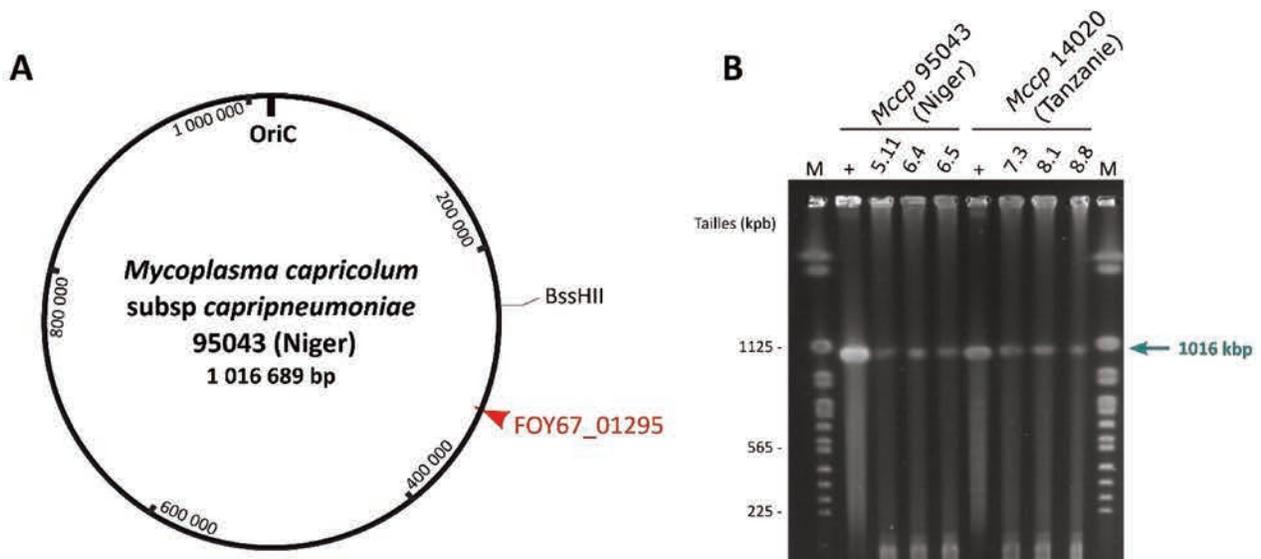


Figure 63 : Criblage par PFGE des transformants levures après clonage des génomes de *Mccp* souches 95043 et 14020 par CReasPy-Fusion. Cette figure regroupe les résultats obtenus pour évaluer le succès de la méthode de clonage avec les génomes de *Mycoplasma capricolum* subsp. *capripneumoniae* (*Mccp*). [A] La carte du génome de *Mccp* 95043 est présentée avec son locus cible FOY67_01295 (flèche rouge). [B] La taille des génomes de *Mccp* 95043 ou 14020 est validée après clonage dans la levure par restriction enzymatique et analyse PFGE. La coupure par l'enzyme BssHIII permet de linéariser le génome en un fragment linéaire de 1016 kpb. "M" : *S. cerevisiae* chromosomal DNA ladder Bio-Rad (225-2200 kpb). "+" : témoin positif avec l'ADNg de *Mccp*.

et 21.3) présentent des bandes intermédiaires synonyme de dégradation du génome bactérien et/ou d'un mélange de population.

Ces premières expériences nous ont permis de démontrer que nous étions capables de cloner des génomes de mycoplasmes par CReasPy-Fusion et de conclure que la méthode était valide.

1.3. CReasPy-Fusion appliquée à deux mycoplasmes pathogènes majeurs du cluster mycoïdes : clonage de génomes entiers ou capture de fragments de génomes

Les résultats obtenus avec *Mcap* CK et *Mmc* GM12 nous ont poussé à étendre la méthode à deux mycoplasmes pathogènes de caprins (*Mccp*) et de bovins (*Mmm*) et dont l'impact sanitaire sur les élevages est majeur^{306,307}. Pour rappel, ces espèces appartiennent au groupe phylogénétique Spiroplasma et, plus particulièrement, elles sont regroupées dans le cluster mycoïdes aux côtés de *Mcap* et *Mmc* (**Figure 59**). Des expériences indépendantes ont à nouveau été produites pour chacune d'entre elles. Des sphéroplastés de levures VL6-48N-pCas9-pARNg ont été fusionnés soit, avec des cellules de *Mccp*, soit avec des cellules de *Mmm*. Les résultats obtenus sont résumés dans le **Tableau 37**.

Une première expérience a permis de démontrer qu'il était possible de cloner le génome de deux souches de terrain de *Mccp* : la souche 14020 (isolée au Niger) et la souche 95043 (isolée en Tanzanie). Ces deux souches nous ont été transmises par la Dr. Lucia Manso-Silvan (CIRAD, INRAE de Montpellier). L'incorporation des lements levures a t effectuee  la place d'un gene cible codant la peptidase S41³⁰⁸. Nous avons represente sur la **Figure 63A** seulement la carte du genome de la souche *Mccp* 14020 et le gene cible FOY67_01295. En effet, bien que nous ayons  disposition au laboratoire la souche *Mccp* 95043, la sequence du genome n'est pas encore disponible. Des analyses de sequenage *de novo* sont en cours afin d'obtenir la sequence complete de ce genome et produire la carte manquante.

Apres clonage par CReasPy-Fusion, 14 et 15 colonies ont t obtenues pour chacune des souches de *Mccp* et analysees par PCR simplexe au niveau du *locus* cible. Trois clones ont t valides pour la souche 95043 (Niger), contre 4 pour la souche 14020 (Tanzanie). La PCR multiplexe a ensuite permis de montrer qu'aucun rearrangement majeur n'avait eu lieu, et

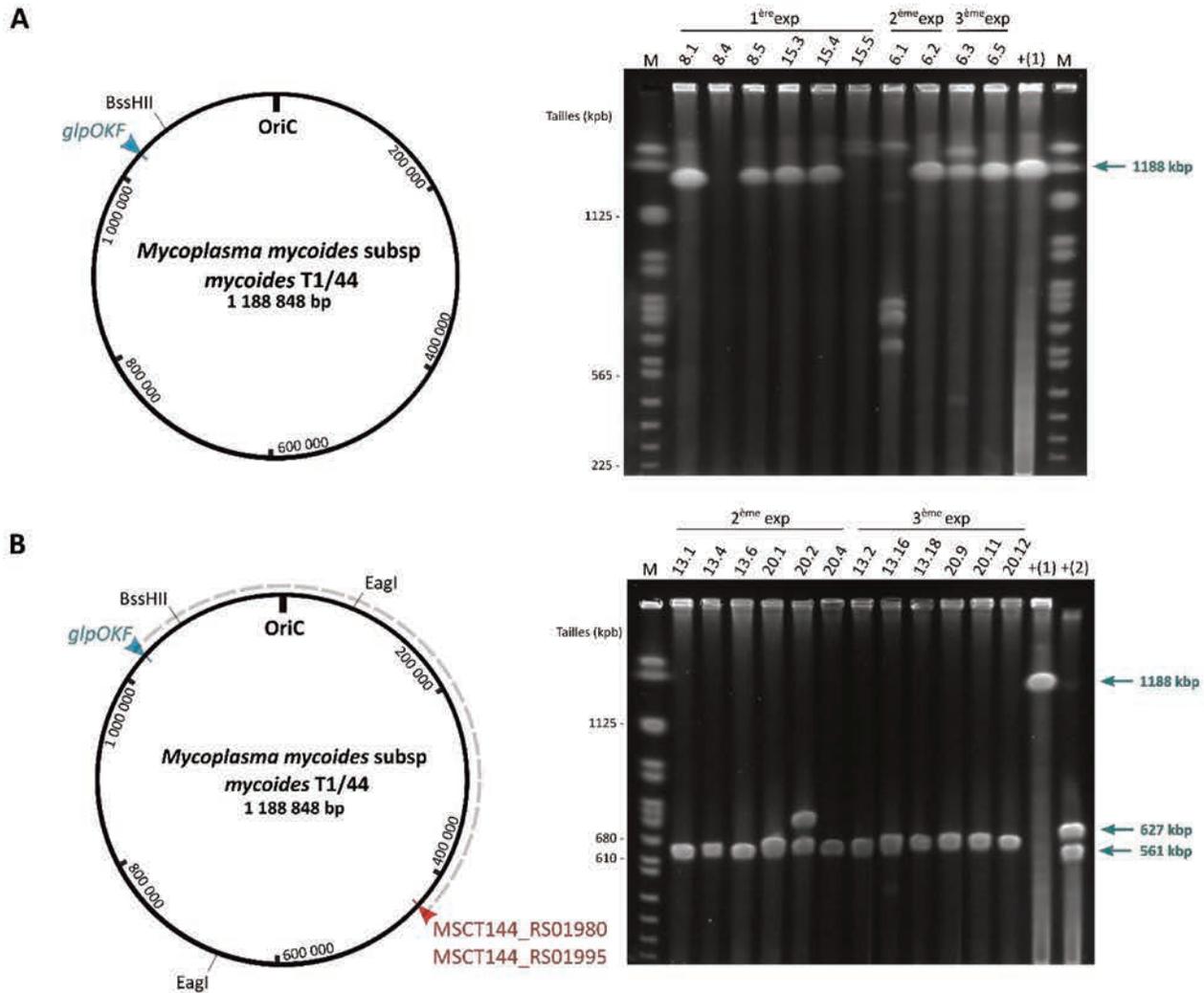


Figure 64 : Criblage par PFGE des transformants levures après clonage du génome ou de la moitié du génome de *Mmm* souche T1/44 par CREasPy-Fusion. Cette figure regroupe les résultats obtenus pour évaluer le succès de la méthode de clonage avec le génome de *Mycoplasma mycoides subsp. mycoides* (*Mmm*). **[A]** La carte du génome de *Mmm* T1/44 est présentée avec son locus cible l'opéron *glpOKF* (flèche bleue). La taille du génome entier de *Mmm* T1/44 est validée après clonage dans la levure par restriction enzymatique et analyse PFGE. La coupure par l'enzyme BssHIII permet de linéariser le génome en un fragment linéaire de 1188 kbp. **[B]** La carte du génome de *Mmm* T1/44 est présentée avec les locus cibles *glpOKF* (flèche colorée bleue) et MSCT144_RS01980 et MSCT144_RS01995 (flèche en rouge). La portion de génome capturée avec double coupure est représentée par le trait gris en pointillé. La taille du fragment de génome de *Mmm* T1/44 capturé est validée après clonage dans la levure par restriction enzymatique et analyse PFGE. La coupure par l'enzyme BssHIII permet de linéariser le génome en un fragment linéaire de 596 kbp. "M" : *S. cerevisiae* chromosomal DNA ladder Bio-Rad (225-2200 kbp). "+(1)" : témoin positif avec l'ADNg de *Mmm* digéré avec BssHIII. "+(2)" : témoin positif avec l'ADNg de *Mmm* digéré avec EagI

que ces génomes étaient probablement intègres dans les clones analysés. Enfin, 6 transformants (3 pour *Mccp* 95043 et 3 pour *Mccp* 14020) ont été analysés par PFGE (**Figure 63B**). Ces clones présentaient des profils de génomes dont la taille était identique à celle du témoin positif (1016 kpb) après digestion par l'enzyme BssHII (linéarisation du chromosome).

Ces expériences montraient que les génomes des deux souches *Mccp* pouvaient être clonés dans la levure par CReasPy-Fusion, et de surcroît avec une efficacité quasi-identique. Environ 20% des clones analysés propageaient un génome *Mccp* entier (pour la souche 95043 et pour la souche 14020). Ce pourcentage est plus élevé que pour les expériences de CReasPy-Fusion effectuées avec les génomes de *Mcap* CK et *Mmc* GM12, montrant que des résultats assez hétérogènes pouvaient être obtenus en fonction de l'espèce manipulée.

Nous avons ensuite travaillé avec la souche vaccinale *Mmm* T1/44, en effectuant le protocole de CReasPy-Fusion en triplicata et de manière à : (i) cloner et modifier le génome entier (1,188 Mb), mais aussi, à (ii) capturer la moitié du génome (0,596 Mb) (**Figure 64A**). Pour les expériences de clonage de génome entier, nous avons choisi de cibler l'opéron *glpOKF*³⁰⁹. Cela nous permettait, en effet, de recycler plusieurs couples d'amorces mais également des levures VL6-48N-Cas9-ARNg produites au cours de précédents travaux, et qui avaient permis de cloner le génome de *Mmm* souche Afadé par CReasPy-Cloning²²⁰. Pour les expériences de capture de fragment de génome, un pARNg-double-target (*cf.* Tableau S1 et §II.1. Matériels et Méthodes), a été produit de manière à cibler simultanément l'opéron *glpOKF* et les gènes MSCT144_RS01980 ou MSCT144_RS01995 codant respectivement pour une lipoprotéine et une protéine hypothétique (**Figure 64B**).

Les résultats résumés **Tableau 37**, montrent que nous avons obtenu un nombre de transformants positifs beaucoup plus élevé que ceux détaillés lors des expériences précédentes (*Mcap* CK, *Mmc* GM12, *Mccp* 95043 et *Mccp* 14020). Sur le nombre de transformants analysés, plus de la moitié se sont avérés positifs. Lors du premier criblage en PCR simplexe, 51 clones sur 60 testés (génome entier) et 64 clones sur 70 testés (moitié de génome) se sont révélés positifs (soit 80 et 90% de clones positifs respectivement). Puis, lors du second criblage en PCR multiplexe, 39 clones sur 51 testés (génome entier) et 39 clones sur 64 testés (moitié de génome) ont été validés (soit 60 et 75% de clones positifs respectivement). Enfin parmi les transformants sélectionnés pour l'analyse par PFGE, 6 clones

sur 10 testés (génomme entier) et 11 clones sur 12 testés (moitié de génome) ont montré le profil attendu (soit 60 et 90% de clones positifs respectivement). Pour cette dernière analyse représentée **Figure 64**, nous avons confirmé que les clones 8.1, 8.5, 15.3, 15.4, 6.2 et 6.5 propageaient le génome entier de *Mmm* T1/44 à 1188 kpb (**Figure 64A**). Les transformants analysés après capture d'un fragment à 596 kpb, sont tous positifs à l'exception du cl 20.2 dont le profil montre, en plus du fragment capturé, une autre bande entre 700 et 800 kpb (**Figure 64B**).

L'ensemble des expériences menées sur *Mccp* et *Mmm* ont permis de confirmer non seulement que (i) nous étions capables d'appliquer la méthode de CREasPy-Fusion pour cloner des génomes issus de souches de terrain (*Mccp*) et d'une souche vaccinale (*Mmm*), mais que (ii) nous pouvions également capturer un fragment de génome en utilisant deux ARNGs. Nous aurions pu penser que l'efficacité de capture d'un fragment de génome serait plus faible, comme c'est le cas en CREasPy-Cloning, lorsque plusieurs *loci* sont ciblés en même temps. Ce n'est pas le cas (du moins sur cet exemple) ; en effet, les résultats obtenus démontraient qu'il était possible de circulariser une portion du chromosome après coupure de deux *locus* distants de 500 kpb avec une efficacité équivalente à celle du clonage d'un génome entier nécessitant seulement une coupure. De plus, dans ces deux cas, le nombre de clones positifs était exceptionnellement haut. Lorsque des expériences de clonage de génomes sont mises en place (quelle que soit la méthode), il est rare d'atteindre des efficacités de clonage aussi élevées. Ces résultats mettaient encore une fois en exergue les grandes variations d'efficacité qu'il peut exister d'une espèce à l'autre en CREasPy-Fusion.

1.4. Mise en évidence d'un facteur génétique influençant l'efficacité de la méthode à travers le clonage du génome de *M. gallisepticum*.

Nous avons vu que la méthode de clonage CREasPy-Fusion, permettait de cloner des génomes provenant de différentes espèces de mycoplasmes, appartenant au même groupe phylogénétique Spiroplasma. Nous avons alors souhaité étendre la méthode en tentant de cloner le génome de *Mycoplasma gallisepticum* (*Mgal*), espèce pathogène de la volaille³¹⁰ et qui appartient à un autre groupe phylogénétique, le groupe Pneumoniae (**Figure 59**). Ce choix était motivé par le fait que le clonage du génome de cette espèce n'avait jamais été décrit dans la littérature. Il était donc intéressant de tenter cette nouvelle approche en travaillant

Tableau 38 : Essai de clonage du génome de *M. gallisepticum* S6 par CReasPy-Fusion

| Condition | Levure | Volume Sphéroplastes | Plasmide ou Pansement | Quantité d'ADN | Mycoplasme | Volume cellules | UFC | Criblage des transformants | |
|-----------|----------------------|----------------------|-----------------------|----------------|----------------|-----------------|-------|----------------------------|--------------|
| | | | | | | | | Analysés | PCR simplexe |
| 1 | VL6-48N-pCas9-pARNg1 | 200 µL | - | - | - | - | 0 | - | - |
| 2 | | | pMT85 | 300 ng | - | - | >1000 | - | - |
| 3 | | | | | <i>Mgal</i> S6 | 100 µL | 1 | - | - |
| 4 | | | PST ARS-CEN-HIS-CysP | 1 µg | - | - | >300 | - | - |
| 5 | | | | | <i>Mgal</i> S6 | 25 µL | 80 | 20 | 0/20 |
| 6 | | | | | | 50 µL | 10 | 10 | 0/10 |
| 7 | | | | | | 100 µL | 0 | - | - |
| 8 | | | | | - | - | 34 | - | - |
| 9 | | | PST CEN-HIS-CysP | 25 µL | 7 | 7 | 0/7 | | |
| 10 | | | | <i>Mgal</i> S6 | 50 µL | 1 | 1 | 0/1 | |
| 11 | | | | | 100 µL | 0 | - | - | |
| 12 | VL6-48N-pCas9-pARNg2 | 200 µL | - | - | - | - | 0 | - | - |
| 13 | | | pMT85 | 300 ng | - | - | >1000 | - | - |
| 14 | | | | | <i>Mgal</i> S6 | 100 µL | 20 | - | - |
| 15 | | | PST ARS-CEN-HIS-CysP | 1 µg | - | - | >300 | - | - |
| 16 | | | | | <i>Mgal</i> S6 | 25 µL | 93 | 20 | 0/20 |
| 17 | | | | | | 50 µL | 26 | 20 | 0/20 |
| 18 | | | | | | 100 µL | 0 | - | - |
| 19 | | | | | - | - | 45 | - | - |
| 20 | | | PST CEN-HIS-CysP | 25 µL | 29 | 20 | 0/20 | | |
| 21 | | | | <i>Mgal</i> S6 | 50 µL | 10 | 10 | 0/10 | |
| 22 | | | | | 100 µL | 0 | - | - | |
| 23 | - | - | - | - | 100µL | 0 | - | - | |

avec ce mycoplasme. Nous avons fusionné des sphéroplastes de levures VL6-48N-pCas9-pARNg avec des cellules de *Mgal* souche S6 (souche type isolée du poulet). Les pARNg ont été construits de sorte à cibler le gène *cysP* (GCW_RS01695) codant la peptidase C1³¹¹. Cette enzyme joue un rôle dans la virulence, plus précisément, dans le clivage des immunoglobulines de l'hôte.

Une première expérience a été planifiée de sorte à comparer deux plasmides pARNg-*cysP* portant deux ARNg différents ciblant tous les deux *cysP* (ARNg 1 et ARNg 2), deux pansements de recombinaison, (ARS-CEN-HIS et CEN-HIS) et plusieurs volumes de cellules de *Mgal* (25 µL, 50 µL, 100 µL) à fusionner à 200 µL de sphéroplastes de levure (**Tableau 38**). Au total, 256 colonies ont été obtenues sur les boîtes pour toutes les conditions confondues. Aussi, plus de colonies ont été obtenues avec le pansement contenant une ARS, plutôt que sans ARS. Cependant, les boîtes contrôles « pansement + ARS » (conditions 4 et 15) contenaient également plus de colonies que les contrôles « pansement – ARS » (conditions 8 et 19). Il était donc probable que ce nombre représente une plus grande proportion de faux positifs. Un second point étonnant, était que la quantité de colonies augmentait, lorsque le volume de cellules de *Mgal* diminuait : ~100 de colonies avec le ratio 25 µL/200 µL, entre 1 et 26 colonies avec le ratio 50 µL/200 µL et aucune colonie avec le ratio 100 µL/200 µL. Ce dernier résultat nous a surpris et nous nous sommes demandés si un facteur biologique ne pouvait pas expliquer cette diminution drastique. De plus, la comparaison entre les contrôles « plasmide » et « plasmide + mycoplasmes », semblait corroborer cette hypothèse, puisque la fusion avec le plasmide seul (pMT85) permettait d'obtenir plus de milles colonies (conditions 2 et 13), tandis qu'en présence de *Mgal*, ce nombre avait chuté à 1 ou 20 UFC (conditions 3 et 14). Nous avons tout de même procédé à l'analyse de 108 clones. Mais, malheureusement, aucun des transformants n'a été validé et ne semblait propager le génome de *Mgal*.

Les travaux de Karas *et al.* avaient mis en évidence en 2013 que, lorsque les systèmes de restriction modification des bactéries fusionnées (*Mcap*, *Mmc*, *Haemophilus influenzae*) étaient inactivés, l'efficacité de la fusion augmentait. Dans le cas précis de *Mgal*, l'analyse du génome via la plateforme Rebase Database³¹² (<http://rebase.neb.com/rebase/rebase.html>), a permis d'identifier uniquement deux systèmes de restriction modification (RM). Cependant, des systèmes de restriction sont aussi retrouvés chez *Mcap*, *Mmc*, *Mccp* et *Mmm* (jusqu'à 6 pour *Mmc* GM12) et nous avons pourtant montré qu'il était possible de cloner le génome de

ces quatre espèces de mycoplasmes en CReasPy-Fusion (cf. Tableau S2, Annexes). Ainsi, tout en gardant à l'esprit une possible implication des systèmes RM, nous avons émis l'hypothèse que d'autres types de nucléases pouvaient entrer en jeu et mettre en échec la méthode. Cette hypothèse était confortée par les résultats obtenus à la fois sur les contrôles « plasmide » (nombreuses colonies) et « plasmide + mycoplasmes » (peu de colonies) mais aussi sur les boîtes essais (augmentation du nombre de transformants levure lorsque le nombre de cellules *Mgal* diminuait). À ce stade nous pouvions en effet imaginer que les molécules de plasmide ou de pansement ajoutées dans le mélange de cellules soient dégradées en amont ou au cours de la Fusion. Ce type de matrice (ADN double brin linéaire ou circulaire), pouvant être un substrat de choix pour des nucléases sécrétées, membranaires ou encore cytoplasmiques.

Nous avons identifié dans la littérature plusieurs candidats potentiels :

- Une endonucléase Cas9 fonctionnelle inhérente au système de défense CRISPR décrit chez *Mgal*, aussi appelée MgCas9^{313,314}. Nous nous sommes demandés dans quelle mesure cette enzyme pouvait avoir un effet délétère lors de la fusion levure/*Mgal* et du mélange des contenus cellulaires (clivage de l'ADNg de levure ou du pansement reconnus comme « étrangers »). Il est important de préciser que ce système n'est pas présent chez les espèces du cluster mycoïdes.
- la protéine MGA_0637 (souche Rlow) ou GCW_RS00070 (souche S6) qui présente quelques similarités de séquences avec la nucléase membranaire MnuA identifiée comme facteur de virulence chez *Mbovis* (MBOVPG45_0215)²⁶⁹. MnuA est capable de digérer des ADNs linéaires ou circulaires et nécessite pour être active des cations divalents comme le Ca²⁺ et le Mg²⁺,^{269,315} présents dans les tampons et solutions du protocole. A noter que ce gène est aussi absent chez les espèces du cluster mycoïdes.
- la protéine MGA_0676³¹⁶ (souche Rlow) ou GSW_RS00180 (souche S6) identifiée chez cet organisme comme une nucléase cytotoxique dépendante du Ca²⁺.³¹⁶ Cette protéine membranaire localisée à la surface des cellules possède un domaine « SNC » retrouvée dans la nucléase de *Staphylococcus aureus*.

Afin d'évaluer dans quelle mesure, une de ces nucléases avait pu inhiber notre expérience de clonage, nous avons travaillé avec différents mutants. Un « mutant Cas9 »

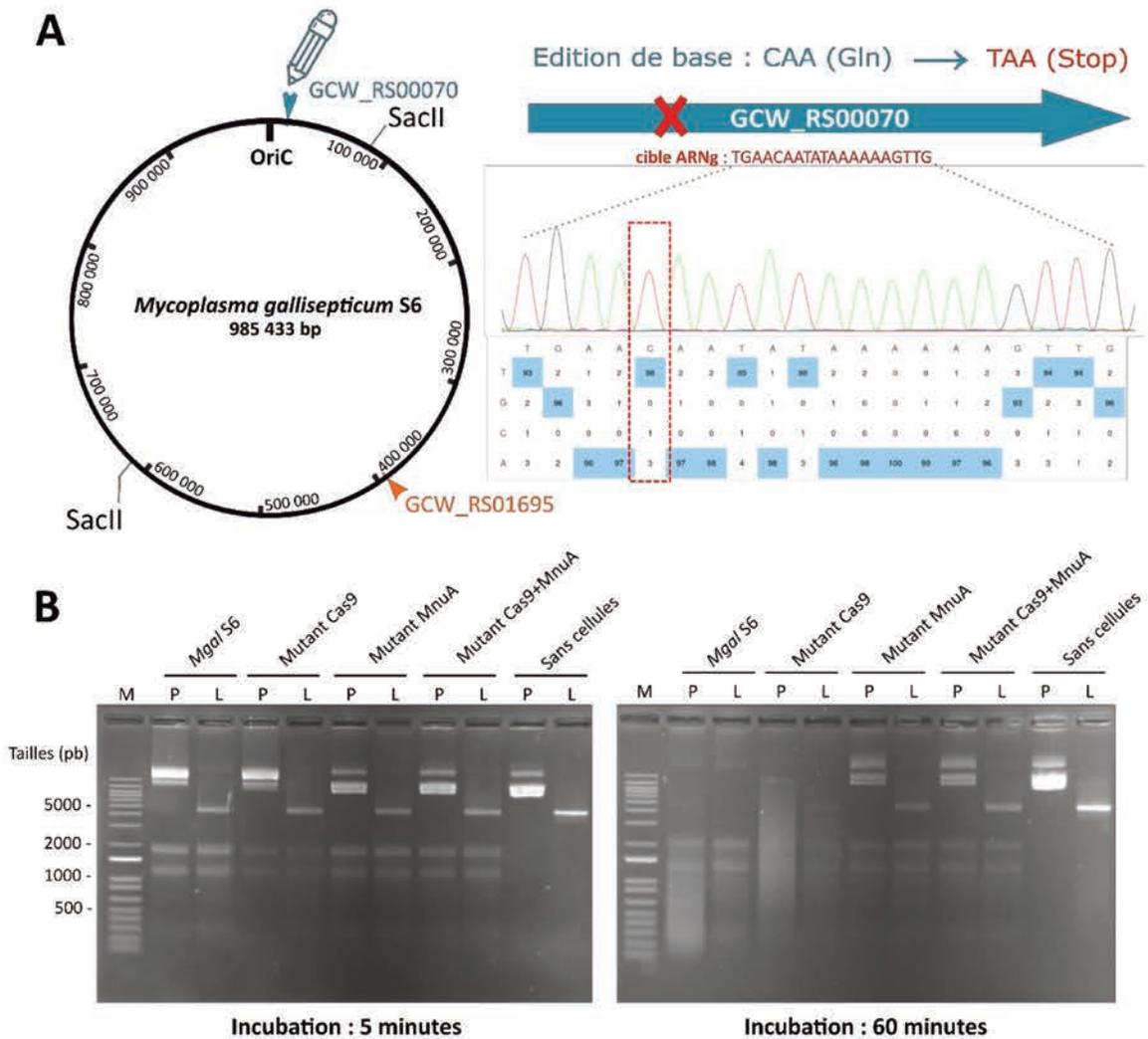


Figure 65 : Inactivation de la protéine homologue à MnuA chez *M. gallisepticum* S6. [A] La carte du génome de *Mgal* S6 est présentée avec le *locus* cible pour la CReasPy-Fusion GCW_RS01695 (flèche orange), mais aussi avec le gène GCW_RS00070 homologue de *mnuA* (flèche colorée en turquoise). L'édition de base avec l'outil CRISPR désaminase est indiquée par le crayon turquoise : l'ARNg choisi (indiqué en rouge) permet d'éditer le codon glycine CAA en codon stop TAA (ou UAA), entraînant l'inactivation du gène cible. Le chromatogramme issu de l'analyse par séquençage sanger, associé avec le logiciel EditR (https://moriaritylab.shinyapps.io/editr_v10/), permet de visualiser la proportion des bases A, T, C et G de chaque pic. La base éditée, c'est-à-dire, un « pic de T » à la place d'un « pic de C » est indiquée avec un rectangle en pointillés rouge. **[B]** Les résultats du test d'activité de la nucléase sont représentés pour la souche *Mgal* S6 WT et les mutants Cas9, MnuA ou Cas9+MnuA, au bout de 5 minutes d'incubation (à gauche) et de 60 minutes d'incubation (à droite). "M": marqueur de taille DNA Ladder 1kb+ Invitrogen (100-12000 bp) ; "P": ADN plasmidique ; "L": ADN linéaire (pansement).

obtenu spontanément au laboratoire (décalage du cadre de lecture au niveau de l'acide aminé 17 de la séquence protéique cf. Thèse Thomas Ipoutcha donnée non publiée), un « mutant MnuA » dont le gène a été inactivé grâce à l'outil d'édition de base CRISPR désaminase, et un double « mutant Cas9+MnuA », issu du mutant Cas9 dans lequel le gène *mnuA* a été inactivé de la même manière que le « mutant MnuA ». Nous avons fait le choix de ne pas inclure de « mutant MGA_0676 » dans un premier temps car (i) une étude menée sur *M. bovis* PG45 avait montré que parmi tous les gènes codant des nucléases chez cette espèce (3 au total), *mnuA* était responsable de la majorité de l'activité nucléasique détectable *in vitro*²⁶⁹ et (ii) nous n'étions pas certains de pouvoir construire un triple mutant dans le temps imparti.

L'outil d'édition de base CRISPR (en anglais CRISPR Base Editor ou CBE) utilisé au cours de cette thèse est issu des travaux du Dr. Thomas Ipoutcha²⁶³, qui a contribué à développer pour les mycoplasmes cette méthodologie de modification de l'ADN sur des bases spécifiques^{317,318}. Ainsi, la nucléase MnuA a été inactivée en suivant le protocole décrit paragraphe IV.2. du Matériel et Méthode. Ce système fonctionne avec une Cas9 inactivée (et donc incapable de cliver l'ADN), fusionnée avec une cytosine désaminase. Dans cette expérience, le sgRNA a été conçu de façon à cibler spécifiquement le *locus* GCW_RS00070 (MnuA), et à modifier le codon CAA (Glutamine) en codon stop TAA par désamination de la base (**Figure 65A**). Le chromatogramme issu de la vérification par séquençage Sanger indique avec le pic rouge (correspondant à un T) que la modification de la base a été obtenue avec succès. L'analyse avec le logiciel EditR a permis d'apprécier une proportion de 96% de base thymine (T) contre 1% de base cytosine (C). Nous avons ensuite entrepris de vérifier le phénotype de nos mutants (**Figure 65B**) en réalisant un test nucléase, suivant le procédé décrit par Sharma *et al.* 2015²⁶⁹ (cf. §IV.3. Matériels et Méthodes). Les cellules de *Mgal* (wild-type (WT) et mutants) ont été incubées avec le plasmide pMT85, ou bien avec le pansement de recombinaison (ARS-CEN-HIS3-CysP). Au bout de 5 minutes d'incubation à 37°C, peu de différence était observée entre les souches testées par comparaison avec le témoin « sans cellules ». La présence de deux bandes surnuméraires (vers 1200 pb et 1800 pb) dans toutes les conditions avec les cellules de *Mgal*, reflétait probablement un artéfact d'ADNg ou de cellules. En revanche au bout d'une heure d'incubation, les plasmides ou pansements mis en contact avec *Mgal* S6 ou le « mutant Cas9 » étaient totalement dégradés ; tandis que ceux incubés avec le « mutant MnuA » ou le double « mutant Cas9+MnuA », restaient intacts avec

Tableau 39 : Essai n°2 de clonage du génome de *M. gallisepticum* S6 mutant Cas9+MnuA par CReasPy-Fusion

| | Condition expérimentale | Colonies (UFC) | Clones positifs / Clones analysés ^a | |
|--------------|---|----------------|--|----------------|
| | | | PCR simplexe | PCR multiplexe |
| Contrôles | Plasmide seul | >1000 | - | - |
| | Plasmide + <i>Mgal</i> 49°C | >1000 | - | - |
| | Plasmide + <i>Mgal</i> 49°C | >1000 | - | - |
| 1 | Resuspension Buffer | 86 | 2/10 | 2/2 |
| 2 | Resuspension Buffer + EDTA | 97 | 0/9 | - |
| 3 | Resuspension Buffer + ADNsb dénaturé | 1 | - | - |
| 4 | Resuspension Buffer + traitement filtre-aiguille | 101 | 0/10 | - |
| 5 | Resuspension Buffer – 49°C | 106 | 2/10 | 2/2 |
| 6 | Resuspension Buffer + EDTA – 49°C | 92 | 2/9 | 2/2 |
| 7 | Resuspension Buffer + ADNsb dénaturé – 49°C | 4 | 0/3 | - |
| 8 | Resuspension Buffer + traitement filtre-aiguille – 49°C | 120 | 1/10 | 1/1 |
| 9 | HBSS | 65 | 3/10 | 3/3 |
| 10 | HBSS + EDTA | 62 | 0/7 | - |
| 11 | HBSS + ADNsb dénaturé | 4 | 0/4 | - |
| 12 | HBSS + traitement filtre-aiguille | 114 | 0/9 | - |
| 13 | HBSS – 49°C | 45 | 1/6 | 1/1 |
| 14 | HBSS + EDTA – 49°C | 48 | 0/7 | - |
| 15 | HBSS + ADNsb dénaturé – 49°C | 4 | 0/3 | - |
| 16 | HBSS + traitement filtre-aiguille – 49°C | 87 | 0/10 | - |
| Total | - | 1036 | 11/117 | 11/117 |

^aLe nombre de transformants levures analysés par PCR et PCR multiplexe sont indiqués, ainsi que le nombre de clones positifs obtenus.

un profil identique au témoin sans cellules. Nous pouvions donc conclure que l'inactivation de la nucléase MnuA avait été effectuée avec succès, et que les mutants pouvaient donc être utilisés pour de nouvelles expériences.

Au cours de l'expérience n°2, nous avons seulement travaillé avec le double « mutant Cas9+MnuA ». Ce choix était principalement motivé par le fait que nous souhaitions d'abord nous affranchir totalement des deux nucléases avant de déterminer si une des deux était plus gênante que l'autre. N'ayant pu déterminer quel pARNg était le plus efficace, nous avons travaillé à nouveau avec les deux souches VL6-48N-Cas9-ARNg1 et VL6-48N-Cas9-ARNg2. Nous avons également choisi de garder seulement le pansement CEN-HIS-CysP (sans ARS), puisque ce dernier permettait de diminuer le nombre de colonies faussement positives. Nous avons également souhaité tester différentes conditions en jouant sur certains paramètres afin d'évaluer si l'on pouvait améliorer le protocole (**Tableau 39**) :

- Ajout ou pas d'EDTA dans le tampon de resuspension des cellules *Mgal* : le but étant de chélater les cations divalents cofacteurs d'enzymes comme MGA_0676 (GSW_RS00180 dans la souche S6) ou d'autres enzymes non identifiées.
- Utilisation d'un tampon autre que le « Resuspension Buffer » issu du protocole original (clonage par Fusion). Nous avons choisi de travailler avec le tampon HBSS 1X dépourvu de certains ions (*e.g.* Ca²⁺, Mg²⁺), et utilisé dans le protocole de transformation de *Mgal*³¹⁹.
- Utilisation d'ADN « cargo » simple brin comme de l'ADN de sperme de saumon dénaturé, dans le but de « camoufler » notre pansement et d'éventuellement le protéger contre des nucléases.
- Dispersion des agrégats cellulaires de la culture de *Mgal*. Ce mycoplasme est capable de produire en fin de phase exponentielle des agrégats voire des biofilms³²⁰. Cette étape est souvent effectuée au laboratoire pour la transformation de *Mgal* et la culture est travaillée à l'aide d'une seringue avec aiguille et passée à travers un filtre 0,45 µm.
- Préchauffage ou pas des cellules de mycoplasmes à 49°C avant Fusion. Karas *et al.* indiquent dans leur protocole qu'il est conseillé de passer par cette étape à 49°C pour inactiver certaines protéines (ex : nucléases).

Tableau 40 : Résumé des quatre expériences effectuées pour cloner le génome de *M. gallisepticum* par CREasPy-Fusion.

| Souches de <i>Mycoplasma gallisepticum</i> | UFC | Clones positifs / Clones analysés ^a | | |
|---|------|--|----------------|-------------------|
| | | PCR simplexe | PCR multiplexe | PFGE ^b |
| 1^{ère} expérience | | | | |
| <i>M. gallisepticum</i> S6 | 256 | 0/108 | - | - |
| 2^{ème} expérience | | | | |
| <i>M. gallisepticum</i> S6 mutant Cas9+MnuA | 1036 | 12/117 | 11/12 | 2/2 |
| 3^{ème} expérience | | | | |
| <i>M. gallisepticum</i> S6 | 134 | 0/25 | - | - |
| <i>M. gallisepticum</i> S6 mutant Cas9 | 82 | 0/27 | - | - |
| <i>M. gallisepticum</i> S6 mutant MnuA | 115 | 3/34 | 2/3 | 1/1 |
| <i>M. gallisepticum</i> S6 mutant Cas9+MnuA | 84 | 7/27 | 4/7 | 1/1 |
| 4^{ème} expérience | | | | |
| <i>M. gallisepticum</i> S6 | 298 | 1/16 | 1/1 | 1/1 |
| <i>M. gallisepticum</i> S6 mutant Cas9 | 412 | 0/20 | - | - |
| <i>M. gallisepticum</i> S6 mutant MnuA | 416 | 1/15 | 1/1 | 1/1 |
| <i>M. gallisepticum</i> S6 mutant Cas9+MnuA | 292 | 7/15 | 6/7 | 1/1 |

^aLe nombre de transformants levures analysés par PCR, PCR multiplexe et PFGE sont indiqués, ainsi que le nombre de clones positifs obtenus.

^bLes analyses par PFGE n'ont pas été effectuées sur l'ensemble des clones positifs en PCR multiplex mais seulement sur un échantillon.

Les résultats obtenus pour chaque condition sont résumés **Tableau 39**. Notre premier réflexe a été de vérifier les contrôles « plasmide » et « plasmide + mycoplasmes ». Cette fois, aucune différence n'a été notée entre ces derniers, révélant que le « mutant Cas9+MnuA », ne semblait pas gêner la transformation du plasmide dans les levures. Au total, 1036 colonies ont été obtenues toutes conditions confondues. Ce nombre variait entre 1 et 120 par condition. Nous pouvions aussi observer que les conditions correspondant au nombre de colonies le plus faible, étaient celles dans lesquelles nous avons rajouté l'ADNs_b dénaturé. Nous avons repiqué 10 colonies par condition quand cela était possible. En prenant en compte que certains transformants repiqués n'ont pas repoussé, nous avons analysé au total 117 colonies. Après extraction de l'ADNg total et analyse par PCR simplexe et PCR multiplexe, nous avons pu identifier 11 transformants sur les 117 propageant un génome de *Mgal* potentiellement intègre. Nous avons pu valider la taille de ce génome cloné dans deux transformants sélectionnés pour analyse par PFGE (cl 6.9 et 10.5 **Figure 66**). Au final, si nous n'avons pas pu identifier de conditions améliorant la méthode, ce résultat fut cependant crucial pour la suite de nos expériences, puisqu'à l'évidence une des deux nucléases mises en cause gênait très probablement le clonage du génome *Mgal* dans la levure en CReasPy-Fusion.

Nous avons donc répété ces expériences deux fois en incluant, en plus du double mutant *Mgal*-Cas9+MnuA, la souche *Mgal* WT ainsi que les simples mutants *Mgal*-Cas9 et *Mgal*-MnuA (expériences n°3 et n°4 résumées dans le **Tableau 40**). Pour chacune de ces expériences, le nombre d'UFC obtenu au total est indiqué, ainsi que les clones positifs validés par PCR simplexe, PCR multiplexe et par PFGE. Au cours de l'expérience n°3 nous avons réussi à cloner le génome de *Mgal* en provenance du « mutant MnuA » et du double mutant « Cas9+MnuA ». Pour ces deux conditions, nous avons obtenu respectivement : (i) 3 clones positifs sur 34, et, 7 sur 27 en PCR simplexe, puis (ii) 2 clones sur 3, et, 4 sur 7 en PCR multiplexe. Finalement, les deux clones sélectionnés (un pour chaque condition) et analysés en PFGE ont été validés avec un profil à deux bandes à la taille attendue après digestion avec l'enzyme SacII (536 kpb et 449 kpb), tout comme le témoin positif (**Figure 66**). En revanche, aucun des clones analysés pour la souche *Mgal* WT et le mutant « Cas9 » se sont révélés positifs et ceci dès l'analyse en PCR simplexe. Après cette expérience, il semblait donc que l'enzyme potentiellement responsable de notre difficulté à cloner le génome *Mgal* dans la levure, soit la nucléase de surface MnuA, plutôt que la protéine Cas9.

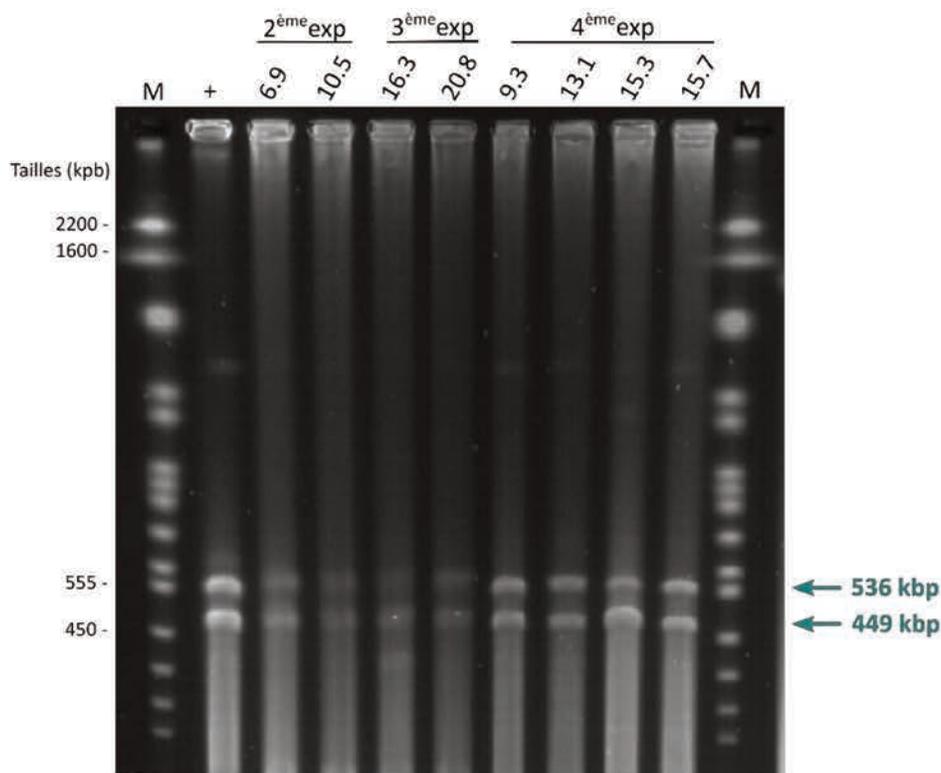


Figure 66 : Criblage par PFGE des transformants levures après clonage du génome *M. gallisepticum* S6 par CREasPy-Fusion. Cette figure regroupe l'analyse de divers transformants levures issus des expériences 2, 3 et 4 de CREasPy-Fusion. "M" : *S. cerevisiae* chromosomal DNA ladder Bio-Rad (225-2200 kpb). "+" : témoin positif avec l'ADNg de *Mgal* digéré avec *SacII*. La taille du génome de *Mgal* S6 (WT ou mutant) cloné est validée après clonage dans la levure par restriction enzymatique et analyse PFGE. La coupure par l'enzyme *SacII* permet d'hydrolyser le génome en deux fragments linéaires à 536 et 449 kpb.

Tableau 41 : Bilan du clonage des génomes de *M. bovis* PG45 et *M. agalactiae* PG2.

| Mycoplasme | Clones positifs / Clones analysés ^a | | |
|--------------------------------------|--|----------------|------|
| | PCR simplexe | PCR multiplexe | PFGE |
| <i>M. bovis</i> PG45 | 0/92 | - | - |
| <i>M. bovis</i> PG45 mutant MnuA | 0/80 | - | - |
| <i>M. agalactiae</i> PG2 | 2/43 | 2/2 | 2/2 |
| <i>M. agalactiae</i> PG2 mutant MnuA | 5/80 | 3/5 | 1/3 |

^aRésultats issus d'une expérience de clonage pour chaque espèce. Les résultats obtenus sont résumés dans ce tableau. Le nombre de transformants levures analysés par PCR, PCR multiplexe et PFGE sont indiqués, ainsi que le nombre de clones positifs obtenus.

Toutefois, la répétition du protocole au cours de l'expérience n°4 nous a forcé à nuancer cette première conclusion. En effet, nous avons obtenu des clones positifs pour le « mutant MnuA » et le double « mutant Cas9+MnuA » comme précédemment, mais aussi pour la souche WT de *Mgal* pour laquelle un clone (cl 9.3) sur les 16 clones criblés, a été validé jusqu'à l'analyse finale par PFGE (**Tableau 40 et Figure 66**). De plus, un seul clone (cl 13.1) sur les 15 analysés à partir du « mutant MnuA » a été validé comme positif. Nous ne pouvions donc pas à ce stade conclure que la délétion de la nucléase MnuA débloquent totalement la méthode. En revanche, nous pouvions dire que l'inactivation de MnuA combinée à celle de la protéine Cas9, permettait d'obtenir un plus grand nombre de transformants levure propageant le génome de *Mgal*. De plus, ces résultats validaient que nous étions capables d'étendre la CReasPy-Fusion à un autre groupe phylogénétique que celui des Spiroplasma. Enfin, nous décrivions ici, pour la première fois, le clonage du génome de *Mgal* souche S6 dans la levure comme un plasmide centromérique surnuméraire.

1.5. Extension de la méthode à d'autres espèces de mycoplasmes appartenant au groupe phylogénétique Hominis

Pour finir d'asseoir la méthode et tenter de l'étendre à un panel encore plus large de mycoplasmes, nous avons décidé de l'appliquer à deux espèces pathogènes de ruminants : *Mycoplasma bovis* PG45 et *Mycoplasma agalactiae* PG2. Ces espèces appartiennent au groupe phylogénétique Hominis (**Figure 59**) et tout comme pour *Mgal*, les génomes de ces dernières n'avaient jamais été clonés dans la levure jusqu'alors. Ce sont des pathogènes de bovins (*M. bovis*) et de caprins (*M. agalactiae*) qui ont un impact sanitaire et économique sur l'élevage^{321,322}. Il était intéressant d'estimer la faisabilité des méthodes de biologie de synthèse maîtrisées au laboratoire sur ces deux espèces d'intérêt vétérinaire, d'autant plus que de nouveaux projets démarraient dans l'équipe pour l'étude de ces espèces.

En nous basant sur les résultats obtenus sur *Mgal* (cf. §II.1.4.), nous avons anticipé le plan des expériences en sélectionnant pour chaque espèce la souche WT, mais aussi, un « mutant MnuA ». En effet, comme décrit plus haut, cette nucléase membranaire a été caractérisée chez *M. bovis*, et des homologues sont retrouvés chez d'autres espèces comme *Mgal*, mais aussi *Maga*. Nous avons à disposition un « mutant MnuA » de *Mbov* issu des

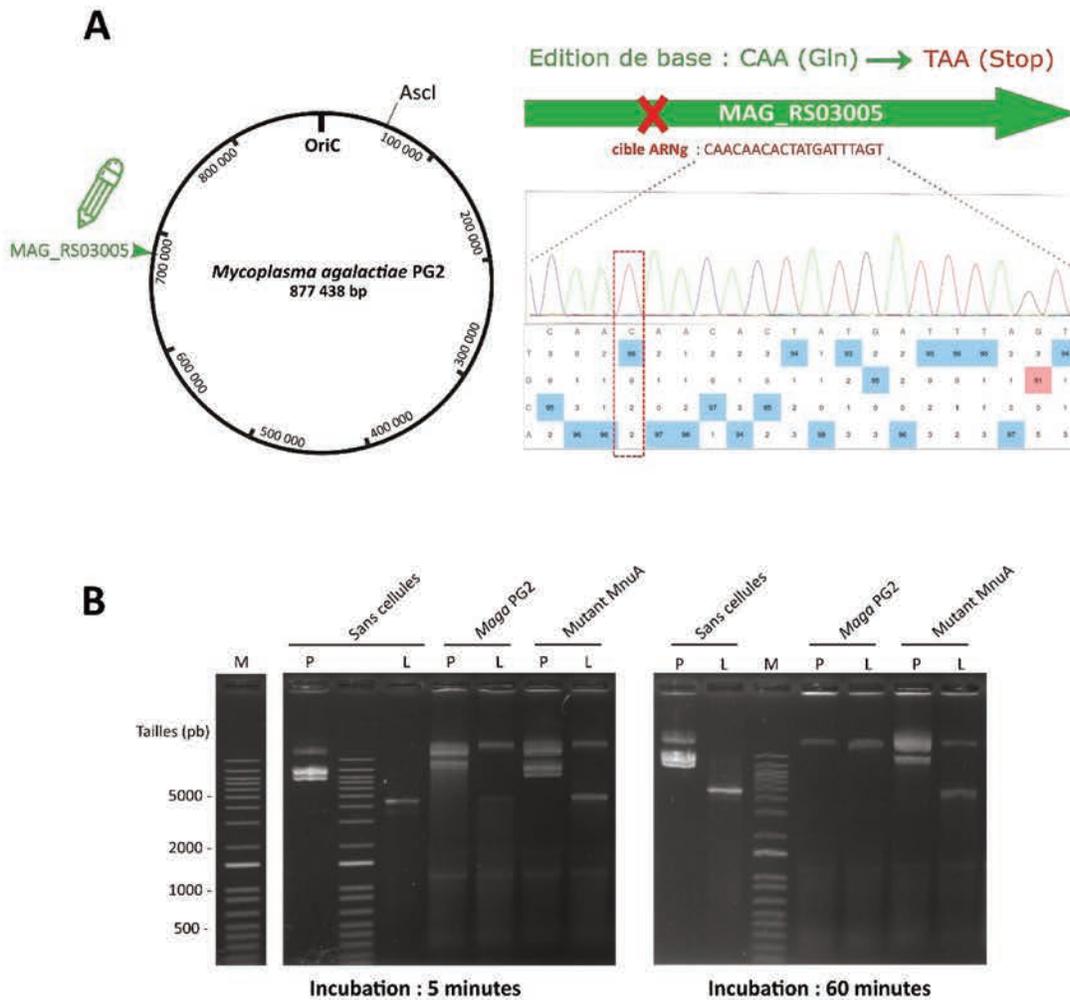


Figure 67 : Inactivation de la protéine homologue à MnuA chez *M. agalactiae* PG2. [A] La carte du génome de *Maga* PG2 est présentée avec le *locus* cible pour la CReasPy-Fusion MAG_RS03005 (flèche verte). L'édition de base avec l'outil CRISPR désaminase est indiquée par le crayon vert sur cette même cible. L'ARNg choisi (indiqué par une croix rouge) permet d'éditer le codon glycine CAA en codon stop TAA (ou UAA), permettant l'inactivation du gène cible. Le chromatogramme issu de l'analyse par séquençage sanger, associé avec le logiciel EditR (https://moriaritylab.shinyapps.io/editr_v10/), permet de visualiser la proportion des bases A, T, C et G de chaque pic. La base éditée, c'est-à-dire, un « pic de T » à la place d'un « pic de C » est indiquée avec un rectangle en pointillés rouge. [B] Les résultats du test d'activité de la nucléase sont représentés pour la souche *Maga* PG2 WT et le mutant MnuA, au bout de 5 minutes d'incubation (à gauche) et de 60 minutes d'incubation (à droite). "M": marqueur de taille DNA Ladder 1 kb+ Invitrogen (100-12000 bp) ; "P": ADN plasmidique ; "L": ADN linéaire (pansement).

travaux effectués au laboratoire pour la mise en place de la méthode d'ingénierie par édition de base²⁶³. En revanche, le « mutant MnuA » pour *M. agalactiae*, a été spécifiquement construit pour nos expériences et le développement de la CREasPy-Fusion. Tout comme pour *Mgal* et décrit paragraphe II.1.4, le « design » du sgRNA a été effectué de façon à cibler spécifiquement le *locus* MAG_RS03005 (homologue de MnuA) et à modifier le codon CAA (Glutamine) en codon stop TAA (**Figure 67A**). Le chromatogramme issu du séquençage Sanger nous a permis de visualiser un pic rouge (correspondant à un T) plutôt qu'un pic violet (correspondant à un C). La modification de la base a donc été obtenue avec succès (96% de base thymine (T) contre 2% de base cytosine (C) visualisé avec le logiciel EditR).

Nous avons ensuite vérifié le phénotype du mutant produit (**Figure 67B**) en exécutant à nouveau le test nucléase²⁶⁹. Les cellules de *Maga* (WT ou mutant) ont été incubées avec un plasmide, ou bien, avec un des pansements de recombinaison (ARS-CEN-HIS3). Au bout de 5 minutes d'incubation à 37°C, nous observions déjà une différence entre la souche WT et le « mutant MnuA », puisque le pansement (condition L sur le gel) était déjà dégradé chez le premier. Au bout d'une heure d'incubation, la différence s'était accentuée puisque seul le « mutant MnuA » arborait un profil similaire au témoin « sans cellules », alors que, la souche WT avait dégradé l'ADN circulaire ou linéaire co-incubé. Nous pouvions donc valider avec le phénotype observé ici, que la nucléase MnuA avait été inactivée chez le mutant.

En amont des expériences de CREasPy-Fusion, nous avons produit, pour ces deux espèces, les pARNg nécessaires pour guider le clivage de Cas9 dans la levure. Les *loci* MBOVPG45_0215 et MAG_RS03005 codant la nucléase MnuA, ont été ciblés pour *M. bovis* PG45 et *M. agalactiae* PG2 respectivement. L'idée était d'éliminer totalement ces gènes au moment du clonage, même si pour les « mutants MnuA » ces derniers avaient été inactivés (mutations ponctuelles). Deux expériences indépendantes ont été effectuées pour chaque espèce. Les sphéroplastés de levures VL6-48N-pCas9-pARNg ont été fusionnés avec des cellules de *Mbov* souche type PG45, ou avec des cellules de *Maga* souche type PG2. Les résultats obtenus sont résumés dans le **Tableau 41**.

L'expérience effectuée pour tenter de cloner le génome de *M. bovis* n'a pas permis de valider des transformants propageant ce génome. Que ce soit avec la souche WT ou bien avec le « mutant MnuA », sur les 92 et 80 clones analysés par PCR simplexe respectivement, nous n'avons pas détecté de signal correspondant à la séquence du génome.

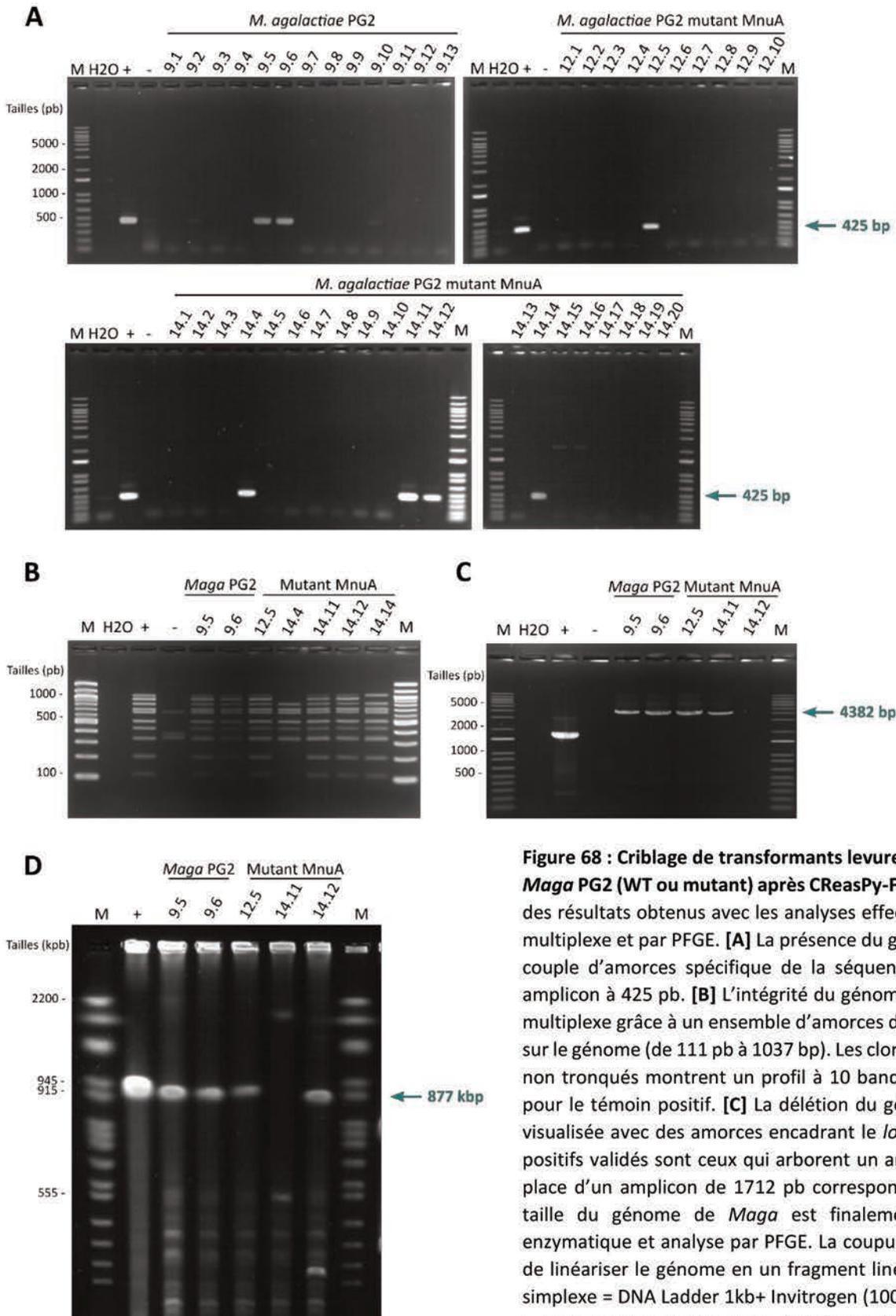


Figure 68 : Criblage de transformants levures propageant le génome de *Maga* PG2 (WT ou mutant) après CREasPy-Fusion. Cette figure regroupe des résultats obtenus avec les analyses effectuées par PCR simplex, PCR multiplexe et par PFGE. **[A]** La présence du génome est détectée avec un couple d'amorces spécifique de la séquence de *Maga* produisant un amplicon à 425 pb. **[B]** L'intégrité du génome cloné est évaluée par PCR multiplexe grâce à un ensemble d'amorces distribué de façon homogène sur le génome (de 111 pb à 1037 pb). Les clones propageant des génomes non tronqués montrent un profil à 10 bandes identique à celui obtenu pour le témoin positif. **[C]** La délétion du gène cible MAG_RS03005 est visualisée avec des amorces encadrant le *locus* cible. Les transformants positifs validés sont ceux qui arborent un amplicon PCR à 4382 pb, à la place d'un amplicon de 1712 pb correspondant au *locus* intact. **[D]** La taille du génome de *Maga* est finalement validée par restriction enzymatique et analyse par PFGE. La coupure par l'enzyme *Ascl* permet de linéariser le génome en un fragment linéaire de 877 kpb. "M" : PCR simplex = DNA Ladder 1kb+ Invitrogen (100-12000 bp) ; PCR multiplexe : DNA Ladder 100bp NEB (100-1517 bp) ; PFGE = *S. cerevisiae* chromosomal DNA ladder Bio-Rad (225-2200 kpb). "H2O" : témoin négatif sans ADN. "+" : témoin positif avec l'ADNg de *Maga*. "-" : témoin négatif levure avec l'ADNg de *S. cerevisiae* VL6-48N.

En revanche, les résultats obtenus pour *M. agalactiae* valident la méthode de CReasPy-Fusion car des transformants levures positifs ont été obtenus (**Tableau 8**). Nous n'avons pas observé de différence entre la souche WT et le « mutant MnuA ». Un peu plus d'une centaine de colonies a été obtenue sur boîtes (43 avec la WT et 80 pour le mutant). La totalité a été repiquée et analysée. L'analyse par PCR simplexe avec le couple d'amorces MagaL_4F/4R, a permis d'obtenir une bande à 425 pb identique au témoin positif pour 2 transformants (cl 9.5 et 9.6) sur 43 pour la souche WT, et 5 transformants (cl 12.5, 14.4, 14.11, 14.12 et 14.14) sur 80 pour le « mutant MnuA » (**Figure 68A**).

Ces clones ont ensuite été analysés par PCR multiplexe avec un ensemble de 10 couples d'amorces répartis sur l'ensemble du génome *Maga*. Pour la souche WT, les 2 clones arboraient un profil de PCR multiplexe complet à l'instar du témoin positif, tandis que pour le « mutant MnuA », seuls 3 clones (cl 12.5, 14.11 et 14.12) ont pu être validés avec un profil correct (**Figure 68B**). Sur ces cinq clones, nous avons vérifié la délétion du *locus* cible MAG_RS03005 avec les amorces Magalactiae_5890_F/Magalactiae_5910_R qui encadrent le gène. Les quatre clones 9.5, 9.6, 12.5 et 14.11 affichaient des profils corrects avec une bande correspondant à l'insertion des éléments levure en lieu et place du *locus* cible (soit 4382 pb contre 1712 pb pour le *locus* WT) (**Figure 68C**). Ces mêmes cinq clones ont été analysés en PFGE. Seuls les transformants levure 9.5, 9.6 et 12.5 ont été validés, le profil de migration correspondant à la taille attendue après digestion et linéarisation du génome par *Ascl* (877 kpb). Le clone 14.11 arborait un profil dégradé et aucune bande à la taille attendue ; le cl 14.12 affichait une bande surnuméraire aux alentours de 300 kpb (**Figure 68D**).

1.6. Bilan et analyses des expériences de CReasPy-Fusion appliquées aux mycoplasmes.

Dans les parties précédentes, nous avons vu que la méthode de CReasPy-Fusion pouvait être appliquée à un certain nombre d'espèces différentes de mycoplasmes avec des efficacités qui pouvaient varier fortement d'une espèce à l'autre. Au total, le génome de 7 souches sur les 10 répertoriées dans le **Tableau 42** ont pu être clonés dans la levure et 3 ont échoué.

Tableau 42 : Bilan des expériences de clonage par CReasPy-Fusion avec plusieurs génomes de mycoplasmes.

| Groupe phylogénétique | Mycoplasme | Souche | Taille Génome | Gène cible | Système de Restriction-Modification * | Homologue MnuA | Clonage Levure |
|--|---|------------------|---|--|---------------------------------------|-----------------------------|----------------|
| Spiroplasma (Cluster mycoides) | <i>M. capricolium</i> subsp <i>capricolium</i> (<i>Mcap</i>) | 27343 (CK) | 1,01 Mb | CCATC restrictase (MCAP_RS00270) | 3 avec 1 restrictase | - | ✓ |
| | | 14232 | 1,03 Mb | Peptidase S41 (Mcapa_s03_0930) | 7 avec 3 restrictases | - | ✗ |
| | <i>M. mycoides</i> subsp <i>capri</i> (<i>Mmc</i>) | GM12 | 1,08 Mb | CCATC restrictase (MMCAP2_RS00520) | 8 avec 6 restrictases | - | ✓ |
| | | 95010 | 1,15 Mb | CCATC restrictase (MLC_0920) | 7 avec 6 restrictases | - | ✗ |
| | <i>M. capricolium</i> subsp <i>capripneumoniae</i> (<i>Mccp</i>) | 95043 (Niger) | 1,01 Mb | Peptidase S41 FOY67_01295 | 5 avec 3 restrictases | - | ✓ |
| | | 14020 (Tanzanie) | IND | IND | IND | - | ✓ |
| <i>M. mycoides</i> subsp <i>mycoides</i> (<i>Mmm</i>) | T1/44 | 1,18 Mb | <i>glpOKF</i> (MSCT144_RS04705/710/715) | 5 avec 3 restrictases | - | ✓ | |
| Pneumoniae | <i>M. gallisepticum</i> (<i>Mgal</i>) | S6 | 0,98 Mb | <i>cysP</i> (GCW_RS01695) | 2 avec 2 restrictases | GCW_RS00070 (<i>mnuA</i>) | ✓ |
| Hominis | <i>M. bovis</i> (<i>Mbov</i>) <i>M. agalactiae</i> (<i>Maga</i>) | PG45 | 1,00 Mb | <i>mnuA</i> (MBOVPG45_0215) | 10 avec 8 restrictases | MBOVPG45_0215 | ✗ |
| | | PG2 | 0,87 Mb | Homologue de <i>mnuA</i> (MAG_RS03005) | 7 avec 1 restrictase + 2 potentielles | MAG_RS03005 | ✓ |

*Nombre de systèmes de restriction et de restrictases identifiées d'après Rebase Database et Molligen 3.0 et 4.0.

IND = indéterminé.

Ainsi, les génomes de *Mcap* CK et *Mmc* GM12 déjà clonés dans la levure avec d'autres stratégies (Classique, Fusion) ont pu être transférés depuis des cellules de mycoplasmes vers la levure et simultanément modifiés avec l'outil CRISPR-Cas9. Malgré des efficacités plutôt faibles, les levures hébergeaient bel et bien des génomes intacts. En effet, les génomes de *Mcap* et *Mmc* clonés et modifiés ont par la suite été transplantés dans des cellules receveuses de mycoplasmes afin de vérifier qu'ils ne contenaient pas de mutations majeures pouvant affecter la viabilité des cellules. Des transplants ont été obtenus pour chaque espèce et ont également été validés par multiplexe PCR. Des analyses par séquençage NGS sont actuellement en cours pour valider entièrement ces clones. Si les séquences des transplants sont correctes, nous pourrions alors valider (i) le cycle complet d'ingénierie de génome bactérien chez la levure, et (ii) l'utilisation de méthode de CREasPy-Fusion pour cloner des génomes bactériens dans la levure.

Nous avons ensuite réussi à cloner le génome de deux souches de terrain de l'espèce *Mccp* (95043 et 14020) et celui de la souche vaccinale *Mmm* T1/44. Si le clonage de la souche *Mccp* Abomsa avec la méthode de clonage CREasPy-Cloning (Gourgues *et al.*, non publié), avait déjà permis de montrer qu'il était possible de cloner le génome de cette espèce dans la levure, nous avons pu montrer ici que la stratégie de CREasPy-Fusion permettait d'obtenir le même résultat, en utilisant deux autres souches (*Mccp* 95043 et *Mccp* 14020). De même que pour *Mcap* CK et *Mmc* GM12, ces génomes ont été transplantés dans ces cellules receveuses et des transplants ont été obtenus. Les analyses NGS sont en cours. Les expériences menées avec *Mmm* T1/44 ont permis, quant à elles, de montrer que l'on pouvait cloner non seulement son génome entier mais aussi capturer une moitié de celui-ci, après coupure avec deux ARNg et circularisation avec la matrice de recombinaison fournie. Tout comme pour la méthode CREasPy-Cloning, il est donc possible de cibler deux *loci* à la fois en CREasPy-Fusion. Les résultats obtenus avec le génome *Mmm* T1/44 de 1,2 Mb sont de loin les meilleurs parmi ceux produits au cours de cette étude. Les génomes entiers n'ont pas été transplantés dans une cellule receveuse pour obtenir des mutants car à l'heure actuelle, cette étape n'a jamais fonctionné chez cette espèce²³⁰.

Nous avons ensuite étendu la méthode à deux groupes phylogénétiques Pneumoniae et Hominis en décrivant, pour la première fois, les clonages des génomes de *Mgal* S6 et de *Maga* PG2 dans la levure. Au cours des expériences avec *Mgal* nous avons montré que

l'inactivation des nucléases Cas9 et MnuA augmentait le nombre de transformants hébergeant le génome d'intérêt. En revanche, pour *Maga* nous n'avons pas observé d'effet particulier en l'absence de MnuA. Nous avons émis des hypothèses pour expliquer ces résultats :

(i) MnuA est une nucléase présente à la surface de quelques espèces de mycoplasmes qui pourrait, en présence des cofacteurs adéquats, dégrader le pansement avant son entrée dans les levures et prévenir la réaction de CReasPy-Fusion. L'alignement des séquences en acides aminés des protéines MnuA indique que MnuA-*Mbov* et MnuA-*Maga* sont relativement bien conservés (75% d'identité) mais différent de MnuA-*Mgal* (30% d'identité). Ceci suffit peut-être à expliquer la différence observée entre *Mgal* et *Maga* lors de la CReasPy-Fusion ; les protéines MnuA ne nécessitant peut-être pas toujours les mêmes conditions pour être fonctionnelles. C'est le cas par exemple pour Mpn491 (MnuA de *Mycoplasma pneumoniae*) qui requiert des ions Mg^{2+} et MBOVPG45_0245 (MnuA de *M. bovis* PG45) qui requiert des ions Ca^{2+} et Mg^{2+} pour être fonctionnelle³¹⁵. De même, certaines MnuA sont potentiellement plus actives sur certains substrats que sur d'autres (*e.g.* ADN circulaire double brin, ADN circulaire simple brin, ADN linéaire double brin, ADN linéaire simple brin). Enfin, il est possible que ces enzymes soient plus ou moins exprimées ou actives en fonction de la phase de croissance. Dans nos expériences, nous avons tenté de travailler avec des cultures en milieu ou fin de phase exponentielle (mesure du pH). Mais ce point n'a pas toujours été bien contrôlé (Expérience n°3 *Mgal* pH entre 6,5 et 6,7 vs Expérience n°4 *Mgal* 6,8 et 7,10) et devra être mieux documenté.

(ii) Cas9 est une endonucléase présente dans *Mgal* et absente de *Maga* qui pourrait, de concert avec un ARNg plus ou moins spécifique, cliver le génome de la levure ou le pansement, et ainsi empêcher la réaction de CReasPy-Fusion. De même, nous ne savons pas si cette enzyme est exprimée de façon constitutive et active quel que soit l'état physiologique des cellules.

(iii) Outre le rôle possible des nucléases MnuA, d'autres nucléases (ex : MGA_0676³¹⁶/GSW_RS00180 chez *Mgal*, MAG_5040 chez *Maga*³²³, MBOVPG45_0310 ou MBOVPG45_0089 chez *M.bovis*²⁶⁹), ou encore, des facteurs non identifiés pourraient également impacter la CReasPy-Fusion. Dans une publication récente (2019), le Dr. Bogumil Karas et ses co-auteurs, ont décrit que l'inactivation du gène *glf* (codant une protéine

impliquée dans l'import de glycérol), multipliait significativement (x 21) le transfert de génome par Fusion²³⁷. Ce gène est présent dans tous les génomes de mycoplasme manipulés dans cette étude. Ce résultat reste peu compris par les auteurs et nous laisse supposer que d'autres facteurs méconnus pourraient jouer sur l'efficacité de notre méthode. Des protéines membranaires pourraient aussi par leur présence ou absence, favoriser la fusion membrane à membrane par exemple.

Au cours de cette étude, le seul génome que nous n'avons pas réussi à cloner est celui de *M. bovis* PG45. Nous avons pourtant réalisé ces expériences avec *M. bovis* sauvage mais aussi avec le « mutant MnuA » de *M. bovis*, un mutant dans lequel l'activité nucléasique est quasi nulle²⁶⁹. Il est possible que la fusion *M. bovis*/levures soit moins efficace qu'avec d'autres espèces de mycoplasme mais cela ne nous paraît pas être la raison principale de cet insuccès. En effet, *M. bovis* PG45 possède un nombre très élevé de système de restriction-modification (RM) et Karas *et al.* ont démontré que la présence de ces systèmes pouvait fortement diminuer l'efficacité de Fusion²⁸³. Dans le but d'affiner cette question, nous avons ajouté à notre étude les deux souches *Mcap* 14232 et *Mmc* 95010. Ces dernières possèdent un ensemble de systèmes de RM différents (et un peu plus complexes) que chez les souches *Mcap* CK et *Mmc* GM12 et leurs génomes ont été clonés par CREasPy-Cloning (résultats non publiés). Pour rappel, c'est une méthode qui met en jeu des chromosomes bactériens isolés en bloc d'agarose et déprotéinisés et non des mycoplasmes entiers. Aucun de ces deux génomes n'a pu être cloné dans la levure par CREasPy-Fusion et ceci malgré plusieurs tentatives. Ces résultats suggèrent que des facteurs présents dans le cytoplasme de ces cellules (tels que des restrictases des systèmes RM) peuvent gêner le processus.

Nous avons initié une étude approfondie des systèmes RM chez ces espèces. Le **Tableau 42** répertorie le nombre total de systèmes RM de type I, II ou III prédits pour ces deux espèces mais également toutes celles manipulées au cours de ce travail. Ils ont été identifiés grâce à l'outil en ligne Rebase DataBase³¹² (<http://rebase.neb.com/rebase/rebase.html>), et aux bases de données Molligen 3.0 et Molligen 4.0³²⁴ (<https://services.cbib.u-bordeaux.fr/molligen/> et <http://www.molligen.org>). Dans ce tableau, nous précisons également si les restrictases ont été décrites dans la littérature ou identifiées comme putatives (sur Molligen). Pour information, deux tableaux plus détaillés sur les systèmes RM sont proposés en annexe (**Tableaux S8 et S9**).

Pour la souche *Mcap* 14232, 7 systèmes RM de type II ou III ont été identifiés, et 3 restrictases (MCAPa_3430, MCAPa_7780 et MCAPa_1280) sont partagées avec *Mmc* GM12, *Mmm* T1/44 et *Mmc* 95010 (**Tableaux S8 et S9**). Les deux premières espèces ont été clonées avec succès par notre méthode, il est donc peu probable qu'individuellement ces enzymes puissent gêner. En revanche, on ne peut pas exclure que combinées elles affectent le processus. De même, la restrictase de Type III (MCAPa_2790), a des homologues chez *Mbov* et *Maga*. Si elle n'a pas strictement gêné la CREasPy-Fusion chez *Maga*, elle pourrait limiter l'efficacité, voire inhiber la réaction en association avec d'autres systèmes.

Pour la souche *Mmc* 95010, nous avons identifié 7 systèmes RM avec 5 restrictases de type II et une restrictase de type III (**Tableaux S8 et S9**). Trois (MLC_RS00485, MLC_RS05455 et MLC_RS00255) sont partagées avec la souche *Mmc* GM12 que nous avons pu cloner et chez qui nous avons répertorié un nombre de systèmes presque équivalent : 8 RM au total avec 6 restrictases. *Mmc* 95010, partage aussi des restrictases avec *Mmm* T1/44 (MLC_RS05065 et MLC_RS01050), espèce dont le génome a été cloné dans la levure avec une grande efficacité. Cependant, un gène codant une restrictase putative de type II (motif Bgcl), MLC_6040, reste sans homologue chez les autres espèces ou souche que nous avons testé. Il pourrait être intéressant de mener une nouvelle expérience de CREasPy-Fusion avec un mutant dans lequel ce gène serait supprimé.

Enfin, pour *M. bovis* PG45, pas moins de 10 systèmes RM ont été répertoriés par Rebase et 9 restrictases ayant des homologues dans des espèces décrites dans les **Tableaux S8 et S9**. Il faut noter que certaines sont tout de même référencées comme tronquées (MBOVPG45_0615 ou MBOVPG45_0617). Il serait intéressant de supprimer plusieurs de ces *loci* afin d'abaisser le nombre de systèmes RM chez *M. bovis* PG45. Nous pourrions par exemple, imaginer supprimer plusieurs gènes, voire tous les gènes codant des restrictases de type II putatives (MBOVPG45_0366, MBOVPG45_0468, MBOVPG45_0560, MBOVPG45_0721, MBOVPG45_0763et) ou bien ceux codant les deux restrictases de type III (MBOVPG45_0167 et MBOVPG45_0439), partagées avec *Maga*. En effet, en construisant un vaste ensemble de mutants, nous pourrions comparer les effets des mutations sur la méthode de CREasPy-Fusion ; si la suppression d'un ou plusieurs systèmes permet de cloner le génome, alors nous démontrerions de façon définitive que les systèmes RM constituent une barrière à la mise en œuvre de cette méthode et par la même occasion identifierions les systèmes en cause. Une

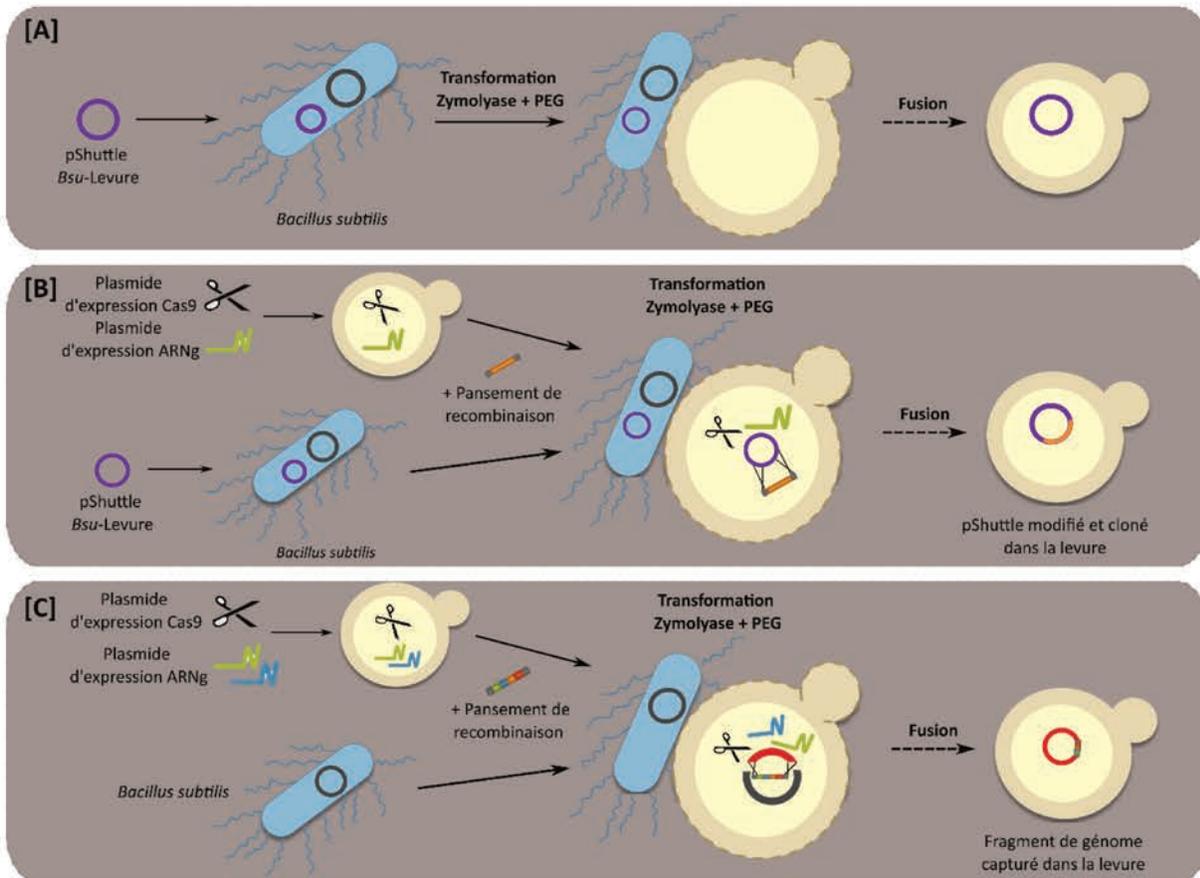


Figure 69 : Stratégie de l'adaptation de la méthode de CRISPR-Cas9 à *B. subtilis*. [A] Des protoplastes de *B. subtilis* porteurs du plasmide navette pShuttle « Bsu-Levure » sont directement fusionnés à des sphéroplastes de levures, suivant un protocole modifié de transformation au PEG par fusion des cellules. [B] Les levures sont transformées avec les plasmides pCas9 (p414-TEF1p-CAS9-CYC1t) et pARNg (p426-SNR52p-gRNA+AarI-SUP4t) permettant respectivement l'expression de la protéine Cas9 et d'un ARNg spécifique du locus cible sur le pShuttle. Puis elles sont fusionnées avec des protoplastes de *B. subtilis* porteurs du plasmide navette, en présence d'un pansement de recombinaison composé d'un marqueur d'auxotrophie (en orange), et qui possède à ses extrémités deux régions identiques au locus ciblé par l'ARNg et à réparer. Le complexe Cas9-ARNg présent dans la levure détecte le site de clivage sur le pShuttle, et génère une coupure double brin. Cette coupure stimule un des systèmes de réparation de l'ADN présents chez la levure. Plus précisément, la présence du « pansement » enclenche le mécanisme de recombinaison homologue « HR ». Après réparation, le pShuttle est modifié et cloné dans la levure. [C] De manière similaire à l'étape B, les levures pré-chargées en pCas9/pgARNs sont fusionnées avec des protoplastes de *B. subtilis*, en présence d'un pansement de recombinaison composé cette fois d'une cassette « éléments levure » et d'un marqueur de sélection, et qui possède à ses extrémités deux régions identiques au locus cible sur le génome *B. subtilis*. Après coupure par le duplexe Cas9-ARNg et réparation grâce au pansement spécifique du fragment à capturer, ce dernier est maintenu stable dans la levure.

stratégie plus simple mais moins démonstrative serait d'identifier une souche *M. bovis* possédant naturellement un nombre réduit de systèmes RM. C'est ce que nous avons montré avec les souches *Mcap* CK et *Mmc* GM12 (« clonables ») et *Mcap* 14232 et *Mmc* 95010 (« non-clonables »). Concernant *M. bovis*, nous avons identifié la souche RM16 qui possède encore 5 systèmes de type II et que nous avons à disposition au laboratoire.

Forts de ces connaissances, et de la maîtrise d'un nouveau protocole de clonage, nous avons décidé de franchir une étape supplémentaire en l'appliquant à la bactérie Gram+ *B. subtilis*, sur laquelle repose tout le projet *Bacillus* 2.0. Nos motivations principales étaient non seulement de (i) montrer la versatilité de la méthode en l'élargissant à des bactéries plus conventionnelles mais également (ii) d'ouvrir des perspectives inédites pour ce projet *Bacillus* 2.0.

II. Adaptation de la méthode de CReasPy-Fusion à *B. subtilis*

Nous avons montré dans la première partie de ce chapitre la mise en place d'une nouvelle méthode de clonage de génome dans la levure : la CReasPy-Fusion. Cette méthode présente deux points forts. D'une part, elle est basée sur un phénomène de fusion cellulaire qui permet l'entrée dans la levure de chromosome entier sous forme condensée ; un processus qui pourrait s'avérer avantageux pour les bactéries avec de grands génomes. D'autre part, elle met en jeu le système d'édition CRISPR-Cas9 et offre, contrairement à la méthode du TAR-Cloning basée sur la présence de site(s) de restriction, une flexibilité importante en termes de capture de fragments d'ADN dans la levure. Ces caractéristiques en faisaient une méthode de choix pour reprendre et progresser sur le projet central de cette thèse : le clonage du génome de *B. subtilis* dans la levure.

Au vu des difficultés rencontrées pour cloner le génome entier (Chapitre 1) ou certains fragments du génome de *B. subtilis* (Chapitre 2), nous avons décidé d'avancer pas à pas et avons donc défini plusieurs étapes pour atteindre notre objectif : mettre en œuvre la CReasPy-Fusion chez *B. subtilis*. La seconde partie du chapitre se décline donc en trois étapes (**Figure 69**) :

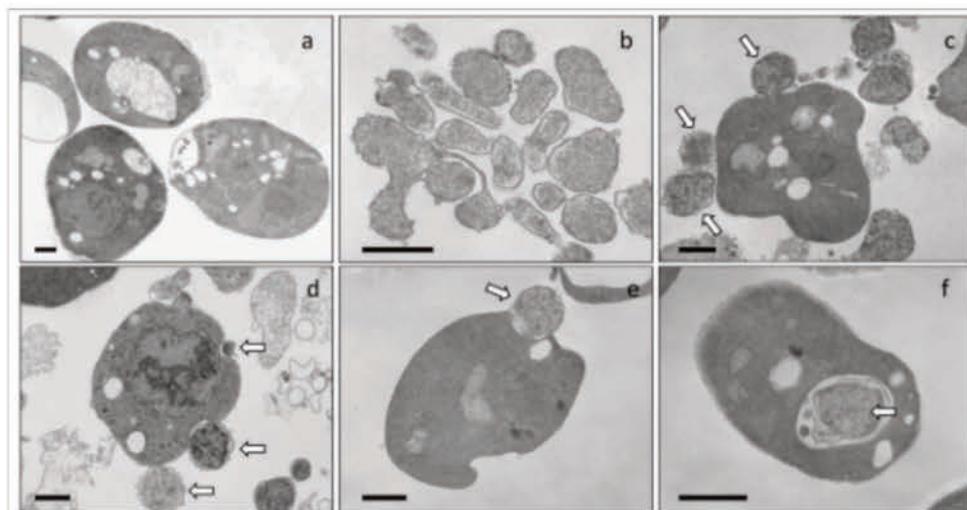
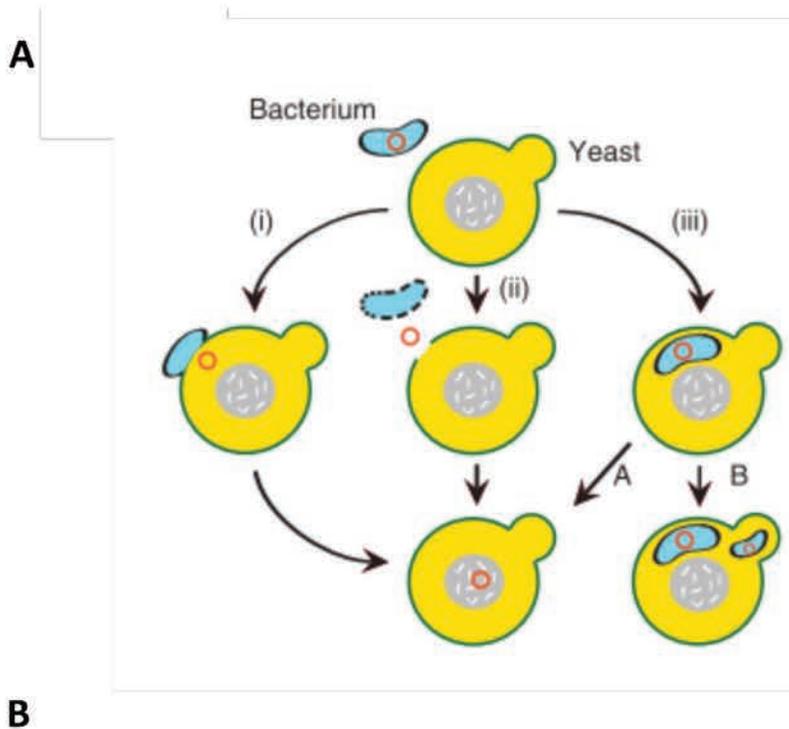


Figure 70 : Scénarios putatifs de la méthode de Fusion et analyses en Microscopie Électronique à Transmission (MET) associées. Cette figure est constituée d'un schéma et d'images de microscopie issus de la publication de Karas *et al.* 2013 **[A]** Les 3 scénarios proposés par les auteurs : (i) Fusion des membranes, (ii) Transformation et (iii) Endocytose pouvant advenir au moment de la Fusion. Après endocytose, les cellules bactériennes sont décomposées et le génome est pris en charge dans le noyau (A) les bactéries survivent en tant que symbiotes (B). Bacterium signifie bactérie et Yeast est traduit de l'anglais par levure. **[B]** Images de Microscopie Électronique à Transmission (MET) après Fusion levure-mycoplasme en présence de PEG. (a) Sphéroplastes de levure seuls, (b) Cellules de mycoplasmes seules, (c à f) les mycoplasmes donneurs sont identifiés par des flèches blanches. L'échelle (bande noire) correspond à 500 nm.

[A] Valider le phénomène de Fusion entre *B. subtilis* (bactérie Gram+) et des levures à l'aide d'un plasmide navette « pShuttle *Bsu*-Levure » porté initialement par *B. subtilis*. A cette étape, le plasmide doit être transféré de *B. subtilis* vers la levure.

[B] Valider le phénomène de Fusion entre *B. subtilis* (bactérie Gram+) et des levures portant le système d'édition CRISPR-Cas9, à l'aide d'un plasmide navette « pShuttle *Bsu*-Levure » cloné initialement chez *B. subtilis*. A cette étape, le plasmide doit être transféré de *B. subtilis* vers la levure dans laquelle il sera modifié grâce au système CRISPR-Cas9.

[C] Capturer de manière précise un fragment de génome de *B. subtilis*, avec un design permettant la double coupure du génome, puis la recircularisation du fragment dans la levure.

II.1. Validation de la Fusion par transfert d'un plasmide navette

Nous nous sommes basés sur les résultats décrits dans le Chapitre 1, et qui montraient que la Fusion entre des protoplastes de *B. subtilis* et sphéroplastes de levures est possible, permettant ainsi le transfert du génome. Nous souhaitons ici valider nos observations à travers une expérience simple mettant en jeu un plasmide navette « pShuttle *Bsu*-Levure », puisque nous avons montré que le génome entier n'était pas stable.

Afin de mettre en place une expérience qui soit la plus démonstrative possible, nous avons pris en compte des hypothèses émises par Karas *et al.* pour tenter d'expliquer les évènements pouvant se produire pendant la Fusion. Pour rappel, les auteurs mettaient en avant trois scénarios possibles (**Figure 70A**) :

- Le scénario (i) « Fusion des membranes » : L'ADNg est transféré entre deux compartiments cellulaires (bactérie vers levure) par fusion membrane à membrane et déversement du cytoplasme bactérien dans la cellule de levure.
- Le scénario (ii) « Transformation » : L'ADN est libéré dans le milieu réactionnel suite à une lyse cellulaire et incorporé par les levures par « transformation » (et non par « fusion » car la molécule d'ADN n'est pas transférée par contact de cellules à cellules).

- Le scénario (iii) « Endocytose ». L'ADN est acquis par endocytose complète de la cellule bactérienne dans la levure, qui pourrait être ensuite dégradée pour en récupérer le contenu génétique, ou bien, se maintenir en tant que symbiote.

Les données de la littérature indiquent que les scénarios (i) et (iii) sont les plus probables. Ils ont été notamment validés par des expériences *in vitro* (+/- DNase) et par microscopie électronique à transmission (MET) où les auteurs observent bien qu'au cours du traitement au PEG des mycoplasmes sont soit accolés à la levure, soit endocytés (**Figure 70B**).

Pour cette expérience, nous avons repris les tests *in vitro*, qui consistent à rajouter ou pas de la DNase I dans le mélange de Fusion. Dans notre cas, nous ne pouvions pas travailler avec le génome entier et donc nous avons construit un plasmide navette capable de se répliquer chez *B. subtilis* grâce aux gènes *cop* et *rep*, et de se maintenir aussi chez la levure grâce aux éléments ARS-CEN-HIS (cf. §II.1.2. Matériel et Méthodes). Ce dernier est nommé « pShuttle *Bsu*-Levure ». Par sa taille et contrairement à un génome entier, ce plasmide a plus de chance, en cas de lyse cellulaire, d'être acquis par la levure via un scénario « Transformation ». Nous avons donc mis en place une expérience de Fusion entre des protoplastes de *B. subtilis* souches MS ou MGP192, au préalable transformés avec le plasmide navette, et des sphéroplastes de levure.

La souche WT de levure VL6-48N a été utilisée pour cette expérience. Nous avons choisi de travailler avec un ratio de 200 μ L de sphéroplastes et 200 μ L de protoplastes de *B. subtilis* (MS ou MGP192) car nous avons obtenu de meilleurs résultats avec ces proportions lors des expériences de Fusion de génome entier (Chapitre 1). Les contrôles suivants étaient inclus dans l'expérience :

- Un « contrôle ADNg total pur » correspondant à la totalité de l'ADNg extrait du clone « MS + pShuttle » ou du clone « MGP192 + pShuttle ». Il s'agissait ici de vérifier (i) qu'en cas de lyse cellulaire, le contenu génétique des protoplastes de *B. subtilis* soit bien dégradé par la DNase I et (ii) que le pShuttle cloné et propagé par les souches MS et MGP192 puisse également se propager chez la levure.

- Des contrôles « levures » qui consistaient à étaler les levures sur milieu riche YPDA ou milieu sélectif SD-HIS après Fusion avec ou sans traitement DNase I. Il s'agissait de montrer que cette enzyme n'avait pas d'activité néfaste pour la levure, nous attendions donc des colonies sur

Sans DNase

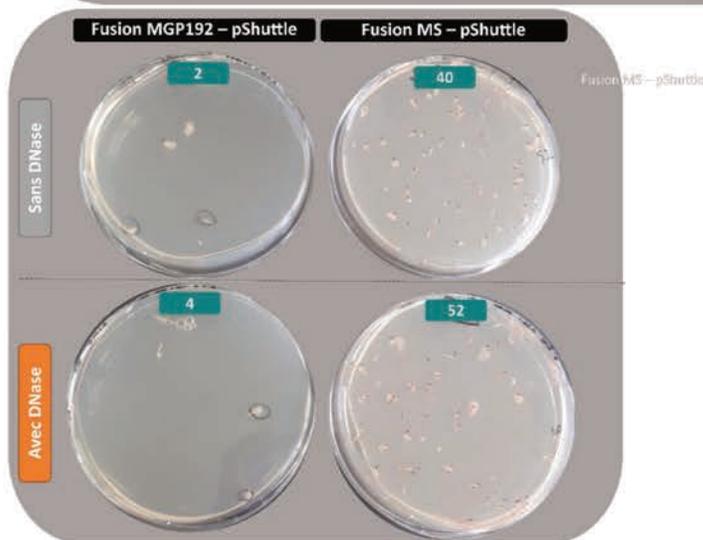
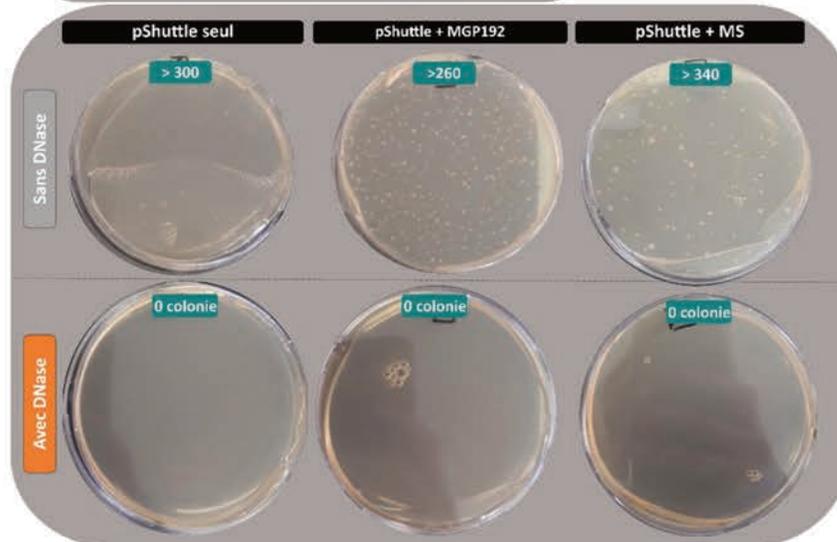
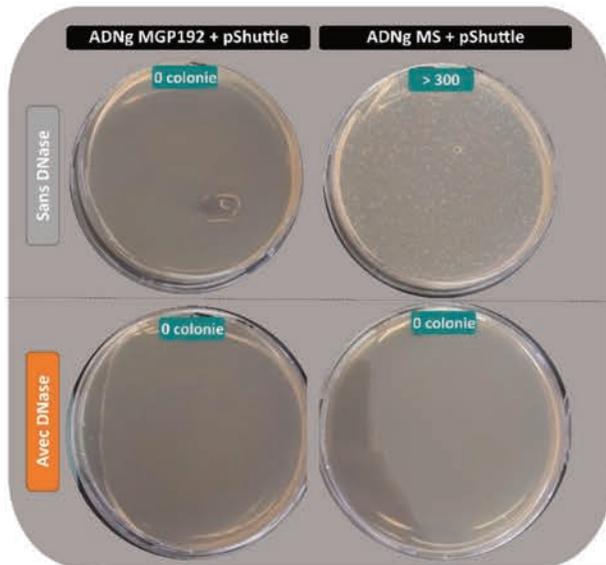


Figure 71 : Photos des boîtes obtenues après l'expérience de Fusion +/-DNase I.
Le nombre de colonies obtenu est indiqué pour chaque boîte dans des cadres verts.

boite YPDA. Il était aussi important d'attester que les sphéroplastes ne sont pas capables de se multiplier seuls dans un milieu de sélection SD-HIS.

- Des contrôles « protoplastes » prouvant qu'après Fusion les protoplastes de *B. subtilis* souche MGP 192 ou MS ne sont pas capables de se multiplier dans un milieu levure.
- Un « contrôle « pShuttle *Bsu*-Levure seul » qui est à la fois un témoin de transformation en absence de DNase I, mais aussi de validation de l'activité de cette dernière. Ce plasmide est issu d'une midiprep (cf. §II.3. Matériels et Méthodes) à partir de cellules d'*E. coli*.
- Un contrôle « pShuttle + protoplastes de *B. subtilis* WT » qui consistait à ajouter indépendamment le plasmide et des cellules de nos deux souches. Ici le pShuttle n'avait pas été transformé au préalable. Nous souhaitons ainsi vérifier que la DNase I dégradait bien le plasmide libre dans l'environnement même en présence de *B. subtilis*. Nous ne savions pas si un effet « de protection » pouvait intervenir dans ce cas. C'est aussi un contrôle de transformation (sans DNase I) afin d'attester qu'aucun facteur biologique chez *B. subtilis* n'inhibe le clonage.

L'expérience a été conduite avec des quantités saturantes de plasmide ou d'ADNg total extrait des clones *Bsu* « MS – pShuttle » ou « MGP 192 – pShuttle ». Le traitement avec l'enzyme DNase I a été effectué avec une quantité en excès (2 Unités finales) afin de dégrader la totalité de l'ADN présent dans le mélange protoplastes-sphéroplastes. Les résultats obtenus sont résumés **Tableau 43** et présentés **Figure 71**.

Les contrôles « sphéroplastes » (essais 1 à 4) effectués ont été validés et ont permis de montrer que la DNase I n'avait pas d'activité néfaste sur la levure, car des colonies sont apparues avec et sans DNase sur milieu riche YPDA. De plus, aucune colonie n'est apparue lorsque les sphéroplastes ont été étalés sur boîte SD-HIS signifiant que la sélection était bien valide. Les contrôles « protoplastes » (essais 5 à 8) ont également été validés. Ceux de la souche MGP192 (avec et sans pShuttle) ont bien permis d'attester qu'aucune colonie n'avait poussé sur milieu de sélection SD-HIS. En revanche, ceux de la souche MS (avec et sans pShuttle) montraient un léger bruit de fond sur les boîtes après transformation. De petites colonies ont été observées et leur présence pouvait être due au fait que cette souche sporule. Si cela n'a pas gêné l'isolement de transformants levure, nous avons systématiquement procédé à des repiquages sur milieu additionné d'antibiotiques (ampicilline et tétracycline).

Tableau 43 : Essai de clonage par Fusion d'un plasmide navette depuis des protoplastes de *B. subtilis* vers la levure avec ou sans DNase.

| | Essai | Condition expérimentale | Colonies (UFC) | Clones positifs / Clones analysés ^a |
|--------------------------------|-------|--|----------------|--|
| | | | | PCR simplexe |
| Contrôles levures | 1 | YPDA | >1000 | - |
| | 2 | YPDA + DNase | >1000 | - |
| | 3 | SD-His | 0 | - |
| | 4 | SD-His + DNase | 0 | - |
| Contrôles protoplastes | 5 | MGP 192 | 0 | - |
| | 6 | MS | Bruit de fond | - |
| | 7 | MGP 192 – pShuttle | 0 | - |
| | 8 | MS – pShuttle | Bruit de fond | - |
| Contrôle ADNg total | 9 | ADNg <i>B. subtilis</i> MGP 192 + pShuttle | 0 | - |
| | 10 | ADNg <i>B. subtilis</i> MGP 192 + pShuttle + DNase | 0 | - |
| Contrôle ADNg total | 11 | ADNg <i>B. subtilis</i> MS + pShuttle | >300 | 5/5 |
| | 12 | ADNg <i>B. subtilis</i> MS + pShuttle + Dnase | 0 | - |
| Contrôle pShuttle | 13 | pShuttle seul | >1000 | 5/5 |
| | 14 | pShuttle seul + Dnase | 0 | - |
| Contrôle pShuttle + <i>Bsu</i> | 15 | pShuttle + <i>B. subtilis</i> MGP 192 | 260 | 5/5 |
| | 16 | pShuttle + <i>B. subtilis</i> MGP 192 + Dnase | 0 | - |
| Contrôle pShuttle + <i>Bsu</i> | 17 | pShuttle + <i>B. subtilis</i> MS | 340 | 5/5 |
| | 18 | pShuttle + <i>B. subtilis</i> MS + Dnase | 0 | - |
| Fusion MGP192 | 19 | Fusion souche MGP 192 – pShuttle | 2 | 0/2 |
| | 20 | Fusion souche MGP 192 – pShuttle + Dnase | 4 | 0/4 |
| Fusion MS | 21 | Fusion souche MS – pShuttle | 40 | 11/15 |
| | 22 | Fusion souche MS – pShuttle + Dnase | 52 | 11/15 |

^aLe nombre de transformants levures analysés par PCR est indiqué, ainsi que le nombre de clones positifs obtenus.

De même, nous avons décidé d'en rajouter systématiquement pour les prochaines expériences.

Pour tous les contrôles réalisés en présence de Dnase I, nous avons observé qu'aucune colonie n'avait poussé sur boîte. Au contraire, en absence de DNase I, nous avons obtenu entre 260 et plus d'un millier de colonies. Cela signifiait que le traitement DNase I avait bien fonctionné et donc que l'ADNg ou le pShuttle libre dans le tube de Fusion étaient bien dégradés. De plus, la présence de colonies sur ces boîtes (sans DNase) permettait de valider l'expérience. Concernant ces contrôles, nous avons obtenu plus de colonies avec l'essai n°13 « pShuttle seul » (>1000), qu'avec le n°11 « ADNg total » pour la souche MS (>300) ; ce résultat s'explique par le fait que le plasmide est en quantité moindre dans l'extrait « ADNg total ». Nous avons également obtenu plus de colonies avec le plasmide seul (n°13) que sur les contrôles « pShuttle + *Bsu* » n°15 souche MGP192 et n°17 souche MS (260 et 340). Nous avons déjà observé ce point en première partie de ce chapitre, avec les expériences sur les mycoplasmes et avons émis pour hypothèse qu'une proportion du plasmide pourrait être captée par les cellules bactériennes (ici *B. subtilis*), diminuant le nombre de transformants levure. Enfin, nous avons noté que les contrôles « ADNg MGP192 + pShuttle » n°9 et n°10 ne semblaient pas avoir bien fonctionnés puisque aucune colonie sur les boîtes avec ou sans Dnase I n'a été observée. Ceci corrélait bien avec certaines de nos observations expérimentales en amont de cette expérience. Nous avons eu de grandes difficultés à obtenir des transformants MGP192 portant le pShuttle « *Bsu*-Levure » et une fois obtenus, les transformants se multipliaient très mal en milieu de sélection. Au cours de l'expérience de Fusion décrite ici, nous avons ensuite rencontré des problèmes pour produire une culture « viable » avec une DO correcte, ce qui, comme nous le redoutions, a certainement affecté le résultat de l'expérience : 2 et 4 transformants sans et avec DNase respectivement (essais n°19 et 20).

Les résultats obtenus avec la souche MS étaient, en revanche, plutôt prometteurs (essais n°21 et 22). Nous avons obtenu un nombre de colonies très proche en absence et en présence de Dnase I (40 et 52 respectivement). Ici la Fusion entre des sphéroplastes de levure et des protoplastes de *B. subtilis* semblait donc validée, et confirmait les résultats de Karas *et al.* selon lesquels, le transfert de ce plasmide se produit plutôt selon les scénarios (i) ou (iii), et passe par une fusion membranaire. Ceci validait aussi que la méthode appliquée et

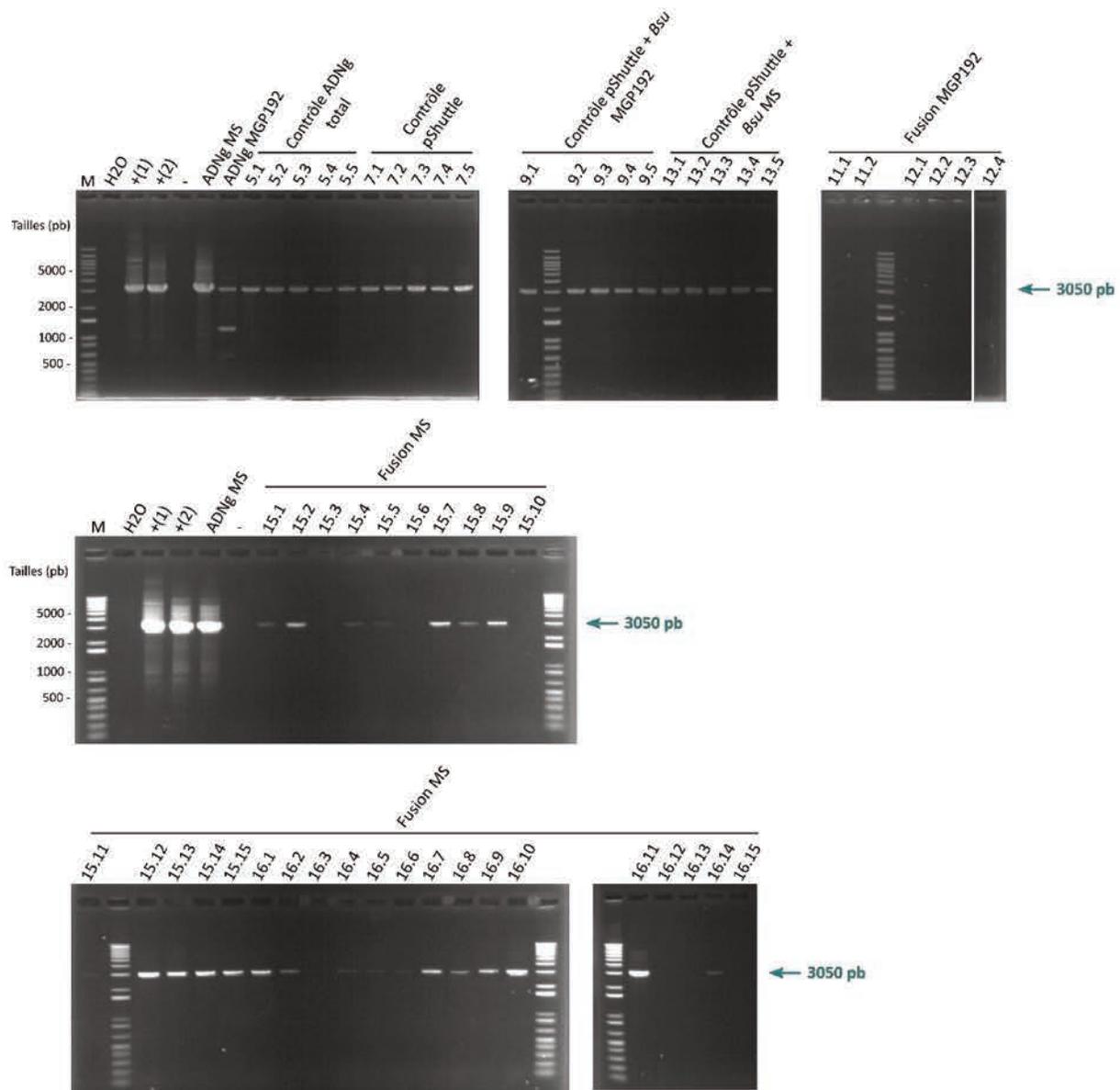


Figure 72 : Criblage de transformants levures propageant le pShuttle « Bsu-Levure » après Fusion. "M" : DNA Ladder 1 kb+ Invitrogen (100-12000 bp). "H2O" : témoin négatif sans ADN. "+(1)" : témoin positif plasmide pJIM. "+(2)" : témoin positif plasmide pShuttle. "-" : témoin négatif levure avec l'ADNg de *S. cerevisiae* VL6-48N. La présence du plasmide navette est détectée avec un couple d'amorces produisant un amplicon à 3050 pb.

les soins apportés à la manipulation des protoplastes de *B. subtilis* induisaient peu de lyse cellulaire.

Ces premières observations devaient tout de même être validées avec des analyses complémentaires, nous avons donc repiqué 5 clones pour les conditions contrôles, 30 clones pour les essais n°21 et 22 « Fusion MS » (15 avec DNase et 15 sans DNase) et enfin 6 pour les essais n°19 et 20 « Fusion MGP192 » (4 avec DNase et 2 sans DNase). La PCR simplexe effectuée avec les amorces Verif_pShuttle_F2/R permettait d'amplifier une partie du pShuttle (celle correspondant aux gènes *cop* et *rep*). Les résultats obtenus sont présentés **Figure 72**. Les clones 5.1 à 5.5 (ADNg total MS), 7.1 à 7.5 (pShuttle seul), 9.1 à 9.5 (pShuttle + MGP192) et 13.1 à 13.5 (pShuttle + MS), issus des conditions contrôles ont été validés : un amplicon PCR de 3050 pb était présent à la taille attendue à l'instar des contrôles positifs.

Concernant les essais sur les 6 clones analysés (cl 11.1, 11.2, 12.1, 12.2, 12.3 et 12.4) correspondants à la « Fusion MGP192 », aucun ne présentait de signal positif correspondant à la présence du plasmide navette. Ce résultat confortait les observations faites lors de la lecture des boîtes. Sur le gel de la **Figure 72**, le résultat obtenu suite à la PCR réalisée avec l'ADNg du clone *B. subtilis* MGP192-pShuttle, montre en plus de l'amplifiat attendu à 3050 pb, une seconde bande à ~1000 pb. Cela pouvait signifier que le plasmide était dégradé dans le clone MGP192 utilisé au cours de la Fusion, et qu'il y avait donc probablement un problème de matériel biologique. Finalement, sur les 15 clones analysés pour la « Fusion MS », que ce soit sans DNase (cl 15.1 à 15.15), ou avec DNase (cl 16.1 à 16.15), 11 transformants affichaient un profil PCR à la taille attendue, et ceci pour les 2 conditions (avec ou sans DNase).

Les résultats de cette expérience, dans son ensemble, validaient bien le fait que nous étions capables d'obtenir des clones levures propageant un pShuttle « *Bsu*-Levure » transféré depuis des protoplastes de *B. subtilis*. Nous démontrions donc ici, pour la première fois, la fusion entre des cellules de levure et des bactéries dont la paroi est Gram+.

II.2. Transfert et modification d'un plasmide navette par CReasPy-Fusion

La Fusion *B. subtilis*-Levure étant validée, nous pouvions passer à l'étape B de notre programme : ajouter le système CRISPR-Cas9 aux levures et réaliser une fusion entre des

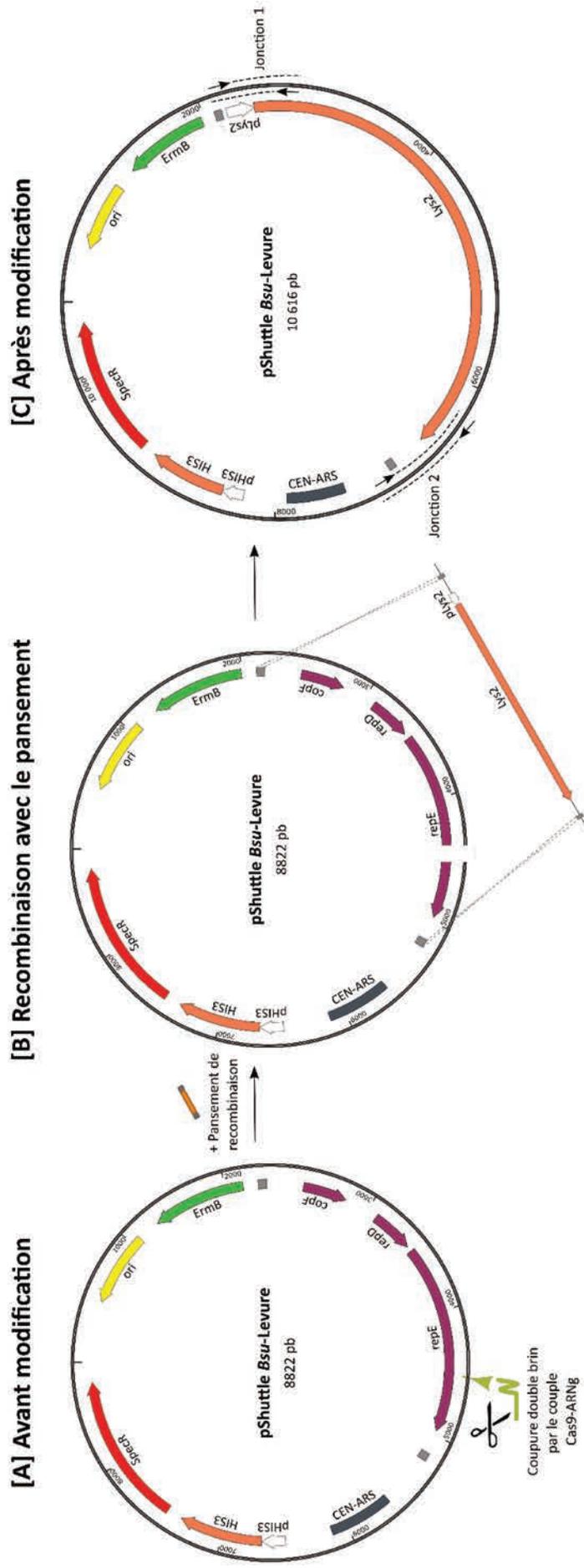


Figure 73 : Modification du pShuttle « Bsu-Levure » au cours de la CRAsPy-Fusion. [A] Avant modification : le duplex Cas9-ARNg cible spécifiquement le gène *repE* et induit une cassure double brin. [B] Recombinaison avec le pansement : le système de réparation par recombinaison homologue est stimulé chez la levure. Le pansement « *plys-Lys2* » co-transformé est utilisé par la levure pour réparer la cassure double brin. [C] Après modification : les éléments « *copF-repD-repE* » sont remplacés par le marqueur *Lys2* et son promoteur : le pShuttle est modifié.

levures pré-chargées en pCas9/pARNg et des protoplastes de *B. subtilis* portant le « pShuttle *Bsu*-Levure ».

De manière identique à ce qui est décrit plus haut (*cf.* §II.1), nous avons pré-transformé la souche levure VL6-48N-Cas9 avec des pARNg spécifiquement construits pour modifier le pShuttle et ciblant le gène *repE* (**Figure 73**). Puis, nous avons mis en contact les levures VL6-48N-Cas9-pgRNAs avec des protoplastes de *B. subtilis* et un pansement de recombinaison spécifique. Puisque nous n'avons pas obtenu de résultats avec la souche MGP192 transformée avec le pShuttle au cours de l'expérience précédente, nous l'avons exclu de notre plan expérimental et n'avons gardé que la souche « MS + pShuttle ». Concernant le pansement de recombinaison, nous avons choisi d'utiliser pour cette expérience un pansement composé d'un marqueur d'auxotrophie pour la lysine (*Lys2*) flanqué de deux cassettes identiques aux régions encadrant les trois gènes « *copF-repD-repE* » que nous souhaitons éliminer. Ce design nous permettait de réaliser une double sélection (milieu SD-HIS-LYS), et de théoriquement ne sélectionner que les transformants propageant des plasmides correctement modifiés (**Figure 73**). L'expérience incluait également divers contrôles :

- Le contrôle « levure » indispensable pour valider le fait que les sphéroplastes n'étaient pas capables de se multiplier sur la sélection SD-HIS-LYS.
- Le contrôle « protoplastes de *B. subtilis* MS » prouvant qu'après CReasPy-Fusion les cellules de *B. subtilis* n'étaient pas capables de se multiplier sur le milieu levure SD-HIS-LYS additionné d'antibiotiques.
- Un contrôle « pansement », qui consistait à transformer les levures-pCas9-pgRNA avec le pansement de recombinaison (composé seulement du marqueur d'auxotrophie *Lys2*) de façon à montrer que les levures ayant acquis le pansement n'étaient pas capables de se multiplier en milieu SD-HIS-LYS.
- Un contrôle positif « pShuttle + pansement » qui consistait à transformer les levures-Cas9-pgRNA avec du plasmide navette purifié permettant de nous assurer que le système d'édition présent dans les levures était fonctionnel (en présence du pansement).

Après transformation, tous les contrôles négatifs « levures », « protoplastes *B. subtilis* » et « pansement » (essais n°1, 2 et 3) ont été validés puisqu'aucune

Tableau 44 : Résumé de l'expérience de clonage du pShuttle par CREasPy-Fusion.

| | Essai | Condition expérimentale | Colonies (UFC) | Clones positifs / Clones analysés ^a |
|------------------------------|-------|---------------------------------|----------------|--|
| | | | | PCR simplexe |
| Contrôle levure | 1 | Sphéroplastes VL6-48N-Cas9-ARNg | 0 | - |
| Contrôle protoplaste | 2 | Protoplastes MS-pShuttle | 0 | - |
| Contrôle pansement | 3 | Pansement Lys2 | 0 | - |
| Contrôle pShuttle | 4 | pShuttle seul | >1000 | 8/10 |
| Modification pShuttle | 5 | CreasPy-Fusion MS – pShuttle | 8 | 7/8 |

^aLe nombre de transformants levures analysés par PCR est indiqué, ainsi que le nombre de clones positifs obtenus.

colonie n'est apparue sur les boites après CReasPy-Fusion (**Tableau 44**). Le contrôle positif « pShuttle + pansement » a également été validé (essai n°4), avec plus d'un millier de colonies ont pu être dénombrées sur cette boite, validant ainsi l'expérience. En plus du contrôle positif, des colonies (au total 8) sont aussi apparues sur la boite essai n°5 « CReasPy-Fusion MS-pShuttle ». 18 transformants ont été repiqués (10 sur la boite contrôle « pShuttle » et 8 sur la boite essai « CReasPy-Fusion MS-pShuttle »), et analysés par PCR avec le couple d'amorces Verif_pShuttle_F2/R (également utilisé §II.2.1).

Les résultats des PCR sont présentés **Figure 74A**. Concernant la boite contrôle n°4 « pShuttle + pansement » : (i) 8 clones sur 10 (clones 6.3 à 6.10) arboraient un amplicon à la taille attendue de 4841 pb, validant l'insertion du marqueur Lys2, en lieu et à la place de l'ensemble « *copF-repD-repE* » et (ii) 2 clones (clones 6.1 et 6.2) présentaient un profil mixte avec une bande à la taille attendue (4841 pb) et une autre identique au T+ correspondant à la taille attendue après amplification sur le pShuttle non modifié (3050 pb). Ce profil signifiait que les clones 6.1 et 6.2 étaient probablement constitués de levures hébergeant le pShuttle modifié et d'autres hébergeant le pShuttle non modifié. L'ensemble de ces résultats validait le fait qu'après entrée dans les levures, le pShuttle était correctement modifié, générant des levures prototrophes pour l'histidine et la lysine.

Concernant la boite essai n°5 « CReasPy-Fusion MS-pShuttle », les résultats de PCR montraient que nous avons un mélange de population dans quasiment tous les clones (**Figure 74A**). En effet, au-delà d'amplifiats aspécifiques, une double bande [3050 pb (pShuttle non modifié) + 4841 pb (pShuttle modifié)] était visible dans 6 clones sur 8 (7.2 ; 8.1 ; 9.1 ; 9.2 ; 9.3 ; 9.4 et 10.1). Parmi les 8 clones analysés, seul le clone 7.1 montrait une bande unique à la taille attendue (4841 pb) et semblait pur ; le clone 9.1 paraissant quant à lui négatif.

Les 8 clones ont été conservés et sous-clonés sur boite sélective SD-LYS afin de ne garder que les clones propageant la bonne construction (**Figure 74B**). Après sous-clonage, tous montraient une unique bande à 4841 pb ((cl 7.11, 7.21, 8.11, 9.21, 9.31, 9.41 et 10.11), excepté le sous-clone 9.11 (dérivant du cl 9.1, identifié négatif lors du précédent criblage). Cet amplicon était la preuve que le plasmide avait été transféré depuis *B. subtilis* et modifié dans la levure. Quelques bandes aspécifiques (déjà observées **Figure 74A**) étaient toujours présentes et correspondaient à des amplifications aspécifiques émanant de l'ADNg levure (voir témoin négatif, **Figure 74B**). Pour asseoir nos résultats de façon définitive, nous avons

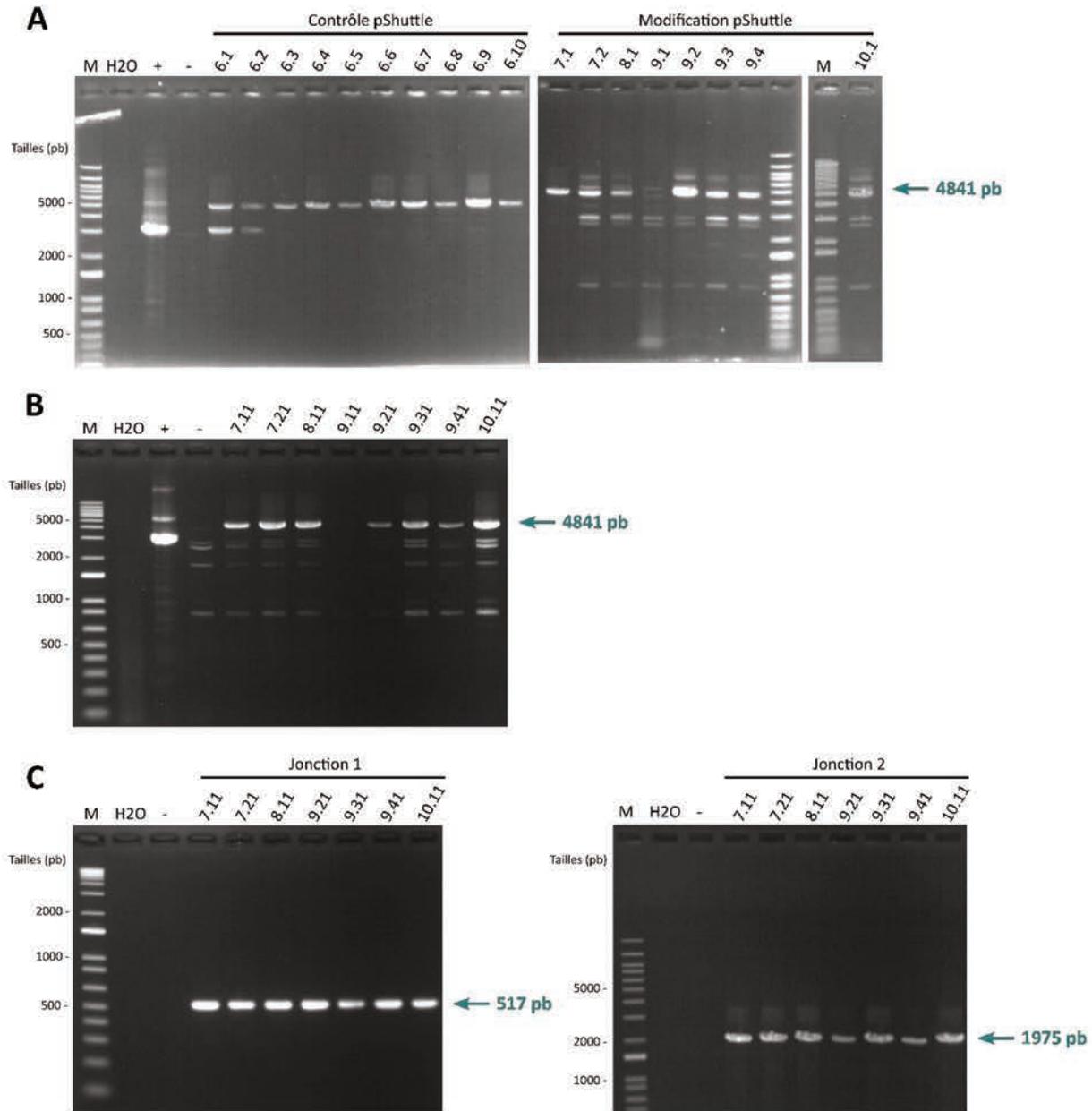


Figure 74 : Criblage de transformants levures propageant le pShuttle « *Bsu*-Levure » après Fusion. [A] La modification du plasmide navette est vérifiée grâce à un couple d'amorces produisant un amplicon à 4841 pb, au lieu de celui produit avec le pShuttle original à 3050 pb (témoin positif). [B] La même analyse est effectuée sur des sous-clones. [C] Les jonctions entre le nouveau pansement (marqueur *Lys2*) et le pShuttle ont été vérifiées avec deux PCR permettant respectivement de produire des bandes à 517 pb (Jonction 1) et à 1975 pb (Jonction 2). "M" : DNA Ladder 1 kb+ Invitrogen (100-12000 bp). "H2O" : témoin négatif sans ADN. "+" : témoin positif plasmide pShuttle. "-" : témoin négatif levure avec l'ADNg de *S. cerevisiae* VL6-48N.

validé les jonctions du pShuttle modifié (avec deux PCRs « Jonction 1 » et « Jonction 2 » schématisées **Figure 74C**). Pour chacun des sous-clones, nous avons obtenu les amplifiats à 517 pb (PCR Jonction 1) et 1975 pb (PCR Jonction 2), correspondant aux bornes souhaitées.

L'ensemble de ces résultats indiquaient clairement que nous étions capables de transférer le pShuttle depuis des protoplastes *B. subtilis* vers la levure, et en même temps de le modifier. La méthode était donc validée sur une petite molécule d'ADN de la taille d'un plasmide, l'enjeu était maintenant de travailler sur le génome de *B. subtilis* et de montrer que nous étions capables de capturer un fragment grâce à cette nouvelle méthode (Étape C du programme). Dans la suite de ce chapitre, nous aborderons les expériences mises en place pour cliver le génome de *B. subtilis* et isoler un fragment donné de manière précise.

II.3. Capture d'un fragment de génome de *B. subtilis*

Après avoir transféré et modifié avec succès un plasmide de *B. subtilis* vers la levure, la dernière étape de notre travail consistait à transférer le génome de *B. subtilis* vers la levure et à en capturer un fragment. Nous avons choisi de capturer un fragment équivalent au « Fragment 4 » (~141 kpb, Figure 46 - Chapitre 2), qui avait été cloné avec succès au cours de nos essais de TAR-Cloning.

Tout le matériel biologique nécessaire à cette première expérience a donc été préparé, à savoir : (i) des protoplastes de *B. subtilis* souche MGP192 et souche MS ; (ii) des sphéroplastes levures VL6-48N hébergeant un plasmide pCas9 ainsi qu'un plasmide pARNg-doubletarget (**Tableaux S1 et S2** – Annexes) de manière à cibler en même temps les gènes *yvcB* et *atpA* bordant le « Fragment 4 » ; et enfin (iii) le pansement de recombinaison composé des éléments ARS-CEN-HIS3-SpecR (gène *aadA* conférant la résistance à la spectinomycine) flanqué de bras d'homologie de 60 pb identiques aux régions bordant le « Fragment 4 ». Après CReasPy-Fusion, des centaines de colonies sont apparues sur les boîtes essais (MGP192 et MS) ainsi que sur la boîte contrôle « pansement de recombinaison + sphéroplastes levures ». Après analyse PCR de plus de 800 colonies, aucun clone portant le « Fragment 4 » n'a pu être isolé. L'expérience a été répétée, le résultat est resté inchangé. À l'issue de ces expériences, nous avons fait le constat que les événements de capture du fragment de génome de *B. subtilis* étaient soit peu fréquents et masqués par un « bruit de fond » trop élevé (colonies ayant

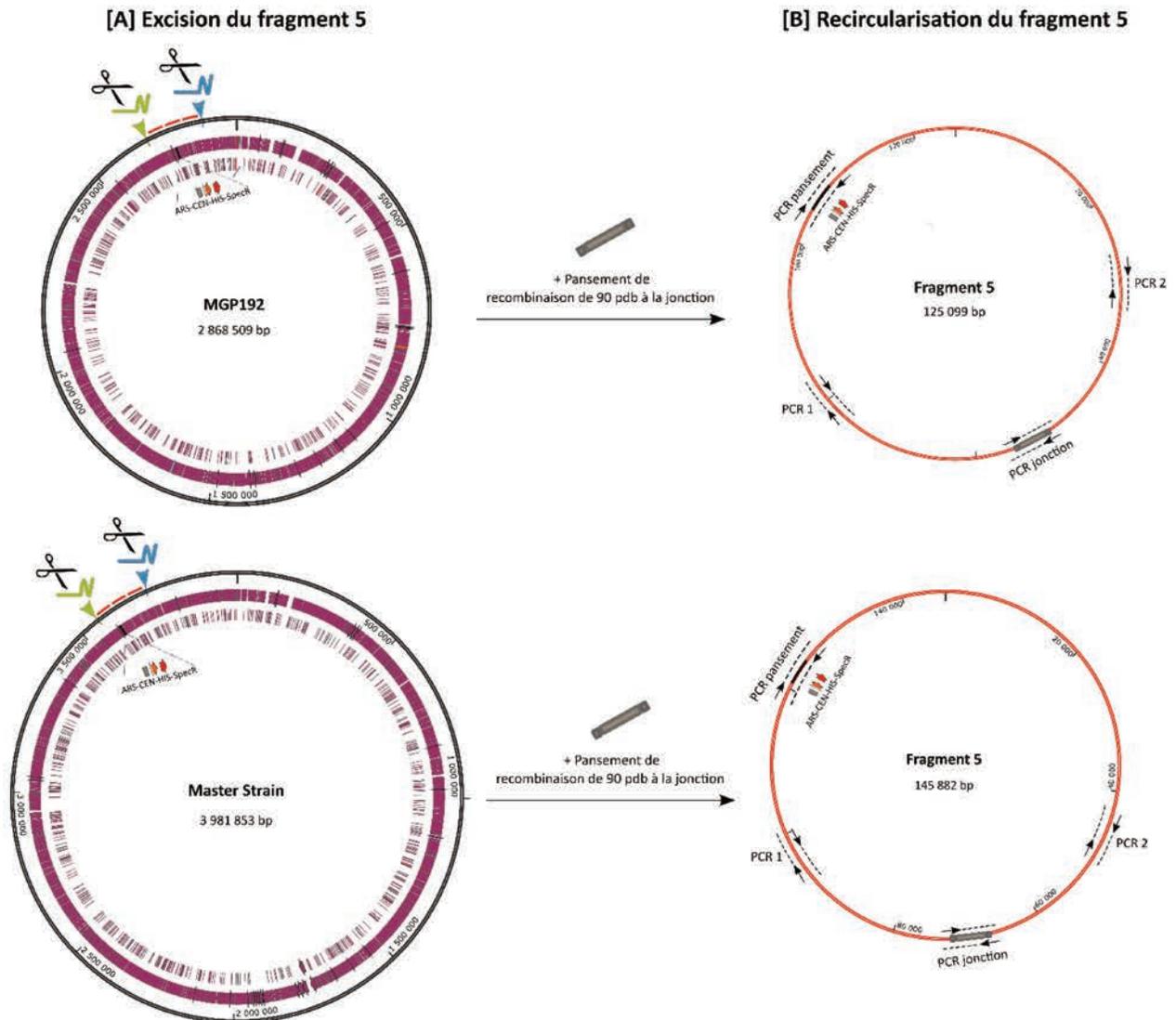


Figure 75 : Capture du « Fragment 5 avec les éléments levure » à partir des génomes de *B. subtilis* souche MGP192 ou souche MS par CReasPy-Fusion. [A] Excision du Fragment 5 : deux duplex Cas9-ARNg ciblent spécifiquement les gènes *lrgB* et *ywE* (flèches verte et bleue). [B] Recircularisation du Fragment 5 : le système de réparation par recombinaison homologue est stimulé chez la levure. Le pansement de « 90 pb à la jonction » co-transformé est utilisé par la levure pour réparer les cassures double brin et former un fragment circulaire.

acquis seulement le marqueur d'auxotrophie), soit inexistant. Ayant démontré précédemment que la fusion entre *B. subtilis* et la levure était possible, et que le « Fragment 4 » était clonable dans la levure, nous avons opté pour la première hypothèse.

Le problème de « bruit de fond » est un problème récurrent dans les expériences de clonage de génomes ou fragments de génomes dans la levure. Comme expliqué au cours du Chapitre 1 (cf. §1.4.), il a été décrit dans la littérature^{234,325} et s'explique la plupart du temps par une circularisation du « pansement » (qui contient tous les éléments nécessaires à sa propre propagation, à savoir une ARS, un centromère et un marqueur d'auxotrophie).

Dans nos expériences nous le mettons en évidence grâce à la boîte contrôle « pansement + sphéroplastes levures » qui est réalisé de façon systématique et qui contient presque toujours des colonies (parfois très nombreuses). Cependant, l'expérience montre que le nombre (plus ou moins grand) de colonies sur ces boîtes ne reflète en rien le succès de l'expérimentation. En effet, dans de nombreux cas nous sommes capables d'isoler les clones levures d'intérêt (et parfois, avec de très bonnes efficacités), alors que des quantités non négligeables de colonies sont présentes sur ces contrôles.

Ce n'était pas le cas apparemment ici et à l'évidence, il nous fallait outrepasser cette barrière pour atteindre nos objectifs. La première idée à laquelle nous avons pensé, a été de travailler avec un pansement sans ARS²²⁰. L'absence d'ARS aide à réduire le nombre de faux positifs car le pansement circularisé, ne peut plus (en théorie) se répliquer. Mais aucun de nos essais n'a été concluant. Après investigation, nous nous sommes rendus compte qu'un motif ACS de 11 nucléotides était encore présent sur notre pansement au niveau du gène *aadA* (SpecR). Nous avons alors tenté de travailler avec un pansement constitué seulement des éléments CEN-HIS et des bras d'homologie de 60 pb. Mais à nouveau, cela n'a eu aucun effet sur nos résultats : aucune colonie positive sur plusieurs centaines de clones analysés. Dans leurs travaux sur le TAR-Cloning, Kouprina *et al.* avaient également fortement insisté sur le fait de vérifier les bras de recombinaison (60 pb) associés au pansement²⁸¹. Nous n'avons pas identifié d'autres ACS pouvant expliquer nos résultats et nous n'avons pas non plus identifié de motifs répétés pouvant expliquer la circularisation du pansement.

Pour démontrer qu'il était possible de capturer des fragments de génomes de *B. subtilis*, notre dernier recours était de trouver une stratégie permettant de nous

Tableau 45 : Bilan du clonage d'un fragment de génome de *B. subtilis* par CREasPy-Fusion.

| Souche de <i>Bacillus subtilis</i> | Colonies (UFC) | Clones positifs / Clones analysés ^a | | |
|------------------------------------|----------------|--|----------------------|-------------------|
| | | PCR simplexe n°1 | Autres PCR simplexes | PFGE ^b |
| MGP192Δ <i>sbo-alb</i> ::ACH-Spec | 78 | 33/78 | 33/33 | 2/2 |
| MSΔ <i>sbo-alb</i> ::ACH-Spec | 2 | 2/2 | 2/2 | 2/2 |

« débarrasser » totalement du « pansement ». Nous avons alors eu l'idée de capturer un fragment de génome *B. subtilis* dans lequel ces éléments génétiques étaient déjà intégrés au chromosome, et de réduire le « pansement » à un simple amplicon de 90 pb (45 pb + 45 pb *cf.* §II.7.2. – Matériels et Méthodes) permettant uniquement de faire la jonction entre les extrémités du fragment à cloner (**Figure 75**).

II.3.1. Capture d'un fragment de génome de *B. subtilis* portant les éléments levures ARS-CEN-HIS3

La mise en place d'une telle expérience requiert d'avoir à disposition des bactéries *B. subtilis* portant dans leurs génomes les éléments levure ARS-CEN-HIS3.

En tout début de thèse (*cf.* §II.2. Chapitre 1), nous avons construit des clones de *B. subtilis* souche MGP192 et souche MS avec une cassette ARS-CEN-HIS3 intégrée à la place de l'opéron *sboAX-albAG*. Il se trouve que cet opéron (position : 2 703 458 pb – 2 710 395 pb dans le génome MGP192) se situe au niveau du « Fragment 5 » (133 kpb) (Figure 46, Chapitre 2) ; un fragment que nous avons réussi à cloner sans aucune difficulté au cours des expériences de TAR-Cloning (Chapitre 2). Au vu de ces informations, nous avons sélectionné ces clones et établi un plan expérimental visant à capturer en CReasPy-Fusion un fragment quasi équivalent au « Fragment 5 ».

Pour cela, la souche levure VL6-48N-Cas9 a été pré-transformée avec un plasmide pARNg portant des séquences permettant l'expression de deux ARNg ciblant respectivement les gènes *IrgB* et *ywbE* bordant le « Fragment 5 » et (**Figure 75**). Parallèlement, une matrice de recombinaison de 90 pb (*cf.* §II.7.2. Matériels et Méthodes) permettant de faire la jonction aux bornes de ce « Fragment 5 » et de le circulariser a été préparée. (**Figure 75**). Finalement, lorsque tout le matériel biologique fut prêt, les sphéroplastes de levures VL6-48N-Cas9-ARNg ont été fusionnés avec la souche MGP192 Δ *sbo-alb*::ACH-Spec ou avec la souche MS Δ *sbo-alb*::ACH-Spec, en présence du pansement (**Figure 75**).

Après sélection des transformants sur milieu SD-HIS les colonies obtenues sur boîte ont été dénombrées puis analysées. Les résultats obtenus sont résumés dans le **Tableau 45**. Nous avons obtenu un total de 78 colonies pour la souche MGP192 contre seulement 2 pour la souche MS. Cette première observation montrait déjà que l'expérience semblait avoir

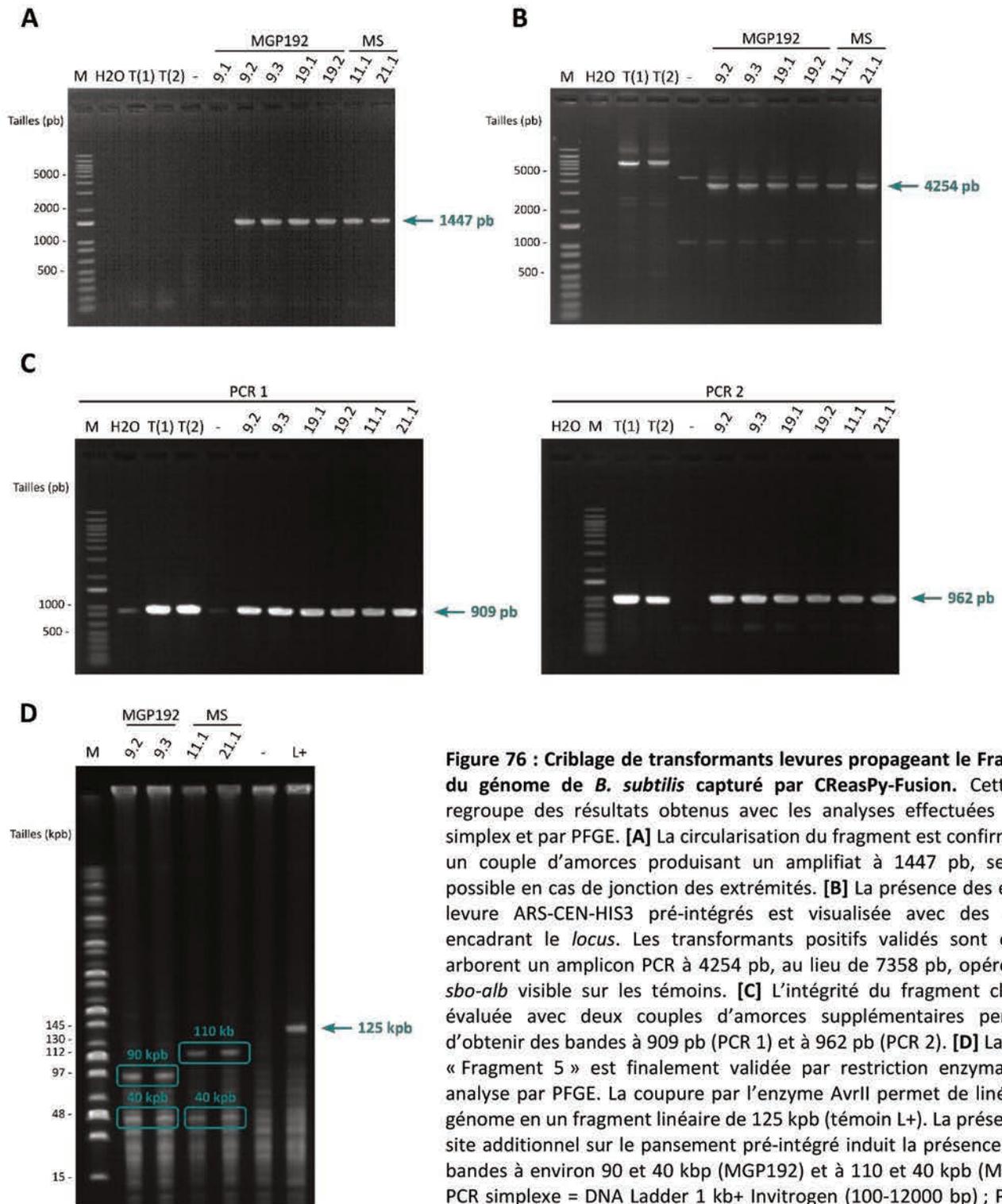


Figure 76 : Criblage de transformants levures propageant le Fragment 5 du génome de *B. subtilis* capturé par CREasPy-Fusion. Cette figure regroupe des résultats obtenus avec les analyses effectuées par PCR simplex et par PFGE. **[A]** La circularisation du fragment est confirmée avec un couple d'amorces produisant un amplifiat à 1447 pb, seulement possible en cas de jonction des extrémités. **[B]** La présence des éléments levure ARS-CEN-HIS3 pré-intégrés est visualisée avec des amorces encadrant le *locus*. Les transformants positifs validés sont ceux qui arborent un amplicon PCR à 4254 pb, au lieu de 7358 pb, opéron intact *sbo-alb* visible sur les témoins. **[C]** L'intégrité du fragment cloné, est évaluée avec deux couples d'amorces supplémentaires permettant d'obtenir des bandes à 909 pb (PCR 1) et à 962 pb (PCR 2). **[D]** La taille du « Fragment 5 » est finalement validée par restriction enzymatique et analyse par PFGE. La coupure par l'enzyme AvrII permet de linéariser le génome en un fragment linéaire de 125 kpb (témoin L+). La présence d'un site additionnel sur le pansement pré-intégré induit la présence de deux bandes à environ 90 et 40 kpb (MGP192) et à 110 et 40 kpb (MS). "M" : PCR simplexe = DNA Ladder 1 kb+ Invitrogen (100-12000 bp) ; PFGE = *S. cerevisiae* chromosomal DNA ladder Bio-Rad (225-2200 kpb). "H2O" : témoin négatif sans ADN. "T(1)" : témoin positif avec l'ADNg de *B. subtilis* MGP192. "T(2)" : témoin positif avec l'ADNg de *B. subtilis* MS. "-" : témoin négatif levure avec l'ADNg de *S. cerevisiae* VL6-48N. "L+" : témoin positif transformant levure de TAR-Cloning portant le Fragment 5.

mieux fonctionné pour la souche MGP192 dont le génome est plus petit (2,86 Mb) que pour la souche MS (3,98 Mb). Une première PCR réalisée au niveau de la jonction du « Fragment 5 » a permis de valider que 33 clones sur 78 étaient positifs pour la souche MGP192, et que les deux clones obtenus pour la MS l'étaient également. En effet, nous attendions un amplifiat PCR à 1447 pb (**Figure 76A**), qui était rendu possible seulement en cas de capture et circularisation du fragment.

Tous ces clones ont ensuite été validés avec des PCRs complémentaires (**Figure 76B**). Nous avons tout d'abord vérifié, la présence des éléments levure ARS-CEN-HIS3 à la place de l'opéron *sboAX-albAG* donnant lieu à une bande à 4254 pb. Puis nous avons cherché à évaluer si la molécule était potentiellement complète (133 kpb) en utilisant deux couples d'amorces situés à des endroits distants sur le fragment. Les résultats présentés **Figure 76C** montrent la présence d'amplifiats à 909 pb (PCR 1) et 962 pb (PCR 2), identiques aux témoins positifs, pour les clones 9.2, 9.3, 19.1 et 19.2 (MGP192), ainsi que les clones 11.1 et 21.1 (MS) Finalement, l'analyse par PFGE a permis de valider de façon définitive la capture du « Fragment 5 » entier dans ces clones (**Figure 76D**). Le fragment 5 porte un site unique AvrII permettant de linéariser le fragment après digestion par cette enzyme et d'obtenir des bandes aux alentours de 125 kpb (MGP192) ou bien de 150 kpb (MS). En accord avec cela, le témoin L+ (à droite sur le gel) qui correspond à un clone levure propageant le « Fragment 5 MGP192 » issu des expériences de TAR-Cloning du chapitre 2, affiche une bande à 125 kpb. En contraste, tous les autres clones arborent un profil à deux bandes : 90 kpb et 40 kpb pour la souche MGP192 et 110 kpb et 40 kpb pour la souche MS. Après analyse, nous nous sommes aperçus qu'un second site AvrII était présent au niveau des éléments levures ARS-CEN-HIS3 ajoutés au génome en lieu et place de l'opéron *sboAX-albAG*. Ce site était également présent dans le plasmide intégratif utilisé pour insérer ces éléments au sein du chromosome. En tenant compte de cette information, nous avons recalculé les tailles théoriques attendues pour les clones issus de la CREasPy-Fusion ; elles étaient totalement conformes aux tailles obtenues sur gel. A l'issue de cette expérience, nous avons donc réussi à capturer, pour la 1^{ère} fois, un fragment du génome de *B. subtilis* par CREasPy-Fusion pour les souches MGP192 et MS.

Ce résultat démontrait de façon définitive que nous étions capables de transférer par fusion un génome à 2,86 Mb (MGP192), mais aussi à 3,98 Mb (MS), puis à isoler un morceau et à le maintenir dans la levure. Cela repoussait un peu plus les limites de notre méthode

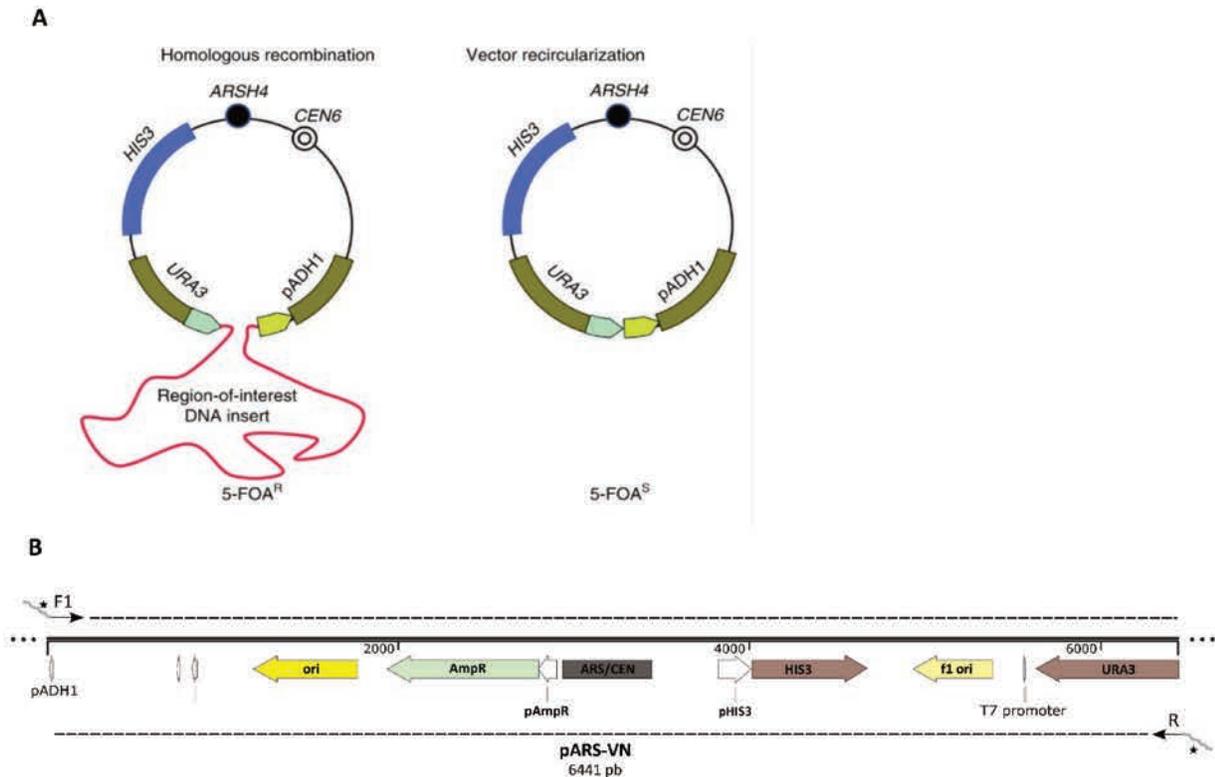


Figure 77 : Principe de la contre sélection avec le marqueur URA3 associé au 5-FOA et design du pARS-VN et criblage de transformants levures propageant le Fragment 5 du génome de *B. subtilis* capturé par CReasPy-Fusion. [A] Ce schéma est issu d'une publication de Kouprina, Noskov et Larionov parue en 2020 et indique le principe de la contre sélection au 5'FOA. En rouge est représenté l'ADN d'intérêt à capturer. Lorsque ce dernier est bien associé au vecteur pARS-VN par recombinaison homologue, les clones levures propageant le vecteur sont résistants au 5-FOA (5-FOA^R). Si le vecteur se circularise sans capturer de fragment, les clones sont alors sensibles (5-FOA^S). Homologous recombination signifie recombinaison homologue. Vector recircularization signifie recircularisation du vecteur. **[B]** La séquence du pARS-VN est utilisée comme matrice pour la production du pansement de recombinaison. Les amorces F/R avec les queues flottantes (en gris avec une étoile), permettent de produire un pansement spécifique du fragment à capturer. Les lignes en pointillées représentent l'amplification PCR.

d'autant qu'au cours du Chapitre 1 (*cf.* §II.3.3.), et particulièrement pour la souche MS, nous n'avions pas réussi à obtenir de colonies sur les boîtes. Cela ouvrait des perspectives inédites pour le projet *Bacillus* 2.0 et plus largement sur le clonage de grands génomes dans la levure.

II.3.2. Capture d'un fragment de génome de *B. subtilis* avec contre sélection au 5-FOA

Suite à ce succès, nous nous sommes demandé s'il n'existait pas une possibilité de cloner des fragments de génomes de *B. subtilis* non marqués. Dans une telle configuration, les éléments levure « ARS-CEN-HIS3 » sont obligatoirement portés par le pansement. Il nous fallait donc trouver une méthode qui permettent de contre-sélectionner ces éléments lorsque ces derniers n'étaient pas rattachés à un fragment d'intérêt.

Après une exploration détaillée de la littérature, nous nous sommes arrêtés sur une stratégie mise au point par Kouprina et Larionov au National Institute of Health (Bethesda)^{232,281}, et centrée sur un plasmide linéaire appelé pARS-VN contenant un jeu de marqueurs de sélection (HIS3) et de contre sélection (URA3/5-FOA). Le design du pARS-VN est tel, que si ce dernier se circularise, alors le gène *URA3* se retrouve sous le contrôle du promoteur constitutif *pADH1*, permettant la synthèse de l'enzyme orotidine 5-phosphate décarboxylase (ou ODCase) capable de convertir le 5-FOA (acide 5-fluoroorotique) en 5-fluorouracil, un composé toxique pour la levure³²⁶(**Figure 77A**). En conséquent, les levures hébergeant le pARS-VN « vide » (sous forme circulaire ou intégré au génome) meurent en présence de 5-FOA dans le milieu alors que les levures hébergeant le pARS-VN recombinant (c'est-à-dire associé à un fragment d'ADN) peuvent se multiplier dans le milieu (**Figure 77A**). À cette étape du travail, nous avons pensé que cette stratégie pouvait nous aider à limiter les faux positifs et à capturer un fragment de génome de *B. subtilis* non marqué. Pour mettre en place cette dernière expérience (**Figure 77**), nous avons dû ajuster plusieurs éléments. En effet, dans ce protocole, le marqueur URA3 est porté par le plasmide pARS-VN, il nous fallait donc changer le marqueur d'auxotrophie (URA3) habituellement porté par le plasmide pARNg (*cf.* §II.1.1- Matériels et Méthodes). Guidés par les géotypes des souches levures que nous avons à disposition au laboratoire, nous avons choisi de travailler avec le marqueur d'auxotrophie pour l'adénine ADE2 et la souche levure W303a (des expériences ayant montré que ce marqueur ne fonctionnait pas dans la souche VL6-48N).

Tableau 46 : Bilan du clonage d'un fragment de génome de *B. subtilis* par CReasPy-Fusion et après contre-sélection avec le 5-FOA.

| Souche de <i>Bacillus subtilis</i> | Colonies (UFC) | Clones positifs / Clones analysés ^a | | |
|------------------------------------|----------------|--|----------------------|------|
| | | PCR simplexe n°1 | Autres PCR simplexes | PFGE |
| <i>B. subtilis</i> MGP192 | 92 | 3/92 | 3/3 | 3/3 |
| <i>B. subtilis</i> MS | 60 | 0/60 | - | - |

^aLe nombre de transformants levures analysés par PCR et PFGE est indiqué, ainsi que le nombre de clones positifs obtenus.

La souche levure W303a-Cas9 a donc été pré-transformée avec des pARNg (p426-gRNA-Ade2-double-target) permettant, comme précédemment, de cliver le génome de *B. subtilis* au niveau des gènes *IrgB* et *ywbE* bordant le « Fragment 5 » (cf. §II.4.1). Parallèlement, un pansement de recombinaison spécifique a été produit par PCR à l'aide du plasmide matrice pARS-VN et d'amorces spécifiques (80 pb) dont les queues flottantes (60 pb) correspondaient aux séquences encadrant le « Fragment 5 » à capturer (cf. §II.7. - Matériels et Méthodes et **Figure 77**).

Lorsque tout le matériel biologique fut préparé, les sphéroplastes de levures W303a-Cas9-ARNg ont été fusionnés avec les souches *B. subtilis* MGP192 ou MS, en présence du pansement de recombinaison « Fragment 5-pARS-VN ». Après sélection des transformants sur milieu SD-HIS + 5-FOA, les colonies obtenues sur boîte ont été dénombrées puis analysées. Les résultats sont résumés dans le **Tableau 46**. Nous avons obtenu un total de 92 colonies pour la souche MGP192 et de 60 pour la souche MS. La totalité des transformants a été extrait et analysée (**Figure 78A**). Nous avons d'abord procédé à l'analyse des clones en « pools » (un pool contenait 5 clones).

La PCR 1 effectuée avec les amorces Bsu_Fgt5.1_F/R spécifique du « Fragment 5 », a permis d'identifier 3 pools positifs (bande attendue à 909 pb). Ces pools correspondaient uniquement à des transformants issus de la Fusion avec des protoplastes de *B. subtilis* souche MGP192. Ceci révélait donc que cette stratégie n'avait pas fonctionné avec la souche MS. Nous avons réextrait individuellement les clones issus des pools 9, 12 et 13 puis avons répété la PCR 1 avec les amorces Bsu_Fgt5.1_F/R (**Figure 78B**). Ce test a permis d'identifier 3 transformants positifs, les clones 4.56, 5.10 et 5.13. Deux autres PCR ont ensuite été réalisées pour confirmer ce résultat : (i) la PCR 2 a été effectuée avec le couple d'amorces Bsu_Fgt5.2_F/R spécifique du « Fragment 5 » et situé à distance du premier couple d'amorces. Un amplicon de taille égale à 962 pb, identique au témoin positif, a été obtenu pour les 3 clones (**Figure 78C**) ; (ii) La PCR 3 a été effectuée avec un couple d'amorces Bsu_Fgt5-jonction_F/R situé de part et d'autre de la jonction (amplification PCR possible si le fragment a été circularisé). Un amplifiat de taille égale à 7888 pb (taille attendue) a également été obtenu pour les 3 clones (**Figure 78D**). L'analyse par PFGE nous a permis de mettre en évidence une bande aux alentours de 125 kpb pour les trois clones (**Figure 78E**) et de valider de façon définitive que nous avons capturé le « Fragment 5 » dans les levures. Cette

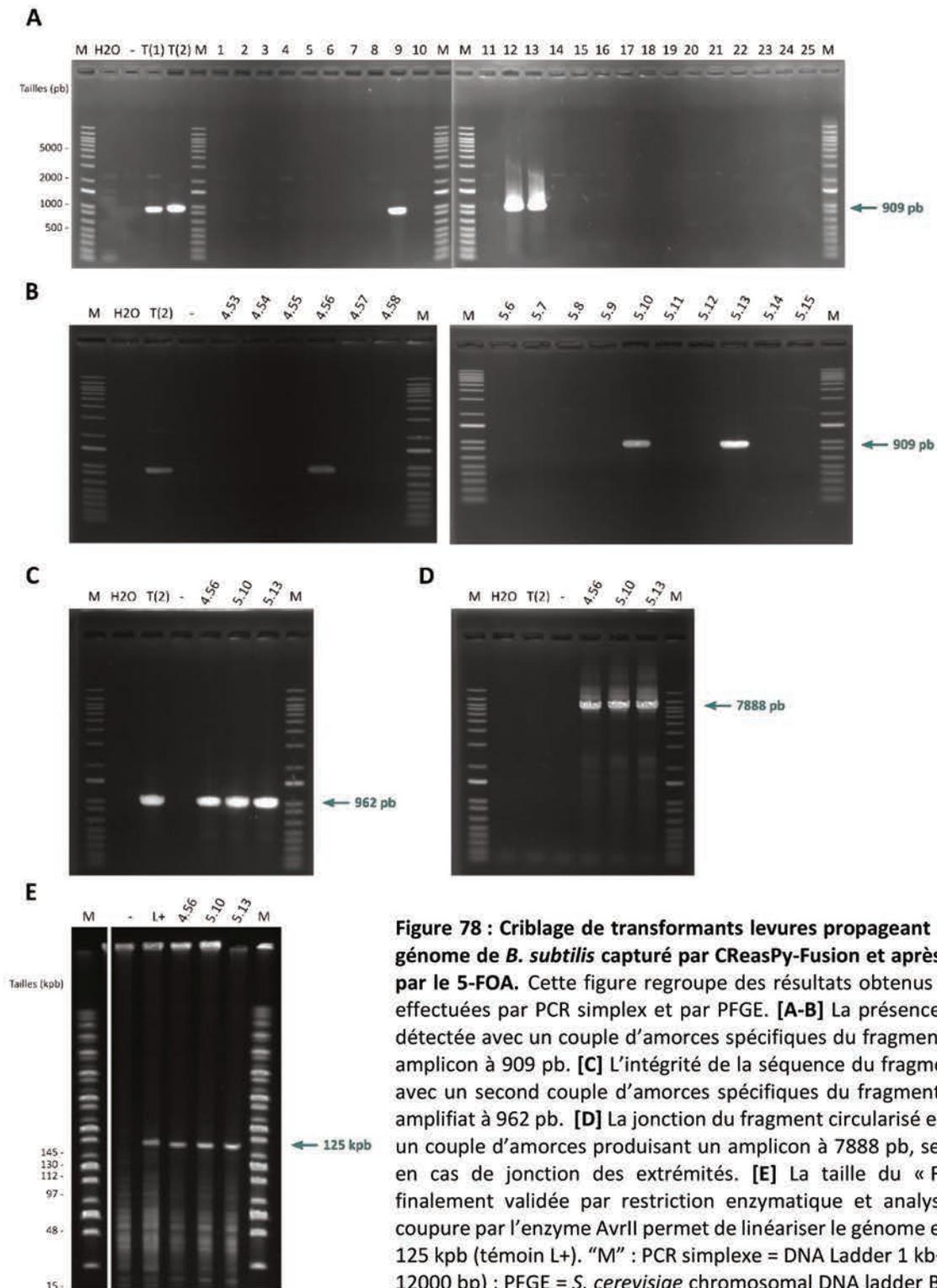


Figure 78 : Criblage de transformants levures propageant le Fragment 5 du génome de *B. subtilis* capturé par CREasPy-Fusion et après contre sélection par le 5-FOA. Cette figure regroupe des résultats obtenus avec les analyses effectuées par PCR simplex et par PFGE. **[A-B]** La présence du fragment est détectée avec un couple d'amorces spécifiques du fragment 5 produisant un amplicon à 909 pb. **[C]** L'intégrité de la séquence du fragment est confirmée avec un second couple d'amorces spécifiques du fragment 5 produisant un amplifiat à 962 pb. **[D]** La jonction du fragment circularisé est confirmée avec un couple d'amorces produisant un amplicon à 7888 pb, seulement possible en cas de jonction des extrémités. **[E]** La taille du « Fragment 5 » est finalement validée par restriction enzymatique et analysée par PFGE. La coupure par l'enzyme AvrII permet de linéariser le génome en un fragment de 125 kpb (témoin L+). "M" : PCR simplexe = DNA Ladder 1 kb+ Invitrogen (100-12000 bp) ; PFGE = *S. cerevisiae* chromosomal DNA ladder Bio-Rad (225-2200 kpb). "H2O" : témoin négatif sans ADN. "T(1)" : témoin positif avec l'ADNg de *B. subtilis* MS. "T(2)" : témoin positif avec l'ADNg de *B. subtilis* MGP192. "-" : témoin négatif levure avec l'ADNg de *S. cerevisiae* VL6-48N. "L+" : témoin positif transformant levure de TAR-Cloning portant le Fragment 5.

expérience montre, que des fragments de génome de *B. subtilis* « non marqués » peuvent être clonés dans la levure.

En conclusion, les trois grandes étapes A, B et C définies au début de la partie II de ce chapitre dans le but d'appliquer la méthode de CReasPy-Fusion à *B. subtilis* ont été exécutées les unes après les autres avec succès. Notre objectif principal est donc atteint. Nous cherchons désormais à consolider nos résultats afin de pouvoir les publier ; nous avons prévu notamment de répéter certaines expériences (qui parfois n'ont été réalisées qu'une seule fois) et d'en produire de nouvelles (capture d'un second fragment de génome plus grand).

Discussion - Conclusion

Au cours de ce chapitre, nous avons décrit des résultats qui montrent qu'il est possible de cloner des génomes bactériens entiers, ou des fragments de chromosomes, avec une nouvelle méthode appelée CReasPy-Fusion.

Pour mettre en place cette méthode, nous avons choisi de travailler, en premier lieu, avec des mycoplasmes pour de multiples raisons. D'une part, ce sont des organismes qui possèdent des génomes de petites tailles (<1,5 Mb), à faible pourcentage G+C ($\leq 40\%$), et qui utilisent un code génétique non standard. Il est désormais acquis que ces caractéristiques facilitent le maintien de génomes dans la levure *S. cerevisiae*^{219,284}. D'autre part, plusieurs génomes de mycoplasmes sélectionnés pour notre étude avaient déjà été clonés dans la levure, dont deux (*Mcap* CK et *Mmc* GM12) par la méthode de Fusion²²³. En se plaçant dans ces conditions, nous savions que ces génomes étaient « clonables » dans la levure, et que l'absence potentielle de transformants sur nos boîtes serait à attribuer (à minima pour *Mcap* CK et *Mmc* GM12) à un problème technique et non biologique.

Cette première série d'expériences (1^{ère} partie du Chapitre 3) nous a fait progresser à bien des égards. En effet, elle nous a permis (i) d'attester de la faisabilité de la méthode (clonage de génomes entiers dans la levure), (ii) de développer une méthode dérivée (capture de fragments de génomes) et enfin, (iii) de mettre en lumière des facteurs dont la présence pouvait avoir un impact direct sur le processus et influencer le résultat final. Au total, pas moins de 7 génomes de différentes espèces et souches de mycoplasmes réparties dans des

groupes phylogénétiques différents ont été clonés dans la levure ; quatre d'entre eux (*Mgal S6* et *Maga PG2*, *Mccp* Tanzanie et *Mccp* Niger) n'avaient jamais été clonés chez cet hôte avant ce travail. Au cours des différentes expériences, nous nous sommes rendus compte que les efficacités de clonage étaient extrêmement variables d'une espèce à l'autre. Nous estimons que cette variation pourrait être due :

- Aux ARNg sélectionnés et utilisés. Les oligonucléotides clonés dans le plasmide p426-SNR52p-gRNA.Aarl.Y-SUP4t sont sélectionnés de façon à cibler spécifiquement un *locus* choisi en utilisant Benchling (<https://www.benchling.com/>). Cet outil permet de lister au sein de ce *locus* donné toutes les séquences de 20 nucléotides (ou « spacers ») présentes en amont d'une séquence PAM (NGG) et de les classer grâce à des scores. Ces derniers reflètent l'efficacité supposée de l'ARNg lors de délétions de gènes par l'outil CRISPR-Cas9. Les séquences ayant des scores « ON target » (reflétant l'efficacité supposée du spacer) et « OFF target » (indiquant la spécificité du spacer) proches de 100, sont prédites comme étant les plus efficaces. L'expérience du laboratoire montre que des ARNg présentant des scores élevés sont parfois totalement inefficaces *in vivo*, c'est pourquoi il est fort probable que certains ARNg utilisés soient plus efficaces que d'autres dans notre étude.
- À la présence de facteurs biologiques actifs comme des enzymes de restriction ou des nucléases. Nous avons notamment mis en évidence que l'inactivation de deux nucléases (MnuA et Cas9) augmentait le nombre de transformants levures propageant le génome de *Mgal*. Mais aussi, que la multiplication des systèmes RM bloquait probablement le clonage de génome pour certains organismes, avec les exemples de *Mcap* 14232, *Mmc* 95010 et *M. bovis* PG45 dont nous n'avons pas réussi à cloner les génomes. Ce point avait d'ailleurs déjà été soulevé par l'équipe du Dr. Karas en 2013²²³. Plus récemment, cette même équipe a montré que la perméase GlpF pouvait également impacter le processus de transfert de génomes de la bactérie vers la levure (sans en comprendre le mécanisme)²³⁷. Ce gène est présent dans de très nombreux génomes. Il reste évident que des facteurs pourraient dans d'autres cas limiter la mise en place de notre méthode voire la mettre en échec. Mais c'est en travaillant avec d'autres espèces ou souches que nous pourrions collecter de nouvelles données. Pour surmonter ces obstacles, deux stratégies peuvent être mises en œuvre : (i) adapter le

protocole : on peut imaginer par exemple ajouter de l'EDTA dans certaines solutions pour inhiber des nucléases de surface ou encore chauffer les échantillons (45 à 55°C) pour inhiber certaines activités enzymatiques (restrictases) ; (ii) produire des mutants (élimination ou mutation du/des gène(s) identifié(s) comme potentiellement problématique(s)). L'avantage de cette seconde approche est qu'elle permet d'identifier de façon claire le/les facteurs en cause. En revanche, le génome cloné contient un ou des gènes mutés.

- Au protocole lui-même. Ce dernier point n'a pas été investigué totalement au cours de ce travail et nous ne pouvons pas en tirer des conclusions claires. Mais il pourrait être important de faire varier certains paramètres de notre protocole en fonction de l'espèce manipulée comme par exemple la phase de croissance des cultures de bactéries donneuses, la nature/composition des tampons et solutions utilisées, ou encore le ratio protoplastes bactériens/sphéroplastes levures. Ces paramètres sont des éléments clés dans la mise au point de protocole de transformations bactériennes³²⁷ ; ils le sont certainement ici aussi. Ces modifications pourraient faciliter la fusion, et potentiellement éviter d'avoir recours à des mutants.

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 Juillet 2022.

Après s'être assurés de la faisabilité de la méthode, nous avons choisi de l'appliquer à *B. subtilis*, organisme central du projet *Bacillus* 2.0. Nous étions conscients que passer des mycoplasmes à *B. subtilis* était audacieux : *B. subtilis* est en effet, une bactérie Gram+, possédant un grand génome (2,86 Mb pour la souche MGP192 et 3,8 Mb pour la souche MS), à fort pourcentage en G+C (44%) et utilisant un code génétique universel. C'est pourquoi, nous avons choisi de transposer la méthode de CReasPy-Fusion à *B. subtilis* par étape (A, B et C) en suivant une stratégie qui reposait sur trois temps forts définies au début de la partie II.2 de ce chapitre. Au cours de la première étape, nous avons documenté, pour la première fois (grâce au transfert d'un plasmide navette de *B. subtilis* vers la levure), la fusion entre une bactérie dont la paroi est Gram+ et des sphéroplastes de levure. La présence de DNase I dans le mélange réactionnel, nous a permis de confirmer qu'au cours de ces expériences, le transfert du plasmide navette se faisait soit par fusion de membrane à membrane, soit par endocytose

(et pas par un phénomène de transformation). Au cours de la deuxième étape, nous avons montré que nous pouvions cloner et simultanément modifier le plasmide navette « *Bsu-Levure* » en utilisant des levures pré-chargées en pCas9/pARNg. Nous validons ici la méthode de CReasPy-Fusion chez *B. subtilis* mais appliquée à un plasmide. Ce n'est finalement qu'au cours de la troisième et dernière étape, que nous avons montré que la CReasPy-Fusion pouvait aussi s'appliquer à la capture d'un fragment de génome préalablement marqué avec les éléments levures, ou encore, à celle d'un fragment vierge de tout marqueur grâce à une contre sélection avec du 5'-FOA. Des différences d'efficacité assez nettes ont été observées lors de la capture d'un génome de *B. subtilis* pré-marqués qui a généré 42% de clones positifs et la capture du même fragment non marqué qui a généré 3% de clones positifs. Cependant, ces expériences sont récentes et nous pensons qu'une marge de progression est possible (notamment pour le fragment non marqué), en optimisant certaines étapes du protocole. Outre la capture du « Fragment 5 » en lui-même, un autre résultat notable (obtenu lors de l'expérience avec le génome *B. subtilis* marqué) est à relever. Des transformants levures cochant tous les critères ont été obtenus après fusion des levures avec la souche MGP192 (2,86 Mb) mais également avec la souche MS (3,98 Mb). Ce résultat indique fortement que des génomes de taille égale ou supérieure à 3,98 Mb peuvent être pris en charge par la levure.

L'ensemble des résultats obtenus sur *B. subtilis* (étapes A, B, C) viennent définitivement consolider ceux produits sur les mycoplasmes et donne une dimension tout autre à l'ensemble de notre travail. Nous montrons, en effet, à travers plusieurs exemples, qu'il est désormais possible de cloner un génome entier, mais aussi, de capturer un fragment d'intérêt en choisissant précisément la localisation du ou des sites de coupure(s) et ceci sans avoir recours aux enzymes de restrictions (*cf.* TAR-Cloning). Cela offre des perspectives uniques pour le projet *Bacillus 2.0*. En effet, nous pouvons désormais imaginer utiliser cette méthode pour cloner des fragments de plus en plus grands, mais aussi, des quarts ou moitiés de génome. Grâce au système d'édition CRISPR-Cas9, nous pouvons aussi imaginer rajouter des ARS faisant défaut à la stabilité du génome dans la levure, ou éliminer les *locus* limitant le clonage de certaines portions (*e.g.* opérons ribosomiques), comme nous l'avons explicité au cours du Chapitre 2.

Avec ce travail, nous ajoutons une nouvelle méthode de clonage de génome dans la levure au catalogue déjà disponible dans la littérature. Comme les autres stratégies, la

CREasPy-Fusion a ses limites, puisque des facteurs biologiques (enzymes, génomes), mais aussi techniques (design, protocole) peuvent influencer son efficacité. Cependant, elle offre également des avantages, qui en fonction de l'organisme que l'on souhaite manipuler, peuvent en faire la méthode de choix. Par exemple, il n'est pas nécessaire de savoir transformer ou modifier l'organisme dont on souhaite cloner le génome, un point crucial pour l'étude de micro-organismes que l'on ne sait pas manipuler génétiquement à l'heure actuelle. Il n'est pas nécessaire non plus, de passer par des étapes de purification des chromosomes en bloc d'agarose, ce qui diminue le temps de préparation et les coûts. De même, lorsque le gène cible est présent dans différents micro-organismes, les plasmides pgRNA (et clones levures correspondants) peuvent être recyclés. Enfin, il est possible que pour les grands génomes (>2 Mb), seules les méthodes de fusion ou la CREasPy-Fusion fonctionnent. Les travaux réalisés sur *B. subtilis*, dans la deuxième partie de ce chapitre, pourraient aboutir à l'écriture d'une seconde publication qui pourrait être soumise à une revue scientifique évaluée par les pairs fin 2022.

DISCUSSION GÉNÉRALE ET CONCLUSION

DISCUSSION GÉNÉRALE ET CONCLUSION

Au début de ces travaux de thèse, nous avons trois objectifs : (i) attester la faisabilité des méthodologies de clonage de génome dans la levure en tentant de cloner un génome réduit de *B. subtilis* à 2,86 Mb, (ii) modifier le génome cloné en utilisant les outils d'ingénierie de génomes bactériens disponibles chez la levure (e.g CRISPR-Cas9), et enfin (iii) transplanter ce génome modifié dans une cellule receveuse adéquate. Mais nous avons été confrontés à des difficultés dès la première étape du projet puisque nous n'avons pas réussi (malgré de nombreuses tentatives) à cloner le génome réduit de *B. subtilis* MGP192 dans la levure en utilisant les stratégies de clonage déjà décrites dans la littérature (Chapitre 1). Face à cette situation, il a fallu revoir nos objectifs de départ et redéfinir le sujet de la thèse. Nous avons choisi de le recentrer sur l'identification des facteurs mettant en échec nos expériences (Chapitre 2, expériences de TAR-Cloning). Ces travaux ont permis de mettre en lumière un certain nombre de barrières (taille du génome, présence/absence de séquences ARS, nombre et distribution des opérons ribosomiques, séquences répétées) qui concourent vraisemblablement à l'instabilité du génome de *B. subtilis* MGP192 dans la levure. Ces données sont cruciales car elles permettent d'éclairer les problèmes auxquels nous étions confrontés, mais aussi de réviser le plan de travail et les stratégies mises en œuvre. Ces investigations nous ont également amené à travailler sur une nouvelle méthode de clonage (CReasPy-Fusion, Chapitre 3). Cette méthode permet de capturer des fragments de génomes ou de cloner des génomes entiers. Nous pensons qu'elle pourrait s'imposer comme « la méthode de choix » pour le clonage de molécules d'ADN de grande taille (>2 Mb), avec des pourcentages de G+C supérieurs à 40% et utilisant un code génétique standard.

Les travaux initiés au cours de cette thèse constituent les fondements du projet *Bacillus* 2.0 qui vise à transférer le cycle complet d'ingénierie de génomes bactériens chez la levure à *B. subtilis* (cf. Figure 1 – Avant-propos). Les méthodes de clonage et d'ingénierie de génomes sont les premières étapes de ce cycle et les données collectées ont permis de progresser sur ces points. Une fois le génome cloné et modifié dans la levure, l'étape suivante consiste à isoler ce dernier et à le transplanter dans une cellule bactérienne receveuse, afin d'obtenir une bactérie génétiquement modifiée. Cette étape développée chez les

mycoplasmes est disponible pour un nombre restreint d'espèces au sein des mollicutes (quelques mycoplasmes et *M. florum*)²⁵³. Elle repose notamment sur la mise au point d'un protocole au PEG, et sur des essais de transplantation de génomes (i) de bactérie à bactérie et (ii) de levure à bactérie. Au cours de ma thèse, des expériences préliminaires de transformation au PEG ont été réalisées avec des plasmides. Quelques transformants ont été obtenus, mais de façon non systématique et avec une efficacité encore faible. Cette étape reste donc entièrement à développer. Une alternative à la transplantation de génomes mise au point pour *Mycoplasma pneumoniae* (méthode RAGE pour « Recombinase-Associated Genomic Exchange »)²⁵⁵, a en revanche été amorcée. Elle permet de transférer, non pas des génomes entiers, mais des sections de génomes de la levure vers la bactérie à éditer.

Dans la suite de cette discussion générale, quatre points seront développés plus en détails : (i) les obstacles rencontrés lors des essais de clonage du génome de *B. subtilis* MGP192 dans la levure, (ii) l'impact de ces résultats sur le projet *Bacillus* 2.0 et les perspectives envisagées pour surmonter ces problèmes, (iii) l'apport de la méthode CReasPy-Fusion pour le projet et au-delà en biologie de synthèse, et enfin (iv) la méthode RAGE comme une alternative possible à la transplantation de génomes.

Identification de verrous biologiques au clonage dans la levure. Les résultats obtenus au cours de mon projet de thèse démontrent qu'un ensemble de facteurs biologiques ne permet pas de cloner en entier le génome réduit de *B. subtilis* souche MGP192 dans la levure. En revanche, la totalité de sa séquence a pu être clonée sous la forme de fragments séparés (25 fragments de 6 à 515 kpb ; cf. Chapitre 2).

Lors de ce projet, nous avons été confrontés en premier lieu, à la taille du génome de la souche MGP192 (2,86 Mb) qui dépassait largement celle des plus gros génomes clonés dans la levure en un seul morceau (*H. influenzae* à 1,8 Mb²³⁷ et *S. citri* à 1,8 Mb²³⁰). Au début de ces travaux, nous ignorions si la taille du chromosome cible était un facteur limitant pour le processus de clonage dans la levure. Des données indiquent qu'une portion du chromosome Y humain dont la taille était de 2,3 Mb avait pu être cloné sous forme de YAC (soit une molécule linéaire avec des télomères à ses extrémités)³²⁸, signifiant qu'il pouvait être possible d'aller au-delà des 1,8 Mb. Dans nos travaux, nous avons tenté de repousser les limites du

clonage de génomes dans la levure en travaillant avec la souche MGP192, mais aussi, avec la souche MS (3,98 Mb). Nous n'avons pas pu démontrer que leurs génomes se maintenaient chez cet hôte mais la taille de la molécule d'ADN ne paraissait pas le seul paramètre en jeu. En effet, au cours du chapitre 1, nous avons observé que le génome pouvait être transféré depuis *B. subtilis* vers la levure par la méthode de Fusion, mais ne s'y maintenait pas puisque seul un fragment d'environ 300 kpb était détecté *in fine*. A l'issue de cette expérience, nous avons donc conclu qu'il paraissait possible d'introduire des molécules de 2,86 Mb dans la levure, mais uniquement par Fusion puisque les méthodes mettant en jeu des génomes isolés en bloc d'agarose (méthode Classique et CREasPy-Cloning) sont restées vaines.

Dans le chapitre 2, nous avons tenté de comprendre qu'elles pouvaient être les raisons de ce résultat. Nous avons mis en évidence qu'il y avait de fortes chances pour que l'instabilité observée soit en partie due à un manque d'ARS. Grâce à des analyses *in silico*, nous avons évalué la proportion d'ACS et d'ARS putatives d'abord sur le « Fragment 1 » à 1,5 Mb (*cf.* expériences de TAR-Cloning). Puis nous avons étendu nos analyses à l'ensemble du génome montrant que la répartition des ARS putatives sur le chromosome de *B. subtilis* MGP192 était raisonnable (36 ARS pour un génome de 2,86 Mb suite à l'analyse par Ori-Finder3), même si quelques sections de ce génome semblaient dépourvues de ces éléments (**Figure 56**). La littérature décrit que les séquences ARS (riches en AT) sont plus rares dans les chromosomes dont le pourcentage en G+C est supérieur à 40%²²⁰. Des chromosomes bactériens ont pu être clonés avec un vecteur contenant une seule ARS lorsque le pourcentage en G+C du fragment est inférieur à 40% ; voire avec un vecteur sans ARS pour les génomes très riches en A+T (comme ceux de certains mycoplasmes par exemple). En revanche, lorsque le génome à une teneur en G+C supérieure à 40%, il a été démontré qu'il était plus judicieux de capturer, dans un premier temps, des fragments chevauchants d'environ 100 kpb avec des ARS, de façon à obtenir dans une seconde étape, une molécule assemblée contenant des origines de répllication de la levure tous les 100 à 150 kpb^{271,284}. Pour le génome de *B. subtilis* MGP192 qui présente un pourcentage en G+C proche des 40%, il n'était pas évident de savoir (avant les expériences de TAR-Cloning menées au cours de cette thèse) si des ARS supplémentaires étaient réellement nécessaires et, le cas échéant combien d'ARS devaient être ajoutées. Désormais et après discussion avec nos collaborateurs, il apparaît primordial d'en ajouter. Des travaux ont donc d'ores et déjà été entamés, et visent à ajouter des ARS que

l'on sait fonctionnelles (*e.g.* ARS416) à des endroits précis du génome de *B. subtilis* MGP192. Cela avait également été fait pour les génomes synthétiques réduits et clonés dans la levure de *Caulobacter crescentus* (0,79 Mb et 57% G+C)²⁷⁰ et d'*E. coli* (1,03 Mb et 51% G+C)²²⁵. Nous avons pour le moment estimé nécessaire d'en ajouter 12 en priorisant les régions du chromosome qui sont les plus pauvres en ARS et les plus instables. Lors de ce design, nous avons veillé, d'une part, à ne pas inactiver des gènes qui seraient essentiels ou importants et d'autre part, à ne pas toucher aux unités de transcription.

Hormis la taille du génome de *B. subtilis* MGP192 et le manque potentiel d'ARS tout au long de ce génome, nos travaux nous ont permis d'identifier un troisième facteur potentiellement bloquant. En effet, lors des essais de capture du « Fragment 7 » (482 kpb), nous avons constaté que les levures propageaient systématiquement un fragment tronqué. Après analyse, nous nous sommes aperçus que les régions « manquantes » étaient les mêmes dans tous les clones levures. Nous avons exploré ces séquences en détail et avons réalisé qu'elles contenaient la majorité des gènes codant les protéines ribosomiques et les ARN ribosomiques. Nous avons émis trois hypothèses : (i) l'expression fortuite de ces gènes aboutit à la formation de machineries hydrides ou dysfonctionnelles gênantes voire létales pour la levure ; (ii) la présence de ces gènes en un même endroit du génome crée une zone fortement répétée pouvant être le siège de remaniements chromosomiques chez la levure et, (iii) une combinaison des deux. Une solution à court terme pourrait être envisagée en éliminant certains de ces gènes. Le ribosome de *B. subtilis* est composé de 53 protéines ribosomiques (33 protéines pour la grande sous-unité et 20 protéines pour la petite sous-unité). Plusieurs de ces protéines sont non essentielles, dans des expériences où l'inactivation ne concerne qu'un seul gène à la fois³²⁹. Il est donc concevable que l'absence simultanée de plusieurs protéines ribosomiques ait un impact significatif sur la viabilité de la bactérie. Lors de tentatives antérieures de construction de cellules bactériennes "minimales", soit entièrement synthétiques ou hautement rationalisées^{169,172,218} des micro-organismes présentant une croissance rapide ont été obtenus. Dans ces exemples, le contenu en gènes codant des protéines ribosomales n'a jamais été modifié. Le projet « Mini*Bacillus* » conduit par l'équipe de J. Stülke ne fait pas exception et tous les gènes ont été conservés lors de la réduction du génome de *B. subtilis*^{183,191}. Il serait donc intéressant d'évaluer dans quelle mesure la suppression de certains de ces gènes permettrait d'obtenir des souches châssis

minimisées de *B. subtilis* viables et dont la fitness ne serait pas impactée dans un milieu de culture défini. Sur le long terme, il pourrait être intéressant de proposer une stratégie qui tendrait plutôt à déplacer ces gènes à d'autres endroits sur le chromosome. A l'instar des travaux réalisés chez *Vibrio cholerae*, les auteurs démontrent que la position du *locus* S10 a un impact sur la croissance de cette bactérie lorsque cette région est éloignée de l'*oriC*³³⁰. Cependant, cela n'est pas rédhibitoire sur la viabilité de la bactérie et laisse penser qu'un compromis pourrait être trouvé entre le nombre de copies de ces gènes à conserver et leur localisation pour maintenir une croissance rapide chez *B. subtilis*.

Finalement, nous avons aussi mentionné que la présence de plusieurs copies des opérons d'ARN ribosomique (ARNr) au niveau de l'*oriC* pourrait engendrer de l'instabilité sur cette portion de génome. *B. subtilis* possède au total 10 opérons qui sont également associés aux gènes pour les ARN de transfert (ARNt). Il est bien établi que la redondance des ARNr et des ARNt est un déterminant important du taux de croissance³³¹. De ce fait (et comme pour les protéines ribosomiques) les travaux sur le génome d'un *MiniBacillus* ont par conséquent conservé la totalité de ces gènes¹⁹¹. Les auteurs soulignent néanmoins, que la réduction du nombre de copies de ces derniers, reste à explorer pour évaluer et obtenir en finalité un mécanisme de traduction minimal qui soit robuste. La structure d'un chromosome bactérien telle qu'elle est décrite, suit un schéma selon lequel les gènes dont l'expression est importante sont situés sur le brin sens mais aussi au plus proche de l'origine de réplication³³². Les opérons *rrn* chez *E. coli*, mais aussi, chez d'autres bactéries sont orientés de sorte qu'ils soient transcrits dans la même orientation que la fourche de réplication des chromosomes³³³, afin de réduire la possibilité de collisions frontales entre les machineries de réplication de l'ADN et de transcription de l'ARN³³⁴. Une étude récente démontre que la réduction de ces copies diminue la vitesse de croissance d'*E. coli* et de *B. subtilis* et augmente le taux de mortalité³³⁵. Il ressort également qu'une des fonctions majeures de la multiplicité des opérons *rrn* dans un chromosome bactérien, est d'assurer une bonne mécanistique de la réplication des chromosomes. Cela permettrait ainsi d'éviter la mort cellulaire en cas de fluctuation majeure des conditions environnementales. Dans le cas des travaux du projet ANR *Bacillus 2.0*, il est nécessaire de pouvoir trouver des solutions pour que cette région se maintienne dans la levure. Une mesure pourrait donc être, comme pour les protéines ribosomiques, de supprimer certaines copies ou bien de les déplacer sur le chromosome (à

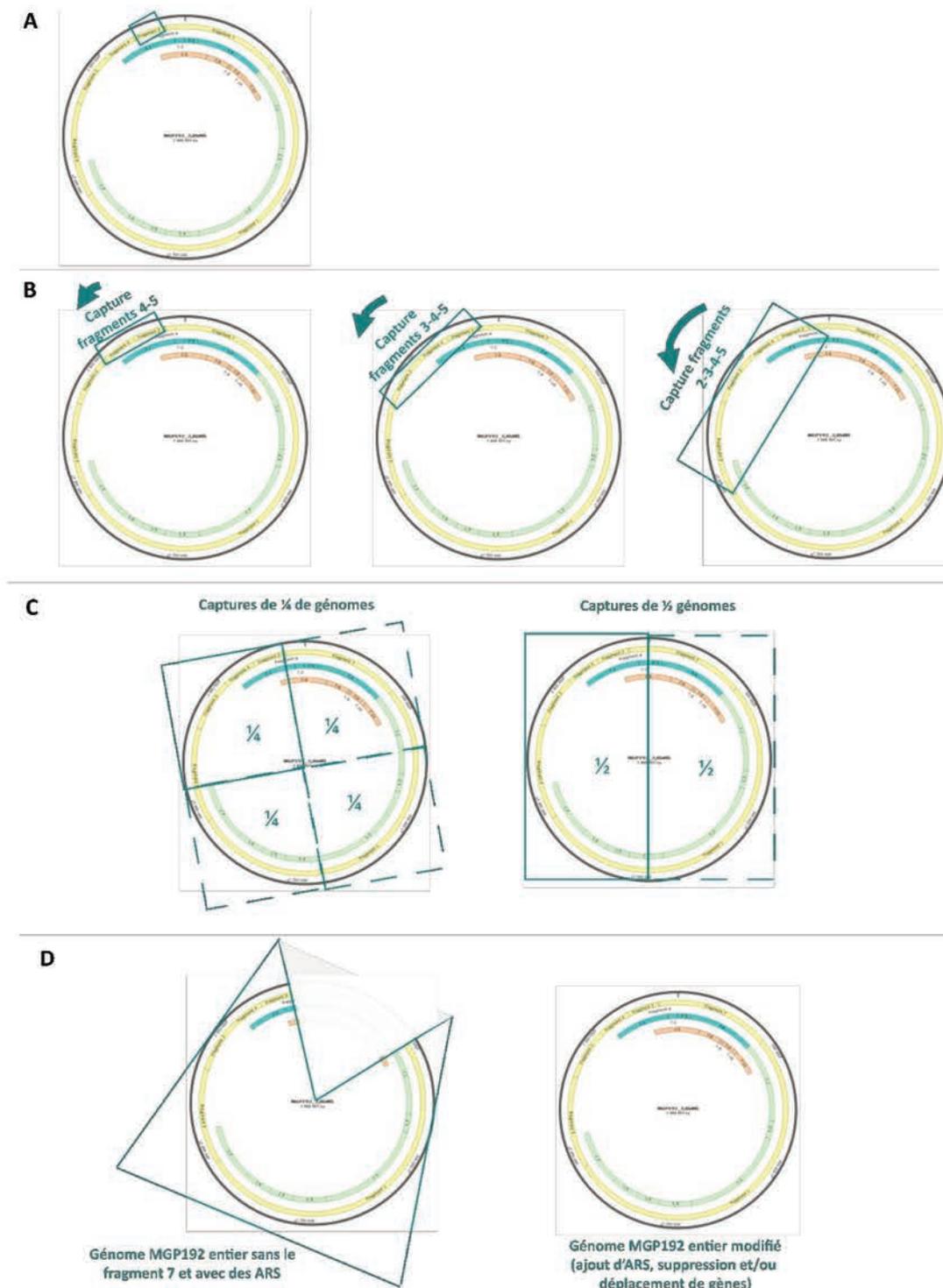


Figure 79 : Plan de travail pour les prochains essais de clonage du génome de *B. subtilis* MGP192 dans la levure en utilisant la méthode de CREasPy-Fusion. [A] Au cours du chapitre 3, nous avons montré que nous pouvions capturer le « Fragment 5 » marqué avec les éléments levure (ARS, CEN, HIS3). **[B]** Dans les prochains mois, nous tenterons de capturer des fragments de plus en plus grands contenant ou non, des ARS supplémentaires. Nous procéderons de manière progressive, comme indiqué sur le schéma, en capturant par exemple le fragment [4 + 5] puis le fragment [3 + 4 + 5], puis le [2 + 3 + 4 + 5]. Ces expériences permettront de démontrer la nécessité d'ajouter des ARS surnuméraires pour stabiliser le clonage de grands fragments du génome de *B. subtilis* MGP192 dans la levure. **[C]** Si les résultats sont satisfaisants, nous tenterons ensuite de capturer des $\frac{1}{4}$ de génomes puis des $\frac{1}{2}$ génomes contenant des ARS supplémentaires (schémas non exhaustifs). **[D]** Enfin, nous tenterons la capture du génome entier (+ARS) sans le « Fragment 7 », puis le génome entier (+ARS) modifié.

titre provisoire et si nécessaire). Il faudrait alors évaluer la dynamique cellulaire et la croissance des mutants produits et en fonction les rajouter en aval ou bien compléter leur expression avec un plasmide auxiliaire.

Perspectives et impacts de nos données sur le Project *Bacillus* 2.0. Les facteurs mis en lumière au cours de ces travaux sont désormais à prendre en compte dans le cadre de nouveaux essais de clonage du génome de *B. subtilis* MGP192 dans la levure. La **Figure 79** présente le plan de travail prévu pour les prochains mois et les jalons intermédiaires. Plus important encore, ils doivent être pris en compte pour réajuster le « design » du génome semi-synthétique *SynBsu* (1,89 Mb) dont l'assemblage a déjà démarré (travail effectué par l'équipe Syber dirigée par le Pr. M. Jules (INRAE Jouy-En-Josas) en parallèle de nos expériences).

Le design de *SynBsu* a été initialement pensé à partir de données de la littérature et d'un ensemble de résultats et d'analyses propres à l'équipe Syber. Nos collaborateurs ont défini un ensemble de règles pour convenir des gènes ou *loci* à conserver ou à exclure. Le génome « désigné » (1,89 Mb) est différent de celui de la souche MGP192 qui a été produite (par cette même équipe) par réduction itérative de portion de génome (approche Top-down)¹⁸⁴. Afin de mieux saisir les incidences de nos résultats sur le design du génome *SynBsu*, nous avons aligné sa séquence (telle que « désigné » au début du projet) avec celle du génome réduit de *B. subtilis* MGP192. La carte présentée **Figure 80** est issue de cet alignement ; elle permet de comparer les deux génomes et de visualiser rapidement les convergences et divergences. Nous pouvons noter en particulier que :

- (1) De nombreuses zones se recoupent, mais il existe aussi des *locus* présents dans MGP192 et absents de *SynBsu*, et à l'inverse (et plus problématique) des *locus* présents dans *SynBsu* et absents de MGP192. Ces derniers représentés par des ronds gris sur la **Figure 80**, s'étalent sur 8 zones et rassemblent 164 gènes au total. Lors de l'assemblage du génome *SynBsu* dans la levure, il sera donc nécessaire de porter une attention particulière à ces régions qui n'auront jamais été clonées dans la levure auparavant.

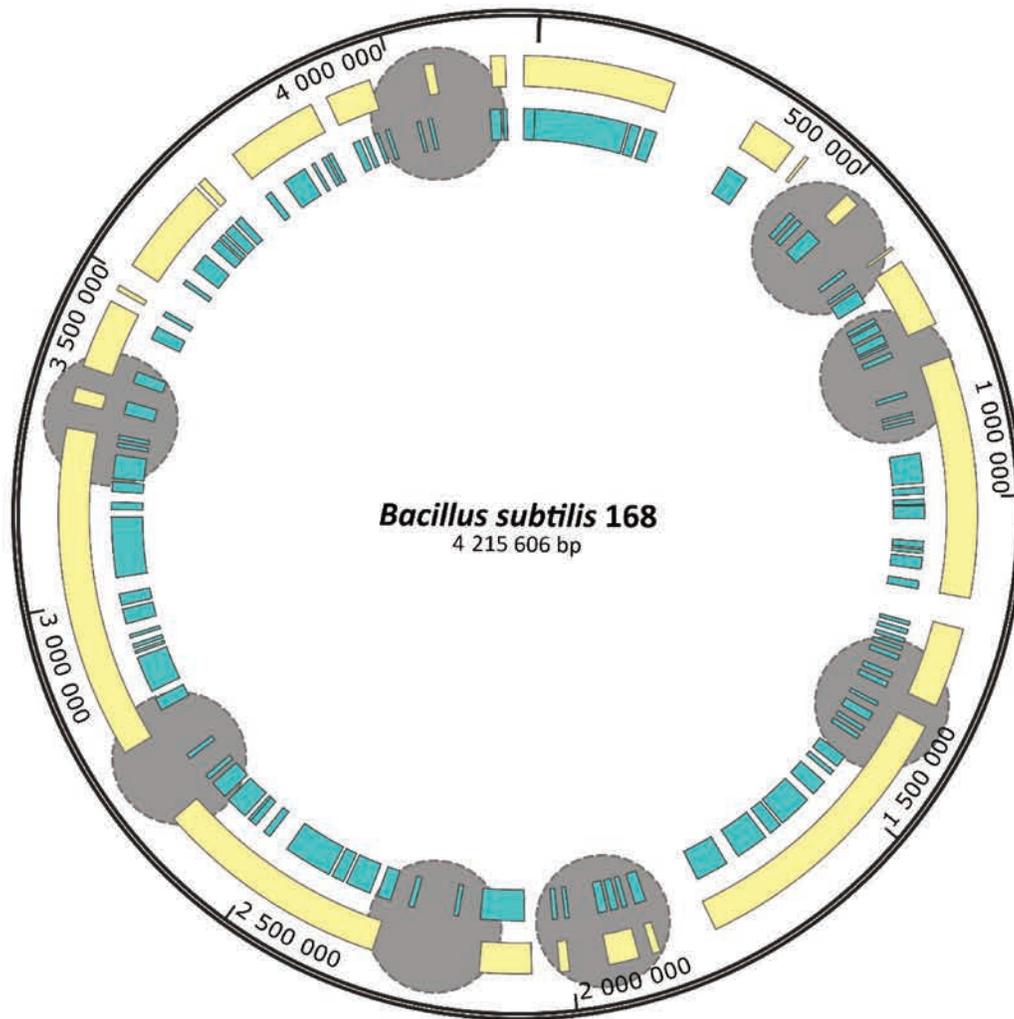


Figure 80 : Mapping des séquences de plusieurs souches de *B. subtilis*. Les séquences des génomes de *B. subtilis* souche 168, MGP192 et SynBsu ont été alignées. En jaune les portions conservées pour la souche MGP192 sont indiquées. En bleu, celles théoriquement conservées et correspondant au design du génome synthétique SynBsu sont représentées. Les zones en gris correspondent aux *locus* qui sont présents chez la souche SynBsu mais absents chez la souche MGP192.

- (2) Le génome synthétique *SynBsu* a une taille de 1,89 Mb ; il est donc bien plus petit que celui de la souche MGP192 (2,86 Mb). Sa taille est proche de celles d'autres génomes clonés avec succès dans la levure (*S. citri*²³⁰ et *H. influenzae*²²³). Aussi, il est fort probable que la taille ne constitue pas une barrière au clonage et à la propagation de génome dans la levure, comme ce fut le cas pour le génome de la souche MGP192.
- (3) Le génome synthétique *SynBsu* (44 GC%) posséderait 18 ARS putatives (analyse Ori-Finder3), soit une ARS environ tous les 112 kpb. Au vu des résultats obtenus au cours de notre étude (mais également des données de la littérature), le design de départ du génome synthétique sera modifié de façon à y intégrer 12 ARS supplémentaires. Ces dernières seront ajoutées exactement aux mêmes positions que les ARS surnuméraires prévues pour le génome de *B. subtilis* MGP192 (cf. paragraphes verrous biologiques).
- (4) Le génome synthétique *SynBsu* (théorique et présenté ici) contient la quasi-totalité du « Fragment 7 ». Comme largement discuté dans la section précédente, cela risque d'être problématique. Des décisions seront prises prochainement pour supprimer ou déplacer ailleurs dans le génome, les opérons codants les protéines ribosomiques et/ou ARN ribosomiques. Il est possible que plusieurs designs soient testés en parallèle.

Les résultats obtenus au cours de cette thèse permettent dorénavant d'aborder le projet avec un œil nouveau. Ils encouragent à revisiter non seulement, les stratégies de clonage, mais aussi à réajuster le design du génome semi-synthétique *SynBsu*.

Développement de la méthode de CREasPy-Fusion. Le choix de la méthode de clonage est primordial en fonction du génome que l'on souhaite cloner et de la nature de l'organisme. Les essais répétés pour adapter les méthodes déjà décrites dans la littérature et tenter de cloner le génome entier de *B. subtilis*, nous ont poussé à innover et à proposer une nouvelle méthodologie dont la mise en place fait l'objet du chapitre 3. La CREasPy-Fusion vient donc étoffer le catalogue des méthodes de clonage disponibles (**Tableau 47**). D'un point de vue propre au projet ANR *Bacillus* 2.0, cette stratégie nous permet d'envisager des perspectives à

Tableau 47 : Comparaison des méthodes de clonage de génomes bactériens dans la levure.

| | Clonage Classique | Clonage par Fusion | TAR-Cloning | CReasPy-Cloning | CReasPy-Fusion |
|--|--|--|--|---|--|
| Génomes bactériens clonés | <i>M. pneumoniae</i> ; <i>M. mycoides</i> subsp. <i>capri</i> <i>M. genitalium</i> (Benders et al. 2010) <i>M. mycoides</i> subsp. <i>mycoides</i> <i>M. leachii</i> ; <i>M. putrefaciens</i> ; <i>S. citri</i> (Labrousseau et al. 2017) <i>M. florum</i> (Baby et al. 2018) | <i>M. capricolum</i> subsp. <i>capricolum</i> , <i>M. mycoides</i> subsp. <i>capri</i> ; <i>H. influenzae</i> (Karas et al. 2013) JCVISyn1.0 et JCVISyn2.0 (mutants) (Karas et al. 2019) | <i>M. genitalium</i> (Benders et al. 2010) <i>A. laidlawii</i> (Karas et al. 2011) <i>P. marinus</i> (Tagwerker et al. 2012) <i>M. hominis</i> (Rideau et al. 2018) | <i>M. pneumoniae</i> ; <i>M. ferruginatoris</i> , <i>M. mycoides</i> subsp. <i>mycoides</i> <i>M. leachii</i> (Ruitz et al. 2019) <i>M. capricolum</i> subsp. <i>capripneumoniae</i> (Gourgues et al. 2022 non publié) <i>M. capricolum</i> subsp. <i>capricolum</i> <i>M. mycoides</i> subsp. <i>capri</i> (MolliLab) | <i>M. capricolum</i> subsp. <i>capricolum</i> <i>M. mycoides</i> subsp. <i>capri</i> <i>M. capricolum</i> subsp. <i>capripneumoniae</i> <i>M. mycoides</i> subsp. <i>mycoides</i> <i>M. gallisepticum</i> <i>M. agalactiae</i> (Guesdon et al. 2022 non publié) |
| Fragments de génomes bactériens clonés | | | Fragments de génome de <i>S. elongatus</i> (Noskov et al. 2012) Fragments de génome de <i>B. subtilis</i> (MolliLab) | | Fragments de génome de <i>B. subtilis</i> (Guesdon et al. 2022-2023 non publié) |
| Avantages | -Ne nécessite pas de connaître la séquence du génome | -Ne nécessite pas de connaître la séquence du génome -Ne nécessite pas de préparer les chromosomes bactériens en bloc d'agarose | -Permet de capturer des fragments de génomes ou des génomes entiers -Méthode de contre sélection URA3-SFOA -Pré-digestion par CRISPR <i>in vitro</i> -Ne nécessite pas de savoir pré-transformer la bactérie, pas besoin d'avoir des outils à disposition | -Permet de cloner des génomes entiers de faire et de faire de la multi-ingénierie -Ne nécessite pas de savoir pré-transformer la bactérie, pas besoin d'avoir des outils à disposition | -Permet de cloner des génomes entiers de faire, de faire de la multi-ingénierie et de capturer des fragments de génome -Ne nécessite pas de savoir pré-transformer la bactérie, pas besoin d'avoir des outils à disposition |
| Inconvénients | -Nécessite de savoir pré-transformer la bactérie. -Outils à disposition pour intégrer les éléments levures (ex. transposon) | -Nécessite de savoir pré-transformer la bactérie. -Outils à disposition pour intégrer les éléments levures (ex. transposon) -Fusion du contenu cellulaire (systèmes RM, nucléases pouvant diminuer/inhiber l'efficacité de la méthode) | -Nécessite de connaître la séquence du génome | -Nécessite de connaître la séquence du génome | -Nécessite de connaître la séquence du génome |

court terme pour (i) capturer des fragments de génomes de plus en plus grands, (ii) éliminer les portions de génomes instables chez la levure (*e.g.* « Fragment 7 ») ou (iii) ajouter des ARS.

D'un point de vue plus large, ces données contribuent également à étoffer les méthodologies de BS disponibles et à les ouvrir pour des génomes pas encore clonés ou très grands grâce à la Fusion. Nous avons observé que pour cloner le génome de *B. subtilis* les méthodes de clonage Classique et par CReasPy-Cloning ne semblaient pas les plus appropriées. Ces dernières ont en commun l'extraction et la purification des chromosomes à cloner en bloc d'agarose (pour éviter les cassures mécaniques). Or, il nous est apparu qu'une barrière technique pouvait limiter le passage du chromosome de *B. subtilis* qui serait alors trop décondensé (déprotéinisation de la molécule d'ADN). En revanche, avec la méthode de Fusion, il semblerait que nous ayons réussi à transférer ce génome (dans sa conformation intacte et condensée depuis une cellule bactérienne) vers des sphéroplastes de levure. Hormis les travaux de Karas *et al.* sur la méthode de Fusion, dont nous nous sommes inspirés pour développer la CReasPy-Fusion, nous n'avons pas d'autres exemple dans la littérature de clonage par Fusion de chromosomes bactériens dans la levure^{223,237}. En revanche, une étude effectuée par Brown *et al.* en 2017, a montré que le transfert de large molécules d'ADN (1,1 Mb) depuis la levure vers des cellules eucaryotes était possible par fusion au PEG. Les auteurs ont défini que le transfert s'opérait de manière plus efficace par cette méthode tout en évitant des cassures mécaniques et la purification en bloc d'agarose³³⁶. Cet exemple est un peu différent que ce que nous souhaitons démontrer mais cela renforce l'idée selon laquelle le transfert de molécules de grandes tailles vers des cellules de mammifères doit être effectué avec une méthode adaptée. Il serait intéressant de tester des gammes de tailles de molécules d'ADN que l'on sait stables dans la levure (1 Mb, 1,5 Mb, 2 Mb et 2,5 Mb, 3 Mb) et de comparer les méthodes de clonage Classique et par Fusion. L'idée étant de pouvoir démontrer à partir de quand (ou de quelle taille) nous n'obtenons plus de transformants levure.

Finalement et à l'instar de ce qui a déjà pu être proposé pour les méthodes de CReasPy-Cloning²²⁰ ou de TAR-Cloning²⁸¹, la CReasPy-Fusion pourrait aussi être utilisée pour cloner des génomes d'organismes non transformables, ou issus d'une population microbienne mixte.

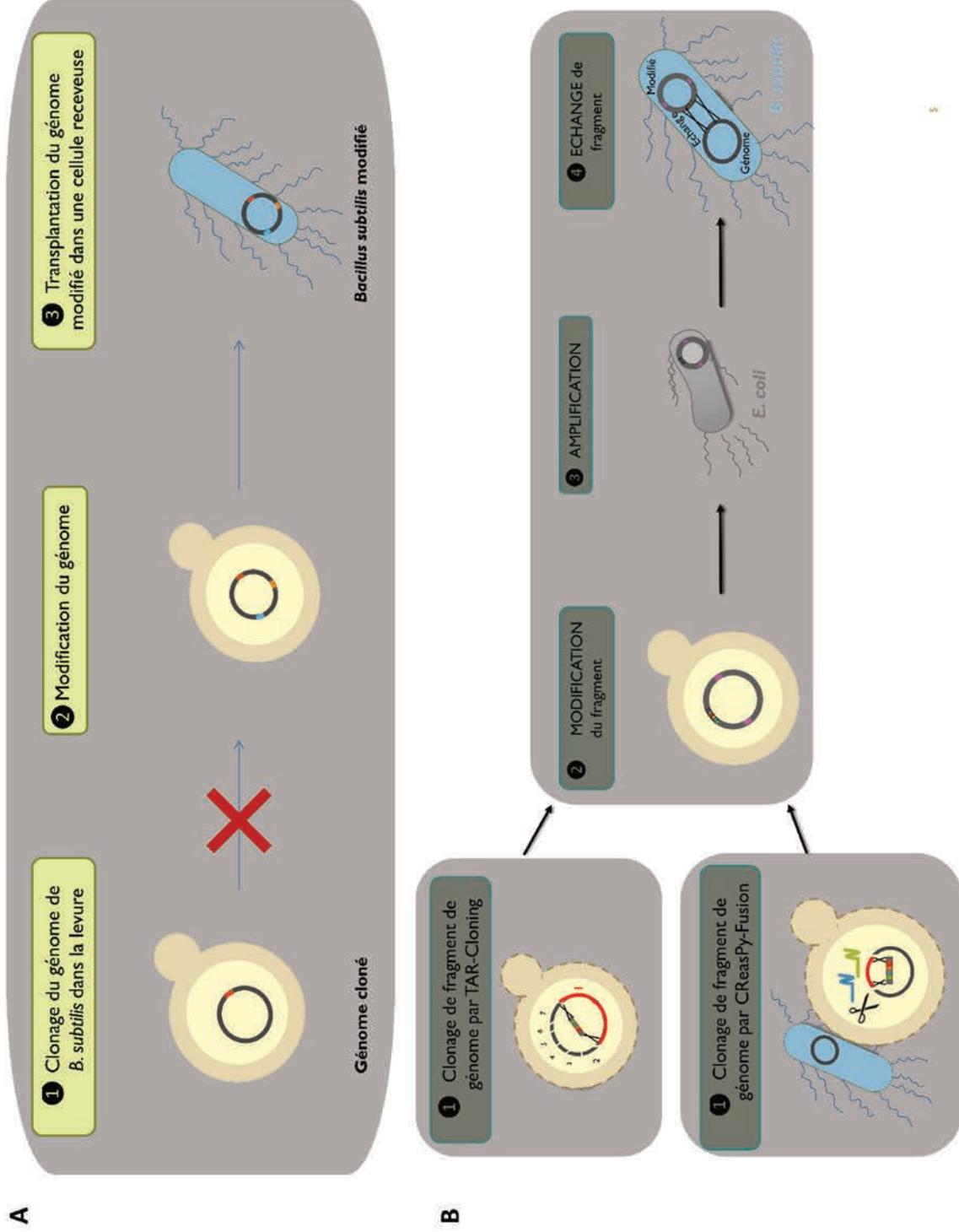


Figure 81 : Alternative proposée à la stratégie originale du projet ANR « *Bacillus 2.0* ». [A] Stratégie initiale représentée avec le clonage du génome dans la levure, sa modification et la transplantation dans une cellule receveuse. [B] Stratégie alternative associant les méthodes de clonage de fragments par TAR-Cloning ou par CRISPR-Cas9, à la méthode RAGE.

Alternatives à la transplantation de génomes et par conséquent à la stratégie initiale du projet *Bacillus* 2.0. Le génome de *B. subtilis*, n'ayant pas pu être cloné entier dans la levure au cours de ces travaux, les essais de transplantation de génome (levures-bactéries) n'ont pas pu être amorcés (**Figure 81A**). En parallèle des expériences sur le développement de méthodologies de clonage de génomes dans la levure, nous avons initié la mise en place d'une stratégie nommée RAGE pour « Recombinase-Associated Genomic Exchange » développée récemment au laboratoire²⁵⁵. Cette méthode permet le transfert, non pas d'un génome entier, mais de portions de génomes de la levure vers la bactérie à éditer. Elle nécessite quatre étapes: (i) la capture entre des sites *lox* incompatibles d'un grand fragment d'ADN génomique bactérien d'intérêt dans un vecteur navette YAC-BAC grâce à la méthode de TAR-Cloning ; (ii) la modification du fragment capturé dans la levure en utilisant les outils génétiques disponibles ; (iii) la transformation du fragment dans *E. coli*, pour permettre son amplification ; (iv) l'échange de fragment d'ADN grâce à une recombinase site-spécifique³³⁷. Cette méthode transposée au projet *Bacillus* 2.0 permet de proposer une stratégie alternative en l'associant avec, soit la méthode de capture de fragments de génome par TAR-Cloning, ou soit la CReasPy-Fusion (**Figure 81B**). C'est dans le cadre du stage de Master 2 de Julien Barret que nous avons pu commencer à mettre en place des expériences qui sont maintenant reprises et poursuivies pour son sujet de thèse. L'idée étant de réutiliser la collection de levures propageant le génome de *B. subtilis* sous forme de fragments et produite par TAR-Cloning. Un des points bloquants est que nous sommes limités par la position des sites de restriction pour le choix du *locus* à capturer. De ce fait, la méthode de CReasPy-Fusion grâce à laquelle nous avons démontré qu'il était possible de capturer n'importe quel fragment de génome offre une voie plus intéressante pour débiter cette stratégie. De même, nous pourrions aussi imaginer travailler avec des fragments synthétiques assemblés par nos collaborateurs. Cette approche itérative pourrait avoir l'avantage de permettre de réaliser plus facilement un cycle « Design – Build – Test – Learn ». Il sera en effet possible de tester l'impact de chaque section modifiée sur la viabilité de la cellule bactérienne en créant une souche « mosaïque » contenant des sections natives et des sections synthétiques. Des approches similaires sans forcément passer par la levure sont retrouvées dans la littérature. Lau *et al.* ont par exemple développé une stratégie nommée SIRCAS (Stepwise Integration of Rolling Circle Amplified Segments) basée sur le système de recombinaison lambda red. Cela a permis le remplacement itératif de fragments synthétiques de 10 à 25 kpb dans le génome de *Salmonella typhimurium*³³⁸. Cette

stratégie a également été utilisée au cours de la construction de cellules contrôlées par des génomes synthétiques (JCVI-Syn1.0 et JCVI-Syn3.0)^{172,218,337}. Très récemment un autre groupe a utilisé la méthode SCRaMble (Synthetic Chromosome Rearrangement and Modification by LoxP-mediated Evolution) pour intégrer un fragment synthétique modifié d'environ 50 kpb chez *Corynebacterium glutamicum*³³⁹. Elles illustrent le potentiel et les avantages de l'ingénierie de génome à grande échelle. Néanmoins, les étapes d'assemblage de fragments synthétiques (dans la levure) ou bien le fait que des bactéries à génome réduit soient moins transformables, justifie de pouvoir développer une approche telle que RAGE dans le cadre du projet *Bacillus 2.0*, en offrant notamment la possibilité de produire des souches de *B. subtilis* mutantes avec un génome réduit et remanié.

Pour conclure, ces travaux sont finalement la genèse de ce qui a pu être initié au sein de l'équipe Mollicutes pour le projet ANR *Bacillus 2.0*. Dorénavant les protocoles nécessaires à la manipulation de ce nouvel organisme dans notre équipe sont disponibles et une routine a donc pu être installée au laboratoire. Il s'agissait aussi d'adapter à *B. subtilis*, les méthodologies de BS maîtrisées au laboratoire sur les Mollicutes et des avancées ont été rendues possibles : des morceaux de génomes de *B. subtilis* ont été clonés et une nouvelle méthode de clonage de génome a été développée. Les données récoltées sont prometteuses et de futures directions sont envisagées pour atteindre l'objectif final du projet vers le développement de châssis semi-synthétiques et minimaux de *B. subtilis*.

ANNEXES

ANNEXES

Tableau S1 : Tableau récapitulatif des séquences spacers et plasmide pARNg associés.

| Expériences | Séquences ARNg « spacer » | PAM | Locus cibles | Organismes | Plasmides | |
|----------------------------------|---------------------------|------------------------|--------------------------------|----------------------------------|-------------------------------|----------------------------|
| CReasPy-Cloning | ACAAGCGATCGATTCCAGG | CGG | <i>sboAX-olbAG</i> | <i>B. subtilis</i> | p426-SNR52p-gRNA1.sbo-alb | |
| | GAAACCGATCGAAAAGCGCGG | CGG | | | p426-SNR52p-gRNA2.sbo-alb | |
| | AATCTGCCCCGATCAGAACAG | CGG | <i>spoVAA-VAf</i> | | p426-SNR52p-gRNA1.spoVAA-F | |
| | ACATCAGGCTGAATGGACCA | GGG | | | p426-SNR52p-gRNA2.spoVAA-F | |
| CReasPy-Fusion | ATTCGTATACAGCTGACCG | CGG | <i>yjcD</i> | | p426-SNR52p-gRNA1.yjcD | |
| | TGTAAAACGATTGGACGCCG | CGG | | | p426-SNR52p-gRNA2.yjcD | |
| | GTTTTGATAGATTCTGACAG | CGG | <i>yvcB</i> | | p426-SNR52p-gRNA1.Fgt4-2,5 | |
| | TGCTGAGATCGAAAAAGCCG | CGG | <i>bmrA</i> | | p426-SNR52p-gRNA2.Fgt4-2,5 | |
| | ACTCAGCAAAAAGTAAACCG | CGG | <i>atpA</i> | | p426-SNR52p-gRNA1.Fgt4-2,64 | |
| | GCAATGTTTCTACTACGCCG | CGG | | | p426-SNR52p-gRNA2.Fgt4-2,64 | |
| | GCATTGATAACGAAATTGCG | CGG | <i>lrgB</i> | | p426-SNR52p-gRNA1.Fgt5-2,78 | |
| | AGACGGACTTTAATGCCATG | CGG | <i>ywbE</i> | | p426-SNR52p-gRNA2.Fgt5-2,78 | |
| | GCTTGGACTGAACGAAGTGA | GGG | <i>repE</i> | p426-SNR52p-gRNA1.repE | | |
| | TCAAGTCTAACGGTTTTAAG | CGG | | p426-SNR52p-gRNA2.repE | | |
| | TTAATTTATCGTCTAACAGG | TGG | <i>glpK</i> | p426-SNR52p-gRNA2.glpOKF | | |
| | CReasPy-Fusion | ACTACAACAAGACCAAGAGA | AGG | RS01980 | <i>Mmm</i> | p426-SNR52p-gRNA1.1980 |
| | | CCACTTCAACTTTCACACGG | TGG | RS01995 | | p426-SNR52p-gRNA2.1995 |
| | | GCTCCACTTCAACTTTCACA | CGG | RS01995 | | p426-SNR52p-gRNA3.1995 |
| | | TTATTTTTACAAAAGAAGT | GGG | RS01995 | | p426-SNR52p-gRNA4.1995 |
| | | AAAATTGAAGTATCATCAAG | TGG | MCAP0050 | | p426-SNR52p-gRNA1.RE-CCATC |
| | | CCTACTTTTACTGCAAAAAGG* | TGG | MCAP0050* | p426-SNR52p-gRNA2.RE-CCATC | |
| | | CCTACTTTTACTGCAAAAAGG* | CGG | MMCAP2_100* | <i>Mmc</i> | p426-SNR52p-gRNA2.RE-CCATC |
| TTAGATGTTAGTTTTAATAG | | TGG | MCCP001_RS01320 FOY67_01295 | <i>Mccp</i> | p426-SNR52p-gRNA2.peptS41 | |
| GTAGGGCTAGTTGAATTTAA | | GGG | MBOVPG45_0215 | <i>M. bovis</i> | p426-SNR52p-gRNA1_bovis0215 | |
| GATAGTCCCGGGAAGAAAAG | | AGG | | | p426-SNR52p-gRNA2_bovis0215 | |
| GTAGGATTGGTCAATTTAA | GGG | MAG_RS03005 | <i>M. agalactiae</i> | p426-SNR52p-gRNA1_agalactiae5900 | | |
| AGTGAATTTATACATCAGCG | GGG | | | p426-SNR52p-gRNA1_agalactiae5900 | | |
| ATAAATCCACCTCTTCGCTT | TGG | <i>cysP</i> | <i>M. gallisepticum</i> | p426-SNR52p-gRNA1_gallicysP | | |
| TGTTCTTTCTTAGTATCAA | GGG | | | p426-SNR52p-gRNA2_gallicysP | | |
| Edition de base (CRISPR Editing) | AACCAAAAATAGACTTAGT | AGG | MBOVPG45_0215 | <i>M. bovis</i> | pMT85_SpdCas9_pmcDA1/0215 | |
| | ACTAAATCATAGTGTGTG | AGG | MAG_RS03005 | <i>M. agalactiae</i> | pMT85_SpdCas9_pmcDA1/RS03005 | |
| | TGAACAATATAAAAAAGTTG | GGG | GCW_RS00070 | <i>M. gallisepticum</i> | pTI4.0_SpdCas9_pmcDA1/RS00070 | |

Tableau S2 : Tableau récapitulatif des plasmides pDT-ARNg.

| Expériences | Organismes | Plasmides | |
|------------------------|--------------------|-----------------------------|---|
| CReasPy-Fusion | <i>B. subtilis</i> | p426-SNR52p-gRNA1.Fgt4-2,5 | p426-gRNA-Ura3-double-target_Fgt4 [A, B, C ou D] |
| | | p426-SNR52p-gRNA2.Fgt4-2,5 | |
| | | p426-SNR52p-gRNA1.Fgt4-2,64 | |
| | | p426-SNR52p-gRNA2.Fgt4-2,64 | |
| | | p426-SNR52p-gRNA1.Fgt4-2,64 | p426-gRNA-Ura3-double-target_Fgt5 [A, B, C ou D] |
| | | p426-SNR52p-gRNA2.Fgt4-2,64 | |
| | | p426-SNR52p-gRNA1.Fgt5-2,78 | p426-gRNA-Ade2-double-target_Fgt5 [E, F, G ou H] |
| | | p426-SNR52p-gRNA2.Fgt5-2,78 | |
| | <i>Mmm</i> | p426-SNR52p-gRNA2.glpOKF | p426-gRNA-Ura3-double-target_GRS [A, B, C ou D] |
| | | p426-SNR52p-gRNA1.1980 | |
| | | p426-SNR52p-gRNA2.1995 | |
| | | p426-SNR52p-gRNA3.1995 | |
| p426-SNR52p-gRNA4.1995 | | | |

Tableau S3 : Tableau récapitulatif des amorces utilisées pour la construction des plasmides.

| Plasmides | Noms des amorces | Séquences amorces 5' → 3' | Taille pb | Matrices |
|---|--------------------------|---|-----------|---|
| p426-gRNA-Ura3-double-target <i>ou</i> p426-gRNA-Ade2-double-target | p426_insertGibson_F | GTACAATCTGCTCTGATGCCGGCTGGAGCTTCTTTGAAAAG | 41 | p426-SNR52p-gRNA+AarI-SUP4t ou p426-SNR52p-gRNA+AarI-SUP4t_Ade2 |
| | p426_insertGibson_R | GTCGGGGCTGGCTTAACCTATGGTACCGGCCCAAATTAAG | 41 | |
| | Div_p426_vecteurGibson_F | CTTTAATTTGCGGCCGGTACCATAGTTAAGCCAGCCCCGAC | 41 | |
| | Div_p426_vecteurGibson_R | CTTTTCAAAGAAGCTCCAGCCGGCATCAGAGCAGATTGTAC | 41 | |
| pMT85-pRS313-pSPEC | Oligo_Spec_NsiI_F1 | GATCATGCATTGCCTCAAGCTAGAGAGTCG | 30 | pDR111 |
| | Oligo_Spec_AgeI_R3 | GATCACCGGTCCGCCAGCCTAAACGGATGA | 31 | |
| pGEM-T-ACH-Spec-1kb | pMT85_F | CAGGGTTATTGTCTCATGAGCG | 22 | pMT85-pRS313-pSPEC |
| | pMT85_R | GGTCATGCCTTCGTAGATC | 19 | |
| | F2'_overlap | GATCTACGAAGGCATGACCTCCTATGAAAACGCATCGTC | 39 | ADNg <i>B. subtilis</i> |
| | R2_overlap | ATGCGGCGGTAATGTTCTCG | 21 | |
| | F3_overlap | GGAGAAGCAAGCGGCATC | 18 | ADNg <i>B. subtilis</i> |
| | R3'_overlap | CGCTCATGAGACAATAACCCTGTTCTGTCGACGATGAGCTG | 40 | |
| pMT85-pRS313-pMiniTetM-pSPEC sans lacZ-sbo | pMT85_ACH_sbo_F | GTACAACATAGATCTGCTAGAAAAGTGCCACCAGGACG | 38 | pMT85-pRS313-pSPEC |
| | pMT85_ACH_SPEC_sbo_R | TCTGCGATTGCGATGTAGCCAGCCTAAACGGATGATC | 38 | |
| | pMT85_Ori-Amp_sbo_F | CCGTCTCCTTGCCATATTGAACCTGTTTGCCAGCC | 36 | pMiniTetM |
| | pMT85_Ori-Amp_sbo_R | CTGATCCGTCTGTACTGGAGGGAAATGTGCGCGGAAC | 38 | |
| | Bsu_sbo_début_F | GTTCCGCGCACATTTCCCTCCAGTACAGACGGAATCAG | 38 | ADNg <i>B. subtilis</i> |
| | Bsu_sbo_début_R | CGTCCTGGTGGCACTTTTCTAGCAGATCTATGTTGTAC | 38 | |
| | Bsu_sbo_fin_SPEC_F | GATCATCCGTTTAGGCTGGGCTACATCGCAATCGCAGA | 38 | ADNg <i>B. subtilis</i> |
| | Bsu_sbo_fin_R | GGCTGGCAAACAGGTTCAATATGGACAAGGAGACGG | 36 | |

En gris sont représentés les sites de restriction inclus dans les séquences des amorces

Tableau S4 : Tableau récapitulatif des amorces utilisées pour la construction des plasmides (suite).

| Plasmides | Noms des amorces | Séquences amorces 5' → 3' | Taille pb | Matrices |
|----------------------------|------------------|------------------------------------|-----------|---------------|
| pShuttle_Bsu-Levure | pMT85_BamHI_F | GATCGGATCCCAGGGTTATTGTCTCATGAGCGG | 33 | pMT85-pRS313- |
| | pMT85_SpeI_R | GATCACTAGTTGGTCATGCCTTCGTAGATC | 30 | pSPEC |
| pGG01 | Cat_EcoRI_F | GATCGAATTCTGACCTCAGAACTCCATCTG | 30 | pBSB9 |
| | Cat_HindIII_R : | GATCAAGCTTCACACAGGAAACAGCTATGAC | 31 | |
| | Ade2_PstI_F1 | GATCCTGCAGGCAGATTGTAAGTACTGAGAGTGC | 30 | pICE-Ade2 |
| | Ade2_EcoRI_R1 | GATCGAATTCTTCTCCTTACGCATCTGTGC | 30 | |
| | ARS413_BamHI_F1 | GATCGGATCCGGAGAAGTCGGAATCGTTGC | 30 | pFL39 |
| | ARS413_PstI_R1 | GATCCTGCAGGAAGTGAGGAAACCTAGTAG | 30 | |
| pGG03 | Ery_EcoRI_F1 | ATGCGAATTCGATCTTTTCTACGGGGTCTG | 30 | pJIM |
| | Ery_HindIII_R1 | GATCAAGCTTTAATCGCCTTGACGACATC | 30 | |
| | Lys2_PstI_F | GATCCTGCAGTTCTCCTTACGCATCTGTGC | 30 | pICE-Lys2 |
| | Lys2_EcoRI_R | GATCGAATTCTTGGTCAGTATCGACCTGTG | 30 | |
| | ARS1_BamHI_F | GATCGGATCCACTGGGTTGGAAGCAAG | 30 | pCM185 |
| | ARS1_PstI_R | GATCCTGCAGATGATAAGCTGTCAAACATGAG | 32 | |

En gris sont représentés les sites de restriction inclus dans les séquences des amorces

Tableau S5 : *Loc*i coupés par l'enzyme SgrDI expérience de TAR-Cloning des Fragments 1 à 7.

| Position du site SgrDI | Gènes | Fonction |
|------------------------|-------------|---|
| 403 003 pb | <i>swrC</i> | Transporteur impliqué dans l'auto-résistance à la surfactine |
| 1 902 489 pb | <i>nadR</i> | Répresseur transcriptionnel de la biosynthèse de NAD <i>de novo</i> |
| 2 293 436 pb | <i>yufO</i> | ABC transporteur de carbohydrate putatif (ATP-binding protein) |
| 2 508 066 pb | <i>yvcB</i> | Protéine hypothétique conservée |
| 2 649 798 pb | <i>atpD</i> | ATP synthetase (sous-unité β composant F1) |
| 2 782 869 pb | <i>ywbD</i> | Methyltransférase ADOMet-dependante putative |
| 2 788 715 | - | Région intergénique entre <i>epr</i> et <i>sacX</i> |

Tableau S6 : *Loc*i coupés par l'enzyme FseI expérience de TAR-Cloning des Sous-Fragments 1.1 à 1.7.

| Position du site SgrDI | Gènes | Fonction |
|------------------------|-------------|--|
| 381 937 pb | <i>yecA</i> | Transporteur |
| 770 029 pb | <i>argD</i> | Acetyl-ornithine transaminase |
| 853 568 pb | <i>yjfA</i> | Inconnue |
| 1 368 407 pb | <i>alsT</i> | Transporteur putatif de la glutamine |
| 1 527 520 pb | <i>cca</i> | Nucleotidyltransférase de l'ARNt Cys |
| 1 639 348 pb | <i>yqjV</i> | Proteine putative impliquée dans la résistance à des médicaments |
| 1 769 662 pb | <i>yqfO</i> | Protéine similaire à la GTP cyclohydrolase |
| 2 039 637 pb | <i>accA</i> | Acetyl-CoA carboxylase |

Tableau S7 : Espèces de mycoplasmes et cibles utilisées pour les expériences de CREasPy-Fusion.

| Espèces de mycoplasmes | | Souches | Type de souche et isolats | GamBank # | Taille génome [Mbp] | % G+C | Hôtes naturels | Mnémonique (ancien "locus tag") | Mnémonique (nouveau "locus tag") | Nom de gène | Produit | Localisation | | |
|--|-----------------------------|-------------------------------------|---------------------------|-----------|---------------------|------------------|----------------|--|----------------------------------|--|--|--------------------------------|-----------------|------------------------------|
| <i>M. capricolium subsp capricolium (Mcap)</i> | ATCC 27343 (California Kid) | Souche type - Etats-Unis (1955) | CP000123.1 | 1,1 | 23,7 | | | MCAP_0050 | MCAP_RS00270 | | Endonucléase de restriction de type II "CCATC" | Genome (63593 to 65170 bp) | | |
| | 14232 | Souche de terrain - France (2004) | JF0000000.000.1 | 1,1 | 23,5 | Ruminants | | Mcapa_s03_0930 (3.0) MCApa_0960 (4.0) | | | Protéine hypothétique : prédiction protéine transmembranaire avec queue spécifique de protéase | Contig 3 (129047 to 131017 bp) | | |
| <i>M. mycoides subsp capri (Mmc)</i> | GM12 | Souche type - Etats-Unis (1979) | CP001668.1 | 1,1 | 23,9 | Ruminants | | MMCAP2_0100 | MMCAP3_RS00520 | | Endonucléase de restriction de type II "CCATC" | Genome (138695 to 140278 bp) | | |
| | 95010 | Souche de terrain - France (1995) | FQ377874.1 | 1,1 | 23,8 | | | | MLC_0920 | | Endonucléase de restriction putative | Genome (123855 to 125435 bp) | | |
| <i>M. capricolium subsp capripneumoniae (Mccp)</i> | 95043 | Souche de terrain - Niger (1995) | CP041705.1 | 1 | 24 | Petits Ruminants | | | FOY67_01295 | | Peptidase S41 | Genome (317661 to 319631 bp) | | |
| | 14020 | Souche de terrain - Tanzanie (2013) | Soon | 1 | 24 | | | | | | Peptidase S41 | IND | | |
| <i>M. mycoides subsp mycoides (Mmm)</i> | T1/44 | Souche vaccinale (1950) | CP054256.1 | 1,2 | 24,0 | Ruminants | | MSCT144_09810-09820-09830 | MSCT144_RS04705-RS04710-RS04715 | glpOKF | Glycerol phosphate oxydase, glycerol kinase, protéine putative entrée glycérol | Genome (1027061 to 1030573 bp) | | |
| | | | | | | | | | | | | MSCT144_04130 | MSCT144_RS01980 | Genome (436854 to 438746 bp) |
| | | | | | | | | | | | | MSCT144_04160 | MSCT144_RS01995 | Genome (440999 to 441838 bp) |
| <i>M. bovis (Mbov)</i> | ATCC 25523 PG45 | Souche type | CP002188.1 | 1 | 29,3 | Ruminants | | MBOYPG45_0215 | mnuA | Nucléase de membrane A | Genome (248878 to 250017 bp) | | | |
| <i>M. agalactiae (Maga)</i> | PG2 | Souche type - Espagne (1952) | CU179680.1 | 0,87 | 29,7 | Petits ruminants | | MAG_5900 | MAG_RS03005 | Protéine hypothétique (Homologue de MBOYPG45_0215) | Genome (681807 to 683021 bp) | | | |
| <i>M. gallisepticum (Mgal)</i> | ATCC 15302 S6 | Souche type | CP006916.3 | 0,98 | 31,4 | Volaille | | GCW_01680 | GCW_RS01695 | cysP | Peptidase C1 | Genome (391878 to 393572 bp) | | |
| | | | | | | | | | | | | GCW_00060 | GCW_RS00070 | mnuA |

Tableau S8 : Récapitulatif des systèmes de restriction (type I et III) chez certaines espèces de mycoplasmes.

*Gène coupé par une transposase ISMbov1 (MBOVPG45_0764). **Gène coupé par une transposase ISMbov7 (MBOVPG45_0616).

***Décalage du cadre de lecture, stop interne à la séquence. §ADN-méthyltransférase site spécifique avec un décalage du cadre de lecture.

†Souches comparatives.

| Mycoplasmes | Souches | Systèmes de restriction-modification | | | | | | |
|-------------|----------|--|----------------------------------|-------------|----------------------------|----------------------------|---------------------------------|---------------|
| | | Type I | | | | | Type III | |
| | | Inconnu | Inconnu | Inconnu | AGGNNGNT | Inconnu | YGAG (Inconnu) | Inconnu |
| <i>Mcap</i> | CK | | | | | | | |
| | 14232 | | | | | | | MCAPa_2790 |
| <i>Mmc</i> | GM12 | | | | | | MMCAP2_RS02925 (MMCAP2_0590) | |
| | 95010 | | | | | | | |
| <i>Mccp</i> | Abomsaf† | MCCP001_RS00260 (MCCP01_0055- 0056-0057-0058)*** | | | | | | |
| | 95043 | | | FOY67_00275 | | | | |
| | 14020 | IND | | | | | | |
| <i>Mmm</i> | T1/44 | | | | | | | |
| <i>Mbov</i> | RM16† | FWK06_001030 | | | | | FWK06_000845§ | |
| | PG45 | MBOVPG45_0239 | MBOVPG45_0615 MBOVPG45_0617** | | | | MBOVPG45_0167 | MBOVPG45_0439 |
| <i>Maga</i> | PG2 | MAG_RS02915 (MAG5700 ou 5710) | | | | | MAG_RS00775 (MAG1520) | MAG_3480 |
| <i>Mgal</i> | S6 | | | | GCW_RS02385 (GCW_02365) | GCW_RS03820 (GCW_03825) | | |

Tableau S9 : Récapitulatif des systèmes de restriction (type II) chez certaines espèces de mycoplasmes.

*Gène coupé par une transposase ISBov1 (MBOVPG45_0764). **Gène coupé par une transposase ISMbov7 (MBOVPG45_0616).

***Décalage du cadre de lecture, stop interne à la séquence. §ADN-méthyltransferase site spécifique avec un décalage du cadre de lecture. †Souches comparatives.

| Mycoplasmes | | Systèmes de restriction-modification | | | | | | | | | | | | |
|-------------|---------|--------------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|----------------------------------|-----------------------------|----------------------------------|----------------------------------|------------------------|-------------|----------------------------------|-------------------------------|
| Souches | | Type II | | | | | | | | | | | | |
| | | CTAG | CCATC (BccI) | CC TTC (HpyAV) | GCATC (SmaI) | GAATC (HinfI) | GATC (Aat256I) | TCTAGA (KbaI) | GGNCC (Sau96I) | Inconnu (putative BglCI) | GGNCC (Eco47II family) | Inconnu | CCTC (MniI) | GNNGC |
| <i>Mcap</i> | CK | | MICAP_RS00270 (MICAP_0050) | | | | | | | | | | | |
| | 14232 | | | MICAPs_3430 | | | | | MICAPs_1280 | | | | MICAPs_7780 | |
| <i>Mmc</i> | GM12 | | MIMCAP2_RS00520 (MIMCAP2_0100) | MIMCAP2_RS03655 (MIMCAP2_0741) | MIMCAP2_RS03730 (MIMCAP2_0756) | MIMCAP2_RS03790 (MIMCAP2_0769) | MIMCAP2_0055 | | | | | | | |
| | 95010 | | MILC_RS00485 (MILC_0920) | | | MILC_RS09455 (MILC_7780) | MILC_RS00255 (MILC_0510) | MILC_RS05085 (MILC_1650) | MILC_RS01050 (MILC_2090) | MILC_6040 | | | | |
| <i>Mcep</i> | Abomsal | | | | | | MCCP001_RS01930 (MCCP01_0438) | | MCCP001_RS01160 (MCCP01_0262) | | | | | |
| | 95043 | | | | | | | | | | | | | |
| 14020 | | | | | | | | | | | FOV67_01140 | FOV67_01885 | | SoqH |
| <i>Mmm</i> | T1/44 | | | | | | | | | | | | | |
| | | | | | | | | | | | | | MST144_RS01115 (MST144_02320) | |
| <i>Mbov</i> | RM16f | | | | | | | | MST144_RS05085 (MST144_10660) | MST144_RS04935 (MST144_10330) | | | | |
| | PG45 | FWK06_001700 MBOVPG45_0366 | | | FWK06_002845 MBOVPG45_0721 | MBOVPG45_0763 MBOVPG45_0765* | MBOVPG45_0560 | | | | | | | FWK06_002380 MBOVPG45_0468 |
| <i>Maga</i> | PG2 | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | |
| <i>Mgal</i> | S6 | | | | | | | | | | | | | |

RÉFÉRENCES BIBLIOGRAPHIQUES

RÉFÉRENCES BIBLIOGRAPHIQUES

1. Paul De Vos, George M. Garrity, D. J., Noel R. Krieg, Wolfgang Ludwig, F. A. R. & Whitman, K.-H. S. and W. B. Bergey's Manual of Systematic Bacteriology: The Firmicutes, 2nd edn. *Springer* **75**, 1058–1061 (2009).
2. Mandic-Mulec, I., Stefanic, P. & van Elsas, J. D. Ecology of *Bacillaceae*. *Microbiol. Spectr.* **3**, (2015).
3. Moyes, R. B., Reynolds, J. & Breakwell, D. P. Differential staining of bacteria: gram stain. *Curr. Protoc. Microbiol.* (2009).
4. M. H. Soule. Identity of *Bacillus subtilis*, Cohn 1872. *J. Infect. Dis.* **51**, 191–215 (1932).
5. Baldwin, V. M. You can't *B. cereus* - A review of *Bacillus cereus* strains that cause anthrax-like disease. *Front. Microbiol.* **11**, 1731 (2020).
6. Jovanovic, J., Ornelis, V. F. M., Madder, A. & Rajkovic, A. *Bacillus cereus* food intoxication and toxicoinfection. *Compr. Rev. Food Sci. Food Saf.* **20**, 3719–3761 (2021).
7. Priest, F. G. *Bacillus subtilis* and other Gram-positive bacteria. *ASM Press* 1–16 (1993).
8. Höfte, H. & Whiteley, H. R. Insecticidal crystal proteins of *Bacillus thuringiensis*. *Microbiol. Rev.* **53**, 242–55 (1989).
9. Nicholson, W. L., Munakata, N., Horneck, G., Melosh, H. J. & Setlow, P. Resistance of *Bacillus Endospores* to Extreme Terrestrial and Extraterrestrial Environments. *Microbiol. Mol. Biol. Rev.* **64**, 548–572 (2000).
10. Errington, J. & van der Aa, L. T. Microbe profile: *Bacillus subtilis*: Model organism for cellular development, and industrial workhorse. *Microbiol. (United Kingdom)* **166**, 425–427 (2020).
11. Zeigler, D. R. *et al.* The origins of 168, W23, and other *Bacillus subtilis* legacy strains. *J. Bacteriol.* **190**, 6983–6995 (2008).
12. Burkholder P R & Giles, N. H. Induced biochemical mutations in *Bacillus subtilis*. *Am. J. Bot.* **34**, 345–8 (1947).
13. Spizizen, J. Transformation of biochemically deficient strains of *Bacillus subtilis* by deoxyribonucleate. *Proc. Natl. Acad. Sci. U. S. A.* **44**, 1072–8 (1958).
14. Anagnostopoulos, C. & Spizizen, J. Requirements for transformation in *Bacillus subtilis*. *J. Bacteriol.* **81**, 741–6 (1961).
15. Kunst, F. *et al.* The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature* **390**, 249–56 (1997).
16. Buescher, J. M. *et al.* Global network reorganization during dynamic adaptations of *Bacillus subtilis* metabolism. *Science (80-)*. **335**, 1099–1103 (2012).
17. Nicolas, P. *et al.* Condition-dependent transcriptome reveals high-level regulatory

- architecture in *Bacillus subtilis*. *Science* **335**, 1103–6 (2012).
18. Barbe, V. *et al.* From a consortium sequence to a unified sequence: *The Bacillus subtilis* 168 reference genome a decade later. *Microbiology* **155**, 1758–1775 (2009).
 19. Becher, D., Büttner, K., Moche, M., Heßling, B. & Hecker, M. From the genome sequence to the protein inventory of *Bacillus subtilis*. *Proteomics* **11**, 2971–2980 (2011).
 20. van Dijl, J. M. & Hecker, M. *Bacillus subtilis*: from soil bacterium to super-secreting cell factory. *Microb. Cell Fact.* **12**, 3 (2013).
 21. Earl, A. M., Losick, R. & Kolter, R. Ecology and genomics of *Bacillus subtilis*. *Trends Microbiol.* **16**, 269–75 (2008).
 22. Porwal, S., Lal, S., Cheema, S. & Kalia, V. C. Phylogeny in aid of the present and novel microbial lineages: Diversity in *Bacillus*. *PLoS One* **4**, (2009).
 23. Nakano, M. M., Dailly, Y. P., Zuber, P. & Clark, D. P. Characterization of anaerobic fermentative growth of *Bacillus subtilis*: identification of fermentation end products and genes required for growth. *J. Bacteriol.* **179**, 6749–55 (1997).
 24. Schallmey, M., Singh, A. & Ward, O. P. Developments in the use of *Bacillus* species for industrial production. *Can. J. Microbiol.* **50**, 1–17 (2004).
 25. Tasaki, S., Nakayama, M. & Shoji, W. Morphologies of *Bacillus subtilis* communities responding to environmental variation. *Dev. Growth Differ.* **59**, 369–378 (2017).
 26. Caulier, S. *et al.* Overview of the Antimicrobial Compounds Produced by Members of the *Bacillus subtilis* Group. *Front. Microbiol.* **10**, 1–19 (2019).
 27. Arnaouteli, S., Bamford, N. C., Stanley-Wall, N. R. & Kovács, Á. T. *Bacillus subtilis* biofilm formation and social interactions. *Nat. Rev. Microbiol.* **19**, 600–614 (2021).
 28. Riley, E. P., Schwarz, C., Derman, A. I. & Lopez-Garrido, J. Milestones in *Bacillus subtilis* sporulation research. *Microb. Cell* **8**, 1–16 (2021).
 29. Mckenney, P. T., Driks, A. & Eichenberger, P. The *Bacillus subtilis* endospore: Assembly and functions of the multilayered coat. *Nat. Rev. Microbiol.* **11**, 33–44 (2013).
 30. Ryter A. [Morphologic study of the sporulation of *Bacillus subtilis*]. *Ann. Inst. Pasteur (Paris)*. **108**, 40–60 (1965).
 31. Tyndall, J. On heat as a germicide when discontinuously applied. *Proc. R. Soc. London* **25**, 569–570 (1877).
 32. Piggot, P. J. & Coote, J. G. Genetic aspects of bacterial endospore formation. *Bacteriol. Rev.* **40**, 908–962 (1976).
 33. Errington, J. Regulation of endospore formation in *Bacillus subtilis*. *Nat. Rev. Microbiol.* **1**, 117–26 (2003).
 34. Fujita, M. & Losick, R. Evidence that entry into sporulation in *Bacillus subtilis* is governed by a gradual increase in the level and activity of the master regulator Spo0A. *Genes Dev.* **19**, 2236–44 (2005).

35. Eijlander, R. T., de Jong, A., Krawczyk, A. O., Holsappel, S. & Kuipers, O. P. SporeWeb: an interactive journey through the complete sporulation cycle of *Bacillus subtilis*. *Nucleic Acids Res.* **42**, D685-91 (2014).
36. Haldenwang, W. G. The sigma factors of *Bacillus subtilis*. *Microbiol. Rev.* **59**, 1–30 (1995).
37. Lee, N. K., Kim, W. S. & Paik, H. D. *Bacillus* strains as human probiotics: characterization, safety, microbiome, and probiotic carrier. *Food Sci. Biotechnol.* **28**, 1297–1305 (2019).
38. Guoyan, Z. *et al.* *Bacillus subtilis* spore surface display technology: A review of its development and applications. *J. Microbiol. Biotechnol.* **29**, 179–190 (2019).
39. Wu, H.-J. & Wu, E. The role of gut microbiota in immune homeostasis and autoimmunity. *Gut Microbes* **3**, 4–14 (2012).
40. Isticato, R. *et al.* Surface display of recombinant proteins on *Bacillus subtilis* spores. *J. Bacteriol.* **183**, 6294–301 (2001).
41. Duc, L. H., Hong, H. A., Fairweather, N., Ricca, E. & Cutting, S. M. Bacterial spores as vaccine vehicles. *Infect. Immun.* **71**, 2810–8 (2003).
42. Pაცეც, J. D. *et al.* Stable episomal expression system under control of a stress inducible promoter enhances the immunogenicity of *Bacillus subtilis* as a vector for antigen delivery. *Vaccine* **24**, 2935–43 (2006).
43. Oh, Y., Kim, J. A., Kim, C.-H., Choi, S.-K. & Pan, J.-G. *Bacillus subtilis* spore vaccines displaying protective antigen induce functional antibodies and protective potency. *BMC Vet. Res.* **16**, 259 (2020).
44. Sung, J. C. *et al.* Expression of SARS-CoV-2 spike protein receptor binding domain on recombinant *B. subtilis* on spore surface: A potential COVID-19 oral vaccine candidate. *Vaccines* **10**, 1–18 (2021).
45. Dubnau, D. DNA uptake in bacteria. *Annu. Rev. Microbiol.* **53**, 217–44 (1999).
46. Dubnau, D. & Blokesch, M. Mechanisms of DNA uptake by naturally competent bacteria. *Annu. Rev. Genet.* **53**, 217–237 (2019).
47. Chen, I. & Dubnau, D. DNA uptake during bacterial transformation. *Nat. Rev. Microbiol.* **2**, 241–249 (2004).
48. Provedji, R. & Dubnau, D. ComEA is a DNA receptor for transformation of competent *Bacillus subtilis*. *Mol. Microbiol.* **31**, 271–80 (1999).
49. Draskovic, I. & Dubnau, D. Biogenesis of a putative channel protein, ComEC, required for DNA uptake: membrane topology, oligomerization and formation of disulphide bonds. *Mol. Microbiol.* **55**, 881–96 (2005).
50. Maier, B. Competence and Transformation in *Bacillus subtilis*. *Curr. Issues Mol. Biol.* **37**, 57–76 (2020).
51. Hamoen, L. W., Venema, G. & Kuipers, O. P. Controlling competence in *Bacillus subtilis* : Shared use of regulators. *Microbiology* **149**, 9–17 (2003).

52. Ramos-Silva, P., Serrano, M. & Henriques, A. O. From root to tips: Sporulation evolution and specialization in *Bacillus subtilis* and the intestinal pathogen *Clostridioides difficile*. *Mol. Biol. Evol.* **36**, 2714–2736 (2019).
53. Romero, D., Aguilar, C., Losick, R. & Kolter, R. Amyloid fibers provide structural integrity to *Bacillus subtilis* biofilms. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 2230–4 (2010).
54. Terra, R., Stanley-Wall, N. R., Cao, G. & Lazazzera, B. A. Identification of *Bacillus subtilis* SipW as a bifunctional signal peptidase that controls surface-adhered biofilm formation. *J. Bacteriol.* **194**, 2781–90 (2012).
55. Bromley, K. M. & MacPhee, C. E. BslA-stabilized emulsion droplets with designed microstructure. *Interface Focus* **7**, (2017).
56. Stanley-Wall, N. R. & MacPhee, C. E. Connecting the dots between bacterial biofilms and ice cream. *Phys. Biol.* **12**, (2015).
57. Kang, S.-Y. *et al.* Engineering *Bacillus subtilis* for the formation of a durable living biocomposite material. *Nat. Commun.* **12**, 7133 (2021).
58. Oslizlo, A. *et al.* Exploring ComQXPA quorum-sensing diversity and biocontrol potential of *Bacillus* spp. isolates from tomato rhizoplane. *Microb. Biotechnol.* **8**, 527–40 (2015).
59. Vejan, P., Abdullah, R., Khadiran, T., Ismail, S. & Nasrulhaq Boyce, A. Role of plant growth promoting rhizobacteria in agricultural sustainability-a review. *Molecules* **21**, (2016).
60. Blake, C., Christensen, M. N. & Kovács, Á. T. Molecular aspects of plant growth promotion and protection by *Bacillus subtilis*. *Mol. Plant. Microbe. Interact.* **34**, 15–25 (2021).
61. Borriss, R. *et al.* *Bacillus subtilis*, the model Gram-positive bacterium: 20 years of annotation refinement. *Microb. Biotechnol.* **11**, 3–17 (2018).
62. Soupene, E. *et al.* Physiological studies of *Escherichia coli* strain MG1655: growth defects and apparent cross-regulation of gene expression. *J. Bacteriol.* **185**, 5611–26 (2003).
63. Moszer, I., Kunst, F. & Danchin, A. The European *Bacillus subtilis* genome sequencing project: current status and accessibility of the data from a new World Wide Web site. *Microbiology* **142**, 2987–91 (1996).
64. Michna, R. H., Zhu, B., Mäder, U. & Stülke, J. SubtiWiki 2.0 - An integrated database for the model organism *Bacillus subtilis*. *Nucleic Acids Res.* **44**, 654–662 (2016).
65. Zhu, B. & Stülke, J. SubtiWiki in 2018: From genes and proteins to functional network annotation of the model organism *Bacillus subtilis*. *Nucleic Acids Res.* **46**, 743–748 (2018).
66. Antelmann, H. *et al.* A proteomic view on genome-based signal peptide predictions. *Genome Res.* **11**, 1484–502 (2001).
67. Fischer, E. & Sauer, U. Large-scale in vivo flux analysis shows rigidity and suboptimal performance of *Bacillus subtilis* metabolism. *Nat. Genet.* **37**, 636–40 (2005).

68. Oh, Y.-K., Palsson, B. O., Park, S. M., Schilling, C. H. & Mahadevan, R. Genome-scale reconstruction of metabolic network in *Bacillus subtilis* based on high-throughput phenotyping and gene essentiality data. *J. Biol. Chem.* **282**, 28791–28799 (2007).
69. Caspi, R. *et al.* The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res.* **44**, D471-80 (2016).
70. Harwood, C. R., Pohl, S., Smith, W. & Wipat, A. *Bacillus subtilis*. Model Gram-Positive Synthetic Biology Chassis. *Methods Microbiol.* **40**, 87–117 (2013).
71. Nijland, R., Burgess, J. G., Errington, J. & Veening, J. W. Transformation of environmental *Bacillus subtilis* isolates by transiently inducing genetic competence. *PLoS One* **5**, 1–7 (2010).
72. Vojcic, L., Despotovic, D., Martinez, R., Maurer, K. H. & Schwaneberg, U. An efficient transformation method for *Bacillus subtilis* DB104. *Appl. Microbiol. Biotechnol.* **94**, 487–493 (2012).
73. Bettinger, G. E. & Young, F. E. Transformation of *Bacillus subtilis*: transforming ability of deoxyribonucleic acid lysates of L forms or protoplasts. *J. Bacteriol.* **122**, 987–993 (1975).
74. Young, F. E., Haywood, P. & Pollock, M. Isolation of L-forms of *Bacillus subtilis* which grow in liquid medium. *J. Bacteriol.* **102**, 867–70 (1970).
75. Errington, J. Cell wall-deficient, L-form bacteria in the 21st century: A personal perspective. *Biochem. Soc. Trans.* **45**, 287–295 (2017).
76. Wyrick, P. B. & Rogers, H. J. Isolation and characterization of cell wall defective variants of *Bacillus subtilis* and *Bacillus licheniformis*. *J. Bacteriol.* **116**, 456–465 (1973).
77. Schaeffer, P. & Hotchkiss, R. D. Fusion of Bacterial Protoplasts. *Methods Cell Biol.* **20**, 149–158 (1978).
78. Chang, S. & Cohen, S. N. High frequency transformation of *Bacillus subtilis* protoplasts by plasmid DNA. *MGG Mol. Gen. Genet.* **168**, 111–115 (1979).
79. Akamatsu, T. Transformation of *Bacillus* Protoplasts by Plasmid pTP4 DNA experiments was improved by the use. **46**, 1617–1621 (1982).
80. Lévi-Meyrueis, C., Fodor, K. & Schaeffer, P. Polyethyleneglycol-induced transformation of *Bacillus subtilis* protoplasts by bacterial chromosomal DNA. *MGG Mol. Gen. Genet.* **179**, 589–594 (1980).
81. Vasileva, D. P. *et al.* Protoplast fusion in *Bacillus* species produces frequent, unbiased, genome-wide homologous recombination. *Nucleic Acids Res.* 1–24 (2022).
82. Yachi, K. *et al.* Biopharmaceutical evaluation of the liposomes prepared by rehydration of freeze-dried empty liposomes (FDELs) with an aqueous solution of a drug. *Biopharm. Drug Dispos.* **17**, 589–605 (1996).
83. Ohta, K. & Ichihashi, N. Liposome fragment-mediated introduction of multiple plasmids into *Bacillus subtilis*. *Biochem. Biophys. Reports* **18**, 100646 (2019).
84. Bron, S., Meima, R., Van Dijl, J. M., Wipat, A. & Harwood, C. R. Molecular Biology and

- Genetics of *Bacillus* species. *Man. Ind. Microbiol. Biotechnol.* 4–49 (1999).
85. Xue, G. P., Johnson, J. S. & Dalrymple, B. P. High osmolarity improves the electrotransformation efficiency of the gram-positive bacteria *Bacillus subtilis* and *Bacillus licheniformis*. *J. Microbiol. Methods* **34**, 183–191 (1999).
 86. Romero, D., Pérez-García, A., Veening, J. W., de Vicente, A. & Kuipers, O. P. Transformation of undomesticated strains of *Bacillus subtilis* by protoplast electroporation. *J. Microbiol. Methods* **66**, 556–559 (2006).
 87. Voelker, L. L. & Dybvig, K. Transposon mutagenesis. *Methods Mol. Biol.* **104**, 235–8 (1998).
 88. Franke, A. E. & Clewell, D. B. 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 (1981).
 89. Courvalin, P. & Carrier, C. Tn1545: a conjugative shuttle transposon. *Mol. Gen. Genet.* **206**, 259–64 (1987).
 90. Youngman, P. J., Perkins, J. B. & Losick, R. Genetic transposition and insertional mutagenesis in *Bacillus subtilis* with *Streptococcus faecalis* transposon Tn917. *Proc. Natl. Acad. Sci. U. S. A.* **80**, 2305–9 (1983).
 91. Youngman, P. J. Plasmid vectors for recovering and exploiting Tn917 transpositions in *Bacillus subtilis* and other gram positives. *Cold Spring Harb.* 79–103 (1985).
 92. Chang, L. K. *et al.* Construction of Tn917ac1, a transposon useful for mutagenesis and cloning of *Bacillus subtilis* genes. *Gene* **150**, 129–34 (1994).
 93. Kleckner N. Transposon Tn10. In: Berg D E, Howe M M, editors. Mobile DNA. *Am. Soc. Microbiol.* 227–268 (1989).
 94. Petit, M. A., Bruand, C., Janni re, L. & Ehrlich, S. D. Tn10-derived transposons active in *Bacillus subtilis*. *J. Bacteriol.* **172**, 6736–40 (1990).
 95. Steinmetz, M. & Richter, R. Easy cloning of mini-Tn10 insertions from the *Bacillus subtilis* chromosome. *J. Bacteriol.* **176**, 1761–3 (1994).
 96. Youngman, P. J., J. B. Perkins, and K. S. Use of Tn917-mediated transcriptional gene fusions to lacZ and cat-86 for the identification and study of spo genes in *Bacillus subtilis* . 47–54 (1985).
 97. Halling, S. M. & Kleckner, N. A symmetrical six-base-pair target site sequence determines Tn10 insertion specificity. *Cell* **28**, 155–63 (1982).
 98. Le Breton, Y., Mohapatra, N. P. & Haldenwang, W. G. In vivo random mutagenesis of *Bacillus subtilis* by use of TnYLB-1, a mariner-based transposon. *Appl. Environ. Microbiol.* **72**, 327–33 (2006).
 99. Pozsgai, E. R., Blair, K. M. & Kearns, D. B. Modified mariner transposons for random inducible-expression insertions and transcriptional reporter fusion insertions in *Bacillus subtilis*. *Appl. Environ. Microbiol.* **78**, 778–785 (2012).
 100. Dong, H. & Zhang, D. Current development in genetic engineering strategies of *Bacillus*

- species. *Microb. Cell Fact.* **13**, 1–11 (2014).
101. Fernández, S., Ayora, S. & Alonso, J. C. *Bacillus subtilis* homologous recombination: Genes and products. *Res. Microbiol.* **151**, 481–486 (2000).
 102. Lopilato, J., Bortner, S. & Beckwith, J. Mutations in a new chromosomal gene of *Escherichia coli* K-12, *pcnB*, reduce plasmid copy number of pBR322 and its derivatives. *Mol. Gen. Genet.* **205**, 285–90 (1986).
 103. Kobayashi, K. *et al.* Essential *Bacillus subtilis* genes. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 4678–4683 (2003).
 104. Vagner, V., Dervyn, E. & Ehrlich, S. D. A vector for systematic gene inactivation in *Bacillus subtilis*. *Microbiology* **144**, 3097–3104 (1998).
 105. Sun, Z. *et al.* A high-efficiency recombineering system with PCR-based ssDNA in *Bacillus subtilis* mediated by the native phage recombinase GP35. *Appl. Microbiol. Biotechnol.* **99**, 5151–62 (2015).
 106. Itaya, M. Effective cloning of unmarked DNA fragments in the *Bacillus subtilis* 168 Genome. *Biosci. Biotechnol. Biochem.* **63**, 602–4 (1999).
 107. Fabret, C., Ehrlich, S. D. & Noirot, P. A new mutation delivery system for genome-scale approaches in *Bacillus subtilis*. *Mol. Microbiol.* **46**, 25–36 (2002).
 108. Yan, X., Yu, H. J., Hong, Q. & Li, S. P. Cre/lox system and PCR-based genome engineering in *Bacillus subtilis*. *Appl. Environ. Microbiol.* **74**, 5556–5562 (2008).
 109. Westbrook, A. W., Moo-Young, M. & Chou, C. P. Development of a CRISPR-Cas9 tool kit for comprehensive engineering of *Bacillus subtilis*. *Appl. Environ. Microbiol.* **82**, 4876–4895 (2016).
 110. Zocca, V. F. B. *et al.* The CRISPR toolbox for the gram-positive model bacterium *Bacillus subtilis*. *Crit. Rev. Biotechnol.* **42**, 813–826 (2022).
 111. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–21 (2012).
 112. Koehler, T. M. & Thorne, C. B. *Bacillus subtilis* (*natto*) plasmid pLS20 mediates interspecies plasmid transfer. *J. Bacteriol.* **169**, 5271–8 (1987).
 113. Bron, S. & Luxen, E. Segregational instability of pUB110-derived recombinant plasmids in *Bacillus subtilis*. *Plasmid* **14**, 235–44 (1985).
 114. Bron, S., Meijer, W., Holsappel, S. & Haima, P. Plasmid instability and molecular cloning in *Bacillus subtilis*. *Res. Microbiol.* **142**, 875–83 (1991).
 115. Gruss, A. & Ehrlich, S. D. The family of highly interrelated single-stranded deoxyribonucleic acid plasmids. *Microbiol. Rev.* **53**, 231–41 (1989).
 116. JanniPre, L., A. Gruss, and S. D. E. Plasmids. In *Bacillus subtilis* and other Gram-positive bacteria: Biochemistry, physiology and molecular genetics. *Am. Soc. Microbiol.* 625–644 (1993).
 117. Bruand, C., Ehrlich, S. D. & Jannièrè, L. Unidirectional theta replication of the

- structurally stable *Enterococcus faecalis* plasmid pAM beta 1. *EMBO J.* **10**, 2171–7 (1991).
118. Simon, D. & Chopin, A. Construction of a vector plasmid family and its use for molecular cloning in *Streptococcus lactis*. *Biochimie* **70**, 559–66 (1988).
 119. Kiewiet, R., Kok, J., Seegers, J. F., Venema, G. & Bron, S. The mode of replication is a major factor in segregational plasmid instability in *Lactococcus lactis*. *Appl. Environ. Microbiol.* **59**, 358–64 (1993).
 120. Wang, P. Z. & Doi, R. H. Overlapping promoters transcribed by *Bacillus subtilis* sigma 55 and sigma 37 RNA polymerase holoenzymes during growth and stationary phases. *J. Biol. Chem.* **259**, 8619–25 (1984).
 121. Song, B. H. & Neuhard, J. Chromosomal location, cloning and nucleotide sequence of the *Bacillus subtilis* *cdd* gene encoding cytidine/deoxycytidine deaminase. *Mol. Gen. Genet.* **216**, 462–8 (1989).
 122. Yansura, D. G. & Henner, D. J. Use of the *Escherichia coli* lac repressor and operator to control gene expression in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. U. S. A.* **81**, 439–43 (1984).
 123. Guiziou, S. *et al.* A part toolbox to tune genetic expression in *Bacillus subtilis*. *Nucleic Acids Res.* **44**, 7495–7508 (2016).
 124. Radeck, J. *et al.* The *Bacillus* BioBrick Box: generation and evaluation of essential genetic building blocks for standardized work with *Bacillus subtilis*. *J. Biol. Eng.* **7**, 29 (2013).
 125. Popp, P. F., Dotzler, M., Radeck, J., Bartels, J. & Mascher, T. The *Bacillus* BioBrick Box 2.0: expanding the genetic toolbox for the standardized work with *Bacillus subtilis*. *Sci. Rep.* **7**, 15058 (2017).
 126. Zeigler, D. R. *Bacillus* Genetic Stock Center. *Biol. Resour. Model Org.* 19 (2019).
 127. Ling Lin Fu *et al.* Protein secretion pathways in *Bacillus subtilis*: Implication for optimization of heterologous protein secretion. *Biotechnol. Adv.* **25**, 1–12 (2007).
 128. Ventura, S. & Villaverde, A. Protein quality in bacterial inclusion bodies. *Trends Biotechnol.* **24**, 179–85 (2006).
 129. Natale, P., Brüser, T. & Driessen, A. J. M. Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane--distinct translocases and mechanisms. *Biochim. Biophys. Acta* **1778**, 1735–56 (2008).
 130. Zhang, K., Su, L. & Wu, J. Recent advances in recombinant protein production by *Bacillus subtilis*. *Annu. Rev. Food Sci. Technol.* **11**, 295–318 (2020).
 131. Su, Y., Liu, C., Fang, H. & Zhang, D. *Bacillus subtilis*: a universal cell factory for industry, agriculture, biomaterials and medicine. *Microb. Cell Fact.* **19**, 173 (2020).
 132. Diao, L. *et al.* Functional implementation of the posttranslational SecB-SecA protein-targeting pathway in *Bacillus subtilis*. *Appl. Environ. Microbiol.* **78**, 651–9 (2012).
 133. Williams, R. C. *et al.* Production of *Bacillus anthracis* protective antigen is dependent on the extracellular chaperone, PrsA. *J. Biol. Chem.* **278**, 18056–62 (2003).

134. Cui, W. *et al.* Exploitation of *Bacillus subtilis* as a robust workhorse for production of heterologous proteins and beyond. *World J. Microbiol. Biotechnol.* **34**, 145 (2018).
135. Acevedo-Rocha, C. G., Gronenberg, L. S., Mack, M., Commichau, F. M. & Genee, H. J. Microbial cell factories for the sustainable manufacturing of B vitamins. *Curr. Opin. Biotechnol.* **56**, 18–29 (2019).
136. Yan, S. & Wu, G. Bottleneck in secretion of α -amylase in *Bacillus subtilis*. *Microb. Cell Fact.* **16**, 124 (2017).
137. Watzlawick, H. & Altenbuchner, J. Multiple integration of the gene *ganA* into the *Bacillus subtilis* chromosome for enhanced β -galactosidase production using the CRISPR/Cas9 system. *AMB Express* **9**, 158 (2019).
138. Lan Thanh Bien, T., Tsuji, S., Tanaka, K., Takenaka, S. & Yoshida, K. Secretion of heterologous thermostable cellulases in *Bacillus subtilis*. *J. Gen. Appl. Microbiol.* **60**, 175–82 (2014).
139. Patel, A. R., Mokashe, N. U., Chaudhari, D. S., Jadhav, A. G. & Patil, U. K. Production optimisation and characterisation of extracellular protease secreted by newly isolated *Bacillus subtilis* AU-2 strain obtained from *Tribolium castaneum* gut. *Biocatal. Agric. Biotechnol.* **19**, 101122 (2019).
140. Revuelta, J. L. *et al.* Bioproduction of riboflavin: a bright yellow history. *J. Ind. Microbiol. Biotechnol.* **44**, 659–665 (2017).
141. Suwannasom, N., Kao, I., Pruß, A., Georgieva, R. & Bäumler, H. Riboflavin: The Health Benefits of a Forgotten Natural Vitamin. *Int. J. Mol. Sci.* **21**, (2020).
142. Hoffmann, M. *et al.* Towards the Anaerobic Production of Surfactin Using *Bacillus subtilis*. *Front. Bioeng. Biotechnol.* **8**, 554903 (2020).
143. Özcengiz, G. & Ögülür, İ. Biochemistry, genetics and regulation of bacilysin biosynthesis and its significance more than an antibiotic. *N. Biotechnol.* **32**, 612–9 (2015).
144. Pierce, J. A., Robertson, C. R. & Leighton, T. J. Physiological and genetic strategies for enhanced subtilisin production by *Bacillus subtilis*. *Biotechnol. Prog.* **8**, 211–8.
145. Stein, T. *Bacillus subtilis* antibiotics: structures, syntheses and specific functions. *Mol. Microbiol.* **56**, 845–57 (2005).
146. Sumi, C. D., Yang, B. W., Yeo, I.-C. & Hahm, Y. T. Antimicrobial peptides of the genus *Bacillus*: a new era for antibiotics. *Can. J. Microbiol.* **61**, 93–103 (2015).
147. Haseloff, J. & Ajioka, J. Synthetic biology: history, challenges and prospects. *J. R. Soc. Interface* **6**, 389–91 (2009).
148. de Lorenzo, V. & Danchin, A. Synthetic biology: discovering new worlds and new words. *EMBO Rep.* **9**, 822–7 (2008).
149. Andrianantoandro, E., Basu, S., Karig, D. K. & Weiss, R. Synthetic biology: new engineering rules for an emerging discipline. *Mol. Syst. Biol.* **2**, 2006.0028 (2006).
150. Meng, F. & Ellis, T. The second decade of synthetic biology: 2010–2020. *Nat. Commun.* **11**, 1–4 (2020).

151. de Lorenzo, V., Krasnogor, N. & Schmidt, M. For the sake of the Bioeconomy: define what a Synthetic Biology Chassis is! *N. Biotechnol.* **60**, 44–51 (2021).
152. What's in a name? *Nat. Biotechnol.* **27**, 1071–1073 (2009).
153. Endy, D. Foundations for engineering biology. *Nature* **438**, 449–453 (2005).
154. Ro, D.-K. *et al.* Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**, 940–3 (2006).
155. Keasling, J. D. Synthetic biology and the development of tools for metabolic engineering. *Metab. Eng.* **14**, 189–195 (2012).
156. Paddon, C. J. & Keasling, J. D. Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. *Nat. Rev. Microbiol.* **12**, 355–67 (2014).
157. Lee, S. K., Chou, H., Ham, T. S., Lee, T. S. & Keasling, J. D. Metabolic engineering of microorganisms for biofuels production: from bugs to synthetic biology to fuels. *Curr. Opin. Biotechnol.* **19**, 556–63 (2008).
158. Gupta, P. & Phulara, S. C. Metabolic engineering for isoprenoid-based biofuel production. *J. Appl. Microbiol.* **119**, 605–19 (2015).
159. Levskaya, A. *et al.* Synthetic biology: engineering *Escherichia coli* to see light. *Nature* **438**, 441–2 (2005).
160. Green, A. A., Silver, P. A., Collins, J. J. & Yin, P. Toehold switches: de-novo-designed regulators of gene expression. *Cell* **159**, 925–39 (2014).
161. Herdewijn, P. & Marlière, P. Toward safe genetically modified organisms through the chemical diversification of nucleic acids. *Chem. Biodivers.* **6**, 791–808 (2009).
162. Wang, K., Neumann, H., Peak-Chew, S. Y. & Chin, J. W. Evolved orthogonal ribosomes enhance the efficiency of synthetic genetic code expansion. *Nat. Biotechnol.* **25**, 770–7 (2007).
163. Neumann, H., Slusarczyk, A. L. & Chin, J. W. De novo generation of mutually orthogonal aminoacyl-tRNA synthetase/tRNA pairs. *J. Am. Chem. Soc.* **132**, 2142–4 (2010).
164. Gibson, D. G. *et al.* Creation of a bacterial cell controlled by a chemically synthesized genome. *Science* **329**, 52–6 (2010).
165. de Lorenzo, V. Beware of metaphors chassis and orthogonality in synthetic biology. *Bioeng. Bugs* **2**, 3–7 (2011).
166. Cardinale, S. & Arkin, A. P. Contextualizing context for synthetic biology - identifying causes of failure of synthetic biological systems. *Biotechnol. J.* **7**, 856–866 (2012).
167. Morowitz, H. J. The completeness of molecular biology. *Isr. J. Med. Sci.* **20**, 750–3 (1984).
168. Cordova, C. M. M., Hoeltgebaum, D. L., Machado, L. D. P. N. & Santos, L. Dos. Molecular biology of mycoplasmas: from the minimum cell concept to the artificial cell. *An. Acad. Bras. Cienc.* **88**, 599–607 (2016).

169. Glass, J. I., Merryman, C., Wise, K. S., Hutchison, C. A. & Smith, H. O. Minimal Cells-Real and Imagined. *Cold Spring Harb. Perspect. Biol.* **9**, 1–12 (2017).
170. Thornburg, Z. R. *et al.* Fundamental behaviors emerge from simulations of a living minimal cell. *Cell* **185**, 345–360 (2022).
171. Lluch-Senar, M. *et al.* Defining a minimal cell: essentiality of small ORFs and ncRNAs in a genome-reduced bacterium. *Mol. Syst. Biol.* **11**, 780 (2015).
172. Hutchison, C. A. *et al.* Design and synthesis of a minimal bacterial genome. *Science* **351**, (2016).
173. Fleischmann, R. D. *et al.* Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **269**, 496–512 (1995).
174. Fraser, C. M. *et al.* The minimal gene complement of *Mycoplasma genitalium*. *Science* **270**, 397–403 (1995).
175. Mushegian, A. R. & Koonin, E. V. A minimal gene set for cellular life derived by comparison of complete bacterial genomes. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 10268–73 (1996).
176. Hutchison, C. A. *et al.* Global transposon mutagenesis and a minimal *Mycoplasma* genome. *Science* **286**, 2165–9 (1999).
177. Glass, J. I. *et al.* Essential genes of a minimal bacterium. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 425–30 (2006).
178. Juhas, M., Eberl, L. & Glass, J. I. Essence of life: essential genes of minimal genomes. *Trends Cell Biol.* **21**, 562–8 (2011).
179. Gil, R., Silva, F. J., Peretó, J. & Moya, A. Determination of the core of a minimal bacterial gene set. *Microbiol. Mol. Biol. Rev.* **68**, 518–37, table of contents (2004).
180. Fehér, T., Papp, B., Pal, C. & Pósfai, G. Systematic genome reductions: theoretical and experimental approaches. *Chem. Rev.* **107**, 3498–513 (2007).
181. Reznikoff, W. S. & Winterberg, K. M. Transposon-based strategies for the identification of essential bacterial genes. *Methods Mol. Biol.* **416**, 13–26 (2008).
182. Hirokawa, Y. *et al.* Genetic manipulations restored the growth fitness of reduced-genome *Escherichia coli*. *J. Biosci. Bioeng.* **116**, 52–8 (2013).
183. Reuß, D. R. *et al.* Large-scale reduction of the *Bacillus subtilis* genome: consequences for the transcriptional network, resource allocation, and metabolism. *Genome Res.* **27**, 289–299 (2017).
184. Etienne Dervyn, Anne-Gaëlle Planson, Kosei Tanaka, Viktor Chubukov, Cyprien Guérin, Sandra Derozier, François Lecoïnte, Uwe Sauer, Ken-Ichi Yoshida, Pierre Nicolas, P. N. * and M. J. Greedy reduction of *Bacillus subtilis* genome yields emergent phenotypes of high resistance to DNA damaging agent and low evolvability. *En cours soumission*.
185. Xu, X., Liu, Y., Du, G., Ledesma-Amaro, R. & Liu, L. Microbial Chassis Development for Natural Product Biosynthesis. *Trends Biotechnol.* **38**, 779–796 (2020).

186. Vickers, C. E., Blank, L. M. & Krömer, J. O. Grand challenge commentary: Chassis cells for industrial biochemical production. *Nat. Chem. Biol.* **6**, 875–7 (2010).
187. Adams, B. L. The next generation of synthetic biology chassis: moving synthetic biology from the laboratory to the field. *ACS Synth. Biol.* **5**, 1328–1330 (2016).
188. Calero, P. & Nikel, P. I. Chasing bacterial chassis for metabolic engineering: a perspective review from classical to non-traditional microorganisms. *Microb. Biotechnol.* **12**, 98–124 (2019).
189. Komatsu, M., Uchiyama, T., Omura, S., Cane, D. E. & Ikeda, H. Genome-minimized *Streptomyces* host for the heterologous expression of secondary metabolism. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 2646–51 (2010).
190. Liu, Y., Liu, L., Li, J., Du, G. & Chen, J. Synthetic Biology Toolbox and Chassis Development in *Bacillus subtilis*. *Trends Biotechnol.* **37**, 548–562 (2019).
191. Reuß, D. R., Commichau, F. M., Gundlach, J., Zhu, B. & Stülke, J. The Blueprint of a Minimal Cell: *MiniBacillus*. *Microbiol. Mol. Biol. Rev.* **80**, 955–987 (2016).
192. Tanaka, K. *et al.* 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 (2013).
193. Manabe, K. *et al.* Combined effect of improved cell yield and increased specific productivity enhances recombinant enzyme production in genome-reduced *Bacillus subtilis* strain MGB874. *Appl. Environ. Microbiol.* **77**, 8370–81 (2011).
194. Li, Y. *et al.* Characterization of genome-reduced *Bacillus subtilis* strains and their application for the production of guanosine and thymidine. *Microb. Cell Fact.* **15**, 94 (2016).
195. Suárez, R. A., Stülke, J. & Van Dijl, J. M. Less is More: Toward a Genome-Reduced *Bacillus* Cell Factory for ‘difficult Proteins’. *ACS Synth. Biol.* **8**, 99–108 (2019).
196. Antelo-Varela, M. *et al.* Membrane modulation of super-secreting ‘*midiBacillus*’ expressing the major *Staphylococcus aureus* antigen - A mass-spectrometry-based absolute quantification approach. *Front. Bioeng. Biotechnol.* **8**, 143 (2020).
197. Broto, A., Gaspari, E., Miravet-Verde, S., dos Santos, V. A. P. M. & Isalan, M. A genetic toolkit and gene switches to limit *Mycoplasma* growth for biosafety applications. *Nat. Commun.* **13**, (2022).
198. Lajoie, M. J. *et al.* Genomically recoded organisms expand biological functions. *Science* **342**, 357–60 (2013).
199. Fredens, J. *et al.* Total synthesis of *Escherichia coli* with a recoded genome. *Nature* **569**, 514–518 (2019).
200. Na, D. DNA steganography: hiding undetectable secret messages within the single nucleotide polymorphisms of a genome and detecting mutation-induced errors. *Microb. Cell Fact.* **19**, 128 (2020).
201. Nielsen, J. Yeast Systems Biology: Model Organism and Cell Factory. *Biotechnol. J.* **14**,

- e1800421 (2019).
202. Sherman, F. Getting started with yeast. *Methods Enzymol.* **350**, 3–41 (2002).
 203. Burke, D. T., Carle, G. F. & Olson, M. V. Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science* **236**, 806–12 (1987).
 204. Larionov, V. *et al.* Specific cloning of human DNA as yeast artificial chromosomes by transformation-associated recombination. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 491–496 (1996).
 205. Kuspa, A., Vollrath, D., Cheng, Y. & Kaiser, D. Physical mapping of the *Myxococcus xanthus* genome by random cloning in yeast artificial chromosomes. *Proc. Natl. Acad. Sci. U. S. A.* **86**, 8917–8921 (1989).
 206. Ketner, G., Spencer, F., Tugendreich, S., Connelly, C. & Hieter, P. Efficient manipulation of the human adenovirus genome as an infectious yeast artificial chromosome clone. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 6186–90 (1994).
 207. Ajioka, J. W. *et al.* Drosophila genome project: One-hit coverage in yeast artificial chromosomes. *Chromosoma* **100**, 495–509 (1991).
 208. Azevedo, V. *et al.* An ordered collection of *Bacillus subtilis* DNA segments cloned in yeast artificial chromosomes. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 6047–6051 (1993).
 209. García-Ramírez, J. J. *et al.* Dominance of Virus over Host Factors in Cross-Species Activation of Human Cytomegalovirus Early Gene Expression. *J. Virol.* **75**, 26–35 (2001).
 210. Larionov, V., Kouprina, N., Solomon, G., Barrett, J. C. & Resnick, M. A. Direct isolation of human *brca2* gene by transformation-associated recombination in yeast. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 7384–7387 (1997).
 211. Kouprina, N. *et al.* Functional copies of a human gene can be directly isolated by transformation-associated recombination cloning with a small 3' end target sequence. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 4469–74 (1998).
 212. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
 213. Gibson, D. G. *et al.* Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science* **319**, 1215–20 (2008).
 214. Gibson, D. G. *et al.* 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 (2008).
 215. Benders, G. A. *et al.* Cloning whole bacterial genomes in yeast. *Nucleic Acids Res.* **38**, 2558–2569 (2010).
 216. Lartigue, C. *et al.* Creating bacterial strains from genomes that have been cloned and engineered in yeast. *Science* **325**, 1693–6 (2009).
 217. Lartigue, C. *et al.* Genome transplantation in bacteria: changing one species to another. *Science* **317**, 632–8 (2007).

218. Gibson, D. G. *et al.* Creation of a bacterial cell controlled by a chemically synthesized genome. *Science* **329**, 52–6 (2010).
219. Vashee, S., Arfi, Y. & Lartigue, C. Budding yeast as a factory to engineer partial and complete microbial genomes. *Curr. Opin. Syst. Biol.* **24**, 1–8 (2020).
220. Ruiz, E. *et al.* 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 (2019).
221. Karas, B. J., Tagwerker, C., Yonemoto, I. T., Hutchison, C. A. & Smith, H. O. Cloning the *Acholeplasma laidlawii* pg-8a genome in *Saccharomyces cerevisiae* as a yeast centromeric plasmid. *ACS Synth. Biol.* **1**, 22–28 (2012).
222. Baby, V. *et al.* Cloning and transplantation of the *Mesoplasma florum* genome. *ACS Synth. Biol.* **7**, 209–217 (2018).
223. Karas, B. J. *et al.* Direct transfer of whole genomes from bacteria to yeast. *Nat. Methods* **10**, 410–412 (2013).
224. Tagwerker, C. *et al.* Sequence analysis of a complete 1.66 Mb *Prochlorococcus marinus* MED4 genome cloned in yeast. *Nucleic Acids Res.* **40**, 10375–10383 (2012).
225. Zhou, J., Wu, R., Xue, X. & Qin, Z. CasHRA (Cas9-facilitated Homologous Recombination Assembly) method of constructing megabase-sized DNA. *Nucleic Acids Res.* **44**, (2016).
226. Venetz, J. E. *et al.* Chemical synthesis rewriting of a bacterial genome to achieve design flexibility and biological functionality. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 8070–8079 (2019).
227. Polo, S., Ketner, G., Levis, R. & Falgout, B. Infectious RNA transcripts from full-length dengue virus type 2 cDNA clones made in yeast. *J. Virol.* **71**, 5366–74 (1997).
228. Thi Nhu Thao, T. *et al.* Rapid reconstruction of SARS-CoV-2 using a synthetic genomics platform. *Nature* **582**, 561–565 (2020).
229. Vashee, S. *et al.* Cloning, assembly, and modification of the primary human cytomegalovirus isolate Toledo by yeast-based Transformation-Associated Recombination. *mSphere* **2**, 1–17 (2017).
230. Labroussaa, F. *et al.* Impact of donor-recipient phylogenetic distance on bacterial genome transplantation. *Nucleic Acids Res.* **44**, 8501–8511 (2016).
231. Kouprina, N. *et al.* Segments missing from the draft human genome sequence can be isolated by transformation-associated recombination cloning in yeast. *EMBO Rep.* **4**, 257–62 (2003).
232. Noskov, V. N. *et al.* A general cloning system to selectively isolate any eukaryotic or prokaryotic genomic region in yeast. *BMC Genomics* **4**, 1–11 (2003).
233. Lee, N. C. O., Larionov, V. & Kouprina, N. Highly efficient CRISPR/Cas9-mediated TAR cloning of genes and chromosomal loci from complex genomes in yeast. *Nucleic Acids Res.* **43**, e55 (2015).
234. Rideau, F. *et al.* Cloning, stability, and modification of *Mycoplasma hominis* genome in

- yeast. *ACS Synth. Biol.* **6**, 891–901 (2017).
235. Karas, B. J. *et al.* Assembly of eukaryotic algal chromosomes in yeast. *J. Biol. Eng.* **7**, 1–12 (2013).
 236. Karas, B. J. *et al.* Transferring whole genomes from bacteria to yeast spheroplasts using entire bacterial cells to reduce DNA shearing. *Nat. Protoc.* **9**, 743–750 (2014).
 237. Karas, B. J. *et al.* 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 (2019).
 238. Talenton, V. *et al.* Genome engineering of the fast-growing *Mycoplasma feriruminatoris* toward a live vaccine chassis. *ACS Synth. Biol.* (2022).
 239. Dicarlo, J. E. *et al.* Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res.* **41**, 4336–4343 (2013).
 240. Fraczek, M. G., Naseeb, S. & Delneri, D. History of genome editing in yeast. *Yeast* **35**, 361–368 (2018).
 241. Noskov, V. N., Segall-Shapiro, T. H. & Chuang, R.-Y. Tandem repeat coupled with endonuclease cleavage (TREC): a seamless modification tool for genome engineering in yeast. *Nucleic Acids Res.* **38**, 2570–6 (2010).
 242. Chandran, S. *et al.* TREC-IN: gene knock-in genetic tool for genomes cloned in yeast. *BMC Genomics* **15**, 1180 (2014).
 243. Lartigue, C. *et al.* The flavoprotein Mcap0476 (RlmFO) catalyzes m5U1939 modification in *Mycoplasma capricolum* 23S rRNA. *Nucleic Acids Res.* **42**, 8073–82 (2014).
 244. Schieck, E. *et al.* Galactofuranose in *Mycoplasma mycoides* is important for membrane integrity and conceals adhesins but does not contribute to serum resistance. *Mol. Microbiol.* **99**, 55–70 (2016).
 245. Jores, J. *et al.* Removal of a subset of non-essential genes fully attenuates a highly virulent *Mycoplasma* strain. *Front. Microbiol.* **10**, 664 (2019).
 246. Lartigue, C. *et al.* Attenuation of a pathogenic *Mycoplasma* strain by modification of the *obg* gene by using Synthetic Biology approaches. *mSphere* **4**, 1–17 (2019).
 247. Barrangou, R. *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709–12 (2007).
 248. Barrangou, R. CRISPR-Cas systems and RNA-guided interference. *Wiley Interdiscip. Rev. RNA* **4**, 267–278 (2013).
 249. Tsarmopoulou, I. *et al.* In-yeast engineering of a bacterial genome using CRISPR/Cas9. *ACS Synth. Biol.* **5**, 104–9 (2016).
 250. Talenton, V. *et al.* Genome engineering of the fast-growing *Mycoplasma feriruminatoris* toward a live vaccine chassis. *ACS Synth. Biol.* (2022) doi:10.1021/acssynbio.2c00062.
 251. Kannan, K. *et al.* One step engineering of the small-subunit ribosomal RNA using CRISPR/Cas9. *Sci. Rep.* **6**, 30714 (2016).

252. Brown, D. M. *et al.* Efficient size-independent chromosome delivery from yeast to cultured cell lines. *Nucleic Acids Res.* **45**, 1–10 (2017).
253. Labroussaa, F., Baby, V., Rodrigue, S. & Lartigue, C. La transplantation de génomes. *Médecine/Sciences* **35**, 761–770 (2019).
254. Lau, Y. H. *et al.* Large-scale recoding of a bacterial genome by iterative recombining of synthetic DNA. *Nucleic Acids Res.* **45**, 6971–6980 (2017).
255. Garcia-Morales, L. *et al.* A RAGE based strategy for the genome engineering of the human respiratory pathogen *Mycoplasma pneumoniae*. *ACS Synth. Biol.* **9**, 2737–2748 (2020).
256. Santos, C. N. S., Regitsky, D. D. & Yoshikuni, Y. Implementation of stable and complex biological systems through recombinase-assisted genome engineering. *Nat. Commun.* **4**, 1–10 (2013).
257. Santos, C. N. S. & Yoshikuni, Y. Engineering complex biological systems in bacteria through recombinase-assisted genome engineering. *Nat. Protoc.* **9**, 1320–1336 (2014).
258. Wang, G. *et al.* CRAGE enables rapid activation of biosynthetic gene clusters in undomesticated bacteria. *Nat. Microbiol.* **4**, 2498–2510 (2019).
259. Freund EA. ‘Culture media for classic mycoplasmas’ in *Methods in Mycoplasmaology*. Elsevier 127–135 (1983).
260. Flagg, M. P., Kao, A. & Hampton, R. Y. Integrating after CEN Excision (ICE) Plasmids: Combining the ease of yeast recombination cloning with the stability of genomic integration. *Yeast* **36**, 593–605 (2019).
261. Dicarolo, J. E. *et al.* Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res.* **41**, 4336–4343 (2013).
262. Pour-El, I., Adams, C. & Minion, F. C. Construction of mini-Tn4001tet and its use in *Mycoplasma gallisepticum*. *Plasmid* **47**, 129–137 (2002).
263. Ipoutcha, T. *et al.* Genome editing of veterinary relevant Mycoplasmas using a CRISPR-Cas Base Editor system. *Appl. Environ. Microbiol.* (2022).
264. Kouprina, N. & Larionov, V. Selective isolation of genomic *loci* from complex genomes by transformation-associated recombination cloning in the yeast *Saccharomyces cerevisiae*. *Nat. Protoc.* **3**, 371–377 (2008).
265. Robertson, W. E. *et al.* Creating custom synthetic genomes in *Escherichia coli* with REXER and GENESIS. *Nat. Protoc.* **16**, 2345–2380 (2021).
266. Zhu, X. *et al.* Mbov_0503 encodes a novel cytoadhesin that facilitates *Mycoplasma bovis* interaction with tight junctions. *Microorganisms* **8**, (2020).
267. Konkol, M. A., Blair, K. M. & Kearns, D. B. Plasmid-encoded ComI inhibits competence in the ancestral 3610 strain of *Bacillus subtilis*. *J. Bacteriol.* **195**, 4085–4093 (2013).
268. Gietz, R. D., Schiestl, R. H., Willems, A. R. & Woods, R. A. Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. *Yeast* **11**, 355–360 (1995).

269. Sharma, S., Tivendale, K. A., Markham, P. F. & Browning, G. F. Disruption of the membrane nuclease gene (MBOVPG45_0215) of *Mycoplasma bovis* greatly reduces cellular nuclease activity. *J. Bacteriol.* **197**, 1549–1558 (2015).
270. Venetz, J. E. *et al.* Chemical synthesis rewriting of a bacterial genome to achieve design flexibility and biological functionality. *Proc. Natl. Acad. Sci.* **116**, 8070–8079 (2019).
271. Noskov, V. N. *et al.* Assembly of large, high G+C bacterial DNA fragments in yeast. *ACS Synth. Biol.* **1**, 267–273 (2012).
272. Broach, J. R. *et al.* Localization and sequence analysis of yeast origins of DNA replication. *Cold Spring Harb. Symp. Quant. Biol.* **47**, 1165–1173 (1982).
273. Van Houten, J. V & Newlon, C. S. Mutational analysis of the consensus sequence of a replication origin from yeast chromosome III. *Mol. Cell. Biol.* **10**, 3917–3925 (1990).
274. Zheng, G., Hehn, R. & Zuber, P. Mutational analysis of the *sbo-alb* locus of *Bacillus subtilis*: identification of genes required for subtilosin production and immunity. *J. Bacteriol.* **182**, 3266–73 (2000).
275. Turner, S. M., Errington, J. & Mandelstam, J. Use of a *lacZ* gene fusion to determine the dependence pattern of sporulation operon *spolIIC* in *spo* mutants of *Bacillus subtilis*: A branched pathway of expression of sporulation operons. *J. Gen. Microbiol.* **132**, 2995–3003 (1986).
276. Deng, H. & Bloomfield, V. A. Structural effects of cobalt-amine compounds on DNA condensation. *Biophys. J.* **77**, 1556–1561 (1999).
277. Withers, H. L. & Nordström, K. Quorum-sensing acts at initiation of chromosomal replication in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 15694–15699 (1998).
278. Zimmerman, S. B. Toroidal nucleoids in *Escherichia coli* exposed to chloramphenicol. *J. Struct. Biol.* **138**, 199–206 (2002).
279. Ouameur, A. A. & Tajmir-Riahi, H. A. Structural analysis of DNA interactions with biogenic polyamines and cobalt(III)hexamine studied by fourier transform infrared and capillary electrophoresis. *J. Biol. Chem.* **279**, 42041–42054 (2004).
280. Hahn, J., DeSantis, M. & Dubnau, D. Mechanisms of Transforming DNA Uptake to the Periplasm of *Bacillus subtilis*. *MBio* **12**, (2021).
281. Kouprina, N., Noskov, V. N. & Larionov, V. 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 (2020).
282. Theis, J. F. & Newlon, C. S. The ARS309 chromosomal replicator of *Saccharomyces cerevisiae* depends on an exceptional ARS consensus sequence. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 10786–91 (1997).
283. Karas, B. J. *et al.* Direct transfer of whole genomes from bacteria to yeast. *Nat. Methods* **10**, 410–412 (2013).
284. Karas, B. J., Suzuki, Y. & Weyman, P. D. Strategies for cloning and manipulating natural and synthetic chromosomes. *Chromosom. Res.* **23**, 57–68 (2015).

285. Natalay Kouprina and Vladimir Larionov. Transformation-associated recombination (TAR) cloning for genomics studies and synthetic biology. *Chromosoma* **176**, 139–148 (2016).
286. Itaya, M., Tsuge, K., Koizumi, M. & Fujita, K. Combining two genomes in one cell: stable cloning of the *Synechocystis* PCC6803 genome in the *Bacillus subtilis* 168 genome. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15971–6 (2005).
287. Pellegrini, O., Mathy, N., Gogos, A., Shapiro, L. & Condon, C. The *Bacillus subtilis* ydcDE operon encodes an endoribonuclease of the MazF/PemK family and its inhibitor. *Mol. Microbiol.* **56**, 1139–1148 (2005).
288. Morimoto, T. *et al.* Six GTP-binding proteins of the Era/Obg family are essential for cell growth in *Bacillus subtilis*. *Microbiology* **148**, 3539–3552 (2002).
289. Nanamiya, H. *et al.* Identification and functional analysis of novel (p)ppGpp synthetase genes in *Bacillus subtilis*. *Mol. Microbiol.* **67**, 291–304 (2008).
290. Sorek, R. *et al.* Genome-wide experimental determination of barriers to horizontal gene transfer. *Science* **318**, 1449–52 (2007).
291. Kimelman, A. *et al.* A vast collection of microbial genes that are toxic to bacteria. *Genome Res.* **22**, 802–9 (2012).
292. Widom, R. L., Jarvis, E. D., LaFauci, G. & Rudner, R. Instability of rRNA operons in *Bacillus subtilis*. *J. Bacteriol.* **170**, 605–610 (1988).
293. Itaya, M. Physical mapping of multiple homologous genes in the *Bacillus subtilis* 168 chromosome: identification of ten ribosomal RNA operon loci. *Biosci. Biotechnol. Biochem.* **57**, 1611–4 (1993).
294. Lauber, M. A., Running, W. E. & Reilly, J. P. *B. subtilis* ribosomal proteins: Structural homology and post-translational modifications. *J. Proteome Res.* **8**, 4193–4206 (2009).
295. Nakano, M. M. & Zuber, P. Identification of genes required for the biosynthesis of the lipopeptide antibiotic surfactin in *Bacillus subtilis*. *Genet. Biotechnol. Bacilli* **170**, 397–405 (1990).
296. Dershowitz, A. & Newlon, C. S. The effect on chromosome stability of deleting replication origins. *Mol. Cell. Biol.* **13**, 391–398 (1993).
297. Wang, D., Lai, F. & Gao, F. Ori-Finder 3: a web server for genome-wide prediction of replication origins in *Saccharomyces cerevisiae*. *Brief. Bioinform.* **00**, 1–13 (2020).
298. Stinchcomb, D. T., Thomas, M., Kelly, J., Selker, E. & Davis, R. W. Eukaryotic DNA segments capable of autonomous replication in yeast. *Proc. Natl. Acad. Sci. U. S. A.* **77**, 4559–63 (1980).
299. Nieduszynski, C. A., Knox, Y. & Donaldson, A. D. Genome-wide identification of replication origins in yeast by comparative genomics. *Genes Dev.* **20**, 1874–1879 (2006).
300. Stinchcomb, D. T., Struhl, K. & Davis, R. W. Isolation and characterisation of a yeast chromosomal replicator. *Nature* **282**, 39–43 (1979).
301. Akanuma, G. *et al.* Inactivation of ribosomal protein genes in *Bacillus subtilis* reveals

- importance of each ribosomal protein for cell proliferation and cell differentiation. *J. Bacteriol.* **194**, 6282–6291 (2012).
302. Galperin, M. Y., Wolf, Y. I., Garushyants, S. K., Alvarez, R. V. & Koonina, E. V. Non essential ribosomal proteins in bacteria and archaea identified using clusters of orthologous genes. *J. Bacteriol.* **203**, 1–17 (2021).
 303. Nanamiya, H. *et al.* *Bacillus subtilis* mutants harbouring a single copy of the rRNA operon exhibit severe defects in growth and sporulation. *Microbiology* **156**, 2944–2952 (2010).
 304. Yano, K. *et al.* Growth and sporulation defects in *Bacillus subtilis* mutants with a single *rrn* operon can be suppressed by amplification of the *rrn* operon. *Microbiol. (United Kingdom)* **162**, 35–45 (2016).
 305. Kouprina, N., Kim, J. & Larionov, V. Highly selective, CRISPR/Cas9-mediated isolation of genes and genomic *loci* from complex genomes by TAR Cloning in yeast. *Curr. Protoc.* **1**, 1–27 (2021).
 306. Di Teodoro, G. *et al.* Contagious Bovine Pleuropneumonia: A comprehensive overview. *Vet. Pathol.* **57**, 476–489 (2020).
 307. Nicholas, R. & Churchward, C. Contagious Caprine Pleuropneumonia: New Aspects of an Old Disease. *Transbound. Emerg. Dis.* **59**, 189–196 (2012).
 308. Ganter, S. *et al.* Proteases as secreted exoproteins in Mycoplasmas from ruminant lungs and their impact on surface-exposed proteins. *Appl. Environ. Microbiol.* **85**, (2019).
 309. Vilei, E. M. & Frey, J. Genetic and biochemical characterization of glycerol uptake in *Mycoplasma mycoides* subsp. *mycoides* SC: Its impact on H₂O₂ production and virulence. *Clin. Diagn. Lab. Immunol.* **8**, 85–92 (2001).
 310. Levisohn, S. & Kleven, S. H. Avian mycoplasmosis (*Mycoplasma gallisepticum*). *Rev. Sci. Tech.* **19**, 425–42 (2000).
 311. Cizelj, I. *et al.* *Mycoplasma gallisepticum* and *Mycoplasma synoviae* express a cysteine protease CysP, which can cleave chicken IgG into Fab and Fc. *Microbiology* **157**, 362–372 (2011).
 312. Roberts, R. J., Vincze, T., Posfai, J. & Macelis, D. REBASE—a database for DNA restriction and modification: Enzymes, genes and genomes. *Nucleic Acids Res.* **43**, D298–D299 (2015).
 313. Delaney, N. F. *et al.* Ultrafast evolution and loss of CRISPRs following a host shift in a novel wildlife pathogen, *Mycoplasma gallisepticum*. *PLoS Genet.* **8**, (2012).
 314. Ipoutcha, T. *et al.* Multiple Origins and Specific Evolution of CRISPR/Cas9 Systems in Minimal Bacteria (*Mollicutes*). *Front. Microbiol.* **10**, 1–14 (2019).
 315. Yiwen, C., Yueyue, W., Lianmei, Q., Cuiming, Z. & Xiaoxing, Y. Infection strategies of mycoplasmas: Unraveling the panoply of virulence factors. *Virulence* **12**, 788–817 (2021).
 316. Xu, J. *et al.* *Mycoplasma gallisepticum* MGA_0676 is a membrane-associated cytotoxic

- nuclease with a staphylococcal nuclease region essential for nuclear translocation and apoptosis induction in chicken cells. *Appl. Microbiol. Biotechnol.* **99**, 1859–1871 (2015).
317. Nishida, K. *et al.* Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science* **353**, (2016).
 318. Komor, A. C. *et al.* Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C:G-to-T:A base editors with higher efficiency and product purity. *Sci. Adv.* **3**, 1–9 (2017).
 319. Ipoutcha, T., Gourgues, G., Lartigue, C., Blanchard, A. & Sirand-Pugnet, P. Genome engineering in *Mycoplasma gallisepticum* using exogenous recombination systems. *ACS Synth. Biol.* (2022).
 320. Chen, H. *et al.* Identification of biofilm formation by *Mycoplasma gallisepticum*. *Vet. Microbiol.* **161**, 96–103 (2012).
 321. Maunsell, F. P. *et al.* *Mycoplasma bovis* infections in cattle. *J. Vet. Intern. Med.* **25**, 772–83 (1967).
 322. Kumar, A., Rahal, A., Chakraborty, S., Verma, A. K. & Dhama, K. *Mycoplasma agalactiae*, an etiological agent of Contagious Agalactia in small ruminants: A Review. *Vet. Med. Int.* **2014**, 286752 (2014).
 323. Cacciotto, C. *et al.* *Mycoplasma agalactiae* MAG_5040 is a Mg²⁺-dependent, sugar-nonspecific SNase recognised by the host humoral response during natural infection. *PLoS One* **8**, e57775 (2013).
 324. Barré, A., de Daruvar, A. & Blanchard, A. MolliGen, a database dedicated to the comparative genomics of *Mollicutes*. *Nucleic Acids Res.* **32**, 307–310 (2004).
 325. Kouprina, N. & Larionov, V. Transformation-associated recombination (TAR) cloning for genomics studies and synthetic biology. *Chromosoma* **125**, 621–632 (2016).
 326. Boeke, J. D., Trueheart, J., Natsoulis, G. & Fink, G. R. 5-Fluoroorotic acid as a selective agent in yeast molecular genetics. *Methods Enzymol.* **154**, 164–75 (1987).
 327. Rideau, F. *et al.* Random transposon insertion in the *Mycoplasma hominis* minimal genome. *Sci. Rep.* **9**, 1–11 (2019).
 328. Marschall, P., Malik, N. & Larin, Z. Transfer of YACs up to 2.3 Mb intact into human cells with polyethylenimine. *Gene Ther.* **6**, 1634–7 (1999).
 329. Akanuma, G. *et al.* Inactivation of ribosomal protein genes in *Bacillus subtilis* reveals importance of each ribosomal protein for cell proliferation and cell differentiation. *J. Bacteriol.* **194**, 6282–91 (2012).
 330. Soler-Bistué, A., Timmermans, M. & Mazel, D. The proximity of ribosomal protein genes to *oriC* enhances *Vibrio cholerae* fitness in the absence of multifork replication. *MBio* **8**, (2017).
 331. Yano, K. *et al.* Multiple rRNA operons are essential for efficient cell growth and sporulation as well as outgrowth in *Bacillus subtilis*. *Microbiology* **159**, 2225–2236 (2013).

332. Rocha, E. P. C. The organization of the bacterial genome. *Annu. Rev. Genet.* **42**, 211–233 (2008).
333. Rocha, E. P. C. Order and disorder in bacterial genomes. *Curr. Opin. Microbiol.* **7**, 519–27 (2004).
334. Merrikh, H., Zhang, Y., Grossman, A. D. & Wang, J. D. Replication-transcription conflicts in bacteria. *Nat. Rev. Microbiol.* **10**, 449–58 (2012).
335. Fleurier, S., Dapa, T., Tenailon, O., Condon, C. & Matic, I. rRNA operon multiplicity as a bacterial genome stability insurance policy. *Nucleic Acids Res.* 1–20 (2022).
336. Brown, D. M. *et al.* Efficient size-independent chromosome delivery from yeast to cultured cell lines. *Nucleic Acids Res.* **45**, e50 (2017).
337. Noskov, V. N., Ma, L., Chen, S. & Chuang, R. Y. Recombinase-mediated cassette exchange (RMCE) system for functional genomics studies in *Mycoplasma mycoides*. *Biol. Proced. Online* **17**, 1–8 (2015).
338. Lau, Y. H. *et al.* Large-scale recoding of a bacterial genome by iterative recombineering of synthetic DNA. *Nucleic Acids Res.* **45**, 6971–6980 (2017).
339. Ye, Y. *et al.* Genomic iterative replacements of large synthetic DNA fragments in *Corynebacterium glutamicum*. *ACS Synth. Biol.* **11**, 1588–1599 (2022).

Développement de méthodes de clonage de génomes entiers chez la levure pour la construction de souches châssis semi-synthétiques de *Bacillus subtilis*.

Un des défis de la biologie de synthèse (BS), est d'apporter des solutions nouvelles à des enjeux globaux majeurs (thérapeutiques/sanitaires, climatiques), en particulier via la construction de souches de production utiles, performantes et respectueuses de l'environnement.

Bacillus subtilis (*Bsu*) est une bactérie Gram+ non pathogène utilisée en biotechnologie comme plateforme de production de molécules d'intérêt. Or, des études récentes ont établi que des souches de *Bsu* génétiquement modifiées permettaient d'obtenir une plus grande production de protéines recombinantes. Cela suggère que la production de châssis *Bsu* réduits pourrait être une étape importante dans l'amélioration des souches à visée industrielle.

Le projet ANR *Bacillus 2.0* dans lequel s'inscrit cette thèse, vise à transposer à *Bsu* des outils récents du domaine de la BS, et à mettre en place un pipeline efficace pour construire à haut débit des châssis modulables selon l'application biotechnologique. Ce pipeline rassemble les technologies suivantes : (i) design d'un génome synthétique (ii) assemblage de fragments d'ADN chevauchants au sein de la levure *Saccharomyces cerevisiae* (iii) isolement du génome et transplantation dans une cellule receveuse bactérienne et (iv) sélection et caractérisation des souches mutantes.

Les objectifs de cette thèse étaient d'attester la faisabilité de ces méthodes, en tentant de cloner et maintenir le génome réduit de *Bsu* MPG192 (2,86 Mb) dans la levure, puis de le modifier avec les outils d'ingénierie disponibles chez *S. cerevisiae*. Dans un premier temps des stratégies déjà décrites dans la littérature ont été déployées afin de cloner le génome entier de *Bsu*, mais sont restées infructueuses. En nous basant sur une approche de TAR-Cloning, nous avons tenté de cloner plusieurs fragments obtenus par restriction du génome de *Bsu*. Initialement, seuls cinq fragments sur sept ont été clonés. L'incapacité à cloner le plus grand de ces fragments (1,50 Mb) est vraisemblablement due à un manque d'ARS et/ou à sa taille. Concernant le second fragment, un ensemble de facteurs peuvent expliquer notre échec : à nouveau le manque d'ARS mais aussi, la présence de nombreux éléments répétés (7 opérons ribosomiques) ou l'expression délétère de ces gènes. Finalement, d'autres expérimentations ont permis de découper le génome en sous-fragments de tailles variables (6kb à 515kb) et ainsi de cloner en 21 morceaux la totalité du génome de *Bsu* MGP 192. Le TAR-Cloning imposant des contraintes dans le choix des sites de restriction, une nouvelle méthode de clonage, appelée CREasPy-Fusion, a été développée. Elle combine le système d'édition CRISPR-Cas9 et la fusion entre des sphéroplastes de levure et des cellules bactériennes. Comme preuve de concept, nous avons d'abord travaillé avec six espèces de mycoplasmes, et démontré qu'il est possible de cloner et modifier des génomes entiers. Cette approche a ensuite été transposée à *Bsu*, validant pour la première fois la fusion entre des sphéroplastes de levure et des protoplastes de bactéries Gram+. Elle a permis la capture précise d'un fragment d'environ 150 kb.

Bien que le génome entier de *Bsu* n'ait pas été cloné dans la levure à ce jour, plusieurs éléments critiques ont pu être identifiés. Tout d'abord, ces travaux soulignent l'importance de la méthode de clonage à adopter en fonction de l'organisme avec lequel on travaille. Ensuite, ils mettent en exergue l'existence de facteurs biologiques et techniques qui expliquent les difficultés actuelles et qui devront être pris en compte dans la suite des expérimentations. Enfin, ils ont permis le développement d'une nouvelle technique de clonage appelée CREasPy-Fusion, qui vient étoffer le catalogue des méthodes déjà décrites. Par sa versatilité, elle ouvre des perspectives pour la capture de grands fragments de génome, pour l'élimination de *loci* problématiques, ou encore, en appui à l'assemblage de fragments synthétiques.

Mots clés : *Bacillus subtilis*, Clonage de génome, Ingénierie génomique, Levure, Châssis bactérien

UMR 1332 Biologie du Fruit et Pathologie – Equipe Mollicutes

Centre INRAE Bordeaux-Aquitaine Bâtiment IBVM

71 Avenue Edouard Bourlaux, CS 2003233882 Villenave d'Ornon Cedex, FRANCE